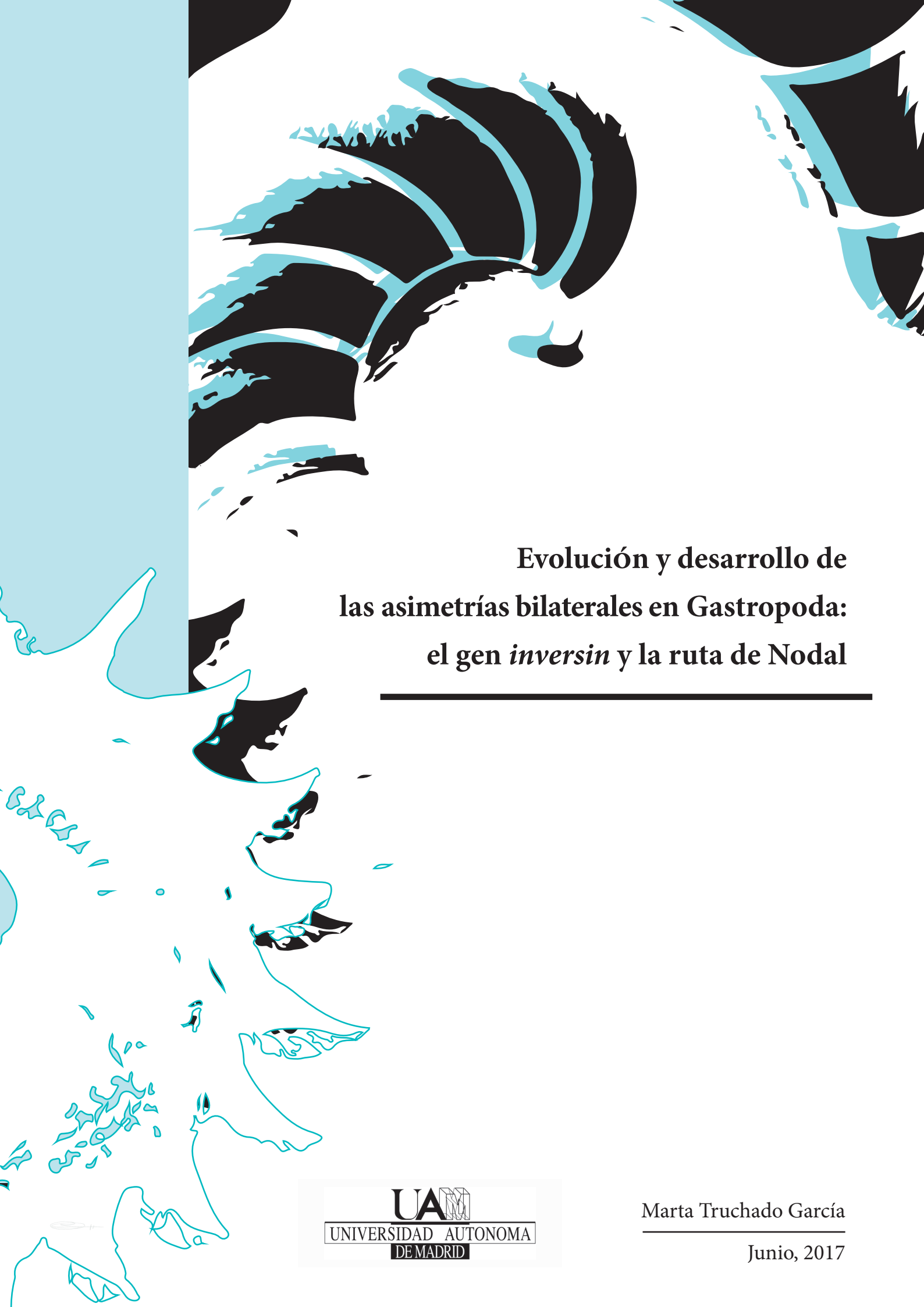


Evolución y desarrollo de las asimetrías bilaterales en Gastropoda: el gen *inversin* y la ruta de Nodal

Marta Truchado García

Junio, 2017

Universidad Autónoma de Madrid
Dpto. Biología Molecular

Evolución y desarrollo de las asimetrías bilaterales en Gastropoda: el gen *inversin* y la ruta de Nodal

Evolution and development of bilateral asymmetries in Gastropoda: the gene *inversin* and the Nodal pathway

Memoria presentada por la licenciada Marta Truchado García para optar al

grado de Doctor en Biociencias Moleculares por la Universidad Autónoma de Madrid

DOCTORANDO:

Marta Truchado García
Facultad de Biología
Departamento de Biología Molecular

DIRECTORA Y TUTORA ACADÉMICA:

Mª Cristina Grande Pardo
Facultad de Biología
Departamento de Biología

TESIS DOCTORAL
Madrid, 2017

Índice

Abreviaturas 7

Resumen 9

Abstract 11

1. Introducción 13

1.1. El concepto de Biodiversidad y los planes corporales en Metazoa 15

1.2. *Genetic toolkit* en genética del desarrollo y su evolución 17

1.2.1. Regulación selectiva de las herramientas génicas 18

1.2.2. Deleción, duplicación y divergencia de genes pre-existentes 18

1.2.3. Co-opción y novedades evolutivas 20

1.3. Relaciones filogenéticas dentro de Metazoa 21

1.4. Asimetrías izquierda/derecha (AID) en organismos bilaterales 24

1.4.1. Eventos tempranos de ruptura de simetría 26

1.4.1.1. Flujo direccional generado por los cilios del nodo de vertebrados 26

1.4.1.2. Flujo direccional de iones a través de bombas 28

1.4.1.3. Dinámica y estructura del citoesqueleto 28

1.4.1.4. La proteína Inversin y su papel en el establecimiento de AID 30

1.4.2. Expresión diferencial e identidad de lado: ruta de Nodal 32

1.4.2.1. La regulación de la ruta Nodal 33

1.4.2.2. EGF-CFC y la ruta Nodal 34

1.5. Spiralia y Gastropoda como sistema modelo en el estudio de AID 35

2. Objetivos 41

3. Materiales y métodos 45

3.1. *C. fornicata* como modelo de estudio 47

3.2. *B. glabrata* como modelo de estudio 50

3.3. Mantenimiento de adultos de las especies de gasterópodos estudiadas; obtención y cultivo de sus embriones 52

3.4. Fijación de embriones 53

3.5. Generación de bases de datos transcriptómicas para las especies *C. fornicata* y *B. glabrata* 54

3.6. Búsqueda de genes candidatos y establecimiento de ortología 55

3.7. Extracción de RNA y producción de cDNA 55

3.8. Amplificación por PCRs y electroforesis en gel de agarosa 56

3.9. Clonaje, transformación y comprobación de colonias positivas 56
3.10. Preparaciones de DNA y secuenciación 57
3.11. Transcripción *in vitro* para sondas e hibridación *in situ* 58
3.12. Selección y diseño de anticuerpos para las proteínas de estudio 58
3.13. Electroforesis en gel de acrilamida y *western blot* 59
3.14. Inmunohistoquímica en embriones completos 60
3.15. Mutagénesis: Generación del molde *wild type* (wt) 60
3.16. Mutagénesis: Generación de mutaciones sobre el molde wt 62
3.17. Transcripción *in vitro* de mRNA para microinyectar en embriones 63
3.18. Diseño y generación de CRISPRs 64
3.19. Diseño y generación de morfolinos 65
3.20. Microinyección de embriones: *C. fornicata* 66
3.21. Microinyecciones de embriones: *D. rerio* 67
3.22. Tratamientos con la proteína Activin en embriones de *C. fornicata* 68
3.23. Microscopía 68
4. Resultados 71
4.1. Caracterización del gen *inversin* en Metazoa y análisis funcional en Gastropoda 73
4.1.1. El gen *inversin* tiene un origen antiguo en la historia evolutiva de los animales 73
4.1.2. El gen *inversin* en *C. fornicata* y en *B. glabrata* 75
4.1.3. Activación espacio-temporal del gen *inversin* durante el desarrollo de *C. fornicata* 77
4.1.3.1. El transcrito de *inversin-L* se detecta durante todo el desarrollo de *C. fornicata* y se expresa de manera asimétrica desde la gastrulación temprana hasta larva 78
4.1.3.2. El transcrito de *inversin-S* se detecta en *C. fornicata* desde la segmentación hasta el estadio de larva y se expresa de manera asimétrica durante todo el desarrollo 78
4.1.4. Detección de la proteína Inversin en Gastropoda 81
4.1.5. Experimentos de sobre-expresión y falta de función del gen *inversin* en zigotos de *C. fornicata* 83
4.2. Caracterización y análisis funcional del gen *nodal* en *C. fornicata* 87

4.2.1. *C. fornicata* tiene dos parálogos del gen *nodal* 87
4.2.2. Activación espacio-temporal del gen *nodal* en *C. fornicata* 88
4.2.2.1. *nodal1* se expresa de manera simétrica en *C. fornicata* 88
4.2.2.2. *nodal2* se expresa de manera asimétrica en *C. fornicata* 88
4.2.3. Detección de la proteína Nodal2 en *C. fornicata* 91
4.2.4. Nodal2 regula el establecimiento de AID en *C. fornicata* 92
4.3. Caracterización de las proteínas EGF-CFC en Metazoa y análisis funcional en Gastropoda 99
4.3.1. Moluscos, anélidos y braquiópodos cuentan con proteínas EGF-CFC 99
4.3.2. Los transcritos sentido y antisentido del gen *egf-cfc* en *C. fornicata* están activos durante todo el desarrollo embrionario 101
4.3.3. La proteína EGF-CFC en *C. fornicata* se localiza en el labio dorsal durante la gastrulación 102
4.3.4. EGF-CFC está implicado en la gastrulación de gasterópodos 104
5. Discusión 111
5.1. La proteína Inversin en Metazoa 113
5.2. El ligando Nodal en *C. fornicata* 120
5.3. La proteína EGF-CFC en Gastropoda 127
5.4. Generación de herramientas y puesta a punto de diferentes metodologías para el estudio de organismos no modelo como los gasterópodos 136
6. Conclusiones (español e inglés) 139
7. Bibliografía 145
8. Apéndices 161
Apéndice 1: Créditos de las imágenes utilizadas en esta tesis 163
Apéndice 2: Bases de datos consultadas en las búsquedas de ortólogos en distintos Metazoa y en Coanoflagelata 165
Apéndice 3: Cebadores (primers) utilizados en esta tesis 167
Apéndice 4: Diseño de los anticuerpos específicos para las especies *C. fornicata* y *B. glabrata* 169
Apéndice 5: Vector utilizado en las construcciones para la generación de mRNA 171
Apéndice 6: Estrategia seguida para la generación de construcciones mutantes 173
Apéndice 7: Secuencias aminoacídicas de la proteína Inversin identificadas en distintas especies 175

Apéndice 8: Secuencias nucleotídicas completas de la región codificante de los genes *inversin, nodal1, nodal2* y *egf-cfc* de las especies *C. fornicata* y *B. glabrata* 189

Apéndice 9: Secuencias aminoacídicas de las proteínas de la familia EGF-CFC de las distintas especies analizadas en esta tesis 193

Apéndice 10: Alineamiento completo de las secuencias aminoacídicas de las proteínas EGF-CFC identificadas 197

Agradecimientos 201

AFSW: *Artificial filtered sea water*; agua marina filtrada artificial.

AID: Asimetrías izquierda-derecha.

BG: *Biomphalaria glabrata*.

CF: *Crepidula fornicata*.

CRISPR: *Clustered regularly interspaced short palindromic repeats*; repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas.

DN: Dominante negativo.

Dominio Ank: Dominio Ankirina.

Dominio D-Box: Dominio de destrucción.

Dominio IQ: Dominio de unión a Calmodulina.

EGF-CFC: Familia de proteínas poseedoras de un dominio tipo EGF y otro CFC.

GPI: Glicosilfosfatidilinositol.

***invL*:** Isoforma larga de *inversin.*

***invS*:** Isoforma corta de *inversin*.

Kb: Kilobase.

LRA: *Left-right asymmetries*.

MHCII: Complejo mayor de histocompatibilidad II.

MO: *Morpholino*.

NATs: *Natural antisense transcripts*; transcritos naturales antisentido.

NCBI: *National Center for Biotechnology Information.*

NLS: *Nuclear localization signal*; señal de localización nuclear.

Oep: *One-eyed pinhead.*

Pb: Pares de bases.

SI: Suero post-inoculación.

SP: *Signal peptide*; péptido señal.

SPI: Suero pre-inoculación.

TFrW: *Treated fresh water*; agua mineral tratada.

WT: *wild-type*; fenotipo silvestre.

ZF: *Danio rerio* o pez cebra.

ΔEGF: Deleción del dominio EGF (*epidermal growth factor*).

Los organismos bilaterados se caracterizan por presentar una línea media denominada eje bilateral, que divide su cuerpo en dos mitades: derecha e izquierda. Estas dos mitades son generalmente simétricas, aunque existen excepciones. La correcta organización y establecimiento de las asimetrías corporales son esenciales para el funcionamiento adecuado de los organismos. Los bilaterados se dividen en tres grandes grupos o superclados: deuteróstomos, espiralados y ecdisozoos, y las asimetrías corporales están presentes, con pocas excepciones, en todos ellos. Las asimetrías morfológicas comienzan a formarse muy tempranamente durante el desarrollo embrionario y continúan hasta el final del mismo. Los primeros eventos de disrupción del patrón de simetría son fundamentales, puesto que establecen las primeras bases del proceso de diferenciación. Los estudios realizados hasta el momento sugieren que diferentes grupos de metazoos presentan diferentes mecanismos de ruptura de simetría, sin que exista un mecanismo general. Tras esta ruptura inicial de la simetría tiene lugar una segunda fase a nivel génico, con la expresión asimétrica de cascadas génicas que amplifican la señal inicial. En esta segunda fase, la ruta clave en el establecimiento de la asimetría izquierda-derecha es la ruta de Nodal. El ligando Nodal pertenece a la familia TGF-β y su señalización activa una cascada génica exclusivamente en un lado del cuerpo dando origen a una organogénesis diferencial.

Los moluscos gasterópodos son organismos bilaterales espiralados cuyas asimetrías morfológicas han sido tradicionalmente descritas en detalle por morfólogos, aunque sólo recientemente se ha profundizado en el estudio de los mecanismos genéticos y moleculares subyacentes. Así, recientemente se ha demostrado que la ruta de Nodal está activa de forma asimétrica en relación con el eje bilateral de los gasterópodos, como ya se conocía en deuteróstomos. Sin embargo, hasta el momento se desconoce tanto su regulación como la cascada de genes activados por esta ruta en gasterópodos. Por otra parte, se desconocen los mecanismos que actúan en los primeros momentos del desarrollo de gasterópodos para romper la simetría inicial y comenzar la señalización diferencial. En esta tesis se han abordado estas dos cuestiones principales: cómo se produce la ruptura inicial de la simetría y cómo se activa, regula y transduce la señalización en gasterópodos.

En primer lugar, esta tesis se ha centrado en el estudio de la proteína Inversin, descrita exclusivamente en deuteróstomos y asociada al citoesqueleto, y cuyos mutantes en ratón muestran una completa reversión tanto de la activación de la ruta de Nodal como del plan corporal final. Los resultados han permitido identificar la proteína Inversin en una gran variedad de animales, lo

cual establece un origen antiguo para esta proteína, anterior al origen de los metazoos. Además, en esta tesis se presentan evidencias que apoyan una potencial implicación de Inversin en el establecimiento de asimetrías bilaterales en gasterópodos gracias a su activación asimétrica durante el desarrollo, así como la presencia de determinados dominios proteicos. Por otra parte, los datos aquí presentados indican que el ligando Nodal controla el establecimiento de asimetrías morfológicas en gasterópodos, puesto que tanto la sobreactivación como el bloqueo de su señalización producen organismos completamente simétricos. Finalmente, en esta tesis se ha identificado una de las proteínas reguladoras clave de la ruta de Nodal en vertebrados, el cofactor EGF-CFC. Esto constituye la primera evidencia de la presencia de este importante regulador en organismos no deuteróstomos. Además, los resultados indican que los transcritos sentido y antisentido del gen *egf-cfc* están activos durante todo el desarrollo embrionario en gasterópodos. Sin embargo, los experimentos realizados sobre su funcionalidad indican que las proteínas EFG-CFC no tienen un papel importante en el establecimiento de asimetrías izquierda-derecha en gasterópodos. Los resultados de esta tesis constituyen un avance en el conocimiento del establecimiento de asimetrías corporales en gasterópodos y su análisis en un contexto evolutivo permite plantear hipótesis e inferir la historia evolutiva de los mecanismos genéticos y moleculares responsables de generar la diversidad de asimetrías morfológicas que observamos en la naturaleza.

Abstract

Bilateral organisms are those that are symmetrical with respect to a midline, called the bilateral axis, which divides their bodies into left and right sides, although there are exceptions. The correct organization and establishment of body asymmetries are essential for the proper functioning of organisms. Bilaterians are divided into three major groups or superclades: deuterostomes, spiralians and ecdysozoans, and body asymmetries are present, with few exceptions, in all of them. Morphological asymmetries start forming very early during development and continue until its completion. The first events of symmetry breaking are essential since they are the first steps of the differentiation process. Previous studies suggest that different metazoan groups display different mechanisms of symmetry breaking, implying there is no universal mechanism of symmetry breaking. After this initial break, a second phase takes place at a genetic level, with the asymmetrical expression of genetic cascades that amplify the initial signal. In this second phase the key pathway for the establishment of the left/right asymmetries is the Nodal pathway. The ligand Nodal belongs to the TGF-β family and its signaling activates a genetic cascade only on one side of the body giving rise to differential organogenesis.

Gastropods mollusks are bilateral spiralian organisms whose morphological asymmetries have been traditionally described in detail by morphologists, although in depth studies on the underlying genetic and molecular mechanisms have only been recently conducted. It has recently been demonstrated that the Nodal pathway is active in an asymmetrical manner with respect to the bilateral axis in gastropods, as previously reported in deuterostomes. However, both its regulation and the target genes and pathways in gastropods are currently unknown. Similarly, the mechanisms that act in the first events of gastropod development for the initial symmetry break and the initiation of differential signaling, are also unknown. In this thesis, we addressed these two major questions: how is this initial symmetry break is produced, and how is this signaling activated, regulated and transduced, in gastropods.

First, this thesis focused on the study of Inversin, exclusively known in deuterostomes and associated with the cytoskeleton. Mouse Inversin mutants show a complete inversion of both the activation of the Nodal pathway and the final body plan. Our results have allowed identification of Inversin in a wide variety of animals, supporting an ancient origin for this protein, previous to the origin of metazoans. Furthermore, we present evidence that suggests that Inversin plays a role in the establishment of the bilateral asymmetries in gastropods due to its asymmetrical expression during the development, as well as the presence of certain

protein domains. On the other hand, we present data showing that the Nodal ligand controls the establishment of the morphological asymmetries in gastropods, because both the over-activation and the blocking of its signaling lead to completely symmetrical individuals. Finally, in this thesis we have identified one of the key regulatory proteins in the Nodal pathway in vertebrates, the co-receptor EGF-CFC. This is the first evidence of the presence of this important regulator in non-deuterostome organisms. Further, our results indicate that the sense and antisense transcripts of the *egf-cfc*- gene are active through the whole embryonic development in gastropods. However, functional experiments show that the EGF-CFC proteins do not play an important role in the establishment of left/right asymmetries in gastropods. This thesis advances our understanding of the establishment of body asymmetries in gastropods, and its analysis in an evolutionary framework provides new hypotheses to infer the evolutionary history of the genetic and molecular mechanisms that generate the diversity of morphological asymmetries that we observe in nature.

A mi madre

"To invent from nothing an animal that can exist (I mean to say that can physiologically grow, nourish itself, resist the environment and predators, and reproduce itself) is an almost imposible feat. It is a project that by far excedes our rational abilities and also that of our best computers: we still know too little about existing vital mechanisms to create other, even only on paper. In other words, evolution has proven itself to be enormously more intelligent than the best evolutionists. Every year that passes confirms the fact that the mechanisms of life are not exceptions to the laws of chemistry and physics, but at the same time the furrow which separates us from the ultimate comprehension of vital phenomena grows ever wider. It is not that problems are not solved and questions not answered, but every solved problem generates dozens of new ones and the process gives no sign of ending".

Primo Levi
"Inventing an Animal" (1985)
Other People's Trades
(tr. Raymond Rosenthal, 1989)

Introducción

El principal objetivo de la biología evolutiva es entender cómo se originó y diversificó la vida. Diferentes disciplinas como la biología molecular, la embriología, la genética, la genómica, la proteómica, la ecología y la epigenética, entre otras, tratan de proporcionar evidencias que ayuden a entender cómo se forman e interaccionan las biomoléculas y cómo las biomoléculas son responsables de producir la diversidad biológica (pasada y actual) que observamos.

En esta tesis se ha profundizado en el estudio de los mecanismos moleculares y genéticos que están implicados en la formación de caracteres morfológicos asimétricos en diferentes grupos animales. Estos datos, discutidos en un marco filogenético, nos ayudan a comprender cuál fue el origen de estos mecanismos y cómo han evolucionado a lo largo de la historia evolutiva de la vida.

A continuación se resume brevemente qué entendemos por diversidad biológica y cuáles son las hipótesis actuales que pretenden explicar su origen desde un punto de vista evolutivo. También se detalla cómo la diversidad morfológica es regulada por determinados genes durante el desarrollo embrionario y cómo las asimetrías morfológicas son un sistema ideal para estudiar la generación y formación de patrón en una estructura embrionaria. A lo largo de esta introducción se detalla qué son las asimetrías morfológicas, cómo se generan, qué procesos las regulan y cuál es su relevancia en el funcionamiento de los organismos. Asimismo, se presenta el sistema modelo escogido, Spiralia y en particular Gastropoda (caracoles), sus asimetrías y el estado actual del conocimiento sobre su formación y regulación.

1.1. El concepto de Biodiversidad y los planes corporales en Metazoa

Desde que el término "**Biodiversidad**" apareciese por primera vez como título de una conferencia de Walter G. Rosen en 1986 ("BioDiversity"), su uso, lejos de quedar obsoleto, se ha extendido y arraigado como uno de los conceptos más importantes del campo de la Biología. El concepto "Biodiversidad" procede de la unión de las palabras "diversidad biológica" y se refiere a la amplia variedad de seres vivos que encontramos en la naturaleza, incluyendo los ecosistemas de los que forman parte y los recursos genéticos que en ellos se hallan. Los planes corporales o arquetipos que observamos en la naturaleza son producto de millones de años de evolución, en los cuales han ido surgiendo innumerables formas que difieren dramáticamente unas de otras (figura 1).

La determinación del plan corporal tiene lugar durante el desarrollo embrionario a través de un complejo programa génico que define el diseño final. Los estudios realizados en

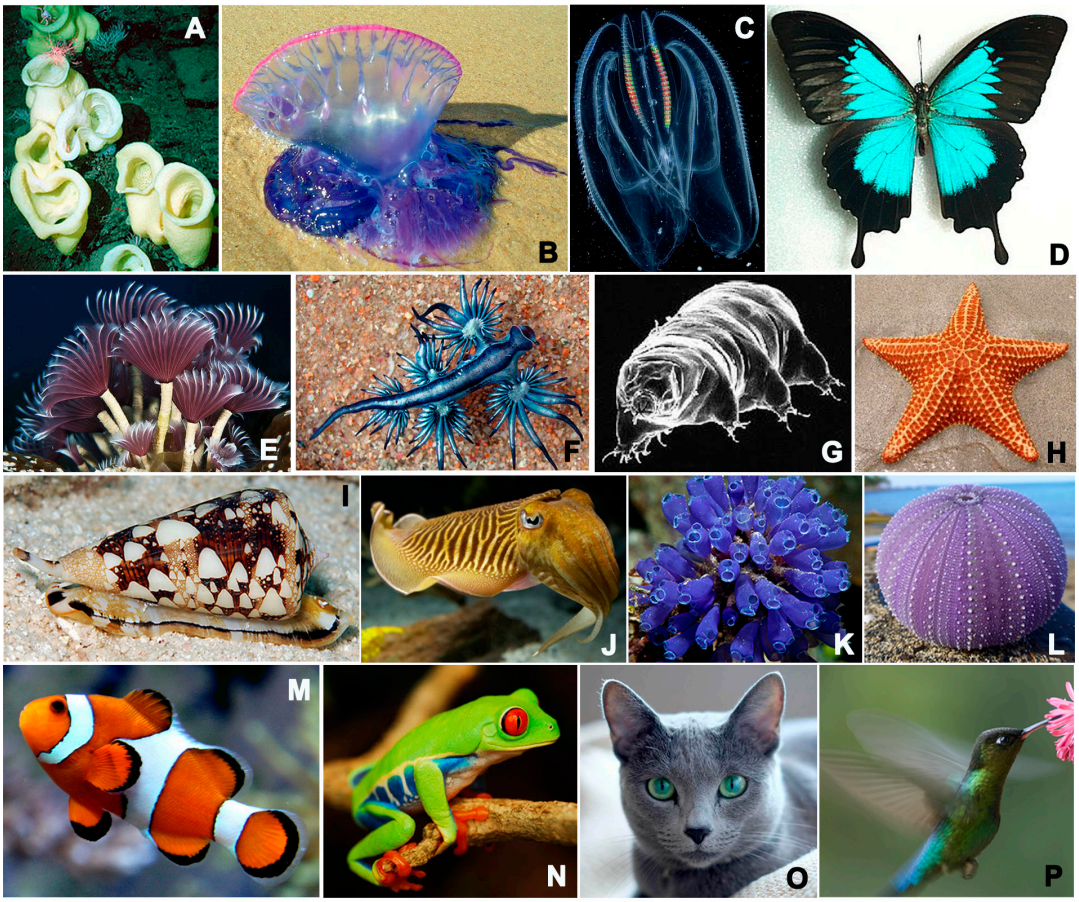


Figura 1. Biodiversidad y planes corporales en Metazoa. Los metazoos comprenden un grupo de organismos con una gran diversidad de planes corporales A) Esponja, *Staurocalyptus sp*; B) Cnidario, *Physalia physalis*; C) Ctenóforo, *Mnemiopsis leidyi;* D) Lepidóptero, *Ulysses telegonus;* E) Poliqueto, *Bispira brunnea*; F) Nudibranquio, *Glaucus atlanticus*; G) Tardígrado, *Hypsibius dujardini*; H) Estrella de mar, *Oreaster reticulatus*; I) Gasterópodo, *Conus ammiralis*; J) Cefalópodo, *Sepia officinalis*; K) Ascidia, *Clavelina moluccensis*; L) Erizo de mar, *Strongylocentrotus purpuratus*; M) Pez payaso, *Amphiprion ocellaris*; N) Rana arborícola, *Agalychnis callidryas*; O) Gato, *Felis catus*; P) Colibrí, *Eupherusa cyanophrys* (créditos de cada imagen en el anexo 1).

los últimos años en el área de la biología del desarrollo han llevado a grandes avances tanto en la identificación de genes importantes involucrados en el programa de desarrollo, como en la comprensión de su relevancia dentro de la determinación de estos diseños. En las últimas dos décadas hemos sido testigos del descubrimiento de uno de los secretos más desconcertantes de la biología del desarrollo; todos los animales, pese a los diferentes planes corporales que presentan, comparten la gran mayoría de las familias génicas, rutas y redes de señalización que regulan los principales eventos del programa de desarrollo; lo que se ha denominado ***genetic toolkit*** o "caja de herramientas génicas" (Cheatle-Jarvela & Pick, 2016). Esta disparidad de morfologías originadas a partir del mismo kit de piezas y herramientas suscita preguntas en diferentes disciplinas originando controversias y debates y promoviendo la propuesta de diferentes mecanismos no excluyentes entre sí para explicar esta cuestión (Cheatle-Jarvela & Pick, 2016).

Por otra parte, el desarrollo de la **genómica** ha permitido realizar estudios comparados entre los distintos filos o grupos de animales y desarrollar el campo de la filogenómica (revisado en Delsuc *et al*., 2005; Telford *et al*., 2015; Shen *et al*., 2017). Los avances que se han ido sucediendo en esta dirección han contribuido no sólo a responder algunas cuestiones tradicionales en cuanto a las relaciones evolutivas entre grupos sino a generar aún más preguntas: ¿Cómo surgieron las diferentes herramientas genéticas? ¿Cuáles eran sus funciones ancestrales y cómo se han ido modulando? ¿De qué antecesor heredamos determinados programas génicos? ¿Qué función poseen estos programas en los distintos grupos animales? Estas aproximaciones han hecho posible ir descifrando cómo modificaciones en la regulación genética generan la diversidad morfológica actual y pasada, así como algunas de las tendencias evolutivas observadas en la diversidad animal.

Bajo este marco ontogénico, molecular y evolutivo nació una nueva disciplina que se conoce como **biología evolutiva del desarrollo** (EvoDevo, del inglés *evolution and development*) (Müller, 2007), que trata de explicar la aparición de nuevas formas de vida a través de la modificación de los mecanismos génicos del desarrollo y su regulación y es asimismo el marco en el que se encuadra esta tesis.

1.2. *Genetic toolkit* en genética del desarrollo y su evolución

De todos los genes con los que cuentan los organismos, sólo una pequeña parte controla la formación de los planes corporales y la morfología. Durante el desarrollo, la mayoría de los procesos que ocurren están controlados por dos tipos de productos génicos: **factores de transcripción**, que son familias génicas que regulan la expresión de otros genes, y miembros de **rutas de señalización** que intervienen en la comunicación celular. Por tanto, para entender cómo se genera la diversidad debemos entender cómo son y cómo se regulan estos factores de transcripción y estas rutas de señalización.

Cambios en la secuencia de los genes que codifican para estas proteínas pueden potencialmente originar cambios evolutivos. Sin embargo, aunque se han descrito casos puntuales (Galant & Carroll, 2002; Ronshaugen *et al*., 2002), esto no parece corresponder con la norma general. Por otro lado, se ha comprobado que secuencias para un mismo gen provenientes de organismos muy divergentes y que cuentan con multitud de cambios entre sí, comparten sin embargo las mismas funciones bioquímicas y sus productos son capaces de actuar de la misma manera durante el desarrollo embrionario (Rincón-Limas *et al*., 1999). No

obstante, hay otras fuentes potenciales de cambio que parecen haber tenido mayor peso en los cambios evolutivos. Por ejemplo, **variaciones en la regulación** de estos productos génicos pueden ocasionar modificaciones en el establecimiento de planes corporales, en la identidad de regiones y en la formación de estructuras (Sanetra *et al*., 2005; Keyte *et al*., 2014; Keyte & Smith, 2014). Asimismo, la **deleción** y la **duplicación y divergencia** de estas herramientas génicas parecen haber sido fuente importante de cambio evolutivo (Zhang, 2003; revisado en Mallo *et al*., 2010 y Pick, 2016). En otras ocasiones, los genes existentes son incluidos en otros **contextos** para generar nuevas estructuras que pueden ser seleccionadas y favorecer el cambio evolutivo (Carrol *et al*., 1994; Monteiro, 2015).

1.2.1. Regulación selectiva de las herramientas génicas

La activación y/o represión de determinados genes en unas regiones y momentos concretos durante el desarrollo del embrión crea diferentes ambientes o contextos celulares en los que distintas rutas se encuentran e interactúan (Zecca *et al*., 1995; Cheatle-Jarvela & Pick, 2016). Asimismo, células que van adquiriendo cierta identidad comienzan a secretar morfógenos (moléculas que difunden en el espacio extracelular y que inducen respuestas en las células vecinas de acuerdo a su concentración local y a la sensibilidad o competencia de aquellas células que los detectan). Un ejemplo que ilustra la importancia de los contextos celulares y los gradientes es el establecimiento del eje antero-posterior en el embrión de *Drosophila melanogaster.* Las proteínas codificadas por los genes *bicoid* y *nanos* generan gradientes opuestos en el embrión; el primero se localiza en la zona más anterior y el segundo, en la zona posterior. Estos factores de transcripción que actúan como morfógenos activan genes de forma selectiva en las regiones en las que se hallan. La sobre-expresión de estos genes en otras zonas donde normalmente no se encuentran cambia el contexto natural y ocasiona cambios en la identidad de regiones y la generación ectópica de estructuras anteriores o posteriores (Driever & Nüsslein-Volhard, 1989; revisado en Rosenberg *et al*., 2009). Por tanto, una respuesta en particular depende del entorno, marco o contexto en el que se esté produciendo la señalización.

1.2.2. Deleción, duplicación y divergencia de genes pre-existentes

El genoma de un organismo es dinámico. En ocasiones, determinados genes pueden duplicarse o eliminarse, lo que modifica la información existente en el genoma y la complejidad real y potencial de los programas génicos (Bennett, 2004). La duplicación de genes se puede

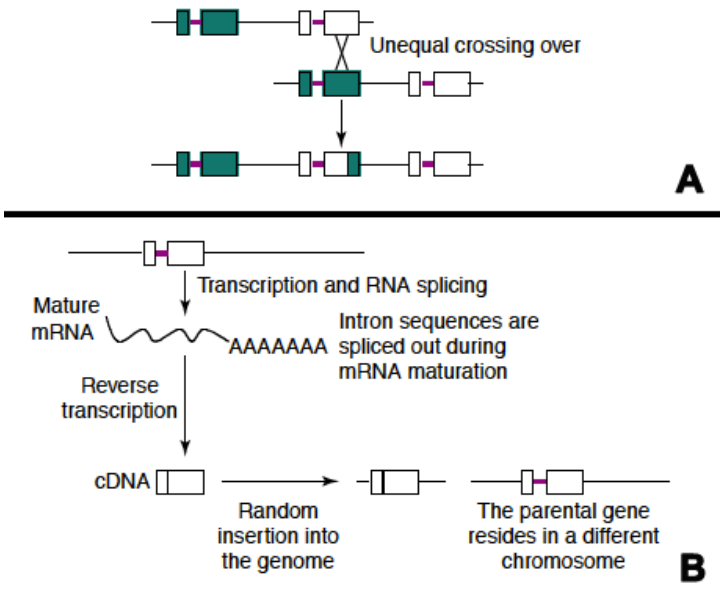


Figura 2. Mecanismos moleculares para la adquisición de nuevos genes a través de la traslocación, deleción, duplicación y divergencia de genes pre-existentes. A) Los entrecruzamientos desiguales entre cromosomas producen inserciones, deleciones y duplicaciones génicas; B) la retrotranscripción de un RNA mensajero maduro que se inserta de forma azarosa en el genoma produce que se duplique un gen existente y que éste se localice en una región y/o cromosoma diferente. Figura tomada y modificada de Zhang (2003).

producir por distintos mecanismos como el entrecruzamiento desigual y la retrotransposición (figura 2). Estos procesos pueden afectar a genes individuales y/o grupos de genes y/o zonas codificantes de determinados dominios proteicos. Por otra parte, también pueden producirse duplicaciones cromosómicas a gran escala o incluso fenómenos de poliploidización (Zhang, 2003; Knytl *et al*., 2017). Las copias generadas por duplicación pueden conservar la función original, proporcionando redundancia y sirviendo como reserva ante mutaciones deletéreas o la pérdida de una de las copias (Long *et al*., 2003). Sin embargo, también pueden incrementar la dosis génica que pudiera ser necesaria en momentos de gran actividad celular (Ohno, 1970). Esto ocurre por ejemplo con los genes codificantes para las histonas o RNAs ribosómicos activos en momentos como la división celular y la transcripción y traducción de material genético (Hurst & Smith, 1998).

Por otro lado, en ocasiones esa duplicación va seguida de una divergencia, lo que lleva a la creación de nuevas familias génicas. Todos los genes de una misma familia génica provienen de un único gen ancestral. Se denominan genes ortólogos aquellos genes de una familia génica presentes en diferentes organismos que provienen de un único gen ancestral. Por otra parte, se conoce como genes parálogos aquéllos que provienen de eventos de duplicación dentro de un mismo genoma. En ocasiones, los genes ortólogos o parálogos se pierden por deleciones cromosómicas, o incluso funcionalmente pueden perder actividad por mutaciones tanto en su secuencia como en sus zonas reguladoras.

El ejemplo clásico para ilustrar este fenómeno son los genes *Hox*, que codifican para factores de transcripción que dan identidad a los segmentos en los que se expresan a lo largo del eje antero-posterior. La hipótesis que pretende explicar la aparición de los complejos *Hox* sugiere que habría un gen *Hox* ancestral que habría sufrido varios procesos de cis-duplicación

(hasta 9 eventos). Posteriormente, se habrían sucedido procesos tanto de pérdida como de trans-duplicación del complejo que darían lugar a los 4 grupos que se observan, por ejemplo, en vertebrados (revisado en Mallo *et al*., 2010 y Cheatle-Jarvela & Pick, 2016)

1.2.3. Co-opción y novedades evolutivas

La co-opción es el proceso por el que zonas reguladoras o genes o grupos de genes relacionados y regulados de forma conjunta son incluidos en otros contextos para generar nuevas estructuras. Este comportamiento modular establece como principal fuente de innovación el uso de bloques de formas pre-existentes más que la creación de nuevos elementos. La modificación de las interacciones entre módulos genera nuevas redes, que pueden a su vez modificar el nivel de organización. Así pues, la co-opción es un poderoso mecanismo para incrementar el conjunto de los módulos que gobiernan el desarrollo.

Se han descrito dos mecanismos génicos principales mediante los cuales puede ocurrir la co-opción: la adquisición de secuencias reguladoras (elementos reguladores en *cis*; CRE) y la ganancia de factores de transcripción a nivel transcripcional. La inclusión de CREs en nuevas regiones provoca que los genes situados *downstream* (o aguas abajo) se incluyan en redes ya existentes al ser ahora competentes, con la consiguiente modificación funcional. Con la ganancia de factores de transcripción *upstream* (o aguas arriba) de determinados genes se puede promover la activación de los mismos en nuevos dominios (Sanetra *et al*., 2005).

