



FACULTAD DE
CIENCIAS

UDELAR | fcien.edu.uy

Universidad de la República

Facultad de Ciencias

Trabajo presentado para acceder al título de Licenciado en Bioquímica

**Genómica de *Echinococcus granulosus*, evolución molecular
de glutatión transferasas de clase Mu (EgGST1)**

Pilar dos Santos

Orientadores: Dra. Verónica Fernández y Dr. Andrés Iriarte

Montevideo, Uruguay

2016

AGRADECIMIENTOS

En primer lugar agradecer a mis orientadores, la Dra. Verónica Fernández y el Dr. Andrés Iriarte por aceptarme como pasante para concluir mis estudios de grado. Por su apoyo y dedicación permanente, por permitirme formar parte de tan enriquecedor proyecto que contribuyo a mi formación.

A los compañeros de la Cátedra de Inmunología por permitirme hacer uso de sus instalaciones, su hospitalidad y su disposición para brindar ayuda con una técnica, un reactivo, una sugerencia oportuna.

A la familia y los amigos que a lo largo de los años me apoyaron y colaboraron para que completara mi formación, estando cerca, dando una mano para cuidar a los “chiquilines”. A l@s chic@s de la facultad por las largas charlas, una “ayudita” para editar una imagen, una idea, una opinión. ¡Gracias!

En especial a Guzmán, Sebastián y Bruno por ser parte de mi vida y bancarme a muerte siempre, no lo hubiera logrado sin ustedes a mi lado.

Tabla de contenido

AGRADECIMIENTOS	2
RESUMEN	4
1. INTRODUCCIÓN GENERAL	5
1.1. Hidatidosis.....	5
1.2 <i>Echinococcus granulosus</i>	5
1.3 Glutación transferasas	7
1.4 Mecanismo catalítico.....	8
1.5 Glutación transferasas en helmintos parásitos	9
2. MATERIALES Y METODOS	13
2.1 ANALISIS INFORMÁTICO	13
2.1.1 Obtención y análisis de secuencias.....	13
2.1.2 Alineamiento de secuencias, modelo evolutivo y análisis filogenético.....	13
2.1.3 Modelado de secuencias	13
2.2 ANALISIS DE LA EXPRESIÓN EN EL PARASITO, DE LOS GENES IDENTIFICADOS	14
2.2.1 Obtención de ARN total del parásito.....	14
2.2.2 Análisis electroforético en geles de Agarosa	14
2.2.3 Digestión de ADN genómico.....	15
2.2.4 Síntesis de ADN copia.....	15
2.2.5 Reacción de amplificación en cadena (PCR)	15
2.2.6 Análisis electroforético en geles de poliacrilamida.....	15
2.2.7 Diseño de <i>primers</i> específicos.....	16
3. RESULTADOS	17
3.1 Identificación de genes y análisis de la estructura exón/intrón.....	18
3.2 Análisis filogenético.....	20
3.3 Análisis de secuencias aminoacídicas	22
3.4 Modelado de secuencias.....	26
3.5 Análisis de la expresión de los genes	27
4. DISCUSIÓN	29
6. CONCLUSIONES	38
Referencias bibliográficas	40
Anexo I	45
Anexo II	48

RESUMEN

La hidatidosis es una enfermedad provocada por el parásito helminto *E. granulosus*. Las enzimas glutatión transferasas (GSTs) son una superfamilia de proteínas multifuncionales con roles esenciales en la detoxificación celular de un amplio rango de moléculas exógenas y endógenas. Los helmintos tienen pocos sistemas de detoxificación, y las GSTs se consideran el principal grupo de enzimas detoxificantes en estos parásitos. Las citosólicas forman el grupo más grande y diverso, donde la clase Mu es una de las 8 clases asignada a esta familia.

Previamente nuestro grupo de trabajo identificó 3 GSTs citosólicas en *E. granulosus*: EgGST1 de clase Mu, EgGST2 de clase Sigma y EgGST3 que no se agrupa con ninguna de las clases hasta ahora descritas. En particular, para EgGST1 se realizó su caracterización bioquímica y se reportó su secuencia completa. Es de nuestro interés, explorar la evolución molecular de esta GST.

A partir de la secuencia de EgGST1, en este trabajo se identificaron en el genoma de *E. granulosus*, varias secuencias relacionadas. Posteriormente se realizó una anotación manual de los *Hits* obtenidos, que dejó cuatro genes ya reportados (*Hit* 1 al 4), el *Hit* 3 se completó en este trabajo. Se analizaron cada secuencia nucleotídica en busca de los exones e intrones, los resultados muestran que la longitud de los exones coincide con lo descrito para los genes de clase Mu. Además, se realizó una búsqueda de motivos y residuos conservados descritos como característicos para las GSTs de esta clase, principalmente aquellos involucrados en su mecanismo catalítico. Por otro lado, se llevó a cabo estudios de filogenia. Resulta interesante haber identificado varios eventos de duplicaciones recurrentes, entre ellos los casos de *H. microstoma* y *E. multilocularis*. Finalmente, se analizó la expresión de dichos genes por el parásito, identificando los ARN mensajeros correspondientes por RT-PCR. En protoescólex, solamente se observó amplicones para los *Hits* 1 y 2, indicando que ambos genes se transcriben.

1. INTRODUCCIÓN GENERAL

1.1. Hidatidosis

La hidatidosis o equinococosis quística o quiste hidático, es una enfermedad zoonótica (enfermedad transmitida al ser humano por los animales) provocada por el parásito *Echinococcus granulosus*. Está distribuida en todo el mundo y se puede observar una alta prevalencia dentro América del Sur en Argentina, Brasil, Chile, Perú y Uruguay (Who, 2014). En el ser humano, la infección por este parásito produce una gran morbilidad, incluso puede producir la muerte del paciente si no se trata. Se presenta de forma asintomática por un período que puede prolongarse por años hasta que se manifiesten los primeros signos clínicos, como ser dolor abdominal, náuseas, vómitos, pérdida de peso, debilidad, entre otros que dificulta su diagnóstico. Para la detección del quiste hidático se utilizan tanto técnicas de imagenología (ecografías, tomografías computadas, etc) como pruebas serológicas, que buscan anticuerpos específicos contra el parásito. No existe un único tratamiento, por lo que se debe adaptar a cada individuo y a cada infección según la morfología, el tamaño y la ubicación del quiste. Requiere un tratamiento farmacológico prolongado y en la mayoría de los casos se debe practicar cirugía para remover los quistes. La cirugía no se puede llevar a cabo en pacientes que presenten quistes en más de un órgano o si están localizados en regiones riesgosas. La ruptura del quiste puede producir una dispersión del material parasitario provocando infecciones secundarias (hidatidosis secundaria). El tratamiento con benzimidazoles como el albendazol facilita su remoción y disminuye el riesgo de infecciones secundarias así como de recurrencia de la parasitosis (Who, 2001).

1.2 *Echinococcus granulosus*

Los helmintos (gusano, organismos pluricelulares complejos) se dividen en dos tipos: platelmintos (cestodos y trematodos) y nematodos. *E. granulosus*

es un cestodo parásito, una clase del filo platelminto (familia Taeniidae). Son endoparásitos, gusanos planos con el cuerpo en forma de cinta, segmentados.

El ciclo de vida *E. granulosus* necesita de dos hospederos, uno definitivo y uno intermediario. El hospedero definitivo que generalmente es el perro, es el que posee el gusano en estado adulto en su intestino delgado, y junto con las heces, va liberando los huevos producidos por el parásito. En las vísceras del hospedero intermediario que suelen ser herbívoros u omnívoros, crecerá la larva (quiste hidático) al ingerir dichos huevos. Los humanos pueden infectarse a través del contacto directo con el hospedero definitivo o por la ingestión de verduras y/o agua contaminada. El huevo eclosiona en el estómago del hospedero intermediario y se libera la oncosfera que es activada en el intestino delgado. La oncosfera penetra la mucosa y entra en la circulación donde se difunde hasta que se aloja en diferentes órganos siendo el hígado y los pulmones los principales, aunque también puede alojarse en los riñones, huesos, corazón y bazo. Una vez alojada la oncosfera, se transforma en quiste hidático o metacestodo que va creciendo lentamente y en su interior se desarrollan los proescólices (PEs). El hospedero definitivo se infecta al ingerir las vísceras de hospederos intermediarios que albergan quistes conteniendo PEs. (Thompson R.C.A., 1995) (Figura 1).

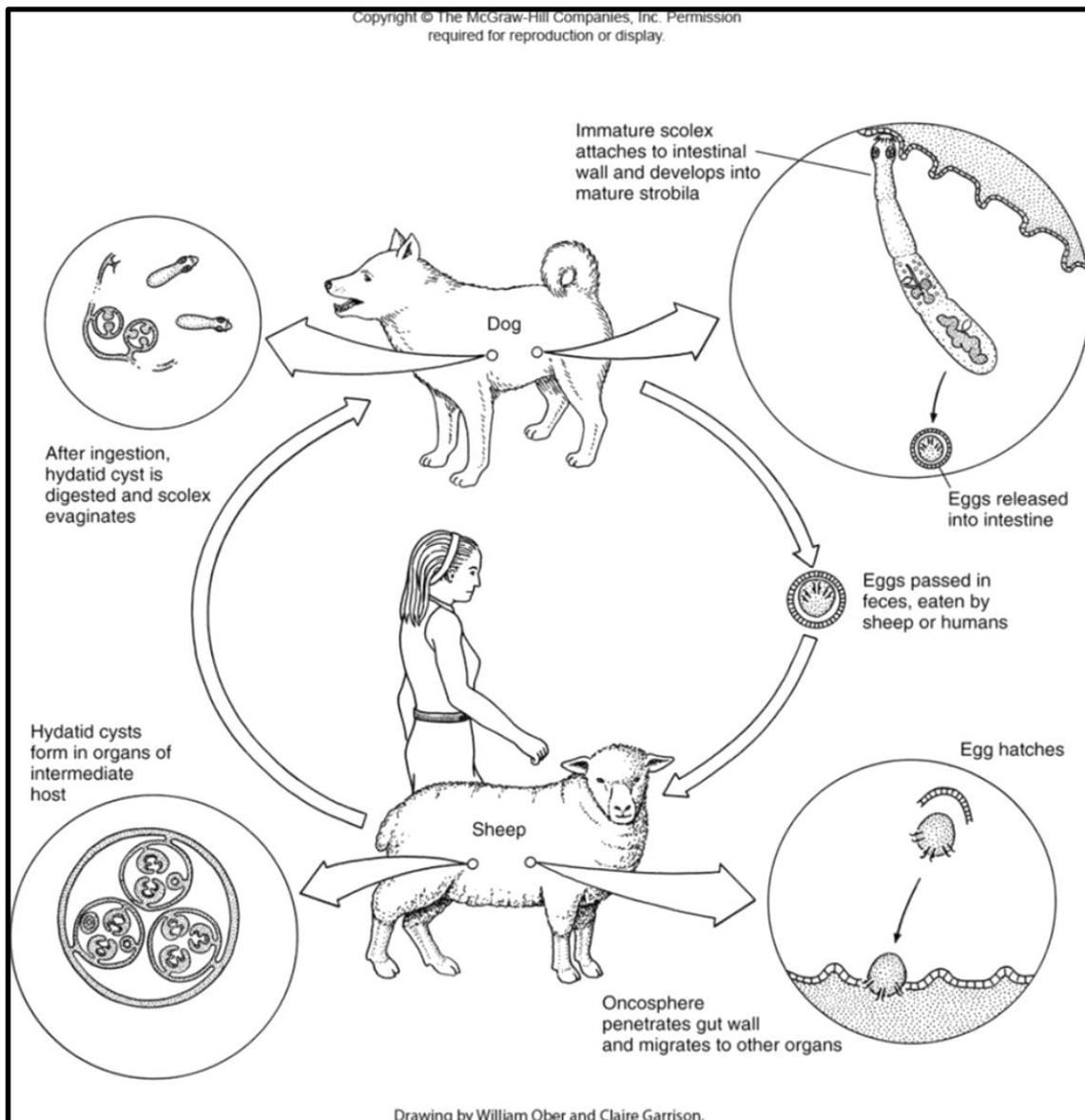


Figura 1 .Ciclo de vida de *E. granulosus*. Extraído de <http://course1.winona.edu/kbates/parasitology/tapeworms.htm>

1.3 Glutación transferasas

Las enzimas glutación transferasas (GSTs) (E.C.2.5.1.18) son una superfamilia de proteínas multifuncionales con roles esenciales en la detoxificación celular de un amplio rango de moléculas exógenas y endógenas. Las GSTs catalizan la conjugación de un nucleófilo celular, el glutación (γ -glutamil-cisteinil-glicina) (GSH), a una amplia variedad de compuestos electrófilos hidrofóbicos. Una vez conjugadas al GSH, las moléculas lipófilicas

reactivas se convierten en sustancias más polares y disminuyen su reactividad, favoreciendo su eliminación (Hayes J.D. y col., 2005). Según estudios evolutivos dicha superfamilia se puede agrupar en 4 familias: 1) citosólicas, 2) mitocondriales o clase Kappa, 3) microsomales (MAPEG) y 4) proteínas bacterianas con resistencia a fosfomicina (Board P.G. y Menon D., 2013).

Las GSTs citosólicas forman el grupo más grande y diverso, contando con miembros en todos los organismos aeróbicos. Existen 8 clases asignadas: Alfa, Mu, Pi, Sigma, Omega, Zeta, Theta y Phi, donde esta última no está presente en mamíferos. Para esta clasificación inicialmente se consideró: especificidad de sustrato, sensibilidad a inhibidores, secuencia de aminoácidos y localización de sus genes en cromosomas (Frova C., 2006; Board P.G. y Menon D., 2013). Más tarde, se observó que sólo las GSTs de mamíferos cumplen con todos los requisitos y dada la amplia variedad de GSTs en otros organismos que no siempre los cumplen, en la actualidad el criterio más fiel para su clasificación es el análisis de la secuencia de aminoácidos. Las GSTs citosólicas son proteínas diméricas tanto con subunidades idénticas como diferentes. La heterodimerización está restringida a subunidades de la misma clase (Board P.G. y Menon D., 2013). Cada subunidad presenta dos dominios: el dominio N-terminal y el C-terminal. En el dominio N-terminal se presenta el sitio de unión para el GSH, denominado sitio G, altamente conservado en las GSTs citosólicas. Es un dominio tipo tiorredoxina compuesto por hélices α y hojas β ($\beta\alpha\beta\alpha\beta\alpha$). Por otro lado, en el dominio C-terminal se presenta el sitio de unión para el sustrato hidrofóbico, denominado sitio H. Está compuesto α -hélices y es menos conservado, lo que le permite la variación en la especificidad de sustrato entre las GSTs (Wilce M.C. y Parker M.W., 1994; Frova C., 2006).

