



UNIVERSIDAD DE CHILE –FACULTAD DE CIENCIAS –ESCUELA DE PREGRADO

**“Identificación y caracterización de SNPs en individuos de  
*Basilichtys microlepidotus* habitando sitios contaminados.”**

**Jorge Cortés Miranda**

Director del Seminario de Título: Dra. Caren Vega Retter

Agosto 2019

Santiago, Chile

## BIOGRAFÍA



Jorge Alejandro Cortés Miranda, nació el 2 de Septiembre de 1995 en la ciudad de Santiago de Chile. Ingreso el año 2010 al liceo Lastarria, donde completo su enseñanza media y comenzó a desarrollar su interés por el área biológica participando del electivo de esta área. En 2014 hace ingreso a la carrera de Ingeniería en Biotecnología Molecular en la Facultad de Ciencias de la Universidad de Chile. En el segundo año de la carrera conoce a la Dra. Caren Vega Retter quién le enseñaría y sería su tutora hasta el día de hoy. Además ese mismo año asume como delegado de carrera. En el tercer año de carrera asumiría como consejero estudiantil en el consejo de Escuela de Pregrado. Durante los últimos años de la carrera continuaría aprendiendo y desarrollando habilidades en el Laboratorio de Ecología y Genética, además incursionaría en el mundo de la difusión científica, formando parte de la organización GECO, constituida por estudiantes de la Facultad de Ciencias con el fin de promover la educación científica fuera de la facultad.

## **DEDICATORIA**

Dedicado a mi familia, mis amigos y amigas, a las buenas personas que conocí en esta facultad y a la preciosa kitty.

## **AGRADECIMIENTOS**

Agradezco a la gente que fue parte de este proceso dentro de la carrera, a aquellos amigos y amigas que conocí a lo largo de estos años y que siempre nos brindamos apoyo cuando lo necesitamos.

Agradezco a mi familia que siempre me apoyo en este camino de diversas formas.

Agradezco a mi tutora Caren, que desde mis primeros años en esta facultad me ha enseñado la mayor parte de cosas que sé, sobre el mundo académico y sobre ser científico.

Agradezco especialmente a todos mis compañeros y compañeras del laboratorio, uno de los lugares más amenos dentro de esta facultad, creado únicamente gracias a su gran disposición y buen humor. Gracias por enseñarme tanto y siempre tenderme una mano.

Agradezco al equipo de CREAR de la Universidad San Sebastián, quienes me brindaron todo su apoyo en esta etapa final y que me han hecho crecer mucho, tanto en lo académico y lo pedagógico, como en lo personal.

Finalmente agradezco a todas aquellas personas que de una u otra forma estuvieron presentes dentro del proceso y que hoy por diversos motivos no nos seguimos acompañando, pero que siguen muy presentes en mis recuerdos.

Gracias a todos y todas por hacer este proceso mucho más ameno y disfrutable.

## ÍNDICE DE CONTENIDOS

<b>RESUMEN</b> .....	1
<b>ABSTRACT</b> .....	3
<b>INTRODUCCIÓN</b> .....	5
<b>MATERIALES Y MÉTODOS</b> .....	11
<b>RESULTADOS</b> .....	18
<b>DISCUSIÓN</b> .....	30
<b>CONCLUSIÓN</b> .....	40
<b>BIBLIOGRAFÍA</b> .....	41
<b>ANEXO</b> .....	48

## ÍNDICE DE TABLAS

<b>Tabla 1: Número de lecturas y bases posteriores al filtro.....</b>	<b>18</b>
<b>Tabla 2: Número de SNPs detectados por comparación.....</b>	<b>20</b>
<b>Tabla 3: Cambio en aminoácidos producto de los SNPs.....</b>	<b>22</b>
<b>Tabla 4: Anotación de contigs con SNPs presentes en comparación PEL-SFM.....</b>	<b>28</b>

## ÍNDICE DE FIGURAS

<b>Figura 1: Mapa de los sitios muestreados .....</b>	<b>12</b>
<b>Figura 2: Diagrama de Venn de contigs compartidos entre comparaciones .....</b>	<b>21</b>
<b>Figura 3: Número de contigs en función de la densidad de SNPs.....</b>	<b>23</b>
<b>Figura 4: Gráfico de GO terms enriquecidos en comparación IM-SFM.....</b>	<b>25</b>
<b>Figura 5: Gráfico de GO terms enriquecidos en comparación MEL-SFM.....</b>	<b>27</b>
<b>Figura A1. Procesos Biológicos enriquecidos comparación IM-SFM.....</b>	<b>48</b>
<b>Figura A2. Componentes celulares enriquecidos en comparación IM-SFM.....</b>	<b>49</b>
<b>Figura A3. Funciones moleculares enriquecidas en comparación IM-SFM.....</b>	<b>50</b>
<b>Figura A4. Procesos biológicos enriquecidos en comparación MEL-SFM.....</b>	<b>51</b>
<b>Figura A5. Componentes celulares enriquecidos en comparación MEL-SFM.....</b>	<b>52</b>
<b>Figura A6. Funciones moleculares enriquecidas en comparación MEL-SFM.....</b>	<b>53</b>

## RESUMEN

La contaminación de los sistemas dulceacuícolas ha generado un gran impacto sobre las poblaciones que habitan en ellos, desde fragmentación de las poblaciones hasta cambios genéticos. Dentro de este ámbito se ha observado que la contaminación puede actuar como fuerza de selección, cambiando las frecuencias alélicas de distintos genes en las poblaciones que habitan sitios contaminados con respecto a las poblaciones que habitan sitios no contaminados. Mediante la utilización de técnicas de secuenciación masiva, tales como RNA-seq ha sido posible observar selección de alelos con hasta una base de diferencia, cuya denominación es polimorfismos de nucleótido simple o SNP.

La cuenca del Río Maipo ha sido una de las más afectadas por la contaminación en Chile, debido al gran número de actividades humanas que se realizan en torno a ella y el gran número de personas que la habitan. En esta cuenca habita el pejerrey endémico *Basilichtys microlepidotus*, registrándose 5 poblaciones, 2 habitando sitios contaminados y 3 habitando sitios no contaminados. El principal objetivo de este trabajo fue la detección y caracterización funcional de SNPs, de *B. microlepidotus*, asociados a la contaminación, al comparar hígados de individuos de sitios contaminados con individuos de un sitio no contaminado. Para esto se compararon individuos de 3 sitios contaminados (IM-MEL-PEL) con individuos de un sitio no contaminado (SFM).

En la comparación IM-SFM se detectaron SNPs asociados a procesos biológicos relacionados con obtención de energía, más concretamente en genes relacionados con la cadena transportadora de electrones. En el caso de la comparación MEL-SFM, también se



observaron SNPs asociados a procesos biológicos relacionados con la obtención de energía y SNPs en genes relacionados con la mantención de la osmoregulación en particular en el gen de la subunidad alfa de la ATPasa sodio potasio. En la comparación PEL-SFM no fue posible realizar un análisis de enriquecimiento, pero también se observaron SNPs asociados al gen de la ATPasa sodio potasio. Estas funciones se han visto alteradas en otros organismos que han sido sometidos a condiciones de contaminación y podrían dar cuenta de selección de mecanismos que permitan sobrellevar el estrés asociado a contaminación en *B. microlepidotus*.

## ABSTRACT

Pollution in freshwater systems has generated a large impact on populations inhabiting these systems, from population fragmentation to genetic changes. It has been observed that pollution could act like selective pressure, modifying genes allelic frequencies in populations who inhabit polluted sites versus populations who inhabit non polluted sites. By using massive sequencing technologies, like RNA-seq, it has been possible to observe allele selection with only differences in one nucleotidic base, which has been denominated as Single Nucleotide Polymorphism or SNP.

The Maipo River basin has been one of the most affected basin by pollution in Chile, due to the high number of human activities that are carried out around it and due to the large human population that inhabit this basin. The endemic silverside *Basilichtys microlepidotus*, inhabits this basin and 5 populations has been registered, 2 inhabiting polluted sites and 3 inhabiting non-polluted sites. The main goal of this research was the detection and functional characterization of SNPs of *B. microlepidotus*, associated to pollution. For this, livers of individuals inhabiting polluted sites (IM-MEL-PEL) were compared with livers of individuals inhabiting a non-polluted site (SFM).

In IM-SFM comparison, SNPs associated to biological process related with energy obtainment, more specifically in genes related to electron transport chain were detected. In MEL-SFM comparison, SNPs associated to biological process related with energy obtainment were also observed, as well as SNPs in osmoregulation related genes, more specifically in alfa subunit of sodium potassium ATPase gene. Similarly, in PEL-SFM SNPs in the alfa subunit of the sodium potassium ATPase gene were observed. These

functions were also altered in other organisms living in polluted environments and could be related to the selection of a mechanism that allows to resist the pollution to which *B. microlepidotus* is exposed.

## INTRODUCCIÓN

Las actividades humanas han transformado los ecosistemas, tanto desde un punto de vista fisicoquímico (Bertness y cols., 2002), como a nivel de la biodiversidad (Dulvy y cols., 2003) e interacción entre organismos (Lürling & Scheffer, 2007). Dentro de los distintos factores que promueven estas alteraciones, se ha determinado que existen 5 que afectan a gran escala: la fragmentación del hábitat, la propagación de especies exóticas, la agricultura, la contaminación y el cambio climático (ver Sih y cols., 2011).

Los ecosistemas acuáticos son unos de los más afectados por los factores mencionados anteriormente. En particular en los ecosistemas de agua dulce, se ha documentado un aumento del impacto de estresores antropogénicos en lugares densamente poblados (Allan y cols., 2013). Uno de los estresores clave es la contaminación, debido a los distintos efectos que produce en estos ecosistemas y a que gran parte de las actividades industriales y domésticas se desarrollan en lugares cercanos a cursos de agua. En este contexto, las actividades asociadas a la agricultura han sido señaladas como unas de las que generan un mayor impacto (Cooper, 1993), observándose problemas en el desarrollo de los organismos debido a contaminantes provenientes de fertilizantes (De Wijer y cols., 2003), un aumento en interacciones de parasitismo en lugares contaminados con metales pesados (Poulin, 1992), disminución del tamaño corporal, cambios en los ciclos de vida (Moiseenko, 2002), cambios en interacciones ecológicas (Scott y Sloman, 2004) e incluso cambios a nivel genético en los organismos que habitan estos sitios (Mitra y cols., 2018). En relación a los cambios genéticos que se han documentado, podemos encontrar ejemplos como diferencias en los patrones de expresión génica (Mitra y cols., 2018), daño

mutagénico (Vanzella y cols., 2007) y selección direccional de determinados loci (Lind & Grahn., 2011; Vega-Retter y cols., 2015), entre otros. Podemos definir la selección direccional como aquella que lleva a la fijación de un alelo, es decir aumenta su frecuencia en una población en particular (Hedrick, 2007). La selección de loci producto de la contaminación ha sido observada tanto en experimentos con condiciones controladas de exposición a contaminantes, como también en experimentos de campo. Un ejemplo con condiciones controladas puede ser encontrado en Tatara y colaboradores (1999), donde utilizando la aproximación de un mesocosmo sometieron a poblaciones de *Gambusia holbrooki* a concentraciones elevadas de mercurio durante 4 generaciones, observándose selección de un alelo en particular del locus Gpi-2, gen implicado en la glucólisis y gluconeogénesis. Por otro lado Lind y Grahn (2011) utilizando la técnica AFLP, observaron que existe selección direccional de 8,4 % de los loci estudiados de *Gasterosteus aculeatus*, al comparar individuos habitando un sitio contaminado producto de una planta de celulosa e individuos de sitios no contaminados, pero debido al tipo de marcador utilizado no es posible determinar las funciones en las que están implicados estos genes.

En la actualidad es posible determinar selección de loci con una alta precisión, debido al desarrollo de marcadores moleculares que constan de diferencias encontradas en una sola base de una secuencia genética, los cuales son conocidos como polimorfismos de nucleótido simple o SNPs por sus siglas. La amplia utilización de estos marcadores ha sido posible gracias al desarrollo de las tecnologías de secuenciación masiva o NGS, las cuales han permitido obtener una gran cantidad de datos a un bajo costo (ver Petterson y

cols., 2009). Dentro de estas técnicas se encuentra RNA-seq, la cual permite obtener millones de copias de secuencias de RNA provenientes de un tejido (Hrdlickova et al., 2016). A partir de los datos obtenidos con esta técnica ha sido posible descubrir SNPs en distintos organismos y bajo distintas condiciones (Cánovas y cols., 2010; Salem y cols., 2012). Una ventaja de estas técnicas, dentro del estudio de los efectos genéticos de la contaminación, es que no es necesario escoger marcadores específicos relacionados con contaminación (ver Bozinovic & Oleksiak, 2011). Es importante mencionar que gran parte de los estudios realizados con RNA-seq asociados a contaminación utilizan tejido de hígado (ver Bozinovic & Oleksiak, 2011), debido a su ya conocida capacidad detoxificante y los numerosos procesos metabólicos que ocurren en él (Grant, 1991; Argiriadi y cols., 1999; Ikemoto y cols., 2004). En el caso particular de SNPs asociados a contaminación podemos mencionar el trabajo de Bélanger-Deschênes y colaboradores (2013), donde se encontró selección direccional de SNPs asociados a respiración celular en poblaciones de *Perca flavescens* habitando zonas contaminadas con metales pesados. Por otro lado Laporte y colaboradores (2015) observaron selección de SNPs asociados a vías metabólicas involucradas en la regulación de esterol, indicando que esto podría contribuir a la supervivencia de anguilas del Atlántico Norte en sitios contaminados con materia orgánica y metales pesados.

En Chile, las problemáticas asociadas a contaminación de sistemas de agua dulce no son menores, ya que muchas de las actividades industriales y agrícolas que se llevan a cabo en el país son desarrolladas en sectores cercanos a ríos y lagos (Narváez y cols., 2007; Ribbe y cols., 2008; Ribeiro y cols., 2014). Un caso particular de este problema se puede

evidenciar en la cuenca del Río Maipo, que ha sido afectada principalmente por contaminantes de origen doméstico (Dirección General de Aguas, 2004), lo que ha conllevado a un enriquecimiento en materia orgánica (Pardo y cols., 2008). Por otro lado también se han detectado distintos pesticidas (Retamal y cols., 2013) y acumulación de distintos metales pesados (Pizarro y cols., 2010).