La co-opción de genes muestra que la función de un gen o una ruta no es rígida sino que los genes pueden ser reclutados en otros contextos para generar estructuras no esperadas ni relacionadas en absoluto con un proceso ya conocido en el que un gen/ruta actúe. Este mecanismo dota de gran versatilidad y plasticidad al sistema, al mismo tiempo que produce un gran ahorro génico. De hecho, algunas innovaciones estructurales tales como la aparición de las alas en los insectos, las modificaciones en los apéndices de los tetrápodos o la aparición de la cresta neural, entre otras, son el resultado del "reciclaje génico" que han sufrido determinados grupos de genes ya existentes al incluirse en nuevas redes, y que han permitido la especificación de nuevas estructuras sin la necesidad de crear sistemas *de novo* (revisado en True & Carrol, 2002).

1.3. Relaciones filogenéticas dentro de Metazoa

El establecimiento de las relaciones filogenéticas entre los distintos grupos de metazoos ha atravesado distintas etapas y se han propuesto numerosas topologías para explicar el camino que siguió la evolución en la generación de dichos grupos. Con el desarrollo de la biología molecular, el estudio de marcadores génicos en los distintos grupos potenció los campos de la sistemática molecular y la filogenómica. Los datos generados en este sentido junto con las clasificaciones clásicas basadas en morfología han mejorado algunas hipótesis existentes haciéndolas más robustas, pero han creado conflictos en otras ramas porque han resultado muy dependientes de los datos y los taxones escogidos. A día de hoy existen nodos muy consistentes y aceptados aunque siguen existiendo inconsistencias y divergencias en otros puntos del árbol de la vida. La gran mayoría de estos nodos inciertos se halla en el origen de los metazoos (Halanych, 2004; Srivastava *et al*., 2008; Schierwater *et al*., 2009; Nosenko *et al*., 2013; Ryan *et al*., 2013; Maldonado, 2012; Taylon-Canon *et al*., 2016) (figura 3).

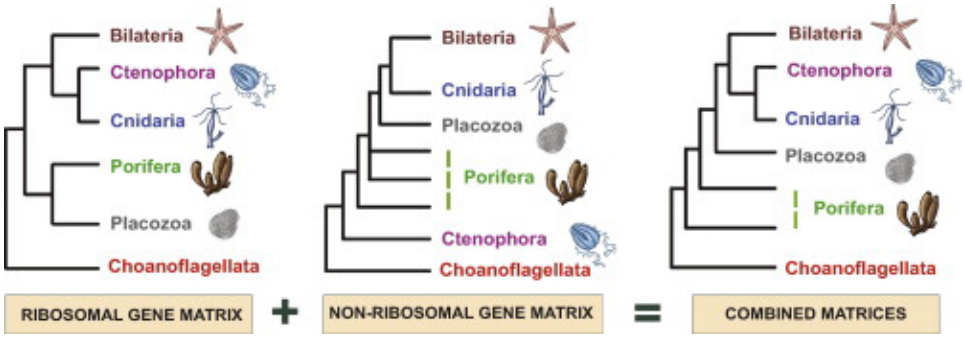


Figura 3. Relaciones filogenéticas en el origen de Metazoa. Varios análisis utilizando diferentes tipos de datos moleculares y taxones dan lugar a diferentes hipótesis filogenéticas inconsistentes entre sí pero todas con alto apoyo estadístico. Se han detectado eventos de saturación y de atracción de ramas largas para determinados genes que dificultan la inferencia filogenética del origen de Metazoa. Figura tomada de Nosenko *et al*. (2013).

Uno de los nodos más sólidos es el que determina que los **coanoflagelados** y los **metazoos** comparten un antecesor común y por tanto forman un grupo monofilético (Carr *et al*., 2008) (figura 4). Los coanoflagelados son organismos unicelulares que recuerdan a los coanocitos, las células características de las esponjas (Porifera). Por su parte, los metazoos incluyen a todos los animales pluricelulares en los que se llevan a cabo procesos complejos como por ejemplo la diferenciación celular que permite la especialización de tejidos para tareas concretas. Uno de los puntos más discutidos hace referencia al origen de los metazoos. Las relaciones entre Porifera, Placozoa, Cnidaria, Ctenophora y el resto de metazoos ha suscitado un gran debate que aún

parece lejos de resolverse (figura 3). Las señales filogenéticas más sólidas parecen considerar a los **poríferos**, las esponjas, como el grupo hermano del resto de metazoos (Srivastava *et al*., 2008; Taylor-Cannon *et al*., 2016). Un grupo de organismos especialmente enigmático lo componen los **placozoos**, animales pequeños en forma de disco sin ejes claramente establecidos, aunque existe una polaridad de tipo dorso-ventral. Los placozoos también parecen haber divergido muy pronto en la historia evolutiva de los metazoos (figura 4). La aparente falta de células musculares o nerviosas, de gastrulación y de otros procesos complejos han llevado tradicionalmente a considerar que tanto poríferos como placozoos presentan caracteres ancestrales (Shierwater *et al*., 2008). Sin embargo, esta afirmación está actualmente bajo debate (Maldonado, 2012). Por otra parte, cnidarios y ctenóforos podrían haber divergido después de poríferos y placozoos, aunque las relaciones entre ellos y con el resto de metazoos es muy controvertida (Ryan *et al*., 2013) (figura 3).

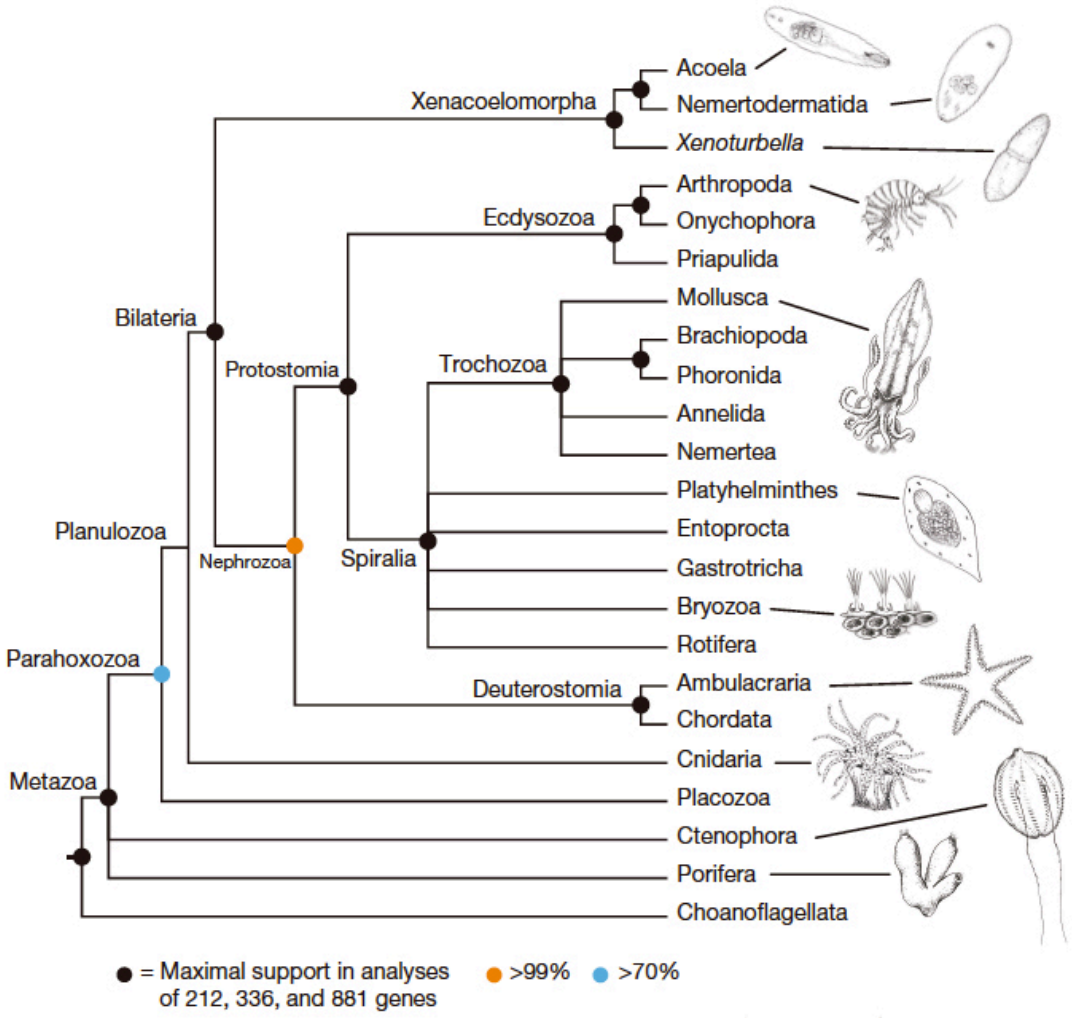


Figura 4. Árbol consenso de las relaciones filogenéticas entre los grupos de metazoos. Choanoflagellata es el grupo hermano de los metazoos. Bilateria comprende dos líneas evolutivas principales: Xenacoelomorpha y Nephrozoa. Nephrozoa comprende dos grupos principales: Protostomia, que incluye Ecdysozoa y Spiralia, y Deuterostomia. Figura tomada de Taylor-Cannon *et al*. (2016).

Si excluimos a poríferos, placozoos, cnidarios y ctenóforos, el resto de metazoos forman un grupo monofilético robusto apoyado por numerosas evidencias: los bilaterados (Bilateria) (figura 4). Los animales bilaterales son aquéllos que poseen un eje dorso-ventral y otro anterior-posterior, el cual define una línea media denominada eje bilateral que divide sus cuerpos en mitades izquierda y derecha. Además, todos los bilaterados son triploblásticos, es decir, presentan tres capas embrionarias: ectodermo, mesodermo y endodermo. La existencia de dichas hojas les ha permitido formar una gran cantidad de órganos, tener una mayor diferenciación y una "cefalización" que dirige y facilita el desplazamiento. Todos estos factores han podido contribuir al gran éxito evolutivo del grupo que actualmente representa la inmensa mayoría de los animales vivos.

El origen y las relaciones evolutivas de los bilaterados también han sido tema de debate durante décadas. Actualmente, las evidencias más robustas apoyan la existencia de dos clados bien diferenciados dentro de los organismos bilaterales: **Xenacelomorpha** (*Xenoturbella* + Acelomorpha) y **Nephrozoa** (Ecdysozoa, Spiralia y Deuterostomia) (figura 4). La existencia de Xenacelomorpha como grupo natural ha sido muy controvertida durante muchos años, principalmente porque Acelomorpha era considerado un grupo de platelmintos (Spiralia) y *Xenoturbella*, un organismo sencillo próximo a Deuterostomia (Philippe *et al*., 2011). Sin embargo, robustas evidencias tanto morfológicas como moleculares han sustentado la cercana relación entre Acelomorpha y *Xenoturbella* y les sitúa como el grupo hermano del resto de organismos bilaterales (Taylor-Cannon *et al*., 2016). Por otra parte, los ecdisozoos contienen el grupo animal con más especies descritas, los artrópodos, además de nematodos, tardígrados y priapúlidos, entre otros grupos. Los espiralados son el grupo con mayor número de phyla y también el más diverso en cuanto a planes corporales; incluye platelmintos, nemertinos, anélidos, moluscos y braquiópodos entre otros grupos. Los grupos Ecdysozoa y Spiralia juntos constituyen el grupo hermano de Deuterostomia (figura 4). Los deuteróstomos incluyen a los equinodermos (como las estrellas de mar), los hemicordados y los cordados, entre los que se encuentran los vertebrados.

En esta tesis se ha decidido seguir la filogenia presentada en Taylor-Cannon *et al*. (2016) una hipótesis robusta que comprende la mayoría de los nodos consistentes y mayoritariamente aceptados en la actualidad.

1.4. Asimetrías izquierda/derecha (AID) en organismos bilaterales

Los animales bilaterales se definen por un único eje de simetría que divide el cuerpo en dos partes simétricamente opuestas: izquierda y derecha. Sin embargo, la mayoría también presenta una gran cantidad de asimetrías concretas tanto en caracteres internos como externos (ej. la posición del corazón o hígado en vertebrados, las pinzas de algunos cangrejos, el sentido de enrollamiento de los caracoles y la torsión del intestino en los artrópodos). Se ha descrito algún tipo de asimetría en adultos o en larvas en representantes de los tres principales grupos de bilaterados (ecdisozoos, espiralados y deuteróstomos) aunque no se han descrito claras asimetrías en otros grupos tales como los acelomorfos. Los órganos impares como el hígado, el estómago, los giros del intestino, e incluso el corazón, generan asimetrías por su posición diferencial a uno u otro lado del eje bilateral. Por otro lado, los órganos pares generan asimetría por su morfología como ocurre por ejemplo con los pulmones de vertebrados (el pulmón izquierdo presenta dos lóbulos a diferencia del derecho que posee tres) o también por su función, como ocurre en el cerebro que pese a ser un órgano impar, es bilateral en su posición y anatomía pero excepcionalmente asimétrico en la distribución de áreas funcionales. Se cree que algunas de estas asimetrías ayudan a mejorar el empaquetamiento de los órganos en las cavidades viscerales y son críticas para el buen funcionamiento de los órganos. Así por ejemplo, el corazón puede alojarse mejor en la cavidad torácica al existir un lóbulo menos en el pulmón izquierdo. La disposición interna anómala de los órganos da lugar a la aparición de defectos en lateralidad conocidos bajo el nombre de heterotaxias, del griego "otro orden", que pueden tener consecuencias severas en la viabilidad de los organismos (Capdevilla *et al*., 2000; Hamada *et al*., 2002; Lobikina *et al*., 2012) (figura 5).

Cuando este cambio de lugar es global y el individuo sufre una reversión completa de todo el plan corporal siendo una imagen especular perfecta del esquema inicial, se obtiene una conformación llamada ***situs inversus*** en contraposición con la conformación normal o *situs solitus* (figura 5). Sin embargo, es más común encontrar situaciones de *situs ambigus* en las que al menos un órgano impar está ubicado incorrectamente. En una tercera categoría se encuentran casos de pérdida de simetría por isomerismo (dos lados izquierdos o derechos, por ejemplo dos pulmones derechos; duplicación de órganos impares como por ejemplo el bazo) (Capdevilla *et al*., 2000) o transposiciones como la que afecta a las arterias grandes del corazón (TAG), en la que las arterias se encuentran intercambiadas entre ventrículos y que requiere cirugía inmediata en neonatos que presentan esta afección (Unolt *et al*., 2013).

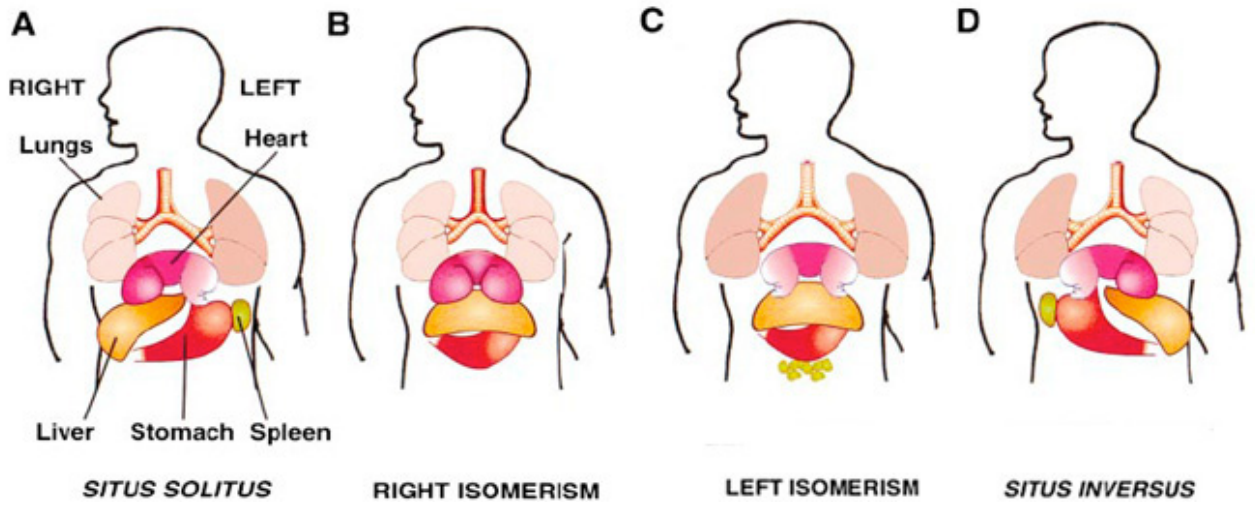


Figura 5. Disposición de los órganos internos en *Homo sapiens.* El correcto empaquetamiento de los órganos en la cavidad visceral es crítico para su buen funcionamiento. La aparición de defectos de lateralidad conocidos bajo el nombre de heterotaxias puede tener consecuencias severas en la viabilidad de los organismos. A) *Situs solitus*; conformación normal en la que por ejemplo el pulmón derecho consta de tres lóbulos, el hígado se coloca en el lado derecho y el bazo en el lado izquierdo; B) isomerismo derecho; dos lados derechos; C) isomerismo izquierdo; dos lados izquierdos; D) *situs inversus*; reversión completa de la conformación normal. Figura tomada y modificada de Capdevila *et al*. (2000)

Pero, ¿cómo se generan estas asimetrías? Por su relevancia médica, el establecimiento de asimetrías ha sido mayoritariamente estudiado en organismos modelo de vertebrados como el ratón (*Mus musculus*), el gallo (*Gallus gallus*), el pez cebra (*Danio rerio*) y la rana (*Xenopus laevis*), aunque también hay algunos estudios en la mosca del vinagre (*D. melanogaster*), en el nematodo (*Caenorhabditis elegans*) y en los caracoles. Todos estos estudios han demostrado que este proceso comienza tempranamente en el embrión y se extiende prácticamente hasta el final del desarrollo. En una primera fase, determinados **eventos tempranos de ruptura de simetría** establecen las primeras bases de diferenciación de lados. Esta información es trasladada en una segunda fase a nivel génico, produciéndose la **expresión asimétrica de cascadas génicas** que a su vez amplifica la señal inicial. La señalización diferencial desemboca en diferencias en la formación del mesodermo y regula la organogénesis. Por tanto, los estudios publicados hasta la fecha apuntan a que el establecimiento de las AID está claramente controlado por mecanismos genéticos. No obstante, mientras algunos mecanismos están evolutivamente conservados, otros son específicos de algunos grupos o incluso especies. A continuación se detalla brevemente cómo se rompe la simetría inicial, cómo se transmite la señal y qué consecuencias tiene en la morfología de los organismos estudiados hasta el momento.

1.4.1. Eventos tempranos de ruptura de simetría

Se han propuesto básicamente tres mecanismos por los cuales se puede romper la simetría inicial: el flujo generado por los cilios del nodo de vertebrados, el flujo diferencial de iones creado a través de bombas y cambios en la dinámica del citoesqueleto que modifican su estructura.

1.4.1.1. Flujo direccional generado por los cilios del nodo de vertebrados

Durante la gastrulación de los vertebrados se forma un grupo de células dilatadas y ciliadas en la parte anterior del embrión a través de las cuales otras células migran para formar diferentes estructuras. Este grupo de células es conocido como el **nodo** y se comporta también como un "organizador", actuando como centro de señalización de multitud de factores que influyen en la síntesis de proteínas de tejidos circundantes para la correcta organización del embrión y la formación de estructuras (figura 6). El nodo se conoce también con el nombre de cápsula de Kupffer en peces, nodo de Hensen en aves y el organizador de Spemann en *Xenopus*. Un mecanismo ampliamente aceptado para romper la simetría inicial implica la rotación de los cilios de las células del nodo y la generación de una corriente de líquido hacia la izquierda del embrión. Tanto la rotación de los cilios como la generación de la corriente de fluido son necesarias para un establecimiento normal de las AID (Nonaka *et al*., 1998; Okada *et al*., 1999; Takeda *et al*., 1999; Shinohara & Hamada, 2017).

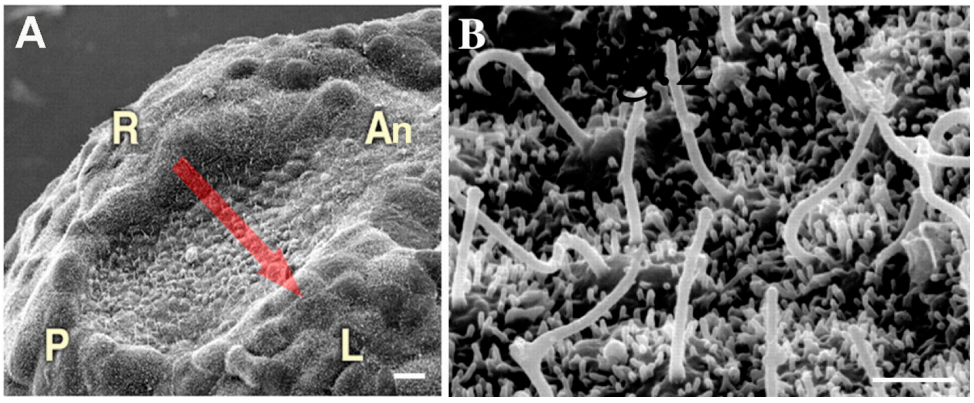


Figura 6. El nodo es una estructura especializada de vertebrados que actúa como "organizador". A) Fotografía de microscopía de barrido del nodo en ratón; An, anterior; P, posterior; L, izquierda; R, derecha; La flecha roja indica la dirección del flujo del nodo; B) detalle de los cilios que se localizan en el nodo; su rotación genera corrientes del líquido extracelular en el nodo. Figura tomada y modificada de Shiratori & Hamada (2006).

Los mecanismos por los cuales la rotación de los cilios y la corriente generada controlan las AID son aún tema de debate, y existen dos hipótesis principales:

- Hipótesis del gradiente químico: algunos autores apoyan la existencia de uno o varios morfógenos que pudieran ser concentrados en la parte izquierda como consecuencia de la corriente generada en el nodo. Algunos candidatos propuestos incluyen las proteínas Sonic-hedgehog (Shh) (Okada *et al*., 1999), Factor de crecimiento fibroblástico 8 (FGF8) (Meyers & Martin, 1999), Factor de crecimiento diferenciador 1 (GDF1) (Rankin *et al*., 2000) y Nodal (Saijoh *et al*., 2003). Sin embargo, hasta el momento no hay evidencias directas que señalen a ninguna de estas proteínas.

- Hipótesis de los dos cilios: hay evidencias que apoyan la existencia de dos tipos de cilios en el nodo, unos con movilidad y otros inmóviles. Los cilios móviles contienen una dineína ("left-right dynein", LRD) que les confiere la movilidad y una policistina ("polycystin" 2, PKD2), una proteína que actúa como un canal de calcio. Se ha propuesto que los cilios móviles serían necesarios para crear el flujo de fluido diferencial en el nodo y que los cilios inmóviles actuarían como mecanorreceptores determinando diferencias en la presión del flujo. Estos mecanorreceptores comenzarían una señalización diferencial por iones de calcio activando la cascada de señalización asimétrica. Apoyando esta hipótesis se ha visto que hay altos niveles de calcio intracelular en el lado izquierdo del nodo y que mutantes para *lrd* o *polycystin2* muestran patrones de asimetrías alterados y la ausencia de altos niveles de calcio intracelular en el lado izquierdo (McGrath *et al*., 2003).

Ya sea por uno u otro mecanismo, parece claro que los cilios tienen un papel relevante en el correcto establecimiento de las AID en vertebrados. Además, este hecho se ha comprobado también en otros deuteróstomos no vertebrados como las ascidias (Nishide *et al*., 2012) y los equinodermos (Tisler *et al*., 2016). Estos datos apoyan un papel ancestral de los cilios para establecer asimetrías en el antecesor común a todos los deuteróstomos.

Sin embargo, en otros deuteróstomos se han descrito signos de lateralidad en etapas muy tempranas del desarrollo, incluso antes de la formación del nodo. Esto abre la posibilidad de que la rotación de los cilios y el flujo del fluido extracelular no constituyan el evento inicial de rotura de simetría, sino más bien un mecanismo de amplificación de una señal anterior.

Además, los organismos bilaterales no deuteróstomos carecen de una estructura homóloga al nodo y son capaces de establecer asimetrías mucho antes de que se formen cilios o incluso careciendo de ellos. Por tanto, debe existir al menos otro mecanismo diferente a la rotación de cilios para romper la simetría inicial en estas otras especies.

1.4.1.2. Flujo direccional de iones a través de bombas

El uso de bombas y canales de iones para establecer gradientes de iones que impliquen diferencias en AID parece ser un mecanismo mucho más extendido en los bilaterados. Se ha demostrado una actividad diferencial de bombas de H^+/K^+- ATPasa y canales de Na^+/K^+ en ambos lados del embrión tanto en vertebrados como en otros deuteróstomos (equinodermos y ascidias) (Rutenberg *et al*., 2002; Shimeld & Levin, 2006; revisado en Raya & Izpisúa-Belmonte, 2006). Esta actividad diferencial ocurre con bastante anterioridad a la formación del nodo y, por tanto, es anterior al flujo en esta zona.

Otros organismos bilaterales no deuteróstomos como el nematodo *C. elegans* o el platelminto *Dugesia japonica* también usan canales de iones para generar AID (Nogi *et al*., 2005; Chang *et al*., 2011; Schumacher *et al*., 2012). El nematodo utiliza canales de calcio para generar diferencias en tipos neuronales entre el lado derecho e izquierdo. Por su parte, el platelminto utiliza bombas H^+/K^+- ATPasa para regenerar cabeza y ojos con diferencias funcionales entre ambos lados del cuerpo.

Se ha propuesto que la actividad polarizada de estos canales de iones produciría la formación de un flujo iónico asimétrico que daría lugar a un gradiente de pH (Walentek *et al.*, 2012). Este gradiente atraería pequeñas moléculas como serotonina, calcio, inositol o ácido retinoico a través de las uniones tipo *gap* entre células, de manera activa y específica hacia un lado. La concentración de estos determinantes en un lado del embrión activaría posteriormente rutas de señalización de forma asimétrica (Walentek *et al*., 2012).

1.4.1.3. Dinámica y estructura del citoesqueleto

El citoesqueleto es un componente crítico a la hora de establecer AID en numerosos organismos bilaterales. El citoesqueleto por sí mismo ya establece una cierta asimetría celular debido a la estructura de los microtúbulos (Xu *et al*., 2007; Inaki *et al.,* 2016). Los filamentos de Actina son hélices que siguen la dirección de las agujas del reloj y pueden generar fuerzas rotacionales que ocasionan una organización asimétrica de toda la red (Tee *et al*., 2015;

Naganathan *et al*., 2014). La quiralidad (propiedad por la cual dos imágenes especulares no pueden superponerse) molecular de los filamentos de Actina se puede transmitir a otras moléculas durante su polimerización, su interacción con proteínas motoras y su despolimerización. Algunos autores defienden que esa quiralidad no sólo se transmite a nivel molecular, sino que también puede transmitirse a nivel celular, tisular y finalmente influir en el plan corporal del organismo (Naganathan *et al*., 2016; McDowell *et al*., 2016).

Se han descrito varios ejemplos que ilustran la relevancia del citoesqueleto de Actina y sus proteínas asociadas en el establecimiento de AID en los principales grupos de organismos bilaterales: espiralados, ecdisozoos y deuteróstomos. Por ejemplo, la correcta polimerización de la Actina y su organización quiral regula la conformación del plan corporal de los caracoles (organismos espiralados) a través de la orientación del huso mitótico en la tercera división del embrión (Shibazaki *et al*., 2004). La alteración física en la dirección de la división celular y por tanto en la orientación del citoesqueleto durante la tercera división en embriones de caracoles tiene efectos en la activación asimétrica de las rutas génicas y consecuentemente en la determinación de las AID (Kuroda *et al*., 2009). Además, en la especie de caracol *Lymnaea stagnalis* se ha relacionado la actividad de la proteína Formin (una proteína que regula la dinámica de los microfilamentos de Actina) con el establecimiento de AID (Davison *et al*., 2016). Sin embargo, esto parece ser particular de esta especie puesto que en otras especies de caracoles Formin no parece tener el mismo papel (Davison *et al.*, 2016).

Por otra parte, se ha descrito el papel de las Miosinas (MyoID), proteínas motoras asociadas a los filamentos de Actina, en la determinación de las AID en el ecdisozoo *D. melanogaster* (Spéder & Noselli, 2006). Las mutaciones provocadas en estas miosinas causan la aparición de la imagen especular del fenotipo silvestre (Spéder *et al*., 2006; Hozumi *et al*., 2006). Sin embargo, en el ecdisozoo *C. elegans* son los microtúbulos y la orientación del huso mitótico los que parecen regular los eventos asimétricos (Bergmann *et al*., 2003).

Del mismo modo, la mutación en el gen que codifica para una proteína asociada al citoesqueleto de vertebrados llamada **Inversin** ocasiona la aparición de la imagen especular del fenotipo silvestre (Yokoyama *et al*., 1993). En esta tesis se ha estudiado de manera detallada la caracterización de esta proteína en organismos bilaterales y por tanto en el apartado 1.4.1.4. se presentan de manera específica los antecedentes de esta proteína.

En resumen, el citoesqueleto parece ser un referente quiral en sí mismo o estar actuando como un distribuidor quiral de moléculas (Qiu *et al*., 2005; Aw *et al*., 2008; Spéder *et al*., 2007; Aw & Levin, 2009). Las divisiones celulares y las uniones célula-célula, procesos controlados

por dinámicas del citoesqueleto, tienen un papel crítico durante el desarrollo embrionario y principalmente en los primeros estadios. Parece claro, pues, que el citoesqueleto puede tener las claves para entender cómo se rompe la simetría para establecer diferencias morfológicas.

1.4.1.4. La proteína Inversin y su papel en el establecimiento de AID

La existencia de un mutante de ratón que mostraba una polaridad en el eje izquierda-derecha completamente invertida (*situs inversus*) con respecto a la forma silvestre llevó a la identificación del gen *inversin* (Yokoyama *et al.*, 1993; Morgan *et al*., 1998; Mochizuki *et al*., 1998) (figura 7). Los mutantes homozigóticos para la mutación recesiva en *inversin* representan el único modelo conocido hasta el momento en el cual las AID están invertidas de manera consistente en contraposición con otras mutaciones que muestran tanto inversión de asimetrías como una distribución azarosa.

La proteína codificada por este gen está formada por unos mil aminoácidos y en ratón se han descrito al menos tres isoformas generadas por procesamiento alternativo que poseen diferente localización subcelular (Nürnberger *et al*., 2002; Eley *et al.*, 2004; Nürnberger *et al*., 2006). Inversin es una proteína con 16 dominios ankirina en su extremo N-terminal, dominios estructurales que están entre los más comunes de las proteínas conocidas. Sin embargo, la segunda mitad de la proteína no muestra similitud con ninguna otra proteína conocida (Mochizuki *et al*., 1998). Este extremo C-terminal tiene dos dominios de unión a Calmodulin (IQ1 y IQ2), dos dominios de localización nuclear (NLS) además de dominios de destrucción (D-Box) y un dominio tipo Nineína (Morgan *et al*., 2002; Eley *et al*., 2004; Shiba *et al.*, 2008; Lienkamp *et al*., 2012).

La mayoría de los descubrimientos sobre Inversin se han desarrollado en ratón por ser el primer organismo en el que esta proteína fue descubierta y posteriormente, a lo largo de los años, se han ido conociendo más aspectos de tipo funcional en cultivos celulares (Nürnberger *et al*., 2004), *Xenopus*, *Gallus* (Morgan *et al*., 2002; Yasuhiko *et al*., 2006; Lienkamp *et al*., 2010) y *Danio* (Otto *et al*., 2003). Por todos estos estudios sabemos que Inversin es una proteína presente en diferentes compartimentos celulares y asociada a varios componentes del citoesqueleto como los microtúbulos, las uniones adherentes, los monocilios y los filamentos de Actina. La presencia de Inversin parece particularmente relevante en los cilios del nodo del ratón (Watanabe *et al*., 2003). Este hecho, junto con la descripción de defectos en el flujo del nodo en los mutantes para *inversin* (un flujo muy débil y turbulento, aunque dirigido hacia la izquierda como en la forma silvestre; Okada *et al*., 1999), ha llevado a diferentes autores

a proponer que la proteína Inversin pudiera requerirse para el correcto movimiento de los cilios del nodo o para la mecanorrecepción de cambios en el flujo (Watanabe *et al.*, 2003). Sin embargo, precisamente el hecho de que el flujo provocado por los cilios del nodo no tenga una dirección inversa en los mutantes con respecto al silvestre, sino que mantenga la dirección de forma levógira, sugiere que la inversión fenotípica descrita en los mutantes para el gen *inversin* no se explica únicamente por defectos en el flujo del nodo (Watanabe *et al*., 2003). Por tanto, el gen *inversin* debe estar relacionado con el establecimiento de AID por algún otro mecanismo.

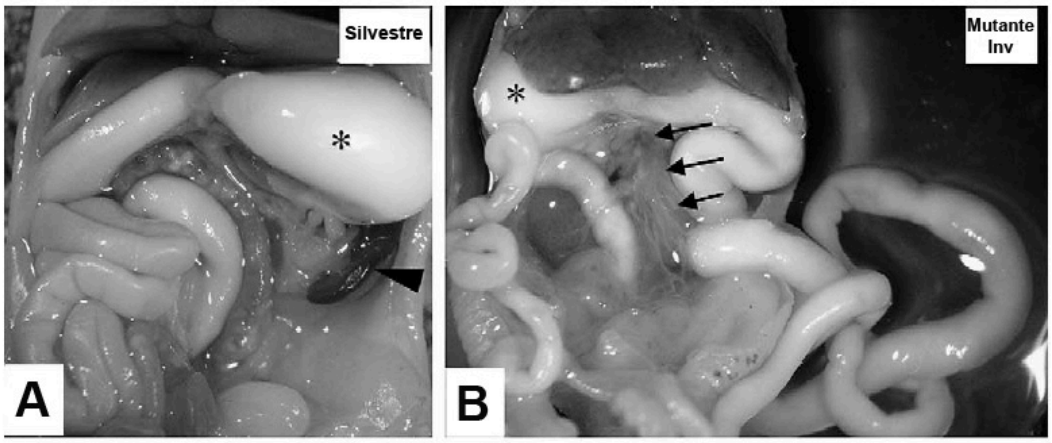


Figura 7. Los mutantes del gen *inversin* presentan una consistente organización corporal de tipo *situs inversus*. A) Anatomía interna de un ratón silvestre, con una organización visceral de tipo *situs solitus*. El asterisco (*) señala el estómago; B) anatomía interna de un ratón mutante para el gen *inversin,* que presenta una reversión completa del plan corporal o *situs inversus*. El asterisco muestra la nueva posición del estómago; las flechas indican la posición de las venas del sistema portal. Figura tomada y modificada de Shimadera *et al*. (2007).

Una pista pudiera estar en los **dominios de unión a Calmodulin** (IQ1 y IQ2) presentes en el extremo C-terminal de Inversin, que interaccionan de manera funcional con la proteína Calmodulin. Calmodulin es una proteína de unión a otras proteínas, con distribución ubicua en las células y que está implicada en una gran variedad de rutas de señalización de una manera calcio-dependiente, ion al cual se une física y reversiblemente (Crivici & Ikura, 1995; Jurado *et al*., 1999; Chin & Means, 2000). El calcio regula la interacción entre Inversin y Calmodulin de manera que Inversin se une a Calmodulin a través de sus dominios IQ en ausencia de calcio, pero es incapaz de unirse en presencia de este ion (Yasuhiko *et al*., 2001). Experimentos hechos en *Xenopus* muestran que la inyección en embriones de dos células de ARN mensajero de una forma mutante carente del dominio IQ de Inversin impide la unión entre Calmodulin e Inversin y produce un fenotipo con un patrón de AID azaroso (Yasuhiko *et al*., 2001; 2006). Curiosamente, este patrón azaroso sólo se recupera cuando se inyecta este mensajero mutante en el blastómero

derecho, mientras que la inyección en el blastómero izquierdo carece de efecto. Por tanto, parece que hay algún componente diferencial entre los blastómeros que provoca que la actividad de Inversin sea dependiente de ello. Se ha propuesto que la concentración de los iones de calcio puede ser este factor diferencial, puesto que se han descrito diferencias en la concentración de calcio entre el lado derecho e izquierdo en varias especies de vertebrados (McGrath *et al*., 2003; Raya *et al*., 2004). En conclusión, al menos en *Xenopus*, la actividad de la proteína Inversin y su relación con AID depende de su dominio de unión a Calmodulin, además del contexto celular en el que se encuentre. Hasta el momento, se desconoce si estos eventos tienen lugar también en otros organismos.

1.4.2. Expresión diferencial e identidad de lado: ruta de Nodal

La información inicial generada por los eventos tempranos de rotura de AID se transmite según avanza el desarrollo y poco después comienza a darse la expresión diferencial de genes a uno u otro lado del cuerpo, generando identidad en los mismos y extendiendo la información a otras partes del embrión. Todos estos procesos, esta amplificación de la señal inicial, desembocan en diferencias en la formación del mesodermo y la organogénesis que desarrollan finalmente el plan corporal.

Los estudios moleculares y genéticos llevados a cabo en este campo señalan a la ruta de señalización del *Transforming Growth Factor-β* (TGF-β) **Nodal** como la ruta crítica para establecer, entre otras cuestiones, las asimetrías corporales en deuteróstomos (Hamada *et al*., 2002; Duboc *et al*., 2005). *lefty*, otro miembro de la familia TGF-β, y el gen homeobox *Pitx* son también genes clave para la determinación de las asimetrías (Boorman & Shimeld, 2002; Hamada *et al*., 2002; Duboc *et al*., 2005). *nodal* y *lefty* se expresan de manera asimétrica y transitoria, y ambos regulan positivamente a *Pitx*, que controla procesos celulares de migración diferencial, proliferación, adhesión y morfogénesis asimétrica en deuteróstomos. La falta de función de estos genes conlleva por ejemplo problemas de heterotaxia (Levin *et al*., 1997). Estudios comparados han mostrado que *nodal*, *lefty* y *Pitx* tienen funciones similares y son regulados de la misma manera en deuteróstomos (Hamada *et al*., 2002; Duboc *et al*., 2005; Shimeld & Levin, 2006).

