

2022

Análisis transcriptómico y participación de la vía canónica
de Wnt en la regeneración de *Lumbriculus variegatus*

Aldo Arturo Téllez García



Universidad Autónoma de Querétaro
Facultad de Ciencias Naturales

Análisis transcriptómico y participación de la vía
canónica de Wnt en la regeneración
de *Lumbriculus variegatus*

Tesis

Que como parte de los requisitos para
obtener el Grado de

Doctor en Ciencias Biológicas

Presenta

Aldo Arturo Téllez García

Dirigido por:

Dr. Fausto Arellano Carbajal

Querétaro, Qro. a junio de 2022



Universidad Autónoma de Querétaro
Facultad de Ciencias Naturales
Doctorado en Ciencias Biológicas

Análisis transcriptómico y participación de la vía
canónica de Wnt en la regeneración
de *Lumbriculus variegatus*

Tesis

Que como parte de los requisitos para obtener el Grado de
Doctor en Ciencias Biológicas

Presenta
Aldo Arturo Téllez García

Dirigido por:
Fausto Arellano Carbajal

Dr. Fausto Arellano Carbajal
Presidente

Dra. Angelina Rodríguez Torres
Secretario

Dr. Roberto Carlos Álvarez Martínez
Vocal

Dr. Edgardo Ulises Esquivel Naranajo
Suplente

Dr. Alejandro Sánchez González
Suplente

Centro Universitario, Querétaro, Qro.
Fecha de aprobación por el Consejo Universitario (junio 2022)
México

Y miré, y he aquí nervios sobre ellos,
y la carne subió,
y la piel cubrió por encima de ellos...

Ezequiel, 37:8

Agradecimientos institucionales

Agradezco a la Universidad Autónoma de Querétaro, a la Facultad de Ciencias Naturales y al Posgrado en Ciencias Biológicas.

Agradezco al Consejo Nacional de Ciencia y Tecnología (CONACYT) por otorgarme la beca de doctorado; No. de CVU / Clave de solicitud: 631374 / 2018-000012-01NACF-04068.

Agradezco a los miembros del Comité Tutor:

- Dr. Alejandro Sánchez Alvarado
- Dra. Angelina Rodríguez Torres
- Dr. Edgardo Ulises Esquivel Naranjo
- Dr. Fausto Arellano Carbajal
- Dr. Roberto Carlos Álvarez Martínez

Agradecimientos personales

Quiero comenzar agradeciendo enormemente a Fausto Arellano Carbajal por darme la oportunidad de trabajar con él en este proyecto. Haber contado con un asesor siempre dispuesto a compartir y a escuchar ideas y preguntas fue, sin duda, una gran motivación en estos cuatro años. Estoy muy agradecido por todo lo que aprendí de él; pero más por contar con alguien que primero se preocupó por la persona y después por el trabajo. Su generosidad es valiosísima, su amistad es invaluable.

También agradezco a los miembros del Comité Tutor por los comentarios que le hicieron a este trabajo, por los consejos que me dieron y por el interés que siempre mostraron para que este proyecto saliera adelante.

Agradezco enormemente a los compañeros e investigadores del Laboratorio de Genética Molecular y Ecología Evolutiva por la convivencia y los buenos ratos; en especial a Juana María López Martínez y Luis Fernando Mendieta Yañez, con quienes compartí el gusto de estudiar a *Lumbriculus variegatus*.

Asimismo agradezco a Luis Alberto Soto García, Nami del Rosario Morales Durán, Carlos Andrés Domínguez Sánchez y Cecilia Barragán Vargas por el apoyo que me brindaron fungiendo como técnicos de laboratorio.

No puedo dejar de agradecer a Alexandra Elbakyan por creer que la ciencia debe ser para todos. Este trabajo se debe, en gran medida, al suyo.

Estaré siempre agradecido con mis padres y hermanos por todo su cariño y apoyo.

Finalmente agradezco a Ana Laura Rodríguez Martínez por siempre estar, por compartir conmigo esta aventura y por ser la sonrisa al final del día.

Índice

Resumen	4
Abstract	5
1. Introducción	6
2. Capítulo 1. Genes y procesos biológicos que intervienen en la regeneración de <i>Lu. variegatus</i> .	10
2.1. Antecedentes	10
2.2. Hipótesis	13
2.3. Objetivos	13
2.4. Métodos	14
2.4.1. Sujetos experimentales	14
2.4.2. Establecimiento de grupos experimentales, extracción de ARN y secuenciación	14
2.4.3. Ensamblaje y anotación del transcriptoma	15
2.4.4. Análisis de expresión diferencial	17
2.4.5. Análisis de enriquecimiento de ontologías genéticas	17
2.5. Resultados	18
2.5.1. Ensamblaje y anotación del transcriptoma	18
2.5.2. Expresión diferencial durante la regeneración	19
2.5.3. Enriquecimiento de ontologías genéticas durante la regeneración	20
2.6. Discusión	22
3. Capítulo 2. Participación de la vía canónica de Wnt en la regeneración de <i>Lu. variegatus</i> .	29
3.1. Antecedentes	29
3.2. Hipótesis	34
3.3. Objetivos	34
3.4. Métodos	35
3.4.1. Patrón de expresión de β -catenin mediante hibridaciones <i>in situ</i>	35
3.4.1.1. Diseño experimental	35
3.4.1.2. Síntesis de las sondas	36
3.4.1.3. Hibridaciones	38
3.4.2. Inhibición farmacológica de la vía canónica de Wnt durante la regeneración	39
3.4.3. Inhibición de la vía canónica de Wnt por medio de ARNi durante la regeneración	40
3.4.4. Sobreactivación farmacológica de la vía canónica de Wnt durante la regeneración	42
3.4.5. Formación del SNC bajo la inhibición y la sobreactivación de la vía canónica de Wnt durante la regeneración	43
3.5. Resultados	44
3.5.1. Patrón de expresión de β -catenin durante la regeneración	44
3.5.2. Efecto de la inhibición farmacológica de la vía canónica de Wnt en el tamaño del blastema durante la regeneración	47
3.5.3. Efecto de la inhibición de la vía canónica de Wnt por medio de ARNi en el tamaño del blastema durante la regeneración	49
3.5.4. Efecto de la sobreactivación farmacológica de la vía canónica de Wnt en el tamaño del blastema durante la regeneración	51
3.5.5. Efecto de la inhibición y sobreactivación farmacológica de la vía canónica de Wnt en la formación del SNC durante la regeneración	52
3.6. Discusión	54
4. Discusión y conclusiones generales	60
5. Referencias	64
Anexo 1. Artículo publicado a partir de los resultados del Capítulo 1.	79
Anexo 2. Nomenclatura de genes, transcritos y proteínas.	91
Anexo 3. Tabla de transcritos diferencialmente expresados con algunos estadísticos del análisis y anotaciones.	92

Índice de Tablas

Tabla 1. Genes candidatos de la regeneración.	11
Tabla 2. Oligonucleótidos utilizados.	36

Índice de Figuras

Figura 1. Transcritos diferencialmente expresados durante la regeneración.	19
Figura 2. Redes de ontologías genéticas enriquecidas.	21
Figura 3. Imagen representativa de las hibridaciones in situ de los grupos control (sonda sentido).	45
Figura 4. Imagen representativa de las hibridaciones in situ de las 0 a las 336 horas post amputación (hpa), cada 12 h, en la regeneración anterior.	46
Figura 5. Imagen representativa de las hibridaciones in situ de las 0 a las 336 horas post amputación (hpa), cada 12 h, en la regeneración posterior.	47
Figura 6. Áreas de los blastemas tratados con IWR-1 de las 48 a las 72 hpa.	48
Figura 7. Áreas de los blastemas tratados con IWR-1 de las 48 a las 96 hpa.	49
Figura 8. Áreas de los blastemas tratados con ARNdc de las 48 a las 72 hpa.	50
Figura 9. Áreas de los blastemas tratados con ARNdc de las 24 a las 72 hpa.	51
Figura 10. Áreas de los blastemas tratados con dos concentraciones de LiCl de las 48 a las 72 hpa.	52
Figura 11. Regeneración del CNV bajo la inhibición y sobreactivación de la vía canónica de Wnt.	53

Resumen

Lumbriculus variegatus es un anélido con una gran capacidad regenerativa. Luego de ser amputados, los fragmentos corporales de este organismo son capaces de regenerar la cabeza y la cola en tan solo dos semanas. Si bien algunos de los aspectos celulares que tienen lugar durante la regeneración han sido estudiados previamente, los mecanismos moleculares que regulan este proceso han recibido menor atención. Con la finalidad de identificar qué genes, procesos celulares y funciones moleculares se encuentran activos durante la regeneración; en este trabajo se llevaron a cabo un análisis de expresión diferencial y un análisis de enriquecimiento de ontologías genéticas de los primeros tres días de regeneración. También, se exploró la participación de la vía canónica de Wnt en la formación del blastema y del sistema nervioso central. Entre los principales hallazgos, se identificó la expresión diferencial de 136 transcritos, algunos de los cuales están asociados con la regeneración y la proliferación celular en otros organismos; así como la sobrerregulación del metabolismo energético, de ARN y de aminoácidos. Por otro lado, se determinó el patrón de expresión espaciotemporal de *β-catenin*, un miembro de la vía canónica de Wnt; y por medio de experimentos de inhibición y sobreactivación farmacológica, se demostró que la adecuada regulación de la actividad de esta vía es necesaria para la formación del blastema y la diferenciación del sistema nervioso central. Finalmente, en esta tesis se generó un transcriptoma anotado de la regeneración que se encuentra publicado y disponible para futuros trabajos.

Palabras clave: *Lumbriculus variegatus*, regeneración, transcriptoma, expresión diferencial, vía canónica de Wnt.

Abstract

Lumbriculus variegatus is an annelid with a great regenerative capacity. After being amputated, the body fragments of this organism are able to regenerate the head and tail in only two weeks. While some cellular aspects that take place during regeneration have been previously studied, the molecular mechanisms that regulate this process have received less attention. In order to identify which genes, cellular processes and molecular functions are active during regeneration, differential expression analysis and gene ontology enrichment analysis of the first three days of regeneration were carried out. The participation of the canonical Wnt pathway in blastema and central nervous system formation was also explored. Among the main findings, 136 transcripts were identified as differentially expressed, some of which are associated with regeneration and cell proliferation in other organisms. It was also found that energy, RNA and amino acid metabolism were upregulated. On the other hand, the spatiotemporal expression pattern of *β-catenin*, a member of the canonical Wnt pathway, was determined; and by means of pharmacological inhibition and overactivation experiments, it was demonstrated that the proper regulation of the activity of this pathway is necessary for blastema formation and central nervous system differentiation. Finally, an annotated regeneration transcriptome was generated in this thesis, which is published and available for future works.

Keywords: *Lumbriculus variegatus*, regeneration, transcriptome, differential expression, canonical Wnt pathway.

1. Introducción

La regeneración, o “la restauración de las partes perdidas del cuerpo”, es una respuesta bastante común entre los animales luego de sufrir una lesión traumática, ya sea por la amputación de alguna parte corporal o por autotomía (Bely & Nyberg, 2010). En algunos grupos de animales, la capacidad de regenerar está limitada a algunos tejidos u órganos específicos; en cambio, otros grupos son capaces de regenerar estructuras corporales completas o, incluso, todo un organismo a partir de un pequeño fragmento corporal (Bely & Nyberg, 2010). Uno de los grupos de animales que exhiben una gran capacidad regenerativa es el de los anélidos, entre los cuales se ubican varias especies capaces de regenerar, en mayor o menor medida, el eje corporal principal (Zattara & Bely, 2016). Algunas especies son incapaces de regenerar, otras pueden regenerar solamente la sección posterior del eje corporal; pero también hay especies capaces de regenerar tanto la sección anterior como la posterior (Özpolat & Bely, 2016; Zattara & Bely, 2016). Asimismo, la regeneración en los anélidos está mediada, principalmente, por un proceso conocido como epimorfosis, el cual involucra la formación de un blastema (*i.e.*, un cúmulo de células multipotentes en el sitio de la lesión) a partir del cual se diferencian los tejidos perdidos (Bely, 2014; Özpolat & Bely, 2016). En general, la epimorfosis en anélidos involucra una serie de pasos bien establecidos: oclusión de la herida; proliferación celular y formación del blastema; migración de células provenientes del tejido no amputado; y diferenciación celular (Bely, 2014). Sin embargo, en algunas especies la epimorfosis está acoplada con otro proceso conocido como morfálaxis (Özpolat & Bely, 2016). En este proceso los tejidos no perdidos son remodelados para adoptar una nueva identidad en el plano corporal o para mantener la proporción del cuerpo (Özpolat & Bely, 2016). De este modo, la regeneración en los anélidos depende, en gran medida, de una oclusión efectiva de la herida y de la formación de un blastema capaz de diferenciarse en nuevos tejidos (Bely, 2014).

Lumbriculus variegatus (Clitellata: Lumbriculida) es una especie de lombriz dulceacuícola de origen holártico y distribución cosmopolita que habita zonas someras de estanques, lagos y arroyos (Cook, 1969; Gustafsson *et al.*, 2009; Timm, 1999). Esta especie hermafrodita es capaz de reproducirse sexualmente, pero su principal modo de reproducción es por medio de fisión, la cual implica un proceso de regeneración anterior y posterior en los fragmentos posterior y anterior, respectivamente (Zoran & Martinez, 2009). De manera similar, esta especie es capaz de regenerar las secciones anterior y posterior del cuerpo cuando sufre autotomía en respuesta a un estímulo mecánico, o bien, cuando el daño mecánico es lo suficientemente severo como para generar una lesión que ampute al organismo y genere dos o más fragmentos (Lesiuk & Drewes, 1999). Esta gran capacidad regenerativa, aunada a la facilidad con la que se puede mantener en el laboratorio, hacen de *Lu. variegatus* un modelo atractivo para estudiar los mecanismos moleculares de la regeneración.

En el caso de una lesión traumática, la regeneración comienza inmediatamente con la oclusión de la herida por medio de una rápida contracción de los músculos circulares de la pared corporal y continúa con la desdiferenciación de las células localizadas cerca de la zona de la lesión (Bely, 2014; Lesiuk & Drewes, 1999). Estas células comienzan a proliferar en las primeras 24 horas post amputación (hpa) y son la principal fuente celular para la formación del blastema (Tweeten & Anderson, 2008); aunque también se ha observado la migración de células que invaden el blastema a las 72 hpa desde los tejidos no amputados (Randolph, 1892; Tweeten & Anderson, 2008; Tweeten & Rainer, 2012). La formación del blastema está seguida por su crecimiento y diferenciación en nuevos tejidos, los cuales comienzan a ser evidentes en el quinto día de regeneración, cuando pueden observarse los nuevos segmentos de la cabeza y la cola (Martinez-Acosta & Zoran, 2015; Tweeten & Reiner, 2012). El crecimiento y la diferenciación continúan hasta aproximadamente dos semanas después de la amputación, cuando se considera que el proceso regenerativo concluye (Martinez-Acosta & Zoran, 2015; Tweeten & Reiner, 2012).

Además de la regeneración epimórfica, *Lu. variegatus* sufre cambios morfolácticos en las zonas remanentes del cordón nervioso ventral (CNV) a partir de las primeras 48 hpa (Drewes & Fourtner, 1990). Estos cambios permiten que las fibras nerviosas adopten una nueva identidad en cuanto a su nueva posición en el plano corporal (Drewes & Fourtner, 1990). Por ejemplo, luego de que una lombriz fue amputada por la mitad, las fibras nerviosas que inicialmente estaban ubicadas en la región medial del cuerpo deben sufrir cambios morfológicos y fisiológicos que les permitan funcionar como fibras anteriores y posteriores en los fragmentos posterior y anterior, respectivamente (Drewes & Fourtner, 1990).

Si bien varios aspectos celulares de la regeneración en *Lu. variegatus* y otros anélidos han sido estudiados ampliamente, los mecanismos moleculares que regulan este proceso han recibido menor atención (Bely, 2014; Özpolat & Bely, 2016). El estudio de estos mecanismos es fundamental para después llevar a cabo análisis comparativos que permitan comprender por qué algunos organismos son capaces de regenerar mientras que otros no. Asimismo, el entendimiento de estos mecanismos podría ayudar a identificar cuáles son los factores que determinan en qué medida pueden regenerar los organismos que poseen esta capacidad.

El objetivo central de esta tesis es determinar cuáles son algunos de los mecanismos moleculares que intervienen en la regeneración de *Lu. variegatus*, haciendo énfasis en aquellos que participan en la formación del blastema y del sistema nervioso central (SNC). Esta tesis se encuentra dividida en dos capítulos. En el primero, se describen los hallazgos de un análisis de expresión diferencial en el que se comparó el perfil de expresión génica de organismos en estado regenerativo (*i.e.*, que sufrieron una lesión traumática y se encuentran en algún punto del proceso regenerativo) con el de organismos en estado no regenerativo. Además, se muestran los resultados de un análisis de ontologías genéticas con el cual se determinaron los procesos biológicos y funciones moleculares más relevantes durante la formación del blastema. Asimismo, se presenta un

transcriptoma de la regeneración que fue anotado con información de homologías de secuencias, ontologías genéticas y el potencial codificante de los transcritos. Por otro lado, en el segundo capítulo se describen los resultados de una serie de experimentos llevados a cabo para determinar el papel que juega la vía canónica de Wnt en la formación del blastema y la diferenciación del SNC. La discusión de ambos capítulos tiene un enfoque comparativo que busca encontrar similitudes y diferencias entre el proceso de regeneración en *Lu. variegatus* y otras especies de anélidos.

2. Capítulo 1. Genes y procesos biológicos que intervienen en la regeneración de *Lu. variegatus*.

Los métodos y resultados de este capítulo se encuentran publicados (Tellez-Garcia *et al.*, 2021); pero aquí se describen nuevamente junto con los antecedentes, hipótesis, objetivos y la discusión de los resultados. También se incluye el artículo completo en el Anexo 1. Por otro lado, la nomenclatura de genes, transcritos y proteínas utilizada en este capítulo y el Capítulo 2 se encuentra descrita en el Anexo 2.

2.1. Antecedentes

La expresión de genes asociados a la regeneración en anélidos ha sido estudiada recientemente por medio de dos aproximaciones principales. Estas aproximaciones han permitido, entre otras cosas, detectar algunos genes cuya participación es común en varias especies y otros que aparentemente se expresan solo en una o en unas cuantas. La detección de estas coincidencias y diferencias es un primer paso para determinar qué mecanismos moleculares son comunes en la regeneración de todo el grupo, así como para comprender las bases genéticas que subyacen las particularidades de la regeneración en las distintas especies (Özpolat & Bely, 2016).

Por un lado, en algunos estudios se ha analizado de manera particular la expresión de genes candidatos de la regeneración (*i.e.*, genes que previamente se han encontrado asociados a este proceso en otros grupos de animales). Principalmente mediante hibridaciones *in situ* de ácidos ribonucleicos (ARN), se ha determinado el patrón de expresión de genes candidatos involucrados en la proliferación, migración y diferenciación de las células del blastema en diferentes especies (Tabla 1) (Özpolat & Bely, 2016).

Tabla 1. Genes candidatos de la regeneración. La expresión de estos genes ha sido determinada por medio de hibridaciones *in situ* de ARN en diferentes especies de anélidos.

Gen	Especies	Sitio de expresión	Referencias
<i>β-catenin</i> (armadillo, <i>Ctnnb1</i> ; β-catenina)	<i>Pristina leidy</i>	Blastema anterior, blastema posterior y zona de fisión.	Nyberg <i>et al.</i> , 2012
<i>frizzled</i> (Frizzled)	<i>Pr. leidy</i>	Blastema anterior, blastema posterior y zona de fisión.	Nyberg <i>et al.</i> , 2012
<i>Gs</i> (<i>Glul</i> ; Glutamina sintetasa)	<i>Enchytraeus japonensis</i>	Blastema anterior y blastema posterior.	Niva <i>et al.</i> , 2008
Varios genes <i>Hox</i>	<i>Alitta virens</i>	Blastema posterior.	Novikova <i>et al.</i> , 2013
	<i>Capitella teleta</i>	Blastema posterior.	de Jong & Seaver, 2016
	<i>Platynereis dumerilii</i>	Blastema posterior.	Pfeifer <i>et al.</i> , 2012
<i>nanos</i> (Nanos)	<i>Paranais litoralis</i>	Zona de fisión.	Bely & Sikes, 2010
	<i>Pr. leidy</i>	Blastema anterior, blastema posterior y zona de fisión.	Bely & Sikes, 2010; Özpolat & Bely, 2015
<i>piwi</i> (Piwi)	<i>Al. virens</i>	Blastema posterior.	Kozin & Kostyuchenko, 2015
	<i>En. japonensis</i>	Blastema anterior.	Tadokoro <i>et al.</i> , 2006
	<i>Pr. leidy</i>	Blastema anterior, blastema posterior y zona de fisión.	Özpolat & Bely, 2015
<i>PL10</i> (<i>DDX3</i> ; Helicasa de ARN dependiente de ATP PI10)	<i>Al. virens</i>	Blastema posterior.	Kozin & Kostyuchenko, 2015
<i>twist</i> (Twist)	<i>Pl. dumerilii</i>	Blastema posterior.	Pfeifer <i>et al.</i> , 2013
<i>vasa</i> (<i>DDX4</i> ; Helicasa de ARN dependiente de ATP vasa)	<i>Al. virens</i>	Blastema posterior.	Kozin & Kostyuchenko, 2015
	<i>Pr. leidy</i>	Blastema anterior, blastema posterior y zona de fisión.	Özpolat & Bely, 2015

Por otro lado, mediante análisis transcriptómicos, se ha logrado identificar una gran cantidad de genes expresados durante la regeneración. Dado que esta aproximación es más amplia —por cuanto no se basa solo en la búsqueda de genes seleccionados *a priori* sino en la identificación de la mayor cantidad posible de genes expresados en un determinado momento—, también ha permitido la identificación de genes candidatos de la regeneración. Por ejemplo, en *Perionyx excavatus* y *Pr. leidy* se ha detectado la expresión de *β-catenin*, *Delta* (*Dll1*; Delta), *Distal-less*

(*DLX1*; proteína Homeobox DLX-1), *JNK* (*basket*, *MAPK8*; Cinasa activada por mitógeno 8) y *Wnt* (*wingless*; Wingless), entre otros (Cho *et al.*, 2009; Nyberg *et al.*, 2012). Pero, además, los análisis transcriptómicos han permitido descubrir nuevos genes de anélidos asociados a la regeneración y cuya expresión en el blastema ha sido confirmada con hibridaciones *in situ*; es el caso de *grimp*, *horu*, *mino* y *Tuba* en *En. japonensis* (Takeo *et al.*, 2008, 2010). Más aún, los datos transcriptómicos se han utilizado para llevar a cabo análisis de expresión diferencial en los que se han comparado los perfiles transcripcionales de organismos en estado regenerativo con los de organismos en estado no regenerativo. Con este tipo de análisis, al ser “ómico” y comparativo, se han logrado identificar grandes cantidades de genes diferencialmente expresados (ya sean sobreexpresados o subexpresados) durante la regeneración en *Eisenia fetida*, *En. japonensis*, *Sphaerosyllis hystrix* y *Syllis gracilis* (Bhambri *et al.*, 2018; Myohara *et al.*, 2006; Ribeiro *et al.*, 2019). Además, con las anotaciones ontológicas de los genes diferencialmente expresados, se han llevado a cabo análisis de enriquecimiento de ontologías genéticas, por medio de los cuales se ha inferido que durante la regeneración incrementan la síntesis de proteínas, la replicación del ácido desoxirribonucleico (ADN), la división celular, la actividad de los ribosomas, la unión de ARN y la organización de la matriz extracelular, principalmente (Bhambri *et al.*, 2018; Ribeiro *et al.*, 2019).

De este modo, actualmente se conocen varios genes con actividad transcripcional y algunos procesos biológicos y metabólicos que participan en la regeneración. Sin embargo, las especies en las que se ha estudiado lo anterior son pocas, considerando que la regeneración se encuentra ampliamente representada entre los anélidos (Özpolat & Bely, 2016). Por esta razón, es necesario ampliar el número de especies en las que se lleven a cabo análisis transcriptómicos que permitan comprender, en términos moleculares, qué de la regeneración es común y qué es particular entre las especies del *phylum* Annelida.

2.2. Hipótesis

- El perfil de expresión génica en *Lu. variegatus* es diferente entre los organismos en estado regenerativo y los organismos en estado no regenerativo.
- Algunos genes que se expresan diferencialmente durante la regeneración en *Lu. variegatus* son los mismos que se expresan diferencialmente durante la regeneración en otras especies de anélidos.
- Algunas ontologías genéticas que se encuentran enriquecidas durante la regeneración en *Lu. variegatus* son la mismas que se encuentran enriquecidas durante la regeneración en otras especies de anélidos.

2.3. Objetivos

Objetivo general

- Determinar qué genes se expresan diferencialmente y qué ontologías genéticas se encuentran enriquecidas durante la regeneración en *Lu. variegatus*.

Objetivos particulares

- Generar un transcriptoma anotado a partir de una mezcla de tejidos de organismos en estado regenerativo y en estado no regenerativo.
- Llevar a cabo un análisis de expresión diferencial en el que se comparen los niveles de expresión génica de organismos en estado regenerativo con los de organismos en estado no regenerativo.
- Llevar a cabo análisis de enriquecimiento de ontologías genéticas con los genes diferencialmente expresados durante la regeneración.

2.4. Métodos

2.4.1. Sujetos experimentales

Este proyecto fue aprobado por el Comité de Bioética de la Facultad de Ciencias Naturales (54FCN2018). Las lombrices con las cuales se llevaron a cabo los experimentos fueron adquiridas por medio de un proveedor comercial (Aquarem) y se mantuvieron dentro del laboratorio en contenedores de plástico (~240 cm² de base) con ~0.6 L de agua purificada a ~5 °C. Las lombrices fueron alimentadas con espirulina en polvo, espolvoreando las algas sobre la superficie del agua una o dos veces por semana, dependiendo de la densidad poblacional. Se cambió el agua de los contenedores por agua limpia cada vez que las lombrices fueron alimentadas y los contenedores se limpiaron una vez por semana.

2.4.2. Establecimiento de grupos experimentales, extracción de ARN y secuenciación

300 lombrices sin signos de fisión ni regeneración fueron sometidas a un periodo de aclimatación de una semana en agua purificada a 20 °C. Después, 240 lombrices que se mantuvieron sin signos de fisión ni regeneración hasta el final del periodo de aclimatación se asignaron aleatoriamente a uno de dos grupos: 120 al grupo de organismos en estado regenerativo (GR) y 120 al grupo de organismos en estado no regenerativo (GN). El resto de las lombrices fueron excluidas del experimento. Las 120 lombrices del GR fueron separadas en 5 subgrupos de réplicas biológicas con 24 lombrices cada uno. Las 24 lombrices de cada réplica fueron amputadas haciendo un corte transversal aproximadamente a la mitad del cuerpo con una navaja de afeitar. Los fragmentos anterior y posterior de 8 de ellas se mantuvieron en un periodo de regeneración (PR) de 24 hpa, los fragmentos de 8 más se mantuvieron en un PR de 48 hpa y los fragmentos de otras 8 se mantuvieron en un PR de 72 hpa. Durante los PR las lombrices permanecieron a 20 °C. Pasados los

periodos de regeneración, los fragmentos fueron colectados en un mismo tubo de microcentrífuga con 1 mL de TRIzol (Ambion) y almacenados en congelación a -80 °C. De esta manera, se formaron 5 réplicas del GR, cada una con 24 lombrices (fragmentos anterior y posterior) provenientes de 3 PR: 8 lombrices de 24 hpa, 8 lombrices de 48 hpa y 8 lombrices de 72 hpa. Las lombrices del GN fueron separadas, tratadas, colectadas y almacenadas de la misma forma que las lombrices del GR; excepto que no fueron amputadas.

Previo a la extracción del ARN, los fragmentos de las lombrices en TRIzol fueron descongelados y sus tejidos homogeneizados dentro del tubo. El ARN se extrajo con el kit de purificación Direct-zol RNA MiniPrep (Zymo Research) siguiendo el manual del fabricante. La integridad de las muestras de ARN fue analizada visualmente en un gel de agarosa al 2 % y se midieron las concentraciones y purezas en un NanoDrop 2000 (Thermo Scientific). Se seleccionaron las tres muestras (réplicas) de cada grupo experimental con las mejores purezas y concentraciones para su secuenciación. Las muestras se enviaron a la Unidad Universitaria de Secuenciación Masiva y Bioinformática (UUSMB) de la Universidad Nacional Autónoma de México (UNAM). Las librerías se prepararon con el protocolo Illumina TruSeq Stranded mRNA y se secuenciaron con la plataforma Illumina NextSeq 500.

2.4.3. Ensamblaje y anotación del transcriptoma

Los análisis cuantitativos de expresión diferencial requieren en primera instancia de un genoma o transcriptoma de referencia para estimar la abundancia de los transcritos (Bray *et al.*, 2016). Dado que en el momento en que se llevó a cabo este trabajo no se contaba con ninguna de estas opciones, se optó por ensamblar *de novo* un transcriptoma. La calidad de las lecturas provenientes de la secuenciación fue verificada en FastQC (v.0.11.9; Andrews, 2010). Los sitios con bajo valor PHRED y las lecturas demasiado cortas fueron eliminados con Trimmomatic (v.0.39;

Bolger *et al.*, 2014). Se ensamblaron dos transcriptomas con Trinity (Galaxy v.0.0.1; Haas *et al.*, 2013) que difirieron en el valor del solapamiento mínimo de las lecturas durante el ensamblaje (prd, path reinforcement distance): 75 y 50 pares de bases (pb). Las secuencias de los transcriptomas ensamblados fueron filtradas con CD-HIT-EST (v.4.7; Fu *et al.*, 2012; Li & Godzik, 2006) para eliminar posibles secuencias duplicadas. Con BUSCO (v.3; Simão *et al.*, 2015) se estimó el porcentaje de presencia de genes ortólogos esperados en cualquier transcriptoma de metazoos y el porcentaje de transcritos fragmentados.

El transcriptoma con menor fragmentación fue anotado funcionalmente con Trinotate (v3.3.0; Bryant *et al.*, 2017). Se identificaron las secuencias de nucleótidos con marcos de lectura abiertos que contuvieran secuencias codificantes de al menos 100 aminoácidos, obteniéndose así las secuencias nucleotídicas de genes y transcritos, y las secuencias peptídicas de sus proteínas predichas. Posteriormente, las secuencias de los transcritos y proteínas se alinearon con la base de datos UniProtKB/Swiss-Prot (The UniProt Consortium, 2019) utilizando BLAST (v.2.10.0; Altschul *et al.*, 1997) con un valor umbral de expectativa (E) igual a 0.001. De esta manera se obtuvo la identidad con anotaciones ontológicas de los transcritos y proteínas, y se predijo la presencia de dominios proteicos, péptidos señal y regiones transmembranales. Esta información fue organizada en una base de datos de tipo Trinotate SQLite. Asimismo, se buscó la presencia de secuencias de ARN no codificante con RNAsamba (Camargo *et al.*, 2020) y de secuencias de ARN largo no codificante (ARNInc) por medio de alineamientos en BLAST (E = 0.001) con una base de datos generada en NONCODE (<http://www.noncode.org/download.php>) a partir de secuencias de *Caenorhabditis elegans*, *Danio rerio*, *Drosophila melanogaster*, *Gallus gallus*, *Homo sapiens*, *Mus musculus*, *Rattus norvegicus* y *Saccharomyces cerevisiae*.

2.4.4. Análisis de expresión diferencial

La cuantificación de los transcritos se llevó a cabo en kallisto (v0.46.1; Bray *et al.*, 2016) utilizando el transcriptoma ensamblado como referencia para el pseudo alineamiento de las lecturas. A partir de esta información se realizó un análisis de expresión diferencial a nivel de transcritos con sleuth (Pimentel *et al.*, 2017). En este análisis se llevaron a cabo dos pruebas estadísticas: una de razón de verosimilitud (LRT) para determinar si hubo diferencias significativas entre los grupos experimentales en cuanto a la abundancia de cada transcrito y una de Wald para determinar la magnitud del cambio en la expresión de cada transcrito. Se consideraron como transcritos diferencialmente expresados aquellos cuyo valor de significancia (valor de q de LRT) fuera ≤ 0.05 y cuyo cambio en la expresión fuera del doble o más (sobrexpresado; valor de beta de la prueba de Wald ≥ 1), o de la mitad o menos (subexpresado; valor de beta de la prueba de Wald ≤ 1).

2.4.5. Análisis de enriquecimiento de ontologías genéticas

Con la finalidad de determinar qué términos ontológicos se encuentran enriquecidos durante la regeneración, se llevaron a cabo análisis de enriquecimiento de ontologías genéticas. Debido a que son pocos los genomas y transcriptomas de anélidos anotados, o estos contienen muy pocos genes anotados con evidencia experimental, se utilizaron los ortólogos de los genes de los transcritos diferencialmente expresados en *Crassostrea gigas*, *Da. rerio*, *Dr. melanogaster* y *Xenopus laevis*. Los ortólogos se buscaron en en OrthoDB (v10.1; Kriventseva *et al.*, 2019), UniProtKB (The UniProt Consortium, 2019) y FlyBase (Thurmond *et al.*, 2019). Los identificadores de los ortólogos se usaron para llevar a cabo los análisis en cada especie con ClueGO (v2.5.4; Bindea *et al.*, 2009) de Cytoscape (v3.7.2; Shannon *et al.*, 2003). El valor estadístico (valor de p) del enriquecimiento de cada término se calculó con una prueba hipergeométrica de dos vías con corrección de Bonferroni, usando todos los genes anotados en cada especie como referencia.

Además, solo se consideraron los términos que tuvieran asociados al menos dos genes diferencialmente expresados y que esos genes representaran al menos el 4 % del total de genes asociados al término. Las categorías ontológicas en las que se buscaron términos enriquecidos fueron: proceso biológico, componente celular, componente del sistema inmune, función molecular y vías KEGG (<https://www.genome.jp/kegg/pathway.html>). Finalmente, se construyeron redes para mostrar las asociaciones de los términos enriquecidos con base en los genes diferencialmente expresados que comparten. El grado de asociación entre los términos se calculó con el coeficiente kappa de Cohen (Bindea *et al.*, 2009; Huang *et al.*, 2007).