En esta cuenca habita el pejerrey *Basilichthys microlepidotus* (Jenyns, 1841), pez endémico de Chile perteneciente a la familia *Atherinopsidae*, que habita en los sistemas de agua dulce desde la latitud 28°S a 39°S (Véliz y cols., 2012). Esta especie es omnívora, alimentándose principalmente de pequeños invertebrados, larvas de insectos, algas filamentosas y detritus (Duarte y cols., 1971). Desova entre los meses de agosto y enero (Comte y Vila, 1992), y actualmente su estado de conservación es considerado vulnerable (Ministerio del Medio Ambiente 2018). Se ha identificado previamente que existen al menos 5 poblaciones de *B. microlepidotus* habitando la cuenca del Río Maipo (Vega-Retter y cols., 2014), dos de las cuales habitan sitios caracterizados como contaminados: Melipilla (MEL) y Pelvin (PEL) y dos que habitan sitios caracterizados como no contaminados: San Francisco de Mostazal – Maipo (SFMMA) e Isla de Maipo - Peñaflores (IMPN). Sin embargo recientemente la calidad del agua en IM ha empeorado, por lo que ha sido catalogado como un sitio contaminado (datos no publicados, Vega-Retter y cols.). En un estudio previo utilizando AFLP se detectaron 8 loci bajo selección (5,37% de los loci totales analizados) al comparar individuos de sitios contaminados y no contaminados. Seis de estos loci están bajo selección direccional, de estos 3 se encontraron en un sitio contaminado (MEL) y los otros 3 en el otro sitio contaminado analizado (PEL), lo que

podría sugerir que debido a las diferencias fisicoquímicas de los sitios existe también una selección diferenciada de loci entre estos (Vega-Retter y cols., 2015). Complementario a lo anterior se ha observado que existe expresión diferencial de genes entre hígados de individuos de *B. microlepidotus* habitando sitios contaminados y no contaminados, dentro de los cuales se encuentran genes asociados a la gluconeogénesis y a cáncer. Uno de estos genes es el gen de la Ornitina descarboxilasa (ODC), la cual está implicada en la progresión de tumores. Además al secuenciar un segmento de la ODC en todos los individuos del estudio, se identificaron 6 alelos y se comprobó que en los sitios contaminados existe un déficit de heterocigotos y una alta frecuencia de un alelo en particular, con respecto a los sitios no contaminados (Vega-Retter y cols., 2018).

Dado estos antecedentes y con el objetivo de entender de manera más global la respuesta genética a la contaminación en este sistema, el propósito de este trabajo es detectar y caracterizar SNPs bajo selección en los sitios contaminados, a partir de datos de RNA-seq de hígado de *Basilichtys microlepidotus*.

**Hipótesis:** La contaminación se relaciona con la selección de alelos en el pejerrey *B. microlepidotus* habitando la cuenca del Río Maipo.

**Objetivo General:** Detectar y caracterizar funcionalmente SNPs presentes en hígados de individuos de *B. microlepidotus* habitando sitios contaminados y no contaminados.

**Objetivos Específicos:**

- a) Detectar SNPs entre hígados de individuos habitando un sitio no contaminado (SFM) y sitios contaminados (PEL, IM y MEL).

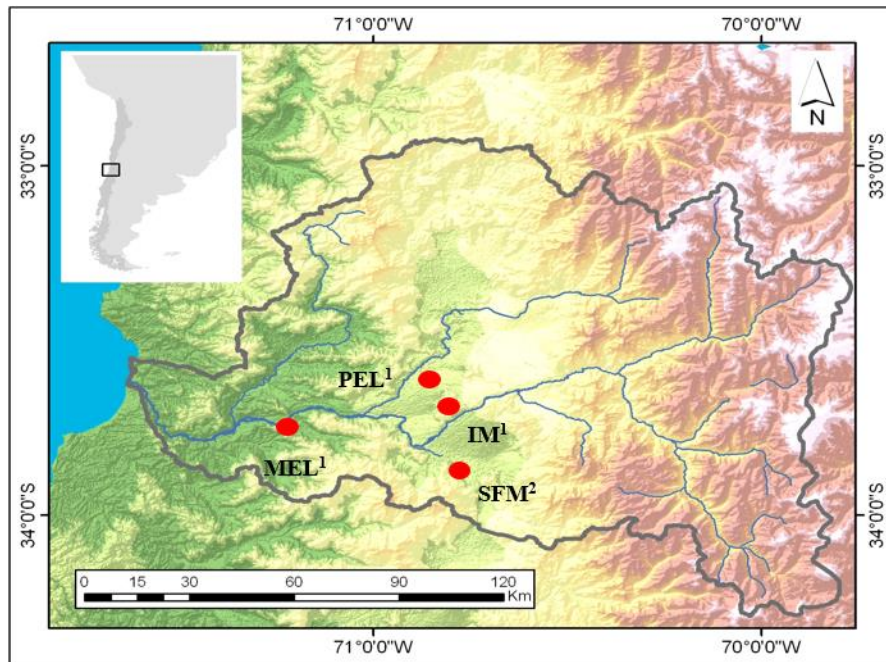


- b)** Determinar qué procesos biológicos, funciones moleculares o componentes celulares están sobre-representados en los SNPs detectados.

## MATERIALES Y MÉTODOS

### 1.- Sitios de muestreo y toma de muestras

Para este trabajo fueron escogidos 4 sitios de muestreo en la cuenca del Río Maipo, Melipilla (MEL), Pelvin (PEL), Isla de Maipo (IM) y San Francisco de Mostazal (SFM). En estudios previos 2 de estos sitios fueron catalogados como contaminados (MEL-PEL) y 2 como no contaminados (IM-SFM) (Vega-Retter y cols., 2014). Esta categorización se realizó utilizando la cuantificación de diferentes parámetros fisicoquímicos y utilizando el índice de calidad de agua de British Columbia (BCWQI, 1996). Sin embargo según datos recientes (2016), la calidad de agua de IM ha empeorado en el tiempo, siendo categorizado como “Marginal” según índice de British Columbia, al igual que los sitios MEL y PEL (datos no publicados), y dado que los individuos fueron muestreados en esa misma campaña, en este trabajo se considerará a IM como un sitio contaminado (Figura 1). Se muestrearon 6 individuos de *Basilichtys microlepidotus* por sitio y se les extrajo el hígado, el cual fue transportado en nitrógeno líquido hasta el laboratorio y posteriormente guardado a -80°C, hasta la extracción de RNA.



**Figura 1: Mapa de los sitios muestreados.** Se indican los sitios muestreados, Melipilla (MEL), Pelvin (PEL), Isla de Maipo (IM) y San Francisco de Mostazal (SFM). El superíndice en cada sitio indica si se trata de un sitio contaminado (1) o no contaminado (2). Imagen modificada de Vega-Retter y cols., (2018).

## **2.- Extracción de RNA, RNA-seq, control de calidad, ensamble *de novo* y mapeo de lecturas**

Se extrajo el RNA total del hígado de los 24 individuos, utilizando PureLink™ RNA Mini Kit (Ambion). Posteriormente se midió la cantidad y calidad del RNA extraído, mediante un Agilent Model 2100 Bioanalyzer. Sólo las muestras con un RIN > 7 fueron utilizadas en los pasos siguientes, debido a esto 3 individuos provenientes de PEL fueron descartados ya que no cumplían con los parámetros mínimos requeridos. Las muestras fueron

guardadas a  $-80^{\circ}\text{C}$  en agua libre de RNAsa y posteriormente enviadas al Centro de Genómica y Bioinformática “Genoma Mayor” de la Universidad Mayor, donde se purificó el RNA para obtener el RNAm, con el cual se realizó un RNA-sequencing paired-end en la plataforma Illumina Hiseq. Los datos brutos obtenidos se filtraron utilizando la herramienta Trim\_Galore (Krueger, 2012) y el script prinseq-lite.pl, se removieron los adaptadores, se eliminaron las lecturas con más de un 10% de bases ambiguas y con un tamaño menor a 50 pares de bases y se descartaron las lecturas que tuvieran una calidad menor a  $Q = 30$ , siendo  $Q = -\log_{10}P$  y  $P$  la probabilidad de que la base en esa posición haya sido leída de forma incorrecta. Al utilizar  $Q = 30$  las probabilidades de que la base leída sea incorrecta es de 1 entre 1000. A partir de estos datos filtrados se realizó un ensamble *de novo*, utilizando el software Bridger (Chang y cols., 2015), generando un total de 74.773 *contigs* de tamaño mayor a 200 pares de bases. Este ensamble fue usado como referencia para realizar un mapeo de las lecturas obtenidas de cada individuo utilizando el software Bowtie2 (Langmead & Salzberg, 2012) utilizando las siguientes opciones: `--no-mixed --no-discordant --end-to-end --no-head --no-sq - p 30 - S`, las opciones `--no-mixed` y `--no-discordant` permiten descartar lecturas que alinean en diferentes transcritos, `--end-to-end` reporta el mapeo sólo si las lecturas se alinean en un 100% y no solo en una parte de ella, mientras que `--no-head` y `--no-sq` eliminan información del archivo SAM resultante. Es importante mencionar que el filtro de los datos, ensamble *de novo* y mapeo fue realizado por el Centro de Genómica y Bioinformática “Genoma Mayor” de la Universidad Mayor.

## 2.- Detección de SNPs

### 2.1 SAMtools

Los archivos SAM obtenidos para cada individuo fueron procesados con el software SAMtools v1.7 (Li y col., 2009), siendo transformados a BAM, ordenados e indexados. Estos archivos finales fueron utilizados para generar archivos *pileup*, mediante la función *mpileup* de SAMtools, para cada una de las comparaciones de los individuos de sitios contaminados con los individuos del sitio no contaminado (IM-SFM, MEL-SFM y PEL-SFM). Estos archivos contienen, para cada individuo, la comparación de cada una de las bases, presentes en cada lectura, con el ensamblaje *de novo* como referencia, determinando si corresponde a la misma base de la referencia u otra distinta indicando, en este caso, a que base corresponde.

### 2.2 Perl

Los archivos en formato *pileup*, obtenidos para cada comparación, fueron filtrados utilizando un script escrito en lenguaje de programación Perl v5.26.1. Se buscaron SNPs fijos, es decir aquellos que están mayormente representados en una población (contaminada o no contaminada) con respecto a otra población y SNPs diversos, los cuales hacen referencia a que en una población ningún alelo tiene una frecuencia muy alta, mientras que en la otra población existe un alelo con mayor frecuencia para esa posición. Estas distinciones se hicieron para poder detectar todos los tipos de cambios de frecuencias alélicas en las poblaciones comparadas. El script permite discriminar entre (i) SNPs fijos, (ii) SNPs diversos en el sitio contaminado y fijos en el sitio no contaminado

y (iii) SNPs diversos en el sitio no contaminado y fijos en el sitio contaminado, además descarta del análisis SNPs que posean menos de 5 lecturas de cobertura y aquellos que presenten más del 5% de lecturas con inserciones o deleciones. Los criterios utilizados para definir cada uno de estos SNPs fueron:

- (i) En el sitio contaminado, se presenta una base en más del 70% de las lecturas, en una posición dada, en al menos 4 individuos y en el sitio no contaminado, se presenta una base distinta en más del 70% de las lecturas, en la misma posición mencionada anteriormente, en al menos 4 individuos.
- (ii) En el sitio no contaminado existe una base fija, es decir presente en más del 70% de las lecturas, en una posición dada, en al menos 4 individuos y en el sitio contaminado no existe una preferencia por una base en dicha posición, es decir ninguna base está representada en un 55% o más de las lecturas en al menos 4 individuos.
- (iii) En el sitio contaminado existe una base fija, es decir presente en más del 70% de los lecturas, en una posición dada, en al menos 4 individuos y en el sitio no contaminado no existe una preferencia por una base en dicha posición, es decir ninguna base está representada en un 55% o más de las lecturas en al menos 4 individuos.

En el caso de PEL, al sólo tener 3 individuos, la restricción impuesta fue la detección de los SNPs, antes mencionados, en todos los individuos de dicha localidad.

Además se determino si los SNPs compartidos entre los distintos sitios contaminados eran sustituciones sinónimas o no. Para esto se utilizó la herramienta blastx de NCBI,

utilizando como referencia la base nr y se compararon los cambios en la secuencia entre el sitio contaminado y el sitio no contaminado.

Por último se calculó la densidad de los SNPs detectados, definiendo densidad como la cantidad de SNPs por tamaño, en número de bases, del contig en el que se encuentran (SNPs/b).

### **3.- Análisis de enriquecimiento**

Se realizó un análisis de enriquecimiento utilizando el software Blast2GO PRO5 (Conesa & cols., 2005). Para esto en primer lugar se realizó un BLAST de las secuencias obtenidas del ensamble *de novo* para hígado, utilizando la función blastx (McGinnis & Madden, 2004), con un *e-value* de  $10^{-5}$  y buscando en un subset de secuencias de vertebrados de la base de datos nr de NCBI, la cual contiene secuencias de proteínas no redundantes y la base de datos de proteínas InterPro. Posteriormente se realizó un mapeo, que consiste en encontrar las ontologías genéticas (GO terms) asociados a las secuencias que lograron al menos un hit en el proceso de BLAST, esto se realiza tomando en cuenta los ID para las secuencias a analizar disponibles en Gene Ontology Consortium y Uniprot. Se llevó a cabo una anotación funcional, la cual consiste en filtrar los GO terms obtenidos anteriormente y asignar estos a las secuencias de hígado analizadas, para este análisis se utilizó un e-value de  $10^{-6}$  y los parámetros predeterminados en el programa. Finalmente utilizando las secuencias que presentan SNPs, detectadas en el paso anterior, se realiza un análisis de enriquecimiento, el cual permitirá determinar que funciones moleculares, procesos biológicos y/o componentes celulares, se encuentran sobre representados en el subset de secuencias entregadas, con respecto a todas las secuencias del ensamble *de novo*.

