



Universidad Andrés Bello
Facultad de Ciencias Biológicas
Escuela de Ingeniería en Biotecnología

**" Evaluación de la diversidad genética en una población de Congrio Colorado
(*Genypterus chilensis*) utilizando microsatélites"**

Proyecto de Tesis presentado como parte de los requisitos para optar
al Grado de **Magíster en Biotecnología**.

Profesor patrocinante: Dr. Alfredo Molina Sirguiado

Facultad de Ecología y Recursos Naturales
Universidad Andrés Bello

Profesor co-patrocinante: Dr. Claudio Meneses

Facultad de Ciencias Biológicas
Universidad Andrés Bello

Pamela González Troncoso

Santiago, Chile

2017.



UNIVERSIDAD ANDRÉS BELLO

Facultad de Ciencias Biológicas

Ingeniería en Biotecnología

**" Evaluación de la diversidad genética en una población de Congrio Colorado
(*Genypterus chilensis*) utilizando microsatélites"**

PAMELA GONZÁLEZ TRONCOSO

Este trabajo fue elaborado bajo la supervisión del Director de Tesis Dr. Alfredo Molina, en el laboratorio de Biotecnología Molecular, Universidad Andrés Bello, aprobado por los miembros de la Comisión de evaluación.

Dr. Alfredo Molina
Director de Tesis

Dr. Claudio Meneses
Co-Director de Tesis

Dra. Patricia Perez
Comisión de Tesis

Dra. Juliana Vianna
Comisión de Tesis

SANTIAGO-CHILE
Año 2017

ÍNDICE

1. RESUMEN.....	5
ABSTRACT	6
2. INTRODUCCIÓN	7
2.1 Acuicultura.....	7
2.2 Estructura poblacional de peces	9
2.2.1 Variabilidad genética.....	10
2.3 Marcadores genéticos.....	11
2.3.1 Microsatélites	12
2.3.1.1 Definición y formación	12
2.3.1.2 Distribución y función.....	13
2.3.1.3 Búsqueda de Microsatélites.....	15
2.3.1.4 Aplicaciones	16
2.4 Hipótesis:.....	19
2.5 Objetivos	19
2.5.1 Objetivo general	19
2.5.2 Objetivos específicos.....	19
3. METODOLOGÍA	20
3.1 Animales y diseño experimental	20
3.2 Búsqueda y caracterización de microsatélites	21
3.3 Evaluación de microsatélites	21
3.4 Análisis de datos	22
3.5 Parámetros estadísticos para estimar la diversidad genética	22
3.5.1 Frecuencias alélicas.....	23
3.5.2 Número de alelos por locus y número efectivo	23
3.5.3 Heterocigosidad.....	24
3.5.4 Índice de fijación o coeficiente de inbreeding.....	25
3.5.5 Estadísticos F de Wright	25
3.5.9 Contenido de información Polimórfico (PIC)	26
3.5.8 Probabilidad de identidad (PI).....	27
4. RESULTADOS.....	28

4.1 Identificación de microsatélites in silico en secuencias del transcriptoma	28
4.2 Validación de microsatélites en un grupo de individuos.....	29
4.3 Estimación de la diversidad genética de las poblaciones	37
5. DISCUSIÓN	47
6. CONCLUSIÓN	53
7. REFERENCIAS	54

1. RESUMEN

El congrio colorado (*Genypterus chilensis*) es una especie nativa de las costas chilenas con una alta demanda gastronómica, siendo un gran candidato para la expansión de la acuicultura nacional. El Centro de Investigaciones Marinas de Quintay (CIMARQ) ha logrado aclimatar al menos dos poblaciones de reproductores. A partir de la primera de ellas (2009) se obtuvo una población de juveniles, sin embargo, esta producción no se ha logrado replicar producto de altas mortalidades asociadas al cultivo larval. Aunque se desconoce el origen de esta mortalidad, y considerando que el cultivo de peces representa una fracción menor en términos de variabilidad genética en comparación a las poblaciones silvestres, es posible que la causa sea un fenómeno de consanguinidad. Consecuentemente, se postuló la siguiente hipótesis: "*La población de congrios colorados de CIMARQ presenta un bajo nivel de polimorfismo de microsatélites, demostrando una escasa variabilidad genética*". En este contexto, el objetivo de este estudio fue desarrollar marcadores de tipo microsatélites, derivados de información transcriptómica, para estimar la diversidad genética de las poblaciones de congrios en el CIMARQ.

Se obtuvieron doce marcadores polimórficos que permitieron analizar parámetros de variabilidad genética como heterocigosidad, índice de fijación, contenido de información polimórfica y diversidad genética *Fst*. En su conjunto estos datos permitieron determinar que la población de reproductores antiguos era la que tenía mejores parámetros de variabilidad y no presentaría el fenómeno de consanguinidad. La población de nuevos reproductores no presentó parámetros ideales de diversidad genética por lo que se sugiere incorporar nuevos ejemplares. Adicionalmente se comprobó que los marcadores escogidos permiten identificar un individuo de congrio colorado con una alta probabilidad.

Este estudio demuestra que los microsatélites derivados de transcriptoma pueden ser eficientes para desarrollar análisis de genética de poblaciones e identificación de individuos. Estos marcadores pueden ser aplicados para monitorear cultivos de *G. chilensis* y pueden ser útiles para el establecimiento de programas de mejoramiento genético. Adicionalmente, debido a la alta conservación que existe en las regiones flanqueantes de los microsatélites, estos marcadores también podrían ser utilizados para análisis genéticos de otras especies del género *Genypterus*.

ABSTRACT

The red cusk-eel (*Genypterus chilensis*) is a native species of the Chilean coasts with a high gastronomic demand, for this reason it is a great candidate for the expansion of national aquaculture. The Center of Marine Research of Quintay (CIMARQ) managed to acclimatize at least two populations of breeders. From the first population (2009) a juvenile population was obtained, nevertheless, this production has not been able to be replicated, due to high mortalities associated with the larval culture. Even though the origin of this mortality remains unknown, and considering that the fish culture represents a small fraction in terms of genetic variability in wild comparison, which can generate the phenomenon of inbreeding. Consequently, the following hypothesis was proposed: “*The population of red cusk-eel of CIMARQ presents a low level of microsatellite polymorphism, showing low genetic variability*”. In this context, the aim of this study was to develop transcriptome-derived microsatellite markers to estimate the genetic diversity of CIMARQ populations.

We obtained twelve polymorphic markers that allowed us to analyze parameters of genetic variability such as expected and observed heterozygosity, fixation index, polymorphic information content, and *Fst* genetic diversity parameter. Together, these data allowed us to determine that the population of old breeders had the best parameters of genetic variability and therefore would not present the phenomenon of inbreeding. The population of new breeders did not present ideal parameters of genetic diversity so it is suggested to incorporate new specimens. The chosen markers have a high probability of identifying an individual of red cusk-eel.

This study demonstrates that transcriptome-derived microsatellites can be efficient in developing population genetics analysis and identifying individuals. These markers can be applied to monitor *G. chilensis* cultures and may be useful for the establishment of breeding programs; and can also be used for the analysis of other *Genypterus* species due to the conservation that exists in flanking regions at microsatellites between species of *Genypterus*.

2. INTRODUCCIÓN

2.1 Acuicultura

En las últimas tres décadas la acuicultura mundial ha incrementado significativamente su productividad, presentando una tasa de crecimiento de 6% anual (FAO, 2016). En Chile la actividad acuícola ha crecido exponencialmente, siendo el mayor productor en América, representando un 27% de la producción total de peces, siendo el noveno exportador a nivel mundial (FAO, 2016). La industria chilena está representada principalmente por el cultivo de salmónidos, particularmente Salmón del Atlántico (*Salmo salar*), Salmón Coho (*Oncorhynchus kisutch*) y Trucha Arcoiris (*Oncorhynchus mykiss*), los cuales representan el 95% de las exportaciones nacionales de peces. Sin embargo, estas especies son susceptibles a enfermedades como la Anemia Infecciosa del Salmón (ISA), la cual el 2008 generó importantes pérdidas económicas para el país (FAO, 2008). Por otra parte, sólo un 26% de las exportaciones marinas nacionales corresponden a especies nativas, principalmente moluscos y algas (López *et al.*, 2012). Debido a la vulnerabilidad de las especies exóticas cultivadas y a la baja exportación de peces nativos, la tendencia actual de la acuicultura chilena es la diversificación de los peces de cultivo basada en especies nativas (Vega *et al.*, 2012), debido a que generan un menor impacto ambiental (Buschmann *et al.*, 1996), además de aumentar y consolidar la participación de la industria en el mercado internacional, permitiendo disminuir los riesgos asociados a la incorporación de especies exóticas (Alvial y Manríquez, 1999).

Una de las especies nativas que se ha considerado para dicha diversificación es el congrio colorado (*Genypterus chilensis*), la cual se distribuye en las costas chilenas desde Arica hasta Cabo de Hornos (Kong *et al.*, 1988). Esta es una especie altamente demandada y cotizada en el mercado local debido a la calidad de su carne blanca y fibrosa (Vega *et al.*, 2012). Sin embargo, existe una demanda insatisfecha, debido a la disminución en los volúmenes de desembarque de congrio colorado, lo cual se atribuye a la sobreexplotación de este recurso por la pesca artesanal provocando un aumento significativo en su valor (Chong y González, 2009). Bajo este contexto, el cultivo de congrio colorado en cautiverio se propone como una gran apuesta para satisfacer la demanda de producción, con el fin de evitar la sobreexplotación de este recurso, resolviendo así

un problema ambiental al permitir la conservación de esta especie y también contribuir con la diversificación de la acuicultura nacional.

A pesar de proyectarse como una especie candidata para la diversificación de la acuicultura nacional, la información sobre la biología del congrio colorado es escasa. Hasta el momento sólo se encuentran disponibles antecedentes respecto a su taxonomía (Chocair *et al.*, 1969), localización geográfica (Avilés, 1979) y fauna parasitaria (George-Nascimento *et al.*, 2004), no habiendo información sobre de la reproducción y primeros estadios de desarrollo en este pez. Hasta la fecha, el cultivo de congrio colorado no ha resultado del todo exitoso, a pesar de los numerosos intentos para reproducir su ciclo de vida en cautiverio. Por ejemplo, el Centro de Investigaciones Marinas de Quintay (CIMARQ) perteneciente a la Universidad Andrés Bello, bajo el proyecto FONDEF D06I-1024 (2008) “Desarrollo de Tecnologías para la Producción de Juveniles de Congrio Colorado”, logró aclimatar a cincuenta y dos ejemplares adultos de congrio colorado (peso promedio 8.8 ± 4.0 Kg) que fueron recolectados de diversos puntos de la V región y establecieron un sistema de incubación de ovas y producción de larvas, lo que en el año 2009 dio como resultado una generación de ciento diez juveniles (peso promedio 0.93 ± 0.3 Kg), sin embargo ésta producción no se ha logrado replicar debido a una alta mortalidad durante el desarrollo larval, y un tercio de los sobrevivientes presentaron malformaciones óseas, principalmente creaneofaciales, lo que les impidió alimentarse. Hasta la fecha no se han obtenido otros ejemplares de congrio colorado nacidos en cautiverio Algunos de los reproductores que se establecieron en un principio han fallecido con en el tiempo por lo que quedan veintiséis individuos que ya no se encuentran en etapa reproductiva, por esta razón se están aclimatando a doce ejemplares de juveniles (peso promedio 3.0 ± 1.0 kg) para ser establecidos como nuevo grupo reproductor.

En este contexto, el cultivo de peces representa por lo general una fracción mucho menor en términos de diversidad genética respecto a la población silvestre, lo cual puede generar el fenómeno de consanguinidad o *inbreeding*, permitiendo el cruce entre individuos relacionados parentalmente, lo que puede conllevar a la generación de anormalidades, infertilidad y mortalidad debido a la formación de homocigotos en genes indeseables (Gjdrem, 2000). En algunas especies de peces como pez zebra (*Danio rerio*) se ha reportado un 43% de mortalidad de larvas de peces consanguíneos en comparación a un 9% de mortalidad en peces no emparentados (McCune *et al.*, 2004). Debido a esta problemática, se desprende la necesidad de

desarrollar herramientas moleculares que permitan analizar la variabilidad genética y la asignación parental en el congrio colorado. Esto con el objetivo de generar peces reproductores fundadores que presenten una amplia variabilidad genética, además de permitir la selección de atributos deseables (ej; crecimiento, resistencia a enfermedades). Esto podría permitir la implementación de programas de cruzamientos para mejorar la calidad de las larvas y disminuir su mortalidad.

2.2 Estructura poblacional de peces

Las poblaciones naturales "de la mayoría de las especie" presentan cierto grado de diferenciación genética que puede estar dada por la interacción de una gran variedad de agentes como la presencia de barreras ambientales, procesos históricos y la distancia geográfica entre subpoblaciones. Dichos factores pueden moldear de alguna manera la información genética entre subpoblaciones de una especie, lo que se conoce como estructuración genética (Balloux y Lugon-Moulin, 2002). Por lo general, en el medio terrestre y dulceacuícola las poblaciones están bien delimitadas unas de otras y regularmente están separadas por barreras geográficas evidentes que evitan el entrecruzamiento y la mezcla entre poblaciones. Sin embargo, en el medio marino estas barreras no están claramente delimitadas o están ausentes. Además, la gran superficie del mar así como el continuo movimiento de las masas oceánicas favorecen el intercambio de individuos entre poblaciones (Knutsen *et al.*, 2003). El desarrollo de la genética de poblaciones ha permitido, por ejemplo, determinar el aislamiento reproductivo o la ocurrencia de flujo genético entre poblaciones de una misma especie.

Existen varios estudios de diferenciación genética de especies marinas, principalmente de peces pelágicos los cuales presentan una elevada capacidad reproductiva, así como un tamaño efectivo poblacional grande, estos factores pueden disminuir la diferenciación genética entre poblaciones de una misma especie. Sin embargo, el principal factor que influye en el grado de diferenciación genética en estas especies es su capacidad de dispersión, ya sea por transporte pasivo de las larvas o por migración activa de los adultos (Rocha-Olivares *et al.*, 2006; González y Zardoya, 2007). Las especies con mayor capacidad de dispersión, por lo general muestran niveles más altos de flujo genético entre poblaciones y por lo tanto tendrán niveles bajos de diferenciación genética (Ward, 2006).

El congrio colorado es una especie intermareal que habita principalmente las cuevas de los fondos rocosos; estudios de diferenciación genética de este tipo de peces es escasa, Kelly y Palumbi (2010) analizaron 50 especies marinas de la zona intermareal, principalmente moluscos y artrópodos, ellos sugieren que las especies entre la zona intermareal media y alta presentan niveles mayores de estructura poblacional que las especies de la zona rocosa. Existen diferencias en la estructura poblacional incluso en especies del mismo género, por ejemplo, la especie intermareal *Gillichthys seta* de la zona intermareal alta presenta menor flujo genético que la especie *G. mirabilis* que habita la zona intermareal baja (Huang y Bernardi, 2001).

2.2.1 Variabilidad genética

Todos los organismos presentan modificaciones en su genoma ya sea debido a mutaciones o por influencia del medio ambiente, dando lugar a la variación genética, la cual también surge entre los individuos permitiendo diferenciación a nivel de población, especie y grupos taxonómicos (Chauhan y Rajiv, 2010). La mayoría de las especies contienen dos sets de cromosomas, uno heredado del padre y el otro de la madre, siendo así una especie diploide (2n). Durante el proceso de meiosis, existen tres procesos que producen variabilidad genética en el gameto: crossing over, segregación y distribución independiente (Abernathy *et al.*, 2010). Debido a que los cromosomas se presentan en pares en especies diploides, cada gen tiene al menos dos copias, las cuales se denominan alelos. Las secuencias de ADN varían dentro de la población, es decir, existen diversos alelos en la misma posición (*locus*). Estas diferencias se denominan polimorfismos los cuales son utilizados como marcadores genéticos debido a la gran variabilidad que presentan (Liu y Cordes, 2004). La diferencia en los alelos dentro de un individuo puede producir variabilidad genética y cambios fenotípicos en una determinada población debido a que existen varios alelos para el gen que codifica a ese carácter (Cifci y Okumus, 2002).

Un análisis genético poblacional de la especie de interés "en base a marcadores genéticos polimórficos" se perfila como una gran estrategia para el establecimiento de un grupo de reproductores óptimo en condiciones de cultivo. Esto puede traer beneficios como la selección de rasgos de interés, pero también permite la mantención de la especie debido a la heterogeneidad genética poblacional. Para la implementación de dicha estrategia se requieren marcadores genéticos polimórficos, en donde los marcadores de ADN son los más populares en

estudios genéticos debido a que proveen mayor información acerca del flujo genético, mutaciones y frecuencia de alelos, entre otros (Neigel, 1997).

2.3 Marcadores genéticos

En distintas especies de organismos cultivados para el consumo humano, incluidos los peces, se ha investigado la identificación de marcadores genéticos, los cuales se han convertido en una herramienta útil en programas de mejoramiento genético, ya que permiten evaluar la variación genética de la población y la realización de estudios parentales, entre otros (Liu y Cordes, 2004).

Existen diversas técnicas moleculares disponibles para identificar marcadores genéticos dependiendo si existe o no la información genética necesaria para su búsqueda. Los marcadores genéticos tipo I son más utilizados en programas de mejoramiento genético, debido a que están asociados a un gen con función conocida por lo cual son de mayor utilidad para rastrear un fenotipo. Bajo esta categoría se encuentran: 1) las alozimas, las cuales son variantes alélicas de proteínas producidas por un mismo *locus*, y 2) los polimorfismos en la longitud de los fragmentos de restricción (RFLP: *Restriction Fragment Length Polymorphism* en inglés), donde se utilizan enzimas de restricción las cuales cortan en regiones específicas del ADN dando así un patrón diferencial de fragmentos producto del polimorfismo que puede encontrarse en las secuencias.