Por el contrario, la red Nodal no está conservada en Ecdysozoa. El hecho de que hasta la fecha no se haya identificado ningún ortólogo de *nodal* o *lefty* en ningún genoma de ecdisozoo, que *Pitx* se exprese de manera completamente simétrica y que su falta de función no repercuta en las asimetrías de los ecdisozoos, sugiere que los mecanismos que generan las asimetrías en este

grupo de organismos bilaterales son claramente diferentes de los descritos para deuteróstomos (Schier, 2009).

Todos estos estudios en deuteróstomos y ecdisozoos sugerían que la red de señalización de Nodal era una novedad evolutiva de los deuteróstomos, que se originó en el ancestro común de este grupo. Sin embargo, primero Grande y Patel (2009), y después estudios posteriores (Grande *et al*., 2014; Martín-Durán *et al*., 2017) han identificado la red de señalización Nodal en casi todos los organismos espiralados y han propuesto un papel relevante de esta ruta en el establecimiento de AID en espiralados al igual que ocurre en deuteróstomos. Este hecho ha supuesto una gran novedad en el campo puesto que ha permitido proponer un nuevo escenario evolutivo en el cual la red de señalización Nodal no es exclusiva de los deuteróstomos, sino que tiene un origen antiguo en la historia evolutiva de los metazoos, situándose concretamente en el antecesor común de deuteróstomos, ecdisozoos y espiralados (Grande *et al*., 2014).

1.4.2.1. La regulación de la ruta Nodal

La regulación de la ruta de Nodal incluye una compleja red de receptores, co-receptores, antagonistas y proteínas intracelulares, además de bucles de regulación negativos y positivos que determinan la proliferación y diferenciación celular y la morfogénesis del tejido (Kenny *et al*., 2014) (figura 8). De forma concisa, estudios principalmente en vertebrados han mostrado que el ligando activo (Nodal) se une a un heterotetrámero compuesto por dos pares de receptores específicos tipo Serina/Treonina kinasas, conocidos como Receptores I y II. La formación del complejo receptor-ligando permite la fosforilación y posterior activación del Receptor I por el Receptor II. Esta activación permite la internalización del complejo por mecanismos de endocitosis, momento en el que el Receptor I fosforila la proteína Smad2. En este estado, la proteína Smad2 recluta a la proteína Smad4, a la cual se une formando un heterodímero que se transloca al núcleo y regula de la expresión de los genes diana (revisado en Massagué, 2012). A pesar de esta aparente simplicidad, esta ruta puede producir unos efectos de activación o represión de genes tremendamente diversos, llegando incluso en algunos casos a tener efectos antagónicos. La respuesta celular a la señalización por Nodal depende del contexto celular definido por los niveles del complejo receptor/ligando, la actividad del receptor, la actividad de las proteínas Smad, así como por la presencia de otros factores de transcripción, modificadores de histonas y remodeladores de cromatina, que determinan qué genes serán susceptibles de ser activados o reprimidos (Massagué, 2012).

Uno de los reguladores críticos para el funcionamiento de la ruta Nodal en vertebrados es el correceptor **EGF-CFC**. Parte de las investigaciones realizadas en esta tesis se han centrado en el estudio de este gen y, por tanto, a continuación se detalla con mayor precisión los antecedentes conocidos de esta proteína.

1.4.2.2. EGF-CFC y la ruta Nodal

Las proteínas EGF-CFC (*Epidermal Growth Factor-CFC*) actúan como co-receptores para diferentes rutas de señalización, lo que les confiere un rol multifuncional. Hasta el momento sólo se han descrito estas proteínas en organismos deuteróstomos y se consideran una innovación evolutiva de este grupo (revisado en Shen & Schier, 2000; Ravisankar *et al.*, 2011). En vertebrados, estas proteínas están implicadas en el establecimiento de asimetrías corporales, la morfogénesis del corazón, movimientos celulares durante la gastrulación, la diferenciación de células madre y el crecimiento y remodelización del tejido adulto. El origen de esta plasticidad reside tanto en su composición aminoacídica como en su estructura, que establece dominios con distintos sitios de reconocimiento que a su vez permiten la interacción con diferentes ligandos y receptores (Minchioti *et al.*, 2002; Strizzi *et al.*, 2005; revisado en Klauzinska *et al.*, 2014). En esta proteína pueden distinguirse tres dominios principales: 1) el **dominio tipo EGF** que reconoce diferentes ligandos como TGFβ-s, BMPs, GDFs y tomorregulina (Harms & Chang, 2003) y permite la regulación de las rutas de Nodal, MAPK y PI3K/Akt; 2) el **dominio CFC** que interacciona con receptores y establece grupos funcionales heteroméricos; 3) el **dominio GPI** en el extremo C-terminal que ancla el co-receptor a la membrana (Strizzi *et al.*, 2005; Gray & Vale, 2012).

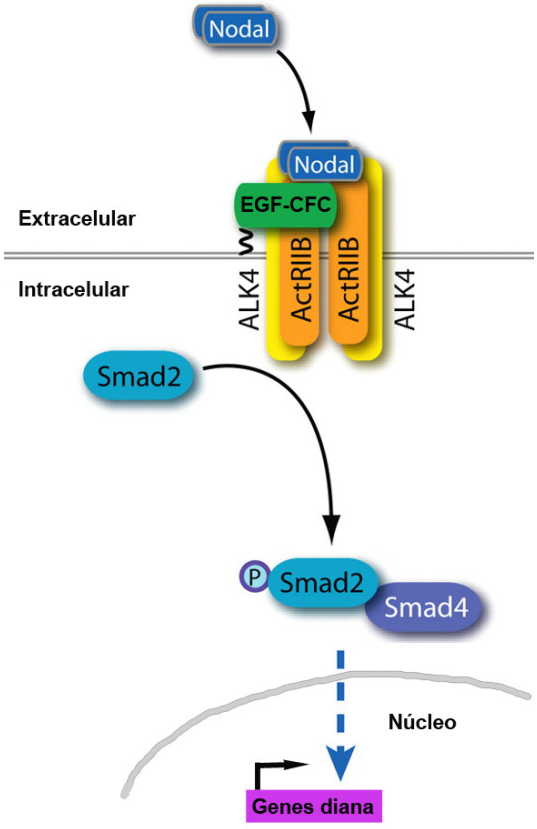


Figura 8. La ruta de Nodal es la ruta clave en el establecimiento de AID. El ligando Nodal se une a sus receptores, que forman un heterotetrámero, al que a su vez se une un co-receptor de la familia EGF-CFC, provocando entonces la transducción de la señal y activando los genes diana. Figura tomada y modificada de Kruithof-de Julio *et al.* (2011).

Una de las funciones más estudiadas de las proteínas EGF-CFC es el establecimiento de AID ya que EGF-CFC es un co-receptor de la ruta de Nodal (apartado 1.4.2.1.). Su participación en la formación

de un complejo junto con los receptores y el ligando es necesaria a nivel estructural, de tal forma que sin este co-receptor, la unión de Nodal a los receptores provoca una disminuida respuesta intracelular (Romano *et al.*, 2012; Yeo & Withman, 2001). En *M. musculus* y *Homo sapiens* hay datos detallados de la interacción entre EGF-CFC y Nodal a nivel molecular (Calvanese *et al.*, 2010). Entre los hallazgos más interesantes se encuentra la caracterización de una treonina (T) dentro del dominio tipo EGF, que se corresponde con la posición T88 en humanos y T72 en ratón, y que ha resultado ser indispensable para la transducción de la señal. También se sabe que la O-fucosilación de este residuo facilita la interacción ligando-co-receptor y ningún otro aminoácido permite la activación de Smad2, ni siquiera aquéllos susceptibles de fucosilación (serina) (Schiffer *et al.*, 2001; Shi *et al.*, 2007). Hasta el momento se desconoce si esta interacción es también necesaria para la activación de la ruta en otros organismos.

1.5. Spiralia y Gastropoda como sistema modelo en el estudio de AID

Los espiralados son organismos que se caracterizan por un tipo de desarrollo embrionario específico: la segmentación espiral. Este tipo de segmentación incluye un programa muy conservado y estereotipado que se caracteriza por la orientación oblicua del plano de división de los blastómeros con respecto al eje animal-vegetal (figura 9). Cuando los blastómeros de un embrión de cuatro células se dividen dan lugar a un embrión con ocho células en el que las nuevas células recién formadas son de un tamaño menor (denominadas micrómeros) con respecto a las originales (macrómeros) (figura 9). La formación de estos micrómeros sigue una rotación de unos 45 grados con respecto al eje animal-vegetal. Los cuatro micrómeros que resultan de cada división son denominados cuartetos. El sentido de la división de los micrómeros comienza siendo en el sentido de las agujas del reloj en la tercera división, pero va alternando sentidos opuestos en las siguientes divisiones. De esta manera, los cuartetos se van formando alternando su disposición y dando un patrón en espiral visto desde el polo animal, lo que da nombre al grupo (Hejnol, 2010; Henry, 2014) (figura 9).

El destino de los linajes celulares que provienen de la división de cada uno de los cuatro macrómeros originales (denominados cuadrantes A, B, C y D) está muy conservado y responde a un programa de inducción también muy conservado en estos organismos (Van den Biggelar & Guerrier, 1979; Martindale *et al.*, 1985, 1986; Freeman & Lundelius, 1992; Dictus & Damen, 1997). La progenie de los macrómeros A, B, C y D da lugar a las partes izquierda, ventral, derecha y dorsal del cuerpo, respectivamente (Verdonk & Van den Biggelar, 1983).

Especialmente importante es la especificación del cuadrante D, que puede considerarse como un "organizador" en estos embriones ya que es un centro de señalización clave para su correcto desarrollo y la correcta especificación del eje dorso-ventral (Henry *et al.*, 2010; Lyons *et al*., 2012; Henry, 2014) (figura 9).

En un momento determinado del desarrollo embrionario las divisiones celulares se realizan no de manera espiral sino bilateral y entonces se puede visualizar un eje de simetría bilateral que en la mayoría de los casos se corresponde con la división simétrica de células del cuadrante D. Por ejemplo, en la especie de caracol *Crepidula fornicata*, es la división bilateral de la célula 4d la que visualiza el eje de simetría izquierda derecha del embrión (Henry *et al*., 2010; Lyons *et al*., 2012; Lyons & Henry, 2014) (figura 9F).

Las asimetrías morfológicas descritas en espiralados son muy diversas. Pueden encontrarse en órganos sensoriales o conexiones nerviosas (nemertinos y platelmintos), en cilios locomotores (platelmintos) y en la dirección de giro de determinadas partes del aparato reproductor y digestivo (anélidos, braquiópodos, rotíferos, platelmintos y moluscos) (Grande, 2010; Martín-Durán *et al*., 2017). Sin embargo, uno de los ejemplos más claros lo constituye el enrollamiento de la concha en los caracoles.

La mayoría de los caracoles tiene el cuerpo enrollado según el sentido de las agujas del reloj (dextrógiros), mientras que en otros caracoles el cuerpo se enrolla en el sentido contrario (levógiros). La quiralidad, es decir, esta propiedad de los caracoles de ser dextrógiros o levógiros, se define muy tempranamente durante el desarrollo embrionario, en concreto durante la tercera división (división de 4 a 8 células). Si un embrión de 4 células sigue una división en espiral en sentido de las agujas del reloj, el adulto resultante será dextrógiro; en caso contrario, será levógiro (figura 10). Una vez definido el sentido de la división, los embriones resultantes desarrollan las asimetrías corporales en los lados contrarios, de modo que dextrógiros y levógiros son imágenes especulares unos de otros. Como se ha comentado en el apartado 1.4.1.3., la orientación del huso mitótico en la tercera división del embrión y la correcta polimerización del citoesqueleto de Actina regulan esta conformación (Shibazaki *et al.,* 2004). La configuración quiral de estos blastómeros es el factor clave para determinar en qué lado del cuerpo se activará la ruta de señalización de Nodal en los embriones y larvas y la consiguiente formación de las asimetrías en los adultos. La ruta Nodal está activa en el lado derecho en caracoles dextrógiros y en el lado izquierdo en levógiros (Grande & Patel, 2009). La alteración física en la dirección en la que se produce la tercera división en embriones de caracoles haciéndola reversa ocasiona la activación asimétrica de las rutas génicas también en el lado reverso, y consecuentemente en la

determinación de las AID, lo que produce un organismo especular (Kuroda *et al*., 2009).

El hecho de que los gasterópodos tengan unas asimetrías corporales muy claramente observables durante el desarrollo ha llamado la atención de naturalistas y biólogos desde hace décadas. Sin embargo, sólo recientemente hemos empezado a comprender algunos de los mecanismos genéticos y moleculares que regulan este proceso y aún quedan cuestiones básicas por esclarecer. En concreto, queda determinar cuál es el primer evento que rompe la asimetría bilateral en los caracoles ocasionando las diferencias detectadas en la configuración del huso mitótico y provocando la configuración quiral de los blastómeros. Además, si bien es cierto que la presencia de la ruta de Nodal en los caracoles y otros espiralados apoya la relevancia de esta ruta en el establecimiento de AID en organismos bilaterales, poco sabemos de su regulación en organismos no vertebrados.

En esta tesis se ha abordado este problema desde el estudio de diferentes genes en dos especies de gasterópodos *C. fornicata* (caracol dextrógiro) y *Biomphalaria glabrata* (caracol levógiro). Se ha caracterizado el gen *inversin* en estos caracoles y en otros metazoos, se ha determinado su patrón de actividad y se han realizado diferentes experimentos para determinar

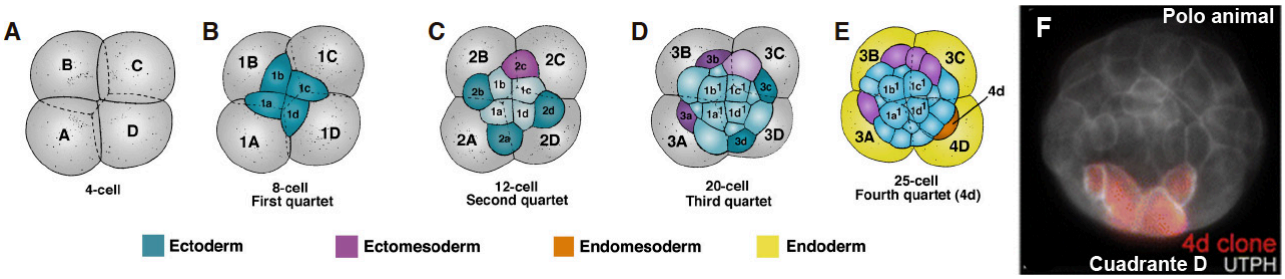


Figura 9. Segmentación espiral en la especie *C. fornicata*. La segmentación espiral es un tipo de programa de división embrionaria muy estereotipada del grupo Spiralia en el que las sucesivas divisiones sufren giros de 45 grados con respecto al cuarteto anterior. Los distintos cuadrantes establecidos, junto con las células que los componen y las identidades que adquieren, siguen destinos celulares muy específicos y conservados en los grupos de Spiralia; A) Embrión de 4 células. Los cuadrantes se designan en esta especie dextrógira siguiendo el sentido de giro de las agujas del reloj; B) embrión de 8 células. Las 4 nuevas células que surgen en cada división se denominan cuartetos. La división desigual genera dos tipos celulares denominados macrómeros (las células más grandes, en el esquema en color gris y amarillo) y micrómeros (las células más pequeñas, en el esquema en distintos colores). Esta división determina la quiralidad del embrión y supone el primer evento de ruptura de simetría, al producirse el primer giro de los micrómeros sobre los macrómeros. En esta especie el giro se produce en el sentido de giro de las agujas del reloj, por lo que se denomina especie dextrógira; C) embrión de 12 células. El nuevo cuarteto de micrómeros proviene de los macrómeros y en este caso sufren un giro levógiro, puesto que el sentido de giro es alternante en las sucesivas divisiones; D) embrión de 20 células; E) embrión de 25 células. En este punto nace la célula 4d, que actúa como organizador y establece el punto a partir del cual se producen divisiones de tipo bilateral; F) seguimiento de la división bilateral de la célula 4d a través del uso de compuestos fluorescentes. Las membranas se han marcado gracias a la expresión ubicua de eutrofina fusionada a una etiqueta fluorescente. Los distintos colores de las figuras A-E muestran los destinos celulares, indicados en la leyenda. A continuación se detalla el convenio tradicionalmente seguido para la nomenclatura de los distintos cuartetos. Los macrómeros se designan con letras mayúsculas (A, B, C y D) y los micrómeros con minúsculas, manteniendo la letra del cuadrante del que provienen. Por ejemplo, en la figura B, Los macrómeros se designan como 1A, 1B, 1C y 1D, mientras que los micrómeros se denominan 1a, 1b, 1c y 1d. Asimismo, los números indican la división de la que proceden. En las sucesivas divisiones, las células hijas que nacen hacia el polo animal reciben el superíndice 1, y aquéllas que se orientan hacia el polo vegetal reciben el superíndice 2 (por ejemplo, $1b^1$ y $1b^2$), y en las sucesivas divisiones se van añadiendo superíndices de acuerdo a este criterio de origen y orientación (por ejemplo, $1b^{11}$ y $1b^{12}$). Figuras tomadas y modificadas de Lyons & Henry (2014) y Lyons *et al*. (2015).

su potencial relevancia en el establecimientos de AID en estas especies de caracoles. Por otra parte, se ha profundizado en el conocimiento de la ruta de señalización Nodal en estas dos especies de gasterópodos a través del estudio de dos componentes básicos de la ruta: el ligando Nodal y el co-receptor EGF-CFC.

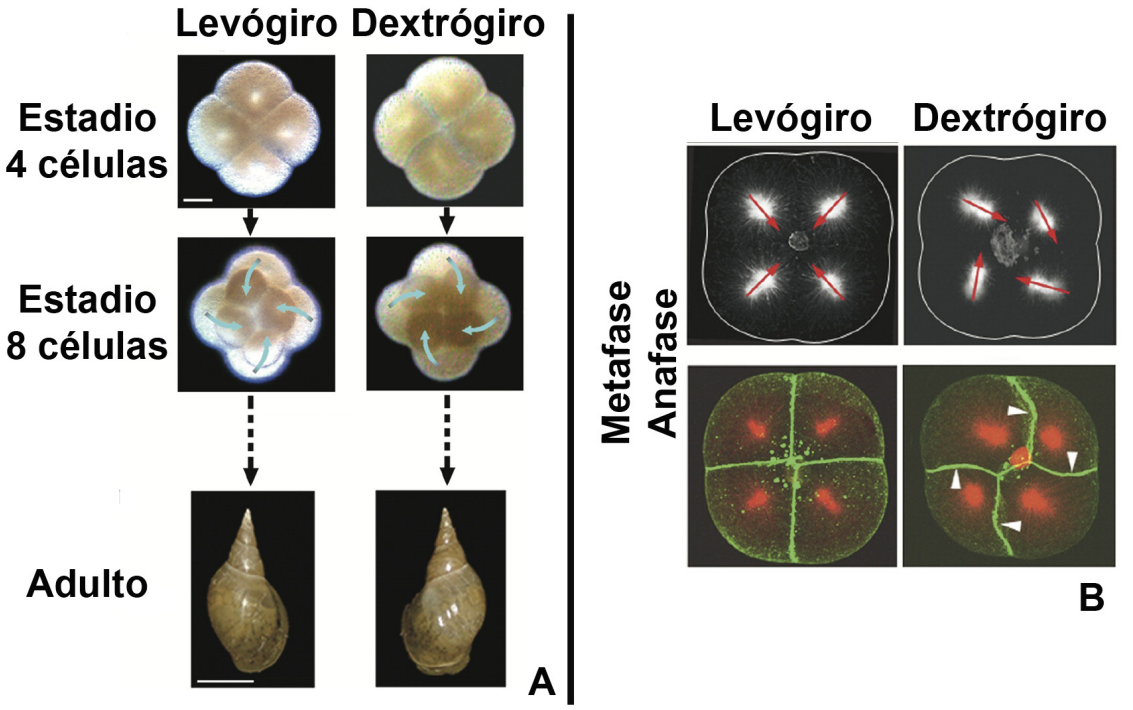


Figura 10. La relación de la tercera división embrionaria con el establecimiento del plan corporal levógiro y dextrógiro en los gasterópodos. A) Los individuos levógiros sufren un giro antihorario de los micrómeros con respecto a los macrómeros, mientras que este giro es de tipo horario en los individuos dextrógiros. Esta conformación define y se corresponde con el plan corporal final del adulto; B) arquitectura del citoesqueleto durante la tercera división en gasterópodos dextrógiros y levógiros. En los organismos dextrógiros, tanto las membranas celulares (cabezas de flecha blancas) como los husos mitóticos (flechas rojas) sufren deformaciones mecánicas que fuerzan esa orientación en la división. Por otro lado, los individuos levógiros no sufren deformaciones en ningún sentido, sino que el citoesqueleto de membranas mantiene una estructura "relajada y recta" y los husos mitóticos se orientan perpendicularmente hacia el centro. Figura tomada y modificada de Shibazi *et al*. (2004)

Objetivos

La presente tesis pretende avanzar en el conocimiento de los mecanismos implicados en el establecimiento de asimetrías corporales durante el desarrollo de los organismos bilaterales, tanto en los primeros momentos del desarrollo cuando se produce la ruptura del patrón simétrico como en las rutas de señalización posteriores que amplifican y mantienen la señalización diferencial.

El gen *inversin* ha sido descrito como un gen clave en este proceso que actúa en el comienzo del desarrollo de los embriones de vertebrados y cuya mutación provoca la aparición de las formas especulares con respecto a las silvestres. Además, se ha propuesto que la actividad de la proteína Inversin y su relación con las AID dependen de su dominio IQ de unión a Calmodulin y del contexto celular en el que se encuentre. Sin embargo, los únicos datos sobre este gen han sido proporcionados por estudios en vertebrados y se desconoce la presencia de *inversin* en ningún otro grupo animal, así como su potencial papel en el establecimiento de sus AID. En esta tesis se planteó contribuir al conocimiento de este gen en metazoos no vertebrados y en particular, se establecieron los siguientes objetivos concretos:

1. Identificar ortólogos del gen *inversin* en metazoos no vertebrados.
2. Establecer el origen de este gen y reconstruir su diversificación en metazoos.
3. Caracterizar la proteína codificada por este gen en metazoos e identificar sus dominios funcionales.
4. Determinar su patrón espacio-temporal de activación en Gastropoda.
5. Comprobar la implicación del dominio IQ en el establecimiento de AID en Gastropoda

Hace casi una década que se describió la presencia de la ruta de Nodal y su implicación en el establecimiento de AID en Gastropoda y otros organismos espiralados. Sin embargo, nada se sabe acerca de los mecanismos de activación y de regulación de esta ruta en estos organismos. Dos componentes críticos para la activación de la ruta en vertebrados son las proteínas Nodal y EGF-CFC. Los co-receptores tipo EGF-CFC sólo se han descrito hasta el momento en deuteróstomos, y se sabe que existe una treonina dentro de su dominio tipo EGF que es indispensable para permitir el acoplamiento correcto del ligando Nodal con sus receptores y en consecuencia, que se produzca la transducción de la señal y activación de la ruta. En esta tesis se ha querido profundizar en el conocimiento de la regulación de la ruta de Nodal en Gastropoda a través de

los siguientes objetivos concretos:

1. Identificar el ortólogo del gen *nodal* en la especie de gasterópodo *C. fornicata*.
2. Describir el patrón de activación espacio-temporal del gen *nodal* en esta especie de Gastropoda.
3. Reconocer posibles consecuencias en el fenotipo de los embriones de *C. fornicata* debidas a la sobre-expresión del gen *nodal* y a la perturbación de la ruta.
4. Identificar ortólogos de los genes *egf-cfc* en metazoos no deuteróstomos.
5. Establecer el origen de estos genes y su diversificación en metazoos.
6. Caracterizar las proteínas codificadas por los genes *egf-cfc* en los diferentes organismos e identificar sus dominios.
7. Determinar el patrón espacio-temporal de activación del gen *egf-cfc* en Gastropoda.
8. Comprobar si la implicación de una treonina concreta dentro de dominio tipo EGF para el funcionamiento de EGF-CFC descrita en vertebrados está conservada también en Gastropoda.

Materiales y métodos

3.1. *C. fornicata* como modelo de estudio

C. fornicata es una especie de gasterópodo marino del género *Crepidula*, que se incluye dentro de la familia Calyptraeidae, perteneciente a su vez al grupo Caenogastropoda. Morfológicamente, esta especie tiene una concha de tipo lapa, aplanada con un ápex posterior y una ligera curva lateral (figura 11). En la parte interna de esta concha existe un septo que define una cavidad en la que se encaja la masa visceral y en donde se ancla un pie ensanchado. *Crepidula fornicata* es una especie sedentaria y filtradora puesto que se alimenta del fitoplancton presente en el agua en movimiento. Los adultos viven anclados a otros adultos, apilándose de forma que el pie se adhiere a la superficie de la concha del individuo subyacente (figura 11). Estas pilas o *stacks* varían en tamaño y pueden llegar a estar formadas por entre 2 y 20 individuos. El animal situado más abajo en la pila se adhiere a una concha vacía, una roca o cualquier otro sustrato sólido que le permita encerrase en su concha, ya que carecen de opérculo (estructura que cierra la apertura de la concha de los gasterópodos).

Los adultos de *C. fornicata* son hermafroditas protándricos, es decir, son machos en su fase juvenil y posteriormente cambian a hembras debido a diferentes factores. Los individuos

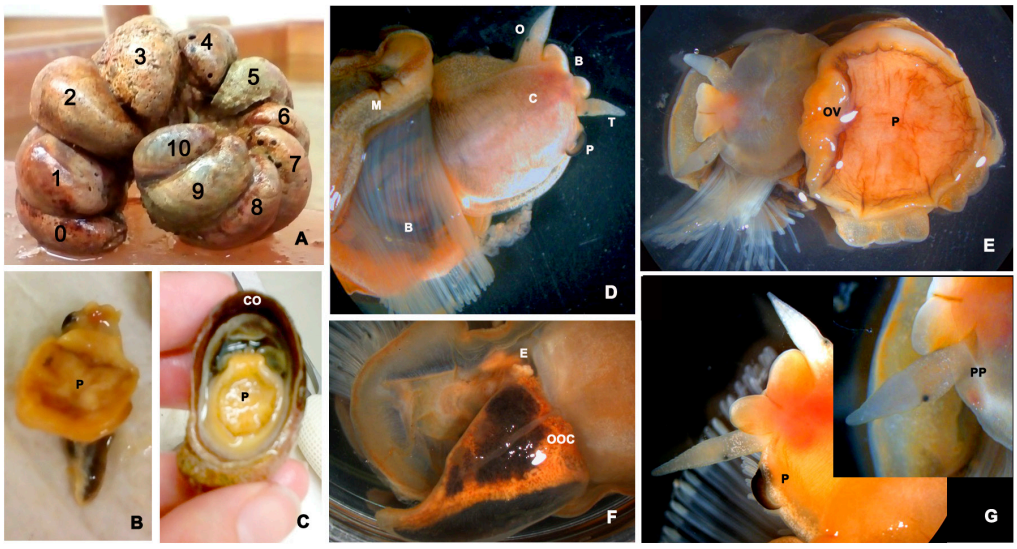


Figura 11. Adultos de la especie *C. fornicata* y su anatomía. A) *Stack* de 10 adultos de la especie *C. fornicata*. El número 0 indica una concha vacía; B) Adulto fuera de la concha (P, pie); C) Adulto dentro de la concha (P, pie; CO, concha). Los adultos poseen un pie ensanchado con el que se adhieren al individuo inmediatamente debajo; D) Vista dorsal de un adulto fuera de la concha, con el manto apartado posteriormente para descubrir la cabeza y visualizar las branquias (M, manto; B, branquias; C, cabeza, O, ojo; B, boca, T, tentáculo, P, pene); E) Vista ventral de una hembra adulta fuera de la concha (OV, oviducto; P, pie); F) Vista dorsal de un adulto, con el manto por encima de la cabeza, para visualizar la masa visceral y las gónadas (E, esperma; OOC, oocitos en maduración); G) Comparación entre machos y hembras adultos. Al ser una especie hermafrodita de tipo protándrico ocurre que, durante el cambio de sexo, el pene de los machos (P) se atrofia hasta quedar reducido a una papila (PP).

más grandes, viejos y por tanto situados en la base de la pila son hembras. Por su parte, los pequeños machos se unen al *stack* y sobre él crecerán y vivirán el resto de su vida. Con el tiempo, estos individuos crecerán y otros nuevos machos se situarán por encima; este hecho provoca un cambio de sexo paulatino hasta que finalmente estos machos se convierten en hembras. La fecundación es interna y tiene lugar entre individuos del mismo *stack*. Los adultos de esta especie son estacionales, siendo los meses de abril a julio los meses reproductivos. Desde mediados de julio a diciembre los adultos producen gametos que se mantienen hasta marzo o abril cuando el clima cambia y comienza la época de reproducción.

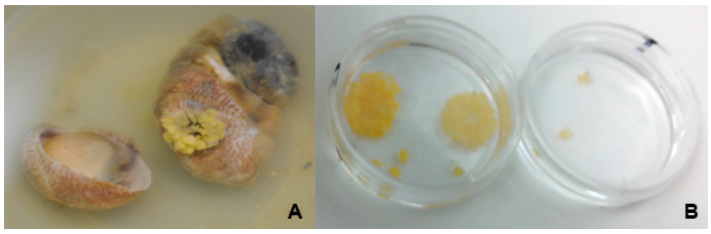


Figura 12. Masa de huevos de la especie *C. fornicata*. Los adultos ovopositan las puestas en la concha sobre la que se anclan; A) La masa de color amarillento es la puesta de embriones que depositó el individuo situado a la parte izquierda de la imagen sobre el otro individuo; B) Puestas retiradas mantenidas en una placa con AFSW con antibióticos. En la placa de la derecha se pueden observar dos bolsas de embriones individuales. Las puestas se componen de un grupo de estas bolsas, unidas en un pedúnculo.

Los huevos se disponen en puestas que se parecen a un racimo de cerezas (figura 12). Hasta 300 embriones síncronos se pueden incluir en un saco (cerezas); un total de 20 sacos se unen a través de un filamento (el pedúnculo de la cereza) para componer una única puesta.

C. fornicata presenta una segmentación espiral de tipo dextrógira, en la que el primer cuarteto de micrómeros experimenta un giro en el sentido de las agujas del reloj (ver apartado 1.5.). A medida que se suceden las divisiones (estadio de blástula), los micrómeros que se sitúan en el polo animal van aumentando en número y cubriendo los macrómeros, proceso conocido con el nombre de epibolia (figura 13 A-B). Los cuadrantes A, B, C y D de una blástula pueden identificarse con un estudio celular detallado: primero se localiza el primer cuarteto que define un cuadrado imaginario; desde cada uno de los vértices de este cuadrado imaginario se traza una línea recta hacia el polo vegetal siguiendo la línea de células; esto delimita los cuatro cuadrantes; el cuadrante con mayor número de células es el cuadrante D y una vez definido éste, se pueden definir los demás cuadrantes A, B y C en el sentido de giro horario.

Posteriormente, empieza la gastrulación (figura 13 C-E). El comienzo del proceso puede distinguirse porque en vista vegetal (desde los macrómeros) los grupos de micrómeros son ya muy numerosos y se establece el inicio del blastoporo (gástrula temprana). A medida que el blastoporo va cerrándose (gástrula media), se produce un abultamiento en la zona posterior-derecha en una vista ventral, que corresponde al cuadrante C. Poco después el embrión se alarga

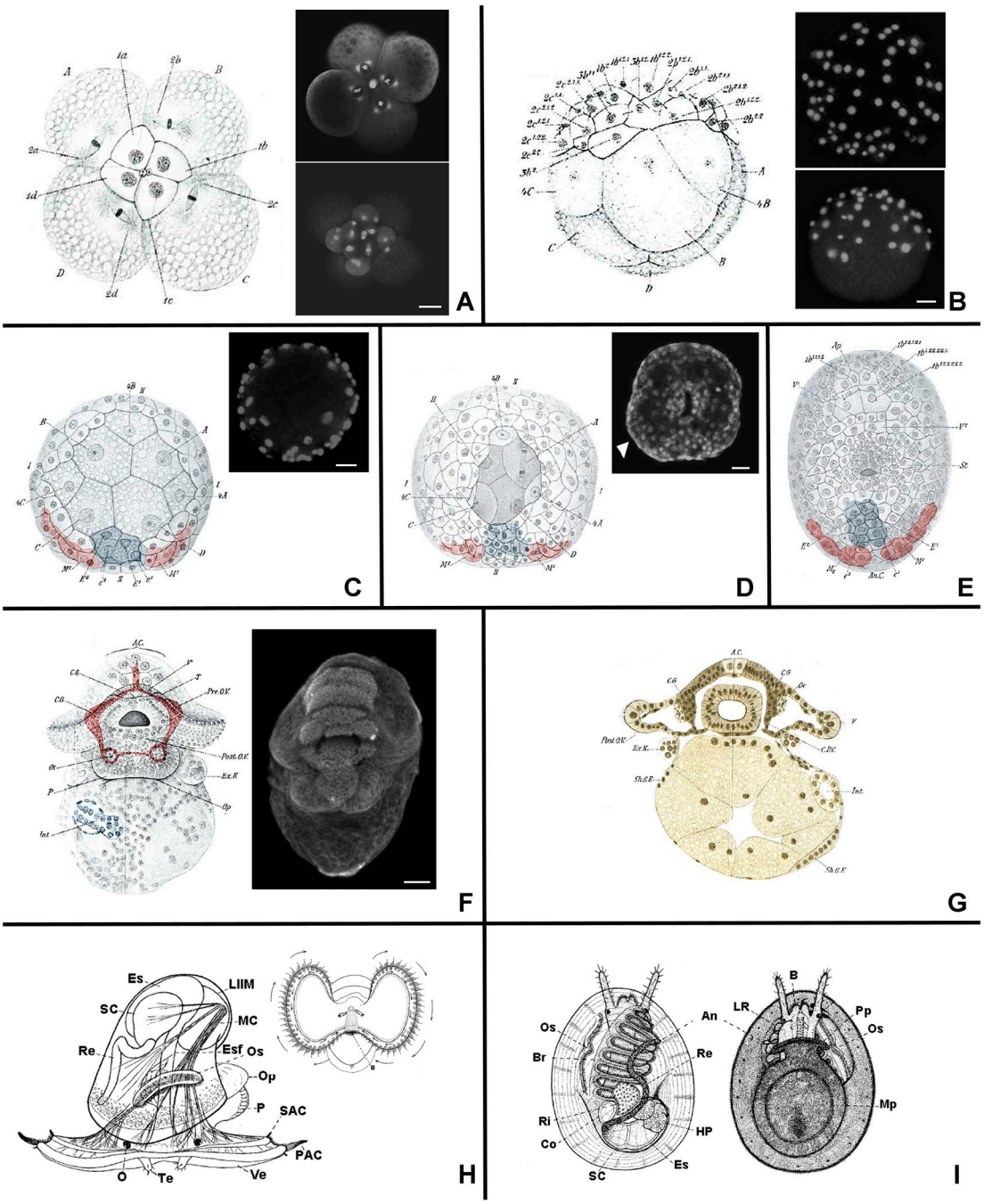


Figura 13. Desarrollo embrionario de *C. fornicata.* A) Estadio de segmentación, en una división de 8 a 12 células. Las fotografías de tinciones de núcleos y Tubulina (a la derecha) muestran un embrión al comienzo de la metafase de la tercera división (panel superior) y un embrión dividiéndose de 12 a 16 células (panel inferior); B) Vista lateral de un embrión en estadio blástula. El casco de micrómeros se sitúa en el polo animal y los macrómeros en el polo vegetal. Las fotografías de tinción de núcleos (a la derecha) muestran una vista desde el polo animal de una blástula (panel superior) en la que puede apreciarse la cruz delimitante de los cuadrantes y una vista lateral de la misma (panel inferior); C) Vista vegetal (futura vista ventral) de un embrión en estadio de gástrula temprana, en epibolia avanzada. La fotografía de tinción de núcleos (a la derecha) muestra un embrión real en este estadio; D) Vista ventral de gástrula media. La fotografía es una tinción de núcleos y Tubulina de este estadio; puede distinguirse el característico abultamiento del cuadrante C (marcado con una punta de flecha); E) Vista ventral de gástrula tardía; F) Estadio preveliger. El blastoporo conforma la boca, se distinguen los velos; la organogénesis está en curso. La fotografía muestra una tinción de Tubulina, que permite observar los cilios; G) Estadio veliger, en proceso de torsión; H) Vista frontal de una larva veliger tras la eclosión. En el detalle pueden verse los velos ciliados extendidos que con su movimiento permiten que esta larva nade libremente; I) Estadio juvenil en vista dorsal (derecha) y ventral (izquierda). Esquemas tomados de Conklin (1897), Werner (1955) y fotografías propias. La escala corresponde a 30 µm.

y el blastoporo, que conformará la futura boca, se cierra y se desplaza anteriormente (estadio de gástrula tardía).

Durante el siguiente estadio, larva preveliger, comienza la organogénesis (figura 13 F); en este momento se forman estructuras como el velo (estructura ciliada que se sitúa alrededor de la parte anterior del embrión y permite que la larva pueda nadar y desplazarse) y el rudimento del pie. Además también se van formando el ano, el intestino y el resto de órganos.