Esto se realizó usando el test exacto de Fisher implementado en el software Blast2GO PRO5, con una tasa de falsos positivos de 5%, teniendo en consideración los resultados obtenidos a partir de la detección de SNPs antes descrita, este análisis se llevó a cabo para los SNPs obtenidos de las comparaciones IM-SFM y MEL-SFM. Para los SNPs obtenidos de la comparación PEL-SFM se realizó una búsqueda bibliográfica de sus funciones, esto debido a su bajo número.



## RESULTADOS

### 1.- RNA-seq y mapeo de lecturas

A partir del experimento de RNA-seq se obtuvo un promedio de 16.332.045 lecturas. Luego de los distintos filtros de calidad aplicados se mantuvieron 16.324.535 lecturas en promedio. Los valores de las lecturas promedio y número de bases secuenciadas en cada sitio se encuentran en la Tabla 1.

**Tabla 1: Número de lecturas y bases posteriores al filtro.** Se muestran los rangos de lecturas obtenidos y el rango de bases secuenciadas para cada sitio. Ya que el largo de cada lectura es de 100 bases, se utilizó este valor para determinar las bases secuenciadas.

Sitio	Rango de lecturas filtradas	Rango de bases secuenciadas (b)
SFM	11.887.612 – 18.735.967	11.887.612.000 – 18.735.967.000
IM	13.073.337 – 13.181.709	13.073.337.000 – 13.181.709.000
MEL	13.399.396 – 17.631.119	13.399.396.000 – 17.631.119.000
PEL	18.316.078 – 21.055.046	18.316.078.000 – 21.055.046.000

Debido a que no el genoma de *B. microlepidotus* no se encuentra secuenciado, las lecturas de todos los individuos se utilizaron para realizar un ensamble *de novo* obteniéndose 74.773 contigs de un tamaño mayor a 200 pares de bases. Posteriormente se realizó el mapeo de las lecturas obtenidas para cada individuo, utilizando como referencia el ensamble *de novo* generado anteriormente. El porcentaje promedio de mapeo de las lecturas de todos los individuos fue de 91,28% con una desviación estándar de  $\pm 4,78\%$ .

## **2.- Detección de SNPs**

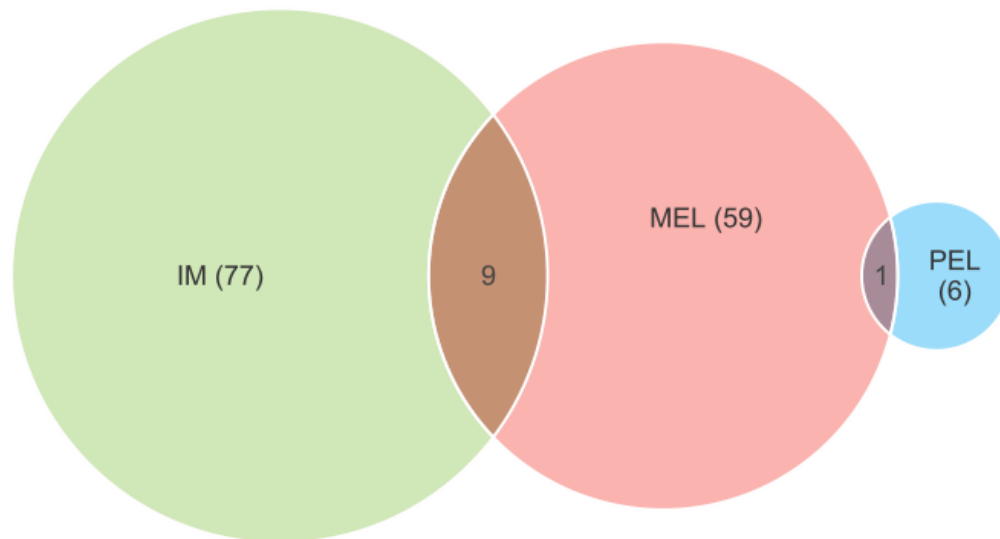
Utilizando el software SAMtools y los archivos SAM obtenidos del mapeo para cada individuo, se generó un archivo en formato pileup para las comparaciones IM-SFM, MEL-SFM y PEL-SFM. Se utilizó un script de Perl para analizar los archivos pileup, permitiendo detectar tres tipos de SNPs: i) SNPs fijos, ii) SNPs diversos en el sitio contaminado y fijos en el sitio no contaminado y iii) SNPs diversos en el sitio no contaminado y fijos en el sitio contaminado. Se detectaron un total de 290 SNPs entre todos los tipos y sitios analizados, distribuidos en 136 contigs distintos. En la comparación IM-SFM se detectaron los 3 tipos de SNPs mencionados, mientras que en la comparación MEL-SFM sólo se detectaron SNPs fijos y en la comparación PEL-SFM se detectaron SNPs fijos y SNPs diversos en el sitio no contaminado. La comparación MEL-SFM es la que presenta el mayor número de SNPs fijos y por su parte la comparación IM-SFM es la que presenta el mayor número de SNPs diversos. A continuación se detalla el número de SNPs detectados para cada comparación y tipo de SNP (Tabla 2).

**Tabla 2: Número de SNPs detectados por comparación.** Se muestra el número de SNPs detectados por cada comparación y entre paréntesis se indica el número de contigs en los que se encuentran. Además se señala el rango de tamaño de los contigs de cada comparación.

<b>Comparación</b>	<b>SNPs fijos</b>	<b>SNPs diversos en contaminado</b>	<b>SNPs diversos en no contaminado</b>	<b>Rango de tamaño de contigs (b)</b>
IM-SFM	93(56)	14(14)	7(7)	1136 - 14210
MEL-SFM	167(59)	0	0	1191 - 9058
PEL-SFM	8(5)	0	1(1)	1191 - 14254

Los contigs corresponden en su mayoría a una función en particular, salvo los contigs Comp49\_seq0 y Comp20\_seq0 que corresponden a más de una función.

Dentro de los contigs que presentan SNPs existen algunos comunes entre las comparaciones de las distintas poblaciones, encontrándose un mayor número de contigs compartidos entre las comparaciones IM-SFM y MEL-SFM, lo que indica selección sobre funciones similares en los distintos pares de sitios mencionados (Figura 2).



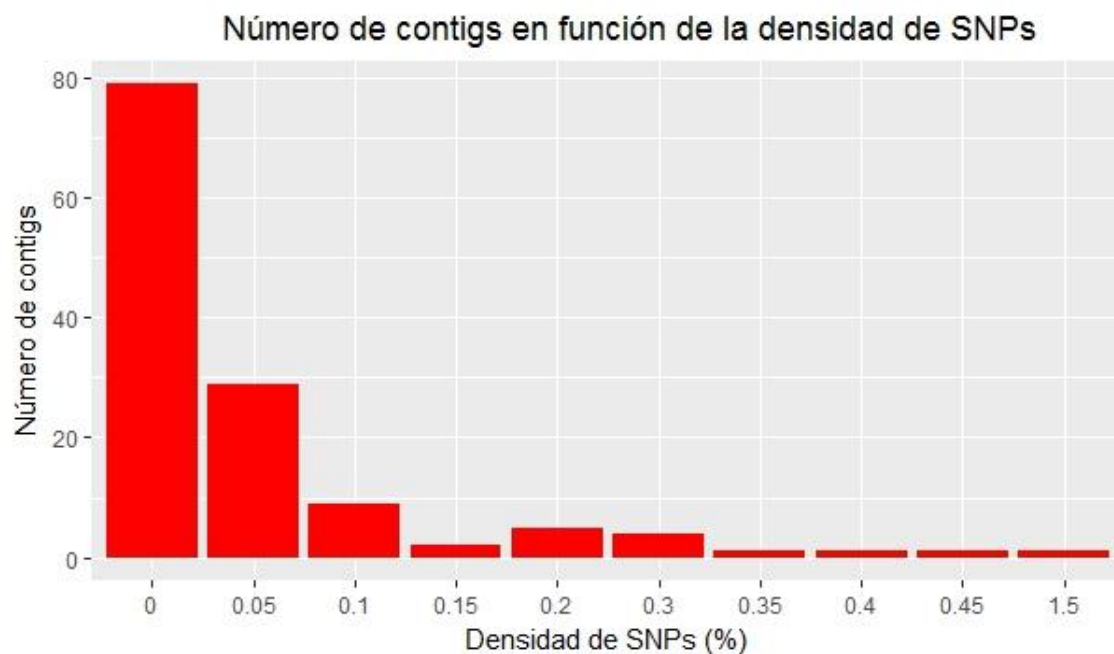
**Figura 2: Diagrama de Venn de contigs compartidos entre comparaciones.** El diagrama muestra el número de contigs que presentan SNPs por comparación entre paréntesis y el número de contigs que presentan SNPs compartidos entre las comparaciones en las intersecciones de los conjuntos.

Se analizaron los aminoácidos para los cuales codifican los codones en los que se encuentran los SNPs compartidos entre los distintos sitios (Tabla 3). Sólo se encontró un SNP que generó un cambio en el aminoácido codificado, este se encontró en las comparaciones IM-SFM y MEL-SFM en el gen Citocromo B (CYTB). El cambio en el aminoácido fue de Arginina en SFM y Triptófano en IM y MEL, este cambio podría comprometer la función o la actividad del producto del gen CYTB.

**Tabla 3: Cambio en aminoácidos producto de los SNPs.** Se muestran los SNPs compartidos entre comparaciones y el aminoácido para el que codifican en el sitio contaminado y en el sitio no contaminado. \* sustitución no sinónima.

<b>Par de comparaciones</b>	<b>Contig</b>	<b>Tamaño (b)</b>	<b>Posición (b)</b>	<b>Aminoácido en contaminado</b>	<b>Aminoácido en no contaminado</b>
<b>(IM-SFM)- (MEL-SFM)</b>	Comp21_seq0	7055	3626	Arginina	Arginina
	Comp49_seq0	6699	5031	Ac.Glutámico	Ac.Glutámico
	Comp49_seq1*	6699	1736	Triptófano	Arginina
	Comp49_seq1	6699	2390	Valina	Valina
	Comp49_seq1	6699	3172	Glicina	Glicina
	Comp49_seq1	6699	5031	Asparagina	Asparagina
<b>(PEL-SFM)- (MEL-SFM)</b>	Comp49_seq1	66999	5889	Arginina	Arginina
	Comp4119_seq0	1180	383	Treonina	Treonina
	Comp4119_seq0	1180	386	Asparagina	Asparagina

Con respecto a la densidad, se observa que la mayoría de los contigs presentan una densidad mayor a 0% y menor a 0,05%, seguido por un grupo de contigs con densidades entre 0,05% y 0,1%, además el contig que presenta la mayor densidad tiene relación con el proceso de respiración celular (Figura 3).



**Figura 3: Número de contigs en función de la densidad de SNPs.** El gráfico muestra número de contigs agrupados según un rango de densidad de SNPs expresada en porcentaje. No se observaron contigs que presenten una densidad de SNPs entre 0,46% y 1,49%.

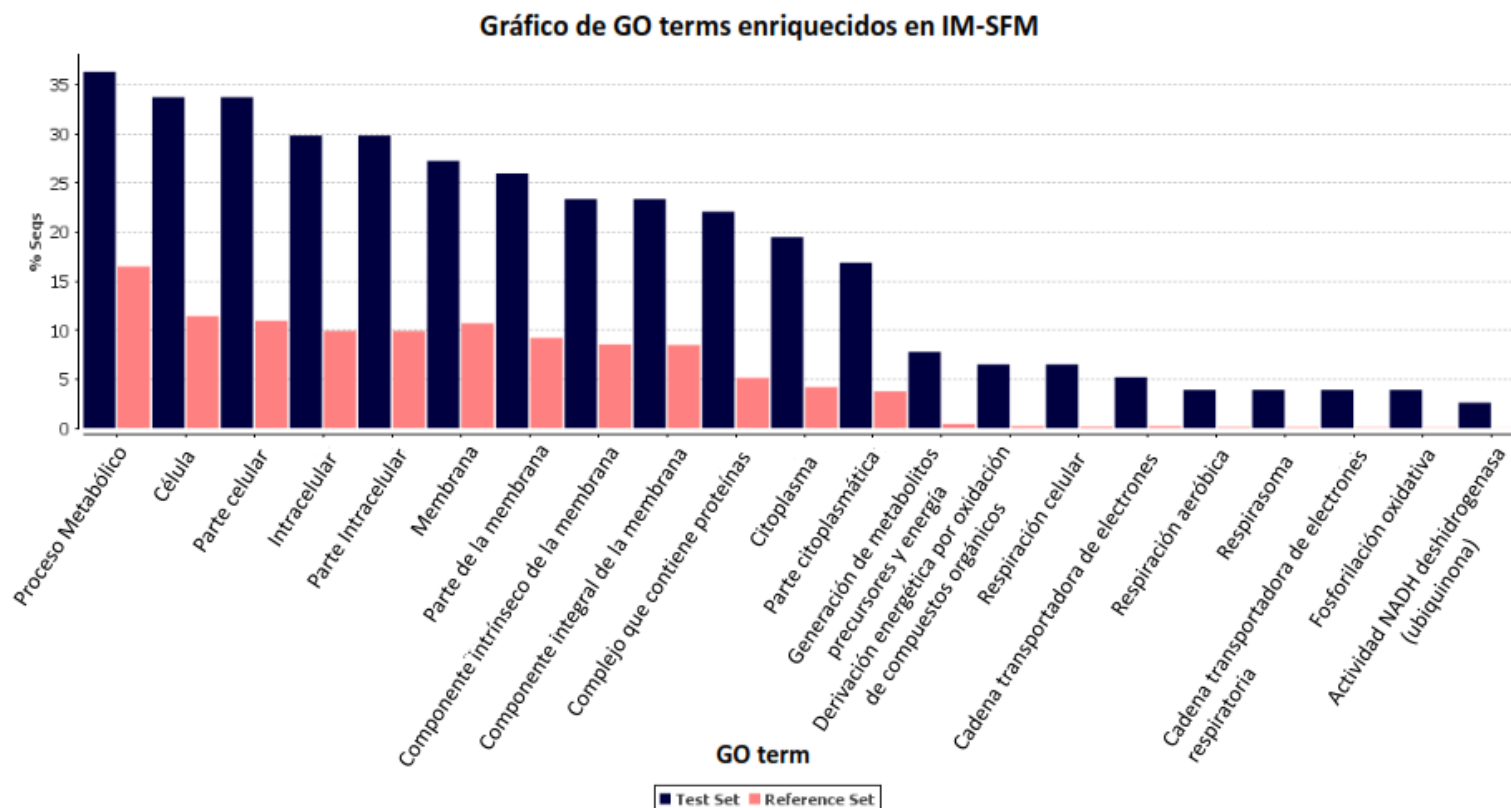
Los valores obtenidos para la densidad se encuentran dentro de los rangos esperables para un estudio de este tipo.