Por otro lado, cuando no hay información del genoma disponible para la especie de interés, se utilizan los marcadores genéticos tipo II, en donde se encuentra: 1) la amplificación aleatoria de ADN polimórfico (RAPD: *Random Amplification of Polymorphic DNA* en inglés), 2) los polimorfismos en la longitud de fragmentos amplificados (AFLP: *Amplified Fragment Length Polymorphism* en inglés), ambos marcadores genéticos son identificados por medio de la reacción en cadena de la polimerasa (PCR), 3) los polimorfismos de nucleótido simple (SNPs), los cuales corresponden a la variación de un nucleótido en la secuencia del gen, por lo cual un SNP puede producir cuatro alelos distintos debido al número de nucleótidos existentes, los SNPs no son muy utilizados en estudios genéticos debido a que presentan una baja herencia y porque además no suelen acumular mutaciones; y 4) los microsatélites, los cuales se presentan en abundancia en el genoma, presentan herencia Mendeliana y alto grado de polimorfismo (Liu y Cordes, 2004). En particular, estos últimos también pueden ser clasificados como marcadores de tipo I ya que también se pueden encontrar en regiones codificantes y son utilizados como

marcadores de excelencia para diversos estudios científicos, como la identificación de especie y determinación de paternidad, ya que son altamente informativos, además pueden ser analizados por medio de sistemas de secuenciación masiva.

La identificación de marcadores genéticos, en una forma eficiente e informativa, es un factor clave en los estudios de variación genética poblacional. Actualmente, existen metodologías como la tecnología de secuenciación de la siguiente generación (*next-generation sequencing technologies* en inglés) o sistemas de secuenciación masiva, que permiten identificar y analizar un gran número de secuencias en forma paralela, permitiendo así la evaluación de marcadores genéticos polimórficos de forma más rápida y eficiente (Guichoux *et al.*, 2011)

2.3.1 *Microsatélites*

2.3.1.1 *Definición y formación*

Los microsatélites o SSR (*simple sequence repeats* en inglés), son pequeñas unidades de 1 a 6 pares de bases de largo que se repiten en tándem, suelen abarcar entre unos pocos nucleótidos hasta cientos de bases de extensión (Beckmann y Weber, 1992). Son considerados como marcadores genéticos debido a que son altamente variables en una población, cuyo polimorfismo se basa en las diferencias en el número de repeticiones que presenta un alelo en un determinado *locus* (Liu y Cordes, 2004)

Se han postulado dos modelos para explicar la generación y mutación de los microsatélites. Uno de éstos es el deslizamiento (*slippage* en inglés) de la ADN polimerasa durante la replicación (Fig.1), en donde ocurre una disociación transiente de la hebra replicativa de ADN seguida por una desalineada reasociación de las hebras formando un pequeño bucle (Levinson y Gutman, 1987), si esta estructura no es reparada, generará el aumento o disminución del número de motivos dependiendo de si el bucle se forma en la cadena naciente o en la parental, respectivamente. (Sia *et al.*, 1997).

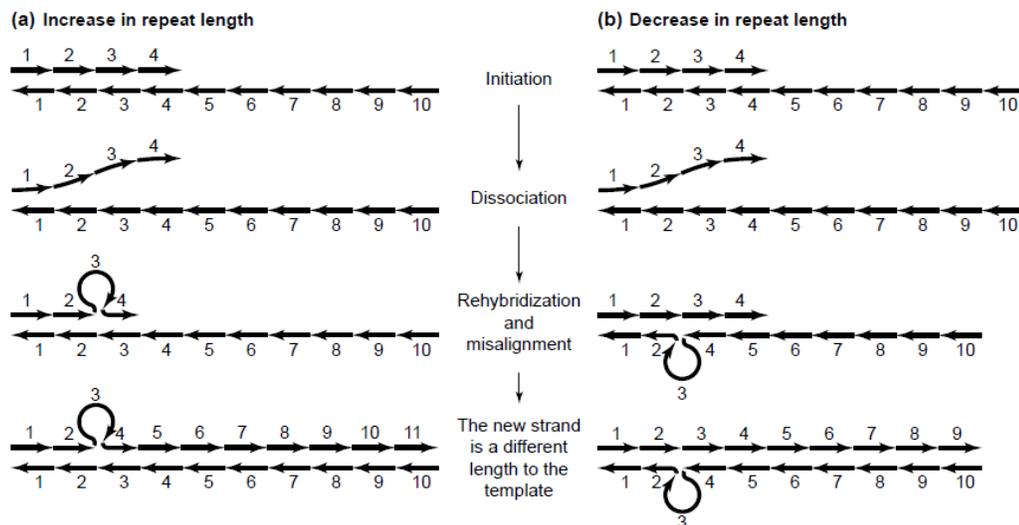


Figura 1. Modelo *slippage* ADN polimerasa (Ellegren, 2000)

Aunque existen suficientes antecedentes a favor del modelo anteriormente mencionado, hay algunos estudios que sugieren que la recombinación no recíproca o conversión genética podría introducir mutaciones en el largo de los microsatélites (Berg *et al.*, 2003). La conversión genética es un proceso meiótico que consiste en la transferencia de material genético entre secuencias homólogas, cromosomas homólogos y también entre la misma cromátida o cromosomas distintos (Chen *et al.*, 2007). La distribución genómica de los microsatélites está relacionada con los sitios de recombinación, siendo más una coincidencia de ubicación que una consecuencia de su generación (Trecó y Arnheim, 1986). Esta relación no ha podido ser comprobada y teniendo en consideración que el cromosoma Y no presenta recombinación meiótica, a pesar de que sí se han encontrado la presencia de microsatélites (Kayser *et al.*, 2000), por lo tanto, el modelo de deslizamiento de la polimerasa sería más asertivo para explicar la evolución de los microsatélites. Además, se ha comprobado *in vitro* "mediante la amplificación por PCR" que la ADN polimerasa es más propensa a tener este deslizamiento en repeticiones más largas de microsatélites (Ellegren, 2004). Esto explicaría que los microsatélites extensos sean más polimórficos, y por ende sean los más utilizados para análisis de variabilidad genética.

2.3.1.2 Distribución y función

Los microsatélites pueden encontrarse en regiones codificantes y no codificantes del ADN y diferir en el largo de sus secuencias y repeticiones, en consecuencia, microsatélites de largo tamaño abundan en regiones no codificantes del genoma, mientras que los microsatélites que se

encuentran en regiones codificantes están representados en menor proporción (Liu *et al.*, 2001), entre 9 a 15% en vertebrados (Van Lith y Van Zutphen, 1996). Esto se debe a una presión selectiva negativa frente a mutaciones en el marco de lectura de las secuencias traducibles (Li *et al.*, 2004).

Los microsatélites tienden a estar distribuidos en el genoma de todos los organismos, tanto procariontes como eucariontes, incluso en pequeños genomas bacterianos (Gur-Arie *et al.*, 2000). En peces se ha estimado que la ocurrencia de un microsatélite es de uno cada 10kb (Wright, 1993). Se encuentran en todos los organismos, en todos los cromosomas y en cada de región de los cromosomas (Liu y Cordes, 2004). No obstante, se encuentran menos presentes en el centrómero debido a la presencia en una alta densidad de elementos transposables, por otro lado, son abundantes en regiones teloméricas, en donde se han encontrado repeticiones de 4 a 10 kb en tilapia del Nilo (Chew *et al.*, 2002).

La mayoría de los microsatélites encontrados se presenta como repeticiones de dinucleótidos, representando entre un 30 a un 67% del total de microsatélites (Liu y Cordes, 2004). En esta categoría el dinucleótido (AC)_n es el motivo más común, siendo 2.3 veces más frecuente que el motivo (AT)_n que le sigue en abundancia (Chistiakov *et al.*, 2006). Los otros tipos de microsatélites tri-, tetra-, penta-, y hexanucleótidos tienen una frecuencia menor y son 1.5 veces menos comunes en el genoma de vertebrados que los dinucleótidos (Toth *et al.*, 2000). Se ha reportado que los microsatélites más polimórficos en una población corresponden a repeticiones de di- y tetranucleotidos (Ellegren, 2004).

La estructura de un gen se compone de una región promotora y de una región transcrita, donde se encuentran la región codificante y regiones que se expresan, pero no se traducen, las cuales ayudan a la estabilidad y vida media de los mRNAs (Brennan y Steitz, 2001). Estas últimas regiones se conocen como regiones no traducidas (UTRs, *untranslated regions* en inglés). Se ha descrito la presencia de microsatélites en todas las regiones nombradas anteriormente, por ejemplo, la presencia de microsatélites en la región promotora genera cambios en la expresión genética debido a que altera la afinidad de un factor de transcripción con su sitio de unión (Edwards *et al.*, 1998). En peces como la dorada (*Sparus aurata*), se han encontrado repeticiones CA_n en el promotor proximal del gen de la hormona de crecimiento, particularmente entre el sitio de unión para los factores de transcripción Pit-1 y HNF-3, en donde se relacionó este marcador con el tamaño de los peces (Almuly *et al.*, 2005). También se ha reportado la existencia de repeticiones de dinucleótidos o trinucleótidos en los UTRs, los cuales

pueden modificar diferencialmente la expresión de un gen alterando el genotipo y fenotipo del organismo (Liu *et al.*, 2001). En humanos se ha encontrado una repetición corta de microsatélite en la región 5' UTR del gen de insulina, la cual es muy polimórfica y en algunos alelos se ha visto que regula la expresión de dicho gen (Melloul *et al.*, 2002). En peces también se ha reportado la presencia de microsatélites en 5'- y 3'UTRs de genes. Por ejemplo, el microsatélite mencionado anteriormente, en la región 5'UTR del gen de insulina humana, también se ha reportado en el gen de insulina de tilapia del Nilo (Chistiakov *et al.*, 2006). Se han identificado microsatélites en 5'- y 3'UTRs de genes de varias especies de peces como, por ejemplo: en el gen del complejo de histocompatibilidad clase I de *Salmo Salar* (Grimholt *et al.*, 2002), y el gen para catalasa de *Danio rerio* (Gerhard *et al.*, 2000) entre otros. También se ha observado que los microsatélites situados en la región 3'UTR podrían alterar la expresión génica debido a su influencia en la estabilidad del mRNA (Chistiakov *et al.*, 2006). En anguila (*Anguilla anguilla*) se ha reportado que la repetición CTG en el 3'UTR del gen *TSHB* forma una estructura en forma de hélice, la cual se asocia con la inestabilidad y heterogeneidad de mRNA de este gen (Pradet-Balade *et al.*, 1998).

Todos los antecedentes descritos anteriormente evidencian la presencia y un rol importante de los microsatélites en regiones génicas, ya sea si se encuentra en las regiones no traducibles como en las codificantes. La variabilidad de microsatélites en estas regiones muestra una importancia fisiológica en términos de que pueden producir cambios en la expresión de algún gen en particular y por ende cambiar el genotipo y fenotipo de un individuo. Así, la identificación en una población de microsatélites polimórficos en estas regiones, podría ser una herramienta útil para programas de manejo genético y selección de reproductores con características de interés debido a la hipervariabilidad que presentan.

.3.1.3 Búsqueda de Microsatélites

El principal inconveniente de trabajar con microsatélites es que estos necesitan ser aislados *de novo* en la mayoría de las especies, involucrando un proceso complejo y laborioso de construcción de una librería genómica (Zane *et al.*, 2002). Brevemente, fragmentos de ADN son digeridos con enzimas de restricción y luego se seleccionan por tamaño y se clonan en un vector plasmídico. El enriquecimiento de las secuencias que contienen microsatélites ocurre antes o

después del paso de clonamiento utilizando sondas con secuencias repetitivas (Zane *et al.*, 2002). Recientemente, la disponibilidad de altos volúmenes de datos de secuencias de numerosos genomas eucariotas ha acelerado las investigaciones basadas en microsatélites (Sharma *et al.*, 2007). La búsqueda de microsatélites desde bases de datos de secuencias existentes ha reducido significativamente el tiempo y el costo involucrado en el aislamiento de estos marcadores.

En la última década, nuevas tecnologías de secuenciación masiva referidas como tecnología de Secuenciación de la Siguiete Generación (NGS, por sus siglas en inglés) han revolucionado el campo de la biología molecular. NGS ha permitido la generación de grandes volúmenes de datos de manera relativamente barata y en tiempos menores (Morozova y Marra, 2008; Wall *et al.*, 2009; Metzker, 2010). Sin embargo, el uso de NGS en genoma o transcriptoma para identificar marcadores moleculares, incluyendo microsatélites, es relativamente nuevo con solo unos pocos estudios que reportan esta aproximación en detalle, especialmente para especies no modelo. No obstante, los autores que usan esta aproximación para el desarrollo de marcadores han publicado resultados alentadores. NGS como un método de obtención de microsatélites comparado con métodos tradicionales basados en librerías genómicas, es más rápido, eficiente y técnicamente menos exigente.

2.3.1.4 Aplicaciones

Debido a su alto grado de polimorfismo, naturaleza multi-alelica, herencia codominante y distribución en el genoma es que los microsatélites han sido aplicados en una amplia variedad de disciplinas e investigaciones (Powel *et al.*, 1996).

Los microsatélites son los marcadores por excelencia para la construcción de mapas genéticos de ligamiento, debido a que son altamente polimórficos e informativos (Chistiakov *et al.*, 2006). Se han generado mapas genéticos de diversas especies de peces de interés comercial como salmón del atlántico (*Salmo salar*)(Moen *et al.*, 2004), trucha arcoíris (*Oncorhynchus mykiss*)(Nichols *et al.*, 2003) y lenguado japonés (*Paralichthys olivaceus*)(Coimbra *et al.*, 2003).

Estos marcadores también pueden ser utilizados en estudios epidemiológicos de patologías. Por ejemplo, se ha encontrado repeticiones de trinucleótidos en la región codificante del gen *rad23b* del sistema de reparación por escisión de nucleótido (*NER*) en tilapia (*Oreochromis niloticus*), en donde se ha involucrado la falla en este sistema con una alta incidencia de mutagénesis y carcinogénesis (Chistiakov *et al.*, 2006).

Los microsatélites también son utilizados en la identificación de un individuo, utilizando un panel con varios *loci* de microsatélites, se puede identificar a un individuo mediante el perfil del genotipo de microsatélites, el cual es altamente discriminativo en la población. También son utilizados en estudios de asignación parental, debido a que los microsatélites representan marcadores de ADN de *locus* codominante. Por cada microsatélite, la progenie hereda un alelo del padre y otro de la madre, debido a este patrón de herencia, este tipo de marcadores polimórficos son ampliamente utilizados en estudios de paternidad. En peces por ejemplo, Herbinger y colaboradores (1995) mediante el uso de 5 marcadores de tipo microsatélites logró asignar el 91% de la progenie de trucha arcoíris (*Oncorhynchus mykiss*) a una o dos parejas de 100 posibles progenitores. Estos marcadores también han sido utilizados para estudios de asignación parental de diversas especies de peces de interés comercial como *Oncorhynchus tshawytscha* (Brentsen *et al.*, 2001), *Scophthalmus maximus* (Castro *et al.*, 2004), entre otros. También se han descrito estudios utilizando los microsatélites en especies que presentan desoves en masa en los estanques de cultivo, en donde hay un gran número de potenciales padres para las larvas, y también se encuentran familias de peces que pueden estar emparentados (Herlin *et al.*, 2007). De esta manera los microsatélites permiten distinguir a los reproductores evitando el fenómeno de consanguinidad antes mencionado.

Gracias al uso de marcadores moleculares, tales como los microsatélites, es posible obtener una gran cantidad de información genética de las poblaciones naturales a partir de la estimación de ciertos parámetros, como pueden ser las frecuencias alélicas de cada subpoblación así como los niveles de heterocigosidad, es decir, la proporción de individuos de la muestra que son heterocigotos (Frankham *et al.*, 2004). Esta información permite estimar el grado de adecuación de la especie a su ambiente y en el caso de que esté presente una estructura, permite valorar la conectividad o los niveles de flujo genético entre poblaciones a partir del estadístico F_{st} propuesto por Wright (1951).

Bajo este contexto, la identificación de microsatélites en regiones codificantes y UTRs como marcadores genéticos a través de técnicas de secuenciación de nueva generación en el congrio colorado, podría proyectarse como una excelente alternativa para evaluar la diversidad genética de las poblaciones de cultivo en Quintay. De esta manera se podría: 1) Evaluar la heterogeneidad de las poblaciones de peces reproductores en centros de cultivo, 2) Seleccionar grupos de peces reproductores que presenten variabilidad genética 3) Genotipificar a estos reproductores y a su descendencia. Esto podría presentarse como una alternativa para solucionar los problemas de cultivo que presenta esta especie, particularmente las mortalidades masivas en las etapas larvales, las cuales podrían ser consecuencia de una alta consanguinidad en la población de congrios existentes en Quintay. Recopilando todos los antecedentes descritos anteriormente surge la siguiente pregunta **¿Será heterogénea la población de congrio colorado presente en Quintay en términos de variabilidad genética?** Esta pregunta permite formular la siguiente hipótesis de trabajo que se postula "*La población de congrios colorados de CIMARQ presenta un bajo nivel de polimorfismo de microsatélites, demostrando una escasa variabilidad genética*"

2.4 Hipótesis:

"La población de congrios colorados de CIMARQ presenta un bajo nivel de polimorfismo de microsátélites, demostrando una escasa variabilidad genética "

2.5 Objetivos

2.5.1 Objetivo general

Estimar la diversidad genética y desarrollar una estrategia para la identificación parental en una población de congrio colorado usando microsátélites.

2.5.2 Objetivos específicos

1. Identificar microsátélites *in silico* en secuencias del transcriptoma de músculo esquelético e hígado de congrio colorado.
2. Validar los microsátélites identificados en un grupo de individuos.
3. Estimar la diversidad genética de una población de congrio colorado utilizando microsátélites.
4. Identificar microsátélites hipervariables para la identificación del origen parental y test de paternidad.

3. METODOLOGÍA

3.1 Animales y diseño experimental

Para este estudio se analizaron tres poblaciones en cautiverio de congrio colorado (*Genypterus chilensis*) pertenecientes al Centro de Investigaciones Marinas de Quintay (CIMARQ) (Valparaíso, Chile. 33°13'00'S 71°38'00'W). La primera población consistía en 26 de los congrios reproductores antiguos (R.A) de 8 años de edad (peso promedio 8.0 ± 2.0 Kg) que fueron aclimatados en el año 2008, estos individuos son los posibles progenitores de una población de 83 congrios juveniles (F1) obtenidos el año 2009 (peso promedio 1.5 ± 0.5 kg). La tercera población contempló 12 individuos de 4 años de edad (peso promedio 3.0 ± 1.0 kg) los cuales son sexualmente inmaduros, pero se proyectan como nuevo grupo reproductor (R.N). También se analizaron 10 individuos de congrio colorado adultos pertenecientes a la región de Coquimbo, Chile (29°57'12'S 71°20'37'W). Las muestras fueron obtenidas el año 2012 y no se identificó el sexo de ninguno de los 131 ejemplares de congrio colorado analizados.