1.4 Mecanismo catalítico

Las enzimas GSTs catalizan el ataque nucleofílico del sustrato fisiológico GSH, a un centro electrofílico de diversos sustratos dando lugar a glutationil conjugados (GS-X). Se representa la reacción de conjugación al sustrato universal 1-cloro-2,4-dinitrobenzeno (CDNB) en la Figura 2:



Figura 2: Sustitución nucleofílica aromática de GSH y CDNB catalizada por las enzimas GST.

Cada GST presenta un residuo característico en su sitio G que es fundamental para la actividad catalítica. Dicho aminoácido participa en la activación del GSH y se ha visto que dependiendo del tipo de GST puede ser Cys, Ser o Tyr. Por ejemplo, para la clase Omega el residuo catalítico es una Cys, que formaría disulfuros mixtos con GSH. Por otro lado, las enzimas de clase Theta y Zeta presentan una Ser, y las de las clases Alfa, Mu, Pi y Sigma una Tyr, donde los hidroxilos de las cadenas laterales de dichos residuos catalíticos, favorecerían la formación y estabilización del GS-anión y posterior ataque nucleofílico sobre el sustrato (Frova C., 2006). Asimismo, otros residuos, altamente conservados, se han descrito como involucrados en la catálisis y/o forman parte del sitio G. Por ejemplo, para la clase Mu de mamíferos, los residuos que forman parte del sitio G son: el Trp8 en la posición adyacente a la Tyr catalítica, y Arg43, Trp46, Lys50, Asn59, Leu60, Gln72, Ser73 y Asp106 (Wilce M.C. y Parker M.W., 1994).

1.5 Glutación transferasas en helmintos parásitos

Los helmintos parásitos parecen tener escasos sistemas de detoxificación. Se han reportado enzimas involucradas en la inactivación de especies oxidantes tales como las peroxirredoxinas, glutación peroxidasas, tioredoxinas (Torres-Rivera A. y Landa A., 2008). Sin embargo, no se han encontrado o tienen poca actividad las enzimas como catalasas y glutación peroxidasas dependiente de selenio (Dzik J.M., 2006). Asimismo, si bien se han identificado algunas enzimas relacionadas con los citocromos P-450, no se ha logrado detectar su actividad hasta el momento (Arbildi P. y col., 2011). Por ello, las GSTs se consideran relevantes en la detoxificación de estos helmintos parásitos. Como se mencionó anteriormente, son enzimas esenciales en la inactivación de un amplio rango de sustancias tanto exógenas como endógenas, que en los parásitos comprendería la inactivación de compuestos generados por el propio metabolismo así como de los generados por el ataque

del sistema inmune del hospedero. Por ejemplo, los componentes oxidantes liberados por las células inmunes efectoras inician la peroxidación de los lípidos de membrana de las células del parásito produciendo hidroxiperóxidos tanto de fosfolípidos como de ácidos grasos. Dichos compuestos son altamente reactivos que se rompen para dar lugar a productos secundarios tóxicos, entre ellos se destaca el 4-hidroxi-2-nonenal (4-HNE) que participa en mecanismos de señalización celular, particularmente en la inducción de proliferación, diferenciación y apoptosis (Arbildi P. y col., 2011). Las GSTs presentan un rol importante en la inactivación de dichas moléculas, las cuales pueden dañar las células del parásito. En mamíferos, las GSTs de clase Alfa estarían a cargo de la detoxificación de los productos primarios y secundarios de la peroxidación lipídica, así como en la regulación de los niveles de 4-HNE (Hayes J.D. y col., 2005). La actividad peroxidasa dependiente de glutatión de la GST inactiva los peróxidos orgánicos, mientras que los productos secundarios los conjuga a GSH a través de la actividad glutatión transferasa. Aunque no se han detectado GSTs de clase Alfa en helmintos, se han descrito que otros tipos de GST parasitarias exhiben actividades capaces de inactivar dichos productos (Torres-Rivera A. y Landa A., 2008). Por lo que, las GSTs de los helmintos parecen ser claves para la biología del parásito, actuando como respuesta al estrés oxidativo, tanto generado de su propio metabolismo como del derivado de las defensas inmunes del hospedero. En la Figura 3 se esquematiza los roles detoxificantes de las GSTs de los helmintos parásitos.

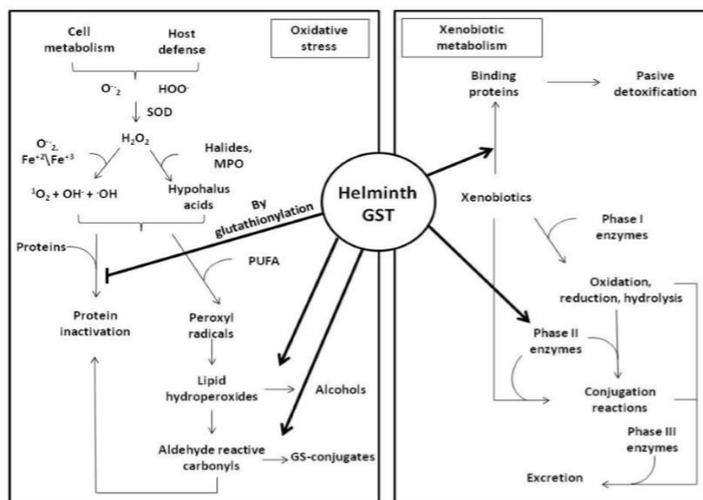


Figura 3. Mecanismos de detoxificación propuestos para las glutatión transferasas de helmintos parásitos. Extraído de Arbildi P. y col., 2011

Glutación transferasa en *E. granulosus*

Así como en otros cestodos, se detectó actividad GST en *E. granulosus* tanto en protoescólex intactos (Repetto Y. y col., 1986) como en la fracción citosólica (Morello y col., 1982). Posteriormente, se purificó por cromatografía de afinidad, una fracción de 25 kDa (EgGST1) con actividad GST (Fernandez C. y Homaeche C.E., 1994) y se reportó su secuencia completa (Fernandez V. y col., 2000). Más recientemente, con el acceso a la información del transcriptoma de *E. granulosus* (Fernandez C. y col., 2002; Parkinson J. y col., 2012), se distinguieron tres GSTs citosólicas y una microsomal. Entre las citosólicas, además de EgGST1 se identificaron otras dos, EgGST2 y EgGST3. Los análisis filogenéticos de las secuencias aminoacídicas deducidas, ubicó inequívocamente a EgGST1 y a EgGST2 dentro de las de la clase Mu y Sigma, respectivamente (Fernandez V. y col., 2000; Iriarte A. y col., 2012; Arbildi P., 2013). Pero los resultados obtenidos para EgGST3 fueron mucho menos claros de interpretar, principalmente debido a que su secuencia no se agrupa con ninguna de las clases de GSTs citosólicas hasta ahora descritas (Iriarte A. y col., 2012; La-Rocca S., 2014). Interesantemente, EgGST3 parece presentar características propias dado que su dominio C-terminal corresponde claramente a uno de clase Sigma mientras que su dominio N-terminal presenta la mayoría de los aminoácidos característicos descritos para la clase Omega (Iriarte A. y col., 2012; La-Rocca S., 2014). Al mismo tiempo, estudios realizados a nivel del genoma de *E. granulosus*, han reportado una expansión de los genes de clase Mu en este parásito (Tsai I.J. y col., 2013), donde uno de ellos corresponde al de EgGST1, obtenido previamente por nuestro grupo (Fernandez V., 2001).

De las GSTs citosólicas identificadas en *E. granulosus*, EgGST1 ha sido la más extensamente estudiada. Se caracterizó bioquímicamente la forma recombinante (EgGST1r) que mostró ser capaz de conjugar GSH a algunos productos secundarios de la peroxidación de lípidos, y exhibir actividad peroxidasa sobre hidroperóxido de cumeno (Harispe L. y col., 2010). Asimismo, se analizó sus parámetros cinéticos y se profundizó en el mecanismo cinético que sigue EgGST1 (Harispe L. y col., 2010; Arbildi P., 2013). Con respecto al

análisis de la secuencia aminoacídica completa de EgGST1, es de destacar la sorprendentemente alta similitud con EMGST1, la enzima descrita en *E. multilocularis* (Liebau E. y col., 1996). Ambas se relacionan fuertemente con las GSTs de clase Mu de mamíferos, ya sea filogenéticamente como por la conservación de aminoácidos descritos como invariantes en el extremo N-terminal, incluyendo la Tyr catalítica (Fernandez V. y col., 2000). Además, los cestodos, pero no los trematodos analizados, conservan el motivo “mu loop” completo (Fernandez V. y col., 2000) que comprende los aminoácidos entre las posiciones 33 y 42 (Ji X. y col., 1992). Dicho motivo es un elemento estructural característico descrito para las enzimas de clase Mu de mamíferos que crea un sitio activo más estrecho y menos expuesto al solvente comparado con las GSTs de otras clases (Wilce M.C. y Parker M.W., 1994). Pero la similitud de EgGST1 con esta clase de GSTs va más allá de la secuencia de aminoácidos: el gen que codifica para EgGST1 conserva la estructura exón/intron descrita para los genes de clase Mu de mamíferos (Fernandez V., 2001; Iriarte A. y col., 2012).

El presente trabajo se centra principalmente en explorar la evolución molecular de las glutatión transferasas de clase Mu de *Echinococcus granulosus* (EgGST1). En este sentido, se obtuvieron todos los genes que codifican para secuencias “homólogas” a EgGST1, en el genoma de *E. granulosus* y se analizó los sitios conservados de cada secuencia para entender su relevancia, en particular la evolución del mu-loop. Asimismo, se realizó una filogenia de las correspondientes secuencias aminoacídicas deducidas para mapear los eventos de duplicación. Finalmente, se analizó la expresión de dichos genes por el parásito, identificando los ARN mensajeros correspondientes por RT-PCR.

2. MATERIALES Y METODOS

2.1 ANALISIS INFORMÁTICO

2.1.1 Obtención y análisis de secuencias

Las secuencias proteicas y nucleotídicas, incluyendo EgGST1, fueron obtenidas en la base de datos de GeneBank (<http://www.ncbi.nlm.nih.gov>). El Scaffold 007, donde se localizaron los genes de *E. granulosus*, fue obtenido de la base de datos del Instituto Sanger, libremente disponible en www.sanger.ac.uk/resources/downloads/helminths/echinococcus-granulosus.html. Se buscaron las secuencias homólogas en el genoma de este organismo utilizando la herramienta online TBLASTN (protein vs. Transl. DNA) en la base antes mencionada. Se utilizó como *query* la secuencia proteica completa de EgGST1. En donde se realizó una anotación manual, se chequeó los sitios dador y aceptor para empalme utilizando el programa BioEdit (Hall T.A., 1999). Finalmente se extrajeron las secuencias de los exones para su traducción y posterior análisis filogenético.

2.1.2 Alineamiento de secuencias, modelo evolutivo y análisis filogenético

Los alineamientos fueron construidos con dos programas: MUSCLE (Edgard R.C., 2004) y PROBCONS (Do C.B. y col., 2005). Se predijo el mejor modelo evolutivo para los alineamientos utilizando el programa MEGA versión 6.0 (Tamura K. y col., 2013). Para la construcción de los arboles filogenéticos se emplearon los métodos *Neighbor-joining* implementado en el programa MEGA 6.0 y *Maximum likelihood* implementado en el programa PHYML versión 3.1 (Guindon and Gascuel, 2003). En archivo anexo I se lista las secuencias utilizadas (Anexo I-Tabla 1).

2.1.3 Modelado de secuencias

Se realizó una predicción de la estructura secundaria y espacial de las secuencias obtenidas con el programa Swiss model <http://swissmodel.expasy.org/interactive>.

2.2 ANALISIS DE LA EXPRESIÓN EN EL PARASITO, DE LOS GENES IDENTIFICADOS

Para las cuatro secuencias identificadas en el ítem anterior, se procedió a analizar la expresión del correspondiente gen en el parásito. Para ello la estrategia seguida fue: obtención de ARN total parasitario, preparación de ADN copia, y amplificación de ADN por PCR convencional, utilizando los cebadores específicos diseñados en este trabajo.

2.2.1 Obtención de ARN total del parásito

Para la obtención de ARN parasitario se partió de protoescólex de *E. granulosus* conservados a -70°C provenientes de quistes hidáticos de origen bovino obtenidos según (Baz A. y col., 1995). La purificación del ARN total se realizó utilizando el protocolo del reactivo de Trizol Reagent (Life Technology) con pequeñas modificaciones. Brevemente, se homogenizaron 250 μl de protoescólex con 750 μl de reactivo. Posteriormente se agregó el Cloroformo agitándose vigorosamente por 15 minutos. Luego de centrifugar, la fase acuosa se precipitó con un volumen de isopropanol y el sedimento se lavó con Etanol 75% y se secó a temperatura ambiente. Finalmente, el ARN total se resuspendió en 100 μl de agua libre de ARNasas, durante 10 min a 60°C . La concentración se determinó mediante absorbancia a 260 nm (A_{260}) en un espectrofotómetro Nanodrop (Thermo Scientific NanoDrop™ 1000 Spectrophotometer). La cuantificación de ácidos nucleicos se asumió que 1 unidad de A_{260} de ARN simple hebra corresponde a 40 $\mu\text{g}/\text{ml}$. También se analizó la contaminación por proteínas determinando la relación A_{260}/A_{280} .

2.2.2 Análisis electroforético en geles de Agarosa

La calidad de la preparación de ARN total se analizó por electroforesis en geles de Agarosa al 2% en TAE 1x (Tris-Acetato 40 mM y EDTA 1 mM) a 100 V (Sambrook J. y Russell D., 2001). La muestra de ARN se mezcló con tampón de carga geles de Agarosa (Glicerol 50%, EDTA 0.2M pH 8.3, Azul de Bromofenol 0.05%). En todos los casos la Agarosa se preparó en dicho buffer conteniendo Bromuro de Etidio 0.5 mg/ml y el ARN se visualizó por transiluminación con luz UV.

2.2.3 Digestión de ADN genómico

Se removió el posible ADN genómico que esté contaminando la preparación de ARN total utilizando DNase I (Fermentas) siguiendo las indicaciones del fabricante. Brevemente, se trató 1 µg de ARN total con 0.5 U de la enzima por 30 minutos a 37 °C y luego se inactivó con 10 µl de EDTA 25 mM durante 10 minutos a 65 °C.