### **3.- Análisis de enriquecimiento**

Utilizando todos los contigs que presentaron SNPs en cada una de las comparaciones se realizó un análisis de enriquecimiento en el software Blast2GO PRO para las comparaciones IM-SFM y MEL-SFM.

En el caso de la comparación IM-SFM se encontraron 21 GO terms sobre representados, observándose GO terms relacionados con los procesos de respiración celular (7) y con la

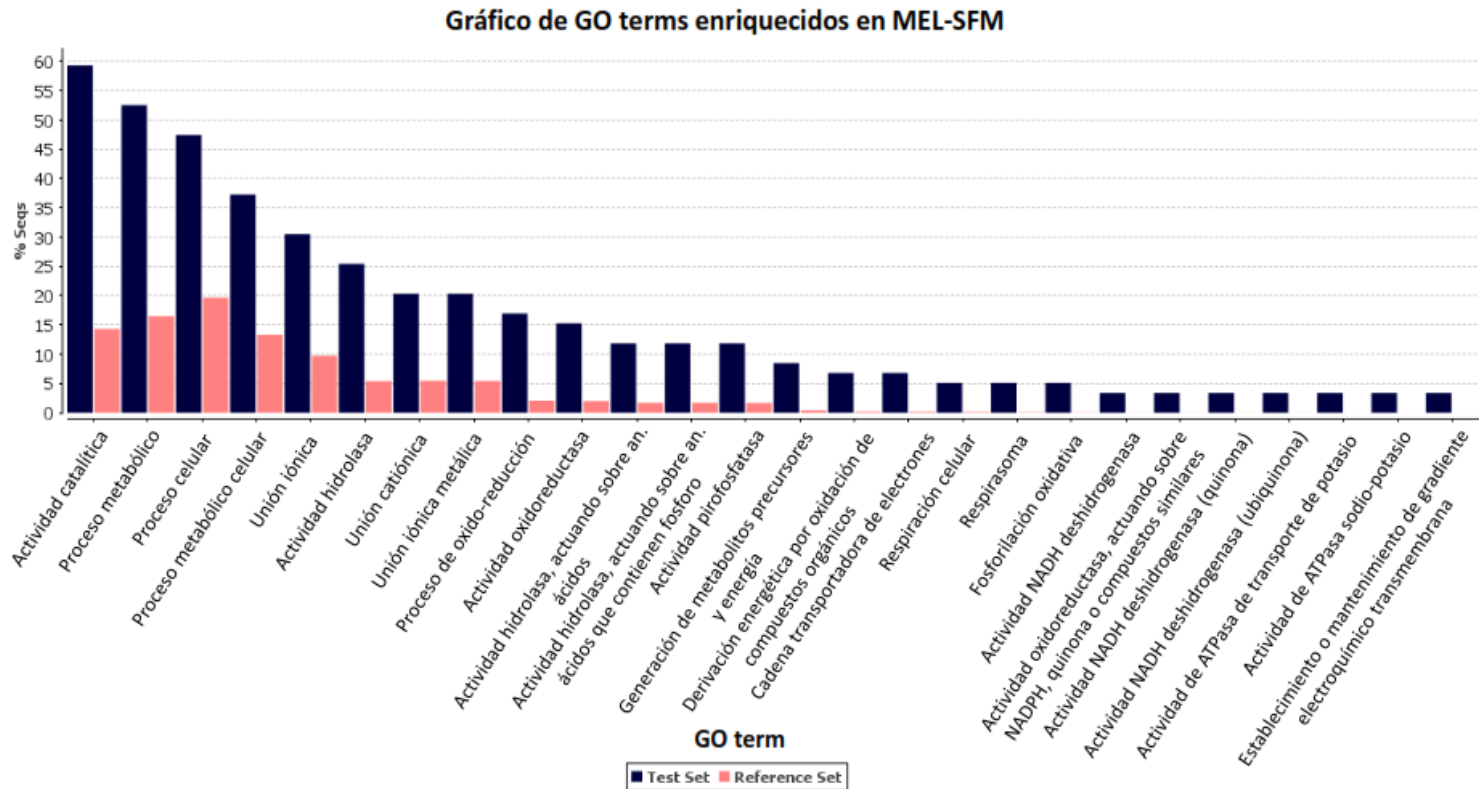
membrana celular (4) (Figura 4) (Figura A1; Figura A2; Figura A3) (Anexo). Dentro de esta categoría de GO terms relacionados con componentes celulares tenemos dos grandes subdivisiones, GO terms asociados a respirasoma y GO terms asociados a citoplasma (Figura A2). La mayor diferencia entre IM y SFM está relacionada con procesos metabólicos, los cuales a su vez tienen relación con los procesos de respiración celular mencionados anteriormente. Por otro lado es importante mencionar que los contigs que presentan mayor densidad de SNPs, están relacionados con las funciones de respiración celular.



**Figura 4: Gráfico de GO terms enriquecidos en comparación IM-SFM.** Se muestra un gráfico de barras con los GO terms que se encuentran sobre-representados en la comparación IM-SFM, representado en el porcentaje de secuencias que presentan dicho GO term. En negro se observa el porcentaje del GO term observado para los contigs con SNPs detectados, mientras que en rosado se observa el porcentaje para el mismo GO term en los contigs de todo el transcriptoma analizado.



En la comparación MEL-SFM se encontraron 26 GO terms sobre representados, destacando 7 relacionados con respiración celular, 1 relacionado con la unión de cationes, 1 relacionado con la unión de metales, 3 relacionados con la ATPasa  $\text{Na}^+/\text{K}^+$  y 1 relacionado con procesos de óxido reducción (Figura 5) (Figura A4; Figura A5; Figura A6). Dentro de este último proceso se encuentra el gen Citocromo P450 2j6, que puede ser de relevancia para el contexto del estudio. Además, los contigs más densos presentes en esta comparación también están relacionados con los procesos de respiración celular.



**Figura 5: Gráfico de GO terms enriquecidos en comparación MEL-SFM.** Se muestra un gráfico de barras con los GO terms que se encuentran sobre-representados en la comparación MEL-SFM, representado en el porcentaje de secuencias que presentan dicho GO term. En negro se observa el porcentaje del GO term observado para los contigs con SNPs detectados, mientras que en rosado se observa el porcentaje para el mismo GO term en los contigs de todo el transcriptoma analizado.

En el caso de la comparación PEL-SFM se procedió a buscar en literatura los GO terms asociados a los contigs que presentaron SNPs, esto debido al bajo número de contigs encontrados (6) (Tabla 4). Dentro de los contigs analizados, destaca el contig comp4119\_seq1, ya que es el contig que presenta mayor densidad de SNPs (3) y está asociado a funciones relacionadas con la ATPasa Na<sup>+</sup>/K<sup>+</sup>.

**Tabla 4: Anotación de contigs con SNPs presentes en comparación PEL-SFM.** Se muestran los contigs que presentan SNPs en la comparación PEL-SFM, la descripción del contig, el e-value que se reportó al momento de realizar el blast de la secuencia, los IDs de los GO terms encontrados. F: función molecular, P: proceso biológico y C: componente celular.

Contig	Tamaño (b)	Descripción	e-value	GO IDs	GO terms
comp95_seq0	14254	Similar a Apolipoproteína B-100	0E0	F:GO:0005319 P:GO:0006869	<b>F:</b> Actividad de transporte de lípidos. <b>P:</b> Transporte de lípidos.
comp4119_seq1	1191	Subunidad alfa ATPasa de transporte de Na <sup>+</sup> /K <sup>+</sup>	0E0	F:GO:0005391 F:GO:0005524 P:GO:0006813 P:GO:0006814 P:GO:0010248 C:GO:0016020 F:GO:0046872	<b>F:</b> Actividad de ATPasa de intercambio sodio/potasio <b>F:</b> Unión de ATP. <b>P:</b> Transporte de ion potasio. <b>P:</b> Transporte de ion sodio. <b>P:</b> Establecimiento mantenimiento de un gradiente electroquímico transmembrana. <b>C:</b> Membrana. <b>F:</b> Unión de ion metálico.
comp5753_seq3	3001	Endonucleasa/Transcriptasa Reversa	5,01E-71	F:GO:0003964 F:GO:0005515 F:GO:0009036 P:GO:0032199 P:GO:0090305	<b>F:</b> Actividad de DNA polimerasa dirigida por RNA. <b>F:</b> Unión de proteínas. <b>F:</b> Actividad desoxiribonucleasa tipo 2 sitio específico. <b>P:</b> Transcripción reversa involucrada en transposición mediada por RNA. <b>P:</b> Hidrólisis de enlace fosfodiéster de ácidos nucleicos.
Comp5829_seq1	2305	Factor de iniciación de la traducción EIF-2B isoforma delta X1	0E0	P:GO:0044237	<b>P:</b> Proceso metabólico celular.
Comp9055_seq0	6206	Pumilio homologo 2 isoforma X1	0E0	F:GO:0003723	<b>F:</b> Unión a RNA.

Comp9743_seq0	2343	Similar a proteína de unión a RNA 4 isoforma X1	1,6E-11	F:GO:0003723	<b>F:</b> Unión a RNA.
---------------	------	---	---------	--------------	------------------------

A partir de los análisis de enriquecimiento de IM-SFM y MEL-SFM, y teniendo en cuenta la tabla presentada para la comparación PEL-SFM, se observa que existen GO terms compartidos entre las distintas comparaciones. Para el caso de IM-SFM y MEL-SFM, hay GO terms asociados a respiración celular compartidos, mientras que entre MEL-SFM y PEL-SFM, hay GO terms asociados a mantenimiento del gradiente electroquímico transmembrana por la ATPasa sodio potasio. Entre la comparación IM-SFM y PEL-SFM sólo se ve un GO term muy general compartido entre ambas comparaciones, el cual corresponde al componente celular de membrana.

## DISCUSIÓN

La contaminación en los sistemas de agua dulce ha demostrado ser un agente selectivo a nivel genético sobre las poblaciones que habitan dichos sistemas. En este contexto el pejerrey endémico *Basilichtys microlepidotus* es de especial interés, ya que habita una de las cuencas más intervenidas por las actividades humanas dentro de Chile.

El principal objetivo de este trabajo fue la detección y caracterización funcional de SNPs presentes en hígados de individuos de *Basilichtys microlepidotus* habitando sitios contaminados y no contaminados. A partir de los análisis realizados fue posible detectar y caracterizar SNPs en individuos de todas las poblaciones que habitan sitios contaminados y asociar funciones en las cuales se encuentran implicados, las que se discutirán a continuación.

Se detectó un total de 290 SNPs bajo selección a partir de las distintas comparaciones, distribuidos en 136 contigs. En trabajos con propósitos similares, donde se utiliza un mayor número de individuos y sitios, se ha detectado un número similar de SNPs asociados a contaminación (Laporte y cols., 2015), así como en otros casos se ha detectado un número mucho mayor de SNPs, siendo estos alrededor de 5000 (Bélanger-Deschênes y cols., 2013). Estas diferencias pueden estar dadas por el tipo de contaminación a la que estuvieron sometidos los organismos y también a la especie en particular con la que se trabaja. Sumado a lo anterior hay al menos dos puntos a tener en consideración, en primer lugar está el número de individuos que fueron analizados en este trabajo, al ser sólo 6 individuos por sitio e incluso 3 para un sitio en particular (PEL), es posible que no se detecte una gran cantidad de SNPs al no encontrarse en los individuos muestreados, esto

se evidencia aún más al comparar los resultados obtenidos para PEL, con los de las comparaciones para IM y MEL. Lo segundo a tener en consideración es el filtro usado para la detección, este es bastante estricto comparado con los parámetros usualmente utilizados en softwares con los mismos propósitos, como por ejemplo vcftools (Danecek y cols., 2011), esto debido a que fue programado para evitar la mayor cantidad posible de falsos positivos.

En cuanto a la densidad de los SNPs, gran parte de los contigs presentan densidades superiores a 0% e inferior a 0,1%. La densidad de SNPs es algo bastante variable, sin embargo en la bibliografía se registra que para casos similares, la mayoría de los contigs no presentan una densidad superior al 2% (Liu y cols., 2011; Wang y cols., 2015). Un punto importante a destacar es que el contig con mayor densidad (1.5%), corresponde a un contig relacionado con funciones asociadas a respiración celular, esto es importante ya que se han encontrado genes mitocondriales asociados a esta función conteniendo SNPs que podrían ser de relevancia para la supervivencia en ambientes contaminados con metales pesados, en particular con Cd, permitiendo que este no bloquee ciertos receptores neuronales y no se acumule dentro de las neuronas (Bélanger-Deschênes y cols., 2013). En el caso de *B. microlepidotus* se encontró una sustitución no sinónima en el gen citocromo B, en los sitios MEL e IM, cuyo producto génico es parte del complejo 3 de la cadena de transporte de electrones, esta sustitución se presenta en una interface entre dominios, regiones que regulan la función y relación entre dominios (Verma y Pandit, 2019), por lo que podría eventualmente afectar la función de esta proteína y de la cadena transportadora de electrones. Se ha encontrado en otro trabajo una sustitución no sinónima

en el gen citocromo B posiblemente a causa de la contaminación, sin embargo este cambio no estaba en una zona que pudiese afectar el funcionamiento de la proteína codificada (Bélanger-Deschênes y cols., 2013).

En la comparación IM-SFM un total de 114 SNPs fueron detectados, gran parte de estos están relacionados con componentes celulares. Los GO terms relacionados con respirasoma, están involucrados con los procesos de respiración celular que también se encuentran sobrerrepresentados en esta comparación. Por otro lado GO terms asociados a citoplasma son bastante generales como parte citoplasmática (GO: 0044444) o intracelular (GO: 0005622) y se han visto sobrerrepresentados en organismos sometidos a distintos tipos de estrés, como en el caso de individuos de *Mya truncata* sometidos a temperaturas elevadas (Sleight y cols., 2018) o en el caso de individuos de *Aequipecten opercularis* sometidos a toxinas de organismos unicelulares (Ventoso y cols., 2019). Por lo que no pueden ser atribuidos exclusivamente a situaciones de contaminación, sino más bien a una situación de estrés en general.