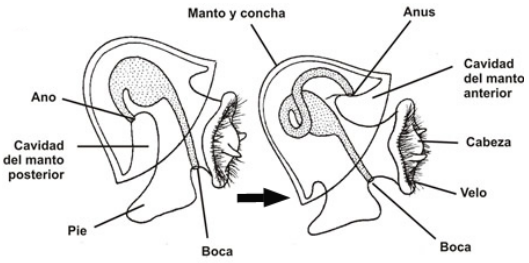


Figura 14. Proceso de torsión en gasterópodos. Durante el desarrollo de la larva veliger hasta la aparición del estadio juvenil, ocurre un proceso en el que la masa visceral rota 180°sobre el pie y la cabeza, en torno a su eje longitudinal. Este sentido de giro se relaciona con la organización dextrógira y levógira del individuo. (Livingstone, © BIODIDAC)

Finalmente, en un estadio más avanzado del desarrollo, se da el proceso de torsión y se forma la larva veliger (figuras 13 H y 14). Esta larva cuenta con una concha más desarrollada, órganos más definidos y cabeza con órganos sensoriales desarrollados que incluye un borde ciliado lleno de sensores. Esta larva alcanza la etapa juvenil con una morfología similar a la del adulto (figura 13 I).

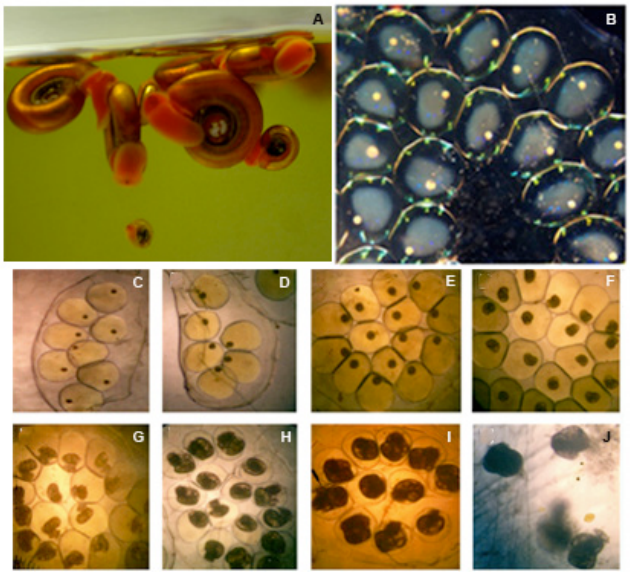


Figura 15. Los adultos de la especie *B. glabrata* y sus puestas de embriones. A) Adultos de la especie levógira *B. glabrata*; B) puesta de embriones en estadio de segmentación. Cada cápsula contiene un embrión individual creciendo envuelto en vitelo; C) blástulas; D) gástrulas; E) larvas trocóforas; F) larvas veliger temprana; G) larvas veliger; H-I) individuos en estadio juvenil; J) juveniles tras la eclosión. Figura tomada y modificada de De Albuquerque *et al.* (2014) y fotografías propias.

3.2. *B. glabrata* como modelo de estudio

B. glabrata es una especie de gasterópodo de agua dulce de la familia Planorbidae, incluida dentro de Pulmonata. Esta especie es conocida por su importancia médica, como hospedador del trematodo *Schistosoma mansoni,* responsable de una grave enfermedad llamada esquistosomiasis (Adema *et al.*, 2017).

B. glabrata tiene una concha aplanada enrollada en sentido contrario a las agujas del reloj (levógira), que siempre cae del revés a un lado del

animal pareciendo dextrógira (figura 15 A). Esta especie es una de las poquísimas especies de gasterópodos exclusivamente levógira.

Los adultos de *B. glabrata* son hermafroditas y pueden autofecundarse, aunque generalmente la fecundación suele ser cruzada. La época reproductiva de esta especie comprende todo el año. Cada puesta está formada por entre 10-30 cápsulas y cada cápsula contiene un embrión (figura 15 B-J). Todos los embriones de una puesta están sincronizados en el desarrollo. A una temperatura de unos 20°C, el tiempo de desarrollo de los embriones es corto (1 semana hasta la eclosión) y los juveniles alcanzan la madurez sexual a los 2 meses de vida.

Los embriones de *B. glabrata* presentan una segmentación espiral de tipo levógira (figura 16), en la que el primer cuarteto de micrómeros experimenta un giro en el sentido contrario al de las agujas del reloj (ver apartado 1.5). En el estadio de blástula, los micrómeros van dividiéndose y la epibolia comienza (figura 16 B). Al igual que en *Crepidula*, las blástula se pueden orientar localizando primeramente el primer cuarteto, que define un cuadrado imaginario. Dado su patrón inverso de segmentación con respecto a uno dextrógiro, el cuadrante C se localiza a la izquierda del cuadrante D en una vista desde el polo animal (figura 16 B).

La gastrulación ocurre a las 26 horas tras la primera división. A medida que se sucede el recubrimiento de los macrómeros por parte de los micrómeros, se forma el blastoporo, que se cierra por completo a las 39 horas tras la primera división. Tras el cierre del blastoporo, éste se desplaza anteriormente y dará lugar a la futura boca. Entonces, el embrión se alarga (estadio de gástrula tardía) (figura 16 C).

En la siguiente etapa se forma la larva trocófora que se caracteriza por poseer la prototroca, un anillo de cilios situado por encima de la boca que permite el movimiento de la larva (figura 16 D-E). La boca se sitúa en la zona ventral de la prototroca, al igual que el rudimento del pie. En la zona dorsal se localiza el glándula de la concha, donde ya puede apreciarse una fina capa mineralizada. Los órganos internos comienzan a diferenciarse. La trocófora se convierte gradualmente en larva veliger a medida que crecen el pie, los velos y la concha, que además se desplaza lateralmente. Más tarde en el desarrollo de la larva veliger tiene lugar la torsión, se desarrolla por completo la cabeza con los ojos y los tentáculos, el velo degenera, la boca se sitúa entre éste y el pie, y la concha se empieza a enrollar (figura 16 F-G). Finalmente, los juveniles, con aspecto de adulto, salen de la cápsula.

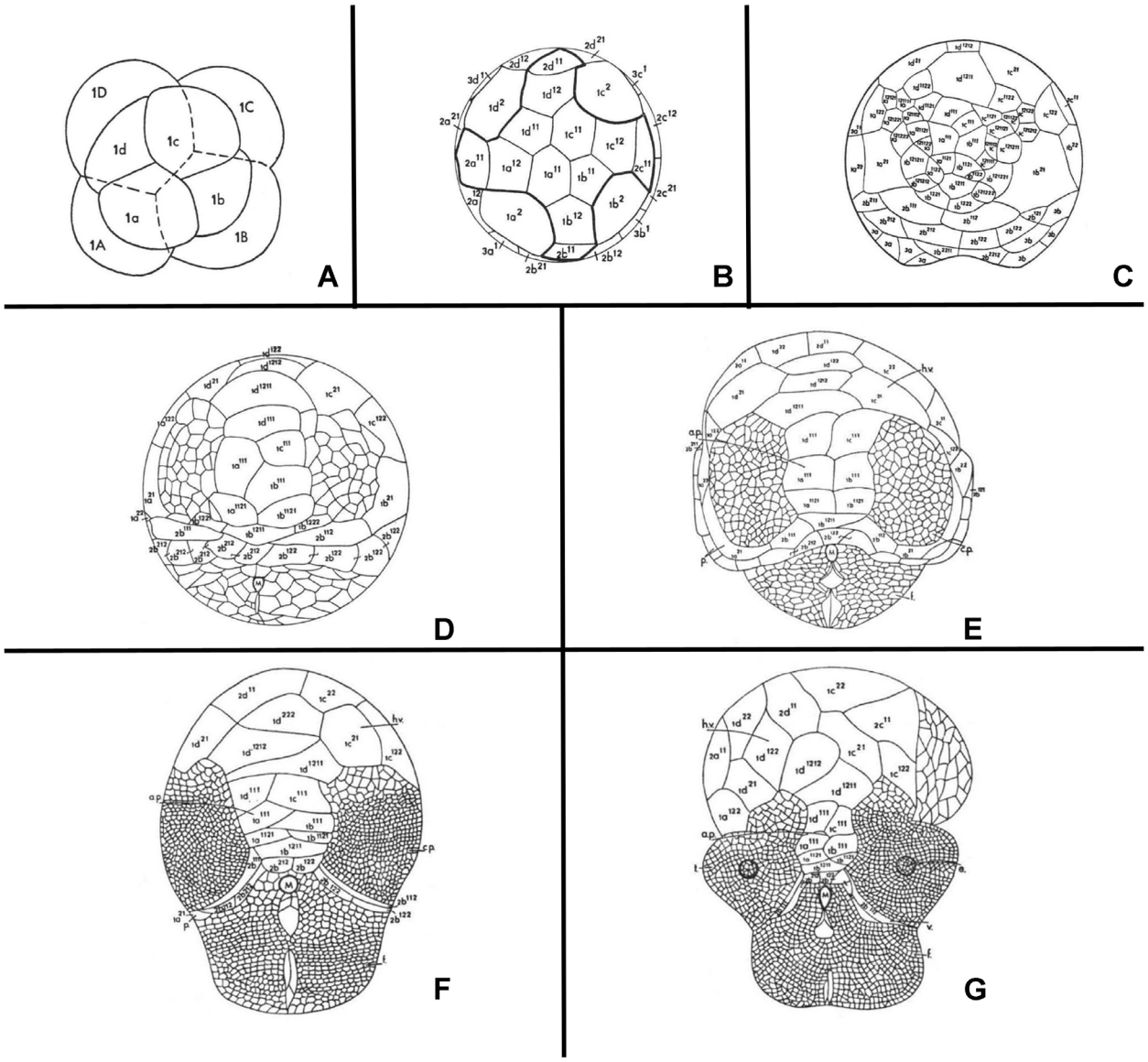


Figura 16. Desarrollo embrionario de *Biomphalaria*. A) Estadio de segmentación en un embrión de 8 células; los micrómeros rotan en sentido antihorario; B) vista animal de un embrión en estadio blástula. El trazo grueso delimita la cruz que define los cuadrantes; C) larva trocófora temprana con la futura boca formada (parte inferior del dibujo); D) vista ventral de una larva trocófora tardía; E) vista ventral de gástrula tardía; F) estadio de larva veliger temprana; G) estadio veliger tardía. Esquemas tomados de Camey (1969).

3.3. Mantenimiento de adultos de las especies de gasterópodos estudiadas; obtención y cultivo de sus embriones

Los adultos de *C. fornicata* utilizados en este estudio se recolectaron por técnicos del Marine Biological Laboratory (MBL, Woods Hole, Massachussets, Estados Unidos) durante los meses de enero y febrero, cuando sus gónadas se encuentran llenas de gametos. Los adultos son mantenidos en agua marina a una temperatura de entre 9 y 11°C prolongando de forma indefinida su estado de latencia. La inducción de los adultos se realiza atemperando gradualmente los

adultos durante una hora en agua entre 18 y 21°C para inducir la puesta. Una vez que un adulto produce huevos, no lo vuelve a hacer hasta la siguiente temporada.

En los meses de verano se trabajó con esta especie en el MBL y en los meses de invierno la experimentación se realizó en el laboratorio del Prof. Henry en la University of Illinois (campus de Urbana-Champaign) en sucesivas estancias breves realizadas en los años 2013, 2014, 2015 y 2016 (18 meses en total). El laboratorio del Prof. Henry en la Universidad de Illinois desarrolló un sistema para mantener los adultos en estado de latencia en agua marina artificial (*Artificial Filtered Sea Water*, AFSW) y a una temperatura entre 9 y 11°C. Durante el transporte de los *stacks* enviados por el MBL se respetó en todo momento la cadena de frío. Una vez en el laboratorio, según la necesidad de embriones, los adultos son transferidos por la tarde a un acuario acondicionado entre 18 y 21°C con un sistema de filtros y corriente de AFSW para inducir la deposición de puestas que son recogidas a la mañana siguiente (ver Henry *et al*., 2006 y 2010 para más detalles). Las puestas pueden mantenerse en placas con AFSW con antibióticos (penicilina-G 6 mg/ml y estreptomicina 20 mg/ml) para que continúen su desarrollo y utilizar los estadios convenientes para cada experimento.

Los adultos de *B. glabrata* de la línea M-line se obtuvieron de Biomedical Research Institute Resources (National Institute of Allergy and Infectious Diseases, NIAID; The National Institutes of Health, NIH). Una población de cerca de 200 individuos se mantuvo en el laboratorio de la Universidad Autónoma de Madrid en tanques de agua dulce a 25°C. El agua (*Treated Fresh Water*, TFrW) utilizada es agua corriente tratada con productos de acuario para regular el pH (JBL) y un suplemento de calcio (Biochem). En estas condiciones la obtención de embriones es continua durante el año y las puestas fueron sistemáticamente recolectadas para los estudios aquí descritos. El desarrollo es completo y normal manteniendo las puestas en las cápsulas y en el mismo agua en el que viven los adultos TFrW. Las puestas pueden mantenerse en placas con TFrW, para que continúen su desarrollo y utilizar los estadios convenientes para cada experimento.

3.4. Fijación de embriones

Para la fijación de embriones de la especie *C. fornicata* se siguió el método descrito en Henry *et al.* (2010) que se resume a continuación. Los embriones se recolectaron abriendo con pinzas las bolsas que los contienen, se lavaron con AFSW y se transfirieron a un tubo tipo eppendorf donde se fijaron en formaldehído 3,7% durante 15 minutos para experimentos de

tinción con anticuerpos y 1 hora para experimentos de hibridaciones *in situ*. A continuación se lavaron en PBS, se deshidrataron en 3 diluciones seriadas de etanol al 30, 70 y 100% y finalmente se conservaron a -20°C para su posterior uso.

Los embriones de *B. glabrata* fueron recolectados rasgando las cápsulas protectoras con agujas afiladas y extrayéndolos con una pipeta Pasteur. Posteriormente fueron transferidos a un recipiente de vidrio con TFrW para limpiar el vitelo. Tras dos lavados rápidos en TFrW, se fijaron en formaldehído 3,7% diluido en agua miliQ durante 15 minutos para experimentos de tinción con anticuerpos y 1 hora para experimentos de hibridaciones *in situ*. Posteriormente, los embriones ya fijados fueron lavados 3 veces con agua miliQ y deshidratados progresivamente en 3 diluciones seriadas de etanol al 30, 70 y 100%. Finalmente los embriones fueron conservados a -20°C para su posterior uso.

Para los experimentos de *western blot* realizados en ambas especies los embriones de cada especie fueron contados y transferidos a un tubo tipo eppendorf. Una vez que los embriones caían al fondo el agua era retirada y los embriones congelados inmediatamente a -80°C para su posterior uso.

También se se conservaron embriones de *C. fornicata* en RNA later (Ambiom, AM7020) a -80°C para posteriormente ser utilizados en protocolos de extracción de RNA.

3.5. Generación de bases de datos transcriptómicas para las especies *C. fornicata* y *B. glabrata*.

Durante el desarrollo de esta tesis se generaron dos bases de datos transcriptómicas de la especie *C. fornicata* (una que incluía RNA procedente de una mezcla de estadios desde huevo hasta larva veliger y otra que incluía RNA exclusivamente de larvas preveliger) y una base de datos transcriptómica de la especie *B. glabrata* (que incluía RNA procedente de una mezcla de estadios desde huevo hasta juvenil). La metodología empleada está descrita en Perry *et al*. (2015) (apéndice 11) y Kenny *et al*. (2016) (apéndice 12) para cada una de las especies, respectivamente. Mi participación se centró en la fijación de embriones en diferentes estadios para ambas especies, la extracción del RNA para cada muestra de embriones y la preparación del material para la posterior realización de las correspondientes genotecas. Las bases de datos fueron instaladas y manejadas en el programa Geneious 6.1.2 (Kearse *et al*., 2012).

3.6. Búsqueda de genes candidatos y establecimiento de ortología

Siguiendo los objetivos propuestos en esta tesis se llevaron a cabo búsquedas en distintas bases de datos (apéndice 2) por medio de tBlastp (Altschul *et al.,* 1990) de los siguientes genes: *inversin*, *nodal* y *egf-cfc*.

Una vez encontrados los candidatos potenciales para cada uno de los genes se comprobó la ortología de los mismos a través de distintos análisis filogenéticos. El tratamiento de secuencias se realizó principalmente en MacVector v12.7 (Olson *et al.,* 1994). Los alineamientos se trabajaron en ClustalX v2.1 (Larkin *et al.,* 2007) y Mesquite v2.75 OSX (Maddison & Maddison, 2011). Los análisis filogenéticos se realizaron en BEAST v1.8.0 (Drummond *et al*., 2012) y MrBayes v3.2.4 (Huelsenbeck & Ronquist, 2001). El modelo óptimo de sustitución se calculó usando Prot-test v.2.4 (Abascal *et al*., 2005). La convergencia de los parámetros se visualizó en Tracer v1.5 (Rambaut & Drummond, 2007). Se realizó un *burn-in* del 25% y los árboles resultantes se consensuaron por medio de TreeAnnotator v1.8.0 (paquete de Beast).

La identificación de los dominios presentes en la proteína Inversin se realizó por medio de búsquedas manuales de los dominios consensos disponibles en la literatura y de la aplicación *Protein Pattern Find* (http://www.bioinformatics.org/sms2/protein_pattern.html).

3.7. Extracción de RNA y producción de cDNA

La extracción de RNA de embriones de *Crepidula*, se realizó a partir de embriones previamente conservados en RNA later (Ambiom, AM7020) a -80°C. Por otro lado, las extracciones realizadas en *Biomphalaria*, se realizaron a partir de embriones frescos. Los embriones de ambas especies se centrifugaron 30 segundos a máxima velocidad para acumularlos en el fondo del tubo y se retiró el líquido en el que se encontraban hasta dejar aproximadamente 10 µl. Se homogenizaron con pequeños émbolos de plástico en 500 µl de Trizol (In vitrogen, 15596026). A continuación, la muestra se incubó 5 minutos a temperatura ambiente, se añadieron 100 µl de cloroformo y se agitó fuertemente durante 15 segundos. Posteriormente, se incubó a temperatura ambiente 2 minutos y se centrifugó a una velocidad no superior a 12000 x g durante 15 minutos a 4°C para separar las fases. Se transfirió la fase acuosa transparente de la parte superior de tubo, que contiene el RNA, a otro tubo limpio. Después se precipitó el RNA añadiendo 250 µl de isopropanol y se dejó incubando durante toda la noche. Posteriormente, se centrifugó a 12000 x g durante 10 minutos a 4°C y se retiró el sobrenadante. El sedimento obtenido se lavó con

900µl de etanol al 75%, dos veces, resuspendiendo con vórtex y centrifugando a 10000 x g durante 5 minutos a 4°C. Una vez que el sedimento se secó brevemente, se disolvió en 20 µl de agua tratada con DEPC. A partir del RNA extraído se generó cDNA, con el kit SMARTer PCR cDNA Synthesis Kit (Clontech), según las instrucciones del fabricante.

3.8. Amplificación por PCRs y electroforesis en gel de agarosa

La amplificación de genes a partir del cDNA generado se realizó a través de PCRs imbricadas con dos parejas de cebadores específicos diseñados a partir de las secuencias ortólogas identificadas previamente (tabla 1; apéndice 3). En una primera PCR se utilizaron los cebadores más externos y el cDNA como molde (25µl de volumen final de reacción). En una segunda PCR se utilizó la pareja de cebadores internos y el producto obtenido de la primera PCR como molde (50 µl de volumen final). Estas amplificaciones se realizaron con Taq platinum (Invitrogen), siguiendo las instrucciones del fabricante. El chequeo de la amplificación se realizó por electroforesis en gel de agarosa (agarosa D-1 Low EEO, Pronadisa, Laboratorios Conda) 1,5% con bromuro de etidio. Los productos deseados se purificaron de banda en gel de agarosa 0,8% por columna con el Kit Zymoclean Gel DNA Recovery (ZymoResearch). La muestras se eluyeron en un volumen final de 10 µl de agua estéril.

3.9. Clonaje, transformación y comprobación de colonias positivas

El clonaje de los productos amplificados por PCR se realizó en el vector pGemTeasy (Promega, A1360) de acuerdo con las instrucciones del kit. La transformación de 50 µl de bacterias quimio-competentes *Escherichia coli DH5α* con 5 µl de producto de ligación se llevó a cabo por choque térmico (30 minutos en hielo, 30 segundos a 42°C, 2 minutos en hielo, incubación a 37°C durante una hora tras la adición de 250 µl de medio SOC (Invitrogen). Tras la incubación, se plaquearon 150 µl del volumen en una placa con LB-agar con ampicilina 100 µg/ml, IPTG (1 mM) y X-Gal (40 µg/ml) y se dejaron crecer durante la noche en una estufa a 37°C. Las colonias fueron seleccionadas por color y posteriormente se crecieron en medio LB con ampicilina a 100 µg/ml durante una hora a 37°C. El tamaño de los insertos se chequearon por PCR (12,5 µl volumen final de reacción), a través de un programa corto de 25 ciclos, a 50°C de temperatura de anillamiento, con los cebadores M13F y M13R. Estas amplificaciones se realizaron con Taq DNA polimerasa (Invitrogen) siguiendo las instrucciones del fabricante.

Gen amplificado	Fragmento	Pareja de cebadores	Condiciones de temperatura y extensión	Tamaño del producto
***inversin*_CF**	Inversin_1	InvCrepi_F//InvCrepi_R InvCrepi_Fin//InvCrepi_Rin	55°C, 1' 55°C, 1'	1,2 Kb
	Inversin_2	InvCrepi_F3//InvCrepi_R2 InvCrepi_F3in//InvCrepi_R2in	50°C, 2' 50°C, 2'	1,4 Kb * 780 pb
	Inversin_3	InvCrepi_F4//InvCrepi_R3 InvCrepi_F4in//InvCrepi_R3in	45°C, 1' 45°C, 1'	693 pb
***inversin*_BG**	InvBG_anks	Biom-Inv-R//GR5'	Programa touch-down; anillamiento 72°C, 70°C, durante 30'' cada uno, 5 ciclos 65°C, 30'', 25 ciclos 62°, 1'30''	1,5 Kb
		Biom-Inv-Rin//GR5'Nested	Programa touch-down; anillamiento 72°C, 70°C, durante 30'' cada uno, 5 ciclos 65°C, 30'', 25 ciclos 62°, 1'30''	
	InvBG_ct	InvBiom_F3//InvBiom_R4 InvBiom_F3in//InvBiom_R4in	55°C, 1'30" 55°C, 1'30"	665 pb
***cripto*_CF**	Cripto_1	CriptoCrepi_F4//CriptoCrepi_R3 CriptoCrepi_F4in//CriptoCrepi_R3in	55°C, 1' 55°C, 1'	560 pb
	Cripto_2	CriptoCrepi_F3// CriptoCrepi_F5 CriptoCrepi_F3in// CriptoCrepi_R5in	55°C, 1' 55°C, 1'	653 pb
	Cripto_3	CriptoCrepi_F2//UMP	Programa touch-down; anillamiento 72°C, 70°C, durante 30'' cada uno, 5 ciclos 69°C, 30'', 25 ciclos 62°, 1'30''	693 pb
		CriptoCrepi_F2in//NUP	Programa touch-down; anillamiento 72°C, 70°C, durante 30'' cada uno, 5 ciclos 65°C, 30'', 25 ciclos 62°, 1'30''	
***cripto*_BG**	CriptoBG_CDS	CriptoBiom_F//CriptoBiom_R CriptoBiom_Fin//CriptoBiom_Rin	55°C, 1' 55°C, 1'	612 pb
***nodal2*_CF**	Nodal2_CDS	Nodal2Crepi_F3//Nodal2Crepi_R2 NodalCrepi_F3in//Nodal2Crepi_R2in	50°C, 1'30" 50°C, 1'30"	1,2 Kb

Tabla 1. Parejas de cebadores, condiciones de amplificación y tamaño esperado para distintas sondas hechas para los genes de interés. CF, *C. fornicata*; BG, *B. glabrata*, Anks, región de dominios ankirina; ct, extremo carboxi-terminal; CDS, región codificante; pb, pares de bases; kb, kilobases; (*)Esta amplificación genera un doble producto correspondiente a la región de cada una de las dos isoformas, L de 1,4 Kb y S de 780 pb. Secuencia de los cebadores en el apéndice 3.

3.10. Preparaciones de DNA y secuenciación

Una vez comprobado el tamaño esperado de los fragmentos clonados, las colonias correspondientes se crecieron en matraces con 10 ml de LB con ampicilina 100 µg/ml, toda la noche a 37°C con agitación. Al día siguiente, se reservaron como *back up* 500 µl del cultivo saturado de cada colonia en un tubo estéril junto con 100 µl de glicerol 85% y se congelaron a -80°C inmediatamente para futuros usos. Del cultivo restante se extrajo el plásmido con el inserto por *miniprep*, utilizando el kit Wizard Plus SV Miniprep DNA Purification System (Promega) de

acuerdo con las instrucciones del fabricante. La muestras se eluyeron en un volumen final de 30 µl de agua estéril.

La concentración y pureza del plásmido extraído fueron cuantificadas con un espectrofotómetro Thermo Scientific NanoDrop 2000. Las muestras fueron enviadas al servicio secuenciación del Parque Científico de Madrid y/o a la empresa Secugen. Para el análisis de las secuencias se utilizaron los programas FinchTV v.1.4.0 (Geospiza, Inc.; Seattle, WA, USA; http://www.geospiza.com), MacVector v12.7 (Olson *et al.,* 1994) y Sequencher v.5.0 (Gene Codes Corporation, Ann Arbor, MI USAhttp://www.genecodes.com).

3.11. Transcripción *in vitro* para sondas e hibridación *in situ*

Para la transcripción se realizó una PCR utilizando una dilución 1:10 de la minipreparación como molde, en un volumen final de 100 µl, con la polimerasa Taq (Invitrogen) y cebadores M13F y M13R. Tras chequear en gel de agarosa 1,5%, la amplificación se limpió por columna utilizando el kit MiniElute PCR purification Kit (Qiagen) y se resuspendió en 30 µl.

A continuación se preparó la reacción de transcripción, de volumen final 21 µl, utilizando 2 µl del producto de PCR como molde [mezcla: 12 µl de agua estéril, 2 µl de buffer de transcripción 10X (Roche), 2 µl de polimerasa T7 o SP6 (Roche), 2 µl de digoxigenina (Roche) y 1 µl de inhibidores de RNAsas (New England, Biolabs)]. La mezcla se incubó a 37°C durante un periodo de 4 a 5 horas, tras el cual se añadió 1 µl de DNAsa (New England, Biolabs) para eliminar el molde y se dejó incubando otros 15 minutos a la misma temperatura. Por último, se limpiaron y concentraron las sondas obtenidas con el kit RNAeasy MInElute Clean-up kit (Qiagen) o con el kit RNA Clean & concentrator -5 (ZymoResearch). Para las reacciones de hibridación *in situ* se utilizó el protocolo indicado en Perry *et al*., (2015) (apéndice 11).

3.12. Selección y diseño de anticuerpos para las proteínas de estudio

Para la detección de las proteínas estudiadas (Inversin, Nodal y Cripto) durante el desarrollo embrionario de *C. fornicata* y *B. glabrata* se usaron diferentes tipo de anticuerpos. En unos primeros análisis se usaron anticuerpos comerciales para las proteínas Inversin (10585-1-AP, Proteintech) y Cripto (ab19917, Abcam).

Posteriormente, también se diseñaron y sintetizaron anticuerpos específicos para Inversin, Nodal y EGF-CFC para ambas especies. Las secuencias de estas proteínas fueron analizadas en busca de: 1) dominios consenso y 2) dominios de unión al complejo de histocompatibilidad tipo

II (MHC II). Además se hicieron análisis de homología de estos dominios con bases de datos genómicas y transcriptómicas de otros moluscos para intentar descartar zonas parecidas entre diferentes especies y elegir así las zonas con mayor especificidad para la especie concreta de trabajo (apéndice 4). El diseño y la posterior síntesis se realizó conjuntamente con la empresa 21st Century Biochemicals, Inc (tabla2).

Especie	Proteína	Péptidos	Conejos	Sueros	Elección
Crepidula	EGF-CFC	2	I6745 (A) I6747 (B)	Pre-inoculación –SPI Post-inoculación – SI (1 a 5)	B potencial
	Inversin	3	I6748 (A) I6749 (B)	Pre-inoculación –SPI Post-inoculación – SI (1 a 4)	B 1:500 Péptido 1
	Nodal2	2	I6751 (A) I6752 (B)	Pre-inoculación –SPI Post-inoculación – SI (1 a 4)	B 1:500
Biomphalaria	EGF-CFC	2	I6753 (A) I6754 (B)	Pre-inoculación –SPI Post-inoculación – SI (1 a 4)	B 1:750
	Inversin	3	I6759 (A) I6760 (B)	Pre-inoculación –SPI Post-inoculación – SI (1 a 4)	B 1:250

Tabla 2. Estrategia de producción de anticuerpos específicos generados contra las proteínas EGF-CFC, Inversin y Nodal2. Para cada proteína se indica el número de péptidos diseñados, el número identificativo de los conejos utilizados y los sueros extraídos de cada conejo. Los sueros se extrajeron a los 15 días de la inoculación de los péptidos (suero 1), a los 30 días (suero 2), a los 45 días (suero 3), a los 60 días (suero 4) y a los 90 días (suero 5).

3.13. Electroforesis en gel de acrilamida y *western blot*

En primer lugar se realizó una extracción de proteína total considerando que se necesitan de 7 a 10 embriones por cada carril usado en el gel de acrilamida. Los embriones se homogeneizaron en solución Laemli 2X de acuerdo a la proporción de 1 µl de tampón por cada embrión. La muestra se agitó en vórtex, se hirvió a 90°C durante 10 minutos y se centrifugó a máxima velocidad durante otros 10 minutos. El sobrenadante se transfirió a un tubo limpio para su posterior uso. El marcador de peso molecular Precision Plus protein Standard dual color (BioRad) fue usado durante todos los experimentos. El porcentaje de acrilamida del gel utilizado dependió del tamaño de la proteína testada; en concreto se utilizaron geles al 6% (Inversin), 10% (Nodal2 y EGF-CFC) y 12% (EGF-CFC). La transferencia a membrana de las proteínas tras la electroforesis fue de tipo semiseca (semy-dry cell, BioRad) con buffer de transferencia Towbin (25mM Tris, 192 mM glicina, 20% metanol v/v). Todos los componentes de la transferencia (papeles, membrana y gel) se empaparon durante 15 minutos en buffer de transferencia Towbin. Las condiciones de transferencia utilizadas fueron 15V, de 15 a 20 minutos. Una vez comprobado

que la transferencia se llevó a cabo correctamente tiñendo con rojo Ponceau, la membrana de nitrocelulosa se lavó en agua y se secó durante 1 hora. Posteriormente la membrana se lavó 5 veces en TBS-T (10 mM Tris pH 7,5; 150 mM NaCl; 0,1%v/v Tween20), se bloqueó en solución de bloqueo en TBS de la casa LiCor, con 0,1% Tween durante 1 hora. Posteriormente se incubó toda la noche en la solución de bloqueo junto con el anticuerpo primario a 4°C. Para el control de carga se utilizó un anticuerpo anti-β-Tubulina a concentración 1:300 (Hybridoma bank, E7). Tras 5 lavados en TBS-T, la membrana se incubó con anticuerpos secundarios con un fluoróforo acoplado (anti-conejo 680 and anti-ratón 800, Oddysey LiCor), usando la solución de bloqueo a una concentración 1:20000 durante 45 minutos, en oscuridad y a temperatura ambiente. Tras 3 lavados en TBS-T y 2 lavados en TBS, la membrana se visualizó con el escáner Oddysey (LiCor).

3.14. Inmunohistoquímica en embriones completos

Los embriones previamente conservados en etanol absoluto fueron rehidratados progresivamente con lavados de 5 minutos en PBS-etanol al 60 y 30%. Para la permeabilización se realizaron 3 lavados de 5 minutos en PBS-tritónX100 (PBTr) al 0,2%. Los embriones se bloquearon en una solución 2% BSA en PBTr durante 1 hora. A continuación se incubaron en solución de bloqueo junto con los anticuerpos primarios durante toda la noche a 4°C. Como control positivo y con el objetivo de marcar el citoesqueleto se utilizó el anticuerpo anti α-Tubulina a 1:20 (Hybridoma bank, 12G10). A continuación, los embriones se lavaron 5 veces en PBTr y se incubaron en solución de bloqueo junto con los anticuerpos secundarios (Goat anti-IgG Rabbit 594, Molecular probes A-11037; Donkey anti-IgG Mouse, Molecular probes A-21202) durante 2 horas en oscuridad y a temperatura ambiente. Posteriormente, los embriones se lavaron 3 veces en PBTr. Los núcleos se tiñeron con Hoestch a una dilución 1:1000 en PBTr durante 30 minutos. Finalmente, tras 5 lavados en PBTr los embriones se incluyeron en glicerol 80% para su posterior observación.

3.15. Mutagénesis: Generación del molde *wild type* (wt)

Para generar mRNA wt de los genes estudiados se usó el vector pCS2+ (Addgene) en el que se insertó la zona codificante completa del gen deseado (apéndice 5). Para preparar el inserto se amplificó mediante PCR la zona codificante completa del gen estudiado incluyendo

el ATG inicial pero no el codón de parada. Los cebadores usados se diseñaron por pares que solapan parcialmente dos a dos. Para su posterior clonaje en el vector pCS2+ se incluyeron en la pareja de cebadores más interna unas dianas de restricción flanqueantes (BamHI, XhoI, EcoRI y SalI) (tabla 3; apéndice3). Las reacciones de PCR se realizaron con la polimerasa Phusion High Fidelity (New England, Biolabs) siguiendo las instrucciones del fabricante, en 25 µl de volumen final para la PCR con cebadores externos y 50 µl de volumen final para la PCR con cebadores internos. Las reacciones de PCR se realizaron según las siguientes condiciones:

Gen amplificado	Pareja de cebadores	Condiciones	Tamaño
***inversinL*_CF**	InvCrepi_F5//InvCrepi_R5 InvCrepi_F5in//InvCrepi_R5in InvCrepi_BI_F//InvCrepi_XI_R	50°C, 4' 50°C, 4'	3'3 Kb y 2,7 Kb
***inversin*_BG**	InvBiom_F5//InvBiom_R5 InvBiom_BI_F//InvBiom_XI-R	45°C, 3' 50°C, 3'	2,4 Kb
***egf-cfc*_CF**	CriptoCrepi_F5// CriptoCrepi_R4 CriptoCrepi_F5in// CriptoCrepi_R4in CriptoCrepi_BI_F//CriptoCrepi_XI_R	50°C, 1' 50°C, 1' 60°C, 1'	627 pb
***oep*_ZF**	Oep_ZF_F//Oep_ZF_R Oep_ZF_BI_F//Oep_ZF_SI_R	55°C, 1' 60°C, 1'	561 pb
***nodal2*_CF**	Nodal2_Crepi_F4//Nodal2_Crepi_R4 Nodal2_Crepi_BI_F//Nodal2_Crepi_R4_XI_R Nodal2_Crepi_Eco_F//Nodal2_Crepi_R4_SI_R	50°C, 2' 50°C, 2' 50°C, 2'	1,4 Kb

Tabla 3. Parejas de cebadores, condiciones de amplificación y tamaño esperado para el clonaje de las regiones codificantes completas de los genes de interés para esta tesis y su posterior incorporación en el vector de expresión pCS2+. CF, *C. fornicata*; BG, *B. glabrata*; ZF, *D. rerio* pb, pares de bases; Kb, kilobases; BI, BamHI; XI, XhoI; SI, SalI; Eco, EcoRI. Secuencia de los cebadores en el apéndice 3.

Los productos se purificaron de banda en gel de agarosa 0,8% por columna con el Kit Zymoclean Gel DNA Recovery (ZymoResearch). La muestras se eluyeron en un volumen final de 10 µl de agua estéril. La elución se digirió con 1,5 µl de las enzimas de restricción mencionadas anteriormente, 2 µl de buffer Cutsmart 10X (New England, Biolabs) en una reacción de volumen final 20 µl a 37°C durante 2 horas (*inversin*_CF, BamHI y XhoI; *inversin*_BG, BamHI y XhoI; *egf-cfc*_CF, BamHI y XhoI; *oep*_ZF, BamHI y SalI; *nodal2*_CF, EcoRI y SalI). El producto se limpió y concentró por columna con el kit QIAquick PCR purification kit (Qiagen), eluyendo en 10 µl de agua estéril.

Por otra parte, el vector se cortó con las mismas enzimas con las que se cortó el inserto. Se digirieron 2 µg de vector con 1,5 µl de cada enzima y 3 µl de buffer Cutsmart 10X en un volumen final de 30 µl (para *inversin*_CF, *inversin*_BG, *egf-cfc*_CF y *oep*_ZF se digirió con BamHI y XhoI, aprovechando que XhoI y SalI son isoesquizómeros; para *nodal2* se utilizaron EcoRI

y SalI). Tras 2 horas y media de digestión a 37°C, se añadió 1 µl de fosfatasa alcalina CIP para defosforilar los extremos y se incubó otra hora. El producto se limpió y concentró a partir de banda en gel de agarosa por columna con el kit Zymoclean Gel DNA Recovery (ZymoResearch) y se eluyó en 15 µl de agua estéril.

Finalmente, 6 µl de la purificación de inserto se ligaron con 100 ng de vector con la enzima T4 ligasa en su buffer correspondiente y en un volumen final de 10 µl. Esta reacción se incubó toda la noche a 17°C. Posteriormente, se transformaron 100 µl de bacterias competentes *E. coli* DH5α con todo el producto de ligación según los pasos descritos en el apartado 3.9. Las colonias obtenidas se chequearon por PCR de colonias y aquéllas positivas se crecieron en medio LB líquido para obtener el plásmido a través de mini preparaciones (apartado 3.10.). Finalmente, los vectores se enviaron a secuenciar con los cebadores SP6 y pCS2+_Rev (apéndice 3) para comprobar la correcta inserción del gen deseado dentro del vector.