2.2.4 Síntesis de ADN copia

Para la obtención del ADN copia (ADNc) se utilizó la transcriptasa reversa M-MVL (SBS) siguiendo las indicaciones del fabricante. Brevemente, luego de tratar por 10 minutos a 70°C el ARN extraído (1 µg) junto con los hexámeros al azar (10 pmoles, SBS), se le agregó una mezcla conteniendo dNTPs (0.5 mM, Fermentas), DTT (10 mM, SBS) y M-MVL (200 U, SBS) y se incubó durante 60 minutos a 37°C. El ADNc así obtenido se diluyó a la cuarta parte y se conservó a – 20 °C hasta su uso.

2.2.5 Reacción de amplificación en cadena (PCR)

Las amplificaciones de ADN se realizaron por PCR convencional utilizando un termociclador automático (GeneAmp PCR System 2400, Perkin Elmer). Cada reacción se realizó en un volumen final de 20 µl conteniendo 1.5 mM MgCl₂ (Fermentas), 0.25 mM de cada dNTP (Fermentas) y 0.2 U de la enzima Taq polimerasa (Fermentas) en su *buffer* de reacción (Fermentas) así como 1 µl ADNc y 1 mM de cada *Primer* específico (ver más adelante). El programa usado se inicia con un paso de desnaturalización de 95°C por 5 minutos, seguido de 40 ciclos de tres temperaturas (94°C, 58°C y 72°C) durante 30 segundos cada una. Finalmente, un período de extensión de 10 minutos a 72°C.

2.2.6 Análisis electroforético en geles de poliacrilamida

El producto de la amplificación se analizó por electroforesis en geles de poliacrilamida al 6% de Acrilamida-Bisacrilamida (38:2) en TAE 1X a 180V.

Cada gel se polimerizó con TEMED 0.1% (v/v) y APS 0.1% (p/v). La muestra se mezcló con tampón de carga (Formaldehído 98%, EDTA 20 mM (pH 8), Azul de Bromofenol 0.05%). El ADN se visualizó por tinción con nitrato de plata según (Sanguinetti C.J. y col., 1994). Brevemente, luego de la electroforesis de las muestras amplificadas, cada gel de poliacrilamida se fijó en Etanol 10% con ácido Acético 0.5% y luego se traspasó a Nitrato de Plata 0.2%. Las muestras se visualizaron con Formaldehído 0.5% en NaOH 3%, deteniéndose la reacción con la solución fijadora.

2.2.7 Diseño de *primers* específicos

Para analizar la expresión por parte del parásito de los genes identificados en el presente trabajo, se diseñó 1 juego de cebadores (*primers*) para cada uno de ellos. En este sentido, se alinearon las secuencias aminoacídicas deducidas de todos los *Hits* y se eligieron las zonas más conservadas (que no contengan *gap*) en las 4 secuencias. En dichas zonas de la secuencia aminoacídica, se indicaron los sitios donde estarían ubicados los intrones, con el fin de diseñar los *primers* en dos exones diferentes para distinguir los amplicones provenientes de ADN genómico. De esta forma, se seleccionó la zona comprendida entre los aminoácidos 108 y 134 de la secuencia *Hit 2* que corresponde a EgGST1 descrita por nuestro grupo (Fernandez V. y col., 2000). Asimismo, se consideraron los siguientes parámetros:

- Temperatura de hibridación: 58-60 °C
- Tamaño de amplicón: 70 - 100 pb
- Tamaño de oligonucleótidos: 20 - 23 nt
- Al menos uno de los dos últimos nucleótidos del extremo 3' del *primer* de un *Hit* sea diferente a los correspondientes de los otros *Hits*.

Como controles se utilizaron los juegos *primers* para genes parasitarios: ARN ribosomal 18s (Egr18s), Actina (EgrActI) y rtEgGST1 diseñados en base a los mismos parámetros y obtenidos previamente por el grupo de trabajo (Arbildi P., 2013). En la Tabla 2 se especifican la secuencia de los *primers* utilizados en este trabajo.

Tabla 2. Secuencias de *primers* específicos de *E. granulosus* para ensayos de PCR convencional.

<i>Gen</i>	<i>Secuencia (5'→3')</i>		<i>Tamaño del amplicón (pb)</i>
Egr18s	Fw	TCAACACGGGAAAACCTCACC	79
	Rv	ACCACCAACCACCAAATCAA	
EgAct1	Fw	CGTGGTCTTGTCTGGTGGAT	72
	Rv	AGGGCAGTCAGGTCCTTGTT	
rtEgGST1	Fw	AGGCCTACTTGTGCGGTTTT	101
	Rv	CCATTTTGCACCTTGCTTCATC	
<i>Hit 1</i>	Fw	TATTCGCAACGGCTTCTACCG	87
	Rv	TTTCATTGGAAGGTTTTTCATG	
<i>Hit 2</i>	Fw	TGCGCATGGCGTTTACGAGG	85
	Rv	TTCTGAGCCAGCGTCTCAAAT	
<i>Hit 3</i>	Fw	TTGCGAATGGCGTTTGGGATG	85
	Rv	TTTTGTGCCAAGGTGTCCAAG	
<i>Hit 4</i>	Fw	TGCGCATCGCGTTTGCGAAG	85
	Rv	TTTTGTGCCAGTGCCTCCAAA	

3. RESULTADOS

En esta sección se identifican los genes de *E. granulosus* que codifican para aquellas GSTs homólogas a EgGST1. Los resultados obtenidos con el programa MUSCLE y el método Neighbor-joining presentaron un mejor ajuste con nuestros datos. Se empleó estos datos para los siguientes análisis. Se presentan los resultados del estudio de la organización de dichos genes (estructura exón/intrón), análisis de las secuencias nucleotídicas y el análisis filogenético de sus secuencias deducidas en el contexto de otras GSTs citosólicas. Finalmente, se analiza la expresión de estos genes por RT-PCR.

3.1 Identificación de genes y análisis de la estructura exón/intrón

Se realizó un BLAST utilizando la secuencia de EgGST1, en la base de datos del genoma de *E. granulosus* como se indica en la sección anterior. Se encontró que el *Scaffold* 007 contiene a todos los genes de la clase Mu de este parásito. Utilizando el programa BioEdit (Hall T.A., 1999) se realizó una anotación manual; las coordenadas del *Hit* obtenido en el blastx se utilizaron como punto de partida. Se identificaron 4 secuencias, denominadas *Hit* 1, 2, 3 y 4, comenzando en todos los casos por el exón 4 y según el gen, se registraron hasta cinco exones (Figura 4B). La longitud de los exones identificados mostró ser absolutamente conservada, el exón 4 es de 82 pb, el exón 5 de 101 pb, el exón 6 de 96 pb, el exón 7 de 111 pb y el exón 8 de 96 pb (Figura 4) y coincide con lo descrito para los genes de GSTs de clase mu. Se observó que los *Hits* 1, 2 y 4 corresponden a secuencias completas ya reportadas, mientras que el *Hit* 3 corresponde a una secuencia reportada parcialmente que se completó en este trabajo. En la Tabla 3 se indican los nombres de dichas secuencias reportadas y su número de acceso al banco de datos (UniprotKB) en relación a cada *Hit*.

Tabla 3: Lista de las secuencias reportadas que corresponden a cada *Hit*.

	Nº de Acceso (UniprotKB)	Longitud en aa
<i>Hit</i> 1	AOA068WTH4	220
<i>Hit</i> 2	O16058	219
<i>Hit</i> 3	A0A068WTJ7	210 @
<i>Hit</i> 4	U6JCG1	201

@corresponde al número de aminoácidos de secuencia deducida completa, es decir, de la secuencia parcial reportada que se completó en este trabajo

A)

	Coordenadas ADN genómico desde	Coordenadas ADN genómico hasta	Coordenadas en el ARN mensajero	Longitud del intrón (pb) *
HIT 1				
Exón 4	2131221	2131302	60-86	217
Exón 5	2131519	2131619	87-120	68
Exón 6	2131687	2131782	121-152	
HIT 2				
Exón 4	2361488	2361569	60-86	91
Exón 5	2361660	2361760	87-120	295
Exón 6	2362055	2362150	121-152	203
Exón 7	2362353	2362463	153-189	93
Exón 8	2362556	2362651	190-220	
HIT 3				
Exón 4	2368935	2369016	60-86	91
Exón 5	2369107	2369207	87-120	76
Exón 6	2369283	2369378	121-152	150
Exón 7	2369528	2369638	153-189	80
Exón 8	2369718	2369813	190-220	
HIT 4				
Exón 4	2374246	2374327	60-86	91
Exón 5	2374418	2374518	87-120	84
Exón 6	2374602	2374697	121-152	222
Exón 7	2374919	2375014	153-189	99
Exón 8	2375113	2375208	190-220	

*Corresponde al intrón posterior del exón

B)

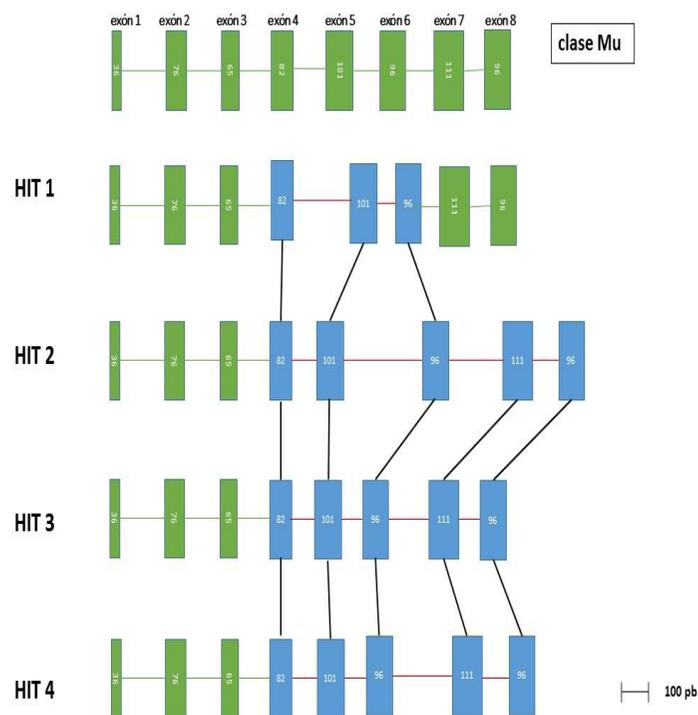


Figura 4. Estructura de los genes identificados. A) Tabla donde se muestran las coordenadas en el *Scaffold 007*, de cada exón identificado para cada *Hit* así como su longitud y la del intrón inmediato. **B)** Representación esquemática de la estructura exón/intrón para cada *Hit*. En color verde se representan los exones y tamaños descritos para clase Mu de mamíferos. En color azul, dibujados a escala según la barra, se indican los exones identificados cuyo tamaño fue confirmado en este trabajo.

3.2 Análisis filogenético

Se alinearon un total de 70 de secuencias proteicas empleando los programas Muscle y Probcons. Para construir los árboles filogenéticos se incluyeron GSTs citosólicas de diferentes clases (Mu, Pi y Sigma) y como grupo externo se utilizaron secuencias de GSTs Microsomales (MAPEG). Todas las secuencias fueron obtenidas de varios organismos, entre ellos mamíferos y helmintos (ver Anexo I).

En la filogenia se observó un grupo principal que incluye las GSTs citosólicas de mamíferos, todas las de platelmintos de la clase Mu y un pequeño grupo de clase Pi con un apoyo estadístico de 77% (Figura 5, ). Por fuera se forman dos grupos de GSTs: uno que reúne a las de clase Sigma y otro al grupo externo (MAPEG). Dentro del grupo principal, con alto apoyo estadístico, 97 %, aparecen juntas las GSTs de clase Mu de platelmintos con las GSTs de mamíferos (Figura 5, ). Las GSTs de cestodos forman un grupo monofilético con apoyo estadístico bajo (33%, Figura 5, ) , quedando por fuera la única GSTs de trematodos incluida en el estudio. A su vez, los *Hits* 1, 2, 3 y 4 se asocian con las GSTs de clase Mu de cestodos (Figura 5, indicado con “Mu de platelmintos”).

Dentro de la Clase Mu se pueden identificar 2 grupos de ortólogos con un apoyo estadístico de 100%: 1) *Hit* 1 con EmW 535300; 2) *Hit* 2 con EmW 538300 y TsM 639300 (Figura 5, ). Si bien no es parte del estudio resulta interesante identificar varios eventos de duplicaciones recurrentes, entre ellos los casos de *H. microstoma* con un apoyo estadístico de 96% y *E. multilocularis* con un apoyo estadístico de 99% (Figura 5, ).