2.5. Resultados

2.5.1. Ensamblaje y anotación del transcriptoma

Las lecturas secuenciadas se depositaron en el NCBI Sequence Read Archive (BioProject: PRJNA662619, SRR12619265-70). De acuerdo con el análisis de fragmentación, el transcriptoma ensamblado con un *prd* de 50 pb contuvo la mayor cantidad de ortólogos esperados en cualquier transcriptoma de metazoos y el menor porcentaje de fragmentación. 97.7 % de los ortólogos esperados estuvieron presentes y 83.8 % de los transcritos estuvieron completos. Este transcriptoma contuvo 138,008 genes, 164,769 transcritos y 65,692 proteínas predichas. El transcriptoma ensamblado y los archivos de anotación se encuentran disponibles en el repositorio de zenodo (<https://doi.org/10.5281/zenodo.4435184>). 40,037 proteínas predichas presentaron alineamientos significativos con la base de datos UniProtKB/Swiss-Prot y 36,494 tuvieron anotaciones de ontologías genéticas. Las secuencias codificantes, proteínas predichas y sus anotaciones se encuentran disponibles en una base de datos de tipo Trinotate SQLite (ver Archivo Suplementario 2 de Tellez-Garcia *et al.*, 2021). En el análisis de RNAsamba se identificaron 44,097 potenciales ARNInc. De estos, el 88.72 % correspondió a

transcritos en los que no se identificaron secuencias codificantes con Trinotate y 83 de ellos tuvieron alineamientos significativos con la base de datos de NONCODE.

2.5.2. Expresión diferencial durante la regeneración

En el análisis de expresión diferencial se identificaron 136 transcritos diferencialmente expresados ($q \leq 0.05$; Figura 1 y Anexo 3); 25 de ellos estuvieron subexpresados ($\beta \leq -1$) y 111 estuvieron sobreexpresados ($\beta \geq 1$). 13 transcritos clasificados como ARNInc estuvieron sobreexpresados.

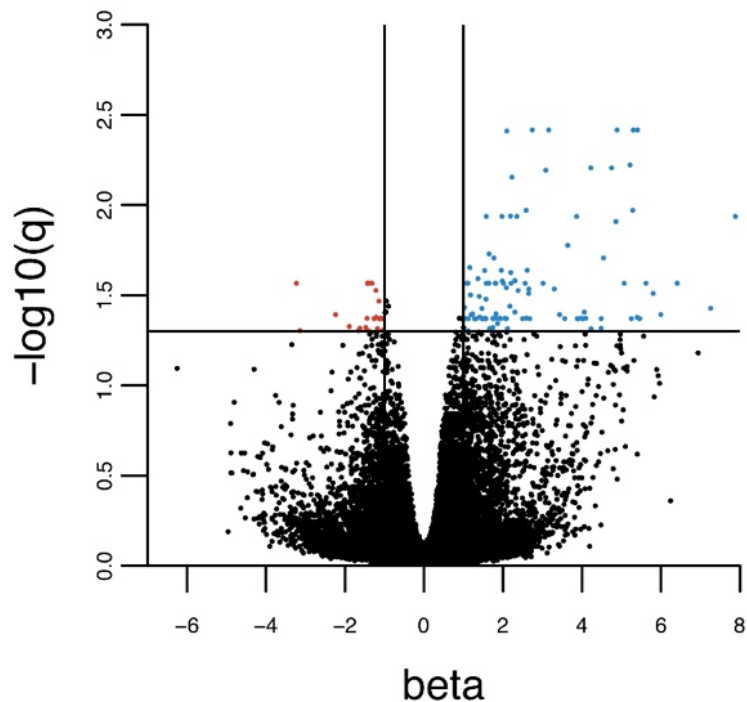


Figura 1. Transcritos diferencialmente expresados durante la regeneración. La gráfica de volcán muestra en \log_2 el cambio en la expresión (β) de cada gen (puntos) con relación a su significancia estadística (q) en una escala $-\log_{10}$. Los puntos ubicados arriba de la línea horizontal representan transcritos con un valor de $q \leq 0.05$ en la LRT. Las líneas verticales señalan el umbral del cambio en la expresión de la mitad ($\beta = -1$) y del doble ($\beta = 1$). Los puntos en rojo representan transcritos subexpresados ($\beta \leq -1$) y los puntos en azul representan transcritos sobreexpresados ($\beta \geq 1$) (Tomado de Tellez-García *et al.*, 2021).

2.5.3. Enriquecimiento de ontologías genéticas durante la regeneración

De las categorías ontológicas en las que se buscaron términos enriquecidos, no se encontraron enriquecimientos de componentes celulares ni de componentes del sistema inmune (Figura 2). Por el contrario, hubo términos enriquecidos de procesos biológicos, funciones moleculares y vías KEGG.

El análisis mostró que la actividad de las descarboxilasas estuvo consistentemente enriquecida en las cuatro especies analizadas debido a la sobreexpresión de *PCK2* (Fosfoenolpiruvato carboxiquinasa [GTP], mitocondrial) y *tdc* (L-tirosina descarboxilasa) (Figura 2). De manera similar, varios términos relacionados con el metabolismo de aminoácidos —principalmente del glutamato y de la glutamina—, estuvieron enriquecidos en las cuatro especies por la subexpresión de *Gs* y la sobreexpresión de *ALDH18A1* (Delta-1-pirrolina-5-carboxilato sintasa), *BHMT* (Betaína--homocisteína S-metiltransferasa 1), *Gls2* (Isoforma hepática de glutaminasa, mitocondrial) y *GLUD2* (Glutamato deshidrogenasa 2, mitocondrial). Finalmente, aunque no fue en todas las especies analizadas, se identificaron enriquecidos el metabolismo de hormonas, de ARN y de compuestos glicosilados; la diferenciación y desarrollo de células endoteliales; el transporte transmembranal de diferentes compuestos; y la actividad de receptores nucleares.

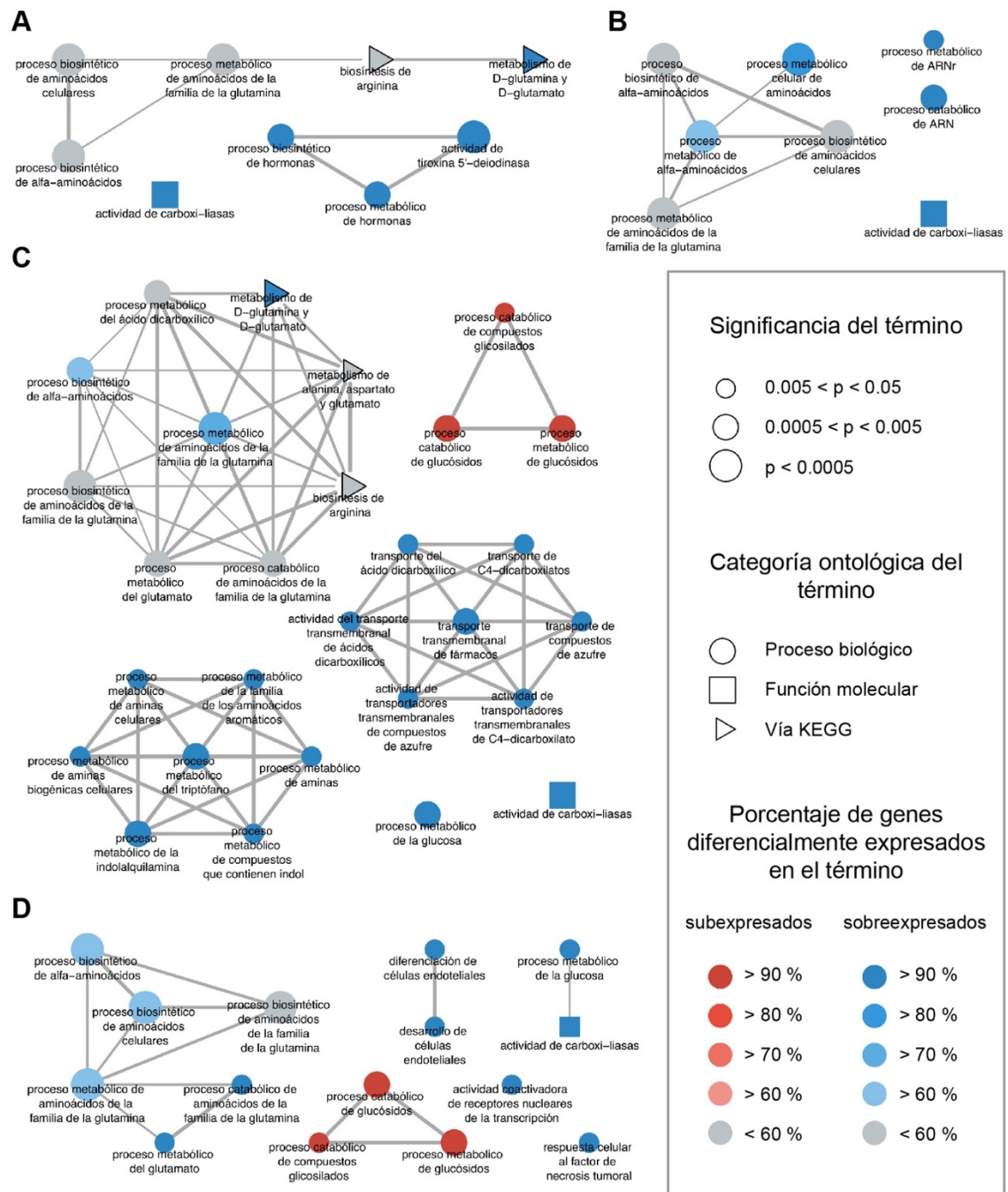


Figura 2. Redes de ontologías genéticas enriquecidas. Los ortólogos de los transcritos diferencialmente expresados fueron utilizados para determinar el enriquecimiento de términos en A, *X. laevis*; B, *Cr. gigas*; C, *Dr. melanogaster*; y D, *Da. rerio*. Los nodos representan los términos enriquecidos, su forma indica la categoría a la que pertenecen y su tamaño la significancia estadística del enriquecimiento. El color de los nodos indica el porcentaje de transcritos subexpresados (rojo) o sobreexpresados (azul) que contribuyeron al enriquecimiento del término. Los conectores de los nodos representan la asociación de dos o más términos por los genes de los transcritos diferencialmente expresados que comparten. El grosor de los conectores indica el grado de asociación (modificado de Tellez-Garcia *et al.*, 2019).

2.6. Discusión

Con la finalidad de conocer el perfil de expresión génica de la regeneración en *Lu. variegatus* y buscar similitudes y diferencias con los mecanismos moleculares que rigen este proceso en otras especies de anélidos, en este trabajo se llevó a cabo un análisis de expresión diferencial en el que se comparó la actividad transcripcional de un GR con el de un GN. Los requerimientos metodológicos del análisis, así como la falta de un genoma o transcriptoma anotado, hicieron necesario el ensamblaje *de novo* de un transcriptoma compuesto por las secuencias expresadas en ambos grupos. El transcriptoma ensamblado fue anotado con los datos de homología para las secuencias nucleotídicas y peptídicas, con información de ontologías genéticas y con el potencial codificante de los transcritos. Este transcriptoma contiene 164,769 transcritos, entre los cuales se encuentran presentes más del 97 % de los transcritos esperados en un transcriptoma de metazoo, y más del 83% están completos. Estos porcentajes indican que se trata de un transcriptoma relativamente bien representado y completo, ya que son similares a los obtenidos en otras especies de anélidos: entre 87.32 y 98.6 % para presencia y entre 66.8 y 98.7 % para genes o transcritos completos (Heikkinen *et al.*, 2017; Ponz-Sagrelles *et al.*, 2018; Ribeiro *et al.*, 2019).

Por otro lado, de los 164,769 transcritos, 99,263 no contenían secuencias codificantes y 44,097 fueron identificados como potenciales ARNInc. Sin embargo, solo 83 de ellos tuvieron alineamientos significativos con una base de datos compuesta por secuencias previamente identificadas como ARN no codificante en varias especies de animales. La presencia de estas secuencias no codificantes en los transcriptomas de anélidos no es rara y su abundancia es muy variable. Por ejemplo, en *Ei. fetida* se identificaron 54,555 transcritos no codificantes de un total de 71,341 transcritos ensamblados (Bhambri *et al.*, 2018), mientras que en *Pygospio elgenas* se encontraron 3,076 de un total de 66,233 (Heikkinen *et al.*, 2017). Además, tal y como se observó en este trabajo (ver Anexo 3), este tipo de transcritos

suelen expresarse diferencialmente en la regeneración de los anélidos (Bhambri *et al.*, 2018; Ribeiro *et al.*, 2019). Por esta razón, resulta importante determinar a qué tipo de ARN no codificante corresponden estas secuencias e identificar si existen grupos de ortólogos que consistentemente se expresen durante la regeneración en varias especies. Esta última información podría servir como punto de partida para llevar a cabo experimentos dirigidos a secuencias de ARN no codificantes candidatas de la regeneración.

En este trabajo se encontraron diferencias entre los perfiles de expresión de los organismos en estado regenerativo y de los organismos en estado no regenerativo. Entre los transcritos diferencialmente expresados que tuvieron alineamientos significativos con la base de datos de UniProtKB, se identificaron algunos cuyos ortólogos participan en la regeneración y en el mantenimiento de células troncales en varios grupos de animales: *bmi1b* (Proteína del complejo Polycomb BMI-1-B), *Hsp60A* (Proteína de choque térmico 60A), *vdr* (HR96; Receptor de vitamina D3) y *VWDE* (Proteína contenedora del factor de von Willebrand D y dominio EGF). *bmi1*, un componente del complejo represor Polycomb 1, regula la expresión de genes *Hox* durante el desarrollo embrionario, promueve la renovación de células troncales de distintos órganos en varios metazoos y está involucrado en la regeneración del páncreas y del epitelio gástrico en mamíferos (Fukuda *et al.*, 2012; Lopez-Arribillaga *et al.*, 2015; Piunti & Shilatifard, 2021; Reinisch *et al.*, 2006). *Hsp60* participa en la oclusión de la herida y en la formación y mantenimiento del blastema de las aletas caudales amputadas en *Da. rerio* y *Paramisgurunus debrayanus* (Li *et al.*, 2014; Makino *et al.*, 2005; Pei *et al.*, 2016). Por su parte, la actividad de *vdr* estimula la proliferación de cardiomiocitos en *Da. rerio*, suprime la proliferación de mioblastos en *M. musculus* e induce la diferenciación de oligodendrocitos en *R. norvegicus* (de la Fuente *et al.*, 2015; Han *et al.*, 2019; Srikuea *et al.*, 2012). Los ortólogos de estos tres genes se encontraron sobreexpresados en el análisis y es posible que durante la regeneración en *Lu. variegatus* estén jugando un papel semejante al que se ha descrito en otras especies (*i.e.*, formación y diferenciación del blastema). En cambio,

VWDE se encontró subexpresado. Leigh *et al.* (2020) han reportado la expresión de este gen en los blastemas de *Ambystoma mexicanum*, *Da. rerio* y *Lepidosiren paradoxa*; así como en la cola en regeneración de *X. laevis*. A pesar de que la función de este gen todavía no ha sido estudiada, algunos experimentos sugieren que no está involucrado en la formación del blastema sino en la determinación del eje proximal distal de esta estructura (Leigh *et al.*, 2020). Sería interesante estudiar su actividad transcripcional en etapas más tardías de la regeneración en *Lu. variegatus* —cuando el blastema ya se haya formado— y, en su caso, buscar la formación de un gradiente morfogénico de su producto proteico.

Además, en este trabajo se encontraron diferencialmente expresados algunos transcritos cuyos ortólogos también se expresan diferencialmente durante la regeneración en otros anélidos. La sobreexpresión de *AIRC* (*paics*; Proteína multifuncional ADE2), un gen involucrado en la síntesis de purinas, ha sido también reportada en *En. japonensis*, *Sp. hystrix* y *Sy. gracilis* (Myohara *et al.*, 2006; Ribeiro *et al.*, 2019). De manera similar, en *En. japonensis* se ha reportado la sobreexpresión de tres genes que aquí también se encontraron sobreexpresados: *BHMT* (Betaína--homocisteína S-metiltransferasa 1), que participa en la síntesis de aminoácidos; *fbl* (ARNr 2'-O-metiltransferasa fibrilarina), que está involucrado en el procesamiento de ARN ribosomal (ARNr); y *GLUD2* —se encontró *GLUD1* en el caso de *En. japonensis*—, que lleva a cabo la conversión de glutamato en alfa-cetoglutarato (Myohara *et al.*, 2006).

La actividad transcripcional de *Gs* que se observó en este trabajo contrasta con lo encontrado previamente en otras especies de anélidos. Mientras que aquí se detectó la subexpresión de este gen durante la formación del blastema, en *En. japonensis* se le observó sobreexpresado durante la misma etapa de la regeneración (Myohara *et al.*, 2006). Además, en esta última especie se corroboró la actividad transcripcional del gen en los blastemas anterior y posterior por medio de hibridaciones *in situ* (Niva *et al.*, 2008). De manera similar, *Gs* no se encontró

diferencialmente expresado en *Sp. hystrix* desde la formación del blastema hasta el inicio de la diferenciación de estructuras externas; pero se encontró sobreexpresado en *Sy. gracilis* (Ribeiro *et al.*, 2019). Estas diferencias requieren de mayor análisis, pero podrían indicar que este gen participa o no en la regeneración, dependiendo de la especie. Asimismo, si se considera que la actividad de la Glutamina sintetasa está relacionada con múltiples procesos celulares que podrían estar ocurriendo durante la regeneración, también es posible que en algunas especies participe en algunos procesos y en otras especies participe en otros; y que las diferencias en los niveles transcripcionales obedezcan a que en algunos procesos se requiere mayor cantidad de la enzima que en otros. Ejemplos de estos procesos son la síntesis de aminoácidos, de nucleótidos y de otros compuestos nitrogenados; así como el reciclamiento del glutamato en el sistema nervioso —si es que estuviera diferenciado— (Castegna & Menga, 2018; Niva *et al.*, 2008). Alternativamente, las disparidades entre este y los trabajos previamente señalados podrían deberse a que la actividad transcripcional de *Gs* es, posiblemente, muy dinámica y a que, a pesar de que en estos trabajos los muestreos se llevaron a cabo durante la formación del blastema, los tiempos de muestreo no fueron exactamente los mismos.

Considerando la subexpresión de *Gs* y la sobreexpresión de *GLS2*, *GLUD2*, *PCK2*, *SLC1A1* (Transportador de aminoácidos excitatorios 3) y *SLC25A10* (Transportador mitocondrial de dicarboxilato); es posible que durante la regeneración en *Lu. variegatus* se esté usando el ciclo de Krebs para producir ATP y otros compuestos o precursores necesarios para satisfacer las necesidades anabólicas. En este escenario, el glutamato estaría siendo importado al citoplasma por medio del Transportador de aminoácidos excitatorios 3 y de ahí a la mitocondria (Kanai & Hediger, 1992). También, dentro de la mitocondria, la glutamina, lejos de sintetizarse por medio de la Glutamina sintetasa, estaría siendo convertida en glutamato por medio de la Glutaminasa (Matés *et al.*, 2019). Esta acumulación de glutamato en la mitocondria permitiría su uso anaplerótico, alimentando el ciclo de

Krebs con alfa-cetoglutarato sintetizado por medio de la Glutamato deshidrogenasa (Plaitakis *et al.*, 2017). La sobreexpresión de *SLC25A10*, un transportador que regula los niveles de malato y succinato en la mitocondria, también da cuenta de la actividad que podría estar teniendo el ciclo de Krebs (Zhou *et al.*, 2015). Además, la Fosfoenolpiruvato carboxiquinasa podría estar utilizando el oxalacetato producido en el ciclo para sintetizar fosfoenolpiruvato. Este último compuesto podría, a su vez, reabastecer el ciclo de Krebs con piruvato, promover la gluconeogénesis o favorecer la síntesis de otros compuestos; como se ha observado bajo condiciones de alta proliferación celular y escasez de glucosa (Vincent *et al.*, 2015).

Recientemente, algunos estudios han demostrado que el estado nutricional, así como el metabolismo energético y de aminoácidos, son cruciales para desencadenar el proceso regenerativo en los animales. Por ejemplo, en los estadios juveniles de *X. laevis* y *X. tropicalis* la capacidad de regenerar la cola decae conforme se agotan las reservas energéticas embrionarias; pero se reestablece conforme los animales son capaces de alimentarse y reabastecer las reservas energéticas por cuenta propia (Williams *et al.*, 2021). De manera similar, el cnidario *Aurelia aurita* es capaz de regenerar morfolácticamente la simetría radial cuando uno o más brazos son amputados durante el estadio juvenil; pero la capacidad de regenerar epimórficamente los brazos es muy limitada (Abrams *et al.*, 2021). Sin embargo, cuando se administra leucina o insulina en el medio, los organismos son capaces de regenerar sustantivamente los brazos perdidos junto con las estructuras musculares y nerviosas que los integran (Abrams *et al.*, 2021). Más aun, algunos organismos típicamente incapaces de regenerar apéndices durante el estadio adulto pueden ganar esta capacidad si se alimentan con compuestos que estimulan el metabolismo energético. *Dr. melanogaster* puede regenerar el fémur, la tibia y su articulación si se enriquece el alimento con leucina e insulina; mientras que *M. musculus* puede regenerar los dígitos desde una distancia anatómica que normalmente no regenera si bebe agua enriquecida con leucina y sacarosa (Abrams *et al.*, 2021). Tomando en cuenta estos estudios, lo anteriormente propuesto con

relación al posible uso anaplerótico del glutamato en el ciclo de Krebs y los resultados de los análisis de enriquecimiento de ontologías genéticas, en los que encontró enriquecimiento del metabolismo de algunos aminoácidos; resulta pertinente llevar a cabo estudios metabolómicos en *Lu. variegatus* que permitan apreciar desde una perspectiva más amplia el papel que podría estar jugando el metabolismo energético y de aminoácidos en la regeneración. Vale la pena preguntarse: ¿Cómo el estado nutricional previo a la amputación influye en la regeneración? ¿Cómo es que los fragmentos posteriores e intermedios resultantes de la amputación satisfacen los requerimientos energéticos de la regeneración cuando no tienen boca? ¿Es posible que la presencia o ausencia de algunos compuestos como la leucina, la sacarosa o la insulina puedan regular la activación o inactivación de la regeneración?

Por otro lado, al análisis de enriquecimiento de ontologías genéticas mostró la existencia de algunas semejanzas con lo reportado en *Sy. gracilis*. En ambas especies se encontró enriquecimiento del metabolismo de ARN, del catabolismo de ARNr y del metabolismo de alfa aminoácidos y aminoácidos de la familia de la glutamina (Ribeiro *et al.*, 2019). Por el contrario, no se encontraron enriquecimientos comunes con *Ei. fetida* ni con *Sp. hystrix* (Bhambri *et al.*, 2018; Ribeiro *et al.*, 2019). Esta variabilidad es resultado de la poca coincidencia que se ha observado en los análisis de expresión diferencial entre las especies de anélidos hasta ahora estudiadas, incluyendo a *Lu. variegatus*. Por ejemplo, los genes candidatos de la regeneración diferencialmente expresados en *Sp. hystrix* y *Sy. gracilis* son poco coincidentes: de 14 genes candidatos, 12 se encontraron diferencialmente expresados en *Sy. gracilis* y solo 3 en *Sp. hystrix* (Ribeiro *et al.*, 2019). Todo esto parece indicar que los procesos moleculares que regulan la regeneración en anélidos son altamente variables entre especies.

Finalmente, cabe señalar que el análisis de expresión diferencial tuvo algunas limitaciones que deben tomarse en cuenta. En primer lugar, dado que el ARN

secuenciado provino de una mezcla de organismos sujetos a tres PR (24, 48 y 72 hpa), es posible que algunos cambios significativos en la expresión de algunos genes dentro de un determinado periodo se hayan diluido entre los tres periodos; impidiendo que estos genes resultaran diferencialmente expresados en el análisis. En consecuencia, también es posible que solo aquellos genes cuya expresión se mantuvo diferencialmente elevada o reducida en los tres PR, o que sus cambios en la expresión fueron lo suficientemente grandes en un determinado PR, pudieran ser identificados como diferencialmente expresados. En segundo lugar, dado que el ARN secuenciado proveniente del GR contuvo fragmentos anteriores y posteriores, es imposible distinguir entre los transcritos diferencialmente expresados que corresponden a la regeneración anterior y a la posterior. Sin embargo, a cambio de estas limitaciones, fue posible ensamblar un transcriptoma con lecturas de tres periodos de regeneración y de las dos secciones corporales, aumentando la probabilidad de hallar transcritos dentro de él y maximizando los recursos disponibles.

En síntesis, los resultados de este capítulo demuestran que existen pocos genes comunes que participan en la regeneración en diferentes especies de anélidos y solo uno (*AIRC*) se sobreexpresa consistentemente en las especies hasta ahora estudiadas. Además, las ontologías genéticas enriquecidas sugieren que las especies analizadas regeneran sin necesariamente preponderar los mismos mecanismos. Si bien esto no apoya la idea de que existen mecanismos moleculares comunes en la regeneración en todos los anélidos, debe considerarse que, sumando este trabajo, solo se cuenta con análisis de expresión diferencial de 5 especies de un *phylum* que agrupa más de 17,000 (Weigert & Bleidorn, 2016). Debido a esto, es necesario seguir explorando esta posibilidad, tratando de abarcar lo mejor posible el árbol filogenético de este grupo.

3. Capítulo 2. Participación de la vía canónica de Wnt en la regeneración de *Lu. variegatus*.

3.1. Antecedentes

La vía canónica de Wnt, también conocida como la vía de Wnt/ β -catenina, es una de las vías de señalización mejor conocidas en términos de transducción de señales y por su participación en diversos procesos biológicos (Cho *et al.*, 2010; Bastakoty & Young, 2016). La activación de esta vía inicia con la unión de un ligando de la familia Wingless (Wnt) al receptor de membrana Frizzled (Fz) y a su correceptor, la Proteína relacionada con el receptor de lipoproteínas de baja densidad 5 o 6 (LRP5/6) (Bastakoty & Young, 2016). La unión de Wnt al complejo Fz-LRP5/6 tiene como consecuencia la translocación de β -catenina del citoplasma al núcleo, en donde β -catenina se une al Factor de células T (TCF) y promueve la transcripción de genes blanco (Bastakoty & Young, 2016). Bajo condiciones de inactivación (*i.e.*, cuando Wnt no se ha unido a Fz), un complejo de degradación, integrado por la Proteína de inhibición del eje (Axin), la Proteína de poliposis adenomatosa coli (APC), la Caseína cinasa 1 (CK1) y la Glucógeno sintasa cinasa 3 (GSK3); induce la degradación de β -catenina mediante su fosforilación en el citoplasma (Amit *et al.*, 2002; Ikeda *et al.*, 1998). β -catenina fosforilada es ubiquitinada por la Proteína contenedora de repeticiones de β -transducina (β -TrCP) y degradada en el proteosoma (Aberle *et al.*, 1997; Marikawa & Elinson, 1998). Asimismo, bajo estas condiciones, la transcripción de genes blanco se encuentra reprimida por la unión de la Proteína potenciadora similar a la transducina (TLE) a TCF (Brantjes *et al.*, 2001). Por el contrario, cuando Wnt se une al complejo Fz-LRP5/6, Dishevelled (Dsh), una proteína presente en el citoplasma, se une a Fz, mientras que Axin es secuestrada por LRP5/6 (Bilic *et al.*, 2007). Esto desestabiliza al complejo de degradación y permite que β -catenina no sea fosforilada, se acumule en el

citoplasma, se transloque al núcleo y active la transcripción de genes blanco junto con TCF (Hanson *et al.*, 2012; Kramps *et al.*, 2002).

En el caso de la regeneración, la vía canónica de Wnt juega diversos papeles, dependiendo de la especie y la estructura en cuestión. En *Hydra*, se ha observado que en mutantes con baja capacidad de regenerar la cabeza y una reducida expresión de *wnt3* en la zona de la lesión, la activación farmacológica de esta vía es capaz de rescatar la regeneración (Lengfeld *et al.*, 2009). Asimismo, la activación de esta vía es necesaria para la formación del blastema en patas amputadas y para la regeneración de la cola en renacuajos de *Xenopus* (Lin & Slack, 2008; Yokoyama *et al.*, 2007). De manera similar, la regeneración en patas amputadas de *Am. mexicanum* se ve afectada por la activación o inactivación de la vía, dependiendo del estadio de regeneración (Kawakami *et al.*, 2006; Wischin *et al.*, 2017). Por otro lado, la inhibición de esta vía por medio del uso de ARN de interferencia (ARNi) contra *wnt1*, β -*catenin* o *Dsh* promueve la formación de cabezas en lugar de colas durante la regeneración posterior de *Schmidtea mediterranea* (Gurley *et al.*, 2008) y *Hofstenia miamia* (Srivastava *et al.*, 2014), lo cual indica que, además de participar en la restitución de tejidos, la vía canónica de Wnt también participa en el establecimiento de los planos corporales durante la regeneración.

En anélidos, la participación de la vía canónica de Wnt durante la regeneración ha sido estudiada desde diversos enfoques. En un análisis transcriptómico llevado a cabo en *Sy. gracilis*, en el cual se compararon los perfiles de expresión génica entre gusanos sujetos a regeneración anterior, regeneración posterior y gusanos no amputados, se encontró que β -*catenin* se encuentra sobreexpresado en la regeneración anterior, comparado con las otras dos condiciones (Ribeiro *et al.*, 2019). También, mediante hibridaciones *in situ*, se ha detectado la expresión de este mismo gen en los blastemas de *Pr. leidyi*, en regeneración anterior y posterior (Nyberg *et al.*, 2012), y *Aelosoma viride*, en regeneración anterior (Chen *et al.*, 2020); observándose señal de expresión desde las 24 hpa y un aumento en la

expresión en todo el blastema entre las 48 y 72 hpa. Por medio de esta misma técnica, se ha detectado también la expresión de *wnt1* en los blastemas de *Ae. viride* (Chen *et al.*, 2020) y de *Pl. dumerilii* (Planques *et al.*, 2019), así como de *wnt4* en *Ae. viride* (Chen *et al.*, 2020).

Desde un punto de vista mecanístico, la modulación de la actividad de la vía canónica de Wnt es crucial para la transición entre la proliferación y la diferenciación del blastema. En *Ae. viride*, el incremento en la proliferación celular durante la formación del blastema está asociado con un aumento en la actividad de la vía a partir de las 24 hpa. Posteriormente, entre las 24 y 48 hpa, una disminución en la actividad, pero no su inhibición total, está relacionada con el inicio de la diferenciación del SNC y la musculatura (Chen *et al.*, 2020). Así, la inactivación farmacológica de la vía tiene como efectos una disminución en la proliferación celular, la formación de blastemas más pequeños, la ausencia de prostomio y la falta de innervación de los órganos anteriores (Chen *et al.*, 2020; Fok *et al.*, 2020). De manera interesante, la sobreactivación farmacológica provoca también una disminución en la proliferación celular y la reducción del tamaño del blastema, así como la inhibición de la diferenciación del neuropilo y las comisuras circumesofágicas que integran la región anterior del SNC (Chen *et al.*, 2020). Si bien una disminución de la proliferación al sobreactivarse la vía resulta contraintuitivo, se ha demostrado que la inactivación farmacológica de la vía disminuye la expresión de genes proapoptóticos y aumenta la de genes antiapoptóticos durante la regeneración (Fok *et al.*, 2020). Esto sugiere que un aumento en la actividad de la vía podría inducir mayor apoptosis y, por lo tanto, un menor número de células en proliferación.

Si bien los trabajos previamente mencionados sugieren que la vía canónica de Wnt podría estar participando en la formación del blastema y su diferenciación, el análisis transcriptómico descrito en el Capítulo 1 no mostró que alguno de los genes de la vía están diferencialmente expresados durante la regeneración temprana en *Lu*.

variegatus (Tellez-Garcia *et al.*, 2021). Sin embargo, sí se encontró sobreexpresión de tres genes cuyos productos participan con esta vía bajo contextos de alta proliferación celular.

Los genes *dio1* y *dio3* codifican las Yodotironina desyodasas tipo 1 (DIO1) y tipo 3 (DIO3), respectivamente (Skah *et al.*, 2017). Ambas enzimas catalizan la degradación de T3, la forma activa de la hormona tiroidea, y de T4, su prohormona, en T2 y rT3, respectivamente. Adicionalmente, DIO1 también cataliza la formación de T3 a partir de T4. T3 participa en varios procesos homeostáticos y del desarrollo, pero, a nivel celular, uno de los efectos que tiene es la modulación de la expresión de genes al unirse con algunos de sus receptores nucleares (Skah *et al.*, 2017). Algunos experimentos en cultivos celulares de cáncer colorrectal han demostrado que las células presentan altos niveles de DIO3 durante la fase de proliferación y que estos niveles descienden durante la fase de diferenciación (Catalano *et al.*, 2016). Los niveles de DIO3 en estas células y en cultivos de células pituitarias están correlacionados con la actividad de la vía canónica de Wnt. Al administrarse T3, los niveles de β -catenina disminuyeron, aumentaron los niveles de algunos reguladores negativos de la vía y se suprimió la expresión de genes blanco (Catalano *et al.*, 2016; Miller *et al.*, 2001). Asimismo, dado que DIO3 es un blanco transcripcional de la vía, la atenuación experimental de β -catenina disminuyó los niveles de DIO3 y aumentó la transcripción dependiente de T3 (Dentice *et al.*, 2012). En ausencia de DIO3, la proliferación se redujo y se promovió la diferenciación por medio de la inducción de E-cadherina, una proteína de membrana plasmática capaz de reducir los niveles nucleares de β -catenina al secuestrarla fuera del núcleo (Dentice *et al.*, 2012). En conjunto, estas observaciones apuntan a que, en ciertos contextos de alta proliferación, existe una retroalimentación positiva entre la actividad de la vía canónica de Wnt y de T3, misma que determina el estado proliferativo o diferenciado de las células.

Por su parte, el Receptor de vitamina D (VDR), tiene un efecto represivo en la actividad de la vía canónica de Wnt y reduce la proliferación celular al ser activado por su ligando, la vitamina D. Este efecto se ha observado, principalmente, en cultivos celulares de cáncer colorrectal (Aguilera *et al.*, 2007; Pálmer *et al.*, 2001; Shah *et al.*, 2006), melanomas (Muralidhar *et al.*, 2019) y cultivos de cardiomiocitos (Hlaing *et al.*, 2014). La represión se da por tres mecanismos diferentes, aunque no excluyentes. Por un lado, VDR induce la expresión de *DKK-1* (Dickkopf-1), un antagonista de Wnt, y de *CK1*, un miembro del complejo de degradación de la vía (Aguilera *et al.*, 2007; Hlaing *et al.*, 2014). También, se ha observado que VDR compite con TCF-4 por la unión de β -catenina, lo que ocasiona una inhibición de la expresión de genes de respuesta del complejo β -catenina/TCF (Pálmer *et al.*, 2001; Shah *et al.*, 2006). Asimismo, VDR induce la expresión de *CDH* (E-cadherina), lo que promueve la reubicación de β -catenina nuclear a la membrana plasmática, lejos de sus genes blanco (Pálmer *et al.*, 2001). Finalmente, se ha observado que, en la epidermis adulta de mamífero, VDR puede actuar como un efector de la vía canónica de Wnt, ya que actúa sinérgicamente con β -catenina para promover la diferenciación de las células del folículo piloso (Pálmer *et al.*, 2008). Cabe señalar que un aumento artificial de los niveles de β -catenina, tiene como consecuencia la formación de tumores (Pálmer *et al.*, 2008). En síntesis, estos trabajos sugieren que VDR, activado por la Vitamina D, tiene un efecto antiproliferativo que se contrapone a la actividad de la vía canónica de Wnt.