Con respecto a los procesos biológicos sobrerrepresentados, en la comparación mencionada, el término Proceso metabólico (GO: 0008152) es el que presenta un mayor porcentaje, este término está relacionado con el proceso de respiración celular (GO: 0045333) y en forma más particular con la cadena respiratoria de transporte de electrones (GO: 0022900). La cadena de transporte de electrones es vital para la obtención de energía en los organismos aeróbicos y medidas de su actividad han sido utilizadas como parte de un indicador del estado energético de organismos que están sometidos a contaminación, conocido como CEA o Asignación energética celular, que además incluye datos sobre la

energía almacenada en las células de estos organismos (Smolders y cols., 2004). En individuos de *Daphnia magna* sometidos a contaminación con metales pesados, se observa un aumento de la actividad de la cadena de transporte de electrones luego de 48 horas (De Coen y Jansen, 2003). Por otro lado en individuos de *Danio rerio* sometidos a distintos contaminantes como mercurio, uranio y nanopartículas de oro, se observaron variaciones en la actividad de la cadena transportadora de electrones, dependiendo del agente contaminante (Bourdineaud y cols., 2013). Además es sabido que frente a situaciones de estrés producidas por contaminantes es necesario realizar una relocalización de la energía utilizada para combatir estos efectos negativos, la cual ya no se encuentra disponible para crecimiento u otros procesos que afectan directamente a la fisiología de los organismos (Calow, 1991). Por otro lado se ha reportado que las actividades de las enzimas de la cadena transportadora de electrones y en general los procesos asociados a la respiración celular están relacionados con el crecimiento en corderos (Hiendleder y cols., 1995). Según trabajos anteriores (Briones, 2019), los individuos de *B. microlepidotus* que habitan sitios contaminados, consumen alimentos de menor calidad y tienen un estado fisiológico deficiente con respecto a aquellos individuos que habitan sitios no contaminados. Además estudios de expresión diferencial en *B. microlepidotus* han indicado que es posible que los individuos que habitan sitios contaminados estén pasando períodos de hambruna, lo que también afectaría a su estado fisiológico (Vega-Retter y cols., 2018). En humanos se han detectado numerosos SNPs asociados a cáncer en genes pertenecientes al complejo I de la cadena transportadora de electrones, dentro de los cuales se encuentra el gen ND5, (Koshikawa y cols., 2017), el cual es uno de los genes que presenta más SNPs en *B. microlepidotus*. En trabajos



anteriores (Briones, 2016) también se ha detectado SNPs en el gen ND5 a partir de datos de RNA-seq en *B. microlepidotus*, no obstante al secuenciar este segmento en una mayor cantidad de individuos se ha descartado como tal. Sin embargo en el trabajo anterior sólo se trabajó con un fragmento de ND5, en el presente trabajo se cuenta con la secuencia completa, además que se presentó más de un SNP en dicho gen, por lo que debe ser estudiado nuevamente. El gen ND5 o subunidad 5 de la NADH deshidrogenasa, es esencial para la actividad del complejo I de la cadena de transporte de electrones (Bai y cols., 2000; Chomyn y cols., 2001). En un estudio de Zhang y colaboradores (2008) se identificaron SNPs en el gen ND5 en vacas, los cuales estaban asociados a una mayor tasa de crecimiento durante los primeros meses de vida. Por otro lado en Bélanger-Deschênes y colaboradores (2013) se ha detectado SNPs en genes mitocondriales que participan de la cadena transportadora de electrones, en individuos de *Perca flavescens* que han estado en sitios contaminados con metales pesados, particularmente en el gen ND2 o subunidad 2 de la NADH deshidrogenasa, sin embargo esta mutación no está involucrada directamente con la función de ND2 en la cadena transportadora de electrones, sino más bien en su función como proteína de andamiaje, donde pudiese tener efectos sobre la sinapsis. En Du y colaboradores (2015) se observó un mejor funcionamiento del metabolismo de fosforilación oxidativa en individuos de *Fundulus heteroclitus* que habitan sitios contaminados con compuestos orgánicos, frente a la exposición a estos compuestos, con respecto a individuos que habitan sitios no contaminados al ser sometidos al mismo contaminante, se plantea que esto permite a los individuos que habitan sitios contaminados mantener la producción de ATP, lo que ayudaría a combatir los altos gastos energéticos que supone estar bajo el estrés provocado por la contaminación.

Además esta característica mostró ser heredable, al comparar individuos de la tercera generación de los mismos sitios muestreados anteriormente (Du y cols., 2016). Recientemente se han identificado SNPs en los genes relacionados con el metabolismo de la fosforilación oxidativa en los individuos de *F. heteroclitus* que habitan los ya mencionados sitios contaminados, además al estudiar parámetros fisiológicos como la tasa de consumo de oxígeno, ésta resultó ser mayor en los individuos provenientes de los sitios contaminados con respecto a los individuos de sitios no contaminados (Jayasundara y cols., 2017). Sin embargo, a pesar de que existe una adaptación (mantención de la producción de ATP) que les permite sobrevivir al estrés causado por la contaminación, estos peces muestran en general un peor desempeño al lidiar con estrés producido por otros agentes como por ejemplo el aumento de temperatura (Jayasundara y cols., 2017). Además, se ha observado expresión diferencial de genes asociados a la cadena transportadora de electrones en organismos que habitan sitios contaminados (Oleksiak, 2008). Por ejemplo en individuos de *Anguilla anguilla* que habitan sitios contaminados con pesticidas y metales, se ha documentado que los genes sub-expresados se encuentran enriquecidos en el proceso biológico de la cadena transportadora de electrones (Pujolar y cols., 2012). Es importante mencionar que el cambio de expresión génica puede ser producto de un cambio a nivel genético por una exposición a largo plazo a la contaminación (Whitehead y cols., 2010). Asociado a todo lo dicho anteriormente, podemos observar que la única función molecular enriquecida en la comparación IM-SFM es justamente la Actividad NADH deshidrogenasa (ubiquinona) (GO:0008137), actividad que corresponde al complejo I de la cadena transportadora de electrones.

Dados estos antecedentes es claro que el proceso biológico de la cadena de transporte de electrones puede verse afectado y puede tener un rol clave en organismos que se encuentran en sistemas contaminados, sin embargo es necesario estudiar en mayor profundidad los SNPs asociados a esta función, para determinar si pudiesen ser beneficiosos o no para los individuos de *B. microlepidotus*.

Con respecto a la comparación MEL-SFM, nuevamente se observan enriquecidos procesos biológicos relacionados con la respiración celular, al igual que en la comparación IM-SFM. Además, dentro de los procesos biológicos enriquecidos aparece el establecimiento o mantenimiento de un gradiente electroquímico transmembrana (GO:0010248), proceso que está directamente relacionado con la función molecular actividad ATPasa de intercambio sodio-potasio (GO:0005391). Este proceso tiene como principal función mantener el equilibrio osmótico y es llevado a cabo por la ATPasa sodio-potasio. Se ha observado que distintos contaminantes tienen un efecto sobre esta actividad, ya sea inhibiéndola en el caso del óxido nítrico (Tipsmark y cols., 2003) o aumentándola en el caso de los hidrocarburos aromáticos policíclicos (Buet y cols., 2006). Además, medidas de su actividad han sido utilizadas como indicadores de estrés osmoregulatorio en tilapia (Sardella y cols., 2004). Se sabe muy poco con respecto a polimorfismos en este gen asociados directamente a contaminación, esto en parte porque es un gen que está relacionado con estrés en general (Jeffries y cols., 2015), por lo que se requiere una investigación más detallada para hacer una correcta relación. Con respecto a otras condiciones, como el aumento de temperatura, ha sido posible detectar al menos un SNP asociado en vacas (Liu y cols., 2010). Por otra parte se ha documentado extensamente que

existen cambios en la expresión de la ATPasa sodio potasio frente a distintas condiciones de contaminación (Sinha y cols., 2012; Jeffries y cols., 2015).

Otro proceso biológico de relevancia que se encuentra sobre-representado en esta comparación es el proceso de oxido-reducción (GO:0055114), dentro de los genes que poseen este GO term se encuentra el gen Citocromo p450 2j6, el cual pertenece a la superfamilia de genes citocromo p450 o CYP. Esta familia de genes tiene una función de oxido-reducción llevada a cabo sobre sustratos lipofílicos, dentro de los cuales se encuentran drogas, pesticidas, hidrocarburos aromáticos, entre otros. En el caso particular del producto del gen Citocromo p450 2j6, este lleva a cabo la oxidación de un alcano a un alcohol y participa de muchos procesos relacionados con el ácido araquidónico (Lee y cols., 2005). Debido a su función, esta superfamilia de genes ha sido utilizada como indicador de contaminación de sistemas acuáticos en peces (Uno y cols., 2012). Estos genes presentan un alto polimorfismo y se cree que esto está relacionado con la susceptibilidad de los individuos por una droga, un agente tóxico del ambiente o un agente endógeno (Nebert y cols., 1996; King y cols., 2002). Se ha documentado que en peces expuestos a contaminación por petróleo cambian los niveles de ácido araquidónico, el cual tiene un rol clave en la señalización celular y además es sustrato del producto del gen Citocromo p450 2j6, por lo que un cambio en la secuencia de esta proteína podría tener relación con la afinidad por este sustrato endógeno (Balk y cols., 2011).

En el caso de la comparación PEL-SFM se observa el más bajo número de SNPs, en parte esto puede ser explicado por el bajo número de individuos que fue utilizado para el análisis. De todas formas fue posible encontrar SNPs en la ATPasa sodio potasio, al igual

que en la comparación MEL-SFM. Además se encontró un SNP en un gen similar a Apolipoproteína B-100, cuyo producto es parte constitutiva esencial de las partículas LDL, las cuales son partículas de baja densidad que transportan lípidos (Brown y Goldstein, 1986). El proceso biológico transporte de lípidos (GO:0006869), al cual está asociado este gen, y otros procesos relacionados con partículas de LDL, se han encontrado en genes que presentan SNPs, en anguilas del Nor-atlántico sometidas a distintos tipos de contaminantes, como metales pesados y materia orgánica (Laporte y cols., 2015). Por otro lado se detectó un SNP en la subunidad delta de EIF-2B, que corresponde a un factor de iniciación de la traducción, que juega un rol esencial al participar en la activación de eIF2, al intercambiar GDP por GTP, dando inicio a la traducción (Ohlenbusch y cols., 2005). No se han vinculado directamente SNPs en este gen con respecto a la contaminación, sin embargo se han encontrado SNPs asociados a contaminación en genes relacionados como EF1A o factor de elongación 1 alfa el cual también participa en la traducción (Bélanger-Deschênes y cols.,2013). Por otro lado se han detectado SNPs en este gen asociados a malformaciones de la materia blanca en el sistema nervioso de humanos (Ohlenbusch y cols., 2005).

En general vemos en todas las comparaciones que los procesos y funciones enriquecidos, donde se encuentran los SNPs, tienen relación con procesos metabólicos relacionados con la obtención de energía o a la mantención de la osmoregulación. Pero, es importante mencionar que se ha documentado que peces adaptados a la contaminación, que presentan selección en genes similares a los encontrados en este trabajo, presentan un peor desempeño en otras condiciones como el estrés provocado por el aumento de temperatura

(Jayasundara y cols., 2017). Por otro lado, como ya se ha mencionado, los individuos de *B. microlepidotus* que habitan sitios contaminados presentan en general un estado fisiológico deficiente y tienen acceso a comida de menor calidad (Briones, 2019). Estas condiciones podrían ser las causantes de la selección de alelos en genes relacionados con procesos metabólicos de obtención de energía u osmoregulación como una forma de adaptación a las condiciones generadas por la exposición crónica a contaminantes, además, las diferencias de expresión génica que se han observado en estos individuos en estudios anteriores (Vega-Retter y cols., 2018), también podrían tener un componente evolutivo. Se ha demostrado en *Fundulus heteroclitus* que al criar en condiciones de laboratorio a individuos de sitios contaminados, estos aún presentan genes expresados diferencialmente con respecto a los individuos de sitios limpios sometidos al mismo tratamiento (Whitehead y cols., 2010). Es necesario estudiar de forma más extensa cada uno de los SNPs detectados en este trabajo para determinar el verdadero impacto que pudiesen tener sobre los aminoácidos para los cuales codifican los codones donde estos cambios se encuentran, vale decir si corresponden a mutaciones sinónimas o no sinónimas, sobre todo en cuanto a la mutación no sinónima encontrada en el gen CYTB en los sitios IM y MEL. Además estos deben ser validados secuenciando un mayor número de individuos. Por último es importante evaluar el impacto sobre las distintas actividades con las que estos SNPs están relacionados, sin embargo, dados los antecedentes expuestos, no se puede descartar que las funciones enriquecidas encontradas en *B. microlepidotus* estén relacionadas con estrategias para enfrentar el estrés provocado por la contaminación.

## CONCLUSIÓN

- El presente trabajo sugiere que la contaminación actúa como fuerza selectiva sobre genes relacionados con el metabolismo de *B. microlepidotus*.
- Se encontró selección de funciones compartidas entre pares de sitios (IM-MEL, PEL-MEL).
- Existe al menos una mutación no sinónima en el gen CYTB, perteneciente a la cadena transportadora de electrones, en los sitios IM y MEL.
- Las funciones enriquecidas donde se detectaron mayor número de SNPs fueron en aquellas relacionadas con la cadena transportadora de electrones y ATPasa sodio potasio.
- Es necesario evaluar la relevancia y el impacto de los SNPs detectados, desde el cambio estructural en la proteína para la cual codifica el gen donde se encuentra dicho SNP, hasta el impacto fisiológico de estos cambios.