Se obtuvo una muestra de la aleta caudal de cada pez y se realizó una extracción de DNA genómico mediante el siguiente protocolo: se homogenizó 500 mg de tejido con 1 mL de buffer de digestión (100 mM NaCl, 10 mM Tris-HCl pH 8, 25 mM EDTA pH 8, 0,5% SDS, 0,1 mg/mL proteinasa K), y se incubó durante 12 h a 50°C, luego se adicionó igual volumen de fenol/cloroformo/alcohol isoamílico y se centrifugó a 13.000 rpm, posteriormente se transfirió la fase acuosa a un nuevo tubo al cual se le agregó 1/2 volumen de acetato de amonio 10 M pH 7,3 y 2 volumen de etanol absoluto, luego se centrifugó a 13.000 rpm durante 5 min, se eliminó el etanol y se agregó 500 μ L de etanol 70% y se centrifugó a 13.000 rpm durante 15 min a 4°C. Luego se eliminó el etanol y se dejó secar el pellet para luego resuspender en 50 μ L de agua libre de nucleasas. Posteriormente se cuantificó el ADN mediante NanoDrop con el sistema espectrofotómetro Epoch multi-volumen y se evaluó su pureza mediante la relación A260/280. La integridad del ADN extraído se evaluó mediante electroforesis en geles de agarosa 1% en buffer TAE 1X conteniendo 0.5 μ g/mL de bromuro de etidio a 100V durante 45 min.

3.2 Búsqueda y caracterización de microsatélites

Se analizaron 48.480 contigs correspondientes a el transcriptoma de referencia anotado de congrio colorado de músculo esquelético e hígado (NCBI número de acceso SRS614525) (Aedo *et al.*, 2014). Se realizó la búsqueda de microsatélites mediante el programa SSRLOCATOR (http://minerva.ufpel.edu.br/lmaia.faem/ssr_install_guide.html) con parámetros de búsqueda establecidos para detectar motivos de di- a deca-nucleotidos con un mínimo de cinco repeticiones de largo.

Posteriormente, se realizó una caracterización de los *contigs* del transcriptoma que contenían microsatélites y que estuviesen anotados, para lo cual se utilizó la herramienta bioinformática BLASTn para identificar los UTRs y la región codificante del gen mediante la comparación del *contig* con el microsatélite y genes en la base de datos de otras especies de peces ya anotados. Se corroboró dicho análisis utilizando también la herramienta ClustalW. También se utilizó el programa bioinformático Ensembl para analizar que los microsatélites encontrados estuviesen en un solo exón, lo que permitió el diseño de los partidores.

3.3 Evaluación de microsatélites

Se realizó un primer monitoreo analizando el largo de los microsatélites con el transcriptoma comparativo de dos congrios colorados juveniles (Aedo *et al.*, 2016), para lo cual se utilizó la herramienta bioinformática TABLET (<http://bioinf.hutton.ac.uk/tablet>) para comparar el número de repeticiones de los microsatélites en los contigs de los tres individuos.

Una vez identificados los microsatélites candidatos, se diseñaron partidores flanqueantes a la región que contiene al microsatélite. Los parámetros para el diseño de los partidores fueron que estos tuviesen un largo de 18 a 22 nucleótidos, un contenido de GC igual a 50% y que el producto de PCR tuviese un tamaño de 120 a 320 pb. Para el diseño de partidores se utilizó el programa Primer3 (<http://frodo.wi.mit.edu/primer3/>) y se sintetizaron en *Integrated DNA Technologies* (IDT; San Diego, CA, USA). En una primera instancia, se analizaron 30 genes candidatos para la posterior selección de microsatélites que contengan repeticiones y que además sean considerados polimórficos entre los dos individuos analizados del estudio antes mencionado.

Primero se analizaron los 30 *loci* con microsatélites con el ADN de 5 congrios colorados del grupo reproductor del CIMARQ mediante un PCR convencional. El PCR se realizó utilizando 20 ng/ μ L de ADN genómico, 5 μ L de GoTaq® Colorless Master Mix (Promega, Southampton, UK), 0,25 μ L de cada partidor a concentración 10 μ M, agua libre de nucleasas hasta completar un volumen final de 10 μ L. Se empleó el siguiente programa de PCR: 95° durante 4 min, seguido de 35 ciclos de 30 seg. a 95°, 20 seg. a temperatura variable dependiendo de la T_m del partidor, 40 seg. a 72° y 5 min de extensión final a 72°. El producto de PCR obtenido primero fue visualizado en un gel de agarosa 2% para corroborar su amplificación y luego fue analizado mediante electroforesis capilar Fragment.

Una vez identificados los *loci* con microsatélites con el mayor número de alelos, se seleccionaron 12 *loci* y se procedió a sintetizar el partidor directo acoplado al fluoróforo FAM o HEX en su extremo 5' (Macrogen. Seoul, Corea del Sur). Se utilizó el programa de PCR antes mencionado y se analizaron todos los individuos de congrio colorado; el producto de PCR primero fue visualizado en un gel de agarosa 2% para determinar su concentración mediante la densitometría del gel usando el HyperLadder II y el software ImageJ; después de obtener la concentración, se procedió a mezclar dos productos de PCR marcados con fluoróforos distintos, con el fin de analizar dos microsatélites en una misma corrida electroforética en el sistema de electroforesis capilar ABI3730XL (Macrogen Inc., Corea del Sur). La identificación del tamaño de los alelos se realizó utilizando el programa GeneMarker v2.6.3 (SoftGenetics, State College, PA) a partir del electroferograma generado por el secuenciador capilar.

3.4 *Análisis de datos*

Antes de realizar el análisis estadístico para la determinación de la diversidad genética, se analizaron los datos obtenidos con el software Micro-Checker v2.2.3 (van Oosterhout *et al.*, 2004) con el fin de evaluar la presencia de alelos nulos y errores de data.

3.5 *Parámetros estadísticos para estimar la diversidad genética*

Para la estimación de la diversidad genética de las cuatro poblaciones de congrio colorado, se analizaron diversos parámetros estadísticos: se realizó un conteo del número de alelos por *locus* (N_a), también se evaluó el número efectivo de alelos (N_e), frecuencias alélicas, el nivel de

heterocigosidad observada (H_o) y esperada (H_e), índice de fijación (F), estadísticos de Wright (F_{is} , F_{st} , F_{it}), distancia estándar de Nei (D_s), coeficiente de similitud genética y probabilidad de identidad (PI) utilizando el software GenaAlex v6.4.1. El contenido de información polimórfico (PIC) se calculó a partir de las frecuencias alélicas mediante website <http://w3.georgikon.hu/pic/english/kezi.aspx>. También se construyó un dendrograma a partir de las distancias genéticas utilizando el software DendroUPGMA (<http://genomes.urv.cat/UPGMA/>).

3.5.1 Frecuencias alélicas

Se puede definir la frecuencia alélica o génica como el cociente resultante de dividir el número de alelos iguales en una población por el número total de alelos. El cálculo de las frecuencias se hace por recuento directo de los alelos presentes, asumiendo que la observación de un solo alelo se corresponde con la condición de homocigosis, por lo tanto, que no hay alelos nulos.

3.5.2 Número de alelos por locus y número efectivo

El número de alelos (N_a) corresponde a la cantidad de alelos diferentes en un *locus* j variando en un intervalo de $1 < k < n_j$. El número total de alelos es la suma total de alelos de los n *loci* estudiados dentro de la población i , obtenido por:

$$N_a = \sum n_j$$

El número efectivo (N_e) se define como la probabilidad de que dos alelos de un *locus* elegido al azar en la población sean idénticos por descendencia (Frankham *et al.*, 2002), es decir, es la proporción de alelos que tienen una mayor probabilidad de transmitirse a la siguiente generación. Esta probabilidad es igual al inverso de la frecuencia esperada de individuos homocigotos en la población y se calcula mediante la expresión:

$$N_e = \frac{1}{\sum p_i^2} = \frac{1}{1 - H_e}$$

Donde:

p_i : frecuencia del i -ésimo alelo en un *locus*.

H_e : heterocigosidad esperada en un *locus*.

3.5.3 Heterocigosidad

Este parámetro se puede estudiar como heterocigosidad observada (H_o) y esperada (H_e). El H_o se define como la frecuencia relativa de individuos heterocigotos observados en la muestra para cualquiera de los *loci* y se calcula por un cómputo directo calculado mediante el cociente entre el número de heterocigotos para el *locus* y el total de la población:

$$H_o = \frac{\text{Número de heterocigotos}}{N}$$

La heterocigosidad esperada (H_e) desde el punto de vista matemático, es la probabilidad de que dos alelos tomados al azar de la población sean diferentes (Crow y Kimura, 1970). La heterocigosidad esperada es calculada usando una formula a partir de frecuencias alélicas:

$$H_e = \frac{2N}{2N - 1} = \left[1 - \sum_{j=1}^1 p_j^2 \right]$$

La heterocigosidad es una medida útil para la obtención de información de un *locus*, es decir los *loci* con heterocigosidad esperada de 0,5 o menos son en general no muy útiles para el análisis de parentesco.

3.5.4 Índice de fijación o coeficiente de inbreeding

Mide la probabilidad de que dos alelos en cualquier *locus* de un individuo sean idénticos por descendencia del ancestro común de los dos padres. Esto significa el grado en el que dos alelos son más propensos a ser homocigotos en lugar de heterocigotos en un individuo, porque los padres están relacionados.

Se calcula en base a las heterocigosidades según:

$$F = \frac{H_e - H_o}{H_e}$$

H_e = heterocigosis esperada media; H_o = heterocigosis observada media.

3.5.5 Estadísticos F de Wright

Los estadísticos de F estiman como está distribuida la variación genética dentro y entre subpoblaciones (Beebee y Rowe, 2008). Las subpoblaciones son entidades más pequeñas dentro de una población, es entonces el nivel más pequeño de estructura poblacional (Balloux & Lugon-Moulin, 2002). Existen tres estadísticos de F (F_{IS} , F_{ST} , F_{IT}) que fueron inicialmente descritos por Wright en 1943-1951 (Nei, 1977). El coeficiente F_{IS} evalúa el grado de endogamia de los individuos dentro de su subpoblación, es en esencia un coeficiente de endogamia. Por otro lado, F_{IT} mide cuanto se aleja una población global del equilibrio Hardy-Weinberg (Allendorf & Luikart, 2007); mientras que F_{ST} (el índice de fijación) es un estimador común de diferenciación entre subpoblaciones. Los estimadores están definidos de la siguiente manera (Beebee & Rowe, 2008):

$$F_{is} = \frac{\overline{H_e} - \overline{H_o}}{H_e}$$

$$F_{it} = \frac{H_t - H_o}{H_t}$$

$$F_{st} = \frac{H_t - \overline{H_e}}{H_t}$$

H_o es la heterocigosidad observada promedio a través de las subpoblaciones.

H_e es la heterocigosidad esperada a través de las subpoblaciones.

H_t es la heterocigosidad esperada para la población total.

El estadístico F_{IS} puede tomar valores que van desde -1 (exogamia) a 1 (endogamia). El índice de fijación F_{ST} toma valores positivos que van de 0 (no hay estructuración poblacional) a 1 (poblaciones separadas totalmente), en donde los valores $0 < F_{ST} < 0.05$ se considera como pequeña diferenciación genética, de $0.05 < F_{ST} < 0.15$ es moderada, de $0.15 < F_{ST} < 0.25$ es una gran diferenciación y $0.25 < F_{ST}$ es considerada una muy buena diferenciación genética (Wright, 1978).

3.5.9 Contenido de información Polimórfico (PIC)

El PIC es un índice que evalúa la información de un marcador en la población de acuerdo a las frecuencias de los alelos (Botstein *et al.*, 1980; Hearne *et al.*, 1992). El valor oscila entre 0 y 1, y se dice que un marcador es altamente informativo cuando el $PIC > 0.5$ razonablemente informativo si se tiene $0.5 > PIC > 0.25$, y poco informativo si ocurre $PIC < 0.25$. Para estos estudios se busca que el marcador tenga muchos alelos y PIC cercanos a 1.

$$PIC = 1 - \sum_{i=1}^j P_i^2 - 2 \sum_{i=j \neq 1}^j \sum_{j=1}^{i=1} P_i^2 P_j^2$$

Donde P_i y P_j son las frecuencias de los alelos i th y j th en el *locus* (Botstein *et al.*, 1980).

3.5.8 Probabilidad de identidad (PI)

La probabilidad de identidad (PI) proporciona una estimación de la probabilidad media de que dos muestras independientes tengan genotipo idéntico. Las estimaciones de PI son más ampliamente utilizadas en estudios forenses para comparar el poder potencial de diferentes *loci* o diferentes conjuntos de *loci*, o para comparar el mismo conjunto de *loci* entre diferentes poblaciones. El parámetro de probabilidad de identidad de cada *locus* se calcula:

$$PI = 2 \left(\sum p_i^2 \right)^2 - \sum p_i^4$$

Donde p_i es la frecuencia del i alelo en un *locus*.

4. RESULTADOS

4.1 Identificación de microsatélites in silico en secuencias del transcriptoma

Con el objetivo de identificar microsatélites en secuencias de congrio colorado, se utilizó la información genética del transcriptoma de referencia de músculo esquelético e hígado de congrio colorado que contempla 48.480 contigs. Mediante el uso del software SSRLOCATOR, se realizó la búsqueda de motivos de di- a deca-nucleotidos con un mínimo de cinco repeticiones de largo. En total, 6.635 (14%) de los contigs tenía al menos un microsatélite en su estructura, de los cuales 3.521 (53%) estaba asociado a un gen y 3.114 (47%) estaba asociado a un contig sin anotación (Figura 1).

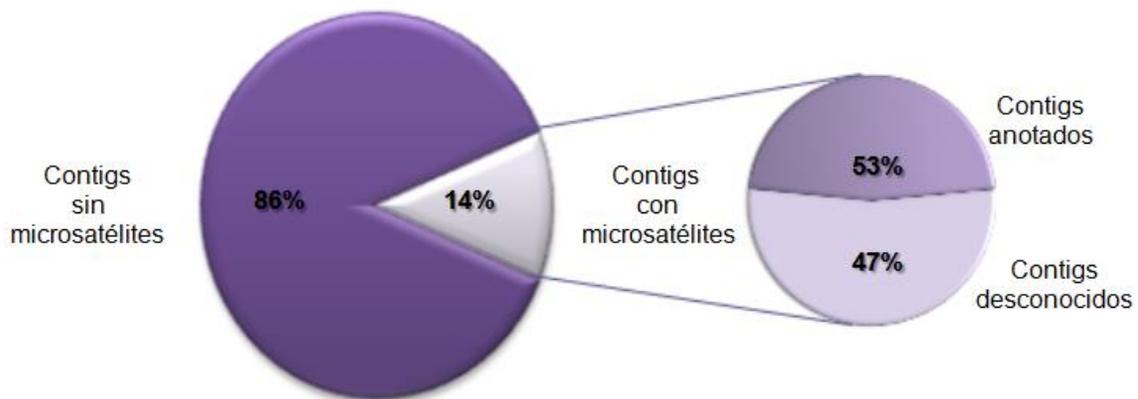


Figura 1. Contenido de microsatélites en transcriptoma de congrio colorado.

Los microsatélites fueron clasificados de acuerdo a su estructura según repeticiones perfectas o compuestas. Del total de microsatélites, sólo el 4.8% eran microsatélites compuestos y el 95.2% eran microsatélites simples ya que no tenían interrupciones en su estructura. Con respecto a la identificación del tipo de microsatélite, se obtuvo que los dinucleótidos eran los más abundantes representando el 53.1% del total de repeticiones y el motivo AC/GT era el más frecuente de ellos (50.4%), los trinucleótidos contemplaban el 16% y los tetranucleótidos el 18.7%, los motivos de diez nucleótidos eran los menos representativos (0.8%) (Tabla 1).

Tabla 1. Distribución de los motivos de microsatélites en el transcriptoma de congrio colorado.

Tipo de motivo	Número	Frecuencia (%)	Motivo más abundante	Frecuencia dentro de su propio tipo
Dinucleótido	3.526	53.1	AC/GT	50.4%
Trinucleótido	1.063	16.0	CTC/GAG	13.6%
Tetranucleótido	1.243	18.7	ACAG/GTCT	5.3%
Pentanucleótido	67	1.0	AGAGG/CCTCT	10.4%
Hexanucleótido	20	0.3	ACCAGG/CCTGGT	10.0%
Heptanucleótido	112	1.7	GAGGTCA/TGACCTC	2.7%
Octanucleótido	126	1.9	ACACACGC/GCGTGTGT	2.4%
Nonanucleótido	112	1.7	CAGCAGGAG/CTCCTGCTG	1.7%
Decanucleótido	50	0.8	AAACACACCC/GGGTGTGTTT	3.5%
Compuesto	316	4.8		
Total	6.635			

4.2 Validación de microsatélites en un grupo de individuos

Con el objetivo de evaluar los microsatélites obtenidos del transcriptoma de referencia de congrio colorado, se analizaron sólo los microsatélites que estuviesen presentes en *contigs* asociados a un gen. En una primera instancia, se evaluó el número de repeticiones del microsatélite del *contig* encontrado en el transcriptoma de referencia y se comparó con secuencias de otros dos congrios colorados utilizando la herramienta bioinformática TABLET. Se logró identificar treinta *contigs* con microsatélites que variaban en el número de repeticiones entre los tres individuos. Se procedió a caracterizar dichos *contigs* evaluando si la secuencia correspondía a la anotación del gen, si el microsatélite presente en dicha secuencia estaba localizado en la región codificante o en los UTRs y que además estuviesen presentes solo en un exón. Ninguno de los microsatélites analizados estaba localizado en la región codificante, sólo uno de ellos estaba en la región 5'UTR y todos los demás en la región 3'UTR. Con respecto al

tipo de microsatélite, veinticinco de los *loci* seleccionados correspondían a microsatélites del tipo dinucleótido, de los cuales en su mayoría correspondía al motivo CA, también había dos tetranucleótidos, un trinucleótido, un hexanucleótido y un nonanucleótido. El número de repeticiones variaba entre los individuos dependiendo del *locus*, hubo diferencias pequeñas entre el número de motivos en los *loci Gch02, Gch04, Gch06, Gch07, Gch08, Gch09, Gch11 y Gch16* los cuales tenían diferencias de un motivo entre los tres individuos de congrio colorado, por el contrario el *locus Gch15* fue el que presentó más diferencias en el número de repeticiones del microsatélite CA, en donde el individuo n°1 presentó veintinueve motivos, el pez n° 2 tenía quince y el congrio colorado del transcriptoma de referencia tenía doce (Tabla 2).