La ausencia de expresión diferencial de los miembros de la vía canónica de Wnt y la sobreexpresión de *vdr* indican que esta vía podría estar siendo reprimida durante la regeneración temprana de *Lu. variegatus*. Sin embargo, los experimentos y observaciones llevados a cabo en la regeneración de otras especies de anélidos, en donde se demuestra que la actividad de la vía puede restringirse a ventanas cortas de tiempo, así como la sobreexpresión de *dio1* y *dio3*, vuelven a esta vía atractiva para investigar su posible participación en este proceso, tanto dentro de los PR en los que se tomaron muestras para ensamblar el transcriptoma como fuera

de estos. Con base en estos antecedentes, en este trabajo se analizó el patrón de expresión de *β-catenin* como un marcador de la activación de esta vía a lo largo de las primeras dos semanas de regeneración. Asimismo, se evaluó el efecto de su inhibición y sobreactivación sobre la formación del blastema y del SNC.

3.2. Hipótesis

- La vía canónica de Wnt regula la formación de los blastemas anterior y posterior en *Lu. variegatus*.
- La vía canónica de Wnt es necesaria para la formación del SNC durante la regeneración en *Lu. variegatus*.

3.3. Objetivos

Objetivo general

- Determinar la participación de la vía canónica de Wnt en la regeneración del blastema y del SNC en *Lu. variegatus*.

Objetivos particulares

- Determinar el patrón de expresión de *β-catenin* durante la regeneración.
- Evaluar el efecto de la inhibición y la sobreactivación de la actividad de la vía canónica de Wnt sobre la formación de los blastemas anterior y posterior.
- Evaluar el efecto de la inhibición y la sobreactivación de la actividad de la vía canónica de Wnt sobre la formación del SNC.

3.4. Métodos

3.4.1. Patrón de expresión de β -catenin mediante hibridaciones *in situ*

Se analizó el patrón de expresión de β -catenin durante la regeneración mediante hibridaciones *in situ* de ARN, tomando como base el protocolo utilizado para hibridaciones *in situ* en *Helobdella* de Weisblat & Kuo (2009). Para esto se utilizó una sonda antisentido del ARN mensajero (ARNm) de β -catenin marcada con Fosfatasa alcalina (AP); la cual, al hibridar con el ARNm, permitiera detectar la expresión de este gen en las lombrices por medio de una reacción de color catalizada por dicha enzima. Para dar cuenta de que la reacción de color solamente se diera en presencia de la sonda antisentido marcada, se utilizó como control negativo una sonda sentido, también marcada, pero que no debería hibridar con el ARNm de β -catenin.

3.4.1.1. Diseño experimental

Un grupo de lombrices se mantuvo en un periodo de aclimatación de 48 h a 20 °C sin alimento antes de ser sometidas a diferentes PR. Luego del periodo de aclimatación, las lombrices fueron cortadas aproximadamente por la mitad del cuerpo y se sometieron a 29 PR (5 lombrices por cada PR) desde las 0 hasta las 336 hpa. Cada PR estuvo separado por 12 h. Una vez concluidos los PR, los fragmentos de las lombrices fueron fijados en paraformaldehído (PFA) al 4 % en *buffer* de fosfatos (PBS) y almacenados en metanol a -20 °C. De esta manera, se lograron obtener muestras de lombrices en estado regenerativo cada 12 h, desde las 0 hpa hasta los 14 días post amputación. En un ensayo preliminar se detectaron picos en la expresión de β -catenin a las 60, 72 y 144 hpa en la regeneración anterior y posterior, por lo que adicionalmente se establecieron 3 grupos control sobre los que se evaluó especificidad de la sonda en estos PR. Las lombrices de los grupos

control fueron tratadas del mismo modo que las lombrices anteriormente mencionadas.

3.4.1.2. Síntesis de las sondas

Dentro del transcriptoma anotado (Capítulo 1), se identificó una secuencia de 2,338 pb homóloga a *β-catenin* (TRINITY_DN157465_c2_g3_i2 en Archivo Suplementario 2 de Tellez-Garcia *et al.*, 2021 o en <https://doi.org/10.5281/zenodo.4435184>). Esta homología fue corroborada en BLAST al presentar similitud con los ortólogos de *β-catenin* de otros metazoos. A partir de esta secuencia se diseñaron los oligonucleótidos bcat_for y bcat_rev (Tabla 2) para amplificar un fragmento de 511 pb por medio de una reacción en cadena de la polimerasa (PCR; 2 min a 94 °C; 30X: 30 s a 94 °C, 40 s a 60 °C, 1 min a 72 °C; 3 min a 72 °C), utilizando como templado el producto de una retrotranscripción llevada a cabo con el kit SuperScript II Reverse Transcriptase (Invitrogen) y ARN purificado de *Lu. variegatus*, siguiendo las especificaciones del fabricante.

Tabla 2. Oligonucleótidos utilizados. Las secuencias subrayadas corresponden al promotor T7.

Nombre del oligonucleótido	Secuencia
bcat_for	GTTGCTGGCTTACGGCAATC
bcat_rev	GGTGAGATGTCGCAAAGCAC
M13	GTAAAACGACGGCCAGT
M13-pUC	CAGGAAACAGCTATGAC
T7_bcat_for	<u>TAATACGACTCACTATAGGGAGAGTTGCTTGGCTTACGGCAATC</u>
T7_bcat_rev	<u>TAATACGACTCACTATAGGGAGAGGTGAGATGTCGCAAAGCAC</u>
GFP_for	TTTTCAAGAGTGCCATGCC
GFP_rev	ATGTGGTCTCTCTTTTCGTTG
T7_GFP_for	<u>TAATACGACTCACTATAGGGAGATTTTCAAGAGTGCCATGCC</u>
T7_GFP_rev	<u>TAATACGACTCACTATAGGGAGAATGTGGTCTCTCTTTTCGTTG</u>

Con una parte del producto de PCR se llevó a cabo una electroforesis en gel de agarosa al 1 % para verificar el peso molecular del producto de la reacción y el resto se insertó dentro del vector TOPO con el kit TOPO TA Cloning kit (Invitrogen). Posteriormente, el vector con el inserto se utilizó para transformar células de *Escherichia coli* de la línea XL1-Blue. Las células transformadas se dejaron crecer en cajas de Petri con medio LB sólido a 37 °C durante una noche. Al día siguiente se tomaron muestras de las colonias que crecieron en el medio sólido picando con una punta de micropipeta cada colonia. La punta se sumergió durante un segundo en un tubo de microcentrífuga con mezcla de reacción para PCR y los oligonucleótidos bcat_for y bcat_rev; se retiró del tubo y después se sumergió dentro de un tubo cónico con 4 mL de medio LB líquido. Con la mezcla de reacción y la muestra de células bacterianas que quedaron dentro del tubo, se llevó a cabo un PCR de colonia bajo las mismas condiciones previamente señaladas. Para verificar la presencia de colonias con células transformadas se llevó a cabo una electroforesis en gel de agarosa al 1 %. Los tubos cónicos con LB líquido y muestras de colonias que, efectivamente, resultaron estar transformadas, se dejaron incubando durante la noche a 37 °C y 200 revoluciones por minuto (RPM). Al día siguiente se llevó a cabo una miniprep por cada colonia con el kit QIAprep Spin Miniprep Kit (QIAGEN) para recuperar el vector con el inserto, siguiendo las instrucciones del fabricante.

Parte del producto de la miniprep se envió a secuenciar (macrogen) con los oligonucleótidos M13F y M13-pUC (Tabla 2), entre cuyos sitios de empalme en el vector se deberían encontrar el inserto de 511 pb y el promotor T7. El resto del producto se almacenó a -20 °C. Una vez entregadas las secuencias, se seleccionó uno de los productos de miniprep en donde el complemento reverso de la secuencia codificante del inserto estuvo ubicado en *cis* y en el extremo 3' con respecto al promotor T7 del vector. Utilizando este producto de miniprep como templado, se llevó a cabo una PCR con los oligonucleótidos M13F y M13-pUC bajo las mismas condiciones previamente señaladas. Con una alícuota de este producto de PCR se

midieron la concentración y la pureza en un NanoDrop 2000 y su peso molecular se verificó mediante electroforesis en gel de agarosa al 1 %. Finalmente, se sintetizó una sonda antisentido —o de ARN complementario— al fragmento de 511 pb con el kit MEGAscript (ambion), siguiendo las especificaciones del fabricante y añadiendo 0.2 µg del producto de PCR y 2.5 µL de digoxigenina-UTP (10 mM). La sonda se diluyó a una concentración de 100 ng/µL en *buffer* de prehibridación para generar una solución stock y se almacenó a -20 °C.

Con el producto de PCR amplificado con los oligonucleótidos M13F y M13-pUC, se llevó a cabo un nuevo PCR bajo las mismas condiciones previamente señaladas utilizando los oligonucleótidos *bcat_for* y *bcat_rev*. El producto de esta reacción fue usado para llevar a cabo una nueva PCR con los oligonucleótidos *bcat_rev* y *T7_bcat_for* (Tabla 2), el cual contiene el promotor T7 en el extremo 5' de *bcat_for*. Esta PCR se llevó a cabo también bajo las mismas condiciones previamente señaladas. Con el producto de esta última PCR se llevó a cabo la síntesis de la sonda sentido (534 pb) del mismo modo que en la síntesis de la sonda antisentido.

3.4.1.3. Hibridaciones

Los fragmentos de las lombrices almacenadas en metanol se rehidrataron secuencialmente en PBS con metanol hasta llegar a PBS al 100 %. Después se incubaron en Proteinasa K (10 µg/mL) por 20 min a 37 °C y se detuvo la acción de la enzima incubando en glicina (2 mg/mL) por 1 min. Posteriormente se incubaron dos veces en *buffer* de trietanolamina (0.1 M; pH 8), agregando dos veces 3 µL de anhídrido acético por cada mililitro de *buffer*, cada 5 min. Luego, se fijaron nuevamente en PFA al 4 % por 20 min y se enjuagaron secuencialmente en PBS con Tween (PBT) al 0.1 %, en PBT:*buffer* de prehibridación (1:1) y en *buffer* de prehibridación. Terminados los enjuagues, las lombrices se dejaron incubando en *buffer* de prehibridación durante la noche a 60 °C en agitación y 60 RPM.

Para la hibridación se preparó una solución de hibridación mezclando la solución stock de la sonda con *buffer* de prehibridación (1:50). Después, la solución de hibridación se precalentó durante 5 min a 90 °C para desnaturalizar la sonda y se dejó enfriar lentamente a temperatura de hibridación (60 °C). Las lombrices se incubaron con la sonda a temperatura de hibridación durante la noche en agitación. Al día siguiente, la sonda se removió enjuagando las lombrices en mezclas de *buffer* de prehibridación: *buffer* de solución salina y citrato de sodio (SSC) 2X (2:1, 1:1, 1:2) a 60 °C, *buffer* SSC (2X, 0.2X, 0.1X) a 60 °C, mezclas de *buffer* 0.1X SSC:PBT (2:1, 1:1, 1:2) a temperatura ambiente y PBT. Antes de llevar a cabo el marcaje con anticuerpo, las lombrices se enjuagaron una vez en solución de bloqueo y se incubaron por 2 h en solución de bloqueo nueva. Posteriormente, se agregó a la solución de bloqueo el anticuerpo Antidigoxigenina conjugado con AP (Roche: 11093274910) en un factor de dilución 1/5,000 y las lombrices se dejaron incubando en esta solución con anticuerpo durante la noche a 4 °C.

A la mañana siguiente, las lombrices se lavaron varias veces en PBT en agitación (80 RPM), se enjuagaron cuatro veces en *buffer* de AP y se incubaron en una solución de NBT/BCIP (Roche: 11681451001) con 50X volúmenes de *buffer* de AP para llevar a cabo la reacción de color. Las lombrices se revisaron bajo el estereoscopio cada 15 min hasta lograr la coloración deseada (entre 1.5 y 2 h) y se enjuagaron varias veces en PBT para detener la reacción. Finalmente, las lombrices se almacenaron en PBT a 4 °C. Para su visualización, las lombrices fueron montadas en un portaobjetos excavado con PBS y fotografiadas de perfil bajo el microscopio.

3.4.2. Inhibición farmacológica de la vía canónica de Wnt durante la regeneración

Para evaluar el efecto de la inhibición de la vía de Wnt durante la regeneración se utilizó el fármaco IWR-1 (Merck: I0161), el cual es un estabilizador de Axin y, por lo

tanto, un promotor de la degradación de β -catenina (Chen *et al.*, 2009). 18 lombrices fueron sometidas a un periodo de aclimatación de 48 h sin alimento y luego cortadas aproximadamente por la mitad del cuerpo. Los fragmentos anteriores y posteriores fueron separados e incubados en IWR-1 disuelto en dimetilsulfóxido (DMSO; 5 mg/mL) como vehículo y a una concentración final de 15 μ M en agua purificada. Como controles negativos se incubaron 18 fragmentos anteriores y posteriores de lombrices en DMSO al 0.12 % en agua purificada y 18 más en agua purificada. Se evaluó el efecto del fármaco sobre dos PR alrededor de las 60 hpa (inicio de la expresión generalizada de β -catenin): i) aplicando el fármaco de las 48 a las 72 hpa y ii) aplicando el fármaco de las 48 a las 96 hpa. Se verificó que las lombrices tuvieran movilidad y circulación sanguínea después de ser incubadas para incluirlas en los análisis. Las lombrices que sufrieron autotomía fueron descartadas. Una vez transcurridos los PR, las lombrices fueron fijadas en PFA al 4 %, deshidratadas en metanol y almacenadas a -20 °C. Para su visualización, las lombrices fueron rehidratadas secuencialmente en PBS:metanol hasta llegar a PBS al 100 %, montadas en un portaobjetos excavado con PBS y fotografiadas de perfil bajo el microscopio. El área del blastema fue medida con ImageJ (Schneider *et al.*, 2012). Para el análisis estadístico, se compararon las medias con un ANOVA cuando los datos cumplieron con los supuestos de normalidad y con un ANOVA de Welch cuando los datos no presentaron homogeneidad de varianzas. Cuando los datos no presentaron distribución normal, las medianas se compararon mediante un análisis de Kruskal-Wallis.

3.4.3. Inhibición de la vía canónica de Wnt por medio de ARNi durante la regeneración

Además de evaluarse el efecto de la inhibición de la vía canónica de Wnt por medio de un fármaco, también se probó el uso de ARNi con este mismo propósito. Wang *et al.* (2015) y Sarker *et al.* (2017) han sumergido en ARN de doble cadena (ARNdc) células cultivadas de *Haliothis diversicolor* e individuos enteros de *Tubifex tubifex*,

respectivamente; y han demostrado su eficacia para silenciar la expresión génica de manera específica. Con base en estos estudios se optó por probar el efecto de sumergir fragmentos de organismos en estado regenerativo en ARNdc de β -catenin. Como control negativo, se sumergieron fragmentos de organismos en estado regenerativo en agua purificada y en ARNdc del gen de la Proteína verde fluorescente (*GFP*), cuya ausencia en *Lu. variegatus* se determinó mediante alineamientos con el transcriptoma usando BLAST.

Con el producto de PCR utilizado en la elaboración de las sondas de las hibridaciones *in situ*, se amplificó y purificó el fragmento de 511 pb de β -catenin bajo las condiciones de reacción previamente mencionadas. Este producto de PCR se utilizó como templado para reamplificar el fragmento, también bajo las mismas condiciones de reacción, pero ahora utilizando los oligonucleótidos T7_bcat_for y T7_bcat_rev (Tabla 2), los cuales contienen el promotor T7 en cada extremo. Para generar el ARNdc de *GFP*, se amplificó un fragmento de 508 pb por medio de PCR (2 min a 94 °C; 30X: 30 s a 94 °C, 40 s a 56 °C, 1 min a 72 °C; 3 min a 72 °C) con los oligonucleótidos GFP_for y GFP_rev (Tabla 2) y utilizando como templado el vector pPD95.75 (<http://www.addgene.org/1494/>). Posteriormente, se reamplificó el producto de PCR con los oligonucleótidos T7_GFP_for y T7_GFP_rev (Tabla 2) bajo las mismas condiciones de reacción. Usando una alícuota de este producto de PCR se midieron la concentración y la pureza en un NanoDrop 2000 y su peso molecular se verificó mediante electroforesis en gel de agarosa al 1 %. Con los productos de PCR etiquetados con el promotor T7 en ambos extremos, se sintetizó ARNdc con el MEGAscript RNAi Kit, siguiendo las instrucciones del fabricante.

De manera similar al experimento de inhibición farmacológica, 20 lombrices fueron sometidas a un periodo de aclimatación de 48 h sin alimento y luego cortadas aproximadamente por la mitad del cuerpo. Los fragmentos anteriores y posteriores fueron separados e incubados en 2 mL de solución de ARNdc de β -catenin (β -catenin-dc) a una concentración de 5 μ g/mL en agua purificada. Para los controles

negativos, se incubaron 20 fragmentos anteriores y posteriores de lombrices en 2 mL de solución de ARNdc de *GFP* (*GFP-dc*) a una concentración de 5 µg/mL en agua purificada; así como 20 fragmentos anteriores y posteriores en 2 mL de agua purificada. Se evaluó el efecto del ARNdc sobre dos PR alrededor de las 60 hpa (inicio de la expresión generalizada de *β-catenin*): i) aplicando el ARNdc de las 24 a las 72 hpa y ii) aplicando el fármaco de las 48 a las 72 hpa. Se verificó que las lombrices tuvieran movilidad y circulación sanguínea después de ser incubadas para incluirlas en los análisis. Las lombrices que sufrieron autotomía fueron descartadas. Una vez transcurridos los PR, las lombrices fueron fijadas en PFA al 4 %, deshidratadas en metanol y almacenadas a -20 °C. Para su visualización, las lombrices fueron rehidratadas secuencialmente en PBS:metanol hasta llegar a PBS al 100 %, montadas en un portaobjetos excavado con PBS y fotografiadas de perfil bajo el microscopio. El área del blastema fue medida con ImageJ (Schneider *et al.*, 2012). Para el análisis estadístico, se compararon las medias con un ANOVA cuando los datos cumplieron con los supuestos de normalidad o las medianas con un análisis de Kruskal-Wallis cuando los datos no presentaron distribución normal.

3.4.4. Sobreactivación farmacológica de la vía canónica de Wnt durante la regeneración

Para evaluar el efecto de la sobreactivación de la vía canónica de Wnt durante la regeneración se utilizó cloruro de litio (LiCl), el cual es un inhibidor de la actividad de GSK3 y, por lo tanto, un activador de la vía (Clément-Lacroix *et al.*, 2005). 20 lombrices fueron sometidas a un periodo de aclimatación de 48 h sin alimento y luego cortadas aproximadamente por la mitad del cuerpo. Los fragmentos anteriores y posteriores fueron separados e incubados en LiCl disuelto en agua purificada a diferentes concentraciones: 1, 2, 5 y 10 mM. Como controles negativos se incubaron 20 lombrices en agua purificada. Se evaluó el efecto del fármaco alrededor de las 60 hpa (inicio de la expresión generalizada de *β-catenin*), aplicando el fármaco de las 48 a las 72 hpa. Se verificó que las lombrices tuvieran movilidad y circulación

sanguínea después de ser incubadas para incluirlas en los análisis. Las lombrices que sufrieron autotomía fueron descartadas. Una vez transcurrido el PR, las lombrices fueron fijadas en PFA al 4 %, deshidratadas en metanol y almacenadas a -20 °C. Para su visualización, las lombrices fueron rehidratadas secuencialmente en PBS:metanol hasta llegar a PBS al 100 %, montadas en un portaobjetos excavado con PBS y fotografiadas de perfil bajo el microscopio. El área del blastema fue medida con ImageJ (Schneider *et al.*, 2012). Para el análisis estadístico, se compararon las medias con un ANOVA cuando los datos cumplieron con los supuestos de normalidad o las medianas con un análisis de Kruskal-Wallis cuando los datos no presentaron distribución normal.

3.4.5. Formación del SNC bajo la inhibición y la sobreactivación de la vía canónica de Wnt durante la regeneración

Para evaluar el efecto de la inhibición y la sobreactivación de la vía canónica de Wnt sobre la formación del SNC durante la regeneración, se observó el crecimiento del CNV en blastemas anteriores (*i.e.*, blastemas de fragmentos posteriores) de lombrices en estado regenerativo incubadas en IWR-1 y LiCl. Las lombrices fueron sometidas a un periodo de aclimatación de 48 h sin alimento y luego cortadas aproximadamente por la mitad del cuerpo. Para evaluar el efecto de la inhibición de la vía, los fragmentos posteriores fueron separados e incubados en IWR-1 disuelto en DMSO (5 mg/mL) a una concentración de 15 μ M en agua purificada. Como control negativo se incubaron fragmentos posteriores en DMSO al 0.12 % en agua purificada. Para evaluar el efecto de la sobreactivación de la vía, fragmentos posteriores fueron separados e incubados en LiCl 2 mM disuelto en agua purificada. Como control negativo se incubaron fragmentos posteriores en agua purificada. Los fragmentos se incubaron en las soluciones de las 48 a las 72 hpa. Se verificó que las lombrices tuvieran movilidad y circulación sanguínea después de ser incubadas para incluirlas en los análisis. Las lombrices que sufrieron autotomía fueron descartadas. Una vez transcurrido el PR, las lombrices fueron fijadas en PFA al 4

%, deshidratadas en metanol y almacenadas a -20 °C. Adicionalmente y con la finalidad de probar la efectividad del marcaje de neuronas, algunas cabezas de lombrices en estado no regenerativo se fijaron, deshidrataron y almacenaron como se describió anteriormente.

Los fragmentos de las lombrices (tres por tratamiento) almacenadas en metanol se rehidrataron secuencialmente en PBS con metanol hasta llegar a PBS al 100 %. Los tejidos se permeabilizaron enjuagándose ocho veces con Tritón X-100 al 0.1 % y 3 veces con PBS; después, se incubaron en Proteinasa K (10 µg/mL) por 25 min a 37 °C y se detuvo la acción de la enzima incubándolos en glicina (2 mg/mL) por 1 min. Luego, las lombrices se fijaron nuevamente en PFA al 4 % por 20 min, se enjuagaron tres veces en PBT y se incubaron por 1 h en solución de bloqueo. El marcaje de los axones del CNV se llevó a cabo mediante inmunohistoquímica de fluorescencia, incubando las lombrices en solución de bloqueo con un anticuerpo primario monoclonal para Tubulina acetilada (Anti-Acetylated Tubulin antibody produced in mouse; Sigma: T7451) diluido 1:500 por 72 h a 4 °C. Posteriormente, las lombrices se enjuagaron en PBT y se incubaron nuevamente en un anticuerpo secundario conjugado con rodamina (Anti-Mouse IgG (H+L)-Rhodamine antibody produced in donkey, Sigma: SAB3701099) diluido 1:400 en solución de bloqueo por ~42 h a temperatura ambiente y en oscuridad. Finalmente, las lombrices se enjuagaron seis veces en PBT, se montaron en un portaobjetos excavado con glicerol al 80 % en PBS y se fotografiaron bajo el microscopio de epifluorescencia (UNICO: G500) con una cámara digital de 5.0 Megapíxeles (Moticam: 580).

3.5. Resultados

3.5.1. Patrón de expresión de *β-catenin* durante la regeneración

Mediante hibridaciones *in situ* se determinó el patrón de expresión de *β-catenin* de las 0 a las 336 hpa cada 12 h. En las lombrices de los grupos control no se observó

señal de la expresión (Figura 3), lo cual indica que en ausencia de hibridación de la sonda no hay marcaje del ARNm. En la regeneración anterior no se observó señal de las 0 a las 48 hpa (Figura 4); en cambio, a las 60 hpa, cuando el blastema se encuentra casi completamente formado, se observó señal intensa en prácticamente todo el tejido en estado regenerativo. Esta señal decayó sin perderse por completo a las 72 hpa y, posteriormente, de las 84 a las 96 hpa se intensificó en la región ventral del tejido en estado regenerativo. De las 108 a las 120 hpa la señal volvió a generalizarse con menor intensidad en todo el tejido en estado regenerativo. A partir de las 132 hpa la expresión de *β-catenin* volvió a regionalizarse, pero esta vez en el prostomio, mientras que la expresión en el resto del tejido en estado regenerativo fue perdiéndose con el tiempo. Finalmente, a partir de las 300 hpa la señal de la expresión fue casi imperceptible.

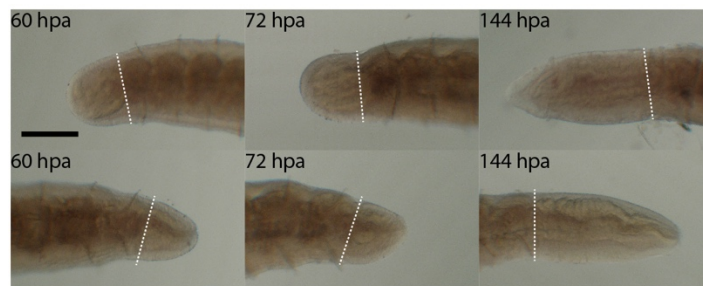


Figura 3. Imagen representativa de las hibridaciones *in situ* de los grupos control (sonda sentido). Los paneles superiores muestran imágenes de regeneración anterior y los paneles inferiores de regeneración posterior. Las líneas punteadas indican el sitio de corte. En todos los cuadros la región anterior de las lombrices está orientada a la izquierda y la región ventral hacia abajo. hpa, horas post amputación. Barra de escala = 0.2 mm (misma para todos los recuadros).

En la regeneración posterior no se observó señal de las 0 a las 24 hpa (Figura 5), mientras que a las 36 y 48 hpa se observó señal de la expresión en la región ventral del blastema. Posteriormente, de las 60 a las 84 hpa, la señal fue paulatinamente aumentando hasta cubrir por completo el tejido en estado regenerativo a las 96 y 108 hpa. La intensidad de la señal se redujo a partir de las 129 hpa, manteniéndose más elevada en la región ventral y reduciéndose paulatinamente en las regiones más proximales del tejido en estado regenerativo. De esta manera, conforme la

sección posterior del animal fue creciendo en longitud, la señal de la expresión solo se mantuvo en la región más distal, cercana a la zona posterior de crecimiento. Finalmente, a partir de las 288 hpa, cuando la morfología externa del tejido en estado regenerativo es muy similar a la del tejido no amputado, la señal fue casi imperceptible. Cabe señalar que a partir de las 192 hpa se observaron unos puntos o manchas con señal de expresión en la región lateral de algunos segmentos en estado regenerativo; un punto por cada segmento. Estos puntos fueron observados solo en los segmentos con estructuras en proceso de diferenciación (determinados así por la presencia de segmentación, el aparente intestino y quetas en desarrollo), estando ubicados en la piel, exactamente alrededor de las quetas en crecimiento; y estuvieron ausentes en los segmentos con quetas completamente desarrolladas.

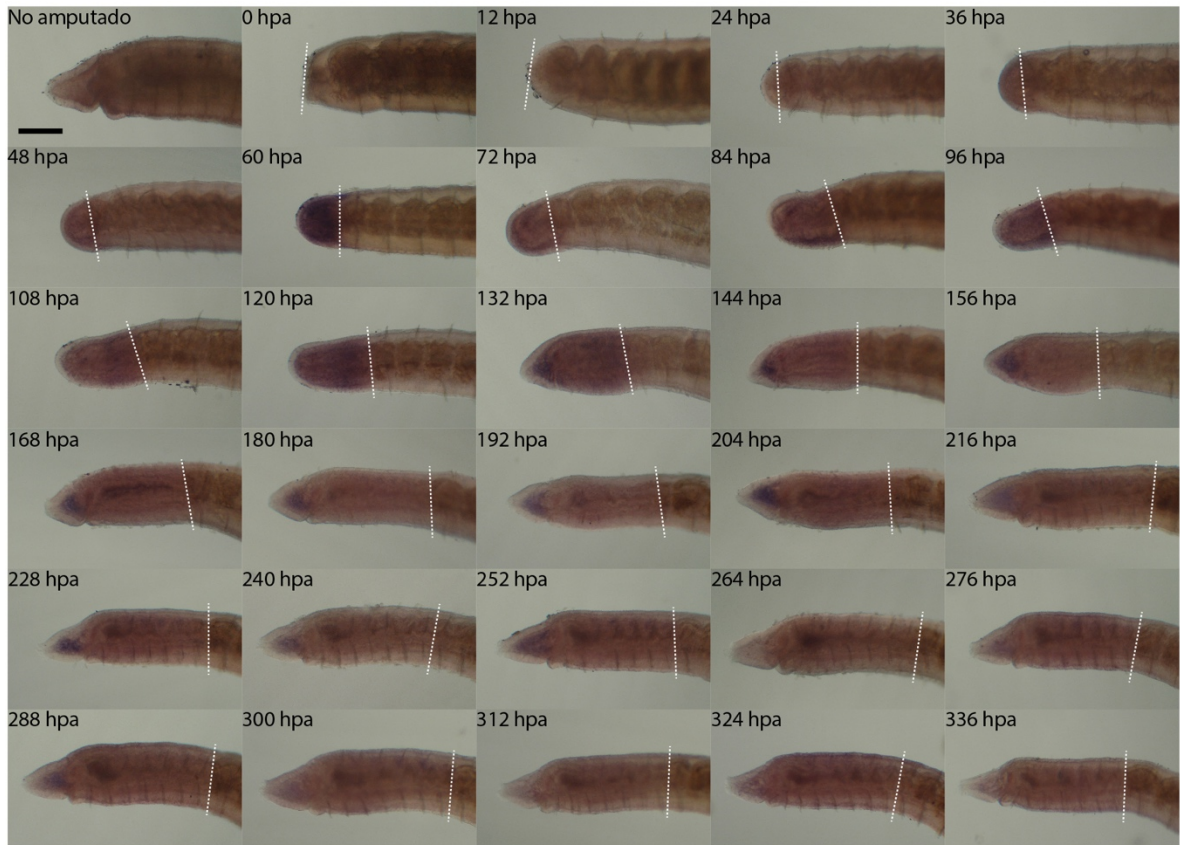


Figura 4. Imagen representativa de las hibridaciones *in situ* de las 0 a las 336 horas post amputación (hpa), cada 12 h, en la regeneración anterior. Las líneas punteadas indican el sitio de corte. En todos los cuadros, la región anterior de las lombrices está orientada a la izquierda y la región ventral hacia abajo. Barra de escala = 0.2 mm (misma para todos los recuadros).

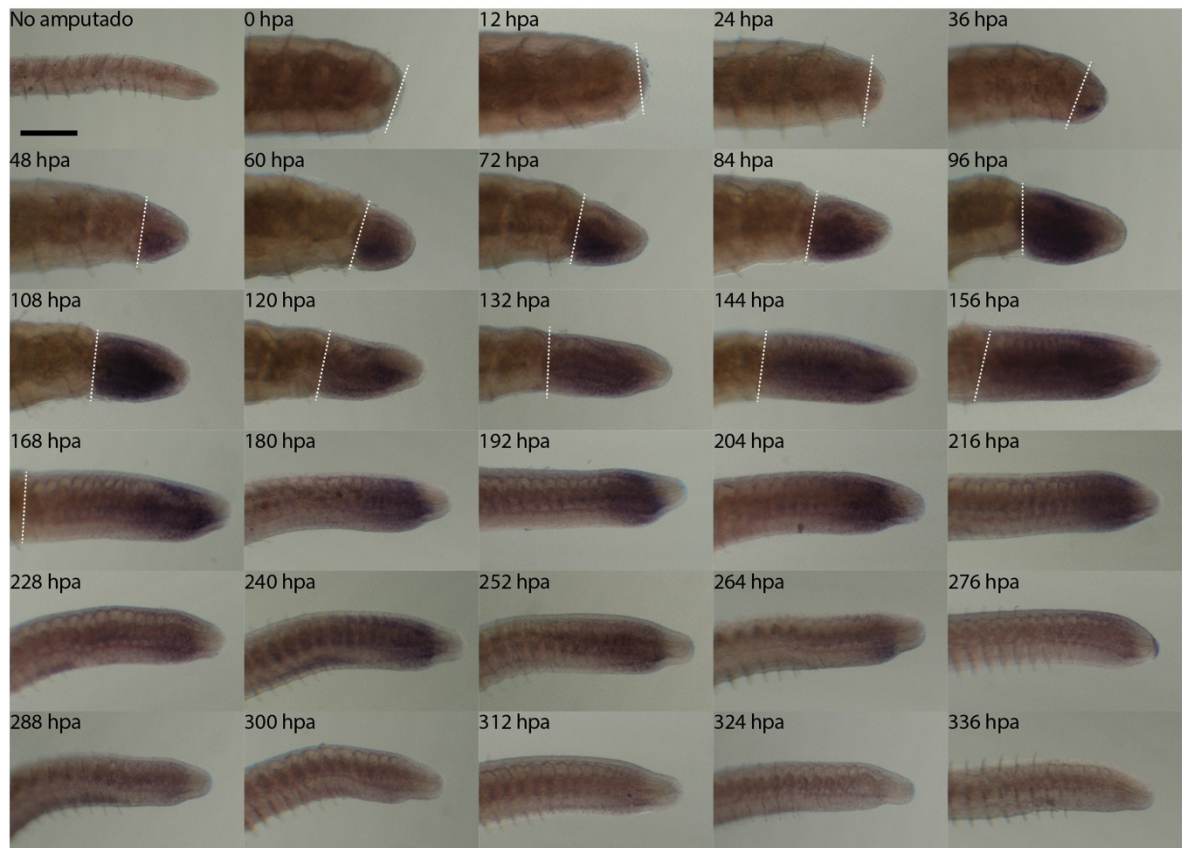


Figura 5. Imagen representativa de las hibridaciones *in situ* de las 0 a las 336 horas post amputación (hpa), cada 12 h, en la regeneración posterior. Las líneas punteadas indican el sitio de corte. A partir de las 180 hpa no se muestra el sitio de corte debido al aumento del tamaño del tejido regenerativo. En todos los cuadros, la región anterior de las lombrices está orientada a la izquierda y la región ventral hacia abajo. Barra de escala = 0.2 mm (misma para todos los recuadros).