3.16. Mutagénesis: Generación de mutaciones sobre el molde wt

Para generar mutaciones dirigidas sobre el molde wt de los genes estudiados se adaptó el protocolo del kit de Q5® Site-Directed Mutagenesis (New England, Biolabs). Las estrategias de diseño de los cebadores se corresponden con las sugeridas en el protocolo del kit según fuesen mutaciones puntuales, delecciones o inserciones (tabla 4; apéndice 6). Los cebadores se fosforilaron utilizando la enzima PNK (New England, BioLabs). La reacción de fosforilación incluyó 20 µl de cebador a concentración 10 µM junto con 4 µl de buffer de la enzima, 1 µl de enzima y 4 µl de ATP 10 mM en un volumen final de 40 µl. Esta reacción se incubó a 37°C durante 1 hora, y posteriormente 20 minutos a 65°C para inactivar la enzima.

A continuación, se prepararon las reacciones de PCR utilizando la polimerasa Phusion High Fidelity (New England, Biolabs) según las instrucciones establecidas por el fabricante. La reacción se ajustó a 35 ciclos, con temperatura de anillamiento a 50°C y tiempo de extensión entre 4 y 5 minutos (45 segundos por Kb). Para un volumen final de reacción de 50 µl se utilizaron 3,5 µl de cada cebador y 10 ng/µl máximo de DNA molde. Tras la amplificación, se añadió 1 µl de dpnI fast-digest (Thermo Fisher) y se incubó 15 minutos. A continuación se limpió por columna QIAquick PCR purification kit (Qiagen) y se eluyó en 10 µl de volumen final en agua miliQ. El producto se ligó utilizando el kit Quick ligation (New England, BioLabs), preparando una reacción de volumen final 20 µl, de acuerdo a las instrucciones del fabricante. La transformación se realizó en bacterias NEB® Turbo Competent *E. coli* (High Efficiency)

(New England, BioLabs), siguiendo el protocolo ya establecido por el fabricante. Los pasos posteriores de siembra, chequeo, cultivo, minipreparación de plásmidos y secuenciación son los ya descritos en los apartados 3.9. y 3.10.

Molde	Construcción mutante	Pareja de cebadores
***inversinL*_CF_wt_pCS2+**	InvL_CF_ΔIQ1	InvCrepi_ DIQ1_F// InvCrepi_ DIQ1_R
	InvL_CF_IQ2AA	InvCrepi_IQ2AA_F//InvCrepi_IQ2AA_R
***inversinL*_CF_ΔIQ1_ pCS2+**	InvL_CF_ΔIQ1+ IQ2AA	InvCrepi_IQ2AA_F//InvCrepi_IQ2AA_R
inversinC*_CF_wt_pCS2+	InvC_CF_ΔIQ1	InvCrepi_ DIQ1_F// InvCrepi_ DIQ1_R
	InvC_CF_IQ2AA	InvCrepi_IQ2AA_F//InvCrepi_IQ2AA_R
***inversinC*_CF_ΔIQ1_pCS2+**	InvC_CF_ΔIQ1+ IQ2AA	InvCrepi_IQ2AA_F//InvCrepi_IQ2AA_R
***inversin*_BG_wt_pCS2+**	INV_BG_IQAA	IQAA_IQ2Biom_F//IQAA_IQ2Biom_R
	CR_CF_ΔEGF	CR_DEGF_F//CR_DEGF_R
***egf-cfc*_CF_wt_pCS2+**	CR_CF_L56A	CR_L56A_F//CR_L56A_R
	CR_CF_L56T	CR_L56T_F//CR_L56T_R
***nodal2*_CF_wt_pCS2+**	NOD2_CF_ΔCS	Nd2_Crepi_DCS_F//Nd2_Crepi_DCS_R

Tabla 4. Moldes y parejas de cebadores utilizados en la generación de construcciones mutantes sobre el vector pCS2+. Los moldes se corresponden con los clonajes de la tabla 3. CF, *C. fornicata;* BG, *B. glabrata*; wt, silvestre; ΔIQ1, deleción del dominio IQ1; IQ2AA, sustitución de los aminoácidos I y Q del IQ2 por 2 alaninas; ΔEGF, deleción del dominio tipo EGF; L56A, sustitución de la leucina en posición 56 por una alanina; L56T, sustitución de la leucina en posición 56 por una treonina; ΔCS, delección de la región de corte. (*) Obtenido por PCR generando una deleción del sitio de splicing con los oligos INVs_crepi_F//INVs_crepi_R. Secuencia de los cebadores en el apéndice 3.

3.17. Transcripción *in vitro* de mRNA para microinyectar en embriones

El fragmento clonado en el vector pCS2+ se amplificó por PCR utilizando los cebadores SP6tr y T3 (apéndice 3) y la polimerasa Phusion High Fidelity (New England, BioLabs) en un volumen final de 50 µl. La temperatura de anillamiento fue de 50°C y la reacción se llevó a cabo durante 35 ciclos. El producto resultante se limpió por columna con el kit DNA Clean & concentrator -25 (ZymoResearch) y se eluyó en 20 µl de volumen final. Posteriormente se utilizaron 3 µl del producto de la PCR como molde en la reacción de transcripción con el kit SP6 mMessage mMachine (Ambion) y se incubó entre 4 y 5 horas. Tras añadir 1µl de DNAsa e incubar a 37°C durante 15 minutos, se recuperó el RNA utilizando un kit de purificación y concentración de RNA (RNA Clean & concentrator -5, zymo research) y se eluyó en un volumen final de 14 µl de agua.

3.18. Diseño y generación de CRISPRs

La preparación la construcción molde para el RNA guía monocatenario así como la preparación de la nucleasa Cas9 se realizó como se describe en Perry & Henry, (2015) y también siguiendo el método facilitado personalmente por Dr. Mali (University of California, San Diego).

Para el diseño del RNA guía se buscó una secuencia de 23 pares de bases del tipo 5'-N19-20NGG-3', cerca de sitio diana de interés (tabla 5); en nuestro caso se diseñó lo más cercana posible al extremo 5' UTR para tratar de eliminar por completo el gen diana. Se incorporaron los 19 nucleótidos de la secuencia encontrada en dos oligonucleótidos de la forma indicada a continuación (en negrita se indican las regiones complementarias, expresado 5'-3'): oligonucleótido *forward*: **TTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCG**N19 y oligonucleótido *reverse* **GACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAAC**N19. Los cebadores anillan entre sí y se extienden para generar fragmentos de DNA bicatenario de 100 pares de bases de longitud, utilizando la polimerasa Phusion (mezcla: 4 µl de buffer; 0,2 µl de polimerasa; 0,4 µl de dNTPs a 10 mM; 11,4 µL de agua; 2 µl de cada cebador a concentración 10 µM. Condiciones: 98°C, 30''; 98°C, 5''; 60°C, 30''; 72°C, 20''; 72°C, 5'; mantener a 4°C).

egf-cfc
FORWARD: CriptoCrepi_CRISPr_F
TTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCG**ATGGACGTGTCGTCTGTCA**
REVERSE: CriptoCrepi_CRISPr_R
GACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAAC**TGACAGACGACACGTCCAT**
FORWARD seq: CriptoCrepi_CRISPr_Fsp6
*TATTTAGGTGACACTATAG***ATGGACGTGT**
inversin
FORWARD: INVCrepi_CRISPr_F
TTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCG**CAGCCACGGCCAATCCCGA**
REVERSE: INVCrepi_CRISPr_R
GACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAAC**TCGGGATTGGCCGTGGCTG**
FORWARD seq: INVCrepi _Fsp6
*TATTTAGGTGACACTATAG***CAGCCACGGC**
nodal2
FORWARD: Nod2Crepi_CRISPr_F
TTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCG**CGGCAGCCATACGTGCGCA**
REVERSE: Nod2Crepi_CRISPr_R
GACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAAC**TGCGCACGTATGGCTGCCG**
FORWARD seq: Nod2Crepi_Fsp6
*TATTTAGGTGACACTATAG***CGGCAGCCAT**

Tabla 5. Cebadores diseñados para la generación del RNA guía (RNAg) y su posterior transcripción para la generación de mutantes en *C. fornicata* a través del uso de tecnología CRISPR. Se diseñaron dichos cebadores para los genes *egf-cfc, inversin* y *nodal2*. Cada uno de estos genes requirió del diseño de 3 oligonucleótidos. En negrita se indican los nucleótidos complementarios, requeridos en el diseño y referidos en el texto.

Por otra parte, se digirieron 3 µg de vector (vector 41824, Addgene) utilizando la enzima de restricción AflII (New England, Biolabs) durante 2 horas a 37°C (mezcla: 6 µl buffer Cutsmart 10X, 3 µl de enzima AflII, 3 µg de plásmido y enrasar a 60 µl de agua). Se purificó de gel de agarosa con el kit Kit Zymoclean Gel DNA Recovery (ZymoResearch), y se eluyó en 14 µl de agua. El fragmento de 100 pares de bases se incluyó en el vector linealizado utilizando el sistema de ensamblaje Gibson (New England, Biolabs), utilizando un ratio de inserto 2X (mezcla: 100 ng del inserto, 50 ng de vector linealizado, 10 µl de Gibson Master Mix y enrasar a 20 µl con agua); se incubó durante 15 minutos a 50°C. Posteriormente, se transformaron 2 µl de la ligación en bacterias DH5α (New England, BioLabs) y se usaron placas de agar con kanamicina 50 µg/ml para crecer las colonias a 37°C durante toda la noche.

Las colonias se chequearon utilizando un primer *forward* específico del vector que consiste en una secuencia común 5'TATTTAGGTGACACTATAGN103' que incluye el promotor SP6 seguido de los 10 primeros nucleótidos de la secuencia guía y como reverse, un cebador común para todos (sgRNA_ R1; apéndice 3). Las colonias positivas (producto de 120 pb) se crecieron para enviar una muestra del plásmido a secuenciar. El método fue el mismo ya descrito en 3.9. y 3.10., salvo que el antibiótico utilizado en este caso fue kanamicina a 50 µg/ml. Por último, la generación del mRNA para inyectar se realizó siguiendo el protocolo del apartado 3.17., utilizando la amplificación anterior como molde.

La preparación de la nucleasa Cas9 se realzó a partir del plásmido de Addgene (No.47929). Este plásmido contiene la secuencia codificante para dicha nucleasa, flanqueada por señales de localización nuclear en ambos extremos amino y carboxilo (Jao *et al*., 2013) y se sintetizó el RNA de la misma forma que la descrita en el apartado 3.17. para la obtención de otros RNAs para inyectar en embriones.

3.19. Diseño y generación de morfolinos

Con el fin de silenciar los genes estudiados se utilizaron oligonucleótidos *morpholino antisense*. Estos oligonucleótidos fueron diseñados por el servicio oficial de la compañía que los comercializa (Gene Tools, LLC) para tener la máxima eficiencia de unión a la secuencia diana. Todos los morfolinos se diseñaron para impedir la transcripción y por tanto, se diseñaron en la región del codón ATG inicial. La concentración final de morfolino empleada en cada experimento fue de 1mM. En concreto, se diseñaron morfolinos para los siguientes genes:

Especie	Gen	Morfolino	Secuencia diana
Crepidula	*inversin*	GCTGCCATTGAAGCGCCTGTCAGCT	CGTGTACTGATCGGAAATGAAGCAGACCGTTGATCTAGCCA ACTACGGAGGTTAA[AGCTGACAGGCGCTTCA(ATG)GCAG C]CACGGCCAATCCCGATG
Crepidula	*egf-cfc*	GAAAAGGTTTGCCATGCTCTTACTC	GAACAGTGGATGTGACCAAACATCCTCGCAGATGTGATGT[G AGTAAGAGC(ATG)GCAAACCTTTTC]ATCATGTGGC
-	Control	GAATACGTTTCCCATCCTGTTACTC	-

Tabla 6. Diseño y secuencia de los morfolinos solicitados a la compañía Gene Tools. Se diseñaron morfolinos para los genes *egf-cfc* e *inversin*. En las secuencias se indican bajo el siguiente código de color: **Secuencia del morfolino** y **ATG inicial**

3.20. Microinyección de embriones: *C. fornicata*

Varios laboratorios han desarrollado diferentes metodologías para microinyectar embriones de caracoles con diversos reactivos. En concreto, los embriones de *C. fornicata* poseen un tamaño muy reducido (200 µM) y la metodología tradicionalmente utilizada para microinyectar embriones de esta especie (inyección en placas Petri rayadas con un cristal) si bien era efectiva conllevaba dificultades técnicas y bajos ratios de supervivencia. A lo largo del desarrollo de esta tesis se diseñó un sistema que facilita el procedimiento de microinyección de embriones de *C. fornicata*. Se imprimió a través de la impresión 3D un molde negativo que pudiese dejar impresiones de surcos sobre agarosa, en los que incluir los embriones. El diseño del molde se hizo en el progama Autodesk 123D (http://www.123dapp.com/design), la preparación de los diseños 3D para su impresión se hizo con el software de FormLabs PreForm (http://formlabs.com/support/software/install-preform-software/)

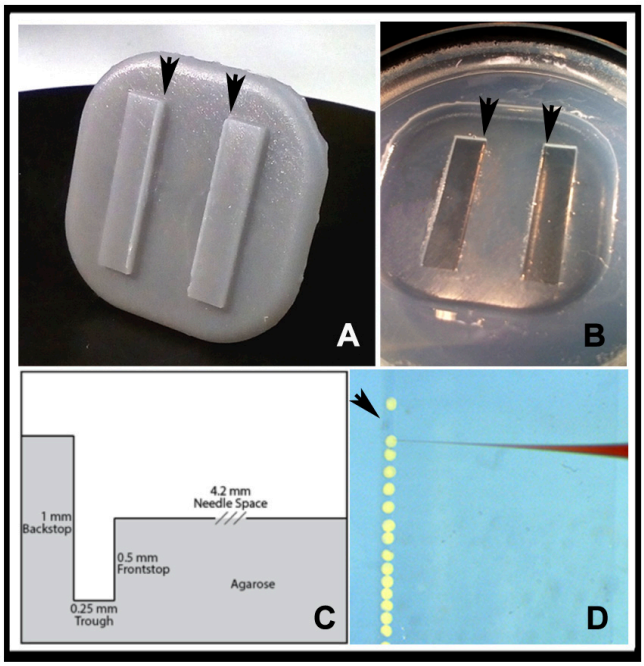


Figura 17. Molde del sistema de microinyección diseñado para *C. fornicata*. A) Molde negativo obtenido por impresión 3D. Las flechas indican los salientes que generarán canales adaptados a los zigotos (200 µM); B) soporte de agarosa en una placa Petri p60. Las flechas indican los surcos en los que se colocan los embriones para ser inyectados; C) diseño del molde negativo con los parámetros y medidas; D) vista a la lupa de zigotos de *C. fornicata* colocados en uno de los canales (flecha). También puede apreciarse la aguja cargada y preparada para microinyectar los embriones.

a resolución 0,025 mm y la impresión se hizo con la impresora 3D de FormLabs Form2 (http://formlabs.com/products/3d-printers/form-2/).

Para la realización de los moldes se disolvió agarosa 2% (D-1 Low EEO, Pronadisa, Laboratorios Conda) en AFSW, y se virtió en una placa petri p60 hasta el borde. El molde impreso se colocó inmediatamente sobre el líquido, dejándolo flotar hasta que se solidificó completamente. Entonces, el molde se retiró dejando dos canales del tamaño exacto para la colocación de los embriones de *C. fornicata*. Tras llenar la placa con AFSW se colocaron los embriones en la agarosa y cuidadosamente se situaron en los carriles para inyectar. Una vez inyectados se crearon corrientes de agua con una pipeta para levantar los embriones de los carriles y recuperarlos, permitiendo así su transferencia a una placa con agua con antibióticos limpia y dejar desarrollar en una cámara húmeda (figura 17).

Para las microinyecciones se fabricaron agujas bajo las siguientes condiciones: *preassure* 500, *heat* 500, *pull* 50, *velocity* 25 y *time* 0. Se utilizó un *puller* de la casa Sutter Instruments (P-97) y capilares de vidrio (World Precision Instruments). El sistema utilizado es el sistema de inyección por presión XenoWorks Digital Microinjector (Sutter Instruments). Los productos para inyectar se diluyeron en agua libre de nucleasas a concentración dependiente del ensayo mezclados con una solución de rojo fenol a 0,5% (Sigma) para visualizar la microinyección (dilución 1:4 del rojo fenol en la solución de RNA). Para visualizar su expresión los mRNAs se coinyectaron con mRNA de *gfp*, procedente de la transcripción *in vitro* del vector pCS2+ vacío. Los mRNA se inyectaron en los zigotos, antes de la primera división y aproximadamente en una cantidad que cubría entre 5-8% del volumen de la célula.

3.21. Microinyecciones de embriones: *D. rerio*

La línea de peces utilizada en esta tesis pertenecen a la línea Oep derivada de la mutagénesis de los 90 realizada en el laboratorio de Nusslein-Volhard, en el alelo tz257 (Zhang *et al*., 1998). La metodología y la línea fueron proporcionados por la Dra. Florencia Cavodeassi (Centro de Biología Molecular Severo Ochoa, UAM-CSIC, Madrid). Esta línea de peces carece del gen *oep*, ortólogo del gen *egf-cfc* estudiado en esta tesis. Los mutantes homozigóticos -/- fueron denominados mutantes Oep, por el fenotipo observado *one-eyed pinhead*.

Los individuos heterozigotos de la línea Oep de *D. rerio* (*oep*+/-) se seleccionaban por la tarde y se ponían en las cubetas separados por una compuerta que era retirada a la mañana siguiente para obtener los embriones. Al cabo de unos 30 minutos, se recolectaban los zigotos

utilizando una malla y se colocaban en una placa con medio E3 (para el stock 60X, añadir 175g de NaCl, 7,6 g de KCl, 29 g de $CaCl_2*2H_2O$, 49 g de $Mg(SO_4)_2*7H_2O$, y enrasar a 10 litros con agua deonizada y añadir 20 ml de azul de metileno 0,01% por cada 10 litros de E3 1X).

Las agujas se hicieron en el mismo tipo de *puller* del apartado 3.20.; los capilares utilizados eran de borosilicato, con filamento interno y 1 mm de grosor (World Precision Instruments, 100F-4). El sistema utilizado es el sistema de inyección por presión XenoWorks Digital Microinjector (Sutter Instruments). Los productos para inyectar se diluyeron en agua libre de nucleasas a concentración dependiente del ensayo. Para visualizar la inyección se mezclaron con una solución de rojo fenol a 0,5% (Sigma), dilución 1:4 del rojo fenol en la solución de RNA. Se inyectaron 200ng de mRNA antes de la primera división directamente en el vitelo. Con una tapa de placa Petri y un portaobjetos, se creó el soporte para colocar los embriones en fila, sin decorionar, e inyectar el mRNA en semi-seco. Posteriormente, los embriones inyectados se crecieron en placas de Petri con medio E3, en un incubador de 28°C.

3.22. Tratamientos con la proteína Activin en embriones de *C. fornicata*

Una serie de embriones de *C. fornicata* en diferentes estadios fueron tratados con proteína Activin (Activin A protein recombinante, 338-AC-010; R&D systems). En concreto se trataron embriones en 2 células (entre 7 y 8 horas post-fertilización), mórulas (32 horas post-fertilización), blástulas (48 horas post-fertilización) y blástulas tardías (61 horas post-fertilización). Los experimentos se realizaron en placas Petri. Por un lado, se trataron embriones con proteína Activin diluida en AFSW a una concentración final de 5 y 10 μM. Por otro lado se pusieron embriones control en AFSW con DMSO al 0,1%. Los embriones se mantuvieron a temperatura y humedad constantes durante 48 horas. Posteriormente, se lavaron 3 veces en AFSW a fin de retirar las soluciones de incubación, y se colocaron en placas limpias para observar su desarrollo.

3.23. Microscopía

El montaje de las preparaciones se llevó a cabo de la siguiente manera: en primer lugar se limpiaron los portaobjetos con el producto Rain-x windscreen protector polish (Rainex) para evitar que los embriones quedaran pegados al cristal. Las preparaciones no eran fijas, para permitir el movimiento de los embriones al hacerlos rodar bajo el cubreobjetos; y con el fin de evitar el aplastamiento de los embriones se cortaron pequeños triángulos de pegatinas que se

colocaron en cada una de las 4 esquinas del cubreobjetos. Los embriones se aplicaron junto con una gota de glicerol en el portaobjetos. En preparaciones fijas se siguió el mismo procedimiento y finalmente se añadió laca de uñas transparente en el borde para sellar.

Para las imágenes tomadas en campo claro, se utilizó el microscopio vertical Axioskop2 plus (Zeiss) acoplado a cámara ccd color tipo Coolsnap FX color (Roper Scientific). Programa metavue 7.10 (Universal Imaging). Objetivos utilizados 5X/0.15 Plan-Neofluar y 10X/0.30 Plan-Neofluar. En campo claro. Filtros de fluorescencia para Hoestch (D315-65 395DCLP D420-70) (49 Zeiss).

Para las imágenes tomadas en fluorescencia y que requerían de mayor definición, se utilizó el microscopio de barrido láser confocal LSM710 acoplado a un microscopio vertical AxioImager M2 (Zeiss). Sistema espectral: Quasar + 2 PMTs. Programa Zeiss ZEN 2010. Objetivos utilizados 5X/0.16 Plan-Neofluar y 10X/0.3 Plan-Apochromat. Láser Diodo 405 nm, Argón 458/488/514 nm y Helio Neón 543nm. Filtros Dapi (G365 FT395 420-470) (Set 49 Zeiss), CFP (426-446 FT455 460-500) (Set 47 Zeiss), FITC (BP450-90 FT 495 BP500-50) (Set 38 Zeiss) y Cy3 (534-559 FT570 570-640) (Set 43 Zeiss).

Los programas utilizados en el procesamiento de las imágenes fueron Helicon Focus versión 6.7.1, ImageJ versión 2.0.0-rc-56/1.51h y Adobe Photoshop CS5 extended versión 12.1 x64.

Las imágenes tomadas de los experimentos de *D. rerio* se realizaron utilizando una lupa M205FA acoplada a una cámara DFC350Fx (Leica Microsystems). Los embriones de 72 horas se anestesiaban en medio con tricaína y se orientaban en un portaobjetos.

Resultados

4.1. Caracterización del gen *inversin* en Metazoa y análisis funcional en Gastropoda

4.1.1. El gen *inversin* tiene un origen antiguo en la historia evolutiva de los animales

Las búsquedas realizadas en las bases de datos genómicas y transcriptómicas de metazoos consultadas se realizaron con la mitad carboxilo-terminal de la proteína Inversin de *Mus*. Esto se debe a que la primera mitad de la proteína contiene dominios ankirina presentes en gran cantidad de otras proteínas diferentes que podrían haber complicado el análisis de los resultados de las búsquedas, mientras que la segunda mitad de la proteína no está presente en ninguna otra conocida. Los resultados de estas búsquedas han permitido identificar ortólogos del gen *inversin* en una especie de coanoflagelado, *Salpingoeca roseta*, y en una gran variedad de metazoos (figura 18; apéndice 7). Concretamente, dentro de metazoos, varias especies de poríferos y cnidarios poseen un ortólogo del gen *inversin*. Sin embargo, las búsquedas realizadas en las bases de datos del ctenóforo *Mnemiopsis leidyi* no ha dado resultados positivos.

Por otra parte, el gen *inversin* está presente en los principales grupos de Bilateria (figura 18). Dentro de Xenacelomorpha tanto Acelos como *Xenoturbella* tienen un ortólogo (figura 18). Aunque hasta el momento las búsquedas realizadas en el NCBI sobre bases de datos genómicas de artrópodos y nematodos no han recuperado ningún ortólogo potencial, se puede confirmar la presencia del gen *inversin* en ecdisozoos gracias a resultados positivos, aunque parciales, en las búsquedas realizadas en los transcriptomas de dos especies de priapúlidos (figura 18). También dentro de Spiralia tanto gastrotricos como briozoos, braquiópodos, moluscos, anélidos y nemertinos poseen ortólogos de este gen (figura 18). Finalmente, dentro de deuteróstomos, ya era conocida la presencia de este gen en vertebrados, aunque las búsquedas realizadas en esta tesis han detectado también la presencia del gen *inversin* en el genoma de un hemicordado (figura 18).

En la mayoría de las bases de datos de especies consultadas se ha encontrado un sólo ortólogo aunque existen dos parálogos en el coanoflagelado *S. rosetta* y el porífero *Oscarella lobularis* (figura 18).

Las proteínas para las que codifican estos genes parecen tener una composición y estructura conservada en todos los grupos (apéndice 7; tabla 7). El tamaño oscila entre 799 aminoácidos (especie *Oopsacas minuta*) y 1459 aminoácidos (especie *Convolutriloba macropiga*). La primera mitad de las proteínas consta de 16 repeticiones ankirina en todos los

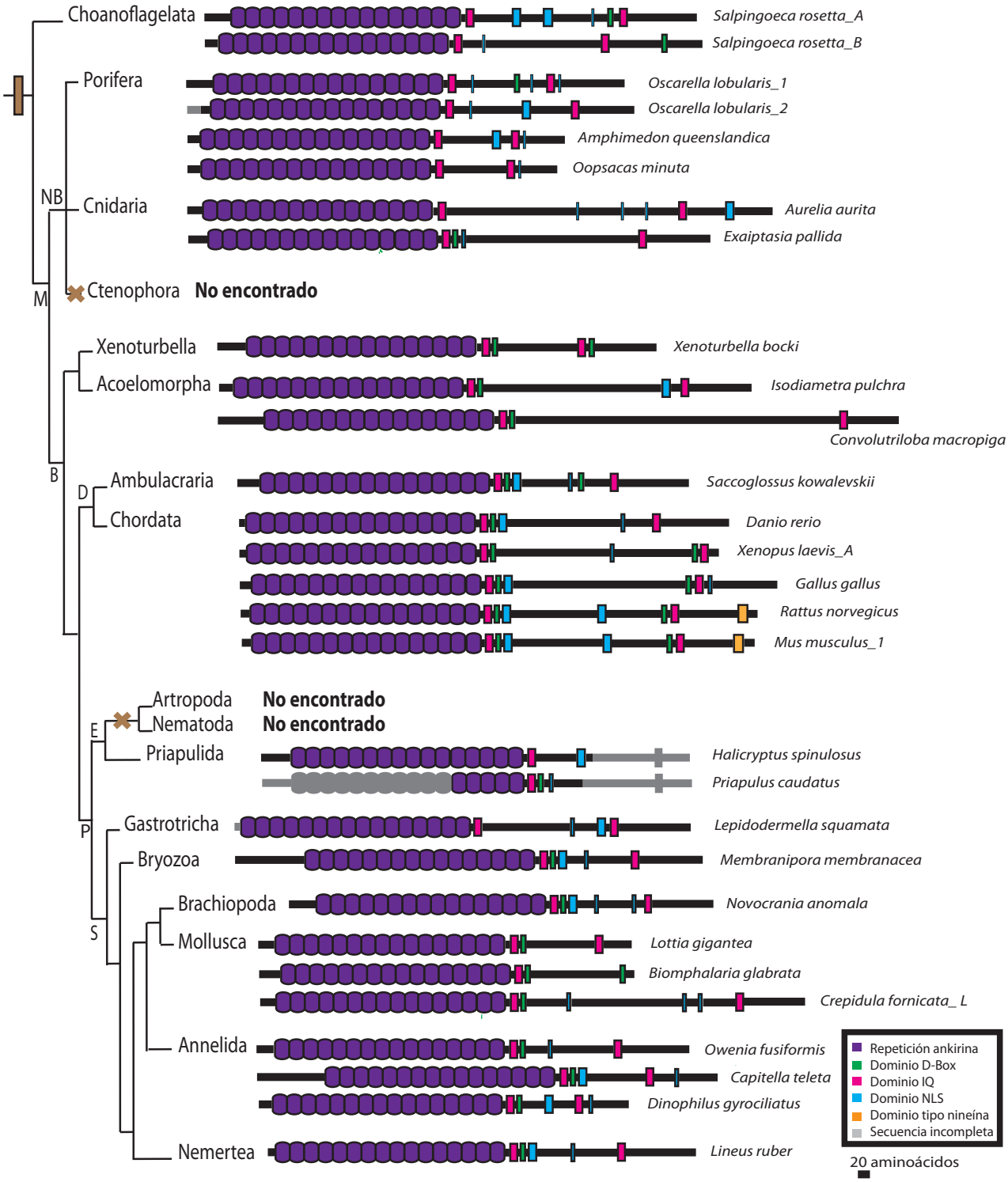


Figura 18. Cladograma de Metazoa mostrando la estructura de la proteína Inversin para cada uno de los principales grupos. En cada grupo se muestra un esquema de la proteína Inversin de cada especie identificada en esta tesis con todos los dominios señalados según se detalla en la leyenda. Todos los esquemas están a escala. El rectángulo marrón en el árbol indica el potencial origen de esta proteína. Las cruces marrones en el árbol indican potenciales deleciones de esta proteína. M, Metazoa; NB, No bilaterados; B, Bilateria; D, Deuterostomia; P, Protostomia; E, Ecdysozoa; S, Spiralia.

ortólogos encontrados. En la segunda mitad, la localización y disposición de los dominios es variable entre ortólogos de diferentes grupos, aunque comparten la presencia de dos dominios IQ de unión a Calmodulina (excepto en *Biomphalaria*, ver más adelante) y un número variable de dominios de destrucción D-Box y de localización nuclear NLS (tanto de tipo canónico monopartido como canónico bipartido). En todos los ortólogos se observa el siguiente patrón en el orden de los dominios: repeticiones de dominios ankirina, seguidos de un dominio IQ (IQ1). Solamente en Bilateria (y en una especie de cnidario) se encuentra el siguiente patrón conservado: repeticiones de ankirina, un dominio IQ (IQ1) y un dominio de destrucción D-Box. Finalmente, el dominio tipo nineína parece ser exclusivo de mamíferos.

Dominio	Consenso	Referencia
Repeticiones ankirina	$XGXTPLHAAX_7VX_2LLXGAX_5(D/N)X$	Morgan *et al.*, 2002
Dominio IQ	$(I/V/L)QX_3RX_4(R/K)X_2(F/I/L/V/W/Y)$	Morgan *et al.*, 2002
Dominio D-Box	$RX_2LX_2(I/V)XN$	Barford, 2011
Dominios NLS:		Barford, 2011
NLS monopartido	$K(K/R)X(K/R)$	Dingwall *et al.*, 1999
NLS bipartido	$(K/R)(K/R)X_{(10\text{-}12)}(K/R)_{(3to5)/5}$	Robbins *et al.*, 1999
Dominio tipo nineína	$S(L/V)X_3L(E/Q)(T/S)IHXL(D/E)X_2G(L/R)(K/S)KK$	Shiva *et al.*, 2008

Tabla 7. Secuencia consenso para los dominios encontrados en la proteína Inversin. X, cualquier aminoácido; subíndice, número de aminoácidos; (*/*), opción entre dos o más aminoácidos en esa posición. Para cada consenso se especifican las referencias correspondientes.

4.1.2. El gen *inversin* en *C. fornicata* y en *B. glabrata*

Las búsquedas informáticas y el posterior trabajo en laboratorio llevados a cabo en esta tesis permitieron la identificación de dos isoformas en el gasterópodo dextrógiro *C. fornicata*: una isoforma larga (*inversin-L*; del inglés *long*) de 1125 aminoácidos y que contiene todo el transcrito y otra más pequeña (*inversin-S*; del inglés *short*) de 918 aminoácidos (figura 19; apéndice 8). Las dos isoformas generadas por procesamiento alternativo se diferencian por la deleción en la forma corta de 207 aminoácidos en la segunda mitad de la proteína, tras el dominio de destrucción D-Box. Se han buscado posibles dominios específicos en esta zona de 207 aminoácidos y sólo se ha encontrado un potencial dominio de localización nuclear NLS (figura 19).

Por otra parte, tanto las búsquedas informáticas como el posterior trabajo en laboratorio han llevado a la identificación de una sola forma de 811 aminoácidos en la especie de gasterópodo

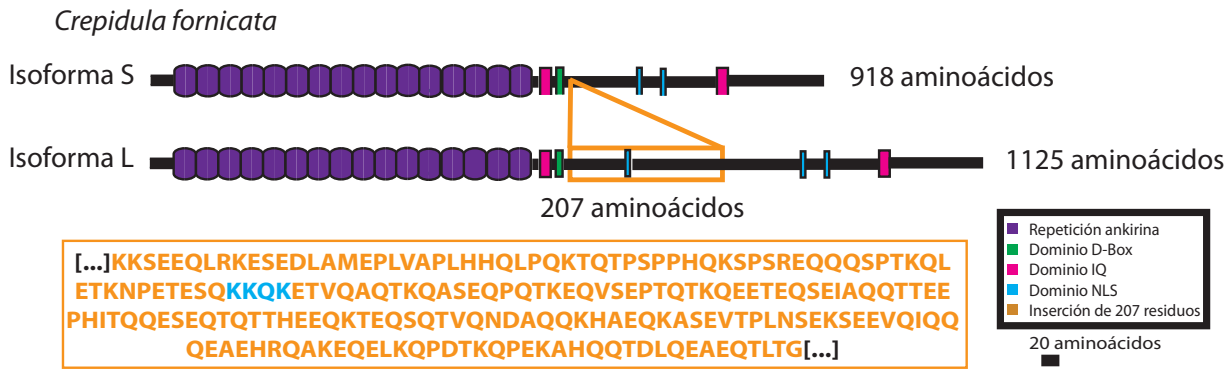


Figura 19. La proteína Inversin en *C. fornicata*. Para la especie dextrógira *C. fornicata* se han identificado dos isoformas: una isoforma larga (*inversin-L*; del inglés *long*) de 1125 aminoácidos y que contiene todo el transcrito y otra más pequeña (*inversin-S*; del inglés *short*) de 918 aminoácidos. Los esquemas muestran todos los dominios identificados en cada isoforma según se detalla en la leyenda. Los esquemas están a escala. El recuadro naranja indica la región de 207 aminoácidos que sufre procesamiento alternativo y que está ausente en la isoforma corta. Dentro del recuadro naranja se marca en azul un potencial dominio NLS.

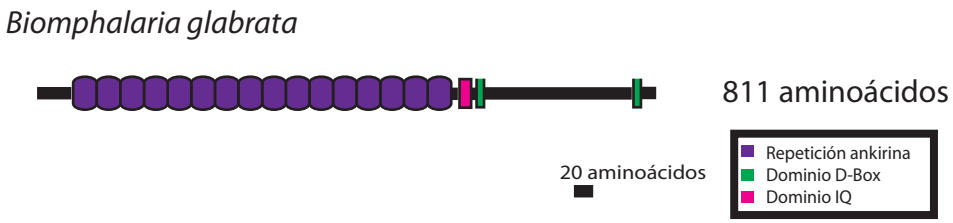


Figura 20. La proteína Inversin en *B. glabrata*. Para la especie levógira *B. glabrata* se ha identificado una isoforma de 811 aminoácidos. El esquema muestra todos los dominios identificados según se detalla en la leyenda (nótese la presencia de un solo dominio IQ). El esquema está a escala.

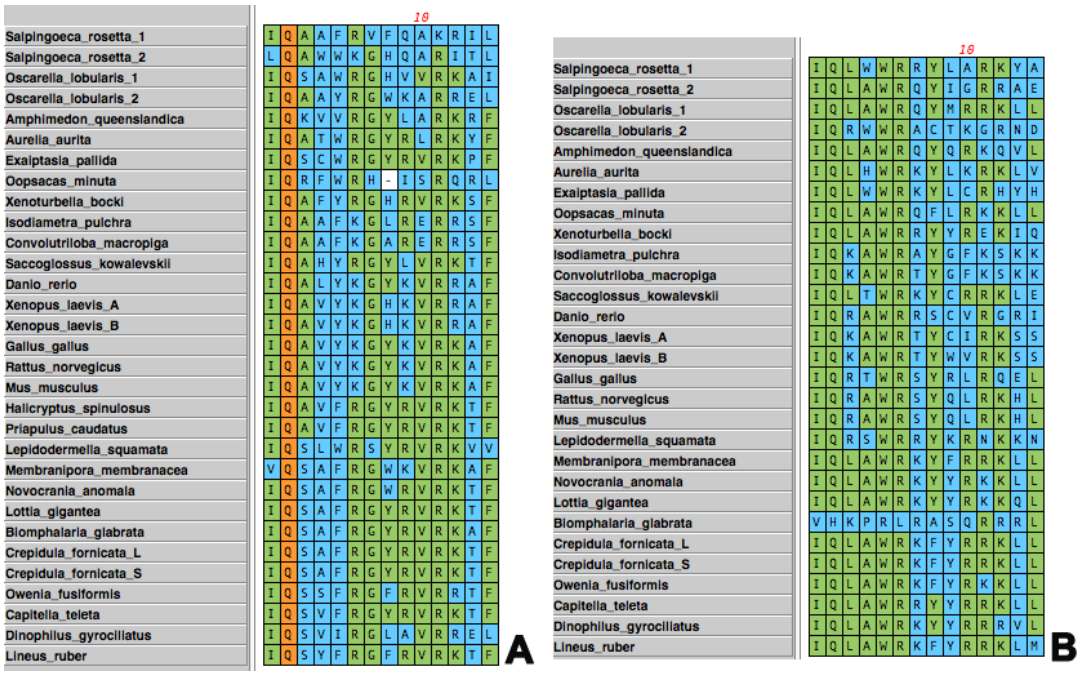


Figura 21. Dominios IQ en la proteína Inversin. A) Fragmento del alineamiento de las proteínas Inversin identificadas que incluye el dominio IQ1. Nótese que este dominio está en todas las proteínas ortólogas identificadas; B) fragmento del alineamiento de las proteínas Inversin identificadas que incluye el dominio IQ2. Nótese que este dominio está en todas las proteínas ortólogas identificadas excepto en *B. glabrata*. Los colores indican que todas las posiciones son idénticas en todas las especies (naranja); que las posiciones son idénticas en el 50% o más de las especies (verde) y posiciones variables (azul). Ver secuencia consenso en la tabla 7.

levógira *B. glabrata* (figura 20; apéndice 8). Como se ha indicado anteriormente, todas las proteínas ortólogas identificadas en este estudio cuentan con la presencia de dos dominios de unión a Calmodulina IQ (IQ1 y IQ2). Sin embargo, en la proteína ortóloga de la especie de gasterópodo levógiro *B. glabrata* solamente se ha podido identificar el dominio IQ1 (figura 21).