Tabla 2. Locus con microsatélites, ubicación y número de motivos en tres individuos.

<i>Locus</i>	Contig	Microsatélite	Ubicación	N° motivos T. referencia	N° motivos pez 1	N° motivos pez 2
<i>Gch01</i>	Ribosomal protein s6 kinase alpha-6	AC	3'UTR	15	17	15
<i>Gch02</i>	Transcription initiation factor TFIID subunit 8	GACCAG	3' UTR	5	6	6
<i>Gch03</i>	Eukaryotic translation initiation factor 2alpha kinase4	TCAGGACCC	5'UTR	7	11	8
<i>Gch04</i>	Calcium uniporter mitochondrial precursor	CAAC	3'UTR	6	7	7
<i>Gch05</i>	Acyl-CoA dehydrogenase family member 9	TTC	3'UTR	9	10	11
<i>Gch06</i>	Ubiquitin-protein ligase hectd1	TTCC	3'UTR	5	6	5
<i>Gch07</i>	Ubiquitin carboxyl-terminal hydrolase isozyme L1	CA	3'UTR	18	19	18
<i>Gch08</i>	Insulin receptor substrate 2-B	CA	3'UTR	15	16	16
<i>Gch09</i>	Ankyrin repeat domain-containing protein 11	GT	3'UTR	19	19	18
<i>Gch10</i>	Vacuolar atpase assembly integral membrane Vma21	TG	3'UTR	17	10	23
<i>Gch11</i>	Opioid growth factor receptor	CA	3'UTR	11	11	12
<i>Gch12</i>	V-type proton atpase 21_kda proteolipid subunit	GT	3'UTR	11	13	15
<i>Gch13</i>	RILP-like_protein_1	CA	3'UTR	16	18	18
<i>Gch14</i>	Apoptosis enhancing nuclease-like	AC	3'UTR	18	25	25
<i>Gch15</i>	Serine-threonine protein kinase MARK2	CA	3'UTR	12	29	15
<i>Gch16</i>	Heme oxygenase 2	TG	3'UTR	13	13	12
<i>Gch17</i>	UPF0687 protein C20orf27 homolog	CA	3'UTR	14	15	15
<i>Gch18</i>	FAS-associated factor 1	CA	3'UTR	18	22	21
<i>Gch19</i>	STE20-related kinase adapter protein alpha	CA	3'UTR	10	17	18
<i>Gch20</i>	Low density lipoprotein receptor adapter protein 1-B	TG	3'UTR	16	19	18
<i>Gch21</i>	Coiled-coil_domain-containing protein 137	TG	3'UTR	10	12	12
<i>Gch22</i>	Probable N-acetyltransferase camello-like	AC	3'UTR	12	12	10
<i>Gch23</i>	Bromodomain-containing protein 7	TG	3'UTR	14	14	18
<i>Gch24</i>	Carboxypeptidase B2 precursor	TG	3'UTR	16	17	20
<i>Gch25</i>	Mitochondrial inner membrane protein isoform 2	CA	3'UTR	11	7	8
<i>Gch26</i>	Polyadenylate-binding protein-interacting 2B	AC	3'UTR	11	13	13
<i>Gch27</i>	Bcl-2-like protein 1	GT	3'UTR	11	13	13
<i>Gch28</i>	Semaphorin-3F	CA	3'UTR	12	14	14
<i>Gch29</i>	Epidermal growth factor receptor kinase substrate 8	GT	3'UTR	11	16	16
<i>Gch30</i>	Protein NOXP20	CA	3'UTR	12	16	16

Posteriormente, se sintetizaron partidores para los primeros veinte *loci* con microsatélites y se realizó PCR convencional para probar dichos partidores (Tabla 3).

Tabla 3. Partidores de los veinte *loci* con microsatélites

<i>Locus</i>	Partidor	Secuencia (5'-3')	% GC	T _m (°C)	Tamaño esperado (pb)
<i>Gch01</i>	CIMARQ.GC.1.F	GAGACGCTCTACCACACTTC	55	55	127
	CIMARQ.GC.1.R	GGTAAAAAAGAGAGACAGTAG	41		
<i>Gch02</i>	CIMARQ.GC.2.F	GACTGAAAACAGACTCGCTC*FAM	50	52	191
	CIMARQ.GC.2.R	AAACGCTCTCAATCTAAACG	40		
<i>Gch03</i>	CIMARQ.GC.3.F	CGATAGAGGCAGGATAACG*FAM	53	52	229
	CIMARQ.GC.3.R	AATAAGGAGAGGGTGGCA	50		
<i>Gch04</i>	CIMARQ.GC.4.F	AATCTGTCCTGCCAAATCT *FAM	42	52	181
	CIMARQ.GC.4.R	GGTCAAGAAGTCAAACGC	50		
<i>Gch05</i>	CIMARQ.GC.5.F	CGTTTACCCTGAATGAAGAG*FAM	45	52	178
	CIMARQ.GC.5.R	AGAAGAAGAAGACGGCTCA	48		
<i>Gch06</i>	CIMARQ.GC.6.F	GCTGGAGGAGGAAGAGGA*FAM	61	54	163
	CIMARQ.GC.6.R	AAGACAGTGAAAGGAAAGATGG	41		
<i>Gch07</i>	CIMARQ.GC.7.F	GCACAAGTACACAGCCTGA*FAM	52	52	124
	CIMARQ.GC.7.R	AGACAACACAGAAAAGTGGG	45		
<i>Gch08</i>	CIMARQ.GC.8.F	TCAGGTAGTTGGGGAGAGT	53	52	293
	CIMARQ.GC.8.R	CTTCCAGAGATTGGGTGTT	48		
<i>Gch09</i>	CIMARQ.GC.9.F	GCGGACATAGATGTTTCAGACC*HEX	50	53	320
	CIMARQ.GC.9.R	GGGAGAAGGATGTGTGTGTA	50		
<i>Gch10</i>	CIMARQ.GC.10.F	CTTCCATTTGGCAGAGTG*HEX	50	53	188
	CIMARQ.GC.10.R	GAGGACGGAGTTTATACAGC	50		
<i>Gch11</i>	CIMARQ.GC.11.F	GACTAACACTGACGCACTGA*HEX	50	53	263
	CIMARQ.GC.11.R	GAAGAGGTGGGGAGTACAT	53		
<i>Gch12</i>	CIMARQ.GC.12.F	TTCCTCTGACCGCTCTTTGT	56	56	212
	CIMARQ.GC.12.R	AACTCTGACAGACCAGGCTAC	52		
<i>Gch13</i>	CIMARQ.GC.13.F	AGGAAGAGAGAGCCTGAAGT*HEX	50	54	288
	CIMARQ.GC.13.R	AGAGGAGGAAGCTGTGTGT	53		
<i>Gch14</i>	CIMARQ.GC.14.F	TACAGCCTGAAGACGCTAC	53	52	293
	CIMARQ.GC.14.R	CACACAACCTGACATGAGAC	50		
<i>Gch15</i>	CIMARQ.GC.15.F	TCCGATCTGCTCTGGCT	59	54	208
	CIMARQ.GC.15.R	ATCTACTGTCCTCCGTCTCC	55		
<i>Gch16</i>	CIMARQ.GC.16.F	AGTGTCAGGTGAGTTACAAGGA	45	51	193
	CIMARQ.GC.16.R	GTCACAGGTCAGACAGGAAG	55		
<i>Gch17</i>	CIMARQ.GC.17.F	TCGTCGAGCCTTCTGTGTA	54	56	208
	CIMARQ.GC.17.R	CTATTCACACTCTGTCAGTCCCT	55		
<i>Gch18</i>	CIMARQ.GC.18.F	GGGACTGACTGACTGATGAA*HEX	50	52	201
	CIMARQ.GC.18.R	CTTGTGTGTCACTGGATGTC	50		
<i>Gch19</i>	CIMARQ.GC.19.F	AGGTAAAAAGCCAGCAGC*HEX	50	53	247
	CIMARQ.GC.19.R	CGTCCCCTCTGTGTATTTAG	50		
<i>Gch20</i>	CIMARQ.GC.20.F	ACAGTCGGTGTGTTACTTCC	50	54	181
	CIMARQ.GC.20.R	GACCTTGACCTCACATGAAC	50		

En una primera instancia, los productos de PCR fueron visualizados en un gel de agarosa 2% para corroborar que existiera amplificado y que fuese del tamaño esperado. De los veinte *loci* analizados, sólo los *loci Gch14* y *Gch16* no generaron ningún producto de PCR y el *locus Gch01* presentó un tamaño mucho mayor al esperado (Figura 2), por lo que se omitieron estos tres *loci* para el posterior análisis.

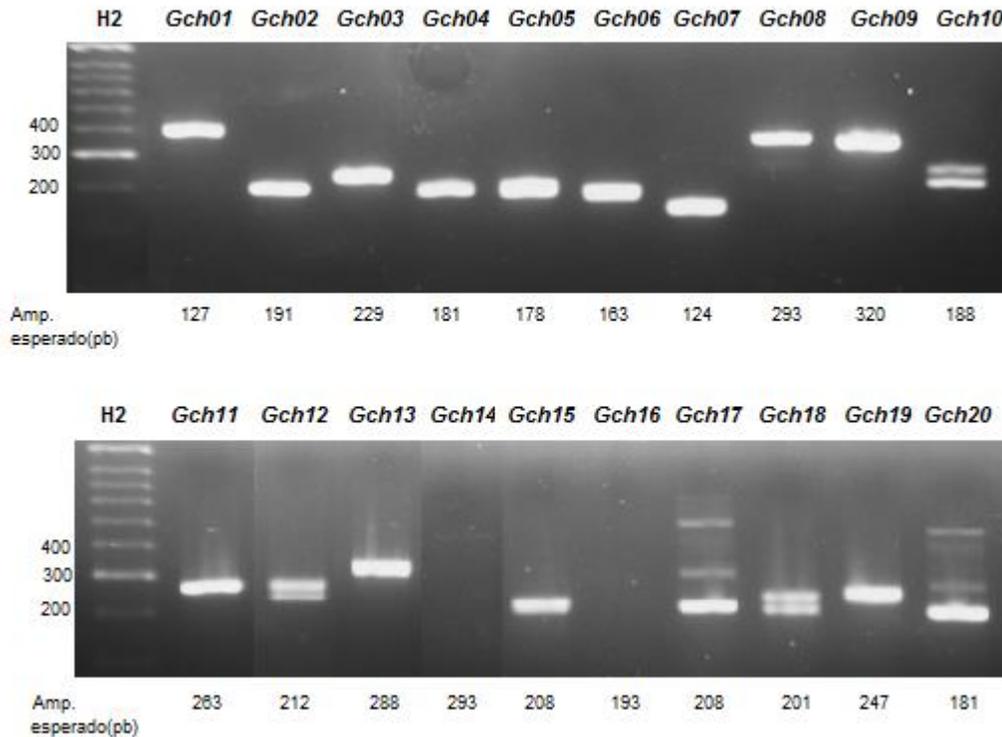


Figura 2. Evaluación de los partidores para los veinte *loci* con microsatélites. Electroforesis en gel de agarosa 2% para el producto de PCR de los veinte *loci* con el DNA genómico del individuo reproductor antiguo n°62. Hyperladder II utilizado como estándar de peso molecular.

Posteriormente, con los diecisiete *loci* que amplificaron de forma correcta, se realizó un PCR convencional con el DNA genómico de cinco individuos reproductores adultos, el producto de PCR fue analizado mediante el sistema de electroforesis capilar Fragment (Figura 3)

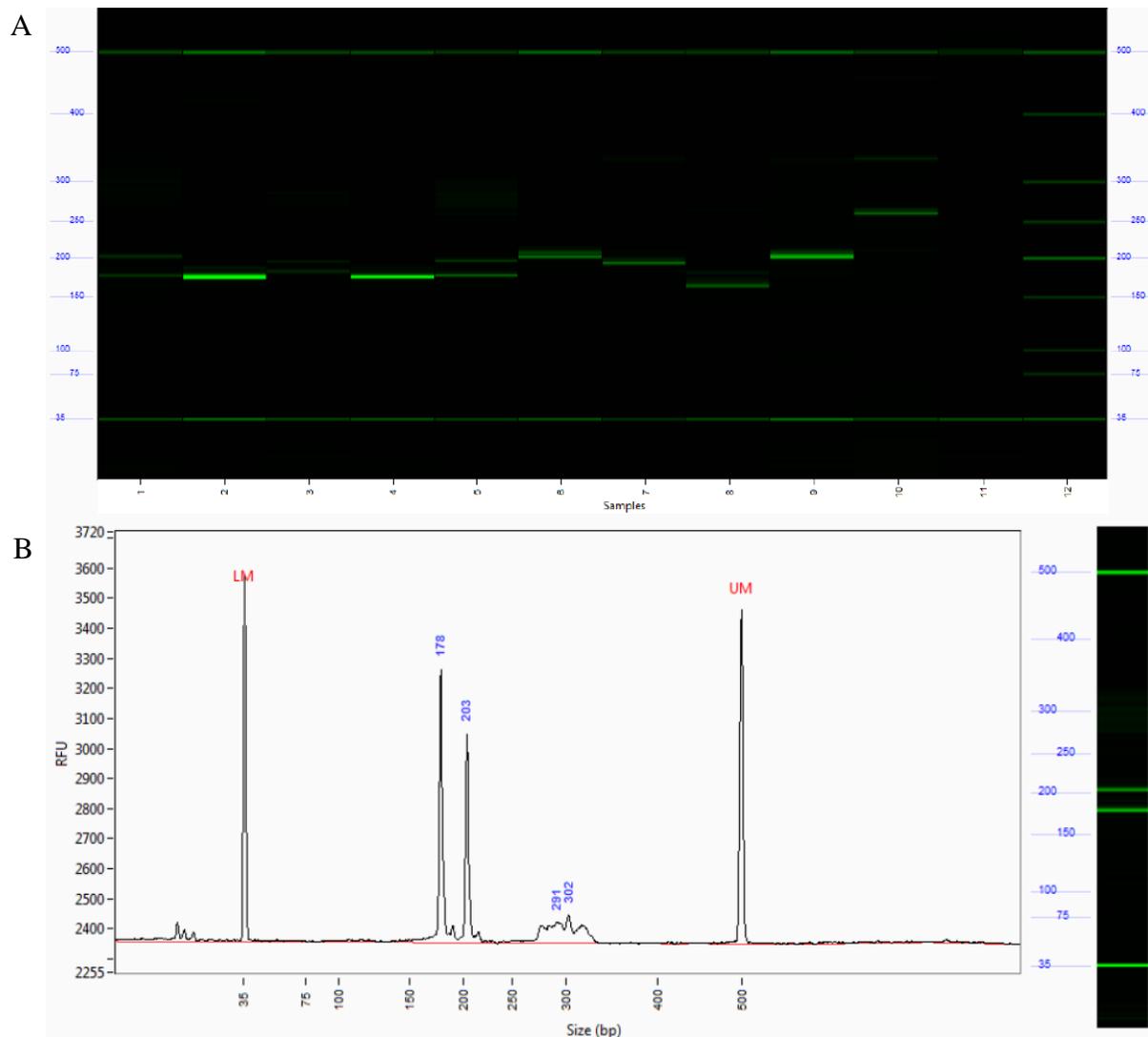


Figura 3. Electroforesis capilar Fragment para la selección de microsatélites polimórficos.

(A) Gel virtual de los productos de PCR en la electroforesis capilar. (B) Electroferograma para el *locus Gch02* del individuo reproductor antiguo n°25.

Con el sistema de electroforesis capilar se obtuvo como resultado un electroferograma del cual se obtienen peaks de intensidad los cuales corresponden al tamaño exacto del producto de PCR, lo que en este caso corresponderían a los alelos para el *locus* con microsatélite. Mediante esta técnica, se obtuvieron los alelos para los diecisiete *loci* de cinco individuos reproductores antiguos (Tabla 4).