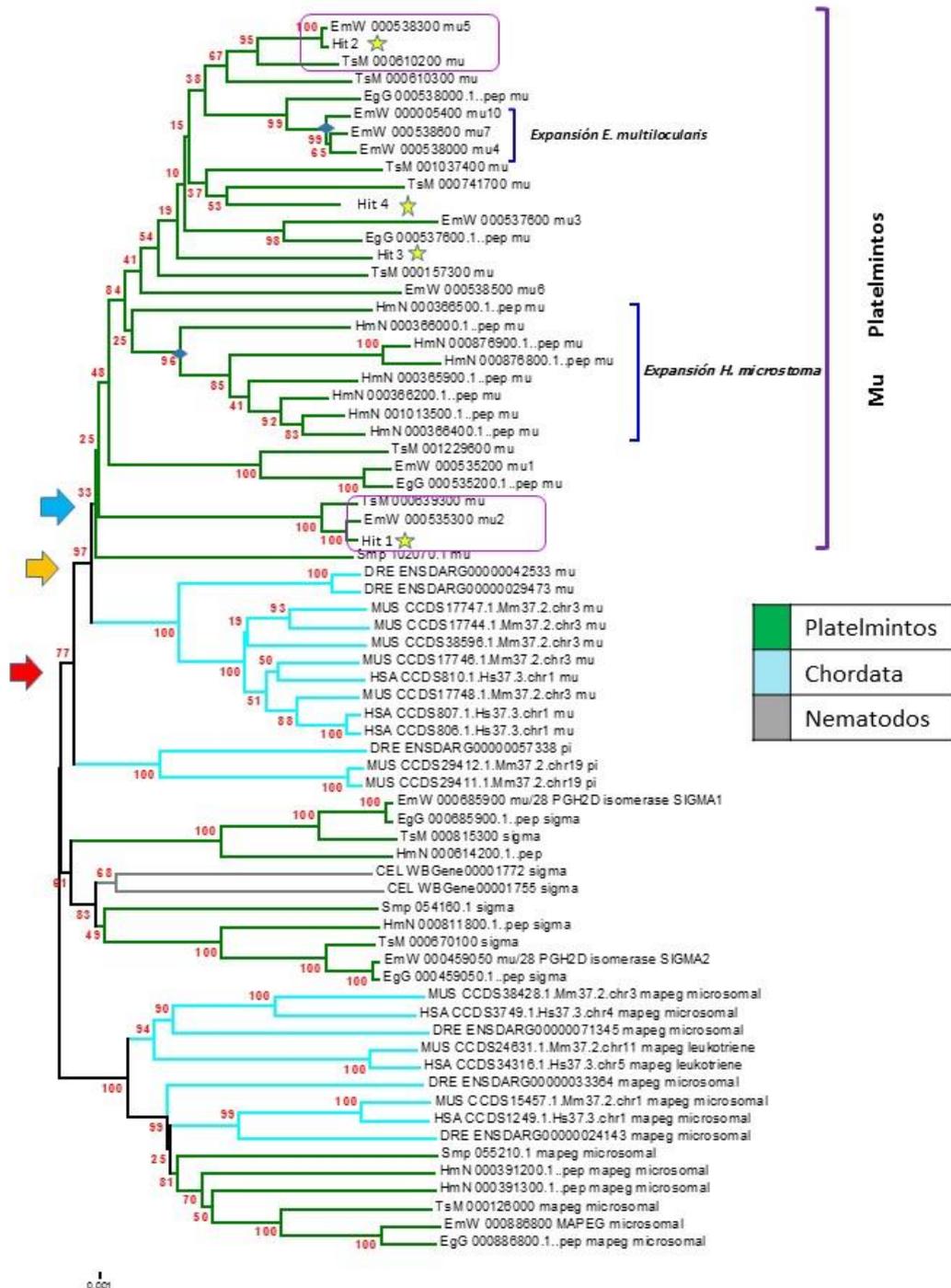


Figura 5. El árbol filogenético fue construido utilizando el programa MEGA 6.0. Se empleó el método Neighbor-Joining a partir de 1000 replicas (*bootstrap*). Se indica las GSTs de clase Mu de platelmintos, GSTs citosólicas ◆, GSTs de clase Mu de platelmintos con las de mamíferos ◆, grupo monofilético de cestodos ◆, ortólogos , duplicaciones ◆. Se señala además las 4 secuencias empleadas en este trabajo ★. En "Anexo" se listan las secuencias incluidas en el análisis y el número de acceso.

3.3 Análisis de secuencias aminoacídicas

Para los alineamientos, se obtuvieron las secuencias nucleotídicas completas. Para ello se completó *Hit 3* y luego se extrajeron las 4 secuencias nucleotídicas del banco de datos del Instituto Sanger. Las secuencias codificantes completas para cada *Hit* se dedujeron usando el programa BioEdit. En base a los alineamientos de secuencias aminoacídicas deducidas completas se buscaron motivos y residuos descritos como característicos para las GSTs de clase Mu, principalmente aquellos involucrados en su mecanismo catalítico (Wilce M.C. y Parker M.W., 1994). Como se esperaba, en el extremo N-terminal se encontró la mayor parte de los residuos identificados como invariantes en las GSTs de clase Mu, destacándose la Tyr considerada esencial para la actividad catalítica (Atkinson H.J. y Babbitt P.C., 2009) en la posición 7 de los cuatro *Hits* (Tyr(n), Figura 6, sombra negra). Asimismo, los *Hits* 1, 2 y 4 mostraron todos los residuos descritos que formarían el sitio de unión a glutatión: Tyr(n), Trp(n+1), Arg(n+36), Trp(n+39), Lys(n+43), Asn(n+52), Leu(n+53), Gln(n+65), Ser(n+66) y Asp(n+99) (Figura 6, indicado ≠) (Wilce M.C. y Parker M.W., 1994). En cambio, la secuencia del *Hit 3* mostró 7 de los 10 residuos conservados. La secuencia deducida de *Hit 3* contiene 210 aminoácidos (Tabla 3), por lo que para optimizar el alineamiento con el resto de las de los *Hits*, fue necesario introducir dos quiebres o *gaps*. Uno de ellos, ubicado entre las posiciones n+35 y n+40 tomando como referencia la secuencia del *Hit 1* (de 220 residuos, Tabla 3), incluye la zona donde debieran ubicarse dos de los residuos faltantes: Arg(n+36) y Trp(n+39). Por otro lado, el residuo invariante Asp(n+99), que para el *Hit 3* se ubicaría en la posición “n+93”, fue reemplazado por una Asn (Figura 6, sombra rojo oscuro).

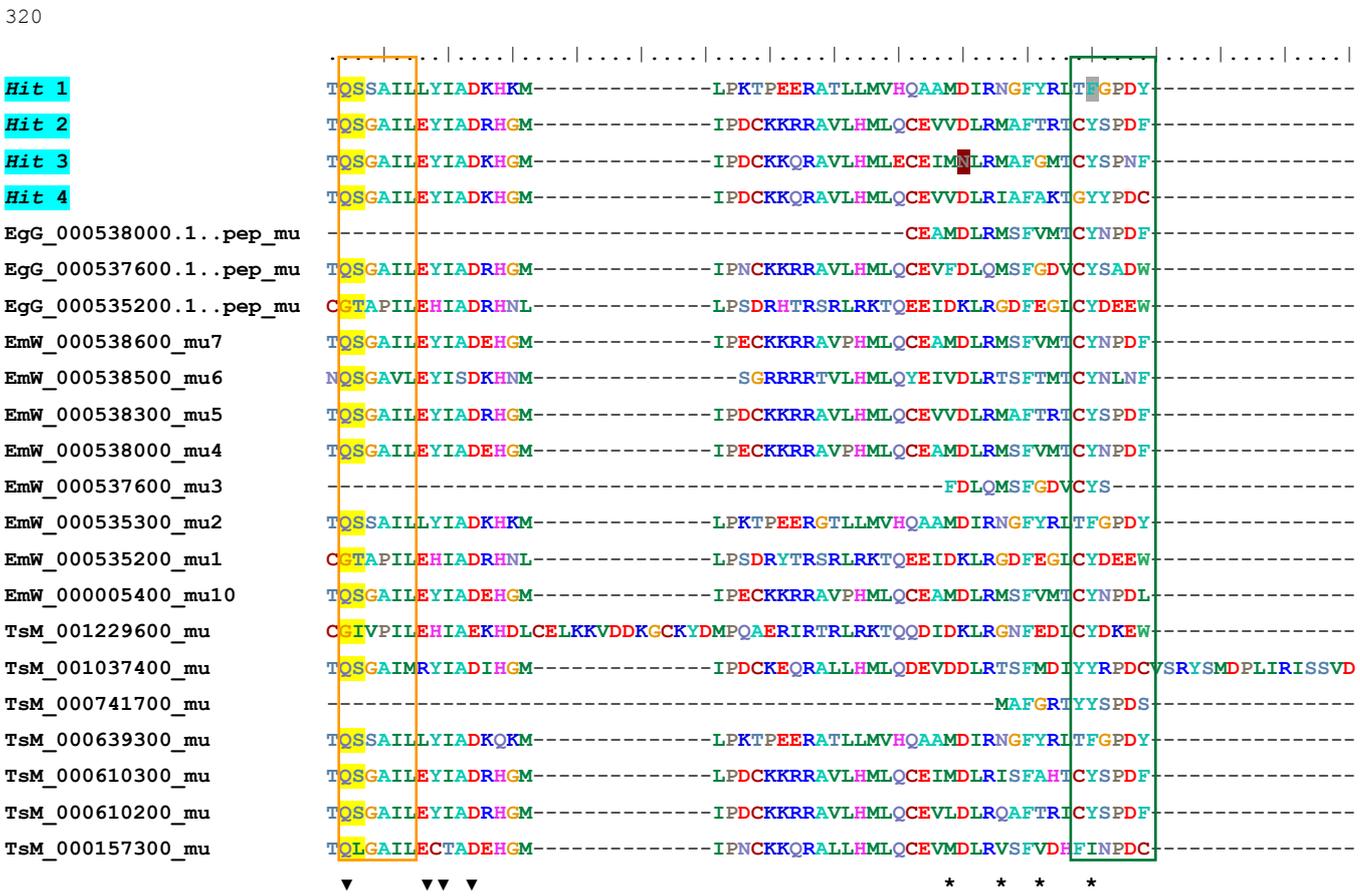
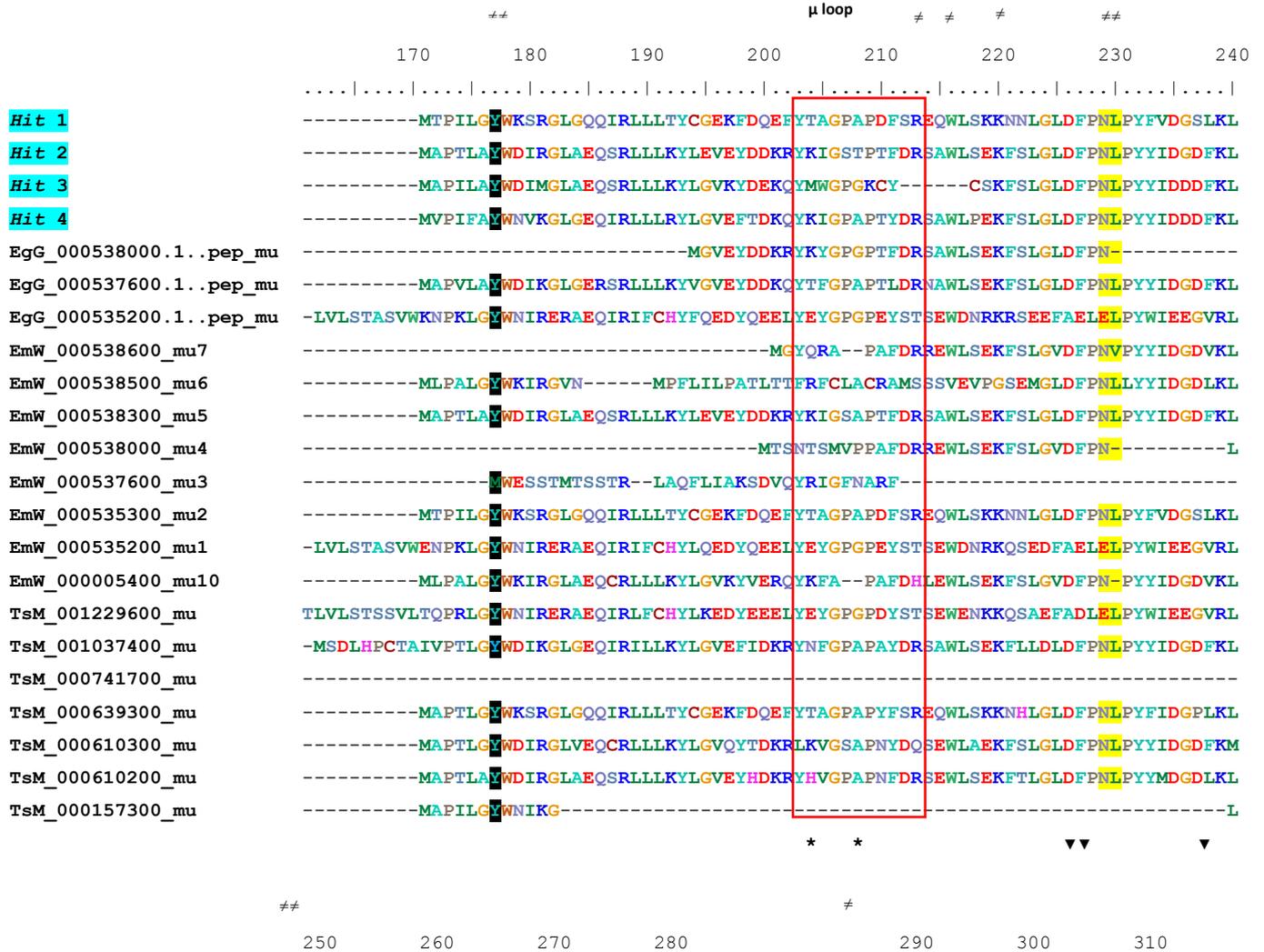
Resultó llamativo encontrar que los residuos conservados Asn(n+52) y Leu(n+53) (Figura 6, sombra amarilla) forman parte de un motivo peptídico (en código de una letra: LGLDFPNLPY) altamente conservado en los cuatro *Hits* de este trabajo, así como también en algunas de las GSTs de otros cestodos incluidos en este estudio. Además, los residuos invariantes Gln(n+65) y Ser(n+66) forman parte de la secuencia (Q)SGAIL (Figura 6, recuadro color naranja) que es homóloga a la secuencia consenso, (Q)[S,T][R,N]AIL,

característica de las GSTs (Wilce M.C. y Parker M.W., 1994) (Figura 6, sombra amarilla).

Asimismo, se han reportado 8 aminoácidos que están dentro o cercanos al sitio de unión al co-sustrato hidrofóbico (sitio H): Tyr(n+109), Ala(n+31), Ala(n+105), Ala(n+206), Met(n+28), Met(n+98), Met(n+102) y Met(n+205), (Wilce M.C. y Parker M.W., 1994). Dichos residuos fueron ubicados en la secuencia de al menos uno de los *Hits*, aunque ninguno de los cuatro *Hits* presentó la totalidad de los aminoácidos señalados. De esta forma se hallaron 4 en el *Hit* 1 (Ala(n+31), Ala(n+206), Met(n+98) y Met(n+205)), 3 en el *Hit* 2 (Tyr(n+109), Ala(n+206) y Met(n+102)), 5 en el *Hit* 3 (Tyr(n+109), Ala(n+206), Met(n+28), Met(n+98) y Met(n+102)) y 3 en el *Hit* 4 (Tyr(n+109), Ala(n+105) y Ala(n+31)). Es de hacer notar que la Tyr(n+109) se encontró en todas las secuencias excepto en la del *Hit* 1 donde está reemplazada por una Phe (Figura 6, sombra gris) y que la Ala(n+206) se encontró en los *Hits* 1, 2 y 3 (Figura 6, sombra violeta), no así en *Hit* 4 dado que la secuencia es más corta (201 aminoácidos, Tabla 3)

Además, tres de los cuatro *Hits* mostraron claramente el *mu loop* (Figura 6, recuadro color rojo). Es un elemento estructural distintiva de las GSTs de clase Mu, que se extiende por los residuos de aminoácidos 33 a 43 (Ji X. y col., 1992), junto al sitio de sustrato de xenobióticos. Es uno de los tres elementos estructurales que crean una hendidura del sitio activo que lo hace más estrecho y menos expuesto a disolvente que en la clase pi (Wilce M.C. y Parker M.W., 1994).