4.1.3. Activación espacio-temporal del gen *inversin* durante el desarrollo de *C. fornicata*

Con el fin de conocer si el gen *inversin* está activo durante el desarrollo embrionario de la especie *C. fornicata* se realizaron reacciones de hibridación *in situ*. Para distinguir el patrón de activación específico de cada una de las dos isoformas se diseñaron sondas en la zona del procesamiento alternativo (figura 22). Así, la sonda que hibrida con la isoforma larga contiene 1,4 Kb, siendo los 700 pb centrales los que codifican para el fragmento ausente en la isoforma corta. Por su parte, la sonda de la forma corta tiene un tamaño de 700pb que hibridan con la parte central del gen, sin el fragmento eliminado en el procesamiento. A continuación se describe el patrón de activación observado para cada una de las dos isoformas.

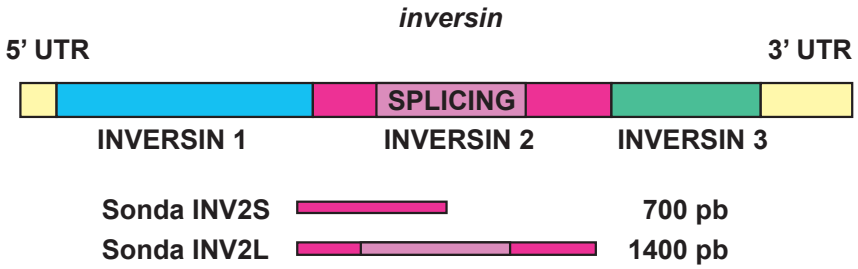


Figura 22. Estrategia de clonaje del gen *inversin* en la especie *C. fornicata* y diseño de las sondas para el estudio de la distribución espacio-temporal del mRNA de ambas isoformas. El cDNA de *inversin* en *C. fornicata* se dividió en 3 regiones para su clonaje (Inversin 1, 2 y 3). Las zonas en color amarillo representan las zonas reguladoras 5' y 3' UTR. La región azul, denominada Inversin 1, de 1,2Kb, contiene la gran mayoría de los dominios ankirina de la proteína. La región rosa, Inversin 2, contiene la zona de *splicing*; el producto amplificado contiene 1,4Kb en la isoforma larga y 700pb en la isoforma corta. La región verde, Inversin 3, es común a las dos isoformas y contiene los 700 pb más terminales de la proteína. Las sondas utilizadas, INV2L y INV2S, se diferencian en los 700 pb centrales que sufren procesamiento alternativo.

4.1.3.1. El transcrito de *inversin-L* se detecta durante todo el desarrollo de *C. fornicata* y se expresa de manera asimétrica desde la gastrulación temprana hasta larva.

El transcrito de la isoforma larga (*inversin-L*) se detecta desde el comienzo del desarrollo; en concreto se ha detectado en huevos fecundados, siendo por tanto un mRNA de origen materno. Durante la segmentación, el transcrito se localiza cerca de los husos mitóticos y su segregación es igual entre todos los blastómeros (figura 23). Sin embargo, en embriones de 37 células, el mRNA sólo se localiza en los macrómeros (A, B, C y D) (figura 23). A partir del estadio de gástrula temprana se observa una disminución de expresión en el macrómero C; esta expresión diferencial se mantiene durante la gastrulación, observándose en gástrula media también (figura 23). En el estadio de organogénesis, la expresión se reduce a células discretas de la zona antero-ventral y disminuye la expresión en el labio dorsal (estomodeo); además, la señal aparece en una célula concreta en el lado derecho (figura 23). Durante el estadio de larva preveliger, la expresión discreta de *inversin-L* en la zona antero-ventral se mantiene y desaparece finalmente del estomodeo (figura 23). En la larva veliger la expresión sigue manteniéndose fuertemente en células discretas de la cabeza y aparece de forma ubicua en la masa visceral (figura 23).

4.1.3.2. El transcrito de *inversin-S* se detecta en *C. fornicata* desde la segmentación hasta el estadio de larva y se expresa de manera asimétrica durante todo el desarrollo.

El transcrito de la isoforma corta (*inversin-S*) se detecta desde estadios de segmentación tardía (37 células) cuando la expresión comienza a ser de tipo zigótica (figura 24). En concreto, el transcrito se detecta mayoritariamente en células del cuadrante A y B, aunque también en células del cuadrante C (células 3a, $1a^2$, $2b^2$,$2b^1$, 3b, $2c^2$ y 4B). Según avanzan las divisiones, la expresión va cambiando y desaparece en los micrómeros del cuadrante A y se activa mayoritariamente en los del cuadrante C (figura 24). Durante la gastrulación la señal se concentra exclusivamente en los macrómeros B y C (figura 24). En el estadio de organogénesis y larva preveliger, la expresión del transcrito *inversin-S* se localiza en el blastoporo y asimétricamente en el ectodermo del lado derecho del velo (figura 24). En el estadio de larva veliger la expresión parece extenderse a todo el velo y concentrarse en determinadas áreas de la masa visceral; aunque poco definidas la señalización es mayoritaria en el lado derecho (figura 24).

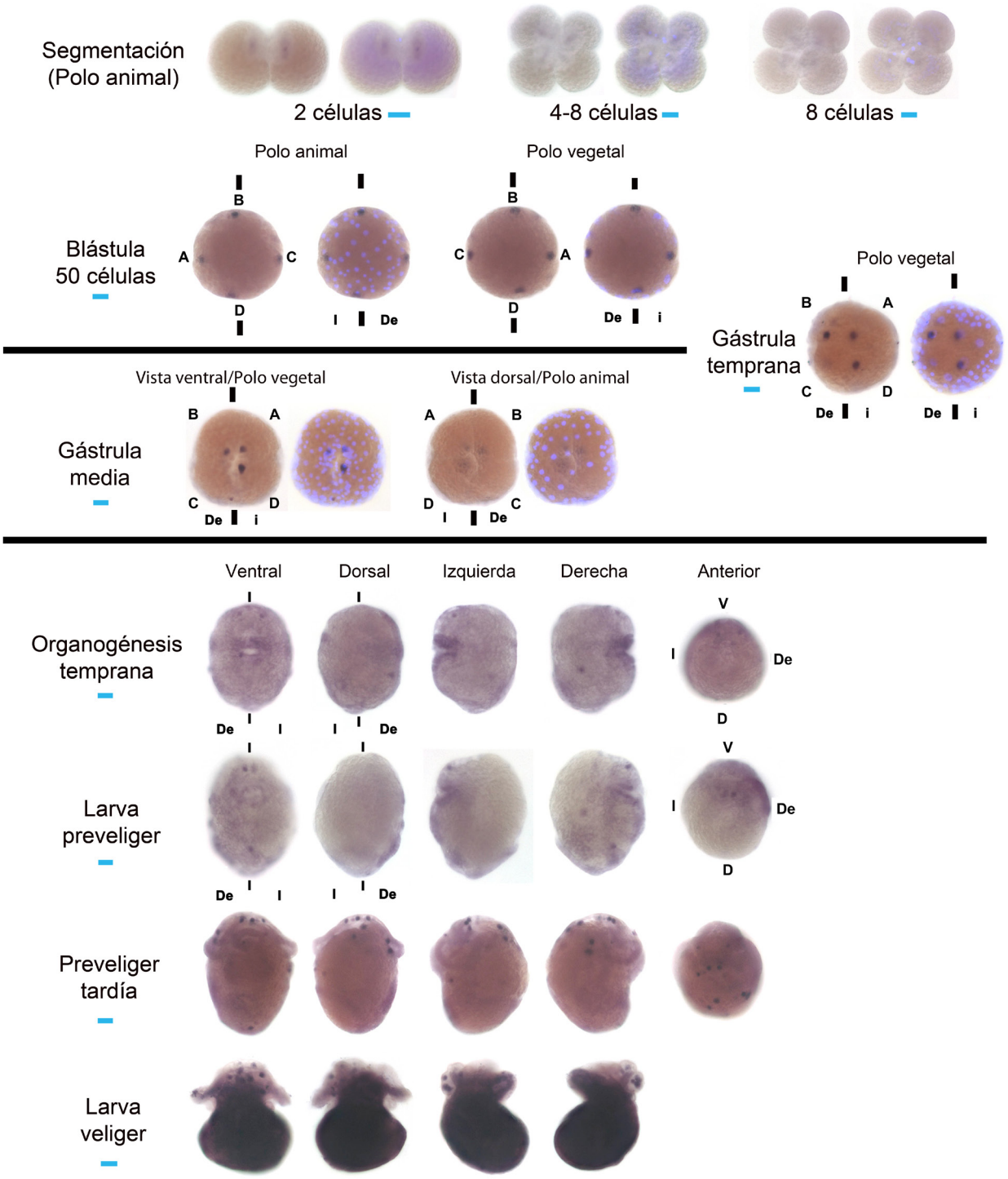


Figura 23. Localización espacio-temporal por hibridación *in situ* del mRNA de *inversin-L* a lo largo del desarrollo de la especie *C. fornicata*. El RNAm de esta forma se encuentra desde los estadios más tempranos del desarrollo (segmentación) hasta larva veliger. En el estadio de blástula y gástrula se indican los cuadrantes A, B, C y D. Cada uno de los estadios del panel presenta varias vistas del mismo embrión indicadas en la figura. A la derecha de cada embrión de los estadios de segmentación, blástula y gástrula, se han realizado superposiciones de las imágenes tomadas en campo claro con imágenes tomadas en fluorescencia para visualizar núcleos teñidos con Hoestch. De: lado derecho; I: lado izquierdo; D: dorsal; V: ventral; las barras representan el eje bilateral. La escala se indica con la línea de color azul, que corresponde a 30 µm.

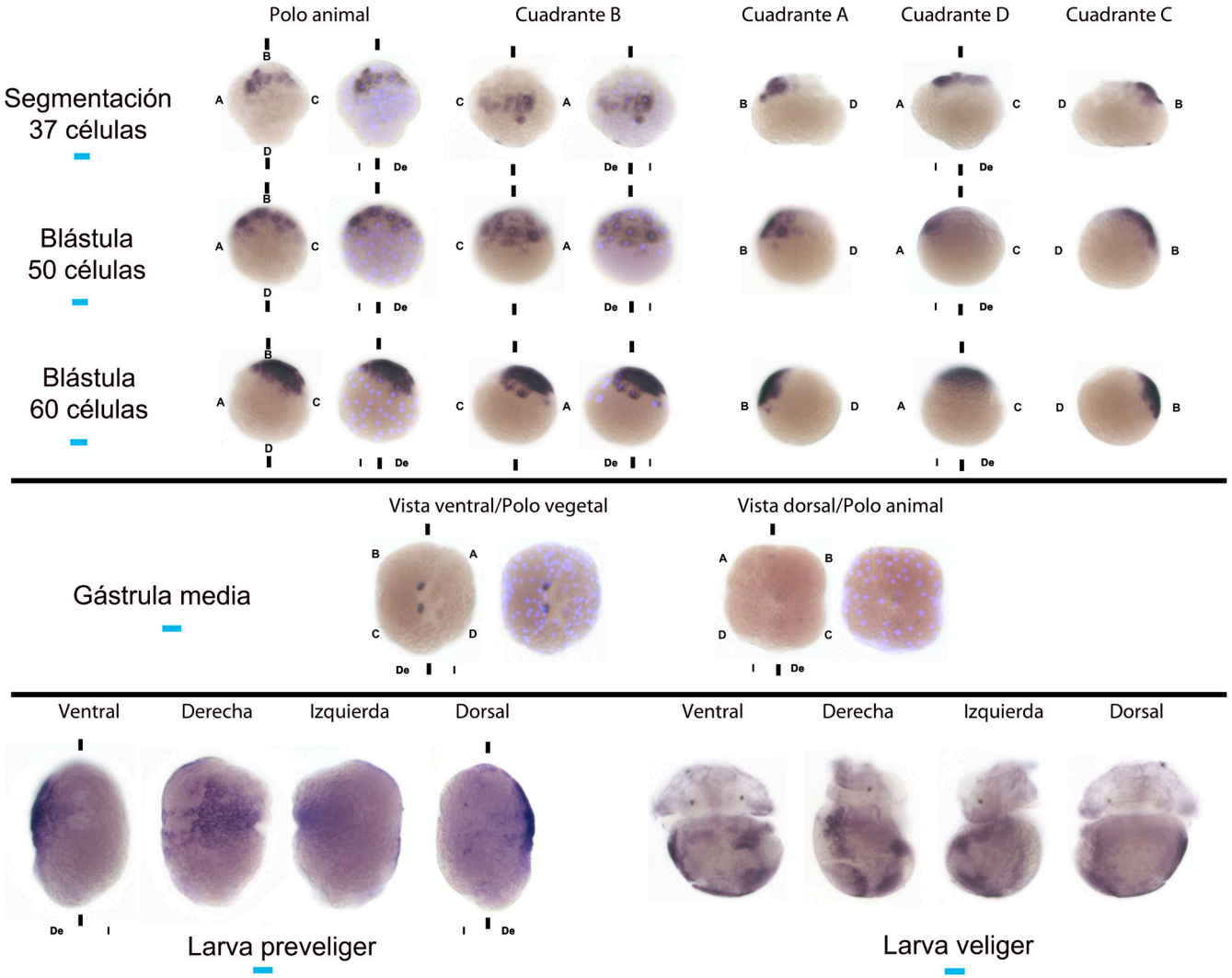


Figura 24. Localización espacio-temporal por hibridación *in situ* del mRNA de *inversin-S* a lo largo del desarrollo de la especie *C. fornicata*. El RNAm de esta forma se encuentra desde estadios tempranos del desarrollo (37 células) hasta larva veliger. En el estadio de segmentación, blástula y gástrula se indican los cuadrantes A, B, C y D. Cada uno de los estadios del panel presenta varias vistas del mismo embrión indicadas en la figura. A la derecha de cada embrión de los estadios de segmentación, blástula y gástrula, se han realizado superposiciones de las imágenes tomadas en campo claro con imágenes tomadas en fluorescencia para visualizar núcleos teñidos con Hoestch. De: lado derecho; I: lado izquierdo; las barras representan el eje bilateral. La escala se indica con la línea de color azul, que corresponde a 30 µm.

4.1.4. Detección de la proteína Inversin en Gastropoda

Los experimentos de inmunohistoquímica realizados utilizando el anticuerpo comercial para Inversin (ver apartado 3.12.) en embriones de *C. fornicata* no produjeron ningún resultado en ninguna de las concentraciones probadas. Por ello, se decidió diseñar un anticuerpo específico para las proteínas Inversin de las especies *C. fornicata* y *B. glabrata* (apéndice 4; tabla 2). Se obtuvieron sueros provenientes de dos conejos (A y B) inoculados con los péptidos diseñados específicamente para cada una de las especies de gasterópodos que fueron probados en experimentos de *western blot* e inmunohistoquímica.

Los resultados de experimentos de *western blot* muestran que el anticuerpo purificado contra el péptido 1 a partir del suero número 3, proveniente de la inoculación de los péptidos de la proteína Inversin de *C. fornicata,* podría estar potencialmente reconociendo dicha proteína, ya que pueden distinguirse bandas intensas del tamaño esperado (99 y 121 KDa) en el experimento (figura 25).

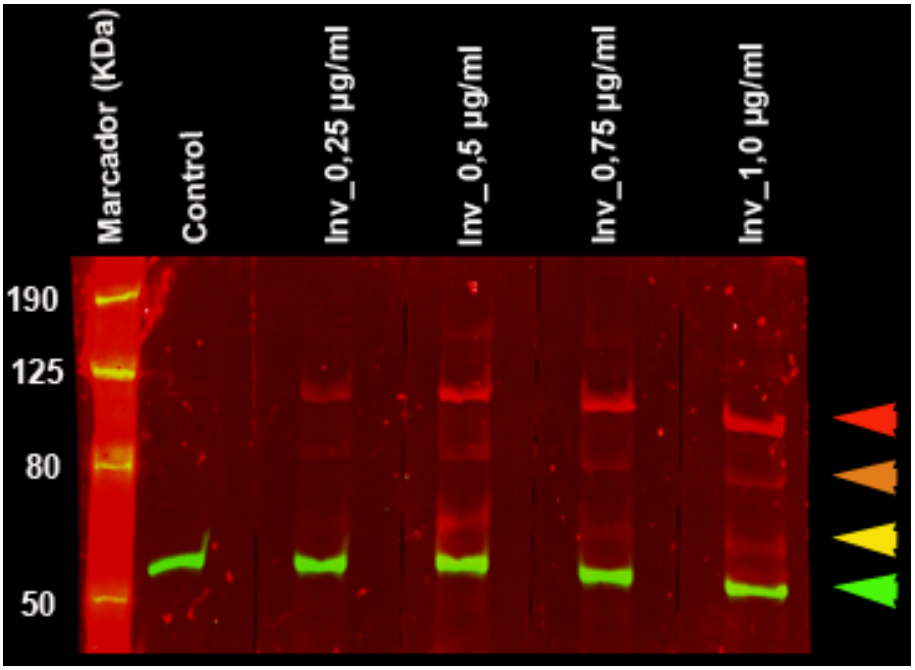


Figura 25. Detección de la proteína Inversin en un extracto de proteína total de larvas preveliger de *C. fornicata* por *western blot*. De izquierda a derecha en cada carril se indica el marcador de peso molecular (con los números a la izquierda que indican la referencia del marcador en KDa), control (anticuerpo anti-βTubulina en un extracto total de larvas preveliger, y ambos anticuerpos secundarios) y experimento (incubación con la purificación del suero 3, del anticuerpo contra el péptido 1 producido en el conejo B, a una concentración de suero 0,25 µg/ml, 0,5 µg/ml, 0,75 µg/ml y 1 µg/ml). La proteína Inversin se reveló en rojo. Los pesos estimados para las isoformas L y S son 122 y 99 KDa, respectivamente señaladas con las flechas roja y naranja. La flecha amarilla indica una banda de tamaño menor que también parece estar siendo reconocida por el suero con el anticuerpo. El control de carga se realizó con β-Tubulina, revelado en verde y señalado con una flecha verde.

Las reacciones de inmunohistoquímica con embriones completos mostraron que mientras en los controles no se apreciaba ningún tipo de señal, los embriones incubados con el suero purificado mostraban una tinción en la membrana celular y perinuclear (foto no disponible). Sin embargo, se precisa un mayor esfuerzo para la optimización de estas reacciones ya que se ha observado bastante ruido de fondo.

Por otra parte, los resultados de experimentos de *western blot* muestran que el suero proveniente de la inoculación de los péptidos de la proteína Inversin de *B. glabrata* podría estar potencialmente reconociendo dicha proteína, ya que pueden distinguirse dos bandas intensas del tamaño esperado (90 KDa) en el experimento correspondiente al uso del suero número 3, proveniente del conejo B (figura 26). Estas bandas no se detectan cuando se incuba con el suero control de pre-inoculación; esto es, proveniente del mismo conejo, pero extraído previamente a la inoculación del antígeno (figura 26).

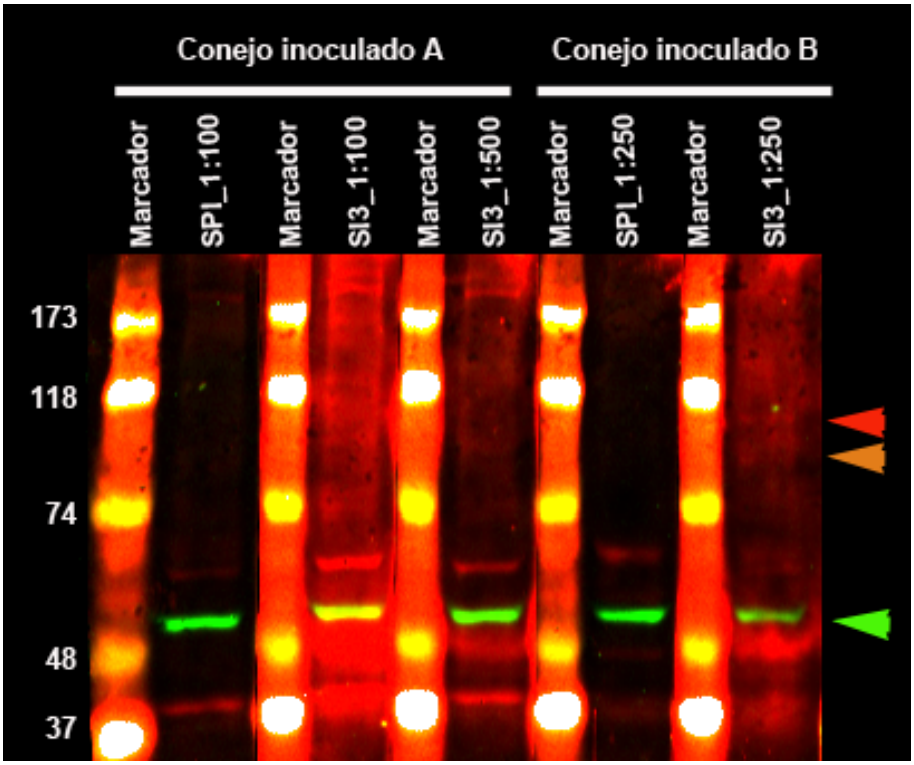


Figura 26. Detección de la proteína Inversin en un extracto de proteína total de larvas de *B. glabrata* por *western blot*. De izquierda a derecha en cada carril se indica el Marcador de peso molecular (con los números a la izquierda que indican la referencia del marcador en KDa), control SPI_1:100 (incubación con el suero del conejo A extraído antes de la inoculación de los péptidos a una concentración 1:100), marcador, experimento_ SI3_1:100 (incubación con el suero 3 del conejo A extraído después de la incubación con los 3 péptidos a una concentración 1:100), marcador, experimento_SI3_1:500 (incubación con el suero 3 del conejo A extraído después de la incubación con los 3 péptidos a una concentración 1:500), marcador, control SPI_1:250 (incubación con el suero del conejo B extraído antes de la inoculación de los péptidos a una concentración 1:250), marcador, experimento_SI3_1:250 (incubación con el suero 3 del conejo B extraído después de la incubación con los 3 péptidos a una concentración 1:250). El peso estimado es 90 KDa, señalado con las flechas roja y naranja, puesto que estimamos que su tamaño podría ser cualquiera de los dos. El control de carga se realizó con β-Tubulina, revelado en verde y señalado con una flecha verde.

4.1.5. Experimentos de sobre-expresión y falta de función del gen *inversin* en zigotos de *C. fornicata*

Para investigar la función del gen *inversin* en el gasterópodo *C. fornicata* se siguieron tres estrategias principales: 1) el uso de un morfolino 2) el uso de CRISPR y 3) el uso de formas silvestres y defectivas (dominante negativo) que fueron inyectadas en embriones de *C. fornicata* (figura 27).

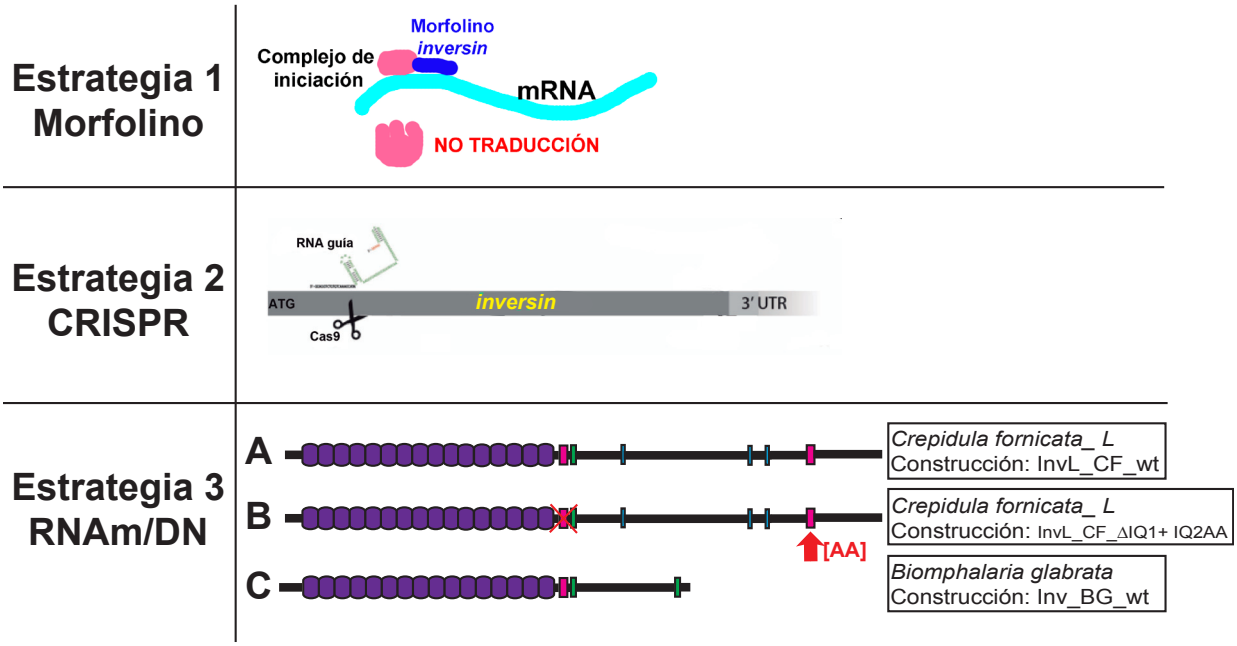


Figura 27. Estrategias seguidas para el abordaje funcional del gen *inversin* en la especie *C. fornicata*. Estrategia 1: uso de morfolinos para evitar la traducción de ambas isoformas de *C. fornicata* (*inversin-L* e *inversin-S*). Estrategia 2: uso de la tecnología CRISPR/Cas9 para silenciar ambas isoformas de *C. fornicata* (*inversin-L* e *inversin-S*). Estrategia 3: uso de mRNA proveniente de construcciones con A) la forma silvestre de la isoforma larga de *C. fornicata* (*inversin-L*); B) la forma de la isoforma larga de *C. fornicata* con la deleción del dominio IQ1 completo y con los aminoácidos IQ del dominio IQ2 sustituidos por AA; C) la forma silvestre del gen *inversin* de *B. glabrata*.

Tanto la microinyección de CRISPRs (apartado 3.18.) para generar individuos carentes del gen *inversin* en *C. fornicata* como la microinyección de morfolinos (apartado 3.19.) para intentar bloquear la traducción del mensajero del gen *inversin* de *C. fornicata*, no mostraron ningún fenotipo anormal ni malformación visible a lo largo del desarrollo (tablas 8 y 9).

En cuanto a la estrategia con construcciones (apartados 3.15. y 3.16.), en primer lugar se preparó una construcción con el mRNA del gen *inversin* silvestre de *C. fornicata* y se inyectó en zigotos de esta especie (experimento A en la figura 27 y en la tabla 10). Estos experimentos pretendían comprobar si la sobre-expresión del gen *inversin* tenía algún efecto en el desarrollo

	Total	Fenotipo WT	Fenotipo simetría
Control	49	100%	0%
Morfolino	70	100%	0%

Tabla 8. Resultados obtenidos tras la inyección del morfolino contra el RNAm de *inversin* en la especie *C. fornicata*. De los 70 zigotos inyectados el 100% presentaba un fenotipo igual al silvestre.

	Total	Fenotipo WT	Fenotipo simetría
Control	279	100%	0%
CRISPR+Mo	95	100%	0%

Tabla 9. Resultados obtenidos tras la inyección del morfolino contra el RNAm de *inversin* junto con el RNAm de *Cas9* y el RNA guía específico de *inversin* en la especie *C. fornicata*. De los 95 zigotos inyectados el 100% presentaba un fenotipo igual al silvestre.

de este gasterópodo. Los resultados no mostraron ningún fenotipo diferente al control en ninguna de las concentraciones utilizadas (figura 28; tabla 10).

A continuación, se preparó una construcción con el mRNA del gen *inversin* de *C. fornicata* sin el dominio IQ1 y con el IQ2 modificado (los aminoácidos IQ fueron sustituidos por dos aminoácidos alanina) y se inyectó en zigotos de esta especie (experimento B en la figura 27 y en la tabla 10). Estos experimentos pretendían comprobar si la falta de los dominios IQ del transcrito del gen *inversin* de *C. fornicata* repercute en el desarrollo de esta especie. En este caso, la mayoría de los embriones resultantes tenían un fenotipo silvestre. Sin embargo, un porcentaje reducido (1,6 %) mostró un fenotipo claramente simétrico tanto en la morfología de la concha como en la anatomía interna (tabla 10).

	RNAm inyectado	Total	Fenotipo WT	Fenotipo simetría
Experimento A (sobre-expresión)	Control	>100	100%	0%
	InvL_CF_wt	98	100%	0%
Experimento B (falta de IQs)	Control	>100	100%	0%
	InvL_ΔIQ1+IQ2AA	123	98,4% (121)	1,6% (2)
Experimento C (especie levógira)	Control	>100	100%	0%
	Inv_BG_WT	133	95,5% (127)	4,5% (6)

Tabla 10. Resultados obtenidos tras la inyección de mRNA proveniente de diferentes construcciones en la especie *C. fornicata*. Experimento A: inyección de la forma silvestre de la isoforma larga de *C. fornicata* (*inversin-L*); de los 98 zigotos inyectados, el 100% presentaba un fenotipo igual al silvestre. Experimento B: inyección de la isoforma larga de *C. fornicata* con la deleción del dominio IQ1 completo y con los aminoácidos IQ del dominio IQ2 sustituidos por AA; de los 123 zigotos inyectados, sólo el 1,6% presentaba un fenotipo de simetría bilateral. Experimento C: inyección de la forma silvestre del gen *inversin* de *B. glabrata*; de los 133 zigotos inyectados, el 4,5% presentaba un fenotipo de simetría bilateral. Los experimentos se corresponden con los mostrados en la figura 27.

Finalmente, se preparó una construcción con el mRNA del gen *inversin* silvestre de *B. glabrata* y se inyectó en zigotos de la especie *C. fornicata* (experimento C en la figura 27 y en la tabla 10). Estos experimentos pretendían comprobar si la presencia del transcrito del gen *inversin* silvestre de *B. glabrata* que carece del dominio IQ2 tiene consecuencias en el desarrollo de la especie *C. fornicata*. Aunque la mayoría de los embriones resultantes tenían un fenotipo silvestre, un porcentaje reducido (4,5 %) mostró un fenotipo claramente simétrico tanto en la morfología de la concha como en la anatomía interna (figura 29 y tabla 10).

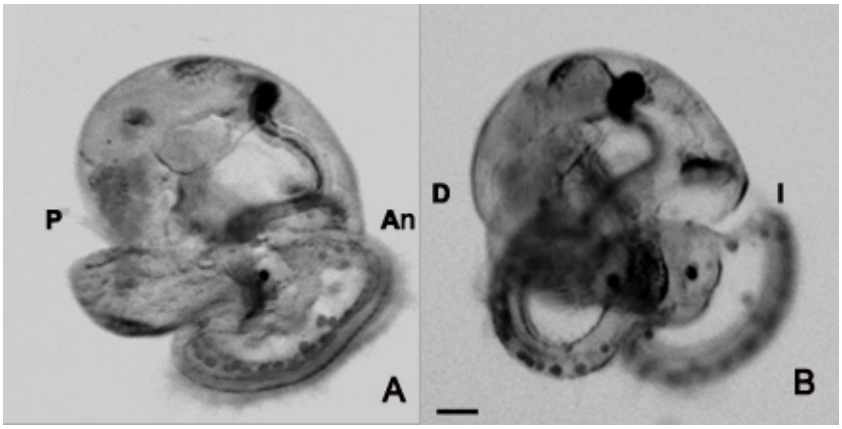


Figura 28. Morfología de la larva veliger de *C. fornicata.* A) larva veliger control; B) larva veliger resultante de un zigoto inyectado con la forma silvestre de la isoforma larga de *C. fornicata* (*inversin-L*); A, vista lateral; B, vista frontal; An, anterior; P, posterior; I, izquierda; D, derecha. La escala corresponde a 30 µm.

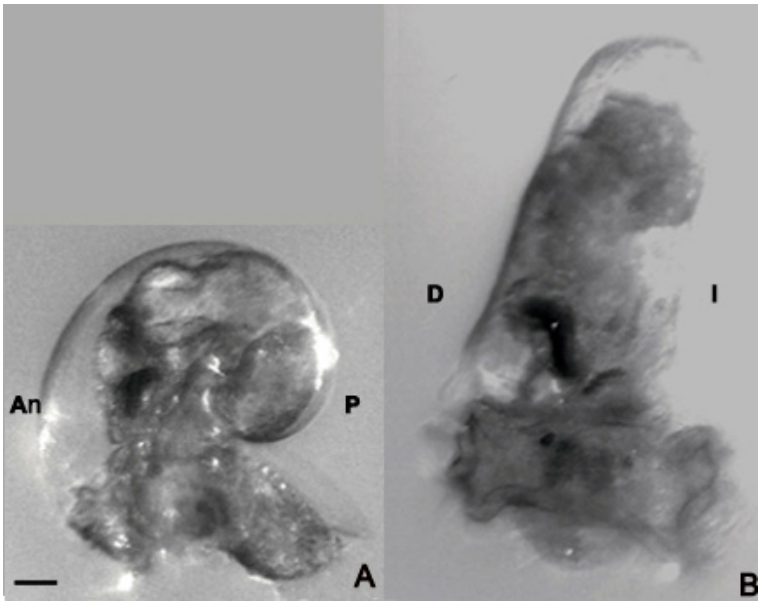


Figura 29. Morfología de la larva veliger de *C. fornicata*. A) larva veliger control; B) larva veliger resultante de un zigoto inyectado con la forma silvestre del gen *inversin* de *B. glabrata*; A, vista lateral; B, vista frontal; An, anterior; P, posterior; I, izquierda; D, derecha. Nótese en B el fenotipo de simetría bilateral. La escala corresponde a 30 µm.

4.2. Caracterización y análisis funcional del gen *nodal* en *C. fornicata*

4.2.1. *C. fornicata* tiene dos parálogos del gen *nodal*

Las búsquedas en los transcriptomas de *C. fornicata* y posteriores análisis de ortología permitieron encontrar dos genes parálogos para el gen *nodal*, aquí denominados *nodal1* y *nodal2* (figura 30; apéndices 8 y 13).

El gen *nodal1* codifica para una proteína de 428 aminoácidos y el gen *nodal2* codifica para una proteína de 472 aminoácidos. Ambas formas presentan las 7 cisteínas características del extremo C-terminal y una región de corte que coincide con ortólogos de otros grupos animales (figuras 30 y 31; apéndice 13).

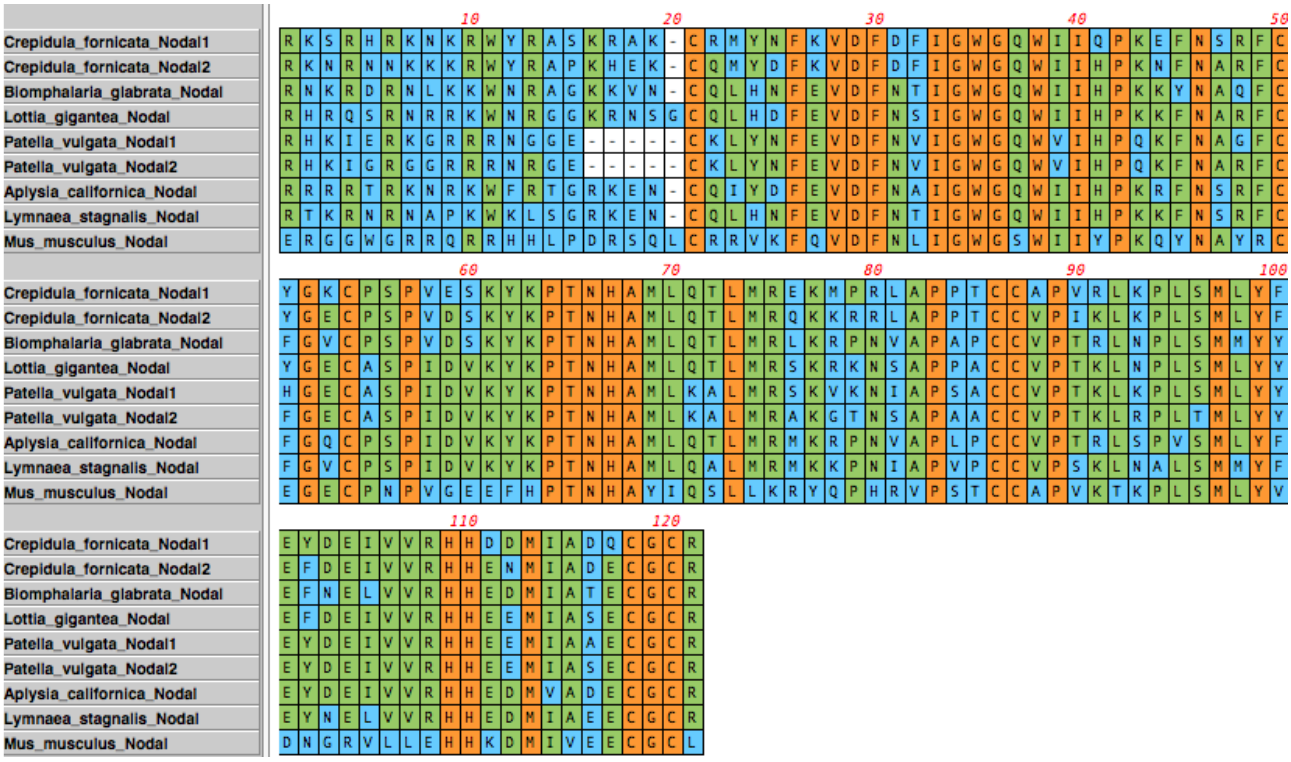


Figura 30. Alineamiento de varias proteínas ortólogas de Nodal de varias especies de gasterópodos (*C. fornicata; B. glabrata; L. gigantea; P. vulgata; Aplysia californica; Lymnaea stagnalis*) y del vertebrado *Mus musculus*. La posición 1-20 corresponde con la zona donde se produce el procesamiento de la pre-proteína. La posición 21 en adelante corresponde a la proteína madura. Los colores indican que todas las posiciones son idénticas en todas las especies (naranja); que las posiciones son idénticas en el 50% o más de las especies (verde) y posiciones variables (azul).