Tabla 4. Alelos de los diecisiete *loci* de cinco congrios reproductores

<i>Locus</i>	Microsatélite	Alelos transcriptoma referencia	Alelos individuo n°25	Alelos individuo n°39	Alelos individuo n°51	Alelos individuo n°62	Alelos individuo n°104	N° alelos totales
<i>Gch02</i>	₅ (GACCAG)	191	178 / 203	176 / 188	183 / 196	176	179 / 197	8
<i>Gch03</i>	₇ (TCAGGACCC)	229	202 / 207	194 / 210	165 / 182	202 / 218	211 / 244	8
<i>Gch04</i>	₆ (CAAC)	181	172	172 / 177	172 / 175	172 / 185	177 / 179	5
<i>Gch05</i>	₉ (TTC)	178	181 / 190	179 / 193	179 / 193	176 / 188	172 / 187	8
<i>Gch06</i>	₅ (TTCC)	163	154	152 / 156	153 / 159	154 / 159	154	5
<i>Gch07</i>	₁₈ (CA)	124	125 / 137	123 / 137	121 / 125	121 / 144	121 / 135	6
<i>Gch08</i>	₁₅ (CA)	293	196 / 205	211 / 220	297	221 / 230	221 / 227	9
<i>Gch09</i>	₁₉ (GT)	320	320 / 331	320 / 327	320 / 337	317 / 335	320 / 330	7
<i>Gch10</i>	₁₇ (TG)	188	185 / 172	236 / 244	187 / 194	171 / 178	183 / 191	10
<i>Gch11</i>	₁₁ (CA)	263	255 / 266	265 / 281	266	265 / 281	265 / 281	4
<i>Gch12</i>	₁₁ (GT)	212	209 / 219	200 / 208	201 / 225	202 / 208	202 / 208	5
<i>Gch13</i>	₁₆ (CA)	288	285 / 322	285 / 322	285 / 318	285 / 306	285 / 318	4
<i>Gch15</i>	₁₂ (CA)	208	191 / 215	214 / 221	219 / 224	205 / 211	207 / 213	10
<i>Gch17</i>	₁₄ (CA)	208	250 / 257	258 / 265	265 / 269	265 / 269	273 / 278	6
<i>Gch18</i>	₁₈ (CA)	201	295 / 301	295 / 299	200 / 210	292 / 297	297 / 301	8
<i>Gch19</i>	₁₀ (CA)	247	248 / 256	251 / 257	247 / 260	252 / 257	244 / 257	5
<i>Gch20</i>	₁₆ (TG)	181	197 / 205	215 / 224	262 / 271	298 / 309	273 / 282	10

Una vez obtenidos todos los alelos, se procedió a seleccionar los doce *loci* que presentaban un mayor número de alelos entre los individuos analizados (*Gch02*, *Gch03*, *Gch04*, *Gch05*, *Gch06*, *Gch07*, *Gch09*, *Gch10*, *Gch12*, *Gch15*, *Gch19* y *Gch20*). Hay *loci* que presentaban un alto número de alelos pero los electroferogramas no eran confiables ya que los peaks de amplificación eran muy bajos por lo que se prefirieron otros *loci*, ésto ocurrió en el caso de los *loci* *Gch08* y *Gch18*. El sistema de electroforesis capilar Fragment no es tan preciso por lo que se procedió a utilizar un sistema más exácto como el sistema de electroforesis capilar ABI3037XL. Para utilizar esta técnica, se mandaron a sintetizar los partidores forward acoplados a un fluoróforo, los primeros seis *loci* seleccionados fueron marcados con el fluoróforo FAM y los otros seis con HEX, esto permitió analizar dos *loci* de un individuo en una misma corrida electroforética (Figura 4).

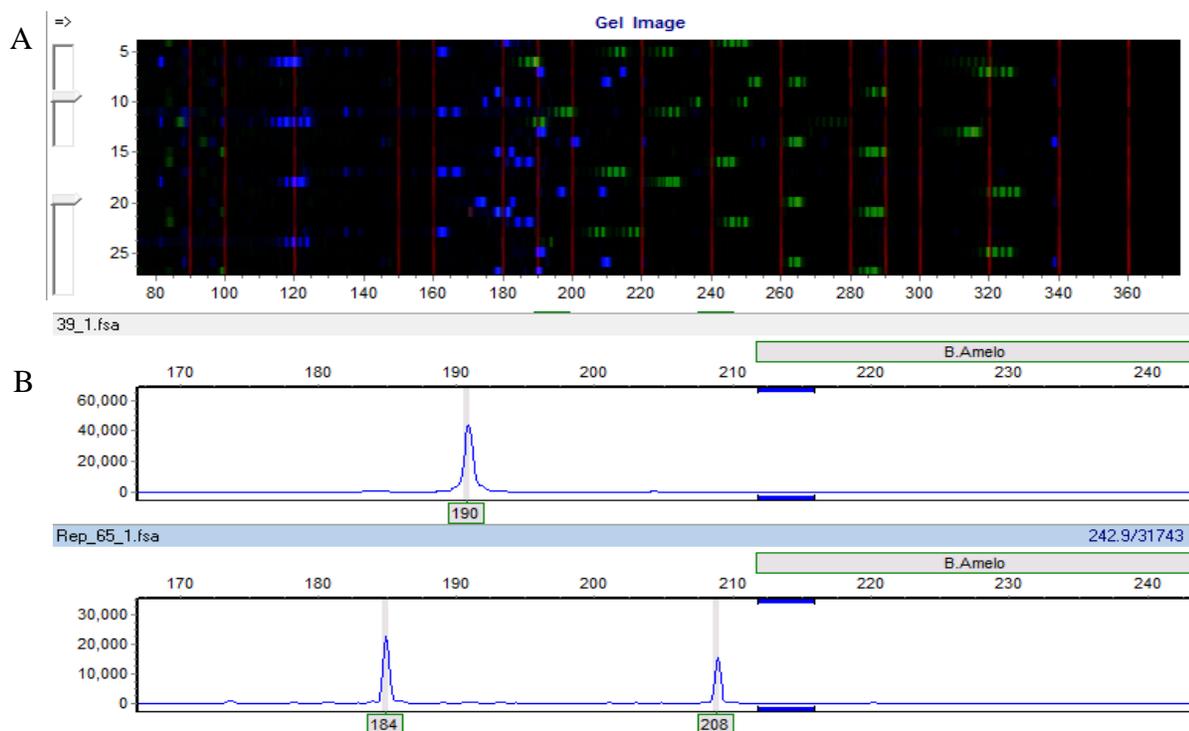
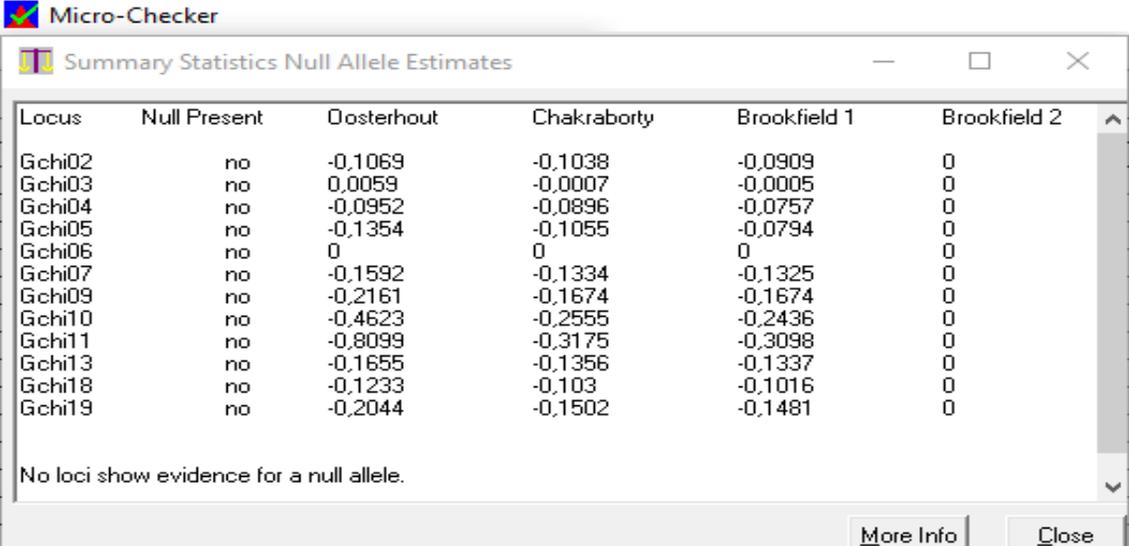


Figura 4. Electroforesis capilar ABI30XL para la identificación de alelos. (A) Gel virtual de los productos de PCR en la electroforesis capilar. (B) Electroferograma para el *locus* *Gch02* de los individuos reproductores antiguos n° 39 y 65.

4.3 Estimación de la diversidad genética de las poblaciones

Mediante electroforesis capilar ABI 3037XL se analizaron las cuatro poblaciones de congreso colorado y a partir de los electroferogramas se obtuvieron los alelos para los doce *loci* por individuo. De esta forma, se pudo identificar individuos homocigotos y heterocigotos para un mismo *locus*, lo que se puede observar en el electroferograma de la Figura 4, en donde el individuo reproductor adulto n°39 presentaba sólo el alelo 190, mientras que el individuo n°65 tenía los alelos 184 y 208 para el *locus Gch02*.

Previo al análisis de diversidad genética, se revisaron todos los datos de los alelos mediante el software Micro-Checker v2.2.3 para identificar la presencia de alelos nulos o errores en la interpretación de los electroferogramas. Según este análisis, no hubo presencia de alelos nulos en los *loci* de las cuatro poblaciones de congreso colorado (Figura 5).



Locus	Null Present	Oosterhout	Chakraborty	Brookfield 1	Brookfield 2
Gchi02	no	-0,1069	-0,1038	-0,0909	0
Gchi03	no	0,0059	-0,0007	-0,0005	0
Gchi04	no	-0,0952	-0,0896	-0,0757	0
Gchi05	no	-0,1354	-0,1055	-0,0794	0
Gchi06	no	0	0	0	0
Gchi07	no	-0,1592	-0,1334	-0,1325	0
Gchi09	no	-0,2161	-0,1674	-0,1674	0
Gchi10	no	-0,4623	-0,2555	-0,2436	0
Gchi11	no	-0,8099	-0,3175	-0,3098	0
Gchi13	no	-0,1655	-0,1356	-0,1337	0
Gchi18	no	-0,1233	-0,103	-0,1016	0
Gchi19	no	-0,2044	-0,1502	-0,1481	0

No loci show evidence for a null allele.

Figura 5. Análisis de la presencia de alelos nulos en las poblaciones de congreso colorado. Resumen del análisis estadístico de los alelos de doce *loci* de las cuatro poblaciones de congreso colorado utilizando el software Micro-Checker v2.2.3

En una primera instancia, se analizó el número de alelos de las cuatro poblaciones, en donde se obtuvo que la población de reproductores adultos fue la que tenía un promedio mayor de número de alelos distintos ($N_a=9.6$) que las otras tres poblaciones y a su vez, presentaba un elevado número de alelos únicos (5.5) y una baja cantidad de alelos comunes (0.9); mientras que la población de reproductores nuevos era la que presentaba valores más bajos de número de alelos (Figura 6).

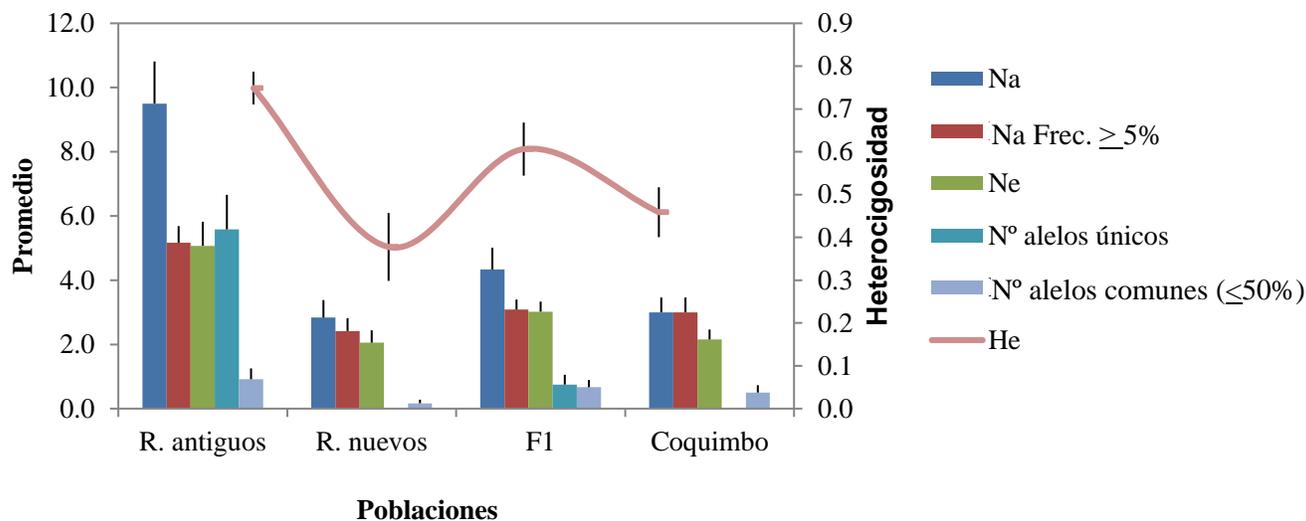


Figura 6. Patrón de alelos de las poblaciones de congrio colorado. N_a : número de alelos por locus, N_e : número efectivo de alelos por locus y H_e : heterocigosidad esperada.

Posteriormente, se procedió a analizar los parámetros de variabilidad genética de las cuatro poblaciones de congrio colorado. La población de reproductores antiguos fue la que presentaba un mayor número de alelos (N_a) por locus, de 3 a 18, mientras que la población de Coquimbo presentaba entre 1 y 9, y las otras dos poblaciones de Quintay tenían entre 1 y 8 alelos por locus; siendo el locus *Gch10* el que presentó el máximo de alelos en la población de reproductores antiguos. El número efectivo de alelos (N_e) registrado en la población de reproductores antiguos variaba entre 1.96 y 8.14, siendo nuevamente los valores más altos entre las cuatro poblaciones.

Con respecto a los niveles de heterocigosidad observada (H_o), la población de reproductores antiguos presentó valores entre 0.38 y 1.0, mientras que en las otras poblaciones la heterocigosidad observada variaba entre 0.0 y 1.0. Los valores de heterocigosidad esperada (H_e) más altos fueron registrados en la población de reproductores antiguos, variando entre 0.49 y 0.91, las otras tres poblaciones presentaron valores entre 0.0 y 0.9, en las cuatro poblaciones

analizadas se observa desequilibrio de Hardy-Weinberg. Con respecto a los niveles de heterocigosidad observada y esperada, los valores más bajos en las cuatro poblaciones siempre correspondían al *locus Gch06* ($\overline{H_o}=0.18$; $\overline{H_e}=0.13$) lo que se debe al bajo número de alelos presentes en este *locus* ya que en tres poblaciones sólo hubo individuos homocigotos con el alelo 162; mientras que los valores más altos de heterocigosidad correspondían principalmente al *locus Gch18* ($\overline{H_o}=0.95$; $\overline{H_e}=0.83$) lo cual coincide con que este *locus* era el que poseía el mayor número de alelos en las cuatro poblaciones de congrio colorado (Tabla 5).

En el caso del índice de fijación (F) todas las poblaciones analizadas presentaron valores negativos o cercanos a 0.0, lo que indicaría un alto número de individuos heterocigotos, sin embargo, el *locus Gch06* presentó valores indeterminados de índice de fijación en tres poblaciones ya que los niveles de heterocigosidad observada y esperada tuvieron valor de 0.0, y al presentar sólo un alelo, todos los individuos serían homocigotos para este *locus* (Tabla 5).

Tabla 5. Variabilidad genética de las cuatro poblaciones de congrio colorado.

Población	Parámetro	Locus											
		<i>Gch02</i>	<i>Gch03</i>	<i>Gch04</i>	<i>Gch05</i>	<i>Gch06</i>	<i>Gch07</i>	<i>Gch09</i>	<i>Gch10</i>	<i>Gch11</i>	<i>Gch13</i>	<i>Gch18</i>	<i>Gch19</i>
Reproductores antiguos	Rango de tamaño(pb)	184-214	173-263	174-190	169-190	162-174	108-132	312-340	176-242	253-267	280-300	195-229	230-260
	Na	7	10	5	7	3	12	11	18	5	10	13	14
	Ne	3.01	5.24	3.13	4.54	2.1	5.66	6.15	11.0	1.96	4.17	8.14	5.75
	Ho	0.65	0.38	0.69	0.38	0.73	1.00	0.77	0.54	0.42	0.77	1.00	0.73
	He	0.67 ^{ns}	0.81 ^{***}	0.68 ^{ns}	0.78 ^{***}	0.52 [*]	0.82 ^{ns}	0.84 ^{ns}	0.91 ^{***}	0.49 [*]	0.76 [*]	0.88 ^{***}	0.83 ^{ns}
	F	0.02	0.52	-0.02	0.51	-0.39	-0.21	0.08	0.41	0.14	-0.01	-0.14	0.12
Reproductores Nuevos	Rango de tamaño(pb)	190-196	173-236	178-182	184-187	162	106-120	324-326	188-194	265	286-290	189-213	236-246
	Na	2	2	2	2	1	4	2	4	1	3	8	3
	Ne	1.28	1.18	2.00	1.09	1.00	2.22	1.95	2.42	1.0	2.07	5.88	2.57
	Ho	0.25	0.17	1.00	0.08	0.00	0.67	0.83	0.92	0.00	0.75	0.92	1.00
	He	0.22 ^{ns}	0.15 ^{ns}	0.50 ^{***}	0.08 ^{ns}	0.00	0.55 ^{ns}	0.49 [*]	0.59 ^{ns}	0.00	0.52 ^{ns}	0.83 ^{ns}	0.61 ^{**}
	F	-0.14	-0.09	-1.00	-0.04	-	-0.22	-0.71	-0.56	-	-0.45	-0.10	-0.64
F1	Rango de tamaño(pb)	190-208	173-236	178-202	181-187	162	108-220	322-328	190-268	263-265	284-294	189-221	194-268
	Na	3	4	3	3	1	6	4	5	2	5	8	6
	Ne	2.82	3.6	2.66	2.03	1.00	4.09	3.49	2.22	2	3.89	4.85	3.58
	Ho	0.75	0.72	0.77	0.58	0.00	0.95	0.99	0.92	0.88	0.94	0.96	0.95
	He	0.63 ^{***}	0.71 ^{***}	0.63 [*]	0.49 [*]	0.00	0.77 ^{***}	0.71 ^{***}	0.55 ^{***}	0.49 ^{***}	0.74 ^{***}	0.81 ^{***}	0.72 ^{***}
	F	-0.18	-0.01	-0.23	-0.18	-	-0.24	-0.39	-0.67	-0.78	-0.27	-0.19	-0.33
Coquimbo	Rango de tamaño(pb)	190-196	173-236	178-184	184-187	162	106-120	320-326	192-194	263-265	284-290	190-218	236-250
	Na	2	4	4	2	1	4	3	2	2	4	9	3
	Ne	1.83	1.71	2.00	1.47	1.00	2.3	2.2	1.98	1.34	2.11	5.26	2.63
	Ho	0.50	0.50	0.80	0.40	0.00	0.70	0.90	0.90	0.30	0.70	0.90	0.80
	He	0.46 ^{ns}	0.42 ^{ns}	0.50 ^{ns}	0.32 ^{ns}	0.00	0.57 ^{ns}	0.55 ^{ns}	0.50 ^{**}	0.26 ^{ns}	0.53 ^{ns}	0.81 ^{ns}	0.62 ^{ns}
	F	-0.10	-0.20	-0.60	-0.25	-	-0.24	-0.65	-0.82	-0.18	-0.33	-0.11	-0.29

Número de alelos (*Na*), alelos efectivos (*Ne*), heterocigosidad observada (*Ho*), heterocigosidad esperada (*He*), índice de fijación (*Fis*); asteriscos en la columna *He* indican los *loci* que no cumplen con el equilibrio de Hardy-Weinberg (ns: no significativo; * P<0.05, ** P<0.01, *** P<0.001).