3.5.2. Efecto de la inhibición farmacológica de la vía canónica de Wnt en el tamaño del blastema durante la regeneración

Se observó un efecto de la aplicación de IWR-1 sobre el tamaño del blastema. En la regeneración anterior, las lombrices tratadas con IWR-1 presentaron blastemas significativamente más pequeños, comparados con los blastemas de las lombrices incubadas en DMSO y agua. Este efecto se observó en las lombrices tratadas de las 48 a las 72 hpa (ANOVA de Welch: $p < 0.001$; Figura 6A) y en las lombrices tratadas de las 48 a las 96 hpa (ANOVA: $p < 0.001$; Figura 7A). En ambos casos,

no se encontraron diferencias significativas entre las lombrices incubadas en DMSO y agua.

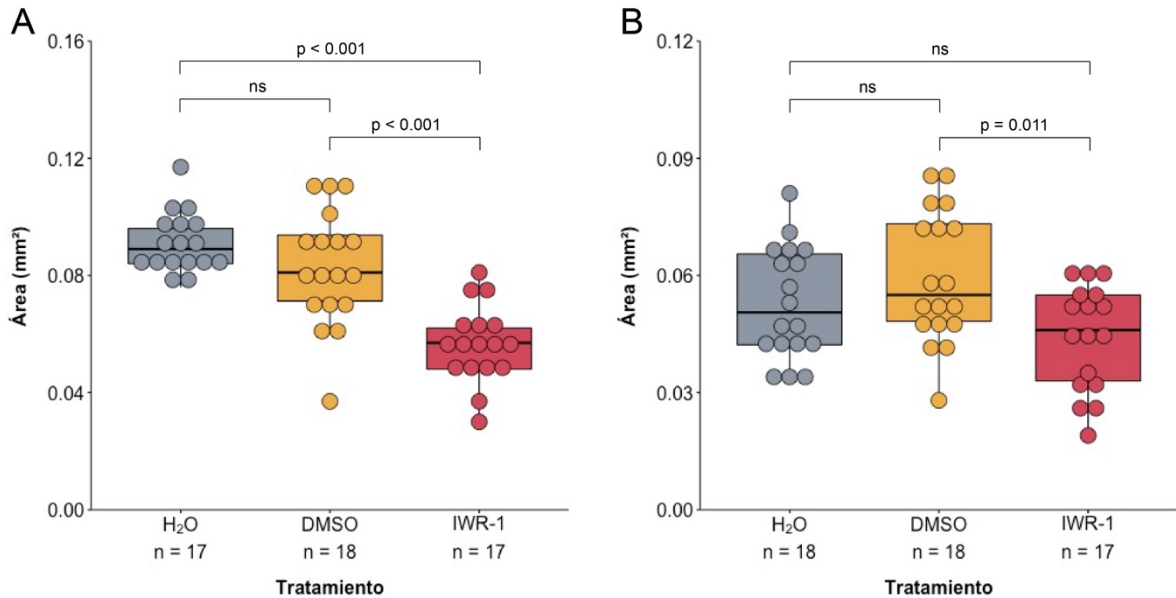


Figura 6. Áreas de los blastemas tratados con IWR-1 de las 48 a las 72 hpa. Tanto en la regeneración anterior (A) como en la posterior (B) se observó que los blastemas tratados con IWR-1 presentaron menor tamaño que aquellos tratados con DMSO. Los valores de p corresponden a las pruebas *post hoc* de Games-Howell; ns, no significativo.

De manera similar, en la regeneración posterior las lombrices tratadas con IWR-1 de las 48 a las 72 hpa presentaron blastemas significativamente más pequeños que los de las lombrices incubadas en DMSO; pero no se observó esta diferencia entre las lombrices tratadas con IWR-1 y aquellas incubadas en agua (ANOVA: $p = 0.0154$; Figura 6B). Finalmente, en el caso de la regeneración posterior, las lombrices tratadas de las 48 a las 96 hpa no mostraron diferencias significativas entre los tres tratamientos (Kruskal-Wallis: $p = 0.9081$; Figura 7B).

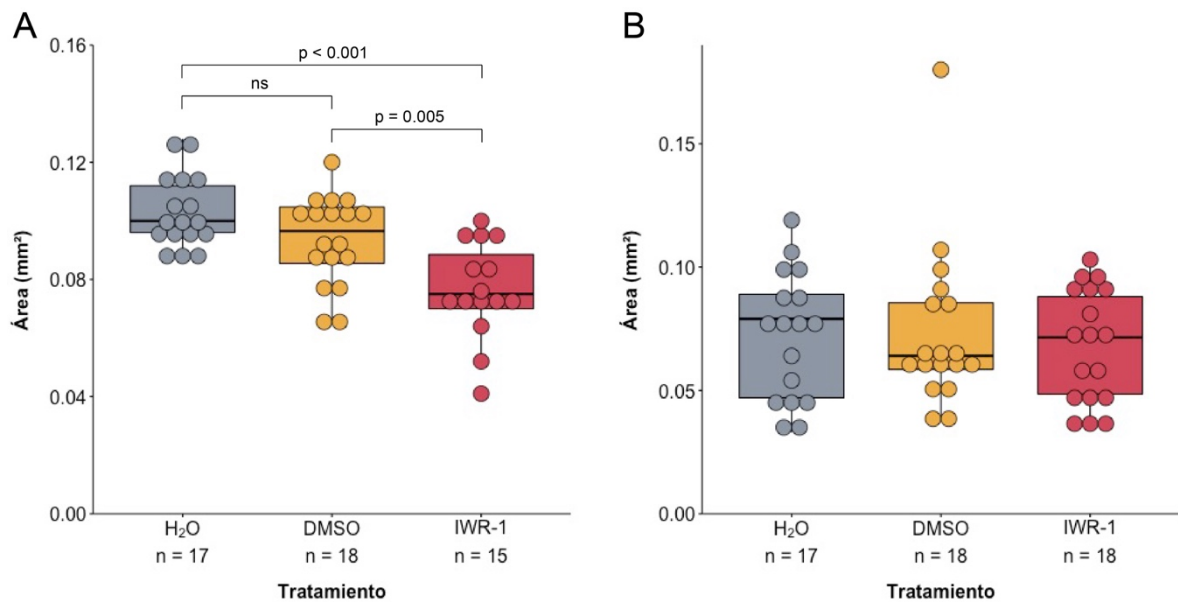


Figura 7. Áreas de los blastemas tratados con IWR-1 de las 48 a las 96 hpa. En la regeneración anterior (A) se observó que los blastemas tratados con IWR-1 presentaron menor tamaño que aquellos tratados con DMSO. Esta diferencia no se observó en la regeneración posterior (B). Los valores de p corresponden a la prueba *post hoc* de Tukey; ns, no significativo.

3.5.3. Efecto de la inhibición de la vía canónica de Wnt por medio de ARNi en el tamaño del blastema durante la regeneración

Dado que con la inhibición farmacológica se observaron las diferencias más significativas en el periodo de incubación de las 48 a las 72 hpa, en este mismo periodo se probó el efecto del ARNdc. Los tamaños de los blastemas no mostraron diferencias significativas entre las lombrices tratadas con *β-catenin-dc*, *GFP-dc* y agua purificada; tanto en la regeneración anterior (ANOVA: $p = 0.366$; Figura 8A) como en la posterior (ANOVA: $p = 0.0619$; Figura 8B).

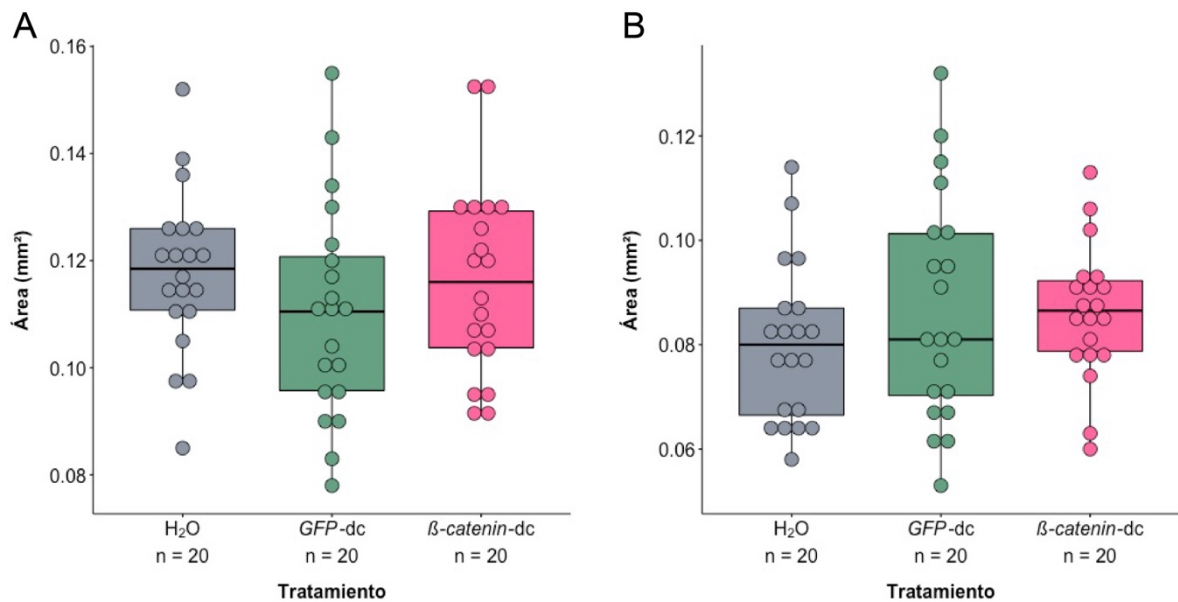


Figura 8. Áreas de los blastemas tratados con ARNdc de las 48 a las 72 hpa. Tanto en la regeneración anterior (A) como en la posterior (B) no se observaron diferencias en los tamaños de los blastemas tratados con β -catenin-dc y GFP-dc.

Al observarse la ausencia de diferencias significativas, se consideró la posibilidad de que el ARNdc requiriera más tiempo para tener un efecto inhibitor. Por este motivo, se incubaron lombrices en ARNdc desde las 24 hasta las 72 hpa. En este segundo ensayo, los tamaños de los blastemas tampoco mostraron diferencias significativas, tanto en la regeneración anterior (Kruskal-Wallis: $p = 0.3899$; Figura 9A) como en la posterior (Kruskal-Wallis: $p = 0.5259$; Figura 9B).

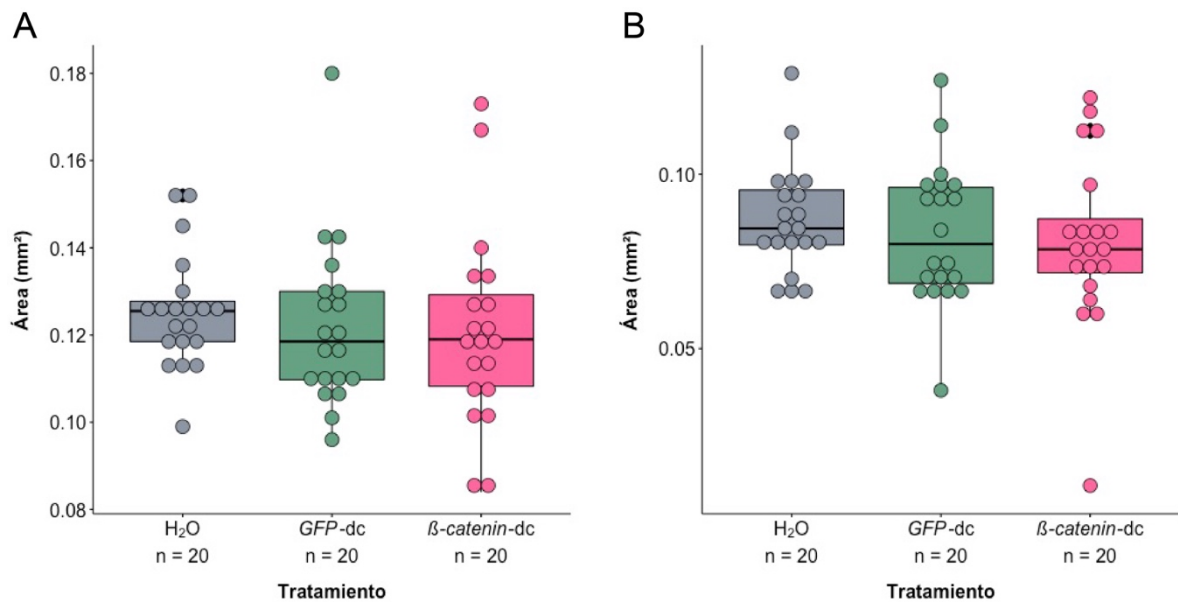


Figura 9. Áreas de los blastemas tratados con ARNdc de las 24 a las 72 hpa. Tanto en la regeneración anterior (A) como en la posterior (B) no se observaron diferencias en los tamaños de los blastemas tratados con β -catenin-dc y GFP-dc.

3.5.4. Efecto de la sobreactivación farmacológica de la vía canónica de Wnt en el tamaño del blastema durante la regeneración

Del mismo modo que en el experimento de inhibición por medio de ARNdc, el efecto del LiCl se probó aplicándolo de las 48 a las 72 hpa por ser el periodo de incubación en que se observaron las diferencias más significativas con el IWR-1. Las lombrices tratadas con LiCl 5 y 10 mM fueron descartadas al observarse degradación de tejidos a las 72 hpa. En contraste, las lombrices tratadas con LiCl 1 y 2 mM no presentaron degradación de tejidos, pero sí movilidad corporal y circulación sanguínea. En la regeneración anterior, las lombrices tratadas con LiCl no presentaron diferencias significativas en el tamaño del área de los blastemas, comparados con los blastemas de las lombrices incubadas en agua (ANOVA: $p = 0.0933$; Figura 10A). En cambio, en la regeneración posterior las lombrices tratadas con LiCl 2 mM presentaron blastemas significativamente más pequeños que los de las lombrices incubadas en LiCl 1 mM y en agua (Kruskal-Wallis: $p = 0.0028$; Figura 10B).

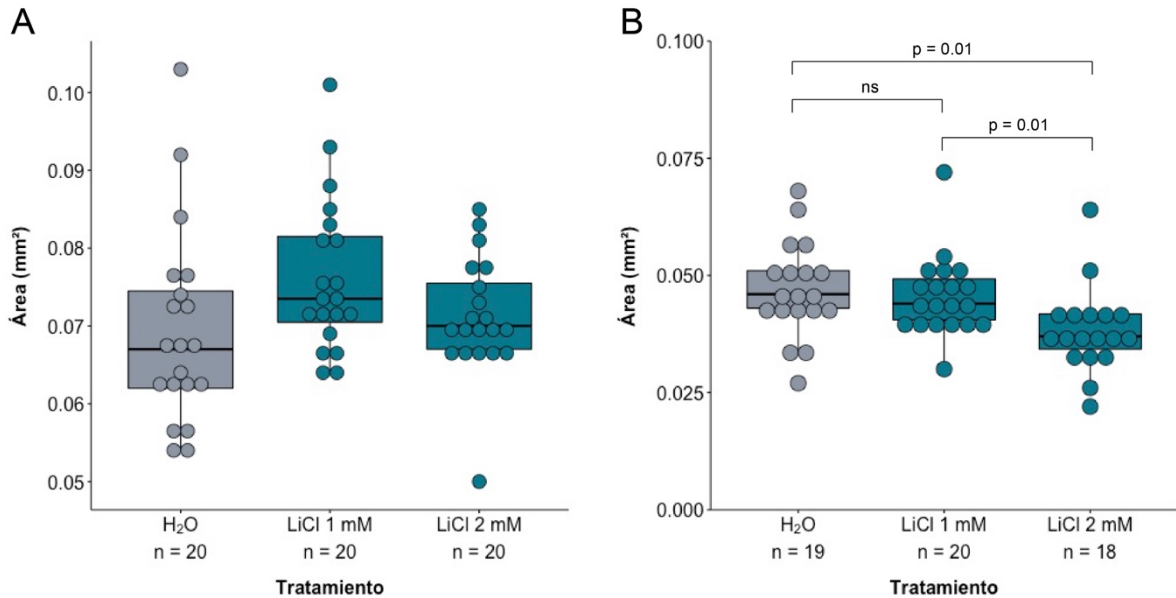


Figura 10. Áreas de los blastemas tratados con dos concentraciones de LiCl de las 48 a las 72 hpa. En la regeneración anterior (A) no se observaron diferencias en los tamaños de los blastemas tratados con LiCl y agua. En cambio, en la regeneración posterior (B), los blastemas tratados con LiCl 2mM presentaron menor tamaño que aquellos tratados con agua. Los valores de p corresponden a la prueba *post hoc* de Wilcoxon; ns, no significativo.

3.5.5. Efecto de la inhibición y sobreactivación farmacológica de la vía canónica de Wnt en la formación del SNC durante la regeneración

Para demostrar que la inmunohistoquímica de fluorescencia, bajo las condiciones descritas, efectivamente funciona para marcar neuronas y sus procesos, primero se probó esta técnica en cabezas de organismos en estado no regenerativo. En estos organismos se observaron neuronas con marcaje fluorescente, principalmente en el prostomio (Figura 11A), lo cual sugiere que la técnica también podría funcionar para marcar el CNV en formación dentro del blastema.

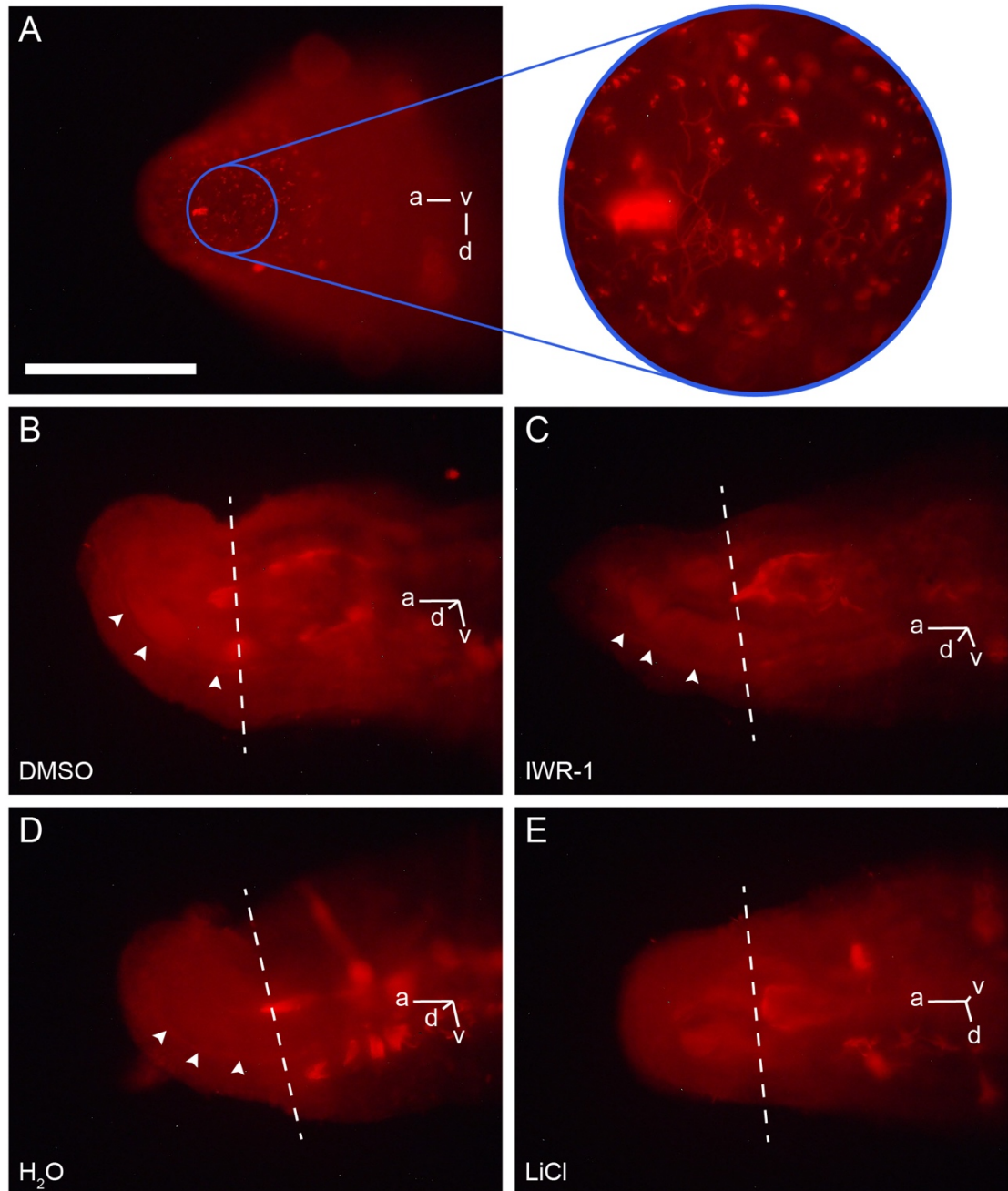


Figura 11. Regeneración del CNV bajo la inhibición y sobreactivación de la vía canónica de Wnt. Marcaje fluorescente con Tubulina acetilada con rodamina. En la cabeza (A), el prostomio cuenta con una abundante población de neuronas con múltiples dendritas. A la derecha se muestra un aumento de la región circulada en azul. En los blastemas de los fragmentos incubados en DMSO (A), IWR-1 (B) y agua (C) puede apreciarse el CNV en formación (cabezas de flecha); pero no en los fragmentos incubados en LiCl (E). Las líneas punteadas indican el sitio de corte. Los esquemas en las fotografías indican la orientación de los fragmentos: a, anterior; d, derecha; v, ventral. Barra de escala = 0.2 mm (misma para todos los recuadros).

En los fragmentos de los organismos experimentales se observó señal fluorescente en algunos tejidos no amputados; específicamente, la autofluorescencia de las quetas y el marcaje de unas estructuras ventrolaterales que podrían ser los metanefridios. A su vez, dentro del blastema se observó el marcaje tenue de una línea ventromedial que va desde el sitio de amputación hasta el extremo más distal del blastema (Figura 11D). La posición y el tamaño de esta línea sugieren que se trata del CNV en formación. Esta marca fluorescente se detectó a las 72 hpa en los blastemas de los fragmentos incubados en DMSO, IWR-1 y agua; pero no en los blastemas de los fragmentos incubados en LiCl (Figura 11B-E).

3.6. Discusión

La expresión de genes de la vía canónica de Wnt en tejidos en estado regenerativo ha sido observada en distintas especies de animales y la actividad de la vía se ha asociado a diferentes procesos (e.g., Srivastava *et al.*, 2014; Wischin *et al.*, 2017). Mediante hibridaciones *in situ*, en este trabajo se detectó la expresión de β -catenin en los blastemas y otros tejidos en proceso de diferenciación de *Lu. variegatus*, tanto en la sección anterior como en la posterior. En los tejidos anteriores en estado regenerativo, la expresión regionalizada de este gen pudo observarse en tres momentos principales: i) a las 60 hpa en todo el blastema; ii) entre las 84 y 96 hpa en la región ventral del tejido en estado regenerativo; y iii) entre las 132 y las 288 hpa en el prostomio. En *Pr. leidy* se ha detectado la expresión de β -catenin en el blastema anterior a las 24 y 72 hpa; pero no se le ha buscado en otros tiempos post amputación (Nyberg *et al.*, 2012). Asimismo, en *Ae. viride* se ha observado su expresión en el blastema anterior a las 24, 48 y 72 hpa; y tampoco se le ha buscado más allá de las 120 hpa (Chen *et al.*, 2020). Si bien la expresión de β -catenin en *Lu. variegatus* comienza después de transcurridas más hpa, debe considerarse que la regeneración en esta especie tarda alrededor de dos semanas en completarse, mientras que en *Pr. leidy* y *Ae. viride* esta tarda aproximadamente 5 días (Chen *et al.*, 2020; Zattar & Bely, 2011). Esta aparente relación entre el momento en que

comienza a expresarse β -catenin y la duración de la regeneración cobra sentido si se considera que, en cualquiera de las especies mencionadas, la expresión del gen coincide con el periodo de formación del blastema. Más aun, en *Ae. viride*, se ha observado que β -catenina se encuentra en los núcleos celulares del tejido proliferativo (Chen *et al.*, 2020). Estas observaciones en conjunto sugieren que la expresión de β -catenin a las 60 hpa podría estar regulando la formación del blastema, quizá promoviendo la proliferación celular. Por otro lado, su expresión en la región ventral del tejido en estado regenerativo (84-96 hpa) podría estar relacionada con la diferenciación de tejidos ventrales; pero la actividad morfogenética que tiene la vía canónica de Wnt en otros organismos también sugiere que podría estar asociada con el establecimiento de ejes (*e.g.*, Gurley *et al.*, 2008; Srivastava *et al.*, 2014), como el eje dorsoventral en alguno de los tejidos de la región ventral. La expresión observada en el prostomio (132-288 hpa) también podría relacionarse con un papel de la vía distinto a la regulación de la proliferación celular, ya que esta expresión tiene lugar en un periodo de tiempo en que la formación de esta estructura ya está prácticamente terminada. El prostomio es un órgano sensorial que contiene una gran cantidad de terminaciones nerviosas (Figura 11A), por lo que la expresión de β -catenin podría estar relacionada con el mantenimiento de esta estructura.

La dinámica de la expresión de β -catenin en el tejido en estado regenerativo posterior fue distinta. Esta comenzó a las 36 hpa en la región ventral del blastema, se extendió dorsalmente hasta cubrir todo el tejido en estado regenerativo a las 96 hpa y, después, se localizó cercana a la zona posterior de crecimiento hasta las 276 hpa. La expresión de β -catenin en el blastema posterior en otros anélidos solo ha sido observada en *Pr. leidyi* a las 48 hpa y no se ha buscado en otros periodos de regeneración (Nyberg *et al.*, 2012). Como en el caso de la regeneración anterior, dada la expresión ventral de este gen en *Lu. variegatus*, no puede descartarse que la vía canónica de Wnt esté involucrada en el establecimiento del eje dorsoventral

del blastema o de algunos tejidos específicos. También, la expresión cercana a la zona posterior de crecimiento sugiere que la vía podría estar promoviendo la proliferación celular durante y después de la regeneración. Sin embargo, todas estas posibilidades requieren de evidencia experimental.

Como se mencionó anteriormente, en el análisis de expresión diferencial del Capítulo 1 no se encontraron diferencialmente expresados ninguno de los genes que componen la vía canónica de Wnt. El patrón de expresión espaciotemporal de *β-catenin* sugiere que esto pudo deberse a que las muestras de ARN se tomaron en PR en los que aparentemente no había actividad significativa de la vía; con excepción de las muestras de regeneración posterior tomadas a las 72 hpa. Las diferencias entre ambos experimentos dan cuenta del dinamismo que tiene la actividad transcripcional de los genes durante la regeneración.

Una vez que se determinó el patrón de expresión espaciotemporal de *β-catenin*, este se tomó como un indicador de la actividad de la vía canónica de Wnt. Con la finalidad de determinar si esta vía juega algún papel en la formación del blastema y del SNC, en este trabajo se llevaron a cabo ensayos para inhibirla y sobreactivarla alrededor de las 60 hpa, punto en el que se comenzó a detectar la expresión generalizada de *β-catenin* en los blastemas anterior y posterior. La aplicación de IWR-1 para inhibir la actividad de la vía tuvo como consecuencia la formación de blastemas más pequeños en la regeneración anterior, tanto cuando se aplicó de las 48 a las 72 hpa como cuando se aplicó de las 48 a las 96 hpa. En la regeneración posterior se observó un efecto similar cuando el fármaco se aplicó de las 48 a las 72 hpa, pero no se observó un cambio en el tamaño de los blastemas cuando se aplicó de las 48 a las 96 hpa. Partiendo del patrón de expresión de *β-catenin* observado en las hibridaciones *in situ*, pareciera que la actividad de la vía canónica de Wnt durante la formación del blastema anterior se encuentra restringida a un periodo de tiempo acotado alrededor de las 60 hpa, la cual decae y se regionaliza a las 72 hpa; por lo que la aplicación de IWR-1 en ambos periodos (48-72 y 48-96

hpa) tuvo un efecto similar al inhibir la vía alrededor de las 60 hpa por igual. En cambio, las hibridaciones *in situ* en la regeneración posterior mostraron que la vía canónica de Wnt se mantiene activa después de las 60 hpa y aumenta paulatinamente su área de actividad hasta cubrir todo el blastema a las 108 hpa. De esta manera, es posible que la aplicación de IWR-1 de las 48 a las 72 hpa haya sido suficiente para inhibir la vía y reducir el tamaño del blastema posterior; pero probablemente la actividad de la vía es tan alta después de las 72 hpa que fue capaz de contrarrestar el efecto inhibitorio del fármaco y compensar su inhibición inicial. En cuanto a la formación del SNC, la inhibición de la vía canónica de Wnt no mostró tener un efecto visible sobre la formación del CNV en los fragmentos en estado regenerativo incubados en IWR-1. En todos los organismos observados pudo apreciarse el CNV en formación siguiendo el mismo patrón que en los organismos incubados en DMSO.

Por su parte, el sumergimiento de fragmentos en estado regenerativo en β -catenin-dc no mostró tener efecto sobre el tamaño de los blastemas. Wang *et al.* (2015) demostraron la efectividad del silenciamiento de la expresión con esta técnica al sumergir células cultivadas por 48 y 72 h en ARNdc de un peso similar y con la misma concentración que se utilizó en este trabajo. Es posible que en este trabajo no se haya observado un efecto similar debido a que, a diferencia de las células cultivadas, la estructura tridimensional de los blastemas pudo haber dificultado la difusión del ARNdc hacia las células del interior. En este sentido, es recomendable intentar la inyección de ARNdc directamente en el blastema para determinar la efectividad de β -catenin-dc en la inhibición de la vía.

Por otro lado, la sobreactivación de la vía canónica de Wnt con LiCl 2 mM no tuvo efecto sobre el tamaño de los blastema anteriores, pero tuvo como consecuencia la inhibición de la formación del CNV. En cambio, la sobreactivación de la vía en la regeneración posterior provocó la reducción del tamaño de los blastemas. Esto significa que la actividad de la vía debe estar finamente regulada en el blastema

anterior para permitir que dos procesos regenerativos distintos sucedan al mismo tiempo; debe ser lo suficientemente alta para permitir la formación del blastema, pero lo suficientemente baja para permitir la formación del CNV. En la regeneración posterior, la regulación de la vía también debe estar regulada, ya que su inhibición o sobreactivación interfieren con la formación del blastema.

Chen *et al.* (2020) y Fok *et al.* (2020) han demostrado que la inhibición farmacológica de la vía canónica de Wnt por medio de la administración de XAV939 durante la formación del blastema anterior en *Ae. viride*, tiene como consecuencia la formación de blastemas más pequeños debido a una disminución de la tasa mitótica y un aumento de la apoptosis. De manera similar, la sobreactivación de la vía por medio de la administración de Azkp también disminuyó el tamaño del blastema como consecuencia de la disminución de la tasa mitótica (Chen *et al.*, 2020). Además, los organismos tratados con XAV939 durante la formación del blastema mostraron un SNC parcialmente diferenciado a los 6 días postamputación; pero aquellos tratados con Azkp fueron incapaces de formar el CNV, el neuropilo y las comisuras circumesofágicas. En conjunto, estos experimentos sugieren que la regulación de la actividad de la vía canónica de Wnt durante la formación del blastema anterior permite un balance entre la proliferación, la apoptosis y la diferenciación celular. Su inhibición disminuye la proliferación celular y aumenta la apoptosis, mientras que su sobreactivación también disminuye la proliferación celular e impide la diferenciación de algunas estructuras.

Asumiendo que, como en *Ae. viride*, el tamaño de los blastemas y la presencia del CNV son indicadores de la proliferación celular y de la diferenciación del SNC, respectivamente; con los resultados aquí presentados se propone que la actividad de la vía canónica de Wnt debe estar finamente regulada y tiene, al menos, dos funciones tejido específicas durante la regeneración temprana en *Lu. variegatus*. Se requiere la activación de la vía para inducir la proliferación celular y la formación del blastema, pero su regulación es necesaria para permitir la diferenciación del SNC.

Esta propuesta debe confirmarse por medio de marcadores de proliferación celular y sería interesante conocer qué papel juega la vía en la formación del SNC durante la regeneración posterior. Asimismo, el patrón de expresión de *β-catenin* después de las 72 hpa sugiere que la vía canónica de Wnt tiene otras funciones durante la regeneración que requieren ser investigadas.

4. Discusión y conclusiones generales

Los anélidos son uno de los grupos de animales con mayor capacidad regenerativa. Algunas especies como *Lu. variegatus* tienen la capacidad de regenerar las secciones anterior y posterior del cuerpo al ser amputadas, incluso al mismo tiempo; y gracias a esta capacidad pueden reproducirse asexualmente por fisión. Sin embargo, los genes que regulan el proceso de regeneración han comenzado a ser estudiados apenas recientemente y se cuenta con información de solo unas cuantas especies (Özpolat & Bely, 2016). Por tal motivo, en este trabajo se determinó el perfil de expresión génica y el enriquecimiento de ontologías genéticas durante la formación del blastema en *Lu. variegatus*. También se investigó la participación de la vía canónica de Wnt, una de las vías de señalización más importantes en la regeneración en anélidos y otros organismos. Los resultados obtenidos permitieron hacer comparaciones con lo que se ha reportado en otras especies de anélidos.

El análisis de expresión diferencial mostró la sobreexpresión de algunos genes que se han asociado a la regeneración o a condiciones de alta proliferación celular en otros organismos: *AIRC*, *BHMT*, *bmi1*, *dio1*, *dio3*, *fbl*, *Hsp60* y *vdr*. Pero también se observó la subexpresión de genes cuya actividad ha sido reportada durante la regeneración: *Gs* y *VWDE*. De estos genes, solo *Gs* ha sido investigado en cuanto al patrón de expresión espaciotemporal que sigue durante la regeneración en *En. japonensis* (Niva *et al.*, 2008); pero no se han llevado a cabo estudios funcionales durante la regeneración en otras especies. Así, estos resultados deben tomarse como un punto de partida para estudiar de manera dirigida el papel específico que juegan estos genes en la formación del blastema.