Figura 31. Estructura genérica de la proteína Nodal. En azul se indica la región del péptido señal. En verde se indica la región pro-proteína, que contiene la región de corte para la generación de la forma madura (RXXR), siendo X cualquier aminoácido. En naranja se indica la región que conforma la proteína madura, con las 7 cisteínas que caracterizan a la familia.

4.2.2. Activación espacio-temporal del gen *nodal* en *C. fornicata*

Con el fin de conocer si los genes *nodal1* y *nodal2* están activos durante el desarrollo embrionario de la especie *C. fornicata* se realizaron reacciones de hibridación *in situ*. Las diferencias a nivel de secuencia entre ambos genes son muy significativas por lo que cada sonda diseñada en la zona codificante reconoce de manera específica a cada uno de los dos parálogos (figura 32).

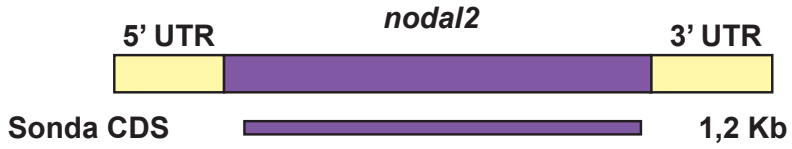


Figura 32. Estrategia de clonaje del gen *nodal2* de *C. fornicata* y diseño de la sonda para el estudio de la distribución espacio-temporal del mRNA. Las zonas en color amarillo representan las zonas reguladoras 5' y 3' UTR; la región morada contiene la región codificante y se diseñó la sonda tomando prácticamente la totalidad de la misma (1,2Kb).

4.2.2.1. *nodal1* se expresa de manera simétrica en *C. fornicata*

La expresión de *nodal1* podría comenzar en un estado avanzado de epibolia, aunque está claramente presente en gástrula media, alrededor del blastoporo. En gástrula tardía u organogénesis, esta expresión se encuentra localizada en la zona anterior y alrededor del estomodeo (región de la boca) y una pequeña región de la zona posterior. En la larva preveliger, esta expresión sigue restringida, aunque de forma más intensa, a la zona anterior y la región de la boca de forma bilateral, así como en la región que dará lugar a la concha (figura 33).

4.2.2.2. *nodal2* se expresa de manera asimétrica en *C. fornicata*

El transcrito de *nodal2* se detecta muy pronto en el desarrollo de *C. fornicata*, en embriones de entre 40-50 células (figura 34). En concreto, en este momento la expresión se detecta en una única célula del cuadrante C (célula $2c^2$). La división celular continua y en estadios posteriores la señal se detecta en la progenie de la célula $2c^2$ (células $2c^{21}$ y $2c^{22}$). Durante la gastrulación la expresión se mantiene en las células descendientes de la celula $2c^2$ y se diferencia en dos dominios (uno anterior y otro posterior) (figura 34). Durante la organogénesis y el desarrollo larvario se observa una expresión completamente asimétrica incluyendo la mitad derecha de la boca, el velo derecho, la mitad derecha del rudimento del pie y un parche en la zona derecha de crecimiento de la concha (figura 34). Esta expresión se mantiene e intensifica a medida que se

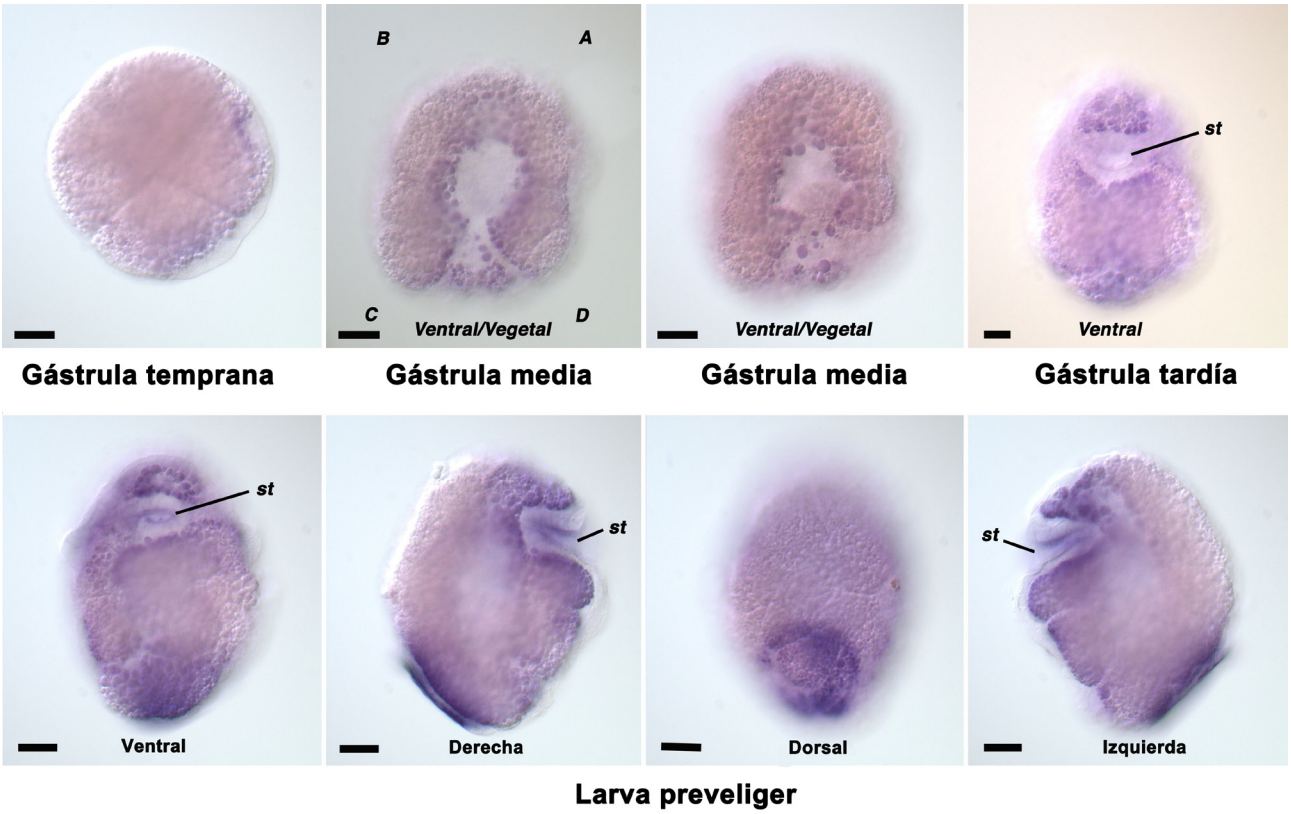


Figura 33. Localización espacio-temporal por hibridación *in situ* del mRNA de *nodal1* a lo largo del desarrollo de la especie *C. fornicata*. El RNAm de *nodal1* se detecta desde gastrulación (paneles superiores) hasta larva veliger (paneles inferiores). En el estadio de gástrula media se indican los cuadrantes A, B, C y D. En la figura se indica cada uno de los estadios así como su orientación. Nótese que la expresión es simétrica en todos los estadios. St, estomodeo. La escala corresponde a 30 µm. Panel cedido por Kimberly J. Perry, laboratorio del Prof. Henry.

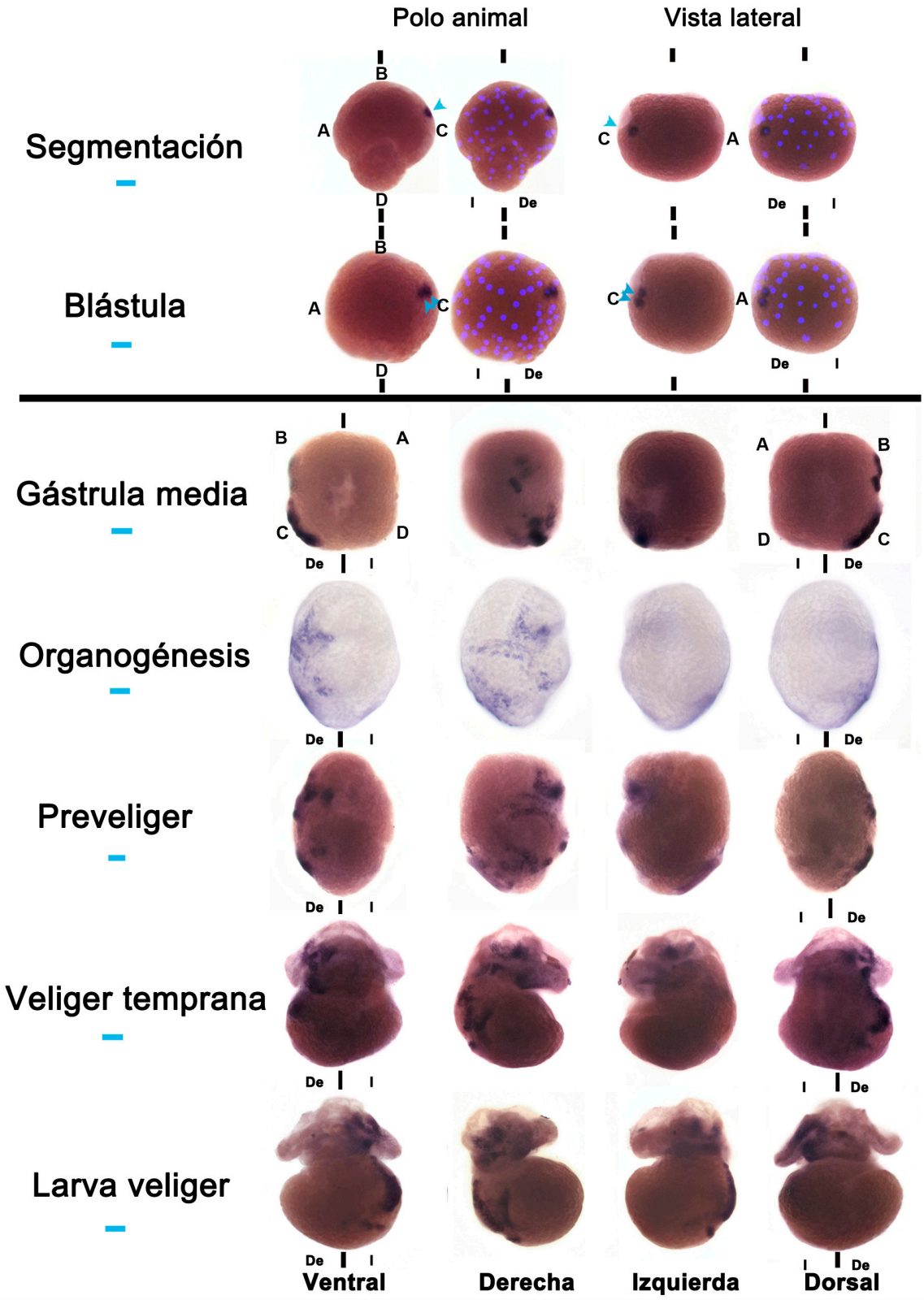


Figura 34. Localización espacio-temporal por hibridación *in situ* del mRNA de *nodal2* a lo largo del desarrollo de la especie *C. fornicata*. El RNAm de *nodal2* se detecta desde los estadios más tempranos del desarrollo (segmentación) hasta larva veliger. En el estadio de segmentación, blástula y gástrula se indican los cuadrantes A, B, C y D. Cada uno de los estadios del panel presenta varias vistas del mismo embrión indicadas en la figura. A la derecha de cada embrión de los estadios de segmentación y blástula, se han realizado superposiciones de las imágenes tomadas en campo claro con imágenes tomadas en fluorescencia para visualizar núcleos teñidos con Hoestch. De: lado derecho; I: lado izquierdo; las barras representan el eje bilateral. Las cabezas de flecha azules indican las células teñidas. La escala se indica con la línea de color azul, que corresponde a 30 µm.

desarrolla la larva preveliger. Finalmente, en la larva veliger la expresión de *nodal2* se limita a la zona derecha de la masa visceral y la cabeza (figura 34).

4.2.3. Detección de la proteína Nodal2 en *C. fornicata*

Durante el desarrollo de esta tesis se diseñó un anticuerpo específico para la proteína Nodal2 de la especie *C. fornicata* (apéndice 4; tabla 2). Se obtuvieron sueros provenientes de dos conejos (A y B) inoculados con los péptidos diseñados específicamente para esta especie de gasterópodo que fueron probados en experimentos de *western blot*.

Los resultados de los experimentos de *western blot* muestran que el suero proveniente de la inoculación de los péptidos de la proteína Nodal2 de *C. fornicata* podría estar potencialmente reconociendo dicha proteína, ya que pueden distinguirse bandas intensas del tamaño esperado (54KDa) en el experimento correspondiente al uso del suero proveniente de los conejos A y B

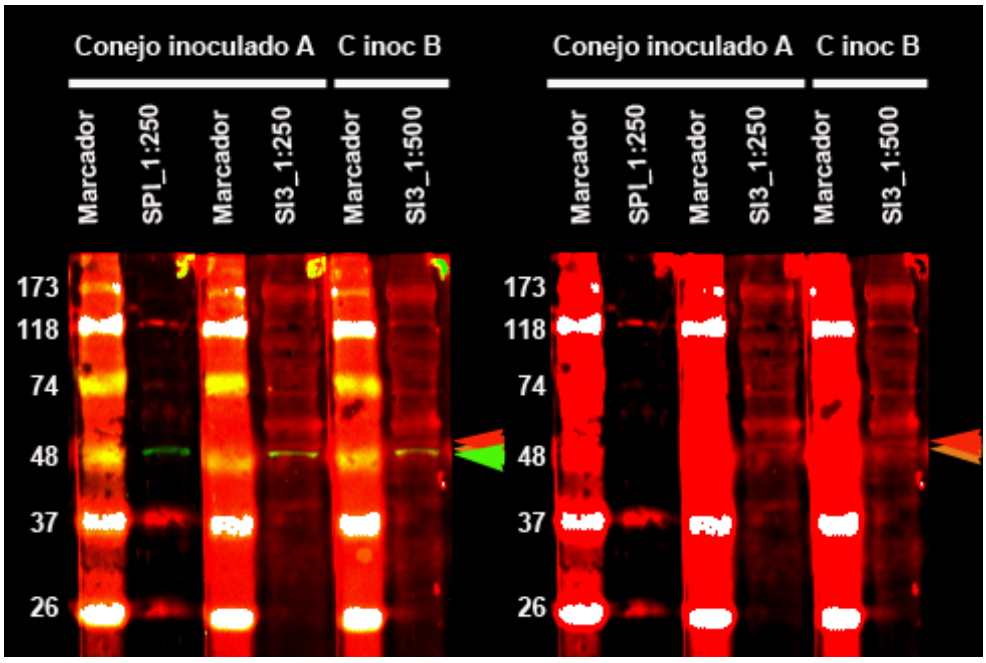


Figura 35. Detección de la proteína Nodal2 en un extracto de proteína total en *C. fornicata* por *western blot*. La imagen muestra dos geles idénticos: uno revelado en los canales rojo y verde (izquierda) y otro revelado con el canal rojo exclusivamente (derecha). De izquierda a derecha en cada carril se indica el marcador de peso molecular (con los números a la izquierda que indican la referencia del marcador en KDa), control SPI_1:250 (incubación con el suero del conejo A extraído antes de la inoculación de los péptidos a una concentración 1:250), marcador, experimento_SI3_1:250 (incubación con el suero 3 del conejo A extraído después de la incubación con los 3 péptidos a una concentración 1:250), marcador, experimento_SI3_1:500 (incubación con el suero 3 del conejo B extraído después de la incubación con los 3 péptidos a una concentración 1:500). El peso estimado es 54 KDa, señalado con las flechas roja y naranja, puesto que estimamos que su tamaño podría ser cualquiera de los dos. El control de carga se realizó con β-Tubulina, revelado en verde y señalado con una flecha verde.

(figura 35). Estas bandas no se detectan cuando se incuba con el suero control, el proveniente de los mismos conejos, pero extraído previamente a la inoculación del antígeno (figura 35). Estas bandas se observan mejor al eliminar el canal que revela el control de carga, con el cual solapa por tamaño.

4.2.4. Nodal2 regula el establecimiento de AID en *C. fornicata*

Para investigar la función del gen *nodal2* en el gasterópodo *C. fornicata* se siguieron tres estrategias principales: 1) el uso de CRISPR 2) el uso de formas silvestres y defectivas (dominante negativo) que fueron inyectadas en embriones de *C. fornicata* y 3) la adición de la proteína recombinante Activin en AFSW donde se crecieron los embriones de *C. fornicata* (figura 36).

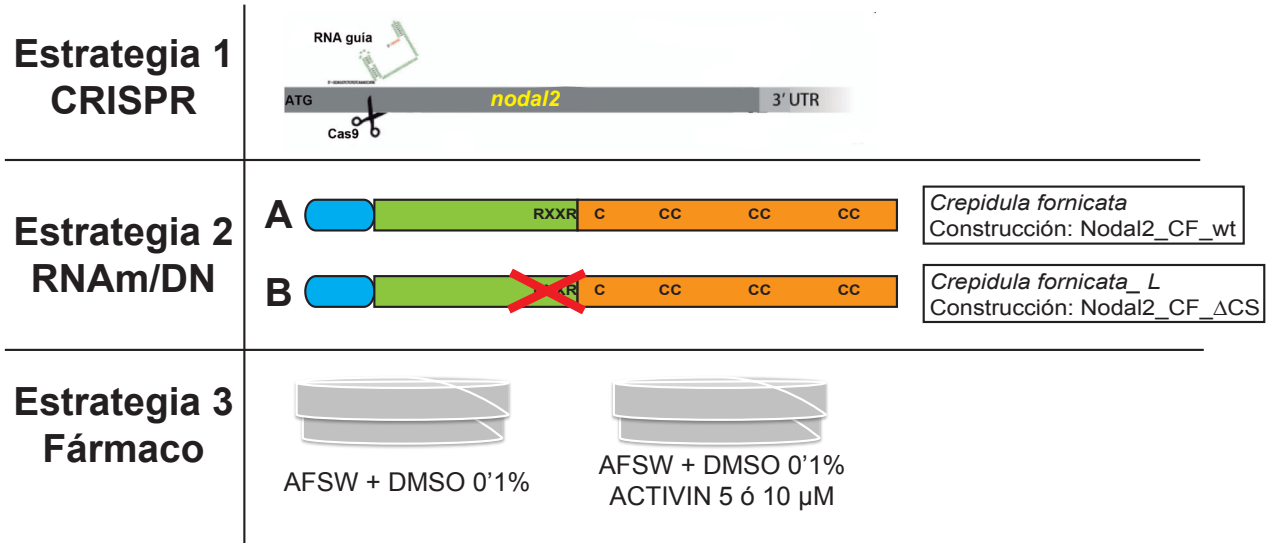


Figura 36. Estrategias seguidas para el abordaje funcional del gen *nodal2* en la especie *C. fornicata*. Estrategia 1: uso de la tecnología CRISPR para silenciar el gen *nodal2* de *C. fornicata*; Estrategia 2: uso de mRNA proveniente de construcciones con A) la forma silvestre del gen *nodal2* de *C. fornicata*; B) la forma del gen *nodal2* de *C. fornicata* con una deleción de la zona de procesamiento de la pre-proteína; Estrategia 3: uso de la proteína recombinante humana Activin disuelta en AFSW a una concentración de 5 y 10 µM.

La microinyección de CRISPRs (apartado 3.18.) para generar individuos carentes del gen *nodal2* en *C. fornicata* no mostró ningún fenotipo anormal ni malformación visible a lo largo del desarrollo (tabla 11).

	Total	Fenotipo WT	Fenotipo simetría
Control	129	100%	0%
CRISPR	29	100%	0%

Tabla 11. Resultados obtenidos tras la inyección del RNAm de *Cas9* y el RNA guía específico de *nodal2* en la especie *C. fornicata*. De los 29 zigotos inyectados el 100% presentaba un fenotipo igual al silvestre.

En cuanto a la estrategia con construcciones (apartados 3.15. y 3.16.), en primer lugar se preparó una construcción con el mRNA del gen *nodal2* silvestre de *C. fornicata* y se inyectó en zigotos de esta especie (experimento A en la figura 36 y en la tabla 12). Este experimento pretendía comprobar si la sobreexpresión del gen *nodal2* tenía algún efecto en el desarrollo de este gasterópodo. Los resultados muestran que un alto porcentaje de los embriones inyectados tienen un fenotipo simétrico tanto en la morfología de la concha como en la anatomía interna (figura 37; tabla 12). El porcentaje de individuos con fenotipo simétrico obtenido está directamente relacionado con la concentración de mRNA inyectado, llegando hasta un 92% en los experimentos con la mayor concentración de mensajero (tabla 12).

A continuación, se preparó una construcción con el mRNA del gen *nodal2* de *C. fornicata* carente de la zona de corte de la pre-proteína y se inyectó en zigotos de esta especie (experimento B en la figura 36 y en la tabla 12). El objetivo de este experimento era comprobar si eliminando

	RNAm inyectado	Total	Fenotipo WT	Fenotipo simetría
Experimento A (Sobre-expresión)	CONTROL	26	100%	0%
	Nd2_CF_WT Stock a 450 ng/µl	20	45% (9)	55% (11)
Experimento A (Sobre-expresion)	CONTROL	30	100%	0%
	Nd2_CF_WT stock a 600 ng/µl	25	8% (2)	92% (23)
Experimento B (Deleción del sitio de corte)	CONTROL	48	100%	0%
	Nd2_CF_ ΔCS	74	100%	0%

Tabla 12. Resultados obtenidos tras la inyección de mRNA proveniente de diferentes construcciones en la especie *C. fornicata*. Experimento A: inyección de la forma silvestre del gen *nodal2* de *C. fornicata* a 450 ng/µl y 600 ng/µl; de los 20 zigotos inyectados a 450 ng/µl de concentración, el 55% presentaba un fenotipo de simetría bilateral; de los 25 zigotos inyectados a 600 ng/µl de concentración, el 92% presentaba un fenotipo de simetría bilateral; Experimento B: inyección de la forma del gen *nodal2* de *C. fornicata* con una deleción de la zona de procesamiento de la pre-proteína; de los 74 zigotos inyectados, el 100% presentaba un fenotipo igual al silvestre.

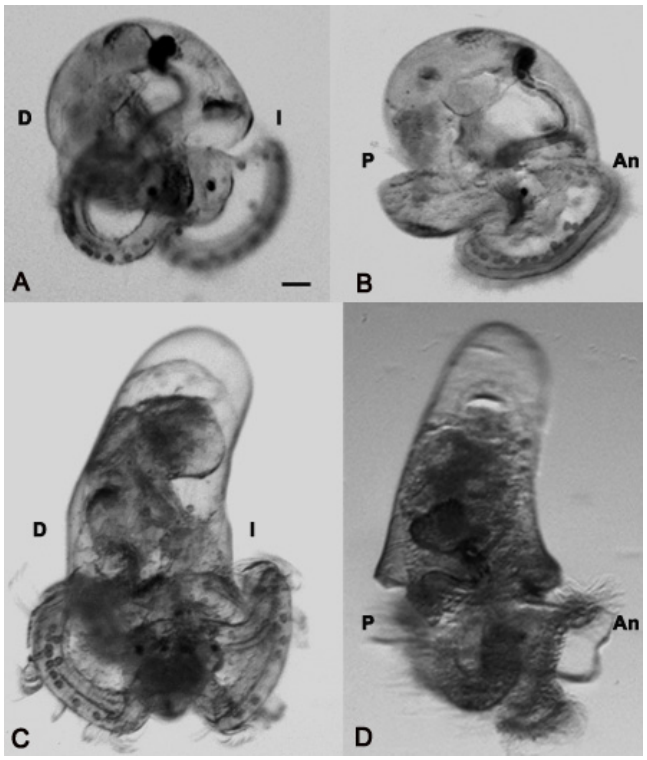


Figura 37. Morfología de la larva veliger de *C. fornicata.* A-B) larva veliger control; A, vista frontal; B, vista lateral. C-D) larva veliger resultante de un zigoto inyectado con la forma silvestre de del gen *nodal2* de *C. fornicata*; C, vista frontal; D, vista lateral; Nótese en C y D el fenotipo de simetría bilateral; An, anterior; P, posterior; I, izquierda; D, derecha. La escala corresponde a 30 µm.

la zona de corte y por tanto evitando potencialmente el procesamiento de la pre-proteína Nodal, se conseguía evitar la activación de la ruta de Nodal de manera específica para *nodal2*. En la secuencia de Nodal2 se detectaron dos potenciales puntos de corte próximos entre sí; ambas zonas fueron delecionadas para asegurar que ninguna de las dos estaba actuando en los embriones inyectados. Cuando se analizaron los fenotipos resultantes no se encontró ninguno diferente al silvestre (tabla 12).

Finalmente se realizaron diferentes experimentos utilizando la proteína Activin recombinante (proteína implicada en la activación de la ruta de Nodal en algunos vertebrados) a modo de fármaco, añadiéndola en AFSW donde se desarrollaron los embriones de *C. fornicata* (apartado 3.22.). Estos experimentos se realizaron para comprobar si esta proteína recombinante era capaz de interferir el correcto funcionamiento de la ruta de Nodal en gasterópodos. Con el fin de establecer la potencial ventana temporal de actuación de este fármaco, se trataron durante dos días embriones de cuatro estadios distintos (7, 32, 48 y 61 horas después de la fertilización) (tabla 13). En embriones de 48 horas se observó un fenotipo de adquisición de simetría (con conchas simétricas y pérdida de torsión de la masa visceral) dependiente de la concentración empleada. El 60% de los individuos sometidos a la concentración más alta mostraba simetría en su plan corporal (figura 38). El tratamiento con Activin en embriones de dos células (7 horas después de la fertilización) durante 48 horas no produjo embriones simétricos (tabla 13). Sin embargo, los embriones de 32 horas después de la fertilización tratados durante las siguientes 48 horas terminaron desarrollando fenotipos simétricos en un 80,4% de los casos (tabla 13). También los embriones de 61 horas después de la fertilización tratados durante las siguientes 48 horas terminaron desarrollando fenotipos

	Concentración	Inicio	Final	Total	Fenotipo WT	Fenotipo de simetría
Experimento A	Control	48h	96h	70	70	0
	5 µM			45	41	4 (8,8%)
	10 µM			43	18	25 (60%)
Experimento B	Control	7h	48h	65	65	0
	10 µM			56	56	0
Experimento C	Control	32h	80h	65	65	0
	10 µM			51	10	41 (80,4%)
Experimento D	Control	61h	109h	30	0	0
	10 µM			32	12	20 (62,5%)

Tabla 13. Resultados obtenidos tras el tratamiento con Activin sobre embriones de la especie *C. fornicata*. Cada experimento A, B, C y D indica un tratamiento en embriones de 48, 7, 32 y 61 horas después de la fertilización, respectivamente. Todos los tratamientos duraron 48 horas. Todos se hicieron a una concentración de Activin 10 µM excepto los embriones de 48 horas que también se trataron con una concentración de 5 µM. De los 51 embriones en estadio de 32 horas después de la fertilización tratados con Activin, el 80,4% presentaba un fenotipo de simetría bilateral.

simétricos en un 62,5% de los casos (tabla 13). Estos datos indican que la presencia de la proteína recombinante Activin interfiere en el correcto establecimiento de AID en el gasterópodo *C. fornicata* y, particularmente, si se añade en embriones en estadio de 32 horas después de la fertilización.

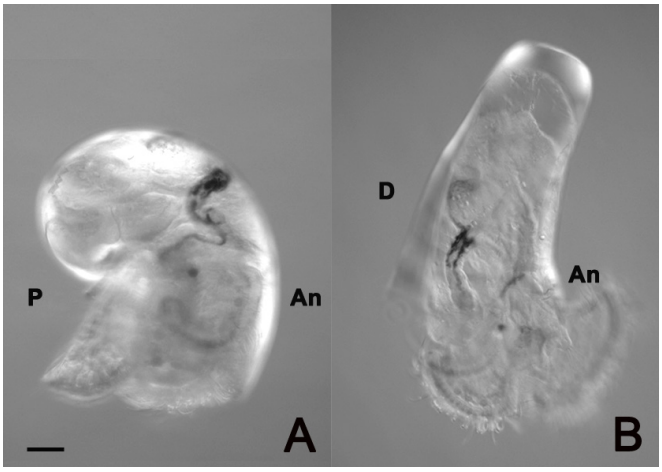


Figura 38. Morfología de la larva veliger de *C. fornicata*. A) larva veliger control; B) larva veliger resultante tras la incubación de los embriones en AFSW con la proteína Activin. An, anterior; P, posterior; D, derecha. Nótese en B el fenotipo de simetría bilateral. Ambas larvas en vista lateral. La escala corresponde a 30 µm.

Con el fin de determinar el mecanismo por el cual la proteína recombinante Activin afecta el establecimiento de AID en los embriones de *C. fornicata* se estudió en detalle un gen diana de la ruta de Nodal en los gasterópodos, *Pitx*. El gen *Pitx* es regulado por Nodal en gasterópodos y se activa de manera asimétrica en el lado derecho siguiendo la activación del gen *nodal* (figura 39). Para comprobar si la ruta de Nodal estaba siendo afectada por el tratamiento con la proteína Activin, se fijaron larvas de embriones tratados y se sometieron a hibridaciones *in situ* para el gen *Pitx*. Los resultados muestran que la expresión asimétrica de *Pitx* desaparece en larvas con tratamiento de Activin (figura 40).

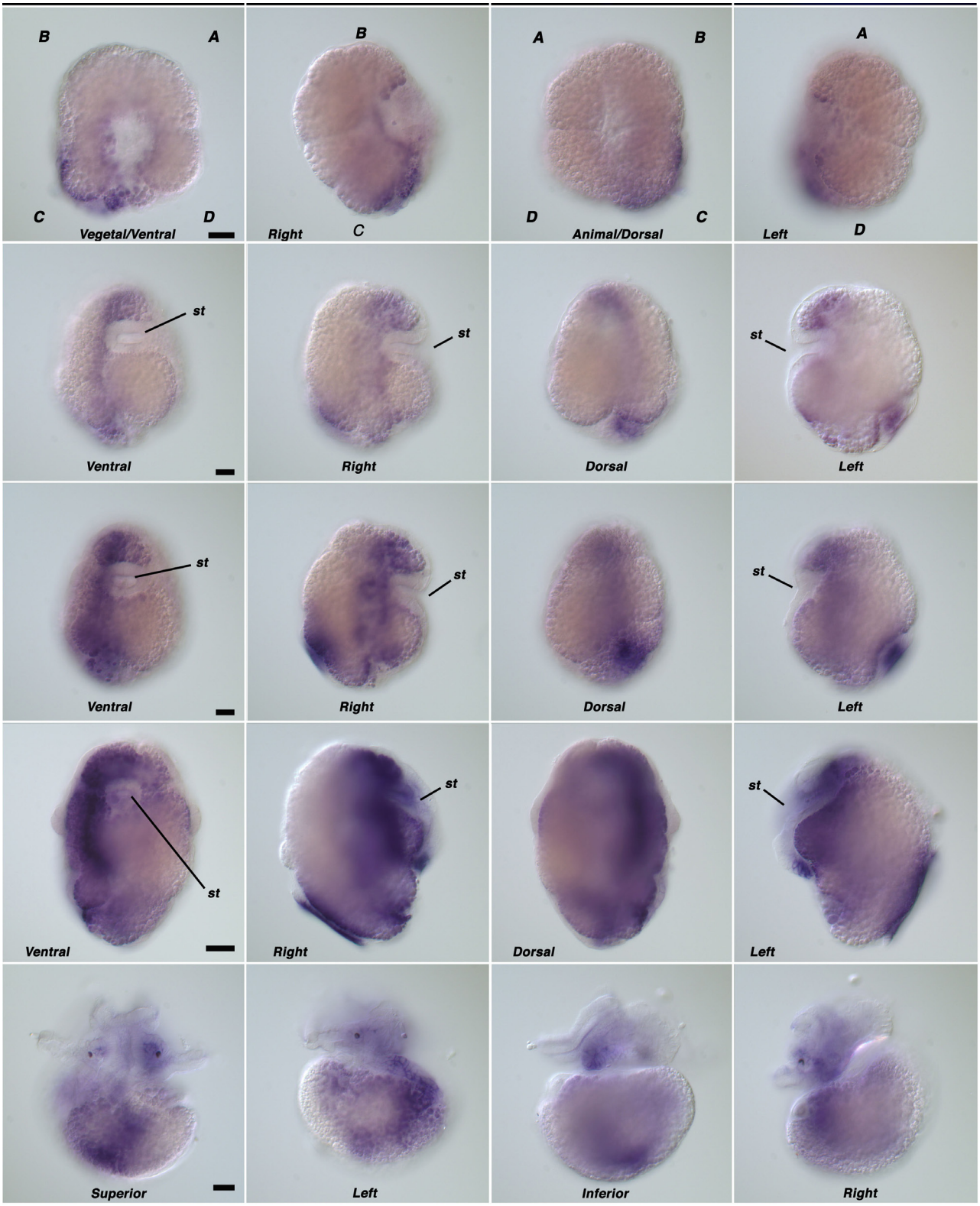


Figura 39. Localización espacio-temporal por hibridación *in situ* del mRNA de *Pitx* a lo largo del desarrollo de la especie *C. fornicata*. El RNAm de *Pitx* se detecta desde gastrulación media (fila superior) hasta larva veliger (fila inferior). En el estadio de gástrula se indican los cuadrantes A, B, C y D. En cada fila se muestran varias vistas del mismo embrión indicadas en la figura. St, estomodeo. La escala corresponde a 30 µm. Panel cedido por Kimberly J. Perry, laboratorio del Prof. Henry.

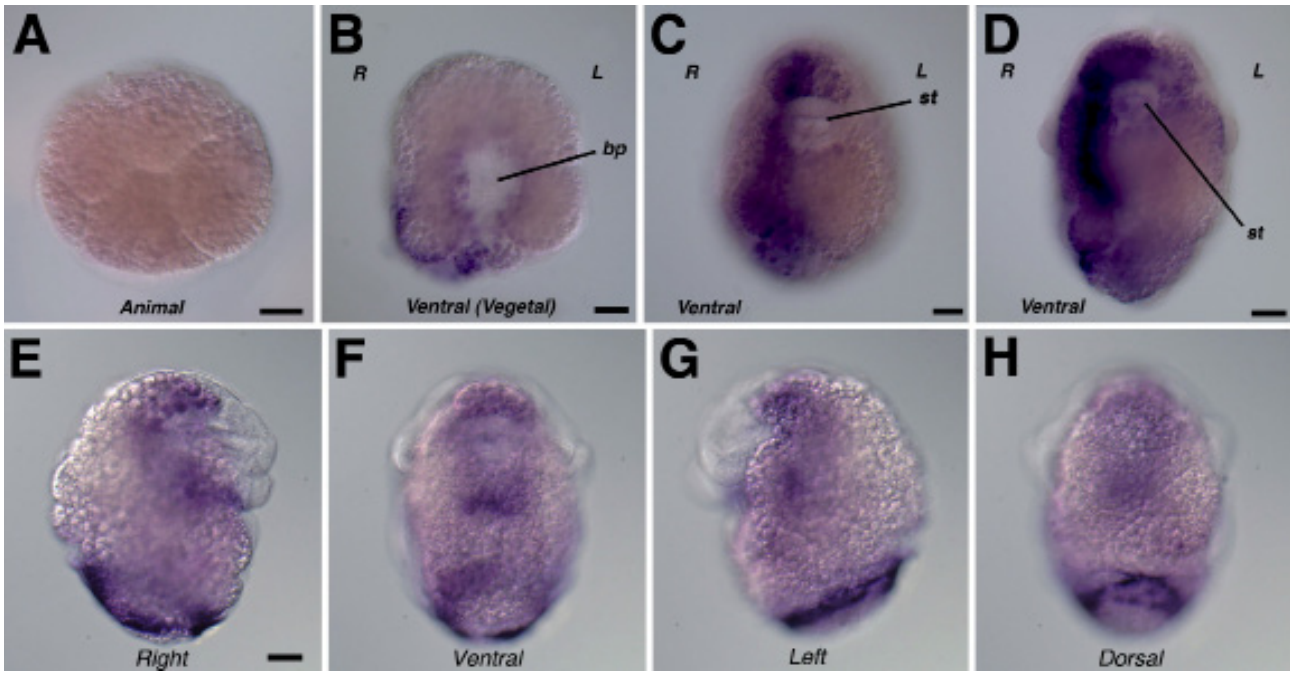


Figura 40. Localización espacio-temporal por hibridación *in situ* del mRNA de *Pitx* a lo largo del desarrollo de la especie *C. fornicata* en embriones tratados con Activin. A-D) Embriones control según la vista que se indica en cada fotografía; A, blástula sin tinción; B, gástrula; C, organogénesis; D, larva veliger; Nótese la tinción en el lado derecho y la tinción simétrica en la zona anterior y en el estomodeo en los estadios B-D; L, izquierda; R, derecha; bp, blastoporo; st, estomodeo; E-H) Larva preveliger en vistas laterales, ventral y dorsal tras el tratamiento con Activin. Nótese como la señal simétrica se mantiene pero se pierde la señal asimétrica del lado derecho. La escala corresponde a 30 μm. Panel cedido por Kimberly J. Perry, laboratorio del Prof. Henry.