En la tabla 6, se indican los valores promedio de variabilidad genética de las cuatro poblaciones de congrio colorado. Para los doce *loci* analizados, se encontraron en promedio de 2.83 de a 9.56 alelos (N_a), mientras que para el número promedio de alelos efectivo (N_e) se encontró entre 2.06 y 5.07. Para todos estos indicadores, la población de reproductores antiguos fue la que obtuvo valores más altos en comparación a las otras poblaciones. Se observaron altos niveles de heterocigosidad observada (H_o) promedio en las cuatro poblaciones entre 0.55 y 0.78, mientras que la heterocigosidad esperada (H_e) promedio fue entre 0.38 y 0.72, nuevamente la población de reproductores antiguos fue la que tuvo valores mayores de variabilidad ($H_o=0.75$; $H_e=0.72$), mientras que la población de nuevos reproductores fue la que obtuvo valores más bajos ($H_o=0.55$; $H_e=0.38$). Los valores de índice de fijación (F) fueron en promedio negativos con excepción de la población de reproductores antiguos que tuvo un valor muy cercano a 0 ($F=0.08$) lo cual está dado por los altos niveles de heterocigosidad observada de las cuatro poblaciones. También se analizó el porcentaje de *loci* polimórficos, en donde se obtuvo que todas las poblaciones presentaron un alto porcentaje de *loci* polimórficos, siendo los reproductores antiguos los que presentaron 100% de sus *loci* con alelos distintos (Tabla 6).

Tabla 6. Índices promedio de variabilidad genética de las cuatro poblaciones de congrio colorado.

Poblaciones	N	N_a	N_e	H_o	H_e	F	%P
R. antiguos	26	9.58	5.07	0.75	0.72	0.08	100%
R. nuevos	12	2.83	2.06	0.55	0.38	-0.36	83%
F1	83	4.17	3.0	0.78	0.60	-0.32	92%
Coquimbo	10	3.33	2.15	0.62	0.46	-0.34	92%

Tamaño de la muestra (N), número de alelos (N_a), alelos efectivos (N_e), heterocigosidad observada (H_o), heterocigosidad esperada (H_e), índice de fijación (F_{is}) y porcentaje de locus polimórficos (%P).

Posteriormente se procedió a analizar la similitud genética y la distancia estándar de Nei de las cuatro poblaciones de congrio colorado, en donde se obtuvieron valores similares de estos parámetros entre los reproductores antiguos y las otras tres poblaciones ($I = 0.67- 0.70$; $D_s = 0.354 - 0.39$). Por otro lado, las poblaciones que compartían mayor similitud genética y menor distancia estándar de Nei fue la población de nuevos reproductores de Quintay con los individuos de la región de Coquimbo ($I = 0.96$, $D_s = 0.03$); ambas poblaciones también compartieron similitud genética con la población F1 (Tabla 7).

Tabla 7. Distancia estándar de Nei (bajo la diagonal) y coeficiente de similitud genética (sobre la diagonal) entre las cuatro poblaciones de congrio colorado

Población	Reproductores antiguos	Reproductores nuevos	Coquimbo	F1
Reproductores antiguos	-	0.672	0.707	0.675
Reproductores nuevos	0.398	-	0.966	0.816
Coquimbo	0.347	0.034	-	0.862
F1	0.394	0.203	0.149	-

Con los resultados de distancia estándar de Nei se construyó un dendrograma según el método UPGMA con el fin de agrupar a las cuatro poblaciones de congrio colorado (Figura 7).

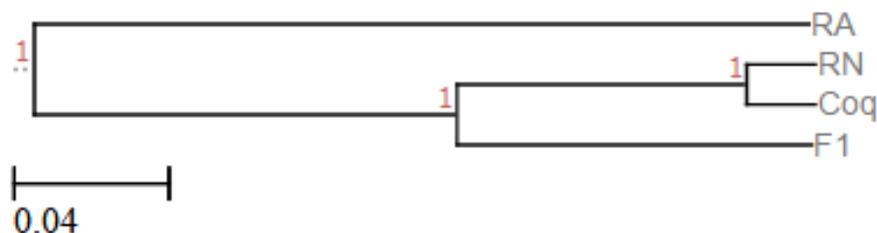


Figura 7. Dendrograma obtenido a partir de la distancia estándar de Nei entre las cuatro poblaciones de congrio colorado utilizando el método de UPGMA.

Con el objetivo de analizar la variabilidad genética de la población de Quintay, se omitió la población de la región de Coquimbo para los siguientes análisis. Primero se evaluó el contenido de información polimórfico (*PIC*) de los doce *loci* en estudio, y se obtuvo que los valores de *PIC* variaban entre 0.1 y 0.8, en donde ocho *loci* resultaron ser marcadores informativos ya que lograban detectar un polimorfismo en la población al tener $PIC \leq 0.5$, siendo el *locus Gch18* el que presentaba el valor más alto de *PIC*. Con respecto a los niveles de heterocigosidad de la población de Quintay, la heterocigosidad observada oscilaba entre 0.18 y 0.95, y la heterocigosidad esperada variaba entre 0.13 y 0.83, para ambos parámetros de heterocigosidad, los *loci Gch06* y *Gch18* eran los de valores más bajos y altos, respectivamente.

También se evaluó el flujo génico de la población, medido como el número de migrantes (*Nm*) entre las poblaciones; se obtuvieron valores entre 0.5 y 4.7 con un promedio de 2.1 individuos, por lo que el flujo génico sería la fuerza evolutiva predominante al ser *Nm* mayor que 1.

Se analizaron los estadísticos *F* de Wright con el fin de evaluar el nivel de endogamia a distintos niveles poblacionales, los valores obtenidos de *Fis* variaban entre -0.41 y 0.15 dando como valor promedio -0.21 por lo que no habría endogamia a nivel de subpoblación, los valores de *Fit* fluctuaron entre -0.29 y 0.36, con un promedio de -0.04 y por ende no habría endogamia como población total. Con respecto al estadístico *Fst*, éste tuvo valores positivos entre 0.05 y 0.34, con un promedio de 0.14 por lo que habría una moderada diferenciación genética en la población de Quintay según la clasificación de diferenciación genética de Wright (Tabla 8).

Tabla 8. Parámetros de variabilidad genética de la población de Quintay.

<i>Locus</i>	PIC	Promedio Ho	Promedio He	Nm	Fis	Fit	Fst
<i>Gch02</i>	0,4	0,54	0,49	1,4	-0,11	0,06	0,15
<i>Gch03</i>	0,5	0,44	0,52	0,8	0,15	0,36	0,24
<i>Gch04</i>	0,5	0,82	0,58	2,8	-0,41	-0,29	0,08
<i>Gch05</i>	0,4	0,36	0,42	0,9	0,11	0,30	0,21
<i>Gch06</i>	0,1	0,18	0,13	0,5	-0,39	0,08	0,34
<i>Gch07</i>	0,6	0,83	0,68	1,6	-0,25	-0,08	0,14
<i>Gch09</i>	0,6	0,87	0,64	2,9	-0,36	-0,25	0,08
<i>Gch10</i>	0,6	0,82	0,64	2,5	-0,29	-0,17	0,09
<i>Gch11</i>	0,3	0,40	0,31	1,1	-0,36	-0,10	0,19
<i>Gch13</i>	0,7	0,79	0,64	3,4	-0,26	-0,17	0,07
<i>Gch18</i>	0,8	0,95	0,83	2,5	-0,15	-0,08	0,06
<i>Gch19</i>	0,6	0,87	0,69	4,7	-0,26	-0,20	0,05
Promedio		0,66	0,55	2,1	-0,21	-0,04	0,14

Contenido de información polimórfico (*PIC*), heterocigosidad observada (*Ho*), heterocigosidad esperada (*He*), número de migrantes (*Nm*), coeficiente de endogamia de un individuo relativo a la subpoblación (*Fis*), coeficiente de endogamia relativo al total de la población (*Fit*), coef. endogamia de la subpoblación relativo a la población completa (*Fst*).

Uno de los objetivos de esta investigación era realizar un sistema de asignación parental, con el fin de identificar a los individuos reproductores antiguos que fueron los progenitores de la población F1, sin embargo, esto no se pudo realizar ya que no se logró identificar el sexo de los reproductores, debido a que las muestras fueron obtenidas en el año 2012 previo a esta investigación y no registraron el sexo de todos los individuos, los cuales algunos fueron falleciendo en el tiempo. Sin embargo, se logró realizar una primera aproximación con un dendrograma de distancia génica entre los reproductores antiguos y la población F1. Se obtuvieron cuatro clusters, en tres de ellos están agrupados casi todos los reproductores antiguos, con excepción de cuatro individuos identificados como reproductor antiguo n°27 (R.A.27), R.A.07, R.A.65 Y R.A.51 los cuales están agrupados en los clusters de los individuos

F1, por lo que estos individuos posiblemente serían los progenitores de estos individuos (Figura 8).

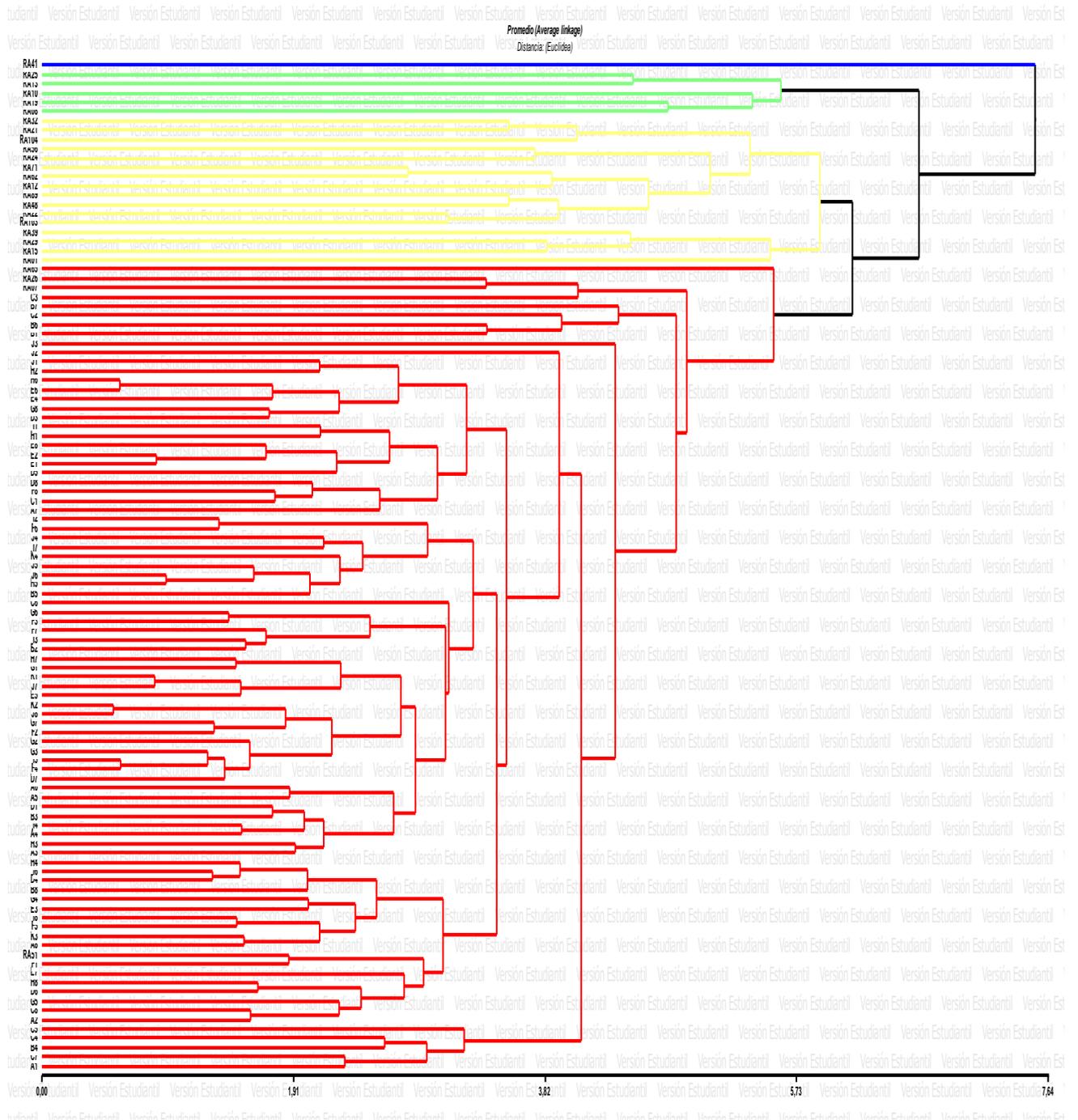


Figura 8. Dendrograma obtenido a partir de la distancia estándar de Nei entre la población de reproductores antiguos y los individuos de F1 de congrio colorado utilizando el método de UPGMA.

Se evaluó el parámetro de probabilidad de identidad (PI) que indica la probabilidad de identificar a dos individuos con genotipo idéntico en una misma población y también permite determinar cuántos *loci* son necesarios para identificar a un individuo con genotipo único. Según los resultados obtenidos, en la población de reproductores nuevos era más probable encontrar individuos con genotipos idénticos ya que los valores de PI eran más altos ($PI=6.3E-01 - 1.2E-05$), seguido por los individuos de Coquimbo ($PI=4.0E-01 - 1.7E-06$), F1 ($PI=2.0E-01 - 3.3E-09$), mientras que los reproductores antiguos fueron los que presentaron la menor probabilidad de identidad ($PI=1.5E-01 - 3.2E-14$) y eran los que menos *loci* requerían para identificar a un individuo ya que con sólo dos *loci* los valores de PI eran lo suficientemente bajos ($PI= 8.8E-03$); y con respecto al global de las cuatro poblaciones, utilizando seis *loci* ya es posible identificar a un individuo de congrio colorado ya que los valores de PI oscilaban entre $3.7E-02$ y $1.8E-06$ en las cuatro poblaciones (Figura 9, Tabla9).

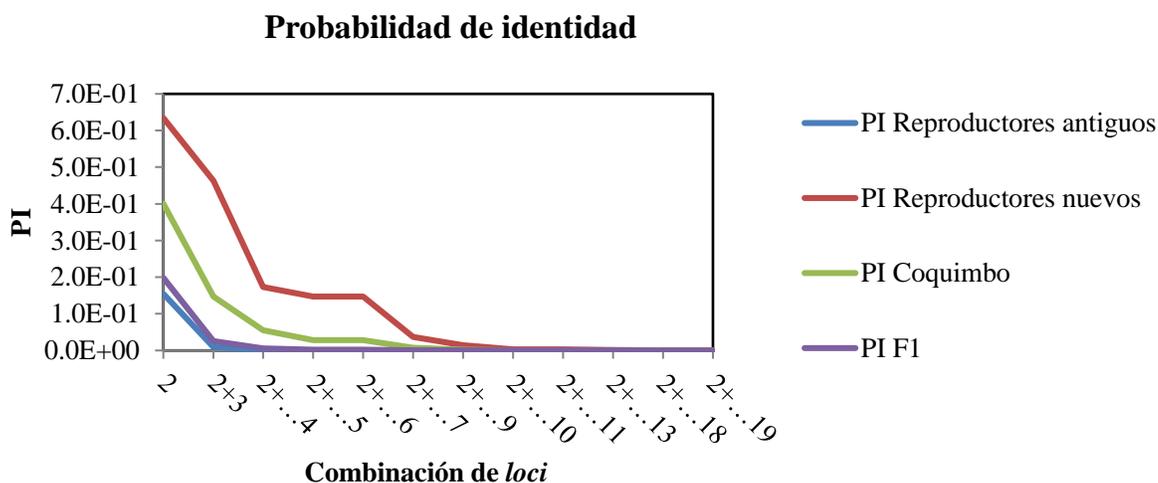


Figura 9. Probabilidad de identidad (PI) de las cuatro poblaciones de congrio colorado utilizando un panel de doce microtatélites.

Tabla 9. Probabilidad de identidad de las cuatro poblaciones de congrio colorado.

Población	N	2	2+3	2+...4	2+...5	2+...6	2+...7	2+...9	2+...10	2+...11	2+...13	2+...18	2+...19
Reproductores antiguos	26	1.5E-01	8.8E-03	1.3E-03	1.1E-04	3.6E-05	1.8E-06	8.2E-08	1.2E-09	3.5E-10	2.7E-11	7.1E-13	3.2E-14
Reproductores nuevos	12	6.3E-01	4.6E-01	1.7E-01	1.5E-01	1.5E-01	3.7E-02	1.4E-02	3.5E-03	3.5E-03	1.1E-03	5.2E-05	1.2E-05
F1	83	2.0E-01	2.5E-02	5.4E-03	1.9E-03	1.9E-03	2.0E-04	2.7E-05	8.2E-06	3.1E-06	3.5E-07	2.6E-08	3.3E-09
Coquimbo	10	4.0E-01	1.5E-01	5.5E-02	2.8E-02	2.8E-02	6.9E-03	2.2E-03	8.1E-04	4.8E-04	1.3E-04	8.0E-06	1.7E-06

5. DISCUSIÓN

En los programas de mejoramiento animal, el efectuar cruces dirigidos permite formar poblaciones homogéneas con características productivas similares. Tal es el caso del Centro de Investigaciones Marinas de Quintay (CIMARQ), cuya formación tuvo como finalidad generar núcleos de reproductores de congrio colorado orientados a obtener juveniles que permitiesen aumentar la población de ejemplares, con el fin de contribuir con la diversificación de la acuicultura nacional y la conservación de esta especie. Sin embargo, a pesar de establecer un grupo reproductor en el año 2009, que logró generar una población de juveniles, esta producción no se ha logrado replicar en el tiempo debido a una alta mortalidad larval y malformaciones en los peces sobrevivientes que al cabo de un tiempo fallecen. Por esta razón es que desde el año 2012 se han aclimatado nuevos ejemplares de congrio colorado que aún no están en etapa reproductiva, para lograr replicar esta producción de juveniles. Debido a que los reproductores se encuentran confinados en un estanque de cultivo y que esta especie presenta fecundación externa es que nace la necesidad de identificar a estos ejemplares, con el fin de identificar a los progenitores de los individuos sobrevivientes y también evaluar la variabilidad genética que presenta la población de reproductores, con el fin de seleccionar a los ejemplares con mayor variabilidad para evitar el fenómeno de consanguineidad.