Se realizó la búsqueda de dominios conservados en la base de datos CDD (Conserved Domain Database) utilizándose como molde los *Hits* 1, 2, 3, y 4. Se observó en el dominio N-terminal los residuos que constituyen el sitio G (Figura 6, ≠), los residuos que constituyen la interface de dimerización (Figura 6, ▼), y los residuos pertenecientes a la interface C-terminal (Figura 6, *).



3.4 Modelado de secuencias

En base a los modelos obtenidos, para *Hit 3* los residuos que corresponderían al motivo *mu-loop* según la Figura 6, están formando parte de la hélice $\alpha 2$ (Figura 7). Se puede apreciar en la estructura tridimensional (Figura 8) que para este *Hit* se reduce considerablemente el *mu-loop* con respecto a los *Hits* 1, 2 y 4. Los valores GMQE, QMEA4NA indican la bondad del modelo obtenido con el programa *Swiss model*.

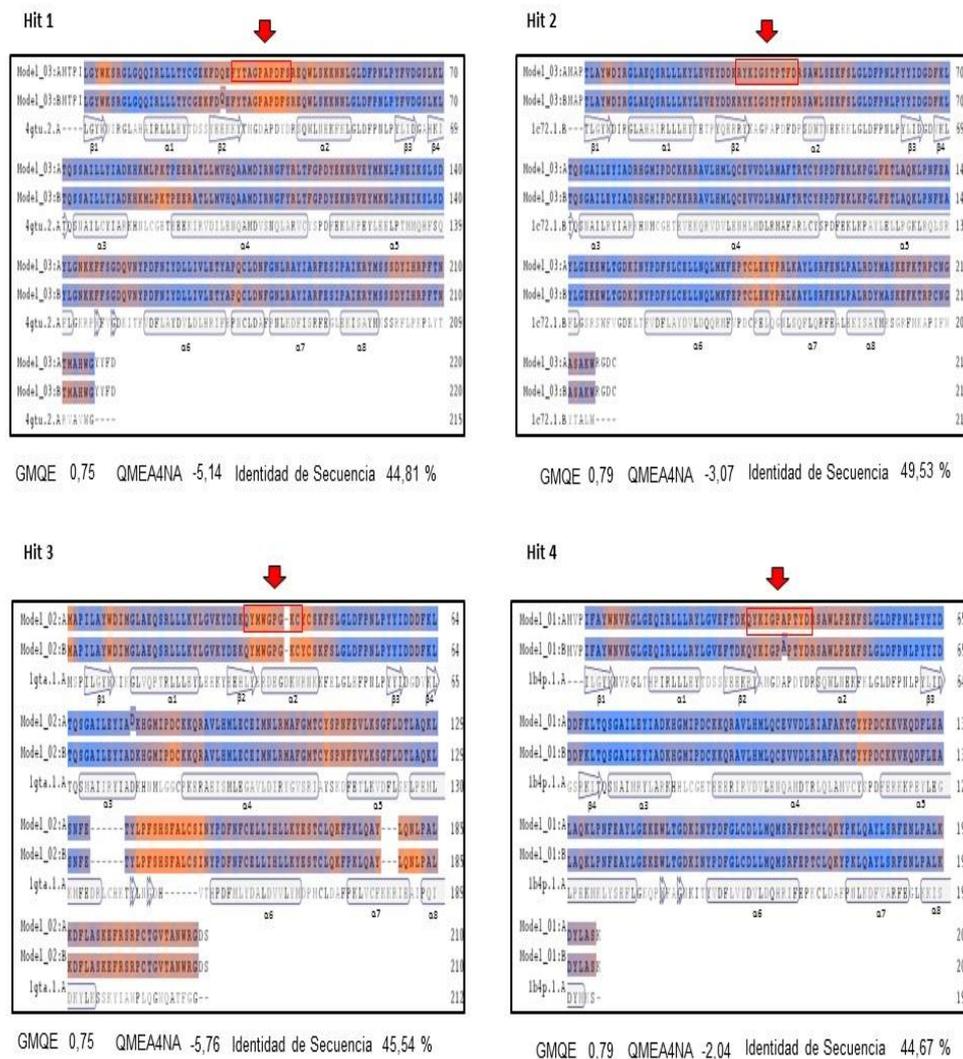


Figura 7 - La estructura secundaria se dedujo con Swiss model <http://swissmodel.expasy.org/interactive>, se representa debajo del alineamiento. Se utilizaron los siguientes moldes: *Hit 1* PDB-ID: 4gtu.2 cadena "A"; *Hit 2* PDB-ID: 1c72.1 cadena "B"; *Hit 3* PDB-ID 1gta.1 cadena "A"; *Hit 4* PDB-ID 1b4p.1 cadena "A". Los valores GMQE, QMEA4NA e Identidad de Secuencia se detallan para cada *Hit*. El recuadro y flecha en rojo indican la secuencia que correspondería al *mu-loop*.

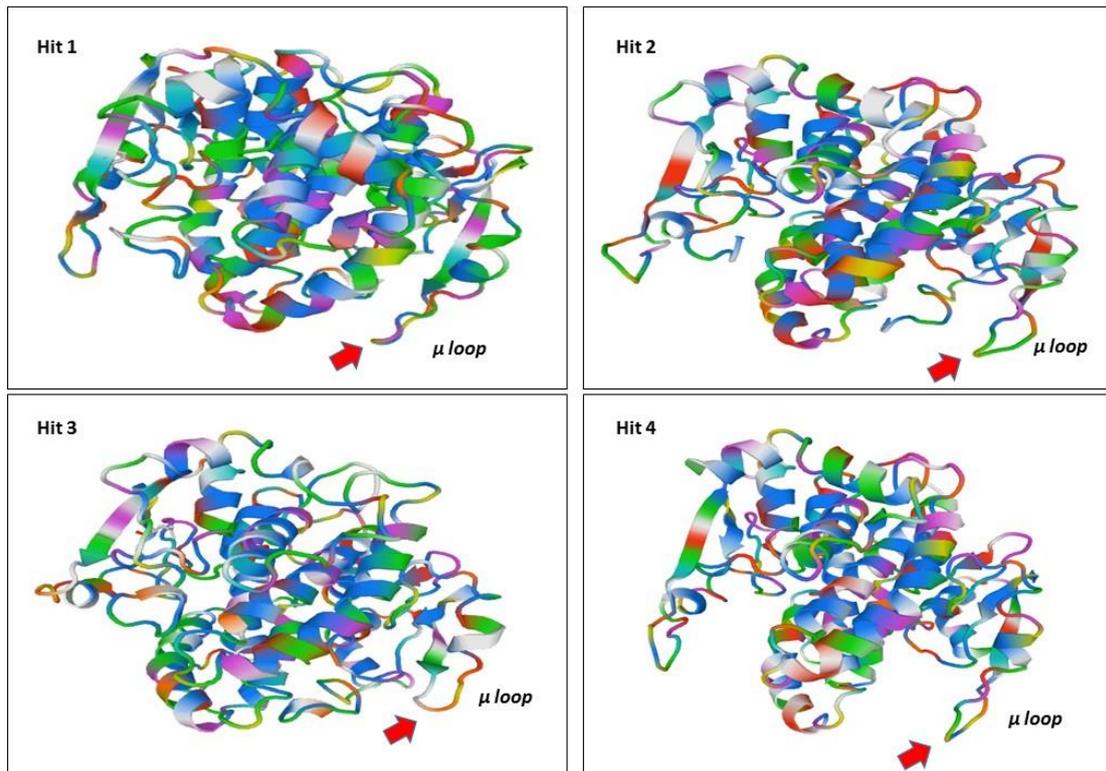


Figura 8 - Modelo de la estructura tridimensional de los *Hits* 1, 2, 3 y 4 obtenida con Swiss model <http://swissmodel.expasy.org/interactive> (figura 11). Se indica en cada recuadro el *mu-loop* (flecha roja).

3.5 Análisis de la expresión de los genes

Se analizó la expresión de los cuatro genes identificados en el presente trabajo. Como se mencionó en la sección anterior, se aisló el ARN de protoescólex y se retrotranscribió. Luego se amplificó por PCR convencional utilizando los juegos de *primers* listados en la Tabla 2 y se analizó en geles de poliacrilamida visualizándose con nitrato de plata. Las amplificaciones se realizaron por cuatruplicados independientes y en las Figuras 7 y 8 se muestran un resultado representativo. Se observaron amplicones correspondientes a los controles internos (Figuras 9 y 10) y a las secuencias de los *Hit 1* e *Hit 2* (Figura 9). No se obtuvo productos de amplificación para los *Hit 3* e *Hit 4* (Figura 10). Vale la pena acotar que, donde se produjo amplificación, se obtuvo un único amplicón con un peso molecular esperado (Figuras 9 y 10).

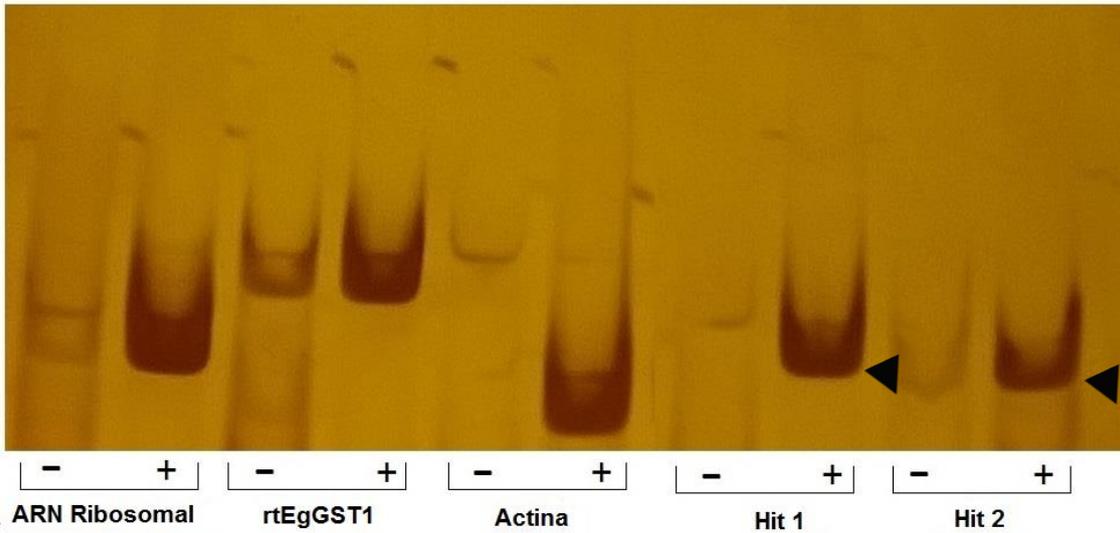


Figura 9. Electroforesis en gel de poliacrilamida. Controles: ARN ribosomal 18s (Egr18s, 79 pb), rtEgGST1 (101 pb) y Actina (EgrAct1, 72 pb); amplicón *Hit 1* (87 pb) y *Hit 2* (85 pb); (-) sin ADNc (+) con ADNc. Las puntas de las flechas indican el producto de amplificación.

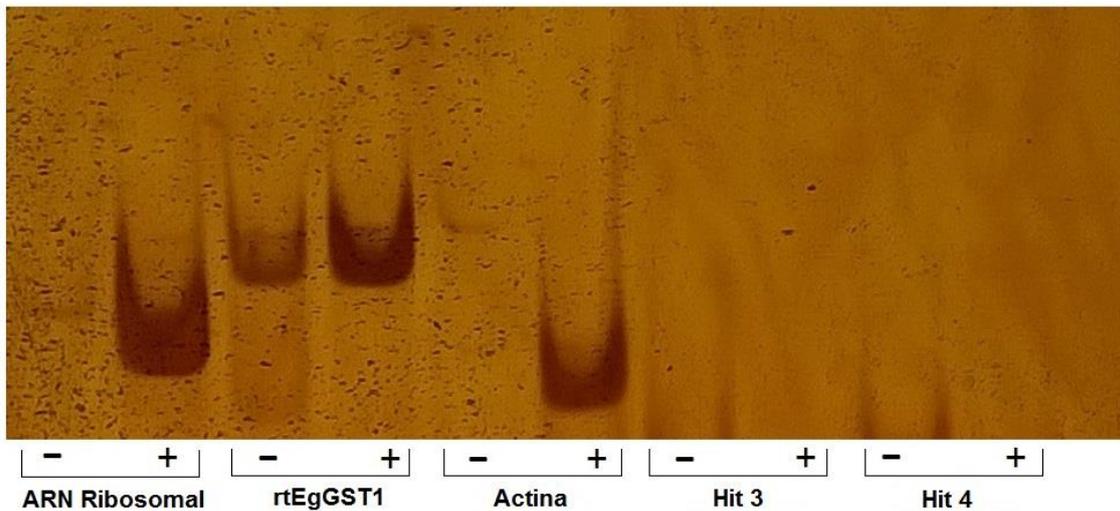


Figura 10. Electroforesis en gel de poliacrilamida. Controles: ARN ribosomal 18s (Egr18s, 79 pb), rtEgGST1 (101 pb) y Actina (EgrAct1, 72 pb); ausencia de producto de amplificación para los *Hit 3* (85 pb) y *Hit 4* (85 pb); (-) sin ADNc (+) con ADNc.

4. DISCUSIÓN

La hidatidosis es una enfermedad provocada por el parásito helminto *E. granulosus*. Está distribuida en todo el mundo provocando grandes problemas sanitarios e importantes pérdidas económicas. Los helmintos tienen pocos sistemas de detoxificación, y las enzimas glutatión transferasas (GSTs) se consideran uno de los principales grupos de enzimas de detoxificación en estos parásitos. Hasta el momento se conocen 8 clases dentro de la superfamilia de GSTs citosólicas: Alfa, Mu, Pi, Sigma, Omega, Zeta, Theta y Phi. Estas enzimas son esenciales en la inactivación de un amplio rango de sustancias, que en los parásitos comprendería la inactivación de compuestos generados por el propio metabolismo así como de los generados por el ataque del sistema inmune del hospedero. Por ello, resulta interesante la identificación y la caracterización bioquímica y molecular de estas moléculas, así como el estudio de su regulación y su expresión.

Las posibles funciones de las GSTs varían entre las diferentes clases. Por ejemplo, se ha descrito que las de clase Sigma tienen la capacidad de sintetizar prostaglandinas (Angeli V. y col., 2001; Herve M. y col., 2003; Sommer A. y col., 2003) y se ha relacionado las prostaglandinas producidas por la GST de *S. mansoni* con la modulación de la respuesta inmune de su hospedero (Angeli V. y col., 2001). También en parásitos, se ha identificado GSTs de clase Omega con actividad tioltransferasa, actividad relacionada con la protección de ciertas proteínas en condiciones de estrés oxidativo a través de la glutationilación-deglutinationilación (Girardini J. y col., 2002; Liebau E. y col., 2008). Además, se han descrito GSTs de clase Mu y Pi en helmintos que participarían ya sea inactivando peróxidos orgánicos como conjugando glutatión a los carbonilos reactivos, moléculas generados en este proceso de lipoperoxidación (Liebau E. y col., 1996a; Liebau E. y col., 1996b; Harispe L. y col., 2010).