Por otro lado, la regeneración es un rasgo ancestral en los anélidos y se ha sugerido la existencia de mecanismos moleculares comunes que regulen este proceso en todo el grupo (Özpolat & Bely, 2016; Zattara & Bely, 2016). Al comparar el perfil de expresión génica de *Lu. variegatus* con el de otras especies se encontraron pocas

coincidencias, apenas cuatro genes diferencialmente expresado se comportan de la misma forma en otras especies: *AIRC*, *BHMT*, *fbl* y *GLUD* (Myohara *et al.*, 2006; Ribeiro *et al.*, 2019). Este resultado no es sorprendente, Ribeiro *et al.* (2019) también encontraron pocas coincidencias al comparar los genes candidatos de la regeneración diferencialmente expresados en *Sp. hystrix* y *Sy. gracilis* (1 de 14). Lo anterior sugiere que los mecanismos moleculares que rigen la regeneración son menos comunes de lo que se pensaba; pero esto requiere confirmarse con un muestreo más exhaustivo en el *phylum* o mediante otras aproximaciones como se hizo inicialmente con los genes candidatos de la regeneración.

Los análisis de enriquecimiento de ontologías genéticas mostraron que el metabolismo de ARN y aminoácidos se encuentra activo durante la regeneración. Esto puede deberse a la necesidad de sintetizar un gran número de proteínas involucradas en procesos relacionados con la regeneración, como son la mitosis, la remodelación de la matriz extracelular o la diferenciación de tejidos, entre otros. (Bely *et al.*, 2014). Adicionalmente, los análisis también mostraron que varios genes relacionados con el metabolismo energético se encontraron sobreexpresados. Al tratarse de genes relacionados con el transporte y degradación del glutamato y otros compuestos del ciclo de Krebs, en este trabajo se propone que la actividad de este ciclo podría estar contribuyendo para satisfacer la alta demanda energética que tienen las células en proliferación (Vincent *et al.*, 2015). En otros organismos se ha observado que el enriquecimiento del alimento o del medio en el que se desarrollan con compuestos esenciales para el metabolismo energético (*e.g.*, sacarosa, insulina y leucina) favorece la regeneración epimórfica, incluso en organismos que típicamente no tienen la capacidad de regenerar estructuras corporales en el estadio adulto (Abrams *et al.*, 2021). Tomando en cuenta el conjunto de estas observaciones, resulta pertinente llevar a cabo estudios enfocados en determinar qué papel juega el metabolismo energético durante la regeneración en *L. variegatus*. Un estudio metabolómico podría arrojar más datos que ayuden a comprender de

qué manera los organismos en estado regenerativo satisfacen sus necesidades energéticas y materiales.

En este trabajo también se investigó la participación de la vía canónica de Wnt durante la regeneración. Si bien el perfil de expresión génica de los organismos en estado regenerativo no mostró genes de esta vía diferencialmente expresados, la sobreexpresión de *dio1*, *dio3* y *vdr* se tomó como un indicativo de su posible actividad. Además, la señalización de Wnt y la actividad de la vía en los blastemas ha sido constatada en varios organismos, incluyendo algunos anélidos (Chen *et al.*, 2020; Nyberg *et al.*, 2012; Planques *et al.*, 2019). Por estos motivos, se decidió investigar el patrón de expresión espaciotemporal de uno de sus genes efectores y el efecto de la alteración de la vía durante la formación del blastema. Las hibridaciones *in situ* mostraron que *β -catenin* se expresa en los blastemas anteriores y posteriores durante la formación y diferenciación de estas estructuras. Además, su expresión es dinámica, cambiando su intensidad a lo largo del tiempo y asociándose con diferentes estructuras de los tejidos en estado regenerativo. A las 60 hpa se observó que *β -catenin* comienza a expresarse tanto en los blastemas anteriores como en los posteriores, por lo que se llevaron a cabo experimentos que dieran cuenta de su participación en la formación del blastema y del SNC alrededor de este PR. La inhibición farmacológica de la vía canónica de Wnt produjo la reducción del tamaño de los blastemas; pero no afectó la formación del CNV. En cambio, su sobreactivación tuvo como consecuencia la disminución del tamaño del blastema posterior e impidió la formación del CNV. Estos resultados son muy semejantes a los reportados por Chen *et al.* (2020) en *Ae. viride* y sugieren que la vía canónica de Wnt es necesaria para que las células del blastema proliferen y mantengan una población celular adecuada; pero su sobreactivación impide la diferenciación del SNC. En este sentido, la modulación de la vía parece crucial para una regeneración adecuada y, por lo tanto, sería interesante determinar qué mecanismos moleculares regulan su actividad. Además, puesto que la expresión de *β -catenin* continua después de las 60 hpa cuando otros tejidos ya se han

diferenciado, es posible que también participe en otros procesos distintos a la proliferación y diferenciación celular.

Por último, como parte de la metodología necesaria para llevar a cabo los análisis aquí descritos, fue necesario ensamblar *de novo* un transcriptoma con secuencias expresadas en el estado no regenerativo en *Lu. variegatus* y secuencias expresadas a las 24, 48 y 72 hpa. Este es el primer transcriptoma de esta especie anotado con datos de homologías de secuencias, ontologías genéticas y del potencial codificante de los transcritos; por lo que se espera que sea una herramienta útil para futuros estudios relacionados con la regeneración y otros temas en *Lu. variegatus* y otros anélidos.

En conclusión, en este estudio se encontraron 136 transcritos diferencialmente expresados durante la regeneración en *Lu. variegatus*, los cuales muestran poca similitud con los transcritos diferencialmente expresados en otras especies de anélidos. Entre estos transcritos, se encuentran algunos relacionados con la regeneración, la proliferación y la diferenciación celular en otras especies. También, se encontró que el metabolismo de ARN, de aminoácidos y de otros compuestos relacionados con la homeostasis energética se encuentran activos durante la regeneración. Finalmente, se determinó que la vía canónica de Wnt es necesaria para la formación del blastema y que su actividad debe estar finamente modulada para permitir la diferenciación del SNC.

5. Referencias

- Aberle, H., Bauer, A., Stappert, J., Kispert, A., Kemler, R. (1997). β -catenin is a target for the ubiquitin-proteasome pathway. *The EMBO Journal*, 16, 3797-3804. <https://doi.org/10.1093/emboj/16.13.3797>
- Abrams, M. J., Tan, F. H., Li, Y., Basinger, T., Heithe, M. L., Sarma, A., Lee, I. T., Condiotte, Z. J., Raffie, M., Dabiri, J. O., Gold, D. A., & Goentoro, L. (2021). A conserved strategy for inducing appendage regeneration in moon jellyfish, *Drosophila*, and mice. *eLife*, 10, e65092. <https://doi.org/10.7554/eLife.65092>
- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, 25(17), 3389-3402. <https://doi.org/10.1093/nar/25.17.3389>
- Aguilera, O., Peña, C., García, J. M., Larriba, M. J., Ordóñez-Morán, P., Navarro, D., Barbáchano, A., López de silanes, I., Ballestar, E., Fraga, M. F., Esteller, M., Gamallo, C., Bonilla, F., González-Sancho, J. M., & Muñoz, A. (2007). The Wnt antagonist *DICKKOPF-1* gene is induced by $1\alpha,25$ -dihydroxyvitamin D_3 associated to the differentiation of human colon cáncer cells. *Carcinogenesis*, 28(9), 1877-1884. <https://doi.org/10.1093/carcin/bgm094>
- Amit, S., Hatzubai, A., Birman, Y., Andersen, A. S., Ben-Shushan, E., Mann, M., Ben-Neriah, Y., & Alkalay, I. (2002). Axin-mediated CKI phosphorylation of β -catenin at Ser 45: a molecular switch for the Wnt pathway. *Genes & Development*, 16(9), 1066-1076. <https://doi.org/10.1101/gad.230302>
- Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>
- Bastakoty, D., & Young, P. P. (2016). Wnt/ β -catenin pathway in tissue injury: roles in pathology and therapeutic opportunities for regeneration. *The FASEB Journal*, 30(10), 3271-3284. <https://doi.org/10.1096/fj.201600502R>

- Bely, A. E. (2014). Early Events in Annelid Regeneration: A Cellular Perspective. *Integrative and Comparative Biology*, 54(4), 688–699. <https://doi.org/10.1093/icb/icu109>
- Bely, A. E., & Nyberg, K. G. (2010). Evolution of animal regeneration: re-emergence of a field. *Trends in Ecology & Evolution*, 25(3), 161–170. <https://doi.org/10.1016/j.tree.2009.08.005>
- Bely, A. E., & Sikes, J. M. (2010). Latent regeneration abilities persist following recent evolutionary loss in asexual annelids. *Proceedings of the National Academy of Sciences*, 107(4), 1464–1469. <https://doi.org/10.1073/pnas.0907931107>
- Bhambri, A., Dhaunta, N., Patel, S. S., Hardikar, M., Bhatt, A., Srikakulam, N., Shridhar, S., Vellarikkal, S., Pandey, R., Jayarajan, R., Verma, A., Kumar, V., Gautam, P., Khanna, Y., Khan, J. A., Fromm, B., Peterson, K. J., Scaria, V., Sivasubbu, S., & Pillai, B. (2018). Large scale changes in the transcriptome of *Eisenia fetida* during regeneration. *PLoS ONE*, 13(9), e0204234. <https://doi.org/10.1371/journal.pone.0204234>
- Bilic, J., Huang, Y., Davidson, G., Zimmermann, T., Cruciat, C., Bienz, M., & Niehrs, C. (2007). Wnt induces LRP6 signalosome and promotes Dishevelled-Dependent LRP6 phosphorylation. *Science*, 316(5831), 1619–1622. <https://doi.org/10.1126/science.1137065>
- Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., Fridman, W.-H., Pagès, F., Trajanoski, Z., & Galon, J. (2009). ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics*, 25(8), 1091–1093. <https://doi.org/10.1093/bioinformatics/btp101>
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>

- Brantjes, H., Roose, J., van de Wetering, M., & Clevers, H. (2001). All Tcf HMG box transcription factors interact with Groucho-related co-repressors. *Nucleic Acid Research*, 29(7), 1410-1419. <https://doi.org/10.1093/nar/29.7.1410>
- Bray, N. L., Pimentel, H., Melsted, P., & Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology*, 34(5), 525–527. <https://doi.org/10.1038/nbt.3519>
- Bryant, D. M., Johnson, K., DiTommaso, T., Tickle, T., Couger, M. B., Payzin-Dogru, D., Lee, T. J., Leigh, N. D., Kuo, T.-H., Davis, F. G., Bateman, J., Bryant, S., Guzikowski, A. R., Tsai, S. L., Coyne, S., Ye, W. W., Freeman Jr, R. M., Peshkin, L., Tabin, C. J., Regev, A., Haas, B. J., Whited, J. L. (2017). A Tissue-Mapped Axolotl De Novo Transcriptome Enables Identification of Limb Regeneration Factors. *Cell Reports*, 18(3), 762–776. <https://doi.org/10.1016/j.celrep.2016.12.063>
- Camargo, A. P., Sourkov, V., Pereira, G. A. G., & Carazzolle, M. F. (2020). RNAsamba: neural network-based assessment of the protein-coding potential of RNA sequences. *NAR Genomics and Bioinformatics*, 2(1), lqz024. <https://doi.org/10.1093/nargab/lqz024>
- Castegna, A., & Menga, A. (2018). Glutamine Synthetase: Localization Dictates Outcome. *Genes*, 9(2), 108. <https://doi.org/10.3390/genes9020108>
- Catalano, V., Dentice, M., Ambrosio, R., Luongo, C., Carollo, R., Benfante, A., Todaro, M., Stassi, G., & Salvatore, D. (2016). Activated Thyroid Hormone Promotes Differentiation and Chemotherapeutic Sensitization of Colorectal Cancer Stem Cells by Regulating Wnt and BMP-4 Signaling. *Cancer Research*, 76(5), 1237-1244. <https://doi.org/10.1158/0008-5472.CAN-15-1542>
- Chen, B., Dodge, M. E., Tang, W., Lu, J., Ma, Z., Fan, C.-W., Wei, S., Hao, W., Kilgore, J., Williams, N. S., Roth, M. G., Amatruda, J. F., Chen, C., & Lum, L. (2009). Small molecule-mediated disruption of Wnt-dependent signaling in tissue regeneration and cancer. *Nature Chemical Biology*, 5, 100-107. <https://doi.org/10.1038/nchembio.137>

- Chen, C., Yueh, W., & Chen, J. (2020) Canonical Wnt Signaling is Involved in Anterior Regeneration of the Annelid *Aelosoma viride*. bioRxiv. <https://doi.org/10.1101/2020.03.01.972448>
- Cho, S.-J., Lee, M. S., Tak, E. S., Lee, E., Koh, K. S., Ahn, C. H., & Park, S. C. (2009). Gene Expression Profile in the Anterior Regeneration of the Earthworm Using Expressed Sequence Tags. *Bioscience, Biotechnology, and Biochemistry*, 73(1), 29–34. <https://doi.org/10.1271/bbb.80391>
- Cho, S.-J., Vellès, Y., Giani Jr, V. C., Seaver, E. C., & Weisblat, D. (2010). Evolutionary Dynamics of the *wnt* Gene Family: A Lophotrochozoan Perspective. *Molecular Biology and Evolution*, 27(2), 1645-1658. <https://doi.org/10.1093/molbev/msq052>
- Clément-Lacroix, P., Ai, M., Morvan, F., Roman-Roman, S., Vayssière, B., Belleville, C., Estrera, K., Warman, M. L., Baron, R. & Rawadi, G. (2005). Lrp5-independent activation of Wnt signaling by lithium chloride increases bone formation and bone mass in mice. *Proceedings of the National Academy of Sciences*, 102(48), 17406-17411. <https://doi.org/10.1073/pnas.0505259102>
- Cook, D. G. (1969). Observations on the life history and ecology of some Lumbriculidae (Annelida, Oligochaeta). *Hydrobiologia*, 34(3–4), 561–574. <https://doi.org/10.1007/BF00045410>
- de Jong, D. M., & Seaver, E. C. (2016). A Stable Thoracic *Hox* Code and Epimorphosis Characterize Posterior Regeneration in *Capitella teleta*. *PLoS ONE*, 11(2), e0149724. <https://doi.org/10.1371/journal.pone.0149724>
- de la Fuente, A. G., Errea, O., van Wijngaarden, P., Gonzalez, G. A., Kerninon, C., Jarjour, A. A., Lewis, H. J., Jones, C. A., Nait-Oumesmar, B., Zhao, C., Huang, J. K., French-Constant, C., & Franklin, R. J. M. (2015). Vitamin D receptor–retinoid X receptor heterodimer signaling regulates oligodendrocyte progenitor cell differentiation. *Journal of Cell Biology*, 211(5), 975–985. <https://doi.org/10.1083/jcb.201505119>
- Dentice, M., Luongo, C., Ambrosio, R., Sibilio, A., Casillo, A., Iaccarino, A., Troncone, G., Fenzi, G., Larsen, P. R., & Salvatore, D. (2012). β -Catenin

- Regulates Deiodinase Levels and Thyroid Hormone Signaling in Colon Cancer Cells. *Gastroenterology*, 143, 1037-1047. <https://doi.org/10.1053/j.gastro.2012.06.042>
- Drewes, C. D., & Fournier, C. R. (1990). Morphallaxis in an aquatic oligochaete, *Lumbriculus variegatus*: Reorganization of escape reflexes in regenerating body fragments. *Developmental Biology*, 138(1), 94–103. [https://doi.org/10.1016/0012-1606\(90\)90179-M](https://doi.org/10.1016/0012-1606(90)90179-M)
- Fok, S. K.-W., Chen, C.-P., Tseng, T.-L., Chiang, Y.-H., & Chen, J.-H. (2020). Caspase dependent apoptosis is required for anterior regeneration in *Aelosoma viride* and its related gene expressions are regulated by the Wnt signaling pathway. *Scientific Reports*, 10, 10692. <https://doi.org/10.1038/s41598-020-64008-1>
- Fu, L., Niu, B., Zhu, Z., Wu, S., & Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics*, 28(23), 3150–3152. <https://doi.org/10.1093/bioinformatics/bts565>
- Fukuda, A., Morris, J. P., & Hebrok, M. (2012). Bmi1 Is Required for Regeneration of the Exocrine Pancreas in Mice. *Gastroenterology*, 143(3), 821-831.e2. <https://doi.org/10.1053/j.gastro.2012.05.009>
- Gurley, K. A., Rink, J. C., & Sánchez Alvarado, A. (2008). β -catenin defines head versus tail identity during planarian regeneration and homeostasis. *Science*, 319(5861), 323-327. <https://doi.org/10.1126/science.1150029>
- Gustafsson, D. R., Price, D. A., & Erséus, C. (2009). Genetic variation in the popular lab worm *Lumbriculus variegatus* (Annelida: Clitellata: Lumbriculidae) reveals cryptic speciation. *Molecular Phylogenetics and Evolution*, 51(2), 182–189. <https://doi.org/10.1016/j.ympev.2008.12.016>
- Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., Couger, M. B., Eccles, D., Li, B., Lieber, M., MacManes, M. D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C. N., Henschel, R., LeDuc, R. D., Friedman, N., & Regev, A. (2013). *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform

- for reference generation and analysis. *Nature Protocols*, 8(8), 1494–1512.
<https://doi.org/10.1038/nprot.2013.084>
- Han, Y., Chen, A., Umansky, K.-B., Oonk, K. A., Choi, W.-Y., Dickson, A. L., Ou, J., Cigliola, V., Yifa, O., Cao, J., Tornini, V. A., Cox, B. D., Tzahor, E., & Poss, K. D. (2019). Vitamin D Stimulates Cardiomyocyte Proliferation and Controls Organ Size and Regeneration in Zebrafish. *Developmental Cell*, 48(6), 853–863.e5. <https://doi.org/10.1016/j.devcel.2019.01.001>
- Hanson, A. J., Wallace, H. A., Freeman, T. J., Beauchamp, R. D., Lee, L. A., & Lee, E. (2012). XIAP Monoubiquitylates Groucho/TLE to Promote Canonical Wnt Signaling. *Molecular Cell*, 45(5), 619–628. <https://doi.org/10.1016/j.molcel.2011.12.032>
- Heikkinen, L. K., Kesäniemi, J. E., & Knott, K. E. (2017). De novo transcriptome assembly and developmental mode specific gene expression of *Pygospio elegans*. *Evolution & Development*, 19(4–5), 205–217. <https://doi.org/10.1111/ede.12230>
- Hlaing, S. M., Garcia, L. A., Contreras, J. R., Norris, K. C., Ferrini, M. G., & Artaza, J. N. (2014). 1,25-Vitamin D₃ promotes cardiac differentiation through modulation of the WNT signaling pathway. *Journal of Molecular Endocrinology*, 53, 303–317. <https://doi.org/10.1530/JME-14-0168>
- Huang, D. W., Sherman, B. T., Tan, Q., Collins, J. R., Alvord, W. G., Roayaei, J., Stephens, R., Basesler, M. W., Lane, H. C., & Lempicki, R. A. (2007). The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. *Genome Biology*, 8, R183. <https://doi.org/10.1186/gb-2007-8-9-r183>
- Ikeda, S., Kishida, S., Yamamoto, H., Murai, H., Koyama, S., & Kikuchi, A. (1998). Axin, a negative regulator of the Wnt signaling pathway, forms a complex with GSK-3 β and β -catenin and promotes GSK-3 β -dependent phosphorylation of β -catenin. *The EMBO Journal*, 17, 1371–1384. <https://doi.org/10.1093/emboj/17.5.1371>

- Kanai, Y., & Hediger, M. A. (1992). Primary structure and functional characterization of a high-affinity glutamate transporter. *Nature*, 360(6403), 467-471. <https://doi.org/10.1038/360467a0>
- Kawakami, Y., Rodriguez Esteban, C., Raya, M., Kawakami, H., Marti, M., Dubova, I., & Izpisua Belmonte, J. C. (2006). Wnt/ β -catenin signaling regulates vertebrate limb regeneration. *Genes & Development*, 20, 3232-3237. <https://doi.org/10.1101/gad.1475106>
- Kozin, V. V., & Kostyuchenko, R. P. (2015). *Vasa*, *PL10*, and *Piwi* gene expression during caudal regeneration of the polychaete annelid *Alitta virens*. *Development Genes and Evolution*, 225(3), 129–138. <https://doi.org/10.1007/s00427-015-0496-1>
- Kramps, T., Peter, O., Brunner, E., Nellen, B., Froesch, B., Chatterjee, S., Murone, M., Züllig, S., & Basler, K. (2002). Wnt/Wingless Signaling Requires BCL9/Legless-Mediated Recruitment of Pygopus to the Nuclear β -catenin-TCF Complex. *Cell*, 109(1), 47-60. [https://doi.org/10.1016/S0092-8674\(02\)00679-7](https://doi.org/10.1016/S0092-8674(02)00679-7)
- Kriventseva, E. V, Kuznetsov, D., Tegenfeldt, F., Manni, M., Dias, R., Simão, F. A., & Zdobnov, E. M. (2019). OrthoDB v10: sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs. *Nucleic Acids Research*, 47(D1), D807–D811. <https://doi.org/10.1093/nar/gky1053>
- Langfeld, T., Watanabe, H., Simakov, O., Lindgens, D., Gee, L., Law, L., Schmidt, H. A., Özbek, S., Bode, H., & Holstein, T. W. (2009). Multiple Wnts are involved in *Hydra* organizer formation and regeneration. *Developmental Biology*, 330(1), 186-199. <https://doi.org/10.1016/j.ydbio.2009.02.004>
- Leigh, N. D., Sessa, S., Dragalzew, A. C., Payzin-Dogru, D., Sousa, J. F., Aggouras, A. N., Johnson, K., Dunlap, G. S., Haas, B. J., Levin, M., Schneider, I., & Whited, J. L. (2020). von Willebrand factor D and EGF domains is an evolutionarily conserved and required feature of blastemas capable of

- multitissue appendage regeneration. *Evolution & Development*, 22, 297-311. <https://doi.org/10.1111/ede.12332>
- Lesiuk, N. M., & Drewes, C. D. (1999). Autotomy reflex in a freshwater oligochaete, *Lumbriculus variegatus* (Clitellata: Lumbriculidae). *Hydrobiologia*, 406(0), 253–261. <https://doi.org/10.1023/A:1003756722019>
- Li, L., Zhai, S., Wang, L., Si, S., Wu, H., & Chang, Z. (2014). *Hsp60* in caudal fin regeneration from *Paramisgurnus dabryanus*: Molecular cloning and expression characterization. *Fish & Shellfish Immunology*, 36(2), 401–408. <https://doi.org/10.1016/j.fsi.2013.12.016>
- Li, W., & Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*, 22(13), 1658–1659. <https://doi.org/10.1093/bioinformatics/btl158>
- Lin, G., & Slack, J. M. W. (2008). Requirement for Wnt and FGF signaling in *Xenopus* tail regeneration. *Developmental Biology*, 316(2), 323-335. <https://doi.org/10.1016/j.ydbio.2008.01.032>
- Lopez-Arribillaga, E., Rodilla, V., Pellegrinet, L., Guiu, J., Iglesias, M., Roman, A. C., Gutarra, S., González, S., Muñoz-Cánoves, P., Fernández-Salguero, P., Radtke, F., Bigas, A., & Espinosa, L. (2015). Bmi1 regulates murine intestinal stem cell proliferation and self-renewal downstream of Notch. *Development*, 142(1), 41–50. <https://doi.org/10.1242/dev.107714>
- Makino, S., Whitehead, G. G., Lien, C.-L., Kim, S., Jhawar, P., Kono, A., Kawata, Y., & Keating, M. T. (2005). Heat-shock protein 60 is required for blastema formation and maintenance during regeneration. *Proceedings of the National Academy of Sciences*, 102(41), 14599–14604. <https://doi.org/10.1073/pnas.0507408102>
- Marikawa, Y., & Elinson, R. P. (1998). beta-TrCP is a negative regulator of *Wnt/beta-catenin* signaling pathway and dorsal axis formation in *Xenopus* embryos. *Mechanisms of Development*, 77(1), 75-80. [https://doi.org/10.1016/S0925-4773\(98\)00134-8](https://doi.org/10.1016/S0925-4773(98)00134-8)

- Martinez-Acosta, V. G., & Zoran, M. J. (2015). Evolutionary Aspects of Annelid Regeneration. In eLS (pp. 1–7). Chichester, UK: *John Wiley & Sons, Ltd.* <https://doi.org/10.1002/9780470015902.a0022103.pub2>
- Matés, J. M., Campos-Sandoval, J. A., Santos-Jiménez, J. de los, & Márquez, J. (2019). Dysregulation of glutaminase and glutamine synthetase in cancer. *Cancer Letters*, 467, 29–39. <https://doi.org/10.1016/j.canlet.2019.09.011>
- Miller, L. D., Park, K. S., Guo, Q. M., Alkharouf, N. W., Malek, R. L., Lee, N. H., Liu, E. T., & Cheng, S.-Y. (2001). Silencing of Wnt Signaling and Activation of Multiple Metabolic Pathways in Response to Thyroid Hormone-Stimulated Cell Proliferation. *Molecular and Cellular Biology*, 21(19), 6626-6639. <https://doi.org/10.1128/MCB.21.19.6626-6639.2001>
- Muralidhar, S., Filia, A., Nsengimana, J., Pozniak, J., O'Shea, S. J., Diaz, J. M., Harland, M., Randerson-Moor, J. A., Reichrath, J., Laye, J. P., van der Weyden, L., Adams, D. J., Bishop, D. T., & Newton-Bishop, J. (2019). Vitamin D-VDR Signaling Inhibits Wnt/ β -catenin-Mediated Melanoma Progression and Promotes Antitumor Immunity. *Cancer Research*, 79(23), 5986-5998. <https://doi.org/10.1158/0008-5472.CAN-18-3927>
- Myohara, M., Niva, C. C., & Lee, J. M. (2006). Molecular approach to annelid regeneration: cDNA subtraction cloning reveals various novel genes that are upregulated during the large-scale regeneration of the oligochaete, *Enchytraeus japonensis*. *Developmental Dynamics*, 235(8), 2051–2070. <https://doi.org/10.1002/dvdy.20849>
- Niva, C. C., Lee, J. M., & Myohara, M. (2008). Glutamine synthetase gene expression during the regeneration of the annelid *Enchytraeus japonensis*. *Development Genes and Evolution*, 218(1), 39–46. <https://doi.org/10.1007/s00427-007-0198-4>
- Novikova, E. L., Bakalenko, N. I., Nesterenko, A. Y., & Kulakova, M. A. (2013). Expression of Hox genes during regeneration of nereid polychaete *Alitta (Nereis) virens* (Annelida, Lophotrochozoa). *EvoDevo*, 4(14). <https://doi.org/10.1186/2041-9139-4-14>

- Nyberg, K. G., Conte, M. A., Kostyun, J. L., Forde, A., & Bely, A. E. (2012). Transcriptome characterization via 454 pyrosequencing of the annelid *Pristina leidy*, an emerging model for studying the evolution of regeneration. *BMC Genomics*, 13(1), 287. <https://doi.org/10.1186/1471-2164-13-287>
- Özpolat, B. D., & Bely, A. E. (2015). Gonad establishment during asexual reproduction in the annelid *Pristina leidy*. *Developmental Biology*, 405(1), 123–136. <https://doi.org/10.1016/j.ydbio.2015.06.001>
- Özpolat, B. D., & Bely, A. E. (2016). Developmental and molecular biology of annelid regeneration: a comparative review of recent studies. *Current Opinion in Genetics & Development*, 40, 144–153. <https://doi.org/10.1016/j.gde.2016.07.010>
- Pálmer, H. G., Anjos-Afonso, F., Carmeliet, G., Takeda, H., & Watt, F. M. (2008) The Vitamin D Receptor Is a Wnt Effector that Controls Hair Follicle Differentiation and Specifies Tumor Type in Adult Epidermis. *PLoS ONE*, 1, e1483. <https://doi.org/10.1371/journal.pone.0001483>
- Pálmer, H. G., González-Sancho, J. M., Espada, J., Berciano, M. T., Puig, I., Baulida, J., Quintanilla, M., Cano, A., García de Herreros, A., Lafarga, M., & Muñoz, A. (2001). Vitamin D₃ promotes the differentiation of colon carcinoma cells by the induction of E-cadherin and the inhibition of β -catenin signaling. *The Journal of Cell Biology*, 154(2), 369-387. <https://doi.org/10.1083/jcb.200102028>
- Pei, W., Tanaka, K., Huang, S. C., Xu, L., Liu, B., Sinclair, J., Idol, J., Varshney, G. K., Huang, H., Lin, S., Nussenblatt, R. B., Mori, R., & Burgess, S. M. (2016). Extracellular HSP60 triggers tissue regeneration and wound healing by regulating inflammation and cell proliferation. *Npj Regenerative Medicine*, 1(1), 16013. <https://doi.org/10.1038/npjregenmed.2016.13>
- Pfeifer, K., Dorresteijn, A. W. C., & Fröbuis, A. C. (2012). Activation of *Hox* genes during caudal regeneration of the polychaete annelid *Platynereis dumerilii*. *Development Genes and Evolution*, 222(3), 165–179. <https://doi.org/10.1007/s00427-012-0402-z>

- Pfeifer, K., Schaub, C., Wolfstetter, G., & Dorresteyjn, A. (2013). Identification and characterization of a *twist* ortholog in the polychaete annelid *Platynereis dumerilii* reveals mesodermal expression of *Pdu-twist*. *Development Genes and Evolution*, 223(5), 319–328. <https://doi.org/10.1007/s00427-013-0448-6>
- Pimentel, H., Bray, N. L., Puente, S., Melsted, P., & Pachter, L. (2017). Differential analysis of RNA-seq incorporating quantification uncertainty. *Nature Methods*, 14(7), 687–690. <https://doi.org/10.1038/nmeth.4324>
- Piunti, A. & Shilatifard, A. (2021). The roles of Polycomb repressive complexes in mammalian development and cancer. *Nature Reviews Molecular Cell Biology*, 22, 326-345. <https://doi.org/10.1038/s41580-021-00341-1>
- Plaitakis, A., Kalef-Ezra, E., Kotzamani, D., Zaganas, I., & Spanaki, C. (2017). The Glutamate Dehydrogenase Pathway and Its Roles in Cell and Tissue Biology in Health and Disease. *Biology*, 6(4), 11. <https://doi.org/10.3390/biology6010011>
- Planques, A., Malem, J., Parapar, J., Vervoot, M., & Gazave, E. (2019). Morphological, cellular and molecular characterization of posterior regeneration in the marine annelid *Platynereis dumerilii*. *Developmental Biology*, 445(2), 189-210. <https://doi.org/10.1016/j.ydbio.2018.11.004>
- Ponz-Segrelles, G., Bleidorn, C., & Aguado, M. T. (2018). Expression of *vasa*, *piwi*, and *nanos* during gametogenesis in *Typosyllis antoni* (Annelida, Syllidae). *Evolution & Development*, 20(5), 132–145. <https://doi.org/10.1111/ede.12263>
- Reinisch, C., Kandutsch, S., Uthman, A., & Pammer, J. (2006). BMI-1: A Protein Expressed in Stem Cells, Specialized Cells and Tumors of the Gastrointestinal Tract. *HISTOLOGY AND HISTOPATHOLOGY*, 21(11), 1143–1149. <https://doi.org/10.14670/HH-21.1143>
- Randolph, H. (1892). The regeneration of the tail in *Lumbriculus*. *Journal of Morphology*, 7(3), 317–344. <https://doi.org/10.1002/jmor.1050070304>
- Ribeiro, R. P., Ponz-Segrelles, G., Bleidorn, C., & Aguado, M. T. (2019). Comparative transcriptomics in Syllidae (Annelida) indicates that posterior regeneration and regular growth are comparable, while anterior regeneration

- is a distinct process. *BMC Genomics*, 20(1), 855.
<https://doi.org/10.1186/s12864-019-6223-y>
- Sarker, S., Mananteau-Ledouble, S., Kotob, M. H., & El-Matbouli, M. (2017) A RNAi-based therapeutic *proof of concept* targets salmonid whirling disease *in vivo*. *PLoS ONE*, 12(6), e0178687. <https://doi.org/10.1371/journal.pone.0178687>
- Schneider, J., Rasband, W. S., & Eliceri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature methods*, 9(7), 671-675.
<https://doi.org/10.1038/nmeth.2089>
- Shah, S., Islam, M. N., Dakshanamurthy, S., Rizvi, I., Rao, M., Herrell, R., Zinser, G., Valrance, M., Aranda, A., Moras, D., Norman, A., Welsh, J., & Byers, S. W. (2006). The Molecular Basis of Vitamin D Receptor and β -Catenin Coexpression. *Molecular Cell*, 21, 799-809.
<https://doi.org/10.1016/j.molcel.2006.01.037>
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., & Ideker, T. (2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Research*, 13(11), 2498–2504. <https://doi.org/10.1101/gr.1239303>
- Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, 31(19), 3210–3212.
<https://doi.org/10.1093/bioinformatics/btv351>
- Skah, S., Uchuya-Castillo, J., Sirakov, M. & Plateroti, M. (2017). The thyroid hormone nuclear receptors and the Wnt/ β -catenin pathway: An intriguing liaison. *Developmental Biology*, 422(2), 71-82.
<https://doi.org/10.1016/j.ydbio.2017.01.003>
- Srikuea, R., Zhang, X., Park-Sarge, O.-K., & Esser, K. A. (2012). VDR and CYP27B1 are expressed in C2C12 cells and regenerating skeletal muscle: potential role in suppression of myoblast proliferation. *American Journal of Physiology-Cell Physiology*, 303(4), C396–C405. <https://doi.org/10.1152/ajpcell.00014.2012>

- Srivastava, M., Mazza-Curll, K. L., van Wolfswinkel, J. C., & Reddien, P. W. (2014). Whole-Body Acoel Regeneration Is Controlled by Wnt and Bmp-Admp Signaling. *Current Biology*, 24, 1107-1111. <https://doi.org/10.1016/j.cub.2014.03.042>
- Tadokoro, R., Sugio, M., Kutsuna, J., Tochinai, S., & Takahashi, Y. (2006). Early Segregation of Germ and Somatic Lineages during Gonadal Regeneration in the Annelid *Enchytraeus japonensis*. *Current Biology*, 16(10), 1012–1017. <https://doi.org/10.1016/j.cub.2006.04.036>
- Takeo, M., Yoshida-Noro, C., & Tochinai, S. (2008). Morphallactic regeneration as revealed by region-specific gene expression in the digestive tract of *Enchytraeus japonensis* (Oligochaeta, Annelida). *Developmental Dynamics*, 237(5), 1284–1294. <https://doi.org/10.1002/dvdy.21518>
- Takeo, M., Yoshida-Noro, C., & Tochinai, S. (2010). Functional analysis of *grimp*, a novel gene required for mesodermal cell proliferation at an initial stage of regeneration in *Enchytraeus japonensis* (Enchytraeidae, Oligochaete). *The International Journal of Developmental Biology*, 54(1), 151–160. <https://doi.org/10.1387/ijdb.082790mt>
- Tellez-Garcia, A. A., Álvarez-Martínez, R., López-Martínez, J. M., & Arellano-Carbajal, F. (2021). Transcriptome analysis during early regeneration of *Lumbriculus variegatus*. *Gene Reports*, 23, 101050. <https://doi.org/10.1016/j.genrep.2021.101050>
- The UniProt Consortium. (2019). UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Research*, 47(D1), D506–D515. <https://doi.org/10.1093/nar/gky1049>
- Thurmond, J., Goodman, J. L., Strelets, V. B., Attrill, H., Gramates, L. S., Marygold, S. J., Matthews, B. B., Millburn, G., Antonazzo, G., Trovisco, V., Kaufman, T. C., Calvi, B. R., & FlyBase Consortium. (2019). FlyBase 2.0: the next generation. *Nucleic Acids Research*, 47(D1), D759–D765. <https://doi.org/10.1093/nar/gky1003>

- Timm, T. (1999). Distribution of freshwater oligochaetes in the west and east coastal regions of the North Pacific Ocean. *Hydrobiologia*, 406(0), 67–81. <https://doi.org/10.1023/A:1003749519662>
- Tweeten, K. A., & Anderson, A. (2008). Analysis of cell proliferation and migration during regeneration in *Lumbriculus variegatus* (Clitellata: Lumbriculidae). *BIOS*, 79(4), 183–190. <https://doi.org/10.1893/0005-3155-79.4.183>
- Tweeten, K. A., & Reiner, A. (2012). Characterization of serine proteases of *Lumbriculus variegatus* and their role in regeneration. *Invertebrate Biology*, 131(4), 322–332. <https://doi.org/10.1111/ivb.12002>
- Vincent, E. E., Sergushichev, A., Griss, T., Gingras, M-C., Samborska, B., Ntimbane, T., Coelho, P. P., Blagih, J., Raissi, T. C., Choinière, L., Bridon, G., Loginicheva, E., Flynn, B. R., Thomas, E. C., Tavaré, J. M., Avizonis, D., Pause, A., Elder, D. J. E., Artyomov, M. N., & Jones, R. G. (2015). Mitochondrial Phosphoenolpyruvate Carboxykinase Regulates Metabolic Adaptation and Enables Glucose-Independent Tumor Growth. *Molecular Cell*, 60(2), 195-207. <https://doi.org/10.1016/j.molcel.2015.08.013>
- Wang, G., Li, N., Zhang, L., Zhang, L., Zhang, Z., & Wang, Y. (2015). IGFBP7 promotes hemocyte proliferation in small abalone *Haliotis diversicolor*, proved by dsRNA and cap mRNA exposure. *Gene*, 571(1), 65-70. <https://doi.org/10.1016/j.gene.2015.06.051>
- Weigert, A., & Bleidorn, C. (2016) Current status of annelid phylogeny. *Organism Diversity & Evolution*, 16, 345-362. <https://doi.org/10.1007/s13127-016-0265-7>
- Weisblat, D. A., & Kuo, D.-H. (2009). *Helobdella* (Leech): A Model for Developmental Studies. *Cold Spring Harbor Protocols*, 2009(4), pdb.emo121. <https://doi.org/10.1101/pdb.emo121>
- Williams, M. C., Patel, J. H., Kakebeen, A. D., & Wills, A. E. (2021). Nutrient availability contributes to a graded refractory period for regeneration in *Xenopus tropicalis*. *Developmental Biology*, 473, 59-70. <https://doi.org/10.1016/j.ydbio.2021.01.005>

- Wischin, S., Castañeda-Patlán, C., Robles-Flores, M., Chimal-Monroy, J. (2017). Chemical activation of Wnt/ β -catenin signalling inhibits innervation and causes skeletal tissue malformations during axolotl limb regeneration. *Mechanisms of Development*, 144(Part B), 182-190. <https://doi.org/10.1016/j.mod.2017.01.005>
- Yokoyama, H., Ogino, H., Stoick-Cooper, C. L., Moon, R. T. (2007). Wnt/ β -catenin signaling has an essential role in the initiation of limb regeneration. *Developmental Biology*, 306(1), 170-178. <https://doi.org/10.1016/j.ydbio.2007.03.014>
- Zattara, E. E., & Bely, A. E. (2011) Evolution of a novel developmental trajectory: fissionis distinct from regeneration in the annelid *Pristina leidyi*. *Evolution and Development*, 13(1), 80-95. <https://doi.org/10.1111/j.1525-142X.2010.00458.x>
- Zattara, E. E., & Bely, A. E. (2016). Phylogenetic distribution of regeneration and asexual reproduction in Annelida: regeneration is ancestral and fission evolves in regenerative clades. *Invertebrate Biology*, 135(4), 400–414. <https://doi.org/10.1111/ivb.12151>
- Zhou, X., Paredes, J. A., Krishnan, S., Curbo, S., & Karlsson, A. (2015). The mitochondrial carrier SLC25A10 regulates cancer cell growth. *Oncotarget*, 6(11), 9271-9283. <https://doi.org/10.18632/oncotarget.3375>
- Zoran, M. J., & Martinez, V. G. (2009). *Lumbriculus variegatus* and the Need for Speed: A Model System for Rapid Escape, Regeneration and Asexual Reproduction. In: D. H. Shain (Ed.), *Annelids in Modern Biology* (pp. 185–202). Hoboken, NJ, USA: John Wiley & Sons, Inc. <https://doi.org/10.1002/9780470455203.ch10>

Anexo 1. Artículo publicado a partir de los resultados del Capítulo 1.