Existen diversos tipos de marcadores genéticos para estudiar la estructura genética de una población, el análisis de la diversidad y los rasgos genéticos adaptativos (Beaumont, 2005). Siendo los marcadores de tipo microsatélite los más utilizados para estudios de variabilidad genética y test de paternidad, debido a su alto grado de polimorfismo y abundancia en el genoma. Sin embargo, la identificación de microsatélites puede ser laboriosa y costosa (Zalapa *et al.*, 2012), no obstante gracias a los sistemas de secuenciación masiva, que permiten identificar y analizar un gran número de secuencias en forma paralela, la búsqueda de estos marcadores genéticos polimórficos es de forma más rápida y eficiente permitiendo así su evaluación (Guichoux *et al.*, 2011). Los marcadores derivados de transcriptoma suelen ser menos polimórficos que los de DNA genómico debido a la localización de las secuencias en los genes (Ellis y Burke, 2007), sin embargo, son ampliamente utilizados en estudios de variabilidad genética.

Este estudio es el primer reporte de marcadores moleculares asociados con genes de función conocida en *Genypterus chilensis* desarrollado a partir de análisis de transcriptoma. Se utilizó el transcriptoma de referencia de músculo esquelético e hígado el cual contempla 48.480 contigs, de los cuales 6.635 (14%) poseía un microsatélite en su estructura, este porcentaje es similar a otros transcriptomas de peces, como el de *Acipenser fulvescens* que presentó un 13.4% de contigs con microsatélites (Doyle *et al.*, 2013). De los 6.635 contigs identificados, 3.521 contigs (53%) estaban asociados a un gen y 3.114 (47%) estaban asociados a un contig sin anotación (Figura.1). Según Weber (1990), los microsatélites pueden clasificarse en tres categorías: microsatélite perfecto, imperfecto y compuesto, en este estudio la mayoría de los microsatélites identificados fueron perfectos (95.2%), resultado similar a otras especies de peces, en donde se obtuvo un 66% en Carpa (*Cyprinus carpio L.*) y un 80% en Salmón del Atlántico (*Salmo salar*) (Crooijmans *et al.*, 1997). El tipo de microsatélite más abundante fue el de los dinucleótidos (53%) siendo el motivo (AC)_n el más frecuente de ellos (50.4%), los trinucleótidos contemplaban el 16% y los tetranucleótidos el 18.7% (Tabla 1), esto es concordante con los datos publicados en otros vertebrados, en donde los dinucleótidos son el tipo de microsatélite más abundante (Toth., 2000). Con respecto a especies de peces, también el dinucleótido es el tipo de microsatélite más abundante en transcriptomas como en la especie *Megalobrama amblycephala* donde un 65% correspondían a dinucleótidos, un 29% a trinucleótidos y un 5% a tetranucleótidos (Tran *et al.*, 2015).

Mediante un análisis comparativo, se logró identificar treinta microsatélites que variaban en el número de repeticiones y que estaban asociados a genes con función conocida. Todos estos microsatélites se ubicaban en las regiones no codificantes, principalmente en el 3'UTR (97%)(Tabla 2), debido probablemente a la selección contra mutaciones de cambio de lectura en los genes (Metzgar *et al.*, 2000; Wren *et al.*, 2000).

Mediante diversos mecanismos de selección, se escogieron doce genes con microsatélites para realizar los análisis de variabilidad genética de las poblaciones de congrio colorado de CIMARQ. Cuando comparamos con otras especies de peces marinos (DeWoody y Avise, 2000), el polimorfismo de los doce *loci* de microsatélites de *G. chilensis* fue intermedio. El número promedio de alelos (5.3) detectado en el congrio colorado fue menor que el reportado en otras especies de peces, con un promedio de 20.4 alelos por *locus* (DeWoody y Avise,

2000). En congrio dorado (*Genypterus blacodes*) ensayos de microsatélites revelaron entre 11 y 56 alelos en nueve *loci* en las costas australianas (Ward y Reilly, 2001), mientras que en las costas chilenas, se obtuvo entre 12 y 21 alelos utilizando tres *loci* (Canales-Aguirre *et al.*, 2010). El bajo número de alelos registrado en las poblaciones de congrio colorado podría deberse al pequeño número de peces analizados en este trabajo ya que los estudios antes mencionados analizaron 206 y 306 individuos respectivamente, mientras que en esta tesis se analizaron 131 individuos.

La variación biológica de la población es una base importante para evaluar los recursos de las especies para que ésta se adapte a los ambientes, mantenga una supervivencia y evolución a largo plazo (Li *et al.*, 2015). N_a , H_o , H_e y PIC son todos parámetros de la genética de poblaciones. En base a este estudio, las variaciones poblacionales de las cuatro poblaciones de congrio colorado fueron altas, a pesar del menor número promedio de alelos ($N_a = 5.3$), los niveles de heterocigosidad fueron altos ($H_o = 0.7$, $H_e = 0.5$) (Tabla 5), similar a lo obtenido en *G. blacodes* ($H_o = 0.84$, $H_e = 0.91$) (Canales-Aguirre *et al.*, 2010). La población de reproductores antiguos fue la que tuvo valores mayores de variabilidad ($H_o = 0.75$; $H_e = 0.72$), mientras que la población de nuevos reproductores fue la que obtuvo valores más bajos ($H_o = 0.55$; $H_e = 0.38$). Con respecto al índice de fijación (F), los valores promedio fueron negativos con excepción de la población de reproductores antiguos que tuvo un valor muy cercano a 0 ($F = 0.08$), lo cual está dado por los altos niveles de heterocigosidad observados en las cuatro poblaciones, lo que implica un alto número de individuos heterocigotos. También se analizó el porcentaje de *loci* polimórficos, en donde se obtuvo que todas las poblaciones presentaron un alto porcentaje de *loci* polimórficos, siendo los reproductores antiguos los que presentaron 100% de sus *loci* con alelos distintos y los reproductores nuevos el valor más bajo con un 83% (Tabla 6). Según estos parámetros de variabilidad genética, la población de reproductores antiguos presentaba valores óptimos de variabilidad genética, observado por los altos niveles de heterocigosidad y un bajo índice de fijación, por lo que esta población no presentaría el fenómeno de consanguineidad, descartando que la mortalidad larval de los congrios colorados fuese por problemas de diversidad genética de esta población. Probablemente, la mortalidad larval estaba asociada a la falta de nutrientes en el alimento suministrado, ya que existen antecedentes que indican que la carencia de vitaminas C y E influyen en el desarrollo y supervivencia en las distintas etapas del desarrollo de peces (Palace V. y Werner J., 2006), larvas de tilapia del nilo (*Oreochromis*

niloticus) alimentadas en carencia de vitamina C presentaban malformaciones y mortalidad en comparación a una dieta normal (Sommani A. y Jindathip R., 2003). Por otro lado, los ejemplares de congrio colorado que se están aclimatando para ser los nuevos peces fundadores en Quintay, fueron los que presentaron niveles más bajos de parámetros de variabilidad, si bien no son valores desfavorables, sería útil incorporar nuevos ejemplares de otras regiones con el fin de aumentar la variabilidad genética de esta población y así tener un stock más enriquecido genéticamente.

En la mayoría de las especies de peces de importancia económica, la diferenciación genética de la población es muy baja ($F_{st} < 0.01$, Ward *et al.*, 1994). Esto puede deberse a dos factores: la presencia de áreas sobrepobladas de desove y/o a un fuerte flujo genético entre las poblaciones que se traduce en una falta de deriva genética, limitando así la divergencia genética entre las poblaciones (Hauser y Ward, 1998). Al analizar el parámetro F_{st} de diferenciación genética de la población, éste tuvo valores positivos entre 0.05 y 0.34, con un promedio de 0.14 por lo que según la clasificación de diferenciación genética de Wright, habría una moderada diferenciación genética en la población de Quintay (Tabla 8), este resultado puede deberse a que el establecimiento del primer grupo reproductor incluía ejemplares de congrio colorado de diversas zonas de la Quinta región, lo que aumentó la diversidad genética de esta población. Por otro lado, estudios en *G. blacodes* indican una baja diferenciación genética de esta especie en poblaciones de las costas chilenas ($F_{st} = 0.032$) (Canales-Aguirre *et al.*, 2010), sin embargo este estudio incluía poblaciones muy cercanas geográficamente, y al considerar que esta especie presenta un fuerte comportamiento migratorio, una alta fecundidad, abundantes huevos y larvas pelágicas, es que se tiende a reducir el aislamiento genético entre las poblaciones naturales en mar (Wirgin y Waldman, 2005)

Uno de los objetivos de esta investigación era identificar a los progenitores de la población F1 mediante el uso de marcadores genéticos de tipo microsatélites, sin embargo, este objetivo no se logró desarrollar al no tener identificado el sexo de los progenitores, por esta razón tampoco se pudo calcular el parámetro de tamaño efectivo de la población. Se logró realizar una primera aproximación, en donde se relacionó la población de reproductores antiguos con la población F1, encontrándose cinco individuos reproductores de sexo indefinido con mayor cercanía a los individuos F1, sin embargo esto no quiere decir que estos individuos sean los

progenitores ya que en la toma de muestras no estaban todos los reproductores, porque el muestreo fue realizado años posteriores a la generación de la población F1, por lo que varios reproductores habían muerto y quizás eran los progenitores de esta población. Este resultado se relaciona con el obtenido al analizar la similitud genética y distancia estándar de Nei, en donde la población de reproductores antiguos con la F1 no presentaron valores altos de similitud genética, por el contrario, la población F1 presentó una alta similitud con la población de Coquimbo. Nei en 1973, determinó el parámetro de distancia estándar, un valor 0 indica que no hay diversidad genética entre las dos poblaciones, mientras que el valor 1 indica que las dos poblaciones son genéticamente diferentes. La población de reproductores antiguos con los individuos F1 presentaron un valor de distancia estándar de 0.394 y 0.675 de similitud genética, por lo que estas poblaciones no estarían tan relacionadas genéticamente, lo que reafirmaría la hipótesis de que los progenitores de la población F1 no fueron analizados (Tabla 7).

También se evaluó el contenido de información polimórfico (*PIC*) de los doce *loci* en estudio para determinar qué tan informativos eran los marcadores seleccionados. Botstein y colaboradores (1980) indicaron que un *locus* era polimórfico cuando $PIC > 0.5$, medianamente polimórfico cuando $0.5 < PIC < 0.25$ y bajamente polimórfico cuando $PIC < 0.25$. Los valores de *PIC* variaban entre 0.1 y 0.8, en donde ocho *loci* resultaron ser marcadores informativos ya que lograban detectar un polimorfismo en la población al tener $PIC \leq 0.5$, siendo el *locus Gch18* el que presentaba el valor más alto de *PIC* (Tabla 8), por lo que la búsqueda y caracterización de *loci* polimórficos fue realizada de forma correcta, ya que se logró identificar marcadores que permitiesen detectar un polimorfismo en la población de congrios colorados. Estos marcadores podrían ser utilizados para analizar otras especies del género *Genyperus*, ya que las regiones flanqueantes al microsatélite suelen ser conservados entre las especies cercanas (Cipriani *et al.*, 1999) o en algunos casos existe amplificación en especies distantes (González-Martínez *et al.*, 2004).

Otro parámetro para la estimación de identificación de un individuo es la probabilidad de identidad (*PI*), o la probabilidad de que dos individuos de una población elegidos de forma aleatoria, tengan el mismo genotipo utilizando los mismos *loci* (Waits *et al.*, 2001). La población de reproductores antiguos era la que presentaba menor probabilidad de encontrar

individuos con genotipo idéntico ya que los valores de PI eran los más bajos ($PI=1.5E-01 - 3.2E-14$), mientras que la población de reproductores nuevos era la que tenía mayor probabilidad ($PI=6.3E-01 - 1.2E-05$), esto confirma que la población de reproductores antiguos fue la de mejores parámetros de variabilidad genética (Figura 9, Tabla 9). Este estimador teórico se utiliza cada vez con mayor confianza para la identificación de un individuo (Mills *et al.*, 2000). En Estados Unidos, en peritajes forenses, utilizan un set de 13 marcadores de tipo microsatélite para la identificación de un individuo (Hares, D. R., 2012), mientras que en Europa utilizan entre 4-7 marcadores (Gill *et al.*, 2006). Al analizar el parámetro de probabilidad de identidad de las cuatro poblaciones, se obtuvo que con tan solo seis marcadores ya es posible identificar a un individuo de congrio colorado por lo que la búsqueda y caracterización de microsatélites provenientes del transcriptoma fue realizada con éxito.

6. CONCLUSIÓN

En este estudio hemos genotipificado por primera vez una población completa de *G. chilensis* utilizando microsatélites. Mediante la búsqueda y caracterización de microsatélites en el transcriptoma de congrio colorado, se aporta con un alto porcentaje de genes con microsatélites. Se identificaron doce marcadores polimórficos que permitieron analizar de forma óptima a cuatro poblaciones de esta especie. Se logró determinar parámetros de variabilidad genética de las poblaciones que permitió descartar el fenómeno de consanguinidad en los reproductores antiguos al presentar parámetros óptimos de variabilidad genética, por lo que la mortalidad larval no es consecuencia de la escasa diversidad genética de los reproductores; por otro lado, gracias a estos marcadores, se identificó que los nuevos reproductores no presentan parámetros ideales de variabilidad genética, por lo que se sugiere incorporar nuevos ejemplares. Seis de los doce *loci* son altamente polimórficos y permiten una alta probabilidad de identidad por lo que éstos marcadores podrían ser utilizados en un PCR Multiplex, permitiendo un análisis más eficiente en a los ejemplares de congrio colorado.

Estos resultados muestran que los marcadores microsatélites transcriptómicos pueden detectar la estructura de la población en *G. chilensis* y pueden utilizarse como marcadores moleculares para estudios genéticos poblacionales. La información sobre la variación genética, la diferenciación del cultivo y las poblaciones silvestres del congrio colorado será útil para el futuro mejoramiento genético, mediante la cría selectiva, y para el diseño de directrices adecuadas para la gestión de estos materiales genéticos. Además del análisis de la diversidad genética, los marcadores de genotipo también podrían utilizarse como marcadores moleculares para el manejo de reproductores, la identificación de individuos y test de paternidad, también podrían ser utilizados para analizar otras especies del género *Genypterus* como el congrio dorado (*Genypterus blacodes*) y congrio negro (*Genypterus maculatus*).

7. REFERENCIAS

Abernathy J., Peatman E. y Liu Z. (2010). Basic Aquaculture Genetics. SRAC publication No. 5001.

Aedo J.E., Maldonado J., Estrada J.M., Fuentes E.N., Silva H., Gallardo-Escarate C., Molina A. y Valdés J.A. (2014) Sequencing and de novo assembly of the red cusk-eel (*Genypterus chilensis*) transcriptome. Mar. Genomics 18 Pt B, 105–107.

Aedo J.E., Maldonado J., Aballai V., Estrada J.M., Bastias-Molina M., Meneses C., Gallardo-Escarate C., Silva H., Molina A. y Valdés J.A. (2016). mRNA-seq reveals skeletal muscle atrophy in response to handling stress in a marine teleost, the red cusk-eel (*Genypterus chilensis*). BMC Genomics 16:1024.

Allendorf F.W. and Luikart G. (2007). Conservation and the Genetics of Populations. Malden: Blackwell Publishing.

Almuly R., Poleg-Danin Y., Gorshkov S., Gorshkova G., Rapoport B., Soller M., Kashi Y. y Funkenstein B. (2005). Characterization of the 5' flanking region of the growth hormone gene of the marine teleost, gilthead sea bream *Sparus aurata*: analysis of a polymorphic microsatellite in the proximal promoter. Fisheries Science 71: 479-490.

Alvial, A., y Manríquez, J. (1999). Diversification of flatfish culture in Chile. Aquaculture, 176: 65-73.

Avilés S. (1979). Congrio colorado, *Genypterus chilensis* (Guichenot) Teleostomi, Perciformes, Ophidiidae. En: IFOP (ed). Estado actual de las principales pesquerías nacionales. Bases para un desarrollo pesquero. Tomo I, Peces 14: 1-13.

Balloux F. y Lugon-Moulin N. (2002). The estimation of population differentiation with microsatellite markers. *Mol. Ecol.* 11: 155-165.

Beaumont, M.A. (2005). Adaptation and speciation: what can FST tell us? *Trends Ecol. Evol.* 20: 435-440.

Beckmann J. y Weber J. (1992). Survey of human and rat microsatellites. *Genomics* 12: 627-631.

Beebe T. y Rowe G. (2008). *An Introduction to Molecular Ecology*. Oxford University Press, Oxford. 400pp.

Berg, I., Neumann, R., Cederberg, H., Rannug, U. y Jeffreys, A. J. (2003). Two modes of germline instability at human minisatellite MS1 (*locus* D1S7): complex rearrangements and paradoxical hyperdeletion. *Am. J. Hum. Genet.* 72:1436–1447.

Bentzen, P., Olsen, J.B., McLean, J.E., Seamons, T.R. y Quinn, T.P. (2001) Kinship analysis of Pacific salmon: insights into mating, homing, and timing of reproduction. *J. Heredity* 92, 127–136.

Botstein D., White R.L., Skolnick M. y Davis R.W. (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. American journal of human genetics, Vol. 32, No 3, p. 314.

Brennan C. y Steitz J. (2001). HuR and mRNA stability. CMLS, Cell. Mol. Life Sci. 58: 266–277.