Nuestro grupo de trabajo, identificó una GST citosólica de *E. granulosus*, llamada EgGST1, de clase Mu, que conjuga GSH a compuestos generados como consecuencia de la peroxidación de lípidos, además presenta actividad glutatión peroxidasa frente a peróxidos orgánicos (Harispe L. y col., 2010). A

partir de la secuencia de EgGST1, en este trabajo se identificaron en el genoma de *E. granulosus*, varias secuencias relacionadas, todas ellas adyacentes en un único *Scaffold* (*Scaffold-007*). El estar ubicadas en un mismo *Scaffold* nos permite tener información precisa sobre la orientación y el orden de unos con respecto a otros. Esto a su vez, nos permite hipotetizar sobre el mecanismo molecular que da lugar a los eventos de duplicación génica, en este caso el tipo de disposición observada, es decir, cercanos unos a continuación de otros sugiere eventos recurrentes de duplicación en tándem. El estar próximos en el mismo *scaffold*, a su vez, minimiza la posibilidad de que las duplicaciones observadas se deban a artefactos asociados al proceso de ensamblaje del genoma. La posterior anotación manual de los *Hits* obtenidos nos dejó cuatro secuencias ya reportadas (*Hit 1* al *4*), donde se editó y completó el *Hit 3*. Además, identificamos en estos *Hits*, hasta cinco exones de los ocho reportados en la clase Mu en *E. granulosus* (Wilce M.C. y Parker M.W., 1994), en todos los casos comenzando por el exón 4. El no haber encontrado los exones correspondientes al inicio para las cuatro secuencias identificadas podría deberse a un incorrecto ensamblado del *scaffold*. Como se esperaba, la longitud de los exones anotados resultó altamente conservada. Por otro lado, en los intrones ubicamos los sitios dador y aceptor para empalme en posiciones idénticas a las observadas en el ARN transcripto de EgGST1, así como otros genes de clase Mu (Fernandez V., 2001). Los genes para ser funcionales deben tener diversas secuencias señales como ser codón ATG inicial, al menos un marco de lectura abierto (es decir regiones que no contengan codones de terminación en su secuencia en alguno de los 3 marcos de lectura) así como los sitios de *splicing* en caso de contener intrones. La conservación estructural observada en los genes, sugiere que las secuencias identificadas corresponderían a genes que se expresan con un rol funcional relevante.

Es interesante mencionar que los resultados de este trabajo son consistentes con los datos obtenidos por nuestro grupo. En este sentido, se observó conservación de la estructura exón-intrón dentro de los genes de una misma clase, pareciendo ser esta una propiedad general de las GSTs citosólicas (Figura 11, Iriarte A. y col., 2012). Por ejemplo, las de clase Sigma

de *T.solium* (TsGST2), *E. multilocularis* (EmGST2) y *E. granulosus* (EgGST2) presentan 4 exones (Iriarte A. y col., 2012; Arbildi P., 2013). Para las de clase Mu, también se ha descrito este tipo de conservación tanto en mamíferos (Daniel V., 1993) como en cestodos (Iriarte A. y col., 2012).

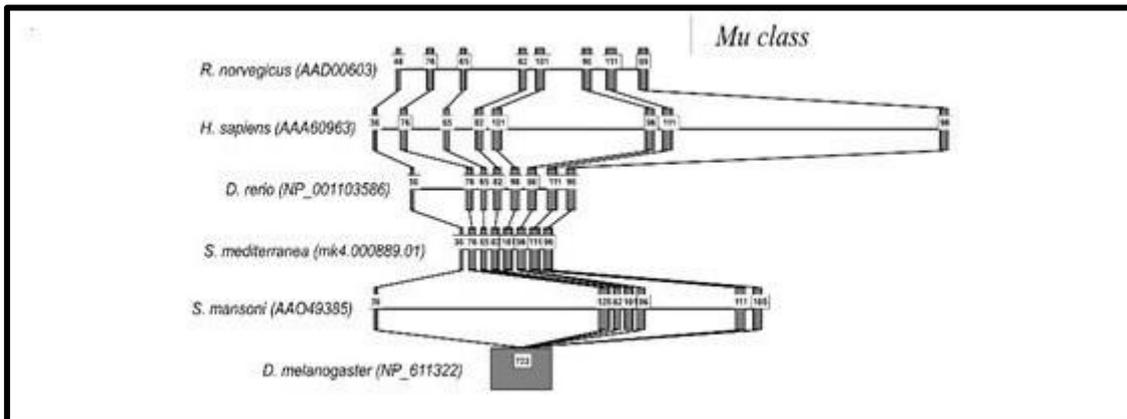


Figura 11. Estructura genómica de los genes de GSTs pertenecientes a la clase Mu (entre especies). Los exones se muestran en recuadros grises, incluye la longitud en pb. Extraído de Iriarte A. y col., 2012

En los alineamientos obtenidos de secuencias aminoacídicas deducidas completas (Figura 6) se presentan los motivos y los residuos característicos de las GSTs de la clase Mu, descritos en otros organismos, respaldando que los genes identificados pertenecen a la clase Mu de las GSTs. Como se esperaba, en el extremo N-terminal se encontró la mayor parte de los residuos identificados como invariantes en las GSTs de esta clase, principalmente los involucrados en el mecanismo catalítico (Wilce M.C. y Parker M.W., 1994). La presencia de una tirosina conservada, Tyr7 (n), considerada esencial para la actividad catalítica de las llamadas Y-GSTs (Atkinson H.J. y Babbitt P.C., 2009), permitió agruparlas inequívocamente en esta categoría. Este residuo es responsable de la producción del tiolato en el GSH (Reinemer P. y col., 1991) necesario para llevar a cabo la conjugación con el sustrato hidrofóbico. Vale la pena mencionar, que estudios llevados a cabo por nuestro grupo de trabajo,

sobre el efecto del pH sobre la actividad de EgGST1, indicaron que EgGST1 presenta un mecanismo catalítico acorde con las Y-GSTs (Arbildi P., 2013). También se observan los residuos aminoacídicos implicados en la formación del sitio de unión a GSH (sitio G), que se encuentran altamente conservados entre las GSTs identificadas en este trabajo, tal como el Trp(n+1), Asn(n+52) y Leu(n+53) (Wilce M.C. y Parker M.W., 1994). Esto sugeriría que las cuatro proteínas homologas duplicadas identificadas tienen una forma de unión al GSH similar.

Las secuencias (Q)SGAIL y (Q)SSAIL halladas (Figura 6) son homólogas a la secuencia consenso conservada identificada en GSTs de mamíferos (Q)[S,T][R,N]AIL en el dominio N-terminal (Wilce M.C. y Parker M.W., 1994). Estas y otras variaciones estructurales que se presentan en las enzimas de los parásitos podrían ser utilizadas para diferenciarlas de las de los mamíferos. La búsqueda de dominios conservados en la base de datos CDD (Conserved Domain Database) relacionó a todas las genes identificaciones con el grupo de las proteínas de tipo Mu. Se observaron aquellos residuos identificados como constituyentes del sitio G, la interface C-terminal y la interface de dimerización. La región de dimerización es una característica propia de cada clase ya que es la que permite la formación de dímeros entre las subunidades de una misma clase (Armstrong R.N., 1997). Esto sugiere que las cuatro proteínas homologas identificadas podrían formar dímeros entre sí, tanto homo como heterodímeros, repercutiendo en su función, específicamente en su rango de sustratos (ver adelante).

A pesar de que la mayoría de los aminoácidos del sitio G están altamente conservados, ciertos residuos fueron sustituidos. El Asp (n+99) que se ubica en una de la α -hélice (α 4) del dominio C-terminal del monómero 2, fue reemplazado por una Asn en el *Hit 3*. Interacciona con el Glu del GSH a través de la formación de puentes de hidrógeno, es decir, el grupo carboxilato de la cadena lateral del Asp interacciona con el grupo α -amino del Glu (Wilce M.C. y Parker M.W., 1994). Es uno de los residuos que son pasibles de interaccionar con el glutatión a través de la formación de puentes de hidrogeno (Wilce M.C. y Parker M.W., 1994), por lo que, cambios en dicho residuo podrían generar un reordenamiento y afectar la interacción de la enzima con el GSH. Asimismo,

otros trabajos han demostrado que el sitio activo del GSH es estabilizado por al menos 15 enlaces de hidrógeno (figura 12) y más de una docena de interacciones electrostáticas (Ji X. y col., 1992; Armstrong R.N., 1997). Colectivamente los cambios en los residuos de aminoácidos que alinean el sitio G causan un cambio local en la red de enlaces de hidrógeno y las interacciones electrostáticas que podrían afectar considerablemente la actividad catalítica de la enzima (Hearne J. y Colman R.F., 2006).

De los residuos descritos que forman parte del sitio de unión al sustrato hidrofóbico (sitio H), la Tyr (n+109) fue reemplazado por una Phe en *Hit* 1. Esta Tyr está contenida dentro de un motivo peptídico (CYSPDF) altamente conservado en tres de los 4 *Hits* de este trabajo (Figura 6). Dicha Tyr se encontraría en el sitio H junto a otros residuos implicados como: Met (n+28), Met (n+98), Met (n+102), Met (n+205), Ala (n+31), Ala (n+105), Ala (n+206) (Figura 9, Wilce M.C. y Parker M.W., 1994). Se ha establecido que la Tyr (n+109) forma puentes de hidrógeno con el principal nitrógeno de la amida de la cadena de Ser (n+201), así como con su hidroxilo de la cadena lateral (Hearne J. y Colman R.F., 2006). Su sustitución podría generar la incapacidad de formar enlaces de hidrogeno con el aminoácido Ser (n+201) modificando la especificidad del sustrato a unir. Por otro lado, el residuo Ala (n+206) no se pudo observar en el *Hit* 4 dado que tiene 201 aminoácidos, siendo más corto que los otros *Hits*, pudiendo verse afectado la unión con los sustratos.

Las diferencias observadas a nivel de secuencias en posiciones clave en la interacción con el sustrato sugiere que las cuatro proteínas homologas identificadas podrían interaccionar con sustratos diferentes o con los mismos sustratos con diferente afinidad. Como forman dímeros, la variabilidad en el sitio de unión al sustrato aumenta exponencialmente, al aumentar el número de combinaciones posibles. Estas diferencias indicarían un proceso de divergencia evolutiva a nivel de secuencias mediada por selección positiva. Por consiguiente diferencias en el sitio de unión al sustrato serían favorecidas por selección, permitiendo al organismo ampliar su capacidad de detoxificación.

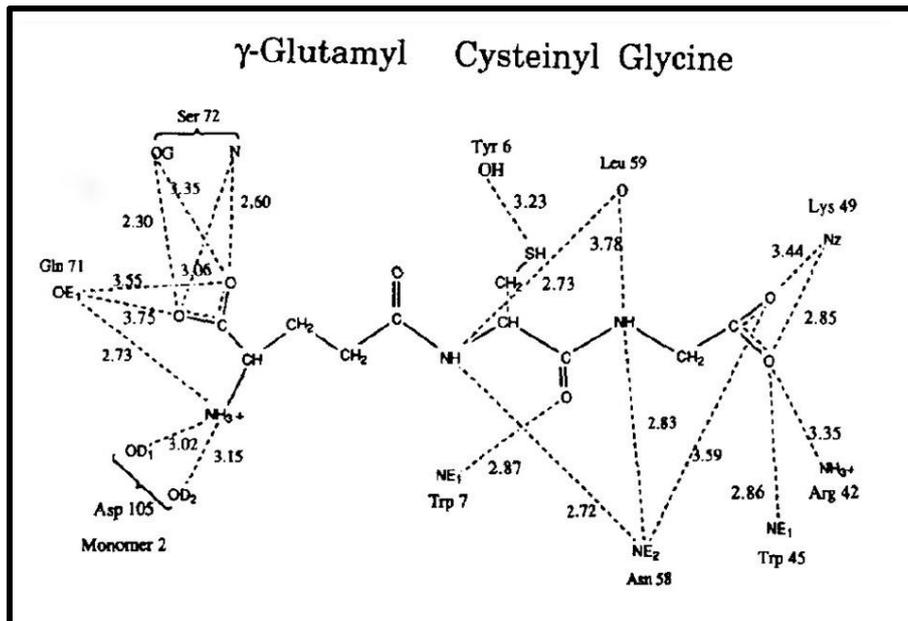


Figura 12 . Esquema interacción residuos-glutatión en el sitio de unión GSH. Extraído de Wilce M.C. y Parker M.W., 1994

El motivo conocido como *mu-loop* (Figura 10) es un elemento estructural característico descrito para las enzimas de clase Mu como ya se mencionó. El motivo está ubicado entre la hoja $\beta 2$ plegada y la hélice $\alpha 2$ del dominio N-terminal. Comprende los aminoácidos entre las posiciones 33 y 42 (Ji X. y col., 1992) y crea un sitio activo más estrecho y menos expuesto al solvente comparado con las GSTs de otras clases (Wilce M.C. y Parker M.W., 1994). Dicho *mu-loop* está presente de forma completa en los *Hits* 1, 2 y 4, pero incompleta en el *Hit* 3 (Figura 6). De hecho, para optimizar el alineamiento con el resto de las secuencias, fue necesario introducir para el *Hit* 3 un quiebre o *gap* entre los aminoácidos 42 y 47 de las secuencias de los otros tres *Hits*, correspondientes al último residuo del *loop* (posición 42) más la región anexa (posiciones 43 al 47) que formarían parte de la hélice $\alpha 2$ del dominio N-terminal (Figura 11). Ante la ausencia de estos residuos, se puede hipotetizar sobre un plegamiento diferente del *loop* sobre el sitio activo, de manera que resulte en una cavidad menos estrecha y más expuesta al solvente (Figura 12). Si bien este cambio no modificaría la estructura global de la enzima, podría afectar su estabilidad y/o la afinidad por los sustratos (Hearne J. y Colman R.F., 2006). La

diferencia observada en el *mu-loop* del *Hit 3* apoya la idea de que podría existir variabilidad en la interacción con los sustratos entre las cuatro proteínas duplicadas en *E. granulosus*, obtenidas en el presente trabajo.