Gene Reports 23 (2021) 101050



Contents lists available at ScienceDirect

Gene Reports

journal homepage: www.elsevier.com/locate/genrep



Transcriptome analysis during early regeneration of *Lumbriculus variegatus*

Aldo Arturo Tellez-García, Roberto Álvarez-Martínez, Juana María López-Martínez, Fausto Arellano-Carbajal*

Natural Sciences School, Autonomous University of Queretaro, Mexico

ARTICLE INFO

Keywords:
Regeneration
Annelida
Lumbriculus variegatus
RNA-seq
Transcriptome

ABSTRACT

Several annelid species exhibit an impressive regenerative capacity. Among them, *Lumbriculus variegatus* can reconstruct its entire body from only three original segments. Due to its high regenerative capability, *L. variegatus* has been proposed as a model for experimental studies of regeneration. However, limited molecular and genetic information is available for *L. variegatus*. Here, we conducted a transcriptome analysis during early regeneration of *L. variegatus* using RNA-sequencing. We generated 61 million Illumina reads during early regeneration of *L. variegatus*. These reads were assembled into two transcriptomes, one of which was selected for annotation. The transcriptome contained 138,008 genes and 164,769 transcripts. We identified 44,097 non-coding transcripts and 65,692 candidate cds with an ORF. Functional annotation with GO categories was possible for 36,494 predicted proteins. One hundred and thirty-six transcripts were differentially expressed during early regeneration, of which 111 were upregulated and 25 were downregulated. GO enrichment analysis showed that carboxylase activity, amino acid metabolism, and metabolism of glycosides were enriched. *vdr*, which encodes for a Vitamin D3 receptor was upregulated during early regeneration. This gene has been implicated in regeneration in vertebrates. On the other hand, *wvde*, a gene associated with the proximal-distal determination of the blastema, was downregulated. Unexpectedly, the Glutamine synthetase coding gene was also downregulated, but this gene was found to be upregulated during regeneration in other oligochaetes. Further functional characterization of these genes in *L. variegatus* will contribute to our understanding of regeneration. This transcriptome will be a resource for studies aimed at the identification of the conserved molecular mechanisms underlying the regeneration.

1. Introduction

Regeneration, “the restoration of lost body parts”, is a common response of animals to traumatic injury, either after amputation or autotomy (Bely and Nyberg, 2010). In annelids, the ability to regenerate the primary body axis is common, but the extent of regeneration is variable (Zattara and Bely, 2016). Some species can only regenerate the anterior portion of the body, others can only regenerate the posterior portion, and others can regenerate both (Ózpolat and Bely, 2016; Zattara and Bely, 2016). In addition, most species restore lost body parts through epimorphosis (*i.e.*, the formation of new tissue by cell proliferation, blastema formation, and differentiation), but in some cases, this type of regeneration is coupled with morphallaxis (*i.e.*, the remodeling of remaining tissues into different ones) (Bely, 2014; Ózpolat and Bely, 2016). Despite these differences, annelid regeneration heavily relies on effective wound healing and the formation of a blastema, capable of

differentiating into new tissues (Bely, 2014). These early stages of regeneration have been extensively studied in some annelid species from a cellular perspective, but the regulatory genes and underlying molecular mechanisms of regeneration remain unclear (Bely, 2014; Ózpolat and Bely, 2016).

In the past two decades, several studies have analyzed the expression of genes that were known to participate in the regeneration and development of other organisms. Mainly using RNA *in situ* hybridization, the expression patterns of *piwi*, *PL10*, *nanos*, *vasa*, *twist*, the glutamine synthetase coding gene, *frizzled*, β -*catenin*, *mino*, *Tuba*, *horu*, *grimp*, and several *Hox* genes were determined during early regeneration for a variety of species (*e.g.* Bely and Sikes, 2010; de Jong and Seaver, 2016; Kozin and Kostyuchenko, 2015; Niva et al., 2008; Nyberg et al., 2012; Ózpolat and Bely, 2015, 2016; Pfeifer et al., 2012; Pfeifer et al., 2013; Tadokoro et al., 2006; Takeo et al., 2008, 2010). Furthermore, high-throughput expression profiling analyses have identified large sets of

* Corresponding author.

E-mail address: fausto.arellano@uaq.mx (F. Arellano-Carbajal).

<https://doi.org/10.1016/j.genrep.2021.101050>

Received 3 October 2020; Received in revised form 19 January 2021; Accepted 31 January 2021

Available online 16 February 2021

2452-0144/© 2021 Elsevier Inc. All rights reserved.

genes with transcriptional activity during early regeneration in *Enchytraeus japonensis* (Myohara et al., 2006), *Perionyx excavatus* (Cho et al., 2009), *Pristina leidyi* (Nyberg et al., 2012), *Sphaerosyllis hystrix* and *Syllis gracilis* (Ribeiro et al., 2019). While most of these studies have reported transcriptional activity of *a priori* selected candidate genes, only Myohara et al. (2006) and Ribeiro et al. (2019) focused on the identification of differentially expressed genes (DEG). This type of approach is very valuable since it allows us to discover new candidate genes based on their relative expression levels, even if they were not expected to participate in regeneration based on the literature. Moreover, DEG analyses may help identify if there is a common genetic program that regulates regeneration in annelids once other species in this phylum are analyzed.

The California blackworm, *Lumbriculus variegatus*, is an annelid species with remarkable regeneration, in which even intermediate fragments of a few segments can fully regenerate both the anterior and posterior portions of the body (Morgulis, 1907). After amputation, regeneration starts with the occlusion of the wound by rapid muscle contractions and the dedifferentiation of cells located near the wound site (Bely, 2014; Lesiuk and Drewes, 1999; Tweeten and Anderson, 2008). These cells begin to proliferate during the first 24 hours post amputation (hpa) and become the main source for the formation of a blastema (Tweeten and Anderson, 2008). Other cellular sources that contribute to blastema formation are migratory cells from non-regenerating tissues, which have been observed to invade the blastema 72 hpa (Randolph, 1892; Tweeten and Anderson, 2008; Tweeten and Reiner, 2012). The formation of the blastema is followed by its growth and subsequent differentiation into new tissues, which become evident on the fifth day of regeneration, when new head and tail segments can be observed (Tweeten and Reiner, 2012). Besides epimorphic regeneration, the ventral nerve cord suffers morphallactic changes within the first two days after amputation, which allows giant fibers to adopt new positional identities (Drewes and Fournier, 1990).

In order to identify the changes in gene expression that occur during early stages of regeneration and to identify candidate genes that might regulate the onset of this process, we carried out a transcriptional analysis on the first 72 h of regeneration in *L. variegatus*. Then, we performed a functional analysis of DEG with the aim to determine which other processes are associated with early regeneration. As part of our methodology, we also generated the first annotated transcriptome of this species that includes a regenerating population. Finally, we compared our results with those obtained through similar analyses in other annelid species.

2. Methods

2.1. Animals

Lumbriculus variegatus worms were acquired from a commercial supplier (Aquarem, Mexico) and maintained in plastic containers (base: 240 cm²) with ~0.6 L of purified bottled water at ~5 °C. Population density was variable, but we removed the excess so that worms never fully covered the bottom of the containers. Worms were fed with spirulina powder twice a week, by sprinkling the algae on the water surface. Containers were cleaned once a week and water was changed every time the worms were fed.

2.2. Sampling and establishment of experimental groups

Three-hundred worms with no signs of fission were distributed in 30 Petri dishes (60 × 15 mm; 10 worms per dish). They were maintained with purified bottled water and no food for one week at 20 °C. Following this starvation period, on day 0 40 worms were assigned to the regenerating group (RG). They were separated in five Petri dishes with purified bottled water (8 worms per dish), transected approximately by the mid-body, and maintained at 20 °C in a regenerating period (RP) of 72 h

without food. An additional 40 worms were assigned to the non-regenerating group (NG). These were treated equally to the RG worms but without transecting them and were left for the same RP. This procedure was repeated on days 1 and 2 after the starvation period, but worms of both experimental groups were maintained in RPs for 48 and 24 h, respectively. In total, 240 worms were used for each experimental condition (Supplementary File 1): 120 belonging to the RG (40 worms from the RP of 72 h, 40 from the RP of 48 h, and 40 from the RP of 24 h) and 120 belonging to the NG (coming from the same RPs as the RG worms). The remaining 60 worms of the initial 300 were used as a backup in case specimens suffered autotomy or died during the starvation period. On day 3 after the starvation period, 8 worms from each RP were pooled in a single sample for both experimental groups, transferred into a microcentrifuge tube with 1 mL of TRIzol (Ambion), and stored at -80 °C. In this way, five replicates with 24 worms each (8 worms from each RP) were established for both groups. Transections were carried out under the stereoscopic microscope placing the worms on a moistened filter paper in a Petri dish and cutting the mid-body of the animals with a razor blade in a single movement.

2.3. RNA purification and sequencing

Samples were thawed on ice before processing and worms were homogenized into the tubes with TRIzol, first by maceration with a Teflon pestle and then by passing the broken tissue several times through an insulin syringe. Total RNA was extracted using the Direct-zol RNA MiniPrep (Zymo Research) kit following the manufacturer instructions. RNA integrity was visually analyzed in a 2% agarose gel electrophoresis while purities and concentrations were measured in a NanoDrop 2000 (ThermoFisher). Those three samples (replicates) of each group with the highest concentration and best purity values were sequenced at the Unit of Massive Sequencing and Bioinformatics of the National Autonomous University of Mexico. Libraries were prepared with the Illumina TruSeq Stranded mRNA protocol and sequencing was performed on an Illumina NextSeq 500 platform.

2.4. Transcriptome assembly

Raw reads were assessed in FastQC (v0.11.9; Andrews, 2010) for the absence of adapters and read quality. Sites with low PHRED scores were trimmed with Trimmomatic (v0.39; Bolger et al., 2014) using the *slidingwindow* option with a window size of 4 bases and a quality average of 20. As a second step, trimmed reads with less than 30 bases were discarded using the *minlen* option. Two transcriptomes were assembled *de novo* in Trinity (Galaxy v0.0.1; Afgan et al., 2018; Haas et al., 2013), feeding the program with the concatenated left (R1) and the concatenated right (R2) reads, specifying an *RF* library type, and using the default parameters, except for the *path reinforcement distance* (*prd*), which was set to 75 base pairs (bp) for the first assembly and to 50 bp for the second assembly in order to evaluate which value produced the most complete assembly. Duplicated sequences were filtered out with CD-HIT-EST (v4.7; Fu et al., 2012; Li and Godzik, 2006) using a *sequence identity threshold* of 0.95 and a *word size* of 10. Transcriptome statistics were obtained by running the TrinityStats.pl script of the Trinity toolkit (v2.10.0; Haas et al., 2013). Completeness of both transcriptomes was assessed in the Benchmarking Universal Single-Copy Orthologs (BUSCO) program (v3; Simão et al., 2015) using the metazoan database (metazoa_odb9), which contains 978 single-copy genes that were expected to be present.

2.5. Functional annotation

The transcriptome assembled with a *prd* of 50 was functionally annotated using the Trinotate suite (v3.3.0; Bryant et al., 2017). First, open reading frames (ORFs) and coding regions (cds) of at least 100 amino acids long were identified with TransDecoder (v5.5.0; Haas et al., 2013).

Then, homology searches against the UniProtKB/Swiss-Prot database (The UniProt Consortium, 2019) were performed in BLAST (v2.10.0; Altschul et al., 1997); for the assembled transcripts using BLASTx and for the TransDecoder-predicted proteins using BLASTp. The *E*-value cutoff was set to 0.001 in both Blast homology searches. TransDecoder-predicted proteins were also used to search for protein domains with HMMER (v3.3; Finn et al., 2011) using the Pfam database (El-Gebali et al., 2019), to predict signal peptides with signalP (v4.1; Nielsen, 2017), and to predict transmembrane regions with tmHMM (v2.0; Krogh et al., 2001). Finally, homology searches and predictions results were loaded into a Trinitate SQLite Database and deposited in zenodo (zenodo.org; DOI: <https://doi.org/10.5281/zenodo.4435184>).

In order to determine the presence of lncRNAs, we analyzed the coding potential of the transcripts assembled in the transcriptome using the RNAsamba program (Camargo et al., 2020), which uses neural networks to classify the transcripts into coding (mRNAs) and non-coding (lncRNAs). As a reference, we used a pre-trained file with genes of *Homo sapiens* included in the program. After that, we blasted (BLASTn) the non-coding transcripts against a multi-species database of lncRNAs (with an *E*-value threshold of 0.001). This database was generated by concatenating files with sequences of *Caenorhabditis elegans*, *Danio rerio*, *Drosophila melanogaster*, *Gallus gallus*, *H. sapiens*, *Mus musculus*, *Rattus norvegicus*, and *Saccharomyces cerevisiae* ncRNAs (especially lncRNAs) downloaded from NONCODE (<http://www.noncode.org/download.php>).

Additionally, to compare the content of the assembled transcriptome with the genomes of other annelids, all TransDecoder-predicted proteins were blasted (tBLASTn; *E*-value threshold of 0.001) against the genomes of *Helobdella robusta* and *Capitella teleta* (Simakov et al., 2013).

2.6. Protein clustering and identification of potential paralogous genes

In order to identify the presence of possible paralogous genes, we performed a hierarchical clustering of all the predicted proteins with CD-HIT (v4.7; Fu et al., 2012; Li and Godzik, 2006). Peptide sequences were sequentially clustered with a sequence identity of 90%, 60%, and 40%, always with an alignment coverage of 75% of the shortest aligned sequence (Devos et al., 2016). Then, cluster files were combined with *clstr_rev.pl* script of CD-HIT and, finally, BLASTp results of the annotated transcriptome were matched with their protein IDs to identify which genes belong to a determined cluster (potential group of paralogs).

2.7. Differential expression analysis

Quantification of transcript abundances was performed for each sample in kallisto (v0.46.1; Bray et al., 2016), using the parameters given by default and the annotated transcriptome as the reference for pseudo alignment. A transcript-level differential expression analysis was carried out in sleuth (Pimentel et al., 2017), where two regression models were generated: one in which gene expression responded to the treatment and another in which gene expression did not respond to any factor. Both models were compared with a Likelihood Ratio Test (LRT) to identify those transcripts that appeared to be significantly affected by the treatment. Finally, the fold change in the expression for each transcript was obtained with a Wald Test (WT), using the parameter $transform_fun_counts = function(x) \log_2(x + 0.5)$. Those transcripts that had a *q*-value < 0.05 in both tests and a beta-value > 1 OR < -1 were considered to be differentially expressed.

Additionally, in order to corroborate the results obtained with sleuth, we carried out a second transcript-level differential expression analysis with DESeq2 (Love et al., 2014). After performing a WT, those transcripts with a *p*-value < 0.05 and a \log_2 fold change > 1 OR < -1 were considered to be differentially expressed. Transcript abundances were imported into DESeq2 with tximport (Soneson et al., 2016).

2.8. Gene ontology (GO) enrichment analysis

Predicted proteins that showed significant BLASTp hits and were encoded by differentially expressed transcripts were selected for the search of protein orthologs in *Crassostrea gigas*, *D. rerio*, *D. melanogaster*, and *Xenopus laevis*. These species were selected because, in contrast with *L. variegatus* and other annelids, they have sequenced genomes with relatively well-annotated GO terms. Orthologs were searched in the OrthoDB (v10.1; Kriventseva et al., 2019), UniProtKB (The UniProt Consortium, 2019), and FlyBase (Thurmond et al., 2019) databases, from where their UniProtKB and Ensembl (Yates et al., 2019) identifiers (IDs; for proteins and coding genes, respectively), were retrieved. These identifiers were used to perform GO enrichment analyses in the ClueGO App (v2.5.4; Bindea et al., 2009) of Cytoscape (v3.7.2; Shannon et al., 2003), where networks of enriched GO terms and signaling pathways were constructed based on the associated genes that two or more terms/pathways shared. The terms and signaling pathways included in the analyses corresponded to the following categories: Biological Processes (BP), Cellular Components (CC), Immune System (IS), Molecular Functions (MF), and KEGG pathways (Kanehisa and Goto, 2000). The analyses were carried out with a Two-sided hypergeometric test and a Bonferroni step down *p*-value correction, using as background all the annotated genes of each species, and separating in different clusters the IDs of the downregulated and upregulated genes. Other parameters were set as follows: a GO tree interval of 4 to 10 GO levels, a minimum of two genes from the input list present in a determined GO term, and that those genes represented at least 4% of the total genes contained in that GO term. For network construction, a Kappa score (Cohen's kappa coefficient) was used to associate terms with a cutoff of 0.4.

3. Results

3.1. RNA-seq and de novo transcriptome assembly

A total of 61 million paired-end reads of 76 bp length were obtained. The depth of the sequenced samples ranged from 8 to 11 million paired-end reads (Table 1). 50 million paired-end reads were retained after trimming low-quality portions of the reads and discarding reads with less than 30 bp.

Two transcriptome assemblies were carried out with Trinity. The first one (assembled with a *prd* value of 75 bp) produced 403,374 trinity transcripts, while the second (assembled with a *prd* value of 50 bp) produced 403,488. Both assemblies had very similar N50 values, median contig lengths, and percentage of retained contigs after filtration in CD-HIT-EST (40%; Table 2). Similarly, completeness analysis in BUSCO showed that both assemblies had the same number of missing orthologous genes, but slightly differed in the number of complete and fragmented sequences, being the second assembly the most complete (Fig. 1). Based on these results, the second assembly was selected for annotation and further analyses. This transcriptome contains 138,008 genes and 164,769 transcripts. The raw reads.

were uploaded at the NCBI Sequence Read Archive (SRA; BioProject: PRJNA662619, SRR12619265–70).

Table 1
Number of paired-end reads before and after trimming and dropping low quality and small read fragments.

Experimental group	Sample ID	Raw paired-end reads	Trimmed paired-end reads
NG	Lv_NG_1	8,438,983	6,831,949
NG	Lv_NG_2	11,369,320	9,643,959
NG	Lv_NG_3	11,769,910	9,754,524
RG	Lv_RG_1	9,127,239	7,126,776
RG	Lv_RG_2	9,997,890	8,162,444
RG	Lv_RG_3	10,323,177	8,538,427

Table 2
Assembly statistics of the two assembled transcriptomes.

Statistics	First assembly (prd = 75 bp)	Second assembly (prd = 50 bp)
Total trinity genes	241,853	241,907
Total trinity transcripts	403,374	403,488
N50	497	497
Median contig length	322	321
Total assembled bases	180,384,000	180,364,240
Number of contigs retained after filtration	164,672	164,769

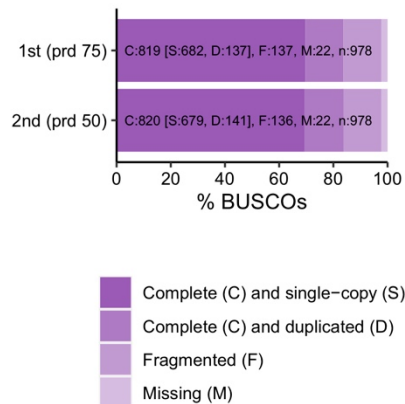


Fig. 1. BUSCO assessment results. Bars represent the percentages of complete, fragmented, and missing orthologs (BUSCOs) of the two assembled transcriptomes compared against the metazoan database. Graph plotted with the ggplot2 package of R (Wickham, 2016; R Core Team, 2020) based on the source code provided by BUSCO.

3.2. Functional annotation of the transcriptome

The second assembled transcriptome was annotated with Trinotate, where predicted cds, protein domains, signal peptides, and transmembrane regions, together with BLAST results, were loaded into a Trinotate SQLite Database (Supplementary File 2) and are also available in zenodo (<https://zenodo.org>; DOI: <https://doi.org/10.5281/zenodo.4435184>). 65,692 candidate cds with an ORF of at least 100 amino acids long were detected with TransDecoder. Around half of these coding regions were identified as part of internal ORFs (53.94%; with no start and stop codons); while the rest was distributed among complete (10.45%; with both start and stop codons), 3' (12.28%; with a stop but not a start codon), and 5' (23.31%; with a start but not a stop codon) ORFs (Fig. 2). 40,037 predicted proteins (candidate cds from TransDecoder) had significant hits in BLASTp, from which 36,494 were functionally annotated with GO categories in Trinotate.

For the RNAsamba analysis, we identified 44,097 non-coding transcripts (potential lncRNAs) of the 164,769 in the transcriptome (Supplementary file 3). 88.72% of these non-coding transcripts corresponded to transcripts where TransDecoder found no cds. We blasted the 44,097 non-coding transcripts against a multispecies lncRNA database and obtained only 83 hits (Supplementary File 3).

By making alignments (tBLASTn) of the amino acid sequences with the genomes of *H. robusta* and *C. teleta*, 57.14% had significant hits with *H. robusta*, 62.69% with *C. teleta*, and 32.89% with none of the genomes (Fig. 3).

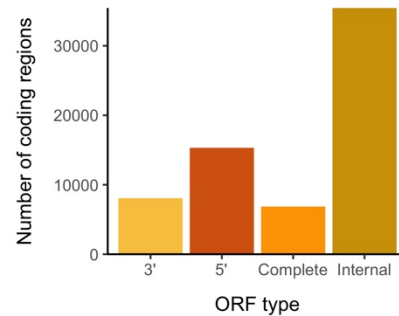


Fig. 2. Number of coding regions with open reading frames (ORF) of at least 100 amino acids. The ORF types are 3' (without start codon), 5' (without stop codon), complete (with both start and stop codons), and internal (without start and stop codons). Graph plotted with the ggplot2 package of R (Wickham, 2016; R Core Team, 2020).

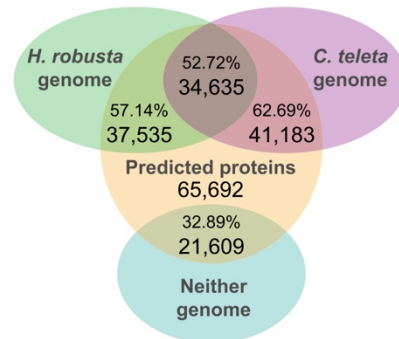


Fig. 3. Venn diagram showing the percentage and number of predicted proteins that had significant hits against the genomes of *H. robusta* and *C. teleta*, both genomes, and neither of them.

3.3. Clusters of potential paralogous genes

31,469 clusters were found with the hierarchical clustering algorithm of CD-HIT (Supplementary File 4). By analyzing the distribution of candidate regeneration genes among the clusters, some of them were found to be clustered with members of the same family while others were clustered with other gene families. For example, the β -catenin genes were clustered all together, except for one, and with none of the other Catenins present in the transcriptome (Clusters 524 and 20,911; Supplementary File 4). Similarly, *AIRC* or *paics* (Cluster 4347; Supplementary File 4), *nanos* (9126 and 21,763; Supplementary File 4), and *piwi* (9215, 16,836, and 360; Supplementary File 4) were grouped in one or more clusters but with no other genes. *PL10* was clustered with other DEAD-box proteins (Cluster 654; Supplementary File 4). In contrast, the two coding genes of *JNK* present in the transcriptome were found to be clustered with several members of the MAPKs gene family and a cyclin-dependent kinase (Cluster 5866; Supplementary File 4). The Glutamine synthetase coding genes were clustered with several lensin, Glutamate-methylamine ligase, and Glutamate-isopropylamine ligase coding genes (Cluster 3067; Supplementary File 4).

3.4. Differential expression during early regeneration

One hundred and thirty-six transcripts were found to be differentially expressed using sleuth (LRT and WT: $q < 0.05$); 25 of them were downregulated ($\beta < -1$) and 111 were upregulated ($\beta > 1$) during early regeneration (Fig. 4). From the total downregulated transcripts, 24 had a predicted ORF and 21 of the predicted proteins had significant hits in BLASTp (Table 3). Similarly, 85 of the total upregulated transcripts had a predicted ORF and 52 predicted proteins had significant hits in BLASTp (Table 3). 13 transcripts classified as non-coding by RNAsamba were upregulated and none of them had significant blast hits against the multispecies lncRNA database (Table 3).

The differential expression analysis carried out with DESeq2 showed very similar results to those obtained with sleuth. 11 transcripts were downregulated and 103 were upregulated (Supplementary File 5). Most of these transcripts and their predicted proteins were also found to be differentially expressed in the analysis with sleuth (Table 3; Supplementary File 5). Therefore, and since the analysis in sleuth provided a bigger number of differentially expressed transcripts, we used the results of sleuth for the GO enrichment analysis.

3.5. Enriched GO terms during early regeneration

Orthologs of the 73 predicted proteins encoded in differentially expressed transcripts and that had significant hits in BLASTp were searched in the OrthoDB, UniProtKB, and FlyBase databases. The orthologs in *C. gigas*, *D. rerio*, *D. melanogaster*, and *X. laevis*, and their IDs from Ensemble and UniProtKB are presented in the Supplementary File 6. IDs were used to carry out enrichment analyses in each species with the ClueGO app of Cytoscape, separating in different clusters the downregulated and upregulated IDs for the construction of networks with the enriched GO terms and KEGG pathways (Supplementary File 7).

Enrichment analyses showed that carboxy-lyase activity (MF category) was consistently enriched across the four reference species, due to the upregulation of the *PCK2* and *tdc* orthologs (Fig. 5). Similarly, amino

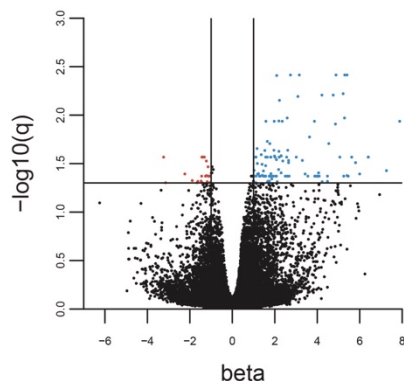


Fig. 4. Differentially expressed transcripts during the first 3 days of regeneration. Volcano plot that shows an approximate value of the \log_2 fold-change in expression (β) for each transcript (dots) against their significance value (q) in a $-\log_{10}$ scale. Dots above the horizontal line correspond to those transcripts with a $q < 0.05$ in the LRT. In red are those transcripts with a half or less-fold change in expression (downregulated; $\beta \leq -1$). Blue dots represent those transcripts with a two or more-fold change in expression (upregulated; $\beta \geq 1$). Vertical lines indicate a half ($\beta = -1$) and a two ($\beta = 1$) fold change in expression. Graph plotted in R (R Core Team, 2020). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

acid metabolism-related terms, such as glutamine family amino acid metabolism (BP category) and alpha-amino acid biosynthesis (BP category), were also enriched across the four species (Fig. 5). Specifically, the metabolism of glutamine-family amino acids was mainly enriched by the downregulation of the Glutamine synthetase coding gene, and the upregulation of the *ALDH18A1* and the *Gls2* orthologs. In the case of alpha-amino acid biosynthesis, this term was enriched by the same orthologous genes plus the upregulation of the ortholog of *BHMT*. Other biological processes that were enriched in more than one species were those related to the metabolism of glycosides, in *D. rerio* and *D. melanogaster*, due to the downregulation of the Alpha-L-fucosidase coding gene and the *NAGA* ortholog (Fig. 5B and C). Similarly, the metabolism of D-Glutamine and D-glutamate was significantly enriched in *D. melanogaster* and *X. laevis* by the upregulation of the *Gls2* and *GLUD2* orthologs (Fig. 5A and B).

4. Discussion

Our transcriptome contained 138,008 genes and 164,769 transcripts. According to the results of the BUSCO completeness analysis, 97.7% of the orthologous genes expected to be present in metazoans were found in this transcriptome (83.8% were complete; see Fig. 1). These percentages were close to those obtained in other annelid transcriptomes assembled with similar methodologies and, since sequences with more than a 95% of similarity were filtered out, this transcriptome also contained a reduced percentage of duplicated copies (14.4%) (Heikkinen et al., 2017; Ponz-Segrelles et al., 2018; Ribeiro et al., 2019). However, the N50 value indicates that the continuity of the assembled transcripts is not as high as would be expected, but it is superior to that obtained in other recently assembled oligochaete transcriptomes (Ponesakki et al., 2017; Thunders et al., 2017). 36,494 of the predicted proteins with significant hits in BLASTp were associated with at least one GO term.

By comparing the contents of the transcriptome with the genomes of *H. robusta* and *C. teleta*, we found that more than 50% of the predicted proteins had significant alignments against both genomes. Using a more stringent *E*-value threshold (2e-15), Mehr et al. (2015) found that 27.1% of the 58,454 predicted proteins present in a transcriptome of the polychaete *Hermodice carunculata* had significant alignments (BLASTp) with both genomes (40.5% with *C. teleta* and 35% with *H. robusta*). Interestingly, even though *L. variegatus* is phylogenetically closer to *H. robusta* than to *C. teleta* (Zattara and Bely, 2016), there were more *L. variegatus* predicted proteins with significant similarities to those of *C. teleta* than to *H. robusta*, but it has to be considered that the genome of *C. teleta* has more coding genes than the genome of *H. robusta* (32,175 vs 23,432, respectively) (Simakov et al., 2013).

The differential expression analysis showed that 136 transcripts were differentially expressed during early regeneration. From these, 73 had predicted ORFs and coded for proteins with significant hits in BLASTp. The remaining transcripts can correspond either to spurious sequences, non-coding RNAs, or novel genes, but these possibilities require further analyses (Bhambri et al., 2018).