## BIBLIOGRAFÍA

- Allan, J. D., McIntyre, P. B., Smith, S. D., Halpern, B. S., Boyer, G. L., Buchsbaum, A., Burton, G. A., Campbell, L. M., Chadderton, W. L., Ciborowski, Jan J.H., Doran, P. J., Eder, T., Infante, D. M., Johnson, L. B., Joseph, C. A., Marino, A. L., Prusevich, A., Read, Jennifer G., Rose, Joan B., Rutherford, Edward S., Sowa, Scott P., y Stienman, Alan D. 2013. Joint analysis of stressors and ecosystem services to enhance restoration effectiveness. *Proceedings of the National Academy of Sciences*. Rev. 110(1): 372-377.
- Argiriadi, M. A., Morisseau, C., Hammock, B. D., y Christianson, D. W. 1999. Detoxification of environmental mutagens and carcinogens: structure, mechanism, and evolution of liver epoxide hydrolase. *Proceedings of the National Academy of Sciences*. 96(19): 10637-10642.
- Arkoosh, M. R., Casillas, E., Clemons, E., Kagley, A. N., Olson, R., Reno, P., y Stein, J. E. 1998. Effect of pollution on fish diseases: potential impacts on salmonid populations. *Journal of Aquatic Animal Health*. Rev. 10(2): 182-190.
- Bai, Y., Shakeley, R. M., y Attardi, G. 2000. Tight control of respiration by NADH dehydrogenase ND5 subunit gene expression in mouse mitochondria. *Molecular and cellular biology*. Rev. 20(3): 805-815.
- Balk, L., Hylland, K., Hansson, T., Berntssen, M. H., Beyer, J., Jonsson, G., Melbye, A., Grung, M., Torstensen, B., Borseth, J., Klungsoyr, J. & Skarphedinsdottir, H. (2011). Biomarkers in natural fish populations indicate adverse biological effects of offshore oil production. *PLoS One*. 6(5): e19735.
- Bélangier-Deschênes, S., Couture, P., Campbell, P. G., y Bernatchez, L. 2013. Evolutionary change driven by metal exposure as revealed by coding SNP genome scan in wild yellow perch (*Perca flavescens*). *Ecotoxicology*. Rev. 22(5): 938-957.
- Bertness, M. D., Ewanchuk, P. J., y Silliman, B. R. 2002. Anthropogenic modification of New England salt marsh landscapes. *Proceedings of the National Academy of Sciences*. Rev. 99(3): 1395-1398.
- Bourdineaud, J. P., Rossignol, R., y Brethes, D. 2013. Zebrafish: a model animal for analyzing the impact of environmental pollutants on muscle and brain mitochondrial bioenergetics. *The international journal of biochemistry & cell biology*. Rev. 45(1): 16-22.
- Bozinovic, G., y Oleksiak, M. F. 2011. Genomic approaches with natural fish populations from polluted environments. *Environmental toxicology and chemistry*. Rev. 30(2): 283-289.
- Buet, A., Banas, D., Vollaire, Y., Coulet, E., y Roche, H. 2006. Biomarker responses in European eel (*Anguilla anguilla*) exposed to persistent organic pollutants. A field study in the Vaccarès lagoon (Camargue, France). *Chemosphere*. Rev. 65(10): 1846-1858.



Briones, M., 2016. Estudio de la variabilidad genética de tres genes funcionales en poblaciones del pejerrey chileno *Basilichthys microlepidotus* presentes en zonas contaminadas y no contaminadas de la cuenca del Río Maipo (tesis de pregrado). Universidad de Chile, Facultad de Ciencias.

Briones, M., 2019. Efecto de la contaminación y el tipo de cauce sobre la estructura trófica del Río Maipo, usando como modelo a los macroinvertebrados bentónicos y al pejerrey *Basilichthys microlepidotus* (Jenyns, 1841) (tesis de magister). Universidad de Chile, Facultad de Ciencias.

Brown MS, Goldstein JL. 1986. A receptor-mediated pathway for cholesterol homeostasis. *Science*. Rev. 232:34-47.

Calow, P. 1991. Physiological costs of combating chemical toxicants: ecological implications. *Comparative biochemistry and physiology. C, Comparative pharmacology and toxicology*. Rev. 100(1-2): 3-6.

Cánovas, A., Rincon, G., Islas-Trejo, A., Wickramasinghe, S., y Medrano, J. F. 2010. SNP discovery in the bovine milk transcriptome using RNA-Seq technology. *Mammalian genome*. Rev. 21(11-12): 592-598.

Chang, Z., Li, G., Liu, J., Zhang, Y., Ashby, C., Liu, D. y Huang, X. 2015. Bridger: a new framework for de novo transcriptome assembly using RNA-seq data. *Genome biology*. Rev. 16(1): 30.

Chomyn, A. 2001. Mitochondrial genetic control of assembly and function of complex I in mammalian cells. *Journal of bioenergetics and biomembranes*. Rev. 33(3), 251-257.

Comte, S., y Vila, I. 1992. Spawning of *Basilichthys microlepidotus* (Jenyns). *Journal of fish biology*. Rev. 41(6): 971-981.

Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., y Robles, M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*. Rev. 21(18): 3674-3676.

Cooper, C. M. 1993. Biological effects of agriculturally derived surface water pollutants on aquatic systems—a review. *Journal of environmental quality*. Rev. 22(3): 402-408.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker R., Lunter G., Marth, G., Sherry, S., Dubin, R. y McVean, G. 2011. The variant call format and VCFtools. *Bioinformatics*. Rev. 27(15): 2156-2158.

De Coen, W. M., y Janssen, C. R. 2003. The missing biomarker link: Relationships between effects on the cellular energy allocation biomarker of toxicant-stressed *Daphnia magna* and corresponding population characteristics. *Environmental Toxicology and Chemistry: An International Journal*. Rev. 22(7): 1632-1641.

DeWijer, P., Watt, P. J., y Oldham, R. S. 2003. Amphibian decline and aquatic pollution: effects of nitrogenous fertiliser on survival and development of larvae of the frog *Rana temporaria*. *Applied Herpetology*. Rev. 1: 3-12.

Dirección general de aguas. 2004. Diagnóstico y clasificación de los cursos y cuerpos de agua según objetivos de calidad, cuenca del Río Maipo. Ministerio de Obras Públicas, Gobierno de Chile.

Du, X., Crawford, D. L., y Oleksiak, M. F. 2015. Effects of anthropogenic pollution on the oxidative phosphorylation pathway of hepatocytes from natural populations of *Fundulus heteroclitus*. *Aquatic Toxicology*. Rev. 165: 231-240.

Du, X., Crawford, D. L., Nacci, D. E., y Oleksiak, M. F. 2016. Heritable oxidative phosphorylation differences in a pollutant resistant *Fundulus heteroclitus* population. *Aquatic Toxicology*. Rev. 177: 44-50.

Duarte, W., Feito, R., Jara, C., Moreno, C., y Orellana, A. 1971. Ictiofauna del sistema hidrográfico del río Maipo. *Boletín del Museo Nacional de Historia Natural*. Rev. 32: 227-268.

Dulvy, N. K., Sadovy, Y., y Reynolds, J. D. 2003. Extinction vulnerability in marine populations. *Fish and fisheries*. Rev. 4(1): 25-64.

Jeffries, K. M., Brander, S. M., Britton, M. T., Fangue, N. A., y Connon, R. E. 2015. Chronic exposures to low and high concentrations of ibuprofen elicit different gene response patterns in a euryhaline fish. *Environmental Science and Pollution Research*. Rev. 22(22): 17397-17413.

Grant, D. M. 1991. Detoxification pathways in the liver. *Journal of inherited metabolic disease*. Rev. 14: 421-430.

Hedrick, P. W. 2007. Balancing selection. *Current Biology*. Rev. 17(7): R230-R231.

Hiendleder, S., Herrmann, M., Dzapo, V., y Wassmuth, R. 1995. Mitochondrial respiratory metabolism and growth performance of lambs, 1: Respiratory-chain enzyme activities. *Journal of Animal Breeding and Genetics (Germany)*. Rev. 112: (381-390).

Hrdlickova, R., Toloue, M., y Tian, B. 2017. RNA-Seq methods for transcriptome analysis. *Wiley Interdisciplinary Reviews: RNA*, 8(1): e1364.

Ikemoto, T., Kunito, T., Tanaka, H., Baba, N., Miyazaki, N., y Tanabe, S. 2004. Detoxification mechanism of heavy metals in marine mammals and seabirds: interaction of selenium with mercury, silver, copper, zinc, and cadmium in liver. *Archives of Environmental Contamination and Toxicology*. Rev. 47(3): 402-413.

Jayasundara, N., Fernando, P. W., Osterberg, J. S., Cammen, K. M., Schultz, T. F., y Di Giulio, R. T. 2017. Cost of tolerance: physiological consequences of evolved resistance to inhabit a polluted environment in teleost fish *Fundulus heteroclitus*. *Environmental science & technology*. Rev. 51(15): 8763-8772.

King, L. M., Ma, J., Srettabunjong, S., Graves, J., Bradbury, J. A., Li, L., Spiecker, M., Liao, J., Mohrenweiser, H. y Zeldin, D. C. 2002. Cloning of CYP2J2 gene and identification of functional polymorphisms. *Molecular pharmacology*. Rev. 61(4): 840-852.

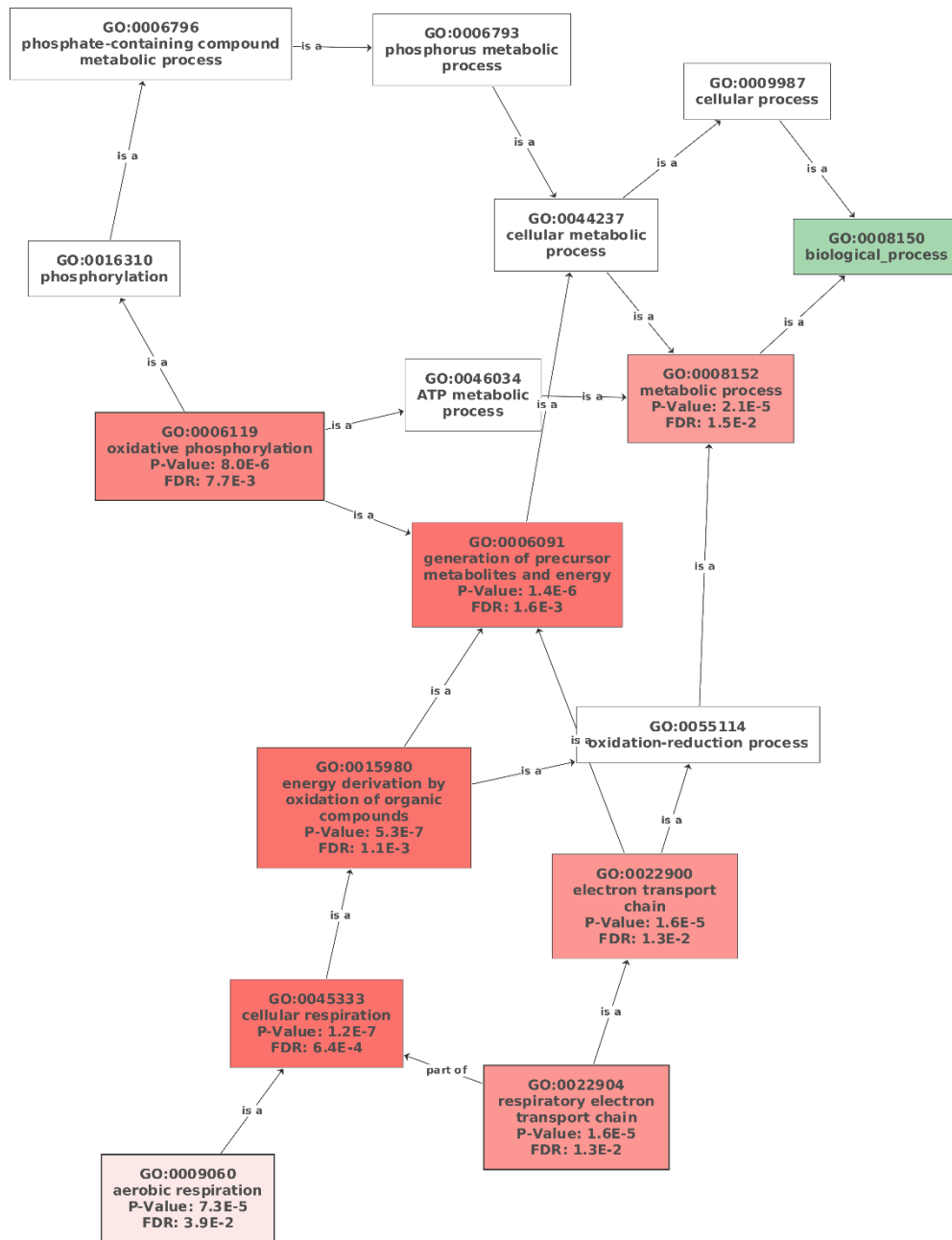
- Koshikawa, N., Akimoto, M., Hayashi, J. I., Nagase, H., y Takenaga, K. 2017. Association of predicted pathogenic mutations in mitochondrial ND genes with distant metastasis in NSCLC and colon cancer. *Scientific Reports*. Rev 7(1): 15535.
- Krueger, F. (2015). Trim galore. A wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files.
- Langmead, B., & Salzberg, S. L. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods*. Rev. 9(4): 357.
- Lauring, A. S., y Andino, R. 2010. Quasispecies theory and the behavior of RNA viruses. *PLoS Pathogens*. Rev. 6(7): e1001005.
- Laporte, M., Pavey, S. A., Rougeux, C., Pierron, F., Lauzent, M., Budzinski, H., y Bernatchez, L. 2016. RAD sequencing reveals within-generation polygenic selection in response to anthropogenic organic and metal contamination in North Atlantic Eels. *Molecular Ecology*. Rev. 25(1): 219-237.
- Lee, S. S., Jeong, H. E., Liu, K. H., Ryu, J. Y., Moon, T., Yoon, C. N., Oh, S., Yun, C., y Shin, J. G. 2005. Identification and functional characterization of novel CYP2J2 variants: G312R variant causes loss of enzyme catalytic activity. *Pharmacogenetics and genomics*. Rev. 15(2): 105-113.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N. y Durbin, R. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics*. Rev. 25(16): 2078-2079.
- Lind, E. E., y Grahn, M. 2011. Directional genetic selection by pulp mill effluent on multiple natural populations of three-spined stickleback (*Gasterosteus aculeatus*). *Ecotoxicology*. Rev. 20(3): 503-512.
- Liu, Y. X., Zhou, X., Li, D. Q., Cui, Q. W., y Wang, G. L. 2010. Association of ATP 1 A 1 gene polymorphism with heat tolerance traits in dairy cattle. *Genetics and Molecular Research*. Rev. 9(2): 891-896.
- Liu, S., Zhou, Z., Lu, J., Sun, F., Wang, S., Liu, H., Jiang, Y., Kucuktas, H., Kaltenboeck, L., Peatman, E. y Liu, Z. 2011. Generation of genome-scale gene-associated SNPs in catfish for the construction of a high-density SNP array. *BMC Genomics*. Rev. 12(1): 53.
- Lürling, M., & Scheffer, M. 2007. Info-disruption: pollution and the transfer of chemical information between organisms. *Trends in Ecology & Evolution*. Rev. 22(7): 374-379.
- McGinnis, S., & Madden, T. L. 2004. BLAST: at the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Research*. Rev. 32(suppl\_2): W20-W25.
- Ministerio del Medio Ambiente Chile. 2018. Inventario nacional de especies. <http://www.mma.gob.cl/clasificacionespecies/listado-especies-nativas-segun-estado-2014.htm>