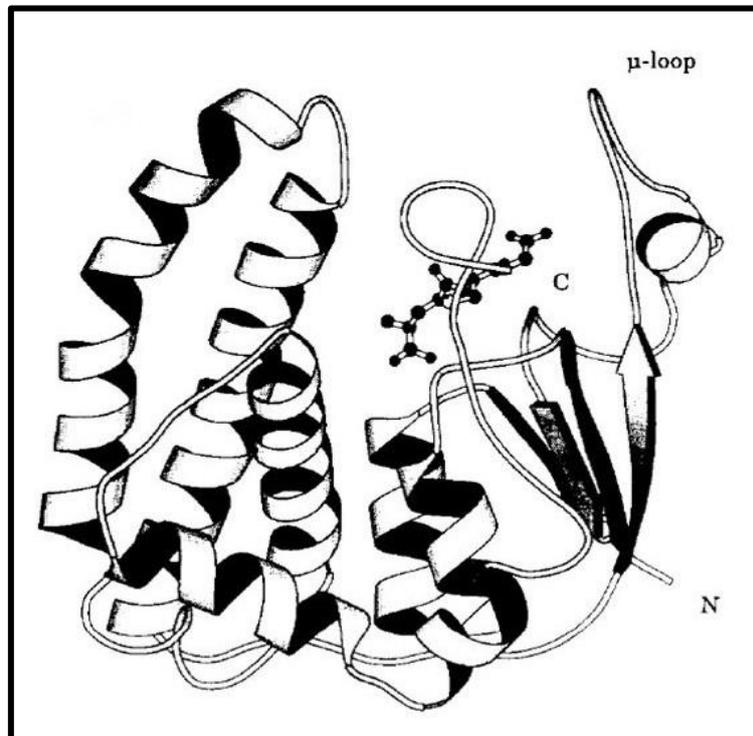


Figura 13. Estructura de la enzima clase Mu en el complejo con el sustrato fisiológico glutatión. Extraído de Wilce M.C. y Parker M.W., 1994.

Para explorar el posible efecto de los cambios en la secuencia observados, en la estructura de la proteína, realizamos una predicción de la estructura secundaria así como espacial de cada *Hit* (figura 8). Si bien habría que confirmarlo con una estrategia diferente, el modelo obtenido para *Hit 3* sugiere que los residuos que corresponderían a motivo *mu-loop* según la Figura 6, estarían formando parte de la hélice α_2 , reduciendo considerablemente el *loop* y por lo tanto, exponiendo el sitio activo al solvente.

Los análisis filogenéticos son concordantes con lo anteriormente discutido, es decir, los genes identificados en este trabajo se ubicaron inequívocamente con las GSTs de clase Mu, formando esta clase un grupo monofilético con un apoyo estadístico de 97%, bien diferenciado de las otras secuencias analizadas (Figura 5). Esto indica una conservación entre las GSTs

citósicas de mamíferos y las correspondientes de los platelmintos en esta clase. Los *Hits* 1, 2, 3 y 4 se encuentran asociados con otras Mu de la familia Taeniidae con diferentes orígenes para cada uno. Esta distribución filogenética observada sugiere que se generaron por eventos de duplicación ancestral. La duplicación un evento frecuente en las familias multigénicas. Es ampliamente aceptado que los procesos de duplicación, de genes y segmentos, promueven la variación y la diversificación génica de diferentes maneras. Los genes duplicados son parte del proceso adaptativo, brindando variación y posiblemente nueva funcionalidad al genoma (Ohno S., 1970). Esto apoya la idea de que las duplicaciones observadas en este trabajo surgen como parte de procesos adaptativos de los distintos linajes. Las enzimas GSTs cumplen roles vitales y de respuesta al ambiente en los parásitos, por lo tanto es de esperar que exista un proceso de adaptación permanente, mediado por selección natural positiva.

Dentro de las GSTs de clase Mu de los cestodos se puede identificar 2 grupos de ortólogos. Resulta interesante identificar 3 grandes linajes, 2 de los cuales se presenta el grupo monofilético sin duplicaciones. En el tercer linaje se observó una expansión específica en *H. microstoma* y en *E. multilocularis* con un apoyo estadístico de 96% y 99% respectivamente (Figura 5). Esta observación sugiere que varios de los genes de GSTs de la clase Mu en platelmintos se generaron por eventos de duplicación reciente. La duplicación es un proceso recurrente en el phylum, que ha ocurrido a lo largo de la evolución del grupo, es decir en varios linajes ancestrales, y que continúa ocurriendo en forma linaje específica.

La divergencia a nivel de secuencia deducida de los genes duplicados de *H. microstoma* y de *E. multilocularis* sugiere que tienen funciones ligeramente distintas (Figura 6). Esta sugerencia es válida para las secuencias identificadas en este trabajo (*Hits* 1 al 4) así como para las otras incluidas en la filogenia. El análisis de conservación de secuencias sugiere un conjunto de residuos conservados así como posiblemente relacionados con el núcleo funcional conservado y otros divergentes. No está claro si estos cambios o sustituciones en los elementos duplicados afectan o no a la dimerización, así como a la preferencia por sustratos o a otras propiedades catalíticas. En un

futuro, esto podría ser analizado una vez expresadas las correspondientes proteínas recombinantes.

Por otro lado, que las secuencias de los genes de los *Hits* se encuentren en los genomas, no implica que sean genes funcionales, es decir que se expresen. Como una aproximación, estudiamos la expresión en *E. granulosus* de los cuatro *Hits* identificados en este trabajo, analizando la transcripción de los genes correspondientes por RT-PCR en protoescólex. Solamente observamos amplicones para los *Hits* 1 y 2, indicando que ambos genes se transcriben. La ausencia de producto de amplificación para los *Hits* 3 y 4 en protoescólex, no excluye que sean expresados en otros estadios del parásito. En este sentido, buscando en la bibliografía, encontramos datos que podrían complementar nuestro trabajo, Zheng H. y colaboradores. Este grupo de investigadores estudió el transcriptoma de *E. granulosus*, partiendo de diferentes materiales parasitarios (gusano adulto, oncoesfera, protoescólex y pared del quiste). Entre las secuencias identificadas por este grupo, encontramos los *Hits* 1, 2 y 4, indicando que son genes que por lo menos se transcriben en alguna de las etapas del ciclo parasitario (Tabla 5). En cambio, el *Hit* 3 no la encontramos en ninguno de los materiales parasitarios analizados por este grupo, lo que es concordante con nuestros resultados. Esto sugiere que, si el parásito no transcribe *Hit* 3, es decir que no lo expresa, lleva a pensar que posiblemente sea un pseudogen, aunque se debieran diseñar otros ensayos concluyentes. De los restantes *Hits*, es interesante destacar que en dicho reporte, tampoco se observó la expresión de *Hit* 4 en protoescólex (Tabla 5), lo que concuerda con los resultados obtenidos en este trabajo. Asimismo, reportan bajos niveles de expresión de *Hit* 4 en los restantes materiales parasitario, sugiriendo una expresión diferencial como se propuso en el párrafo anterior. Por su lado, la expresión de *Hit* 1 sólo fue reportada en la pared del quiste (Tabla 5) aunque en este trabajo la amplificamos de protoescólex. Muy probablemente, esta inconsistencia podría ser debida a diferencias en el material parasitario utilizado, es decir, la procedencia y/o cepa del parásito, el lote y/o las condiciones del material de partida, etc., aunque más ensayos debieran diseñarse. Por último, en concordancia con lo obtenido en esta tesis y con lo observado por nuestro grupo de trabajo previamente (datos no

publicados), la expresión de *Hit 2* se encontró en todos los materiales analizados por dicho grupo de investigadores (Tabla 5).

Tabla 5. Niveles de expresión para los genes anotados por Zheng H. y colaboradores equivalentes a los *Hits* 1,2 y 4. Ver Anexo II.

Hits	Identificador gen	Adulto	Oncoesfera	Protoescólex	Pared quiste hidático
1	EG_07247	Sin expresión	Sin expresión	Sin expresión	Muy baja expresión
2	EG_07276	Alta expresión	Alta expresión	Baja expresión	Alta expresión
4	EG_09218	Muy baja expresión	Muy baja expresión	Sin expresión	Muy baja expresión

6. CONCLUSIONES

En suma, los estudios de conservación estructural del gen, conservación de dominio y los análisis filogenéticos nos llevan a concluir que las cuatro secuencias estudiadas en este trabajo (*Hits* 1 al 4) pertenecen a la familia multigénica de GSTs de la clase Mu y sugieren que se generaron por eventos de duplicación.

En particular, el *Hit 3* si bien su estructura exón-intrón coincide con lo descrito para los otros genes se observaron diferencias a nivel de secuencias en posiciones clave en el estudio de conservación de dominio. Además, la secuencia aminoacídicas que corresponde al *mu-loop* está incompleta y se aprecia una considerable reducción del *loop* en su estructura tridimensional. Por otro lado, no se obtuvo expresión para este *Hit* y no se encontró en ninguno de los materiales parasitarios analizados por Zheng H. y colaboradores. Esto sugiere que posiblemente el *Hit 3* sea un pseudogen.

Los cambios observados en las secuencias para los cuatro *Hits* sugieren que podría existir variabilidad en la interacción con los sustratos y se observó una expresión diferencial de los mismos.

Como complemento de los resultados presentados en este trabajo se podría actualizar la base de datos incorporando nuevos genomas para continuar con estudios de conservación estructural del gen, conservación de

dominio y análisis filogenéticos. Asimismo, sería interesante analizar expresión de estos genes en diferentes materiales parasitarios (gusano adulto, oncoesfera, protoescólex y pared del quiste) por PCR en tiempo real.

Referencias bibliográficas

Angeli V., C. Faveeuw, O. Roye, J. Fontaine, E. Teissier, A. Capron, I. Wolowczuk, M. Capron y F. Trottein (2001). Role of the parasite-derived prostaglandin D2 in the inhibition of epidermal Langerhans cell migration during schistosomiasis infection. *J Exp Med* 193(10): 1135-1147.

Arbidi, P. (2013). Tesis de Maestría en Química: Glutación transferasas en *Echinococcus granulosus*: Identificación, estudios cinéticos y de expresión. Cátedra de Inmunología, Facultad de Química. Montevideo, Universidad de la República.

Arbidi, P., S. La-Rocca y V. Fernández (2011). Glutathione Transferases in helminth parasites. *Research in helminths*. A. Esteves. Kerala, India, Transworld Research Network: 57 - 72.

Armstrong, R.N. (1997). Structure, catalytic mechanism, and evolution of the glutathione transferases. *Chem Res Toxicol* 10, 2-18.

Atkinson, H. J. y P. C. Babbitt (2009). Glutathione transferases are structural and functional outliers in the thioredoxin fold. *Biochemistry* 48: 11108-11116.

Baz, A., A. Hernandez, S. Dematteis, H. Carol y A. Nieto (1995). Idiotypic modulation of the antibody response of mice to *Echinococcus granulosus* antigens. *Immunology* 84: 350-354.

Board, P. G. y D. Menon (2013). Glutathione transferases, regulators of cellular metabolism and physiology. *Biochim Biophys Acta* 1830: 3267-3288.

Daniel, V. (1993), Glutathione S-transferases: gene structure and regulation of expression. *CRC Crit. Revs. Biochem. Mol. Biol.*, 28: 173-207.

Do, C. B., M. S. Mahabhashyam, M. Brudno y S. Batzoglou (2005). ProbCons: Probabilistic consistency-based multiple sequence alignment. *Genome Res* 15: 330-340.

Dzik, J. M. (2006). Molecules released by helminth parasites involved in host colonization. *Acta Biochim Pol* 53: 33-64.

Edgar, R.C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5, 113.

Fernández, C., W.F. Gregory, P. Loke y R.M. Maizels (2002). Full-length-enriched Cdna libraries from *Echinococcus granulosus* contain separate populations of oligo-capped and trans-spliced transcripts and a high level of predicted signal peptide sequences. *Mol Biochem Parasitol.* 122: 171-180.

Fernandez, C. y C. E. Hormaeche (1994). Isolation and biochemical characterisation of a glutathione S-transferase from *Echinococcus granulosus* protoscoleces. *Int J Parasitol* 24: 1063-1066.

Fernandez, V., C. Chalar, C. Martinez, H. Musto, A. Zaha y C. Fernandez (2000). *Echinococcus granulosus*: molecular cloning and phylogenetic analysis of an inducible glutathione S-transferase. *Exp Parasitol* 96: 190-194.

Fernández, V. (2001). Tesis de Doctorado en Química: Estudio del gen de una glutathione transferasa inducible de *Echinococcus granulosus*. Cátedra de Inmunología, Facultad de Química, Montevideo, Universidad de la República.

Frova, C. (2006). Glutathione transferases in the genomics era: new insights and perspectives. *Biomol Eng* 23: 149-169.

Girardini, J., A. Amirante, K. Zemzoumi y E. Serra (2002). Characterization of an omega-class glutathione S-transferase from *Schistosoma mansoni* with glutaredoxin-like dehydroascorbate reductase and thiol transferase activities. *Eur J Biochem* 269(22): 5512-5521.

Hall, T. A. (1999). BioEdit : a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp* 41: 95-98.

Harispe, L., G. Garcia, P. Arbildi, L. Pascovich, C. Chalar, A. Zaha, C. Fernandez y V. Fernandez (2010). Biochemical analysis of a recombinant glutathione transferase from the cestode *Echinococcus granulosus*. *Acta Trop* 114: 31-36.

Hayes, J. D., J. U. Flanagan y I. R. Jowsey (2005). Glutathione transferases. *Annu Rev Pharmacol Toxicol* 45: 51-88.

Hearne J. L. y R. F. Colman (2006), Contribution of the mu loop to the structure and function of rat glutathione transferase M1-1. *Protein Science* 15: 1277-1289.

Herve M., V. Angeli, E. Pinzar, R. Wintjens, C. Faveeuw, S. Narumiya, A. Capron, Y. Urade, M. Capron, G. Riveau y F. Trottein (2003). Pivotal roles of the parasite PGD2 synthase and of the host D prostanoid receptor 1 in schistosome immune evasion. *Eur J Immunol* 33(10): 2764-2772.

Iriarte, A., P. Arbildi, S. La-Rocca, H. Musto y V. Fernandez (2012). Identification of novel glutathione transferases in *Echinococcus granulosus*. An evolutionary perspective. *Acta Trop* 123: 208-216.

Ji, X., P. Zhang, R.N. Armstrong y G.L. Gilliland (1992). The three-dimensional structure of a glutathione S-transferase from the mu gene class. Structural analysis of the binary complex of isoenzyme 3-3 and glutathione at 2.2-Å resolution. *Biochemistry* 31:10169-10184.