Among the differentially expressed transcripts, we found three upregulated whose coding gene orthologs have been reported to be involved in regeneration or stem cell maintenance in other animals: *bmi1b* (Polycomb complex protein BMI-1-B), *Hsp60A* (Heat shock protein 60A), and *vdr* (Vitamin D3 receptor). As part of the Polycomb repressive complexes, BMI-1 promotes the self-renewal of stem cells in several metazoan organs, and it has been shown to be implicated in the pancreas and gut epithelial regeneration of mammals (Fukuda et al., 2012; Lopez-Arribillaga et al., 2015; Reinisch et al., 2006). Hsp60 is a stress-response protein that has been demonstrated to participate in wound healing and blastema formation and maintenance in amputated caudal fins of *Paramisgurnus dabryanus* and *D. rerio* (Li et al., 2014; Makino et al., 2005; Pei et al., 2016). The Vitamin D3 receptor (*HR96* in several invertebrates; Litoff et al., 2014) can play different tissue-dependent roles in regeneration; while it stimulates cardiomyocyte

Table 3
Differentially expressed transcripts with their coding potential (RNAsamba) and their predicted protein identities according to the BLASTp results.

Transcript ID	q value ^a	beta	Coding potential	Protein coordinates	UniProtKB ID	E-value	Gene	Protein
TRINITY_DN107164_c0_g1_i2	0.042480926	1.836549578	coding					
TRINITY_DN118089_c0_g1_i1	0.042480926	2.585215768	coding	153-824[+]				
TRINITY_DN119931_c0_g1_i1	0.01234777	4.858756433	non-coding					
TRINITY_DN126101_c0_g1_i1	0.028738752	2.081652193	coding	85-477[+]				
TRINITY_DN127537_c0_g1_i1	0.023672617	2.196215335	coding	2-898[+]	DER_CHICK	1.95E-81	DER	D-erythrose reductase
TRINITY_DN128414_c0_g1_i7	0.042637618	5.24558697	coding					
TRINITY_DN131814_c0_g1_i1	0.027154302	3.012153423	coding	104-532[+]				
TRINITY_DN131814_c0_g2_i2	0.027154302	2.558505061	coding	106-738[+]	TENA_CHICK	1.54E-24	TNC	Tenascin
TRINITY_DN132073_c0_g2_i5	0.029772101	2.382378947	coding	120-971[+]				
TRINITY_DN132086_c0_g2_i2	0.030946663	2.648450292	coding	140-1069[-]	MRC1_MOUSE	2.15E-06	Mrc1	Macrophage mannose receptor 1
TRINITY_DN132581_c0_g2_i7	0.045410469	1.861925798	coding	3-521[-]				
TRINITY_DN132821_c1_g1_i1	0.027154302	5.066803871	coding					
TRINITY_DN133355_c0_g2_i1	0.042480926	2.046329277	coding	2-343[-]				
TRINITY_DN133765_c0_g1_i1	0.040154043	1.821896144	coding	270-785[+]	RNC_MOUSE	2.07E-29	Drosha	Ribonuclease 3
TRINITY_DN134295_c1_g1_i2	0.03152603	1.176590666	coding					
TRINITY_DN135002_c0_g1_i6	0.040504243	1.197139031	coding	113-523[+]				
TRINITY_DN135551_c0_g2_i3	0.042637618	3.880456614	coding					
TRINITY_DN135612_c0_g1_i5	0.042637618	4.018399851	non-coding					
TRINITY_DN136715_c0_g3_i4	0.003839525	2.741313977	coding	104-1243[+]				
TRINITY_DN137232_c1_g1_i4	0.070708421	2.226023779	coding	102-854[+]	SCA_DROME	4.72E-25	sca	Protein scabrous
TRINITY_DN137442_c1_g1_i4	0.027154302	1.818777726	coding	55-489[+]				
TRINITY_DN138956_c0_g2_i2	0.048181466	1.673943568	coding	172-828[+]	STAR5_MOUSE	1.18E-45	Stard5	STAR-related lipid transfer protein 5
TRINITY_DN139156_c0_g1_i1	0.016738305	3.639943829	coding					
TRINITY_DN139156_c1_g1_i2	0.042637618	5.453108707	coding	3-455[+]				
TRINITY_DN139156_c1_g1_i7	0.005998649	5.215429158	coding					
TRINITY_DN139156_c1_g1_i8	0.042637618	3.907882366	non-coding					
TRINITY_DN139686_c0_g1_i2	0.039234618	4.058088235	coding	88-678[-]				
TRINITY_DN139686_c0_g1_i4	0.048448131	2.115158541	coding	19-783[+]	XDH_DROSO	7.73E-65	Xdh	Xanthine dehydrogenase
TRINITY_DN140112_c0_g4_i6	0.042637618	5.458878121	coding					
TRINITY_DN140818_c0_g1_i1	0.042480926	1.752060858	coding	238-1743[+]	RN217_XENLA	1.47E-50	rnf217	Probable E3 ubiquitin-protein ligase RNF217
TRINITY_DN140866_c0_g1_i1	0.037454652	1.467082032	coding	2-340[+]				
TRINITY_DN140949_c0_g1_i1	0.011575834	7.882464106	coding					
TRINITY_DN142268_c0_g1_i1	0.027154302	5.619820311	coding	206-1213[-]	CNOT9_RAT	1.49E-84	Cnot9	CCR4-NOT transcription complex subunit 9
TRINITY_DN142838_c1_g2_i3	0.00622477	4.222214553	non-coding	191-532[-]				
TRINITY_DN143316_c1_g2_i2	0.042637618	1.562028235	coding					
TRINITY_DN143360_c1_g2_i12	0.003890196	2.096281624	coding	224-3043[+]	TYRDC_ENTFC	6.14E-85	tdc	L-tyrosine decarboxylase
TRINITY_DN143360_c1_g2_i7	0.042637618	3.56297818	coding	3-353[+]	TYRDC_ENTFC	1.24E-08	tdc	L-tyrosine decarboxylase
TRINITY_DN143477_c1_g1_i3	0.029200349	3.299343203	coding	3-1028[+]	LIPR1_RAT	6.06E-45	Pnlpp1	Inactive pancreatic lipase-related protein 1
TRINITY_DN143688_c0_g1_i1	0.042637618	1.745398642	coding	1-399[-]	TYRDC_ENTFC	3.48E-11	tdc	L-tyrosine decarboxylase
TRINITY_DN143865_c0_g1_i5	0.047747385	1.745985562	non-coding					
TRINITY_DN144159_c0_g1_i3	0.011575834	2.350824922	coding	1-696[-]	AB22G_ARATH	2.65E-47	ABCG22	ABC transporter G family member 22
TRINITY_DN144384_c0_g2_i6	0.042637618	2.676921632	coding	129-938[-]	CNMD_DANRE	4.03E-06	cnmd	Leukocyte cell-derived chemotaxin 1
TRINITY_DN145599_c0_g1_i3	0.027154302	1.084402319	coding	61-972[+]	FBRI_XENLA	3.21E-145	fbf1	rRNA 2'-O-methyltransferase fibrillar
TRINITY_DN145656_c0_g2_i3	0.027154302	1.803025966	coding	221-925[-]	PPIB_RAT	2.58E-52	Ppib	Peptidyl-prolyl cis-trans isomerase B
TRINITY_DN146187_c0_g1_i4	0.042637618	1.494385147	coding	114-752[+]				
TRINITY_DN146805_c0_g1_i1	0.027154302	1.116740538	coding	223-891[+]				
TRINITY_DN147249_c0_g1_i1	0.036386268	2.176304958	coding	108-632[+]	UBCD2_DROME	4.8E-104	Ubc2	Ubiquitin-conjugating enzyme E2-24 kDa
TRINITY_DN147672_c0_g1_i3	0.011575834	3.863359295	non-coding	3-2360[-]				
TRINITY_DN147710_c0_g1_i1	0.049897086	1.706317613	coding	109-873[+]	HDHD3_BOVIN	1.34E-46	HDHD3	Haloacid dehalogenase-like hydrolase domain-containing protein 3
TRINITY_DN148419_c1_g2_i1	0.03325253	1.569590355	coding	130-516[-]	LOXH1_HUMAN	2.48E-16	LOXH1	Lipoxygenase homology domain-containing protein 1
TRINITY_DN149785_c2_g1_i2	0.023000595	2.613752942	coding	3-587[+]	LECM_OXYSU	0.0000066	N/A	C-type lectin mannose-binding isoform
TRINITY_DN149785_c2_g1_i4	0.026216117	2.301636807	coding	1-471[+]				
TRINITY_DN149833_c0_g1_i11	0.037368021	7.254803295	coding	479-835[-]				

(continued on next page)

Table 3 (continued).

Transcript ID	q value ^a	beta	Coding potential	Protein coordinates	UniProtKB ID	E-value	Gene	Protein
TRINITY_DN149833_c0_g1_i4	0.041959542	5.396879173	coding					
TRINITY_DN149833_c0_g1_i5	0.019636382	4.546602658	coding					
TRINITY_DN149833_c0_g1_i6	0.042637618	4.016994448	coding					
TRINITY_DN150530_c0_g1_i3	0.042637618	1.0354762	coding	344-1171[-]	SYK_MOUSE	5.79E-155	Kars1	Lysine-tRNA ligase
TRINITY_DN150716_c0_g2_i9	0.040504243	3.433564263	coding	178-840[-]	PNPH_HUMAN	1.54E-77	PNP	Purine nucleoside phosphorylase
TRINITY_DN150786_c0_g1_i5	0.042637618	1.12189264	coding	122-835[+]				
TRINITY_DN151033_c0_g2_i4	0.027154302	1.57670141	non-coding	40-420[-]				
TRINITY_DN151731_c3_g1_i7	0.049651906	1.113948885	coding	2-1417[+]	LOXH1_HUMAN	1.4E-13	LOXHD1	Lipoxygenase homology domain-containing protein 1
TRINITY_DN151731_c3_g1_i9	0.042637618	1.276173995	coding	2-1417[+]	LOXH1_HUMAN	1.05E-13	LOXHD1	Lipoxygenase homology domain-containing protein 1
TRINITY_DN151950_c2_g3_i11	0.040199145	1.230941755	coding	1-399[+]				
TRINITY_DN152166_c0_g1_i2	0.02635097	1.98371186	coding	217-1125[+]	PNPH_HUMAN	2.67E-100	PNP	Purine nucleoside phosphorylase
TRINITY_DN152358_c0_g1_i1	0.042480926	1.541413276	non-coding	108-635[-]	TL11_CHICK	2.3E-35	TL11	Tollid-like protein 1
TRINITY_DN152704_c0_g2_i1	0.003839525	3.157435079	coding	259-1449[-]	BHMT1_BOVIN	0	BHMT	Betaine-homocysteine S-methyltransferase 1
TRINITY_DN152735_c0_g1_i7	0.029200349	2.652276647	coding	61-939[-]	DIC_HUMAN	1.07E-110	SLC25A10	Mitochondrial dicarboxylate carrier
TRINITY_DN152950_c0_g1_i2	0.042637618	1.319490775	coding	82-1626[+]	CHIA_RAT	1.69E-148	Chia	Acidic mammalian chitinase
TRINITY_DN153121_c0_g2_i1	0.006419726	3.084193726	coding	141-1331[+]	BHMT1_BOVIN	0	BHMT	Betaine-homocysteine S-methyltransferase 1
TRINITY_DN153303_c0_g3_i1	0.041818452	1.413947888	coding	661-1162[-]	SBSPO_MOUSE	3.61E-07	Sbspon	Somatostatin-B and thrombospondin type-1 domain-containing protein
TRINITY_DN153303_c0_g3_i5	0.025596788	1.366811086	coding	568-1416[-]	SBSPO_HUMAN	4.26E-27	SBSPO	Somatostatin-B and thrombospondin type-1 domain-containing protein
TRINITY_DN153581_c0_g2_i2	0.042480926	1.805744365	coding	94-1566[+]	VDR_XENLA	3.8E-50	vdv	Vitamin D3 receptor
TRINITY_DN153945_c0_g3_i3	0.02635097	2.302249823	coding	221-1099[+]	DIC_HUMAN	3.84E-114	SLC25A10	Mitochondrial dicarboxylate carrier
TRINITY_DN154043_c0_g1_i1	0.022171906	1.161758876	coding	131-1753[+]	DH4A_PONPY	0	GLUD2	Glutamate dehydrogenase 2, mitochondrial
TRINITY_DN154555_c0_g1_i1	0.011518196	2.188233475	coding	117-1472[+]	PTN1_CHICK	1.03E-113	PTPN1	Tyrosine-protein phosphatase non-receptor type 1
TRINITY_DN154645_c0_g1_i6	0.042637618	2.495793164	coding	81-1349[+]	PUR6_CHICK	0	AIRC	Multifunctional protein ADE2
TRINITY_DN154690_c1_g1_i11	0.042480926	4.093641951	coding	43-459[-]	IOD3_XENLA	8.6E-37	dio3	Thyroxine 5-deiodinase
TRINITY_DN154690_c1_g1_i16	0.027154302	6.407688072	coding	2-409[-]				
TRINITY_DN154690_c1_g1_i6	0.042637618	4.486328221	coding					
TRINITY_DN154802_c0_g1_i7	0.023000595	1.957463848	coding	248-1054[-]				
TRINITY_DN155100_c1_g2_i11	0.027154302	1.644838888	coding	3-764[+]				
TRINITY_DN155100_c1_g2_i3	0.027154302	2.202976874	coding	2-898[+]				
TRINITY_DN155331_c0_g1_i3	0.048448131	4.230273735	coding	132-518[+]	IOD1_ORENI	5.83E-10	dio1	Type I iodothyronine deiodinase
TRINITY_DN155331_c0_g2_i6	0.042480926	3.982547414	coding					
TRINITY_DN155483_c0_g2_i1	0.019636382	1.771303412	coding	120-1763[-]	EAA3_RABIT	3.77E-160	SLC1A1	Excitatory amino acid transporter 3
TRINITY_DN155522_c0_g1_i6	0.042637618	1.934406572	coding	31-2088[-]				
TRINITY_DN155538_c1_g2_i4	0.037017376	1.026284243	coding	84-1148[-]	CH60A_DROME	0	Hsp60A	Heat shock protein 60A
TRINITY_DN155561_c0_g1_i2	0.042480926	1.920765495	coding	1-1596[-]	PCKGM_HUMAN	0	PCK2	Phosphoenolpyruvate carboxylase [GTP], mitochondrial
TRINITY_DN155665_c0_g2_i3	0.00622477	4.75030702	coding	1-1365[-]				
TRINITY_DN156117_c1_g1_i6	0.023063413	1.53561334	non-coding	82-825[-]	GLSL_MOUSE	4.84E-104	Gls2	Glutaminase liver isoform, mitochondria
TRINITY_DN156161_c0_g2_i2	0.042637618	2.084033048	coding	3-887[+]				
TRINITY_DN156204_c0_g1_i9	0.011575834	1.577842334	coding	309-1652[+]	RH8GB_XENLA	1.16E-152	rhbg-b	Ammonium transporter Rh type B-B
TRINITY_DN156223_c0_g2_i2	0.010708602	2.586333437	coding	1-1398[-]	AB22G_ARATH	3.22E-60	ABCG22	ABC transporter G family member 22
TRINITY_DN156451_c0_g2_i1	0.039234618	2.299358468	coding	1-495[+]	AB21G_ARATH	1.32E-22	ABCG21	ABC transporter G family member 21
TRINITY_DN156515_c0_g2_i1	0.027154302	2.011271031	coding	2-3373[+]				
TRINITY_DN156569_c0_g1_i3	0.032101349	1.399806596	coding	72-2474[+]	P5CS_HUMAN	0	ALDH18A1	Delta-1-pyrroline-5-carboxylate synthase
TRINITY_DN156783_c1_g2_i7	0.048447399	1.459730816	coding					
TRINITY_DN156887_c0_g1_i2	0.018670002	1.647213202	coding	291-1916[+]	S17A5_HUMAN	3.54E-137	SLC17A5	Sialin
TRINITY_DN157209_c0_g1_i1	0.011575834	1.974123543	coding	350-3562[+]	XDH_CHICK	0	XDH	Xanthine dehydrogenase/oxidase
TRINITY_DN157209_c0_g1_i3	0.003839525	5.297031104	coding	555-1895[-]				
TRINITY_DN157209_c1_g2_i1	0.048181466	4.481362931	coding	2-442[-]				
TRINITY_DN157209_c1_g4_i2	0.030946663	5.80110287	coding	1-399[-]				
TRINITY_DN157209_c1_g5_i1	0.027154302	6.408115289	coding					
TRINITY_DN157697_c0_g1_i1	0.010708602	5.283846682	coding	2-1861[+]				
TRINITY_DN157697_c0_g2_i2	0.003839525	4.887248388	coding	183-1325[+]	BMI1B_XENLA	6.16E-41	bmi1b	Polycomb complex protein BMI-1-B
TRINITY_DN2847_c0_g1_i1	0.040504243	5.995032406	non-coding					

(continued on next page)

Table 3 (continued)

Transcript ID	q value ^a	beta	Coding potential	Protein coordinates	UniProtKB ID	E-value	Gene	Protein
TRINITY_DN67135_c0_g1_j1	0.042637618	1.569265796	non-coding	2-433[-]	CO6A4_MOUSE	4.63E-12	Col6a4	Collagen alpha-4(VI) chain
TRINITY_DN72222_c0_g1_j1	0.048435074	1.700345252	non-coding					
TRINITY_DN76637_c0_g1_j1	0.042480926	2.592262612	non-coding					
TRINITY_DN81011_c1_g1_j1	0.042637618	1.466339824	coding					
TRINITY_DN88164_c0_g2_j1	0.003839525	5.40715304	coding					
TRINITY_DN117858_c0_g3_j2	0.040504243	-2.23776774	coding	2-451[+]				
TRINITY_DN136322_c1_g3_j2	0.042480926	-1.08528247	coding	174-1259[+]	LECM_THRJA	3.87E-10	N/A	C-type lectin lectosin-Thr1
TRINITY_DN129389_c0_g1_j5	0.049651906	-1.418789311	coding	165-590[-]	FABPM_SCHRG	5.29E-10	N/A	Fatty acid-binding protein, muscle
TRINITY_DN136164_c0_g1_j3	0.042480926	-1.439273969	coding	2-754[+]	GILT_CARAU	2.61E-63	ifl30	Gamma-interferon-inducible lysosomal thiol reductase
TRINITY_DN141551_c1_g1_j3	0.049651906	-1.65844818	coding	3-680[+]				
TRINITY_DN143759_c0_g2_j7	0.027154302	-1.429093207	coding					
TRINITY_DN144202_c0_g1_j1	0.027154302	-1.340042415	coding	185-1234[+]	CATL_DROME	9.84E-130	Cp1	Cathepsin L
TRINITY_DN144762_c0_g1_j1	0.042637618	-1.056300441	coding	445-1173[+]	CDQ_CAEER	3.63E-63	cdo-1	Cysteine dioxygenase
TRINITY_DN148815_c0_g1_j1	0.047149105	-1.889933188	coding	223-1743[+]	BDS1_YEAST	3.18E-55	BDS1	Alkyl/aryl-sulfatase BDS1
TRINITY_DN149083_c3_g5_j2	0.048447399	-1.176657241	coding	53-1243[+]	ASAH1_BOVIN	1.13E-151	ASAH1	Acid ceramidase
TRINITY_DN151389_c0_g1_j4	0.049651906	-3.140677669	coding	1-609[-]	CATL_DROME	3.82E-64	Cp1	Cathepsin L
TRINITY_DN151389_c1_g3_j7	0.027154302	-3.22980076	coding	1-951[-]	CATL2_HUMAN	1.22E-74	CTSV	Cathepsin L2
TRINITY_DN151656_c1_g1_j1	0.042637618	-1.125724656	coding	185-1186[+]	CATL1_HUMAN	2.96E-114	CTSL	Cathepsin L1
TRINITY_DN152891_c1_g2_j1	0.027154302	-1.42551057	coding	3-767[-]	FUCO_BRAF1	2.51E-122	BRAF1DRAFT_56888	Alpha-L-fucosidase
TRINITY_DN152937_c0_g2_j3	0.048181466	-1.622320003	coding	1-906[+]	GLNA_PANAR	1.8E-149	N/A	Glutamine synthetase
TRINITY_DN154069_c0_g1_j2	0.027154302	-1.316968665	coding	245-1600[-]	FUCO_BRAF1	0	BRAF1DRAFT_56888	Alpha-L-fucosidase
TRINITY_DN154732_c0_g3_j8	0.04951684	-1.045062189	coding	1-654[-]	NAGAB_CHECK	1.55E-54	NAGA	Alpha-N-acetylgalactosaminidase
TRINITY_DN154913_c1_g1_j2	0.042637618	-1.262050457	coding	1-540[-]				
TRINITY_DN155041_c0_g1_j1	0.029709192	-1.216549344	coding	234-1628[-]	FUCO_BRAF1	0	BRAF1DRAFT_56888	Alpha-L-fucosidase
TRINITY_DN156917_c1_g2_j7	0.027154302	-1.425794851	coding	3-1574[-]	BXL7_ARATH	7.14E-90	BXL7	Probable beta-D-xylosidase 7
TRINITY_DN156976_c1_g1_j1	0.042637618	-1.12538291	coding	123-1250[-]	BXL2_ARATH	1.74E-48	BXL2	Probable beta-D-xylosidase 2
TRINITY_DN157399_c3_g1_j2	0.041944616	-1.208770152	coding	1231-4809[+]	VWDE_MOUSE	3.83E-24	Vwde	von Willebrand factor D and EGF domain-containing protein
TRINITY_DN157557_c1_g3_j3	0.027154302	-1.39354781	coding	186-2144[-]	VWDE_HUMAN	4.34E-26	VWDE	von Willebrand factor D and EGF domain-containing protein
TRINITY_DN157560_c0_g1_j2	0.034103341	-1.137852965	coding	42-2267[+]	BXL7_ARATH	1.93E-107	BXL7	Probable beta-D-xylosidase 7
TRINITY_DN157676_c1_g1_j2	0.047747385	-1.473970682	coding	3-935[+]	VWDE_HUMAN	1.64E-06	VWDE	von Willebrand factor D and EGF domain-containing protein

^a q value from the Likelihood Ratio Test.

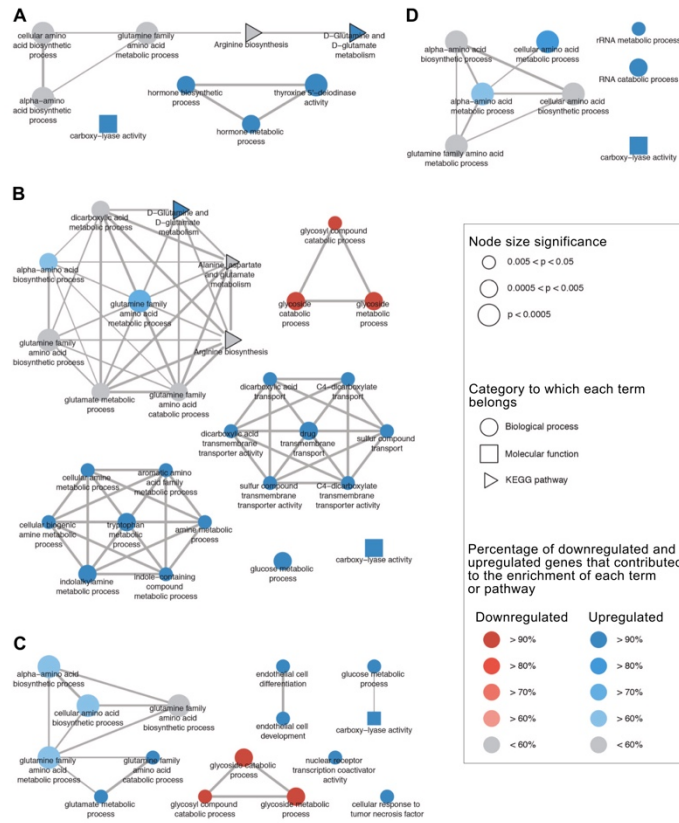


Fig. 5. Networks of enriched GO terms and KEGG pathways. Networks were constructed using the annotations of the differentially expressed gene orthologs for: A, *X. laevis*; B, *D. melanogaster*; C, *D. rerio*; and D, *C. gigas*. Nodes represent enriched terms or pathways (Two-sided hypergeometric test; $p < 0.05$). The width of the edges is defined by the Kappa score and represents the interrelations between two terms or pathways based on the number of genes they share. Since each term or pathway can be enriched by both downregulated and upregulated genes, the colour of the nodes indicate the percentage of genes from each cluster that contribute to the enrichment. Nodes in the red scale correspond to terms or pathways in which 60% or more of the genes that accounted for the enrichment were downregulated genes. Nodes in the blue scale correspond to terms or pathways in which 60% or more of the genes that accounted for the enrichment were upregulated genes. Nodes in gray correspond to terms or pathways in which both downregulated and upregulated genes contributed almost equally to the enrichment (between 60% and 40% each). Networks were obtained from the ClueGO app and redrawn with the igraph package of R (Csardi and Nepusz, 2006; R Core Team, 2020). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

proliferation in *D. rerio*, it also suppresses myoblast proliferation in *M. musculus* and induces oligodendrocyte differentiation in *R. norvegicus* with demyelinating lesions (de la Fuente et al., 2015; Han et al., 2019; Srikuea et al., 2012). In addition, we identified three upregulated genes that have been reported to be upregulated during regeneration in other annelid species: specifically, *BHMT*, *AIRC* or *paics* (Multifunctional protein ADE2), and *fbl* (rRNA 2'-O-methyltransferase fibrillarlin) (Myohara et al., 2006; Ribeiro et al., 2019).

Interestingly, glutamine/glutamate-metabolism related genes have been found to be differentially expressed during regeneration in other annelid species. Upregulation of the Glutamine synthetase coding gene has been reported in *E. japonensis* and *S. gracilis* (but not in *S. hystrix*), while *GLUD1* (Glutamate dehydrogenase 1, mitochondrial) upregulation has been reported in *E. japonensis* (Myohara et al., 2006; Niva et al., 2008; Ribeiro et al., 2019). In this work, we found downregulation of the Glutamine synthetase coding gene and upregulation of the *GLUD2* and *Gls2* orthologs. These contrasting differences in the expression of the Glutamine synthetase coding gene among these species deserve further analysis. However, based on our results we suggest that during early regeneration in *L. variegatus*, glutamine is degraded into glutamate, and further into α -ketoglutarate. While enrichment analyses indicate that the metabolism and synthesis of amino acids might be activated by these and other DEGs, another non-exclusive possibility is that glutamine and glutamate degradation is fueling the citric acid cycle through the

production of α -ketoglutarate, an intermediate of this metabolic pathway, as occurs under high proliferation conditions (Matés et al., 2019; Plaitakis et al., 2017).

A recent work has reported the expression of *vwde* in the blastemas of *Ambystoma mexicanum*, *D. rerio*, *Lepidosiren paradoxa*, and *Polypyterus senegalus*; and in the regenerating tail of *X. laevis* (Leigh et al., 2020). While the role of this gene is still unknown, knockdown experiments suggest that it is not involved in the formation of the blastema, but in the proximal-distal determination of its structure (Leigh et al., 2020). Our analysis showed that this gene ortholog is downregulated during early regeneration, while the blastema is beginning to form, but it would be interesting to analyze its expression in later stages.

It should be noted that, although our analysis allowed us to identify several differentially expressed transcripts, our experimental design has some limitations that must be considered. First, since we pooled in a single sample the worm fragments of the three RPs, it is likely that significant changes in the expression of some genes smoothed out across periods, making those genes unable to be recovered as differentially expressed. Thereby, only those genes whose expression was consistently down- or up-regulated within the three RPs, or whose change in expression was enough extreme in a determined RP, would be recovered as differentially expressed. Second, the sequenced mRNAs of the RG came from both regenerating and non-regenerating tissues (i.e., the blastema and the rest of the body, respectively), therefore it is

impossible to distinguish between regeneration-related and amputation-induced changes in expression. Third, since anterior and posterior regenerating fragments were also pooled in the same sample, the differentially expressed transcripts that we report here can be down- or up-regulated in either regenerating heads, regenerating tails, or both.

The enrichment analysis showed that the main BPs and Kegg pathways affected during early regeneration are those related to the metabolism of amino acids, glycosyl compounds, glucose, RNA, and hormones. Myohara et al. (2006) also showed that several enzymes related to the metabolism of carbohydrates were up and downregulated during the regeneration of *E. japonensis*. Based on the results of our study and the aforementioned study, we suggest that a metabolomic analysis of regeneration would be relevant to understand this phenomenon.

In this work, we analyzed the changes in gene expression that occur during early regeneration in *L. variegatus*. For this, we compared expression levels between a population of regenerating worms (24, 48, and 72 hpa) against a non-regenerating population. Our methodological approach required (for quantifying transcript abundances) a reference genome or a transcriptome that contained transcript sequences of genes involved in regeneration, either for this species or a closely related one. Since, to the best of our knowledge, there are no such references and we wanted to avoid losing information of genes that could be unique to *L. variegatus*, we conducted a *de novo* assembly and annotation of a transcriptome of the two populations. This transcriptome will be a resource that will help to elucidate the conserved molecular mechanisms underlying regeneration in *L. variegatus*. It is necessary to continue studies on *L. variegatus* regeneration to develop its potential as a model system for furthering our understanding of the cellular and molecular mechanisms of regeneration. Further work should focus on the role that DEGs and other candidate genes play in the regeneration of this worm from a more mechanistic approach, focusing mainly on the reestablishment of the nervous system.

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.genrep.2021.101050>.

Abbreviation

bp	base pairs
BP	Biological Processes
BUSCO	Benchmarking Universal Single-Copy Orthologs
CC	Cellular Components
cds	coding regions
DEG	Differentially Expressed Genes
GO	Gene Ontology
hpa	hours post amputation
ID	identifier
IS	Immune System
LRT	Likelihood Ratio Test
MF	Molecular Functions
NG	non-regenerating group
ORF	open reading frame
prd	path reinforcement distance
RG	regenerating group
RP	regenerating period
SRA	Sequence Read Archive
WT	Wald Test

CRedit authorship contribution statement

Aldo Arturo Telles-Garcia: Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft. **Roberto Álvarez-Martínez:** Formal analysis. **Juana María López-Martínez:** Investigation. **Fausto Arellano-Carbajal:** Conceptualization, Methodology, Resources, Writing – original draft, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by a Grant of FOFI-UAQ, Mexico (FNB201603). A.A.T.-G. was endowed with a graduate fellowship from CONACYT, Mexico (Scholarship number-CVU: 631374).

References

- Afgan, E., Baker, D., Batut, B., van den Beek, M., Bouvier, D., Čech, M., Chilton, J., Clements, D., Coraor, N., Grüning, B.A., Guerler, A., Hillman-Jackson, J., Hiltmann, S., Jallili, V., Rasche, H., Soranzo, N., Goecks, J., Taylor, J., Nekruteno, A., Blankenberg, D., 2018. The galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.* 46 (W1), W537–W544. <https://doi.org/10.1093/nar/gky379>.
- Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25 (17), 3389–3402. <https://doi.org/10.1093/nar/25.17.3389>.
- Andrews, S., 2010. FastQC: A Quality Control Tool for High Throughput Sequence Data. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
- Bely, A.E., 2014. Early events in annelid regeneration: a cellular perspective. *Integr. Comp. Biol.* 54 (4), 688–699. <https://doi.org/10.1093/icb/ictu109>.
- Bely, A.E., Nyberg, K.G., 2010. Evolution of animal regeneration: re-emergence of a field. *Trends Ecol. Evol.* 25 (3), 161–170. <https://doi.org/10.1016/j.tree.2009.08.005>.
- Bely, A.E., Sikes, J.M., 2010. Latent regeneration abilities persist following recent evolutionary loss in asexual annelids. *Proc. Natl. Acad. Sci.* 107 (4), 1464–1469. <https://doi.org/10.1073/pnas.0907931107>.
- Bhambri, A., Dhautia, N., Patel, S.S., Hardikar, M., Bhatt, A., Srikakulam, N., Shridhar, S., Vellarikkal, S., Pandey, R., Jayarajan, R., Verma, A., Kumar, V., Gautam, P., Khanna, Y., Khan, J.A., Fromm, B., Peterson, K.J., Scaria, V., Sivasubbu, S., Pillai, B., 2018. Large scale changes in the transcriptome of *Eisenia fetida* during regeneration. *PLoS One* 13 (9), e0204234.
- Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., Fridman, W.-H., Pagès, F., Trajanoski, Z., Galon, J., 2009. ClueGO: a Cytoscape plugin to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* 25 (8), 1091–1093. <https://doi.org/10.1093/bioinformatics/btp101>.
- Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30 (15), 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
- Bray, N.L., Pimentel, H., Melsted, P., Pachter, L., 2016. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* 34 (5), 525–527. <https://doi.org/10.1038/nbt.3519>.
- Bryant, D.M., Johnson, K., DiTommaso, T., Tickle, T., Couger, M.B., Payzin-Dogru, D., Lee, T.J., Leigh, N.D., Kuo, T.-H., Davis, F.G., Bateman, J., Bryant, S., Guzikowski, A. R., Tsai, S.L., Coyne, S., Ye, W.W., Freeman Jr., R.M., Peshkin, L., Tabin, C.J., Regev, A., Haas, B.J., White, J.L., 2017. A tissue-mapped axolotl *De novo* transcriptome enables identification of limb regeneration factors. *Cell Rep.* 18 (3), 762–776. <https://doi.org/10.1016/j.celrep.2016.12.063>.
- Camargo, A. P., Sourkov, V., Pereira, G. A. G., & Carazzolle, M. F. (2020). RNAsamba: neural network-based assessment of the protein-coding potential of RNA sequences. *NAR Genomics and Bioinformatics*, 2(1), lqz024. doi:<https://doi.org/10.1093/narg/ab/lqz024>.
- Cho, S.-J., Lee, M.S., Tak, E.S., Lee, E., Koh, K.S., Ahn, C.H., Park, S.C., 2009. Gene expression profile in the anterior regeneration of the earthworm using expressed sequence tags. *Biosci. Biotechnol. Biochem.* 73 (1), 29–34. <https://doi.org/10.1271/bbb.80391>.
- Csardi, G., Nepusz, T., 2006. The igraph software package for complex network research. *InterJournal, Complex Sy* 1695.
- Devos, N., Szövényi, P., Weston, D.J., Rothfels, C.J., Johnson, M.G., Shaw, A.J., 2016. Analyses of transcriptome sequences reveal multiple ancient large-scale duplication events in the ancestor of Sphagnopsida (Bryophyta). *New Phytol.* 211 (1), 300–318. <https://doi.org/10.1111/nph.13887>.
- Drewes, C.D., Fournier, C.R., 1990. Morphallaxis in an aquatic oligochaete, *Lumbriculus variegatus*: reorganization of escape reflexes in regenerating body fragments. *Dev. Biol.* 138 (1), 94–103. [https://doi.org/10.1016/0012-1606\(90\)90179-M](https://doi.org/10.1016/0012-1606(90)90179-M).
- El-Gebali, S., Mistry, J., Bateman, A., Eddy, S.R., Luciani, A., Potter, S.C., Qureshi, M., Richardson, L.J., Salazar, G.A., Smart, A., Sonnhammer, E.L.L., Hirsh, L., Paladin, L., Piovesan, D., Tosatto, S.C.E., Finn, R.D., 2019. The Pfam protein families database in 2019. *Nucleic Acids Res.* 47 (D1), D427–D432. <https://doi.org/10.1093/nar/gky995>.
- Finn, R.D., Clements, J., Eddy, S.R., 2011. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* 39 (suppl), W29–W37. <https://doi.org/10.1093/nar/gkr367>.