- Mitra, T., Mohanty, B. P., Mohanty, S., Purohit, G. K., y Das, B. K. 2018. Expression patterns and mutation analysis of p53 in fish *Rita rita* from polluted riverine environment. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*. Rev. 832: 41-51.
- Moiseenko, T. I. 2002. Change in the life cycle strategy of fish under the effect of chronic water pollution. *Russian Journal of Ecology*. Rev. 33(1): 45-55.
- Narváez, J., Richter, P., y Toral, M. 2007. Preliminary physical chemical characterization of river waters and sediments affected by copper mining activity in Central Chile: application of multivariate analysis. *Journal of the Chilean Chemical Society*. Rev. 52(3): 1261-1265.
- Nebert, D. W., Mckinnon, R. A., y Puga, A. 1996. Human drug-metabolizing enzyme polymorphisms: effects on risk of toxicity and cancer. *DNA and Cell Biology*. Rev. 15(4): 273-280.
- Ng, P. & Henikoff, S. 2003. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Research*. 31(13): 3812-3814.
- Ohlenbusch, A., Henneke, M., Brockmann, K., Goerg, M., Hanefeld, F., Kohlschütter, A., y Gärtner, J. 2005. Identification of ten novel mutations in patients with eIF2B-related disorders. *Human Mutation*. 25(4): 411-411.
- Oleksiak, M. F. 2008. Changes in gene expression due to chronic exposure to environmental pollutants. *Aquatic Toxicology*. Rev. 90(3): 161-171.
- Pardo, R., Vila, I., y Bustamante, R. 2008. Demographic response of *Stratiodrillus aeglyphilus* (Anelida, Histriobdellidae) to organic enrichment: experimental assessment. *Hydrobiologia*. Rev. 598(1): 11-16.
- Pettersson, E., Lundeberg, J., y Ahmadian, A. 2009. Generations of sequencing technologies. *Genomics*. Rev. 93(2): 105-111.
- Pizarro, J., Vergara, P. M., Rodríguez, J. A., y Valenzuela, A. M. 2010. Heavy metals in northern Chilean rivers: spatial variation and temporal trends. *Journal of Hazardous Materials*. Rev. 181(1-3): 747-754.
- Poulin, R. 1992. Toxic pollution and parasitism in freshwater fish. *Parasitology Today*. Rev. 8(2): 58-61.
- Pujolar, J. M., Marino, I. A., Milan, M., Coppe, A., Maes, G. E., Capoccioni, F., Ciccotti, E., Bervoets, L., Covaci, A., Belpaire, C., Cramp, G., Patarnello, T., Bargelloni, L., Bartoluzzi, S., y Zane, L. 2012. Surviving in a toxic world: transcriptomics and gene expression profiling in response to environmental pollution in the critically endangered European eel. *BMC Genomics*. Rev. 13(1): 507.
- Retamal, M., Costa, C., Suárez, J. M., y Richter, P. 2013. Multi-determination of organic pollutants in water by gas chromatography coupled to triple quadrupole mass

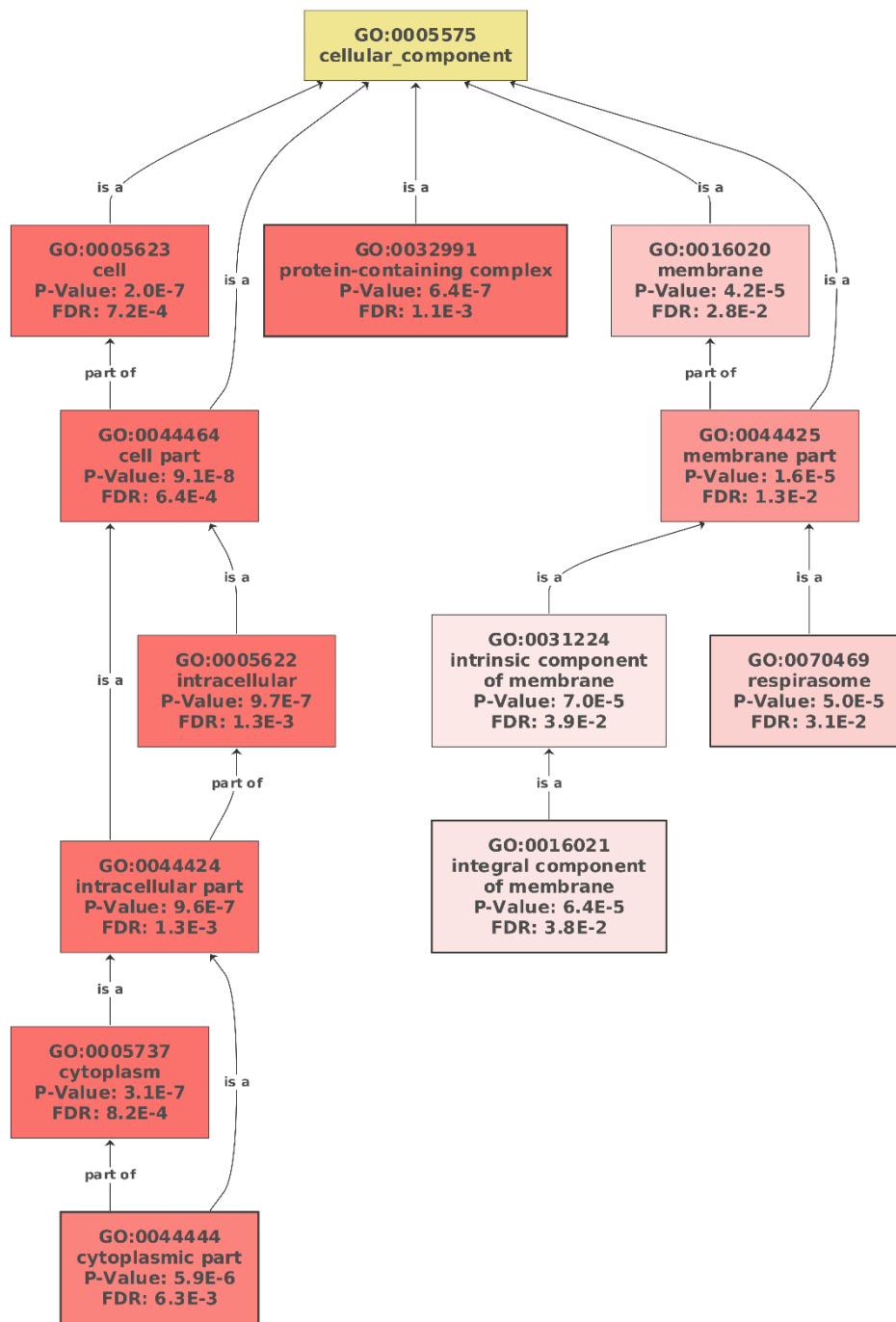
- spectrometry. *International Journal of Environmental Analytical Chemistry*. Rev 93(1): 93-107.
- Ribbe, L., Delgado, P., Salgado, E., y Flügel, W. A. 2008. Nitrate pollution of surface water induced by agricultural non-point pollution in the Pochay watershed, Chile. *Desalination*. Rev. 226(1-3): 13-20.
- Ribeiro, L., Kretschmer, N., Nascimento, J., Buxo, A., Rötting, T. S., Soto, G. y Oyarzun, R. 2014. Water quality assessment of the mining-impacted Elqui River Basin, Chile. *Mine Water and the Environment*. Rev. 33(2): 165-176.
- Salem, M., Vallejo, R. L., Leeds, T. D., Palti, Y., Liu, S., Sabbagh, A. y Yao, J. 2012. RNA-Seq identifies SNP markers for growth traits in rainbow trout. *PLoS One*. Rev. 7(5): e36264.
- Sardella, B. A., Matey, V., Cooper, J., Gonzalez, R. J., y Brauner, C. J. 2004. Physiological, biochemical and morphological indicators of osmoregulatory stress in California Mozambique tilapia (*Oreochromis mossambicus* × *O. urolepis hornorum*) exposed to hypersaline water. *Journal of Experimental Biology*. Rev. 207(8): 1399-1413.
- Scott, G. R., & Sloman, K. A. 2004. The effects of environmental pollutants on complex fish behaviour: integrating behavioural and physiological indicators of toxicity. *Aquatic Toxicology*. Rev. 68(4), 369-392.
- Sih, A., Ferrari, M. C., y Harris, D. J. 2011. Evolution and behavioural responses to human-induced rapid environmental change. *Evolutionary Applications*. Rev. 4(2): 367-387.
- Sinha, A. K., Diricx, M., Chan, L. P., Liew, H. J., Kumar, V., Blust, R., y De Boeck, G. 2012. Expression pattern of potential biomarker genes related to growth, ion regulation and stress in response to ammonia exposure, food deprivation and exercise in common carp (*Cyprinus carpio*). *Aquatic toxicology*. Rev. 122: 93-105.
- Sleight, V. A., Peck, L. S., Dyrinda, E. A., Smith, V. J., y Clark, M. S. 2018. Cellular stress responses to chronic heat shock and shell damage in temperate *Mya truncata*. *Cell Stress and Chaperones*. Rev. 23(5): 1003-1017.
- Smolders, R., Bervoets, L., De Coen, W., y Blust, R. 2004. Cellular energy allocation in zebra mussels exposed along a pollution gradient: linking cellular effects to higher levels of biological organization. *Environmental Pollution*. Rev. 129(1): 99-112.
- Tatara, C. P., Mulvey, M., y Newman, M. C. 1999. Genetic and demographic responses of mosquitofish (*Gambusia holbrooki*) populations exposed to mercury for multiple generations. *Environmental Toxicology and Chemistry: An International Journal*. Rev. 18(12): 2840-2845.
- Tipsmark, C. K., y Madsen, S. S. 2003. Regulation of Na<sup>+</sup>/K<sup>+</sup>-ATPase activity by nitric oxide in the kidney and gill of the brown trout (*Salmo trutta*). *Journal of Experimental Biology*. 206(9): 1503-1510.

- Uno, T., Ishizuka, M., y Itakura, T. 2012. Cytochrome P450 (CYP) in fish. *Environmental Toxicology and Pharmacology*. Rev. 34(1): 1-13.
- Vanzella, T. P., Martinez, C. B. R., y Cólus, I. M. S. 2007. Genotoxic and mutagenic effects of diesel oil water soluble fraction on a neotropical fish species. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*. Rev. 631(1): 36-43.
- Ventoso, P., Pazos, A. J., Pérez-Parallé, M. L., Blanco, J., Triviño, J. C., y Sánchez, J. L. (2019). RNA-Seq Transcriptome Profiling of the Queen Scallop (*Aequipecten opercularis*) Digestive Gland after Exposure to Domoic Acid-Producing *Pseudo-nitzschia*. *Toxins*. Rev. 11(2): 97.
- Vega-Retter, C., Munoz-Rojas, P., Vila, I., Copaja, S., y Véliz, D. 2014. Genetic effects of living in a highly polluted environment: the case of the silverside *Basilichthys microlepidotus* (Jenyns)(Teleostei: *atherinopsidae*) in the Maipo River basin, central Chile. *Population Ecology*. Rev. 56(4): 569-579.
- Vega-Retter, C., Vila, I., y Véliz, D. 2015. Signatures of directional and balancing selection in the silverside *Basilichthys microlepidotus* (Teleostei: *Atherinopsidae*) inhabiting a polluted river. *Evolutionary Biology*. 42(2): 156-168.
- Vega-Retter, C., Rojas-Hernandez, N., Vila, I., Espejo, R., Loyola, D. E., Copaja, S. y Véliz, D. 2018. Differential gene expression revealed with RNA-Seq and parallel genotype selection of the ornithine decarboxylase gene in fish inhabiting polluted areas. *Scientific Reports*. Rev. 8(1): 4820.
- Véliz, D., Catalán, L., Pardo, R., Acuña, P., Díaz, A., Poulin, E., y Vila, I. 2012. The genus *Basilichthys* (Teleostei: *Atherinopsidae*) revisited along its Chilean distribution range (21° to 40° S) using variation in morphology and mtDNA. *Revista Chilena de Historia Natural*. Rev. 85(1): 49-59.
- Verma, R., & Pandit, S. B. 2019. Unraveling the structural landscape of intra-chain domain interfaces: Implication in the evolution of domain-domain interactions. *PloS One*. 14(8): e0220336.
- Wang, J., Qi, H., Li, L., Que, H., Wang, D., y Zhang, G. 2015. Discovery and validation of genic single nucleotide polymorphisms in the Pacific oyster *Crassostrea gigas*. *Molecular Ecology Resources*. Rev. 15(1): 123-135.
- Whitehead, A., Galvez, F., Zhang, S., Williams, L. M., y Oleksiak, M. F. 2010. Functional genomics of physiological plasticity and local adaptation in killifish. *Journal of Heredity*. Rev. 102(5): 499-511.
- Zhang, B., Chen, H., Hua, L., Zhang, C., Kang, X., Wang, X., Pan, C., Lan, X. y Lei, C. 2008. Novel SNPs of the mtDNA ND5 gene and their associations with several growth traits in the Nanyang cattle breed. *Biochemical genetics*. Rev. 46(5-6): 362-368.