Buschmann A., Lopez D. y Medina A. (1996). A review of the environmental effects and alternative production strategies of marine aquaculture in Chile. Aquacultural Engineering vol. 15. No 6: 397-421.

Canales-Aguirre C.B., Ferrada S., Hernández C.E. y Galleguillos R. (2010). Population structure and demographic history of *Genypterus blacodes* using microsatellite *loci*. Fish Res. Elsevier B.V.; 106: 102–106.

Canales-Aguirre C., Ferrada S., Hernandez C. y Galleguillos R. (2010) Usefulness of heterologous microsatellites obtained from *Genypterus blacodes* (Schneider 1801) in species *Genypterus* of Southeast Pacific. Gayana 71(1): 74-77.

Castro J., Bouza C., Presa P., Pino-Querido A., Rianza A., Ferreiro I., Sánchez L. y Martínez P. (2004). Potential sources of error in parentage assessment of turbot (*Scophthalmus maximus*) using microsatellite *loci*. Aquaculture 242: 119-135.

Chauhan T. y Rajiv K. (2010). Molecular markers and their applications in fisheries and aquaculture. Advances in bioscience and biotechnology 1: 281-291.

Chen J., Cooper D., Chuzhanova N., Férec C. y Patrinos G. (2007). Gene conversion: mechanisms, evolution and human disease. *Nature* vol 8 762-775.

Chew N., Oliveira C., Wright J. y Dobson M. (2002). Molecular and cytogenetic analysis of the telomeric (TTAGGG)_n repetitive sequences in the Nile tilapia *Oreochromis niloticus* (Teleostei: Cichlidae). *Cromosoma* 111: 45-52.

Chistiakov D., Hellemans B. y Volckaert F. (2006). Microsatellites and their genomic distribution, evolution, function and applications: A review with special reference to fish genetics. *Aquaculture* 255 1-29.

Chocair J., Orellana F. y Serra J. (1969). Estudio del género *Genypterus* (congrios) en aguas chilenas (Pisces: Ophidiidae). Tesis de Biología Marina, Universidad de Chile, Valparaíso, 48.

Chong J. y González P. (2009). Ciclo reproductivo y talla media de madurez del congrio colorado, *Genypterus chilensis* (Guichenot, 1881) en el litoral de Talcahuano, Chile. *Rev. Biol. Mar. Oceanogr.*, 44 (1): 257-262.

Ciftci Y. y Okumus I. (2002). Fish Population Genetics and Applications of Molecular Markers to Fisheries and Aquaculture: I- Basic Principles of Fish Population Genetics. *Turkish Journal of Fisheries and Aquatic Sciences* 2: 145-155.

Cipriani G., Lot G., Huang W-G., Marrazo M., Peterlunger E. y Testolin R. (1999). AC/GT and AG/CT microsatellite repeats in peach [*Prunus persica* (L.) Batsch]: isolation, characterisation and cross-species amplification in *Prunus*. *Theor Appl Genet.* 99: 65 - 72.

Coimbra M., Kobayashi, K., Koretsugu S., Hasegawa O., Ohara E. y Ozaki A. (2003). A genetic linkage map of the Japanese flounder, *Paralichtys olivaceus*. *Aquaculture* 220: 203–218.

Crooijmans R.P.M.A., Poel J.J.V., Groenen M.A.M., Bierbooms V.A.F. y Komen J. (1997) Microsatellite markers in common carp (*Cyprinus carpio L.*) *Animal Genetics* 28: 129-134.

Crow J.F. y Kimura M. (1970). *An Introduction to Population Genetics Theory*. Harper & Row, New York.

DeWoody J.A. y Avise J.C.(2000). Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. *J. Fish Biol.* 56:461-473.

Doyle J.M., Siegmund G., Ruhl J.D., Eo S.H., Hale M.C. y Marra N.J. (2013). Microsatellite analyses across three diverse vertebrate transcriptomes (*Acipenser fulvescens*, *Ambystoma tigrinum*, and *Dipodomys spectabilis*). *Genome*; 56(7):407–414.

Edwards Y.J., Elgar G., Clark M.S. y Bishop M.J. (1998). The identification and characterization of microsatellites in the compact genome of the Japanese pufferfish, *Fugu rubripes*: perspectives in functional and comparative genomic analysis. *J. Mol. Biol.* 278: 843–854.

Ellegren H. (2000). Microsatellite mutations in germline: implications for evolutionary inference. *Trends Genet.* 16:551-558.

Ellegren H. (2004). Microsatellites: simple sequence with complex evolution. *Nature* vol. 5 435-445.

Ellis J. R. y Burke J. M. (2007). EST-SSRs as a resource for population genetic analyses. *Heredity*. 99: 125-132.

FAO. 2008. The State of World Fisheries and Aquaculture.

FAO. 2016. The State of World Fisheries and Aquaculture.

Frankham, R., Ballou J.D. y Briscoe, D.A. (2002). Introduction to conservation genetics. Cambridge University Press. U.K.

Frankham, R., Ballou J.D. y Briscoe, D.A. (2004) A Primer of Conservation Genetics, Cambridge: Cambridge University Press.

George-Nascimento M., Lobos V., Torrijos C. y Rasul Khan. (2004). Species composition of assemblages of ceratomyxa (*Myxozoa*), parasites of the lings *Genypterus* (ophiidae) in the southeastern pacific ocean: an ecomorphometric approach. *Journal of Parasitology*. Vol. 90, No. 6, pp. 1352-1355.

Gerhard G., Kauffman E. y Grundy M. (2000). Molecular cloning and sequence analysis of the *Danio rerio catalase* gene. *Biochem Mol. Biol*, 127: 447-457.

Gill P., Fereday L., Morling N. y Schneider P. (2006). The evolution of DNA databases recommendations for new European STR *loci*. *Forensic Sci. Int.* 2006;156:242–244.

Gjdrem, T. (2000). Genetic improvement of cold-water fish species. *Aquaculture* 31, 25-33.

González-Martínez S., Robledo-Arnuncio J., Collada C., Díaz A., Williams C., Alía R. (2004). Cross-amplification and sequence variation of microsatellite *loci* in Eurasian hard pines. *Theor Appl Genet.* 109:103-111.

González, E. y R. Zardoya. (2007). Relative role of life-history traits and historical factors in shaping genetic population structure of sardines (*Sardina pilchardus*). *Evolutionary Biology.* 7: 197.

Grimholt U., Drablos F., Jorgensen S., Hoyheim B. y Stet R. (2002). The major histocompatibility class I *locus* in Atlantic salmon (*Salmo salar L.*): polymorphism, linkage analysis and protein modeling. *Immunogenetics* 54: 570-581.

Guichoux E., LAGache L., Wangner S., Chaumeil P., Leger P., Lepais O., Lepoittevin C., Malausa T., Revardel E., Salin F. y Petit R. (2011). Current trends in microsatellite genotyping. *Molecular Ecology Resources* 11:591-611.

Gur-Arie R., Cohen C., Eitan L., Shelef E., Hallerman E. y Kashi Y. (2000). Abundance, non-random genomic distribution, nucleotide composition and polymorphism of simple sequence repeats in *E. coli*. *Genome Res.* 10: 61-70.

Hares D. (2012). Expanding the CODIS Core *Loci* in the United States. *Forensic Sci Int Genet.* 6, e52–54.

Hauser L. y Ward R.D. (1998). Population identification in pelagic fish: the limits of molecular markers. In: Carvalho, G.R. (Ed.). *Advances in Molecular Ecology.* IOSPress, Amsterdam, pp. 191–224.

Hearne C. M., Ghosh S. y Todd J. A. (1992). Microsatellites for linkage analysis of genetic traits. *Trends in Genetics*. Vol. 8, No 8, p. 288-294.

Herbinger C.M., Doyle R.W., Pitman E.R., Paquet D., Mesa K.A., Morris D.B., Wright J.M. y Cook D. (1995). DNA fingerprint based analysis of parental and maternal effects on offsprings growth and survival in communally reared rainbow trout. *Aquaculture* 137, 245-56.

Herlin M., Taggart J., McAndrew B. y Penman D. (2007). Parentage allocation in a complex situation: A large commercial Atlantic cod (*Gadus morhua*) mass spawning tank. *Aquaculture* 272S1 S195–S203.

Huang D. y Bernardi G. (2001). Disjunct Sea of Cortez-Pacific Ocean *Gillichthys mirabilis* populations and the evolutionary origin of their Sea of Cortez endemic relative, *Gillichthys seta*. *Marine Biology*, 138, 421–428.

Kayser M., Roewer L., Hedman M., Henke L., Henke J., Brauer S., Kruger C., Krawczak M., Naggy M., Dobosz T., Sibor R., Knijff P., Stoneking M. y Sajantila A. (2000). Characteristics and frequency of germline mutations at microsatellite *loci* from the human Y chromosome, as revealed by direct observation in father/son pairs. *Am. J. Hum. Genet.* 66: 1580–1588.

Kelly R.P. y Palumbi S.R. (2010). Genetic structure among 50 species of the northeastern Pacific rocky intertidal community. *PLOS One*, 5, e8594.

Knutsen H., P.E. Jorde, C. André y C. Stenseth. (2003). Fine-scaled geographical population structuring in a highly mobile marine species: the Atlantic cod. *Mol. Ecol.* 12: 385-394.

Kong I., Meléndez R. y Henríquez G. (1988). Los peces Ophidiformes de aguas profundas entre Arica e isla Mocha. *Estud. Oceanol.*, 7: 1-15.

Levinson G. y Gutman G. (1987). High frequencies of short frameshifts in poly-CA/TG tandem repeats borne by bacteriophage M13 in *Escherichia coli* K-12. *Nucleic Acids Research* vol.15 No 13.

Li D., Yang H., Zou Z., Xiao W., Zhu J. y Luo Y. (2015). Polymorphic microsatellite differences among four cultured populations of two selected tilapia strains. *Aquaculture Research*, 46:492-498.

Li Y., Korol A., Fahima T. y Nevo E. (2004). Microsatellites within genes: structure, function and evolution. *Mol. Biol. Evol.* 21: 991-1007.

Liu Z. y Cordes J. (2004). DNA marker technologies and their applications in aquaculture genetic. *Aquaculture* 238:1-37.

Liu Z., Li P., Kocabas A., Karsi A. y Ju Z. (2001). Microsatellite-containing genes from channel catfish brain: Evidence of trinucleotide repeat expansion in the coding region of nucleotide excision repair gene *rad23b*. *Biochem Biophys Res Commun* 289: 317-324.

Lopez D., Lopez B., Arriagada S., Mora O., Bedecarratz P., Pineda M., González M., Andrade L., Uribe J. y Riquelme V. (2012). Diversification of Chilean aquaculture: the case of the giant barnacle *Austromegabalanus psittacus* (Molina, 1782). *Latin American Journal of Aquatic Research* 40: 596-607.

McCune A.R., Houle D., McMillan K., Annable R., Kondrashov A.S. (2004). Two classes of deleterious recessive alleles in a natural population of zebrafish, *Danio rerio*. *Proc Biol Sci.* 271(1552):2025–2033.

Melloul D., Marshak S. y Cerasi E. (2002). Regulation of the insulin gene transcription. *Diabetologia* 45: 309-326.

Metzgar D., Bytof J. y Wills C. (2000). Selection against frameshift mutations limits microsatellite expansion in coding DNA. *Genome Research* 10, 72-80.

Metzker M.L. (2010) Sequencing technologies, the next generation. *Nature Reviews Genetics* 11, 31-46.

Mills L., Citta J., Lair K., Schwartz M. y Tallmon D. (2000). Estimating animal abundance using non-invasive DNA sampling: promise and pitfalls. *Ecological Applications*, 10, 283–294.

Moen T., Hoyheim B., Munck H. y Gomez-Raya L. (2004). A linkage map of Atlantic salmon (*Salmo salar*) reveals an uncommonly large difference in recombination rate between the sexes. *Anim. Genet.* 35: 81-92.

Morozova O. y Marra M.A. (2008). Applications of next-generation sequencing technologies in functional genomics. *Genomics* 92, 255-264.

Nei M. (1977). F-statistics and analysis of gene diversity in subdivided populations. *Annals of Human Genetic*, 41:225-233.

Nichols K., Young W., Danzmann G., Robison D., Rexroad C., Noakes M., Phillips B., Bentzen, P., Spies I., Knudsen K., Allendorf W., Cunningham B.M., Brunelli J., Zhang H., Ristow S., Drew R., Brown K.H., Wheeler A. y Thorgaard G. (2003). A consolidated linkage map for rainbow trout (*Oncorhynchus mykiss*). *Anim. Genet.* 34:102–115.

Niegel J. (1997). A comparison of alternative strategies for estimating gene flow from genetic markers. *Ann Rev ecol syst.* 28:105-128.

Palace V.P. y Werner J. (2006). Vitamins A and E in the maternal diet influence egg quality and early life stage development in fish: a review. *Sci Mar* **70S2**, 41–57.

Powel W., Machray G. y Provan J. (1996). Polymorphism revealed by simple sequence repeats. *Trends Plant Sci.* 1: 215-222.

Pradet-Balade B., Salmon C., Hardy A. y Querat B. (1998). Heterogeneity of eel thyrotropin beta mRNA is due to a minisatellite in the 3' untranslated region of the gene. *Gene* 215: 251-257.

Rocha-Olivares A., Bobadilla-Jiménez M., Ortega-García S., Saavedra-Sotelo N. y Sandoval-Castillo J.R. (2006). Mitochondrial variability of dolphinfish *Coryphaena hippurus* populations in the Pacific Ocean. *Ciencias Marinas.* 32(3): 569-578.

Sharma P.C., Grover A. y Kahl G. (2007). Mining microsatellites in eukaryotic genomes. *Trends in Biotechnology* 25, 490-498.

Sia E., Koskoska R.J., Dominska M., Greenwell P. y Petes T.D. (1997). Microsatellite instability in yeast: dependence on repeat unit size and mismatch repair genes. *Molecular and Cellular Biology* 17: 2851-2858.

Sommani A. y Jindathip R. (2003). Effect of vitamin C on growth, mortality and deformities of Nile tilapia larvae. *Thai Fisheries Gazette*. 56 (1): 29-33.

Toth G., Gaspari Z. y Jurka J. (2000). Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Res.* 10, 967-981.

Tran N.T., Gao Z.X., Zhao H.H., Yi S.K., Chen B.X., Zhao Y.H., Lin L., Liu X.Q. y Wang W.M. (2015). Transcriptome analysis and microsatellite discovery in the blunt snout bream (*Megalobrama amblycephala*) after challenge with *Aeromonas hydrophila*. *Fish Shellfish Immunol.* 45, 72–82.

Treco D., y Arnheim N. (1986). The evolutionarily conserved repetitive sequence d(TG.AC)_n promotes reciprocal exchange and generates unusual recombinant tetrads during yeast meiosis. *Molecular and Cell Biology* 6:3934-3947.

van Lith H. y Van Zutphen L. (1996). Characterization of rabbit DNA microsatellites extracted from EMBL nucleotide sequence database. *Anim. Genet.* 27: 387-395.

van Oosterhout C., Hutchinson W.F., Wills D.P.M. y Shipley P. (2004). Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes*; 4:535–538.

Vega R., Pradenas M., Estrada J., Ramirez D., Valdevenito I., Mardones A., Dantagnan P., Alfaro D., Encina F. y Pichara C. (2012). Evaluation and comparison of the efficiency of two incubation systems for *Genypterus chilensis* (Guichenot, 1848) eggs. Latin American Journal of Aquatic Research 40, 187-200.

Waits L.P., Luikart G. y Taberlet P. (2001). Estimating the probability of identity among genotypes in natural populations: Cautions and guidelines. Molecular Ecology, 10, 249–256.

Wall P.K., Leebens-Mack J., Chanderbali A.S., Barakat A., Wolcott E., Liang H., Landherr L., Tomsho, L.P., Hu Y., Carlson J.E., Schuster S.C., Soltis D.E., Soltis P.S., Altman N. y dePamphilis C.W. (2009). Comparison of next generation sequencing technologies for transcriptome characterization. BMC Genomics 10, Article no.347.

Ward R.D., Woodwark M. y Skibinski D.O.F. (1994). A comparison of genetic diversity levels in marine, freshwater and anadromous fish. J. Fish Biol. 44, 213–232.

Ward R.D. (2006). The importance of identify spatial population structure in restocking and stock enhancement programmes. Fish. Res. 80: 9-18.

Ward R. y A. Reilly. (2001). Development of microsatellite *loci* for population studies of the pink ling, *Genypterus blacodes* (Teleostei: Ophidiidae). Molecular Ecology Notes 1: 173-175.

Weber J.L. (1990). Informativeness of human (dC-dA)_n (dG-dT)_n polymorphisms. Genomics 7:524-530.

Wirgin I. y Waldman J.R., (2005). Use of nuclear of DNA in stock identification: single-copy and repetitive sequence markers. In: Cadrin, S.X., Friedland, K.D., Waldman, J.R. (Eds.), Stock Identification Methods: Applications in Fishery Science. Elsevier Academic Press, Oxford, UK, pp. 331–370.

Wren J.D., Forgacs E., Fondon J.W., Pertsemlidis A., Cheng S.Y., Gallardo T., Williams R.S., Shohet R.V., Minna J.D. y Garner H.R. (2000). Repeat polymorphisms within gene regions: phenotypic and evolutionary implications. *American Journal of Human Genetics* 67, 345-356.

Wright S. (1951). The genetical structure of population. *Annals of Eugenics*. 15. 323-354.

Wright S. (1978). *Evolution and the Genetics of Populations - variability within and among natural populations*. The University of Chicago Press; Chicago, USA.

Wright J. (1993). DNA fingerprinting in fishes. *Molecular Biology of fishes* 58-91.

Zalapa J.E., Cuevas H., Steffan S., Senalik D., Zeldin E., McCown B., Harbut R. y Simon P. (2012). Using next-generation sequencing approaches to isolate simple sequence repeat (SSR) *loci* in the plant sciences. *Amer. J. Bot.* 99, 193–208.

Zane L., Bagelloni L. y Patarnello T. (2002). Strategies for microsatellite isolation: a review. *Molecular Ecology* 11, 1-16.