- Fu, L., Niu, B., Zhu, Z., Wu, S., Li, W., 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28 (23), 3150–3152. <https://doi.org/10.1093/bioinformatics/bts565>.
- de la Fuente, A.G., Errea, O., van Wijngaarden, P., Gonzalez, G.A., Kerinon, C., Jarjour, A.A., Lewis, H.J., Jones, C.A., Nait-Oumesmar, B., Zhao, C., Huang, J.K., ffrinch-Constant, C., & Franklin, R. J. M., 2015. Vitamin D receptor-retinoid X receptor heterodimer signaling regulates oligodendrocyte progenitor cell differentiation. *J. Cell Biol.* 211 (5), 975–985. <https://doi.org/10.1083/jcb.201505119>.
- Fukuda, A., Morris, J. P., & Hebrok, M. (2012). Bmi1 Is Required for Regeneration of the Exocrine Pancreas in Mice. *Gastroenterology*, 143(3), 821–831.e2. doi:<https://doi.org/10.1053/j.gastro.2012.05.009>.
- Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., MacManes, M.D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C.N., Henschel, R., LeDuc, R.D., Friedman, N., Regev, A., 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* 8 (8), 1494–1512. <https://doi.org/10.1038/nprot.2013.084>.
- Han, Y., Chen, A., Umansky, K.-B., Oonk, K. A., Choi, W.-Y., Dickson, A. L., Ou, J., Cigliola, V., Yifa, O., Cao, J., Tornini, V. A., Cox, B. D., Tzahor, E., & Poss, K. D. (2019). Vitamin D Stimulates Cardiomyocyte Proliferation and Controls Organ Size and Regeneration in Zebrafish. *Developmental Cell*, 48(6), 853–863.e5. doi:<https://doi.org/10.1016/j.devcel.2019.01.001>.
- Heikkinen, L.K., Kesäniemi, J.E., Knott, K.E., 2017. De novo transcriptome assembly and developmental mode specific gene expression of *Pygospio elegans*. *Evolution & Development* 19 (4–5), 205–217. <https://doi.org/10.1111/ede.12230>.
- de Jong, D.M., Seaver, E.C., 2016. A stable thoracic hox code and epimorphosis characterize posterior regeneration in *Capitella teleta*. *PLoS One* 11 (2), e0149724. <https://doi.org/10.1371/journal.pone.0149724>.
- Kanehisa, M., Goto, S., 2000. Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28 (1), 27–30. <https://doi.org/10.1093/nar/28.1.27>.
- Kozin, V.V., Kostyuchenko, R.P., 2015. *Vasa*, *Pl10*, and *Phwi* gene expression during caudal regeneration of the polychaete annelid *Alitta virens*. *Dev. Genes Evol.* 225 (3), 129–138. <https://doi.org/10.1007/s00427-015-0496-1>.
- Kriventseva, E.V., Kuznetsov, D., Tegenfeldt, F., Manni, M., Dias, R., Simão, F.A., Zdobnov, E.M., 2019. OrthoDB v10: sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs. *Nucleic Acids Res.* 47 (D1), D807–D811. <https://doi.org/10.1093/nar/kyk1053>.
- Krogh, A., Larsson, B., von Heijne, G., Sonnhammer, E.L., 2001. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *J. Mol. Biol.* 305 (3), 567–580. <https://doi.org/10.1006/jmbi.2000.4315>.
- Leigh, N.D., Sessa, S., Dragalzew, A.C., Payzin-Dogru, D., Sousa, J.F., Aggouras, A.N., Johnson, K., Dunlap, G.S., Haas, B.J., Levin, M., Schneider, I., Whited, J.L., 2020. von Willebrand factor D and EGF domains is an evolutionarily conserved and required feature of blastemas capable of multitissue appendage regeneration. *Evolution & Development* 22, 297–311. <https://doi.org/10.1111/ede.12332>.
- Lesiuk, N. M., & Drewes, C. D. (1999). Autotomy reflex in a freshwater oligochaete, *Lumbriculus variegatus* (Clitellata: Lumbriculidae). *Hydrobiologia*, 406(0), 253–261. doi:<https://doi.org/10.1023/A:1003756722019>.
- Li, W., Godzik, A., 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22 (13), 1658–1659. <https://doi.org/10.1093/bioinformatics/btl158>.
- Li, L., Zhai, S., Wang, L., Si, S., Wu, H., Chang, Z., 2014. *Hsp60* in caudal fin regeneration from *Parasigurnus dabryanus*: molecular cloning and expression characterization. *Fish & Shellfish Immunology* 36 (2), 401–408. <https://doi.org/10.1016/j.fsi.2013.12.016>.
- Litoff, E.J., Garriott, T.E., Ginjupalli, G.K., Butler, L., Gay, C., Scott, K., Baldwin, W.S., 2014. Annotation of the *Daphnia magna* nuclear receptors: comparison to *Daphnia pulex*. *Gene* 552 (1), 116–125. <https://doi.org/10.1016/j.gene.2014.09.024>.
- Lopez-Arribillaga, E., Rodilla, V., Pellegrinet, L., Guiu, J., Iglesias, M., Roman, A.C., Gutarra, S., González, S., Muñoz-Cánoves, P., Fernández-Salguero, P., Radtke, F., Bigas, A., Espinosa, L., 2015. Bmi1 regulates murine intestinal stem cell proliferation and self-renewal downstream of notch. *Development* 142 (1), 41–50. <https://doi.org/10.1242/dev.107714>.
- Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15 (550) <https://doi.org/10.1186/s13059-014-0550-8>.
- Makino, S., Whitehead, G.G., Lien, C.-L., Kim, S., Jhawar, P., Kono, A., Kawata, Y., Keating, M.T., 2005. Heat-shock protein 60 is required for blastema formation and maintenance during regeneration. *Proc. Natl. Acad. Sci.* 102 (41), 14599–14604. <https://doi.org/10.1073/pnas.0507408102>.
- Mates, J. M., Campos-Sandoval, J. A., Santos-Jiménez, J. de los, & Márquez, J. (2019). Dysregulation of glutaminase and glutamine synthetase in cancer. *Cancer Lett.*, 467, 29–39. doi:<https://doi.org/10.1016/j.canlet.2019.09.011>.
- Mehr, S., Verdes, A., DeSalle, R., Sparks, J., Pieribone, V., Gruber, D.F., 2015. Transcriptome sequencing and annotation of the polychaete *Hermodice carunculata* (Annelida, Amphinomidae). *BMC Genomics* 16, 445. <https://doi.org/10.1186/s12864-015-1565-6>.
- Morgulis, S., 1907. Observations and experiments on regeneration in *Lumbriculus*. *J. Exp. Zool.* 4 (4), 549–574. <https://doi.org/10.1002/jez.1400040405>.
- Myohara, M., Niva, C.C., Lee, J.M., 2006. Molecular approach to annelid regeneration: cDNA subtraction cloning reveals various novel genes that are upregulated during the large-scale regeneration of the oligochaete, *Enchytraeus japonensis*. *Dev. Dyn.* 235 (8), 2051–2070. <https://doi.org/10.1002/dvdy.20849>.
- Nielsen, H., 2017. Predicting secretory proteins with SignalP. In: Kihara, D. (Ed.), *Protein Function Prediction*. Humana Press, New York, NY, pp. 59–73. https://doi.org/10.1007/978-1-4939-7015-5_6.
- Niva, C.C., Lee, J.M., Myohara, M., 2008. Glutamine synthetase gene expression during the regeneration of the annelid *Enchytraeus japonensis*. *Dev. Genes Evol.* 218 (1), 39–46. <https://doi.org/10.1007/s00427-007-0198-4>.
- Nyberg, K.G., Conte, M.A., Kostyun, J.L., Forde, A., Bely, A.E., 2012. Transcriptome characterization via 454 pyrosequencing of the annelid *Pristina leidyi*, an emerging model for studying the evolution of regeneration. *BMC Genomics* 13 (1), 287. <https://doi.org/10.1186/1471-2164-13-287>.
- Özpolat, B.D., Bely, A.E., 2015. Gonad establishment during asexual reproduction in the annelid *Pristina leidyi*. *Dev. Biol.* 405 (1), 123–136. <https://doi.org/10.1016/j.ydbio.2015.06.001>.
- Özpolat, B.D., Bely, A.E., 2016. Developmental and molecular biology of annelid regeneration: a comparative review of recent studies. *Curr. Opin. Genet. Dev.* 40, 144–153. <https://doi.org/10.1016/j.cde.2016.07.010>.
- Pei, W., Tanaka, K., Huang, S.C., Xu, L., Liu, B., Sinclair, J., Idol, J., Varshney, G.K., Huang, H., Lin, S., Nussenblatt, R.B., Mori, R., Burgess, S.M., 2016. Extracellular HSP60 triggers tissue regeneration and wound healing by regulating inflammation and cell proliferation. *Npj Regenerative Medicine* 1 (1), 16013. <https://doi.org/10.1038/npregenmed.2016.13>.
- Pfeifer, K., Dorrestein, A.W.C., Fröbisch, A.C., 2012. Activation of *Hox* genes during caudal regeneration of the polychaete annelid *Platynereis dumerilii*. *Dev. Genes Evol.* 222 (3), 165–179. <https://doi.org/10.1007/s00427-012-0402-z>.
- Pfeifer, K., Schaub, C., Wolfstetter, G., Dorrestein, A., 2013. Identification and characterization of a *twist* ortholog in the polychaete annelid *Platynereis dumerilii* reveals mesodermal expression of *Pdu-twist*. *Dev. Genes Evol.* 223 (5), 319–328. <https://doi.org/10.1007/s00427-013-0448-6>.
- Pimentel, H., Bray, N.L., Puente, S., Melsted, P., Pachter, L., 2017. Differential analysis of RNA-seq incorporating quantification uncertainty. *Nat. Methods* 14 (7), 687–690. <https://doi.org/10.1038/nmeth.4324>.
- Plaitakis, A., Kalaf-Ezra, E., Kotzamani, D., Zaganas, I., Spanaki, C., 2017. The glutamate dehydrogenase pathway and its roles in cell and tissue biology in health and disease. *Biology* 6 (4), 11. <https://doi.org/10.3390/biology6010011>.
- Ponesakki, V., Paul, S., Mani, D.K.S., Rajendiran, V., Kanniah, P., Sivasubramanian, 2017. Annotation of nerve cord transcriptome in earthworm *Eisenia fetida*. *Genomics Data* 14, 91–105. <https://doi.org/10.1016/j.gdata.2017.10.002>.
- Ponz-Segrelles, G., Bleidorn, C., Aguado, M.T., 2018. Expression of *vasa*, *piwi*, and *nanos* during gametogenesis in *Tysoyllis antoni* (Annelida, Syllidae). *Evolution & Development* 20 (5), 132–145. <https://doi.org/10.1111/ede.12263>.
- R Core Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Randolph, H., 1892. The regeneration of the tail in *Lumbriculus*. *J. Morphol.* 7 (3), 317–344. <https://doi.org/10.1002/jmor.1050070304>.
- Reinisch, C., Kandutsch, S., Uthman, A., & Pammer, J. (2006). BMI-1: A Protein Expressed in Stem Cells, Specialized Cells and Tumors of the Gastrointestinal Tract. *HISTOLOGY AND HISTOPATHOLOGY*, 21(11), 1143–1149. doi:10.14670/HH-21.11.43.
- Ribeiro, R.P., Ponz-Segrelles, G., Bleidorn, C., Aguado, M.T., 2019. Comparative transcriptomics in Syllidae (Annelida) indicates that posterior regeneration and regular growth are comparable, while anterior regeneration is a distinct process. *BMC Genomics* 20 (1), 855. <https://doi.org/10.1186/s12864-019-6223-y>.
- Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., Ideker, T., 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13 (11), 2498–2504. <https://doi.org/10.1101/gr.1239303>.
- Simakov, O., Marletaz, F., Cho, S.J., Edsinger-Gonzales, E., Havlak, P., Hellsten, U., Kuo, D.H., Larsson, T., Lv, J., Arendt, D., Savage, R., Osogawa, K., de Jong, P., Grimwood, J., Chapman, J.A., Shapiro, H., Aerts, A., Otillar, R.P., Terry, A.Y., Boore, J.L., Grigoriev, I.V., Lindberg, D.R., Seaver, E.C., Weisblat, D.A., Putnam, N. H., Rokhsar, D.S., 2013. Insights into bilaterian evolution from three spiralian genomes. *Nature* 493, 526–531.
- Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., Zdobnov, E.M., 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31 (19), 3210–3212. <https://doi.org/10.1093/bioinformatics/btv351>.
- Soneson, C., Love, M. I., & Robinson, M. D. (2016). Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *FI000Research*, 4(1521).
- Srikuea, R., Zhang, X., Park-Sarge, O.-K., Esser, K.A., 2012. VDR and CYP27B1 are expressed in C2C12 cells and regenerating skeletal muscle: potential role in suppression of myoblast proliferation. *Am. J. Phys. Cell Phys.* 303 (4), C396–C405. <https://doi.org/10.1152/ajpcell.00014.2012>.
- Tadokoro, R., Sugio, M., Kutsuna, J., Tochinali, S., Takahashi, Y., 2006. Early segregation of germ and somatic lineages during gonadal regeneration in the annelid *Enchytraeus japonensis*. *Curr. Biol.* 16 (10), 1012–1017. <https://doi.org/10.1016/j.cub.2006.04.036>.
- Takeo, M., Yoshida-Noro, C., Tochinali, S., 2008. Morphallactic regeneration as revealed by region-specific gene expression in the digestive tract of *Enchytraeus japonensis* (Oligochaeta, Annelida). *Dev. Dyn.* 237 (5), 1284–1294. <https://doi.org/10.1002/dvdy.21518>.
- Takeo, M., Yoshida-Noro, C., Tochinali, S., 2010. Functional analysis of *grimp*, a novel gene required for mesodermal cell proliferation at an initial stage of regeneration in *Enchytraeus japonensis* (Enchytraeidae, Oligochaeta). *The International Journal of Developmental Biology* 54 (1), 151–160. <https://doi.org/10.1387/ijdb.082790mt>.

- The UniProt Consortium, 2019. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* 47 (D1), D506–D515. <https://doi.org/10.1093/nar/gky1049>.
- Thunders, M., Cavanagh, J., & Li, Y. (2017). De novo transcriptome assembly, functional annotation and differential gene expression analysis of juvenile and adult *E. fetida*, a model oligochaete used in ecotoxicological studies. *Biological Research* 50(1), 7. doi: <https://doi.org/10.1186/s40659-017-0114-y>.
- Thurmond, J., Goodman, J.L., Strelets, V.B., Attrill, H., Gramates, L.S., Marygold, S.J., Matthews, B.B., Millburn, G., Antonazzo, G., Trovisco, V., Kaufman, T.C., Calvi, B. R., FlyBase Consortium, 2019. FlyBase 2.0: the next generation. *Nucleic Acids Res.* 47 (D1), D759–D765. <https://doi.org/10.1093/nar/gky1003>.
- Tweeten, K.A., Anderson, A., 2008. Analysis of cell proliferation and migration during regeneration in *Lumbriculus variegatus* (Clitellata: Lumbriculidae). *BIOS* 79 (4), 183–190. <https://doi.org/10.1893/0005-3155-79.4.183>.
- Tweeten, K.A., Reiner, A., 2012. Characterization of serine proteases of *Lumbriculus variegatus* and their role in regeneration. *Invertebr. Biol.* 131 (4), 322–332. <https://doi.org/10.1111/ivb.12002>.
- Wickham, H., 2016. *ggplot2: Elegant Graphics for Data Analysis* (Springer-Verlag New York).
- Yates, A.D., Achuthan, P., Akanni, W., Allen, J., Allen, J., Alvarez-Jarreta, J., Amode, M. R., Armean, I.M., Azov, A.G., Bennett, R., Bhai, J., Billis, K., Boddu, S., Marugán, J. C., Cummins, C., Davidson, C., Dodiya, K., Fatima, R., Gall, A., Giron, C.G., Gil, L., Grego, T., Haggerty, L., Haskell, E., Hourlier, T., Izuogu, O.G., Janacek, S.H., Juetteman, T., Kay, M., Lavidas, I., Le, T., Lemos, D., Gonzalez Martinez, J., Maurel, T., McDowall, M., McMahon, A., Mohanan, S., Moore, B., Nuhn, M., Oheh, D.N., Parker, A., Parton, A., Patricio, M., Sakhivel, M.P., Abdul Salam, A.I., Schmitt, B.M., Schuilenburg, H., Sheppard, D., Sycheva, M., Szuba, M., Taylor, K., Thormann, A., Threadgold, G., Vullo, A., Walts, B., Winterbottom, A., Zadissa, A., Chakiachvili, M., Flint, B., Frankish, A., Hunt, S.E., Ilesley, G., Kostadima, M., Langridge, N., Loveland, J.E., Martin, F.J., Morales, J., Mudge, J.M., Muffato, M., Perry, E., Ruffier, M., Trevanion, S.J., Cunningham, F., Howe, K.L., Zerbino, D.R., Flicek, P., 2019. Ensembl 2020. *Nucleic Acids Res.* <https://doi.org/10.1093/nar/gkz966>.
- Zattara, E.E., Bely, A.E., 2016. Phylogenetic distribution of regeneration and asexual reproduction in Annelida: regeneration is ancestral and fission evolves in regenerative clades. *Invertebr. Biol.* 135 (4), 400–414. <https://doi.org/10.1111/ivb.12151>.

Anexo 2. Nomenclatura de genes, transcritos y proteínas.

En este trabajo, los nombres de genes y transcritos están escritos en letra cursiva, mientras que los nombres de proteínas están escritos en letra sin estilo y con la inicial en mayúscula. Con la finalidad de ubicar fácilmente la relación que hay entre los genes codificantes y sus productos, la primera vez que se menciona un gen aparecen enseguida y entre paréntesis otros nombres con los que se conoce al gen y el nombre de la proteína que codifica. Por ejemplo: *β -catenin* (*armadillo*, *Ctnnb1*; β -catenina). Sin embargo, en algunas ocasiones solo aparece el nombre de la proteína entre paréntesis, ya que solo se incluyeron los otros nombres del gen si se detectó que estos eran utilizados con frecuencia en las fuentes bibliográficas consultadas. Por su parte, la primera vez que se menciona una proteína aparece enseguida y entre paréntesis su símbolo o abreviación. Por ejemplo: Frizzled (Fz).

Anexo 3. Tabla de transcritos diferencialmente expresados con algunos estadísticos del análisis y anotaciones.

ID del transcrito	Valor de <i>q</i>	<i>beta</i>	Gen	Proteína
TRINITY_DN107 164_c0_g1_i2	0.0425	1.8365		
TRINITY_DN118 089_c0_g1_i1	0.0425	2.5852		
TRINITY_DN119 931_c0_g1_i1	0.0123	4.8588		
TRINITY_DN126 101_c0_g1_i1	0.0287	2.0817		
TRINITY_DN127 537_c0_g1_i1	0.0237	2.1962	<i>DER</i>	D-eritruosa reductasa
TRINITY_DN128 414_c0_g1_i7	0.0426	5.2486		
TRINITY_DN131 814_c0_g1_i1	0.0272	3.0122		
TRINITY_DN131 814_c0_g2_i2	0.0272	2.5585	<i>TNC</i>	Tenascina
TRINITY_DN132 073_c0_g2_i5	0.0298	2.3824		
TRINITY_DN132 086_c0_g2_i2	0.0309	2.6485	<i>Mrc1</i>	Receptor de manosa de macrófago 1
TRINITY_DN132 581_c0_g2_i7	0.0454	1.8619		
TRINITY_DN132 821_c1_g1_i1	0.0272	5.0668		
TRINITY_DN133 355_c0_g2_i1	0.0425	2.0463		
TRINITY_DN133 765_c0_g1_i1	0.0402	1.8219	<i>Drosha</i>	Ribonucleasa 3
TRINITY_DN134 295_c1_g1_i2	0.0315	1.1766		
TRINITY_DN135 002_c0_g1_i6	0.0405	1.1971		
TRINITY_DN135 551_c0_g2_i3	0.0426	3.8805		
TRINITY_DN135 612_c0_g1_i5	0.0426	4.0184		
TRINITY_DN136 715_c0_g3_i4	0.0038	2.7413		
TRINITY_DN137 232_c1_g1_i4	0.0070	2.2260	<i>sca</i>	Scabrous
TRINITY_DN137 442_c1_g1_i4	0.0272	1.8188		

TRINITY_DN138 956_c0_g2_i2	0.0482	1.6739	<i>Stard5</i>	Proteína de transferencia de lípidos relacionada a StAR
TRINITY_DN139 156_c0_g1_i1	0.0167	3.6399		
TRINITY_DN139 156_c1_g1_i2	0.0426	5.4531		
TRINITY_DN139 156_c1_g1_i7	0.0060	5.2154		
TRINITY_DN139 156_c1_g1_i8	0.0426	3.9079		
TRINITY_DN139 686_c0_g1_i2	0.0392	4.0581		
TRINITY_DN139 878_c0_g1_i4	0.0484	2.1152	<i>Xdh</i>	Xantina deshidrogenasa
TRINITY_DN140 112_c0_g4_i6	0.0426	5.4589		
TRINITY_DN140 818_c0_g1_i1	0.0425	1.7521	<i>rnf217</i>	Probable E3 ubiquitín ligasa RNF217
TRINITY_DN140 866_c0_g1_i1	0.0375	1.4671		
TRINITY_DN140 949_c0_g1_i1	0.0116	7.8825		
TRINITY_DN142 268_c0_g1_i1	0.0272	5.6198	<i>Cnot9</i>	Subunidad 9 del complejo de transcripciónCCR4-NOT
TRINITY_DN142 838_c1_g2_i3	0.0062	4.2222		
TRINITY_DN143 316_c1_g2_i2	0.0426	1.5620		
TRINITY_DN143 360_c1_g2_i12	0.0039	2.0963	<i>tdc</i>	L-tirosina descarboxilasa
TRINITY_DN143 360_c1_g2_i7	0.0426	3.5630	<i>tdc</i>	L-tirosina descarboxilasa
TRINITY_DN143 477_c1_g1_i3	0.0292	3.2993	<i>Pnliprp1</i>	Proteína inactiva relacionada con la lipasa pancreática 1
TRINITY_DN143 688_c3_g1_i1	0.0426	1.7454	<i>tdc</i>	L-tirosina descarboxilasa
TRINITY_DN143 865_c0_g1_i5	0.0477	1.7460		
TRINITY_DN144 159_c0_g1_i3	0.0116	2.3506	<i>ABCG22</i>	Miembro de la familia G del transportador ABC 22
TRINITY_DN144 384_c0_g2_i6	0.0426	2.6769	<i>cnmd</i>	Quimiotaxina derivada de células leucocitarias 1
TRINITY_DN145 539_c0_g1_i3	0.0272	1.0844	<i>fbl</i>	ARNr 2'-O-metiltransferasa fibrilarina
TRINITY_DN145 656_c0_g2_i3	0.0272	1.8030	<i>Ppib</i>	Peptidil-prolil cis-trans isomerasa B
TRINITY_DN146 187_c0_g1_i4	0.0426	1.4944		

TRINITY_DN146 805_c0_g1_i1	0.0272	1.1167		
TRINITY_DN147 249_c0_g1_i1	0.0364	2.1763	<i>Ubc2</i>	Enzima cinjugadora de ubiquitina E2-24 kDa
TRINITY_DN147 672_c0_g1_i3	0.0116	3.8634		
TRINITY_DN147 710_c0_g1_i1	0.0499	1.7063	<i>HDHD3</i>	Proteína contenedora de dominio de hidrolasa similar a la deshalogenasa de holoácidos 3
TRINITY_DN148 419_c1_g2_i1	0.0333	1.5696	<i>LOXHD1</i>	Proteína contenedora de dominio de homología de lipoxigenasa 1
TRINITY_DN149 785_c2_g1_i2	0.0230	2.6138	N/A	Isoforma de unión a manosa de la lectina tipo C
TRINITY_DN149 785_c2_g1_i4	0.0262	2.3016		
TRINITY_DN149 833_c0_g1_i11	0.0374	7.2548		
TRINITY_DN149 833_c0_g1_i4	0.0420	5.3969		
TRINITY_DN149 833_c0_g1_i5	0.0196	4.5446		
TRINITY_DN149 833_c0_g1_i6	0.0426	4.0170		
TRINITY_DN150 530_c0_g1_i3	0.0426	1.0355	<i>Kars1</i>	Lisina--ARNt ligasa
TRINITY_DN150 716_c0_g2_i9	0.0405	3.4336	<i>PNP</i>	Purina nucleósido fosforilasa
TRINITY_DN150 786_c0_g1_i5	0.0426	1.1219		
TRINITY_DN151 033_c0_g2_i4	0.0272	1.5767		
TRINITY_DN151 731_c3_g1_i7	0.0497	1.1139	<i>LOXHD1</i>	Proteína contenedora de dominio de homología de lipoxigenasa 1
TRINITY_DN151 731_c3_g1_i9	0.0426	1.2762	<i>LOXHD1</i>	Proteína contenedora de dominio de homología de lipoxigenasa 1
TRINITY_DN151 950_c2_g3_i11	0.0402	1.2309		
TRINITY_DN152 166_c0_g1_i2	0.0264	1.9834	<i>PNP</i>	Purina nucleósido fosforilasa
TRINITY_DN152 658_c0_g1_i1	0.0425	1.5414	<i>TLL1</i>	Proteína similar a Tolloid 1
TRINITY_DN152 704_c0_g2_i1	0.0038	3.1574	<i>BHMT</i>	Betaína--homocisteína S- metiltransferasa 1
TRINITY_DN152 735_c0_g1_i7	0.0292	2.6523	<i>SLC25A10</i>	Transportador mitocondrial de dicarboxilato
TRINITY_DN152 950_c0_g1_i2	0.0426	1.3195	<i>Chia</i>	Quitinasa ácida de mamíferos

TRINITY_DN153 121_c0_g2_i1	0.0064	3.0842	<i>BHMT</i>	Betaína--homocisteína S- metiltransferasa 1
TRINITY_DN153 303_c0_g3_i1	0.0418	1.4139	<i>Sbspon</i>	Proteína contenedora de dominios de somatomedina B y trombospondina tipo 1
TRINITY_DN153 303_c0_g3_i5	0.0256	1.3668	<i>SBSPON</i>	Proteína contenedora de dominios de somatomedina B y trombospondina tipo 1
TRINITY_DN153 581_c0_g2_i2	0.0425	1.8057	<i>vdr</i>	Receptor de vitamina D3
TRINITY_DN153 945_c0_g1_i3	0.0264	2.3022	<i>SLC25A10</i>	Transportador mitocondrial de dicarboxilato
TRINITY_DN154 043_c0_g1_i1	0.0222	1.1618	<i>GLUD2</i>	Glutamato deshidrogenasa 2, mitocondrial
TRINITY_DN154 555_c0_g1_i1	0.0115	2.1882	<i>PTPN1</i>	Proteína tirosina fosfatasa de tipo no receptor 1
TRINITY_DN154 645_c0_g1_i6	0.0426	2.4958	<i>AIRC</i>	Proteína multifuncional ADE2
TRINITY_DN154 690_c1_g1_i11	0.0425	4.0936	<i>dio3</i>	Yodotironina desyodasa de tipo III
TRINITY_DN154 690_c1_g1_i16	0.0272	6.4077		
TRINITY_DN154 690_c1_g1_i6	0.0426	4.4863		
TRINITY_DN154 802_c0_g1_i7	0.0230	1.9575		
TRINITY_DN155 100_c1_g2_i11	0.0272	1.6448		
TRINITY_DN155 100_c1_g2_i3	0.0272	2.2030		
TRINITY_DN155 331_c0_g1_i3	0.0484	4.2303	<i>dio1</i>	Yodotironina desyodasa de tipo I
TRINITY_DN155 331_c0_g2_i6	0.0425	3.9825		
TRINITY_DN155 483_c0_g2_i1	0.0196	1.7713	<i>SLC1A1</i>	Transportador de aminoácidos excitatorios 3
TRINITY_DN155 522_c2_g1_i6	0.0426	1.9344		
TRINITY_DN155 538_c1_g2_i4	0.0370	1.0263	<i>Hsp60A</i>	Proteína de choque térmico 60A
TRINITY_DN155 561_c0_g1_i2	0.0425	1.9208	<i>PCK2</i>	Fosfoenolpiruvato carboxiquinasa [GTP], mitocondrial
TRINITY_DN155 665_c0_g2_i3	0.0062	4.7503		
TRINITY_DN156 117_c1_g1_i6	0.0231	1.5356	<i>Gls2</i>	Isoforma hepática de glutaminasa, mitocondrial
TRINITY_DN156 161_c0_g3_i2	0.0426	2.0840		

TRINITY_DN156 204_c3_g1_i9	0.0116	1.5778	<i>rhbg-b</i>	Transportador de amonio Rh tipo B-B
TRINITY_DN156 223_c0_g2_i2	0.0107	2.5863	<i>ABCG22</i>	Miembro de la familia G del transportador ABC 22
TRINITY_DN156 451_c2_g2_i1	0.0392	2.2994	<i>ABCG21</i>	Miembro de la familia G del transportador ABC 21
TRINITY_DN156 515_c0_g2_i1	0.0272	2.0113		
TRINITY_DN156 569_c0_g1_i3	0.0321	1.3998	<i>ALDH18A1</i>	Delta-1-pirrolina-5-carboxilato sintasa
TRINITY_DN156 783_c1_g2_i7	0.0484	1.6597		
TRINITY_DN156 887_c1_g1_i2	0.0187	1.6472	<i>SLC17A5</i>	Sialina
TRINITY_DN157 200_c0_g1_i1	0.0116	1.9741	<i>XDH</i>	Xantina deshidrogenasa/oxidasa
TRINITY_DN157 209_c0_g1_i3	0.0038	5.2970		
TRINITY_DN157 209_c1_g2_i1	0.0482	4.4814		
TRINITY_DN157 209_c1_g4_i2	0.0309	5.8011		
TRINITY_DN157 209_c1_g5_i1	0.0272	6.4081		
TRINITY_DN157 697_c2_g1_i1	0.0107	5.2838		
TRINITY_DN157 697_c2_g2_i2	0.0038	4.8872	<i>bmi1b</i>	Proteína del complejo Polycomb BMI-1-B
TRINITY_DN284 7_c0_g1_i1	0.0405	5.9950		
TRINITY_DN671 35_c0_g1_i1	0.0426	1.5693	<i>Col6a4</i>	Cadena de colágeno alfa-4(VI)
TRINITY_DN722 22_c0_g1_i1	0.0484	1.7003		
TRINITY_DN766 37_c0_g1_i1	0.0425	2.5923		
TRINITY_DN810 11_c1_g1_i1	0.0426	1.4663		
TRINITY_DN881 64_c0_g2_i1	0.0038	5.4072		
TRINITY_DN117 858_c0_g3_i2	0.0405	-2.2378		
TRINITY_DN126 522_c1_g1_i2	0.0425	-1.0853	<i>N/A</i>	Lectina tipo C lectoxin-Thr1
TRINITY_DN129 389_c0_g1_i5	0.0497	-1.4188	<i>N/A</i>	Proteína de unión a ácidos grasos, músculo
TRINITY_DN136 164_c0_g1_i3	0.0425	-1.4393	<i>ifi30</i>	Tiol reductasa lisosomal inducible por interferón gamma

TRINITY_DN141 551_c1_g1_i3	0.0497	-1.6584		
TRINITY_DN143 759_c0_g2_i7	0.0272	-1.4291		
TRINITY_DN144 202_c0_g1_i1	0.0272	-1.3400	<i>Cp1</i>	Catepsina L
TRINITY_DN144 762_c0_g1_i1	0.0426	-1.0563	<i>cdo-1</i>	Cisteína dioxigenasa
TRINITY_DN148 815_c0_g1_i1	0.0471	-1.8899	<i>BDS1</i>	Alquil/arilsulfatasa BDS1
TRINITY_DN149 083_c3_g5_i2	0.0484	-1.1767	<i>ASAH1</i>	Ceramidasa ácida
TRINITY_DN151 389_c0_g1_i4	0.0497	-3.1407	<i>Cp1</i>	Catepsina L
TRINITY_DN151 389_c1_g1_i7	0.0272	-3.2298	<i>CTSV</i>	Catepsina L2
TRINITY_DN151 656_c1_g1_i1	0.0426	-1.1257	<i>CTSL</i>	Catepsina L1
TRINITY_DN152 861_c1_g2_i1	0.0272	-1.4255	<i>BRAFLDRA FT_56888</i>	Alfa-L-fucosidasa
TRINITY_DN152 937_c0_g2_i3	0.0482	-1.6223	<i>N/A</i>	Glutamina sintetasa
TRINITY_DN154 069_c0_g1_i2	0.0272	-1.3170	<i>BRAFLDRA FT_56888</i>	Alfa-L-fucosidasa
TRINITY_DN154 732_c0_g3_i8	0.0495	-1.0451	<i>NAGA</i>	Alfa-N-acetilgalactosaminidasa
TRINITY_DN154 913_c1_g1_i2	0.0426	-1.2621		
TRINITY_DN155 041_c0_g1_i1	0.0297	-1.2165	<i>BRAFLDRA FT_56888</i>	Alfa-L-fucosidasa
TRINITY_DN156 917_c1_g2_i7	0.0272	-1.4258	<i>BXL7</i>	Probable beta-D-xilosidasa 7
TRINITY_DN156 976_c1_g1_i1	0.0426	-1.1254	<i>BXL2</i>	Probable beta-D-xilosidasa 2
TRINITY_DN157 399_c3_g1_i2	0.0419	-1.2088	<i>Vwde</i>	Proteína contenedora del factor de von Willebrand D y dominio EGF
TRINITY_DN157 557_c1_g3_i3	0.0272	-1.3935	<i>VWDE</i>	Proteína contenedora del factor de von Willebrand D y dominio EGF
TRINITY_DN157 560_c1_g1_i2	0.0341	-1.1379	<i>BXL7</i>	Probable beta-D-xilosidasa 7
TRINITY_DN157 676_c4_g1_i2	0.0477	-1.4740	<i>VWDE</i>	Proteína contenedora del factor de von Willebrand D y dominio EGF