## ANEXO

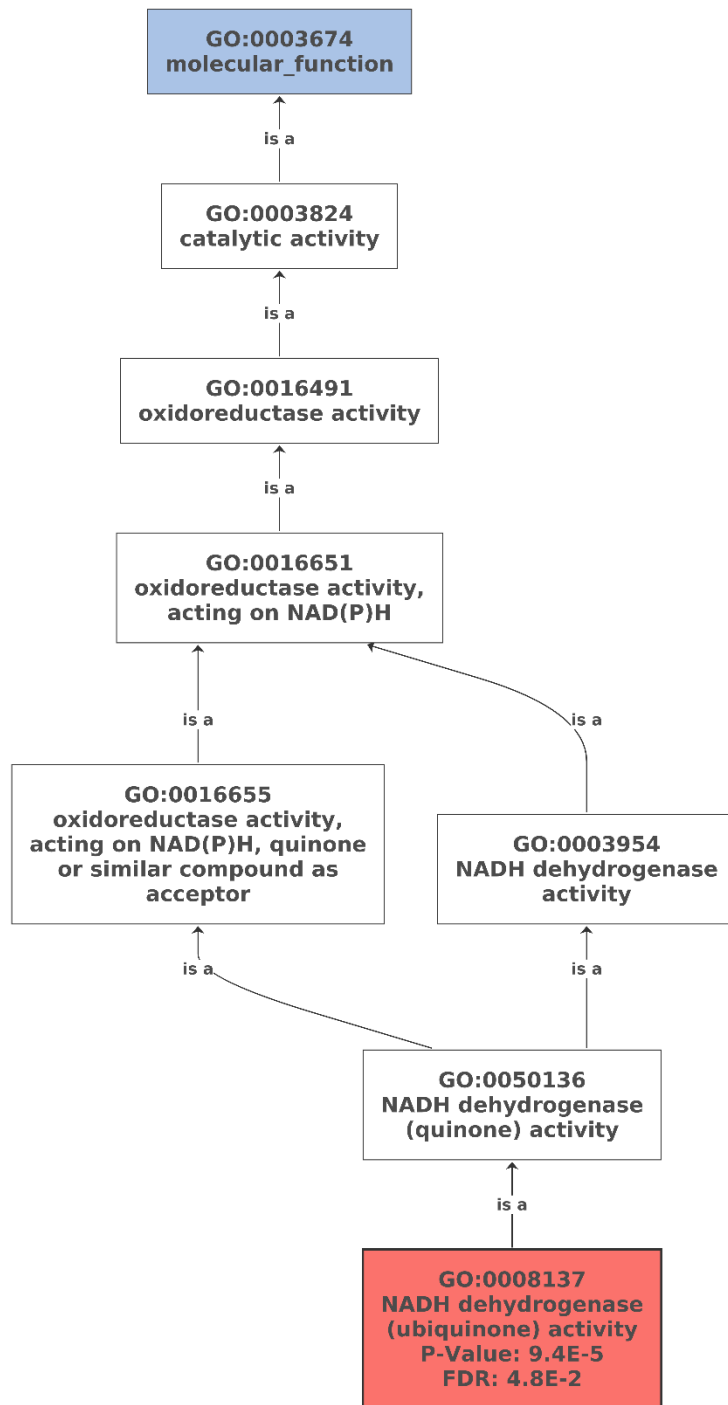


**Figura A1. Procesos Biológicos enriquecidos comparación IM-SFM.** Se muestra los GO term de los procesos biológicos enriquecidos (rojo) y su relación entre si para la comparación IM-SFM.

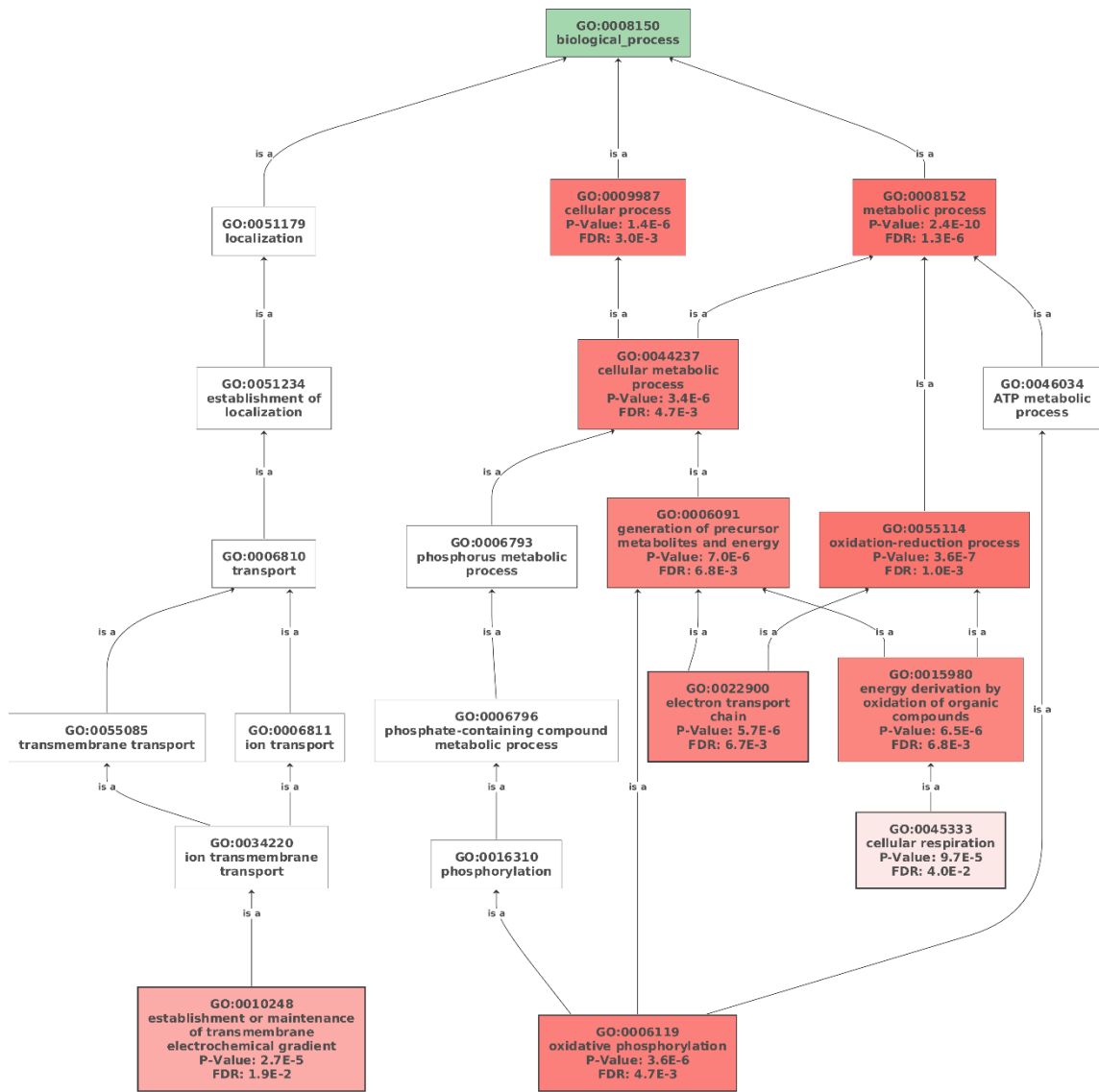


**Figura A2. Componentes celulares enriquecidos en comparación IM-SFM.** Se muestran los GO term correspondientes a los componentes celulares enriquecidos (Rojo) y su relación entre si para la comparación IM-SFM.

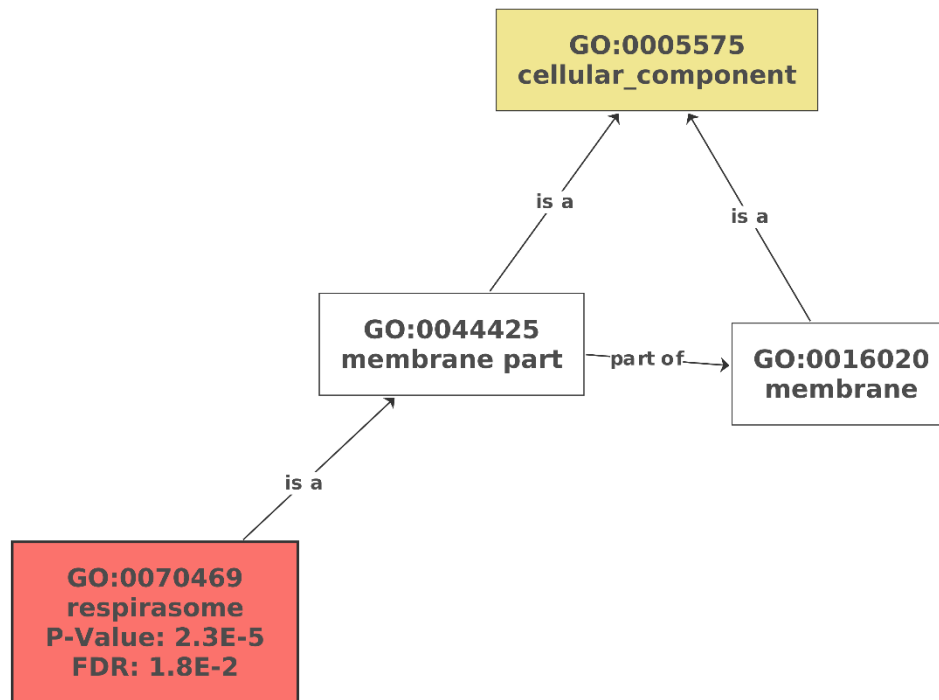




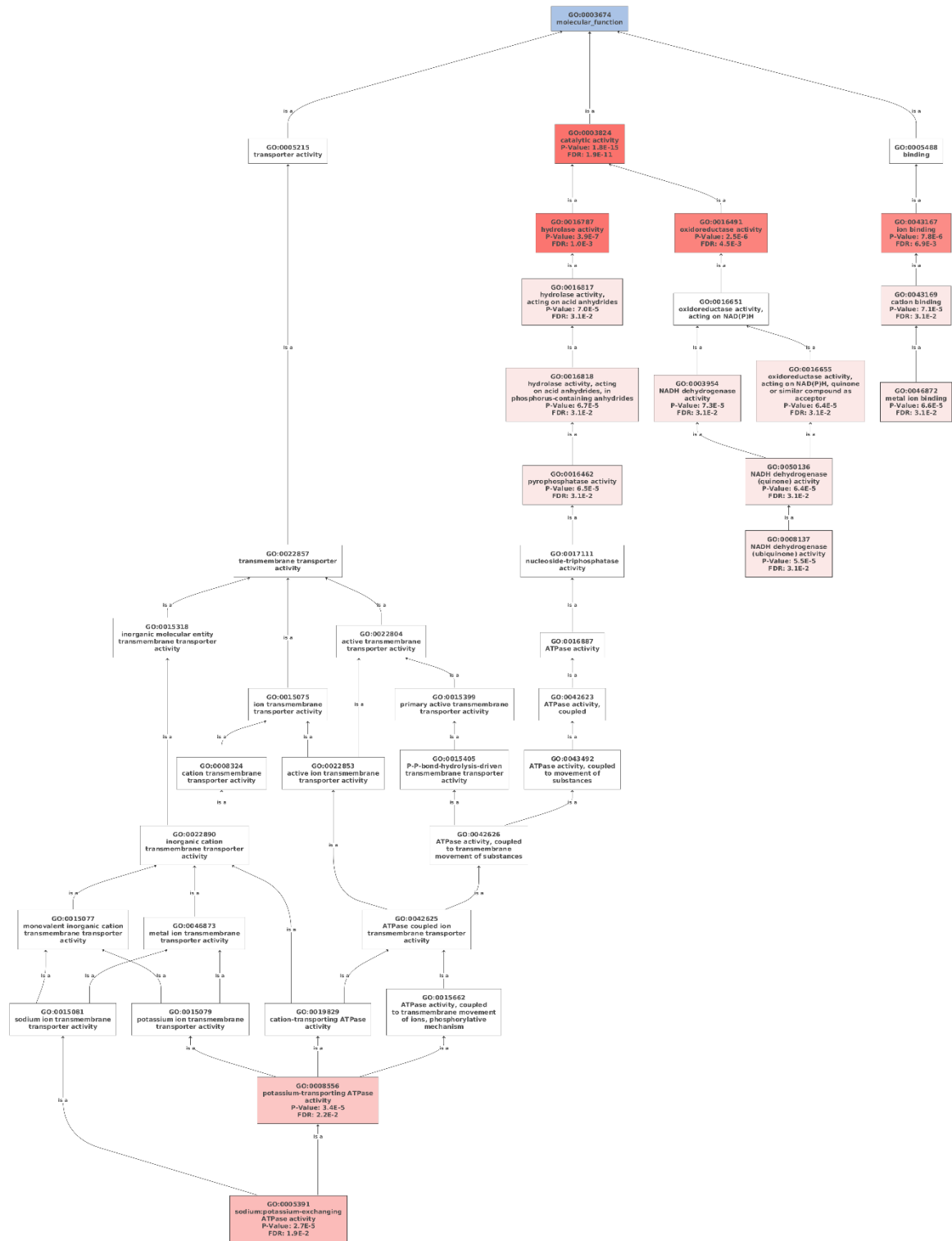
**Figura A3. Funciones moleculares enriquecidas en comparación IM-SFM.** Se muestran los GO term asociados a las funciones moleculares enriquecidas (rojo) y su relación entre si para la comparación IM-SFM.



**Figura A4. Procesos biológicos enriquecidos en comparación MEL-SFM.** Se muestran los GO term asociados a los procesos biológicos enriquecidos (rojo) y su relación entre sí para la comparación MEL-SFM.



**Figura A5. Componentes celulares enriquecidos en comparación MEL-SFM.** Se muestran los GO term asociados a los componentes celulares enriquecidos (rojo) y su relación entre si para la comparación MEL-SFM.



**Figura A6. Funciones moleculares enriquecidas en comparación MEL-SFM.** Se muestran los GO term asociados a las funciones moleculares enriquecidas (rojo) y su relación entre si para la comparación MEL-SFM.