La Rocca S., (2014) Tesis de Maestría en Biología: Estudio de la distribución de las glutatión transferasas en *Echinococcus granulosus* y caracterización de una nueva enzima heterodimérica. Cátedra de Inmunología, Facultad de Ciencias, Montevideo, Universidad de la República.

Liebau, E., J. Hoppner, M. Muhlmeister, C. Burmeister, K. Luersen, M. Perbandt, C. Schmetz, D Buttner y N. Brattig (2008). The secretory omega-class glutathione transferase OvgST3 from the human pathogenic parasite *Onchocerca volvulus*. *FEBS J* 275(13): 3438-3453.

Liebau, E., V. Muller, R. Lucius, R. D. Walter y K. Henkle-Duhrsen (1996a). Molecular cloning, expression and characterization of a recombinant glutathione S-transferase from *Echinococcus multilocularis*. *Mol Biochem Parasitol* 77: 49-56.

Liebau, E., G. Wildenburg, P. M. Brophy, R. D. Walter y K. Henkle-Duhrsen (1996b). Biochemical analysis, gene structure and localization of the 24 kDa glutathione S-transferase from *Onchocerca volvulus*. *Mol Biochem Parasitol* 80(1): 27-39.

Morello, A., Y. Repetto y A. Atias (1982). Characterization of glutathione S-transferase activity in *Echinococcus granulosus*. *Comp Biochem Physiol B* 72: 449-452.

Ohno. S (1970), *Evolution by gene duplication*. Springer- Verlag, New York.

Parkinson, J., J. D. Wasmuth, G. Salinas, C. V. Bizarro, C. Sanford, M. Berriman, H. B. Ferreira, A. Zaha, M. L. Blaxter, R. M. Maizels, et al. (2012). Correction: A Transcriptomic Analysis of *Echinococcus granulosus* Larval Stages: Implications for Parasite Biology and Host Adaptation. *PLoS Negl Trop Dis* 6: e1897.

Reinemer, P., H.W. Dirr, R. Ladenstein, J. Schaffer, O. Gallay y R. Huber (1991). The three-dimensional structure of class pi glutathione S-transferase in complex with glutathione sulfonate at 2.3 Å resolution. *EMBO J* 10, 1997-2005.

Repetto, Y., M. E. Letelier y A. Morello (1986). Glutathione S-transferase activity in intact protoscolices of *Echinococcus granulosus*. *Biochemical Society Transactions* 14: 132-133.

Sanguinetti, C. J., E. Dias-Neto y A. J. Simpson (1994). Rapid silver staining and recovery of PCR products separated on polyacrylamide gels. *Biotechniques* 17: 914-921.

Sommer, A., Rickert R., Fischer P., Steinhart H., Walter R. D. y Liebau E. (2003). A dominant role for extracellular glutathione S-transferase from *Onchocerca volvulus* is the production of prostaglandin D₂. *Infect Immun* 71(6): 3603-3606.

Tamura, K., G. Stecher, D. Peterson, A. Filipski y S. Kumar (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol.* 30: 2725-2729.

Thompson, R. C. A. (1995). *Biology and systematics of Echinococcus*. The biology of *Echinococcus* and hydatid disease. Editores R. C. A. Thompson, Lymbery A.J. Wallingford, UK, CAB International.

Torres-Rivera, A. y A. Landa (2008). Glutathione transferases from parasites: a biochemical view. *Acta Trop* 105: 99-112.

Tsai I. J., M. Zarowiecki, N. Holroyd, A. Garciarrubi, A. Sanchez-Flores, K. L. Brooks, A. Tracey, R. J. Bobes, G. Fragoso, E. Sciotto, et al. (2013). The genomes of four tapeworm species reveal adaptations to parasitism. *Nature*. 496: 57-63.

WHO (2001). WHO/OIE Manual on Echinococcosis in Humans and Animals: a Public Health Problem of Global Concern, World Health Organization, World Organisation for Animal Health.

WHO, (2014). Neglected tropical Diseases. A statistical update – Latest data available.

Wilce, M. C. y M. W. Parker (1994). Structure and function of glutathione S-transferases. *Biochim Biophys Acta* 1205: 1-18.

Zheng H., W. Zhang, L. Zhang, Z. Zhang, J. Li, G. Lu, Y. Zhu, Y. Wang, Y. Huang, J. Liu, H. Kang, J. Chen, L. Wang, A. Chen, S. Yu, Z. Gao, L. Jin, W. Gu, Z. Wang, L. Zhao, B. Shi, H. Wen, R. Lin, M. K. Jones, B. Brejova, T. Vinar, G. Zhao, D. P. McManus, Z. Chen, Y. Zhou y S. Wang (2013). The genome of the hydatid tapeworm *Echinococcus granulosus*. *Nature Genetics* 45(10): 1168-1175.

Anexo I

Nombre	Número de acceso	Organismo	Filum	Base de datos
clase mapeg				
DRE_ENSDARG00000024143	ENSDARG00000024143	Danio rerio	chordata	Bgee
DRE_ENSDARG00000033364	ENSDARG00000033364	Danio rerio	chordata	Bgee
DRE_ENSDARG00000071345	ENSDARG00000071345	Danio rerio	chordata	Bgee
EgG_000886800.1..pep	A0A068W9V4	Echinococcus granulosus	plathyhelminthes	Uniprot
EmW_000886800_MAPEG_microsomal	A0A068Y9K2	Echinococcus multilocularis	plathyhelminthes	Uniprot
HmN_000391200.1..pep	A0A068XP16	Hymenolepis microstoma	plathyhelminthes	Uniprot
HmN_000391300.1..pep	A0A068XPN3	Hymenolepis microstoma	plathyhelminthes	Uniprot
HSA_CCDS1249.1.Hs37.3.chr1	O14880	Homo sapiens	chordata	Uniprot
HSA_CCDS34316.1.Hs37.3.chr5	Q16873	Homo sapiens	chordata	Uniprot
HSA_CCDS3749.1.Hs37.3.chr4	Q99735	Homo sapiens	chordata	Uniprot
MUS_CCDS15457.1.Mm37.2.chr1	Q9CPU4	Mus musculus	chordata	Uniprot
MUS_CCDS24631.1.Mm37.2.chr11	Q60860	Mus musculus	chordata	Uniprot
MUS_CCDS38428.1.Mm37.2.chr3	NP_778160	Mus musculus	chordata	GeneBank
Smp_055210.1	G4VH65	Schistosoma mansoni	plathyhelminthes	Uniprot
TsM_000126000	TsM_000126000	Taenia solium	plathyhelminthes	Wormbase
clase mu				
DRE_ENSDARG00000029473	ENSDARG00000029473	Danio rerio	chordata	Bgee
DRE_ENSDARG00000042533	ENSDARG00000042533	Danio rerio	chordata	Bgee
EgG_000535200.1..pep	A0A068WLI9	Echinococcus granulosus	plathyhelminthes	Uniprot
EgG_000535300.1..pep	A0A068WTH4	Echinococcus granulosus	plathyhelminthes	Uniprot
EgG_000537600.1..pep	A0A068WR90	Echinococcus granulosus	plathyhelminthes	Uniprot
EgG_000538000.1..pep	A0A068WLL3	Echinococcus granulosus	plathyhelminthes	Uniprot
EgG_000538900.1..pep	U6JCG1	Echinococcus granulosus	Plathyhelminthes	Uniprot
EmW_000005400_mu10	A0A087VWB2	Echinococcus multilocularis	plathyhelminthes	Uniprot
EmW_000535200_mu1	A0A068Y038	Echinococcus multilocularis	plathyhelminthes	Uniprot
EmW_000535300_mu2	A0A068Y6Y2	Echinococcus multilocularis	plathyhelminthes	Uniprot
EmW_000537600_mu3	A0A068Y4J2	Echinococcus multilocularis	plathyhelminthes	Uniprot
EmW_000538000_mu4	A0A068Y062	Echinococcus multilocularis	plathyhelminthes	Uniprot
EmW_000538300_mu5	Q27653	Echinococcus multilocularis	plathyhelminthes	Uniprot
EmW_000538500_mu6	A0A068Y0T5	Echinococcus multilocularis	plathyhelminthes	Uniprot
EmW_000538600_mu7	A0A068Y4J5	Echinococcus multilocularis	plathyhelminthes	Uniprot

EmW_000538700_mu8	A0A068Y066	Echinococcus multilocularis	plathyhelminthes	Uniprot
EmW_000538900_mu9	A0A068Y7D9	Echinococcus multilocularis	plathyhelminthes	Uniprot
Hit 1	A0A068WTH4	Echinococcus granulosus	plathyhelminthes	Uniprot
Hit 2	O16058	Echinococcus granulosus	plathyhelminthes	Uniprot
Hit 3	A0A068WTJ7	Echinococcus granulosus	plathyhelminthes	Uniprot
Hit 4	U6JCG1	Echinococcus granulosus	plathyhelminthes	Uniprot
HmN_000365900.1..pep	A0A068XNR1	Hymenolepis microstoma	plathyhelminthes	Uniprot
HmN_000366000.1..pep	A0A068XH77	Hymenolepis microstoma	plathyhelminthes	Uniprot
HmN_000366200.1..pep	A0A068XGM8	Hymenolepis microstoma	plathyhelminthes	Uniprot
HmN_000366400.1..pep	A0A068XNR4	Hymenolepis microstoma	plathyhelminthes	Uniprot
HmN_000366500.1..pep	A0A068XH82	Hymenolepis microstoma	plathyhelminthes	Uniprot
HmN_000876800.1..pep	U6HCF7	Hymenolepis microstoma	plathyhelminthes	Uniprot
HmN_000876900.1..pep	U6H733	Hymenolepis microstoma	plathyhelminthes	Uniprot
HmN_001013500.1..pep	U6HAV6	Hymenolepis microstoma	plathyhelminthes	Uniprot
HSA_CCDS806.1.Hs37.3.chr1	Q03013	Homo sapiens	chordata	Uniprot
HSA_CCDS807.1.Hs37.3.chr1	Q03013	Homo sapiens	chordata	Uniprot
HSA_CCDS810.1.Hs37.3.chr1	P09488	Homo sapiens	chordata	Uniprot
MUS_CCDS17744.1.Mm37.2.chr3	P19639	Mus musculus	chordata	Uniprot
MUS_CCDS17746.1.Mm37.2.chr3	P15626	Mus musculus	chordata	Uniprot
MUS_CCDS17747.1.Mm37.2.chr3	P10649	Mus musculus	chordata	Uniprot
MUS_CCDS17748.1.Mm37.2.chr3	EDL01940	Mus musculus	chordata	GeneBank
MUS_CCDS38596.1.Mm37.2.chr3	O35660	Mus musculus	chordata	Uniprot
Smp_102070.1	P35661	Schistosoma mansoni	plathyhelminthes	Uniprot
TsM_000157300	TsM_000157300	Taenia solium	plathyhelminthes	Wormbase
TsM_000610200	TsM_000610200	Taenia solium	plathyhelminthes	Wormbase
TsM_000610300	TsM_000610300	Taenia solium	plathyhelminthes	Wormbase
TsM_000639300	TsM_000639300	Taenia solium	plathyhelminthes	Wormbase
TsM_000741700	TsM_000741700	Taenia solium	plathyhelminthes	Wormbase
TsM_001037400	TsM_001037400	Taenia solium	plathyhelminthes	Wormbase
TsM_001229600	TsM_001229600	Taenia solium	plathyhelminthes	Wormbase
clase pi				
HmN_000614200.1..pep	A0A068X6X8	Hymenolepis microstoma	plathyhelminthes	Uniprot
DRE_ENSDARG00000057338	ENSDARG00000057338	Danio rerio	chordata	Bgee
MUS_CCDS29411.1.Mm37.2.chr19	P19157	Mus musculus	chordata	Uniprot
MUS_CCDS29412.1.Mm37.2.chr19	P46425	Mus musculus	chordata	Uniprot
clase sigma				
CEL_WBGene00001755	P91253	Caenorhabditis elegans	plathyhelminthes	Uniprot

CEL_WBGene00001772	Q93694	Caenorhabditis elegans	plathyhelminthes	Uniprot
EgG_000459050.1..pep	A0A068WGU0	Echinococcus granulosus	plathyhelminthes	Uniprot
EgG_000685900.1..pep	U6JFS3	Echinococcus granulosus	plathyhelminthes	Uniprot
EmW_000459050_mu/28_PGH2D_isomerase_SIGMA2	A0A068Y500	Echinococcus multilocularis	plathyhelminthes	Uniprot
EmW_000685900_mu/28_PGH2D_isomerase_SIGMA1	A0A068Y7Y4 A0A068XEH4	Echinococcus multilocularis	plathyhelminthes	Uniprot
HmN_000811800.1..pep		Hymenolepis microstoma	plathyhelminthes	Uniprot
Smp_054160.1	P09792	Schistosoma mansoni	plathyhelminthes	Uniprot
TsM_000670100	TsM_000670100	Taenia solium	plathyhelminthes	Wormbase
TsM_000815300	TsM_000815300	Taenia solium	plathyhelminthes	Wormbase

Tabla 1. Secuencias utilizadas en la construcción del árbol filogenético de las GSTs presentado en la Figura 5. Se indica el número de acceso de cada proteína, la base de datos, el organismo del que proviene la misma y el filum.

Anexo II

Gene ID	Length	Adult	Onc	PSC	Cyst	Annotation	GOs
		Reads	Reads	Reads	Reads		
EG_07247 - Hit 1	1701	0	0	0	2	Glutathione S-transferase Mu 1	F:glutathione transferase activity; F:binding; P: xenobiotic catabolic process; P:nitrobenzene metabolic process; C:cytosol; P:glutathione metabolic process; P:glutathione conjugation reaction
EG_07276 - Hit 2	660	211	181	20	182	Glutathione S-transferase	C:cell part; F:glutathione transferase activity; F:protein binding; P:glutathione metabolic process; P:glutathione conjugation reaction
EG_09218 - Hit 4	606	8	4	0	1	Glutathione S-transferase Mu 1	C:cell part; F:glutathione transferase activity; F:protein binding; P:glutathione metabolic process; P:glutathione conjugation reaction

Tabla 6. Fracción de la tabla de los perfiles de transcripción para los genes anotados por Zheng H. y colaboradores, partiendo de diferentes materiales parasitarios: gusano adulto, oncoesfera, protoescólex y pared del quiste. Se ubican los *Hits* 1, 2, y 4 entre las secuencias identificadas por este grupo, no se ubica el *Hit* 3. En color se indican las columnas con los valores tomados como referencia para la elaboración de la Tabla 5; 0 = sin expresión, 1-20 Muy Baja expresión, 21-90 Baja expresión, > 91 Alta expresión.