4.3. Caracterización de las proteínas EGF-CFC en Metazoa y análisis funcional en Gastropoda

4.3.1. Moluscos, anélidos y braquiópodos cuentan con proteínas EGF-CFC

Las búsquedas realizadas en las bases de datos genómicas y transcriptómicas de metazoos consultadas se realizaron a partir de los genes *cripto* y *cryptic* de *Mus* y *oep* de *D. rerio*. Los resultados de estas búsquedas y los subsiguientes análisis filogenéticos han permitido

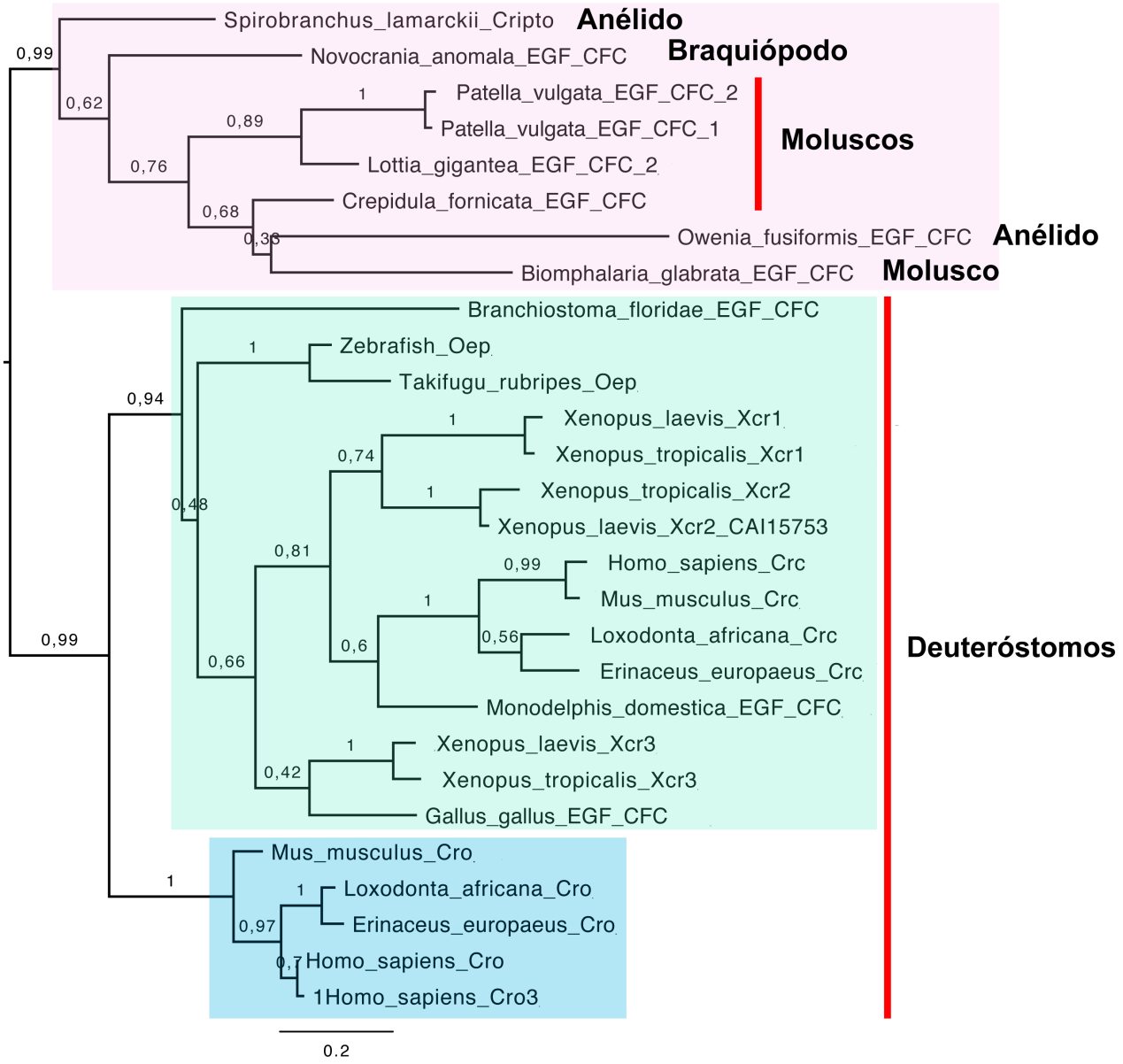


Figura 41. Análisis filogenético de las proteínas EGF-CFC en metazoos. Árbol consenso 50%-*majority rule* post *burn-in* resultado de un análisis de inferencia bayesiana basado en la secuencia de aminoácidos de las proteínas EGF-CFC de varios metazoos. El modelo de sustitución aminoacídica usado es WAG+I+G. Los números sobre los nodos representan la probabilidad posterior bayesiana.

identificar ortólogos del gen *egf-cfc* en moluscos, anélidos y braquiópodos (figura 41). Se encontró un único representante del gen *egf-cfc* en todos los bilaterados, excepto en *Xenopus*, *Patella* y mamíferos (figura 41; apéndices 8 y 9). También se buscó de manera activa en los genomas de cnidarios, ctenóforos, ecdisozoos y xenacelomorfos pero no se obtuvo ningún resultado positivo.

El análisis de la estructura molecular de las proteínas ortólogas identificadas muestra que todas ellas tienen una estructura muy conservada que incluye un dominio de tipo EGF, un dominio CFC y un dominio GPI (figuras 42, 43 y apéndice 10).

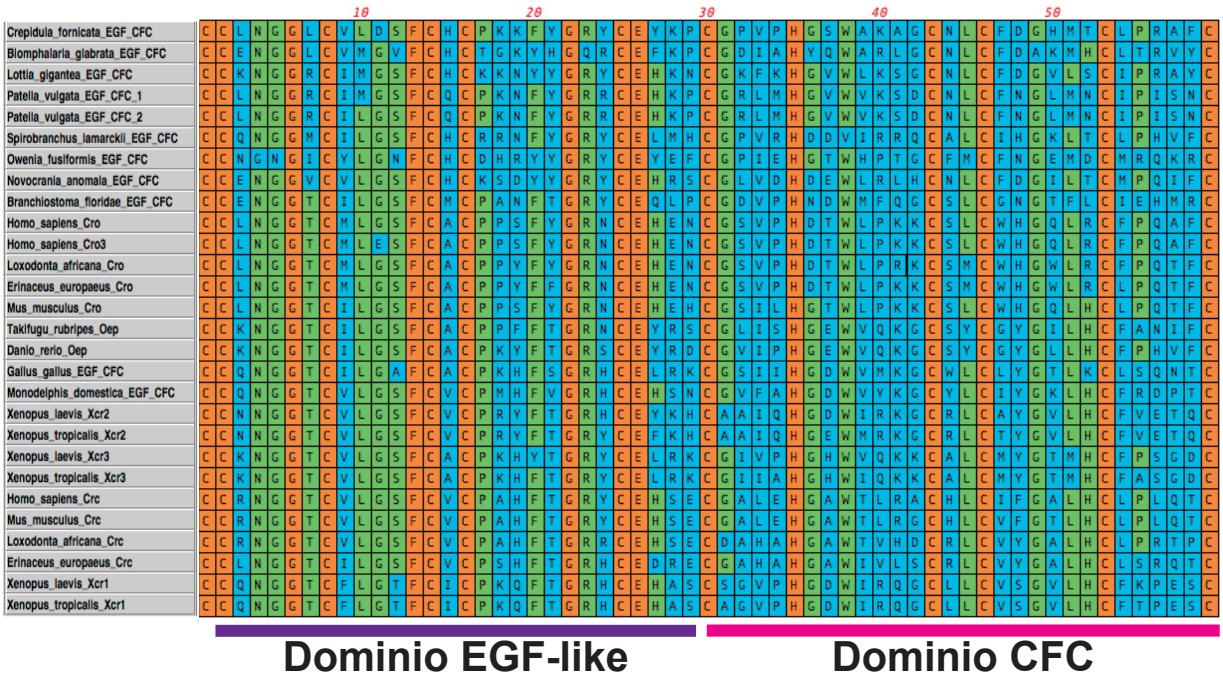


Figura 42. Alineamiento de proteínas ortólogas EGF-CFC de varias especies de bilaterados. Se muestran los dominios EGF y CFC. Los colores indican que todas las posiciones son idénticas en todas las especies (naranja); que las posiciones son idénticas en el 50% ó más de las especies (verde) y posiciones variables (azul).

Figura 43. Estructura genérica de la proteína EGF-CFC. En azul se indica la región del péptido señal. En morado se indica la región del dominio tipo EGF. En rosa se indica la región que conforma el dominio CFC. En verde se indica la región del dominio GPI de anclaje a membrana.

4.3.2. Los transcritos sentido y antisentido del gen *egf-cfc* en *C. fornicata* están activos durante todo el desarrollo embrionario

Con el fin de detectar el patrón de activación espacio-temporal del gen *egf-cfc* durante el desarrollo embrionario de *C. fornicata* se diseñaron tres sondas (sonda CR_1, CR_2 y CR_3) (figura 44). La sonda CR_1 incluía casi la totalidad de la región codificante; la sonda CR_2 incluía desde la mitad del dominio EGF hasta un fragmento de la zona 3'UTR; la sonda CR_3 incluía el dominio CFC y la zona 3'UTR con la zona de poliadenilación incluida. Las señales con las sondas CR_1 y CR_2 fueron congruentes, mientras que no se obtuvo ninguna señal utilizando la sonda CR_3.

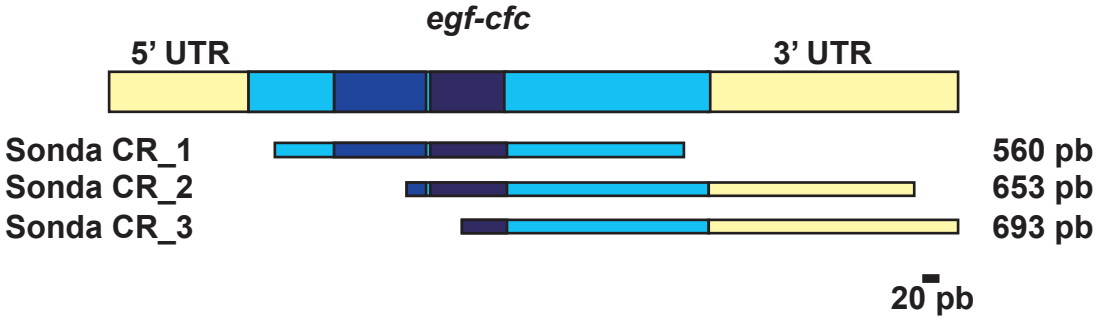


Figura 44. Estrategia de clonaje del gen *egf-cfc* en la especie *C. fornicata* y diseño de las sondas para el estudio de la distribución espacio-temporal del mRNA. Las zonas en color amarillo representan las zonas reguladoras 5' y 3' UTR; la región azul contiene la región codificante. En los colores más oscuros se representan las regiones codificantes para los dominios tipo EGF y CFC. La sonda CR1 contiene la práctica totalidad de la región codificante; la sonda CR2 comprende gran parte de la zona 3' UTR hasta un tercio del dominio tipo EGF. La sonda CR3 contiene toda la región 3'UTR y casi toda la región CFC. La barra indica la escala.

El transcrito del gen *egf-cfc* se detecta desde las primeras divisiones en todos los blastómeros del embrión. Durante el estadio de blástula se mantiene la expresión en todos los blastómeros aunque en las gástrulas se restringe a la zona del labio dorsal principalmente. En el estadio de preveliger, la señal se hace muy intensa y se restringe a la zona del estomodeo y parte del velo siempre de forma simétrica (figura 45). En este estadio también se observan señales discretas en células a ambos lados, en la zona posterior y en la zona de la concha (figura 45), aunque cabe destacar que estas señales no eran consistentes entre individuos del mismo estadio.

Por otra parte, cuando se hibridó con la sonda sentido (realizada en el sentido de transcripción del RNA) se obtuvieron los mismos resultados que con la sonda antisentido, aunque los dominios de expresión eran ligeramente más amplios (figura 45).

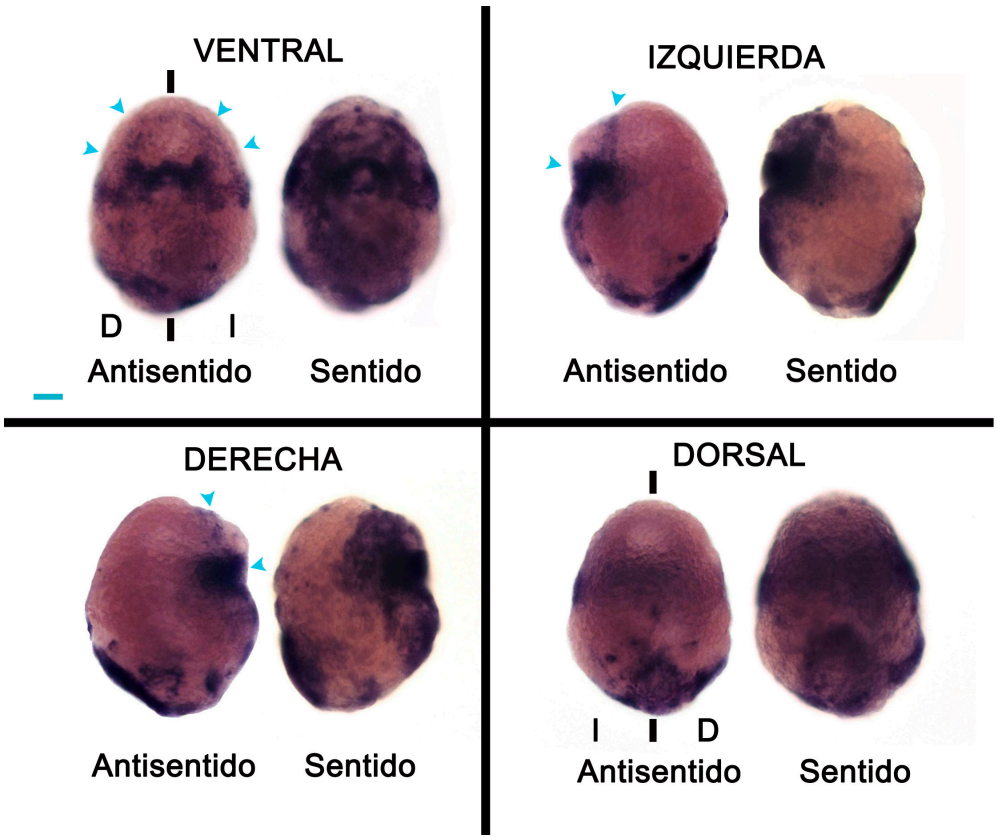


Figura 45. Localización espacio-temporal por hibridación *in situ* del mRNA de *egf-cfc* en larva preveliger de la especie *C. fornicata*. En el panel se muestra larvas preveliger en distintas vistas (ventral, izquierda, derecha, dorsal). Para cada vista se muestran larvas incubadas con la sonda antisentido y la sonda sentido. Las flechas azules indican las señales simétricas, las barras indican el eje bilateral. I, izquierda; D, derecha. La escala se indica con la línea de color azul, que corresponde a 30 µm.

4.3.3. La proteína EGF-CFC en *C. fornicata* se localiza en el labio dorsal durante la gastrulación

Los experimentos de inmunohistoquímica realizados utilizando el anticuerpo comercial para Cripto de humanos (ver apartado 3.12.) en embriones de *C. fornicata* no produjeron ningún resultado en ninguna de las concentraciones probadas. Por ello, se decidió diseñar un anticuerpo específico para las proteínas EGF-CFC de las especies *C. fornicata* y *B. glabrata* (apéndice 4; tabla 2).

Se obtuvieron sueros provenientes de dos conejos inoculados con los péptidos diseñados específicamente para cada una de las especies de gasterópodos que fueron probados en experimentos de *western blot*. Los resultados de estos experimentos muestran que el suero proveniente de la inoculación de los péptidos de la proteína EGF-CFC de *B. glabrata* podría estar potencialmente reconociendo dicha proteína, ya que puede distinguirse una intensa banda del tamaño esperado (23KDa) en el experimento correspondiente al uso del suero proveniente

del conejo B (figura 46). Esta banda no se detecta cuando se incuba con el suero control, el proveniente del mismo conejo, pero extraído previamente a la inoculación del antígeno (figura 46).

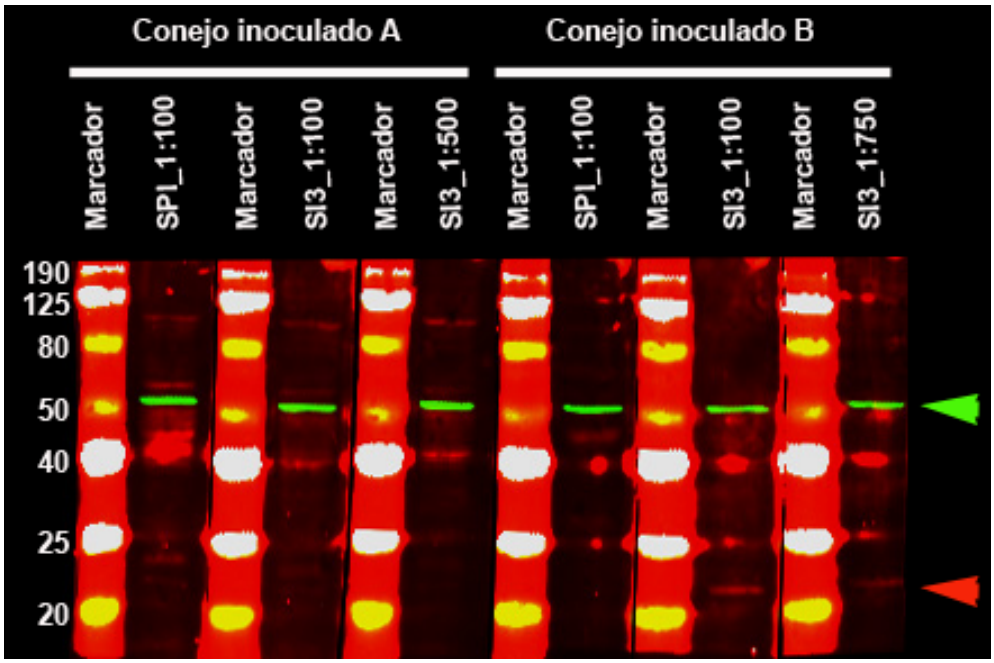


Figura 46. Detección de la proteína EGF-CFC en un extracto de proteína total en *B. glabrata* por *western blot*. De izquierda a derecha en cada carril se indica el marcador de peso molecular (con los números a la izquierda que indican la referencia del marcador en KDa), control_SPI_1:100 (incubación con el suero del conejo A extraído antes de la inoculación de los péptidos a una concentración 1:100), marcador, experimento_SI3_1:100 (incubación con el suero 3 del conejo A extraído después de la inoculación con los 3 péptidos a una concentración 1:100), marcador, experimento_SI3_1:500 (incubación con el suero 3 del conejo A extraído después de la inoculación con los 3 péptidos a una concentración 1:500), marcador, control_SPI_1:100 (incubación con el suero del conejo B extraído antes de la inoculación de los péptidos a una concentración 1:100), marcador, experimento_SI3_1:100 (incubación con el suero 3 del conejo B extraído después de la inoculación con los 3 péptidos a una concentración 1:100), marcador, experimento_SI3_1:750 (incubación con el suero 3 del conejo B extraído después de la inoculación con los 3 péptidos a una concentración 1:750). El peso estimado es 23 KDa, señalado con la flecha roja. El control de carga se realizó con β-Tubulina, revelado en verde y señalado con una flecha verde.

Por otra parte, el suero proveniente de la inoculación de los péptidos de la proteína EGF-CFC de *C. fornicata* se testó en reacciones de inmunohistoquímica con embriones completos. Los resultados muestran que, mientras en los controles no se apreciaba ningún tipo de señal, las larvas preveliger incubadas con el anticuerpo purificado a partir del suero del conejo B sí mostraban una fuerte tinción. En concreto, esta señal se localiza exclusivamente en las membranas de las células del labio dorsal y en las células localizadas en una posición inmediatamente anterior a ellas (figura 47). La señal es completamente simétrica.

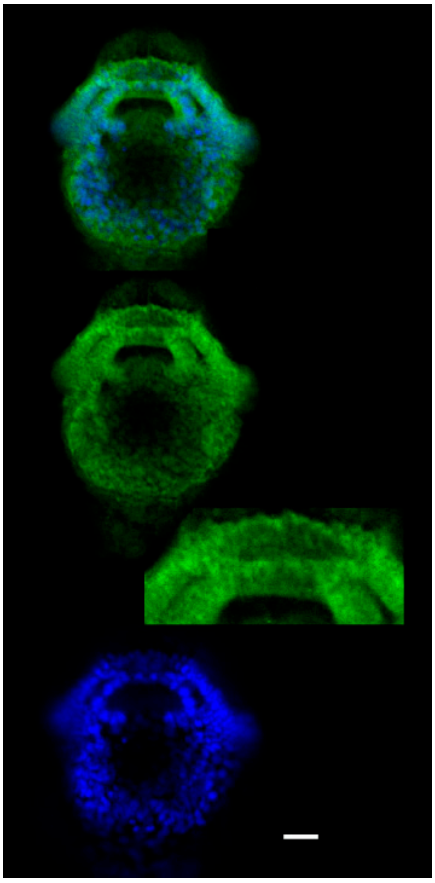


Figura 47. Detección de la proteína EGF-CFC en larvas de *C. fornicata* por inmunofluorescencia. En la imagen se muestra la tinción de una larva preveliger en vista ventral. En el canal verde se muestra la tinción con el anticuerpo específico contra EGF-CFC y en azul la tinción de núcleos con Hoestch. La imagen superior muestra ambos canales al mismo tiempo. En la imagen ampliada, la proteína EGF-CFC puede observarse en la membrana celular de las células que componen el labio dorsal de forma simétrica. La escala corresponde a 30 µm.

4.3.4. EGF-CFC está implicado en la gastrulación de gasterópodos

Para investigar la función del gen *egf-cfc* en el gasterópodo *C. fornicata* se siguieron dos estrategias principales: 1) el uso de un morfolino y 2) el uso de formas silvestres y defectivas (dominante negativo) que fueron inyectadas en embriones de *C. fornicata* y en embriones mutantes para el gen *oep* en *D. rerio* (figura 48).

La microinyección de morfolinos (apartado 3.19.) para intentar bloquear la traducción del mensajero del gen *egf-cfc* de *C. fornicata* no mostró ningún fenotipo anormal ni malformación visible a lo largo del desarrollo (tabla 14).

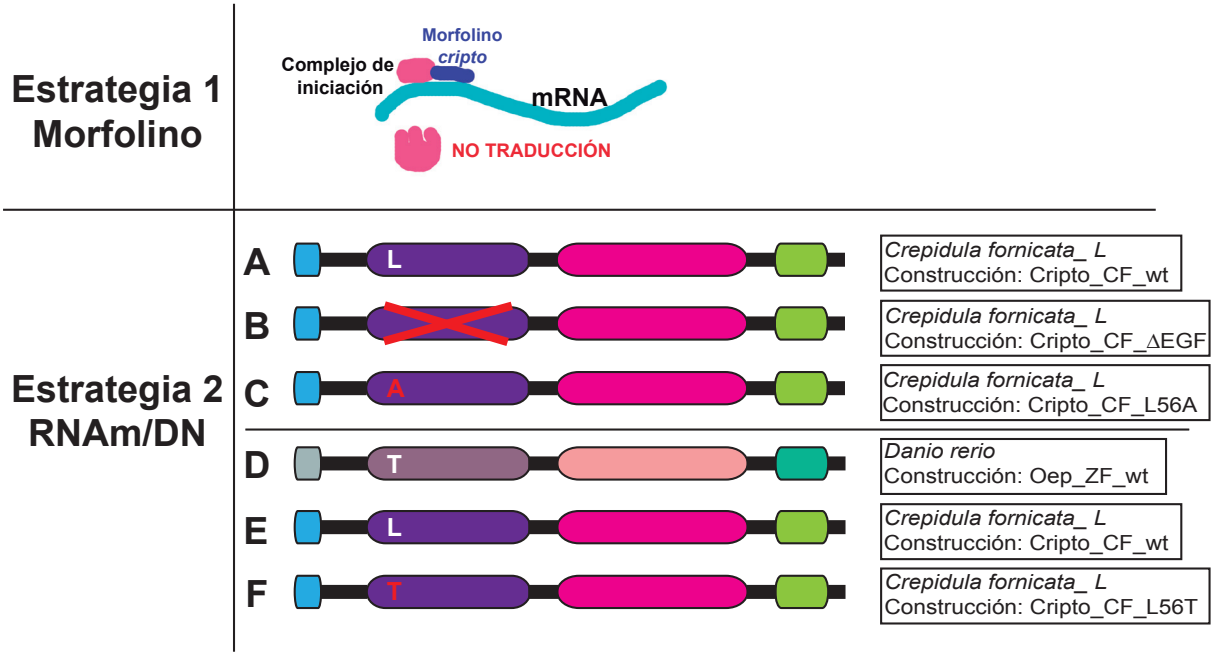


Figura 48. Estrategias seguidas para abordaje funcional del gen *egf-cfc* en la especie *C. fornicata*. Estrategia 1: uso de morfolinos para evitar la traducción del gen *egf-cfc* de *C. fornicata*. Estrategia 2: uso de mRNA proveniente de construcciones con A) la forma silvestre del gen *egf-cfc* de *C. fornicata*; B) la forma del gen *egf-cfc* de *C. fornicata* con una deleción del dominio EGF; C) la forma del gen *egf-cfc* de *C. fornicata* con una sustitución de la leucina 56 por alanina; D) la forma silvestre del gen *oep* de *D. rerio*; E) la forma silvestre del gen *egf-cfc* de *C. fornicata*; F) la forma del gen *egf-cfc* de *C. fornicata* con una sustitución de la leucina 56 por treonina.

	Total	Fenotipo WT	Fenotipo mutante
Control	60	100%	0%
Morfolino	80	100%	0%

Tabla 14. Resultados obtenidos tras la inyección del morfolino contra el RNAm de *egf-cfc* en la especie *C. fornicata*. De los 80 zigotos inyectados, el 100% presentaba un fenotipo igual al silvestre.

En cuanto a la estrategia con construcciones (apartado 3.15. y 3.16.), en primer lugar se preparó una construcción con el mRNA del gen *egf-cfc* silvestre de *C. fornicata* y se inyectó en zigotos de esta especie (experimento A en la figura 48 y en la tabla 15). Estos experimentos pretendían comprobar si la sobreexpresión del gen *egf-cfc* tenía algún efecto en el desarrollo de este gasterópodo. Los resultados no mostraron ningún fenotipo diferente al control (tabla 15).

A continuación, se preparó una construcción con el mRNA del gen *egf-cfc* de *C. fornicata* sin el dominio de tipo EGF y se inyectó en zigotos de esta especie (experimento B en la figura 48 y en la tabla 15). En este caso, los embriones en desarrollo presentaban una mortalidad elevada y los que sobrevivían tenían grandes defectos mostrando un fenotipo radializado sin ningún eje o diferenciación celular (figura 49). La larva radializada es una masa de células transparentes

	RNAm inyectado	Total	Fenotipo WT	Fenotipo radializado
Experimento A (Sobre-expresión)	Control	70	100%	0%
	CR_CF_WT	84	100%	0%
Experimento B (Deleción del dominio tipo EGF)	Control	34	100%	0%
	CR_CF_ΔEGF Stock a 600 ng/µl	62	52 (83,9%)	10 (16,1%)
	CR_CF_ΔEGF Stock a 800 ng/µl	61	10 (16,4%) mortalidad 45,4%	20 (38,2%)
Experimento C (mutación L56A)	Control	34	100%	0%
	CR_CF_L56A	72	100%	0%

Tabla 15. Resultados obtenidos tras la inyección de mRNA proveniente de diferentes construcciones en la especie *C. fornicata*. Experimento A: inyección de la forma silvestre del gen *egf-cfc* de *C. fornicata*; de los 84 zigotos inyectados, el 100 % presentaba un fenotipo silvestre; experimento B: inyección del gen *egf-cfc* de *C. fornicata* con el dominio EGF delecionado a dos concentraciones (600 y 800 ng/µl); de los 123 zigotos inyectados, un porcentaje variable dependiente de concentración presentaba un fenotipo silvestre; experimento C: inyección del gen *egf-cfc* de *C. fornicata* en el que se ha sustituido la leucina de la posición 56 por una alanina. De los 72 zigotos inyectados, el 100 % presentaba un fenotipo silvestre.

elongada en la que se pueden diferenciar tres partes principales que están adheridas a una masa de vitelo (figura 49). Estas larvas no han gastrulado y no se ha producido una diferenciación de tipos celulares.

La treonina situada en la posición 88 (T88) en el ortólogo de Cripto en humanos ha resultado ser crítica para la interacción con Nodal en vertebrados. Un análisis detenido de esta posición en el alineamiento de todos los ortólogos de las proteínas EGF-CFC identificadas en esta tesis muestra que la T88 solamente está conservada en deuteróstomos, mientras que los moluscos poseen una leucina (L) o una arginina (R), los anélidos una metionina (M) o una isoleucina (I) y el braquiópodo una valina (V) (figura 50). La posición homóloga a la T88 de humanos en *C. fornicata* es una leucina situada en la posición 56 (L56) (Figura 50). Con el fin de

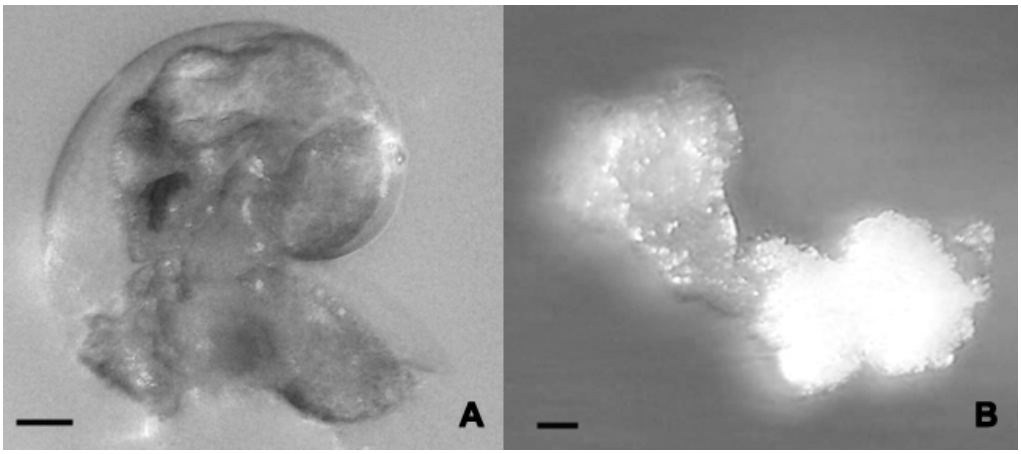


Figura 49. Morfología de la larva veliger de *C. fornicata*. A) larva veliger control, vista lateral; B) larva radializada resultante de un zigoto inyectado con la forma del gen *egf-cfc* de *C. fornicata* con una deleción del dominio EGF. La escala corresponde a 30 µm.

comprobar si esa posición de la proteína EGF-CFC es una posición crítica para el funcionamiento de esta proteína en gasterópodos al igual que lo es en vertebrados, se hicieron diferentes construcciones. En primer lugar, se preparó una construcción del gen *egf-cfc* de *C. fornicata* sustituyendo la L56 por una alanina, y se inyectó en zigotos de esta especie (experimento C en la figura 48 y en la tabla 15). No se observó ningún fenotipo en estos embriones diferente al observado en los embriones control (tabla 15).

A continuación, se realizaron diferentes experimentos en una línea mutante de pez cebra (*D. rerio*) para *oep.* Esta línea mutante es una línea heterozigótica (oep+/-), puesto que los individuos homozigóticos (oep-/-) para este gen no son viables. Los cruces de esta línea heterozigótica poseen una típica segregación mendeliana, en la que un 25% de embriones presentan potencialmente el fenotipo mutante Oep (*oep*-/-) y un 75% de embriones presentan fenotipo de tipo silvestre (*oep* +/-; *oep*+/+) (tabla 16). Esta línea está extensamente caracterizada en la literatura y los fenotipos mutantes Oep muestran ciclopia (la presencia de un único ojo por la ausencia de la división de la cavidad orbitaria en dos órbitas en el prosencéfalo), desviación y acortamiento del eje anterior-posterior, hipertrofia cardiaca y reversión en la posición del corazón (figura 51).

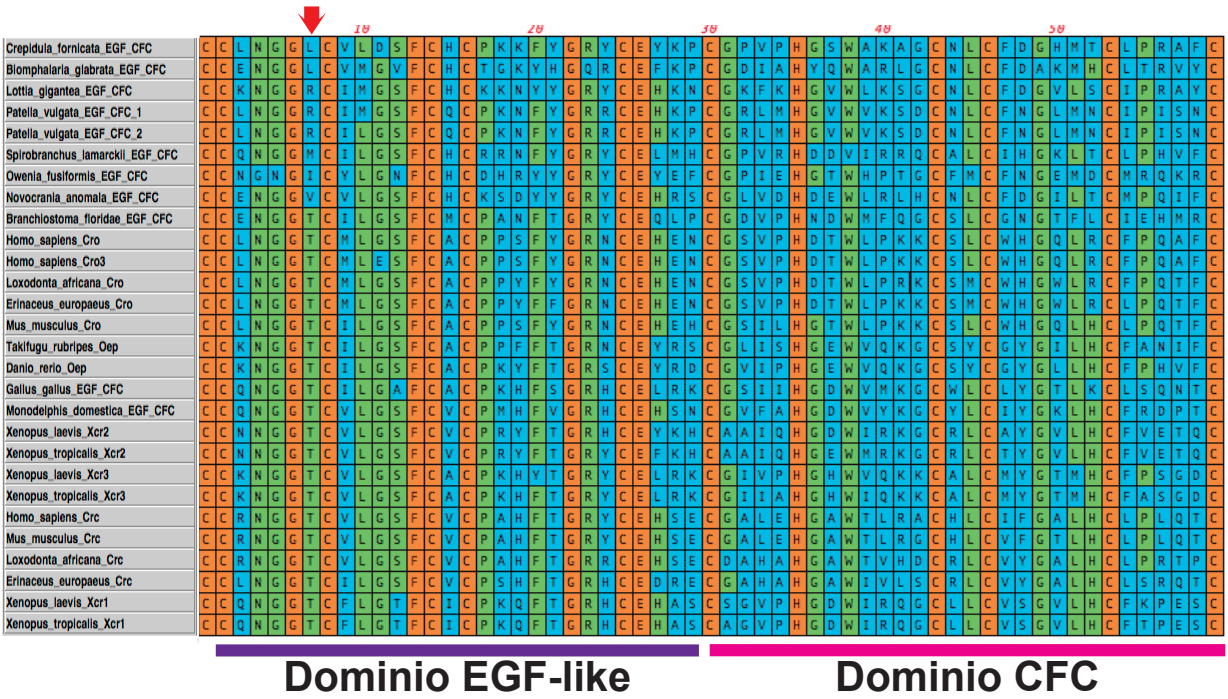


Figura 50. Alineamiento de la región que contiene los dominios tipo EGF y CFC de varias proteínas ortólogas EGF-CFC de varias especies de bilaterados. La flecha roja indica la posición conservada en deuteróstomos que presenta una treonina (T88). Esta posición varía en el resto de representantes no deuteróstomos. Los colores indican que todas las posiciones son idénticas en todas las especies (naranja); que las posiciones son idénticas en el 50% ó más de las especies (verde) y posiciones variables (azul).

Con el fin de comprobar si la inyección del gen *oep* silvestre era capaz de rescatar el fenotipo silvestre en los embriones homozigóticos, se generó mRNA del gen *oep* de *D. rerio* silvestre y se inyectó en zigotos procedentes del cruce de individuos *oep*+/- (experimento D en la figura 48 y en la tabla 16). Los resultados obtenidos muestran que la inyección de mRNA de *oep* sobre dichos zigotos genera un incremento en el recuento de fenotipos silvestres (89%) con respecto al esperado en condiciones de un cruce normal heterozigótico (75,9%) (tabla 16).

	RNA	Total	Fenotipo WT	Fenotipo Oep clásico	Fenotipo Minion
Experimento base (contaje de fenotipos en cruce *oep* +/-)	Control	54	41 (75,9%)	13 (24,1%)	0
Experimento D (rescate con el gen propio)	*oep*	73	65 (89%)	8 (11%)	0
Experimento E (rescate con el gen de *Crepidula*)	*cripto*_CF_WT	97	76 (78,4%)	21 (21,6%)	0
Experimento F (rescate con el gen de *Crepidula* modificado)	*cripto*_CF_L56T	39	33 (89,7%)	4 (10,3%)	2

Tabla 16. Resultados obtenidos tras la inyección de mRNA proveniente de diferentes construcciones en mutantes *oep* de la especie *D. rerio*. Experimento base, contaje de las proporciones de un cruce entre adultos heterozigóticos *oep* +/-: siguiendo una típica segregación mendeliana, el porcentaje de fenotipo silvestre es un 75 % y el porcentaje de fenotipo Oep clásico es un 25 %; experimento D, inyección de la forma silvestre del gen *oep* de *D. rerio*; el porcentaje de fenotipo silvestre sube hasta un 89 %; experimento E, inyección de la forma silvestre del gen *egf-cfc* de *C. fornicata*; el porcentaje de fenotipo silvestre es un 78,4 %; experimento F, inyección del gen *egf-cfc* de *C. fornicata* en el que se ha sustituido la leucina de la posición 56 por una treonina; el porcentaje de fenotipo silvestre sube hasta un 89,7 %.

A continuación se preparó un mRNA del gen *egf-cfc* silvestre de *C. fornicata* y se inyectó en zigotos de la línea mutante Oep (experimento E en la figura 48 y en la tabla 16). En este caso, los resultados muestran que el gen *egf-cfc* silvestre de *C. fornicata* no es capaz de rescatar el fenotipo silvestre (21,6 % de mutantes) (tabla 16; figura 51). Sin embargo, la inyección de zigotos procedentes del cruce entre adultos *oep*+/- en *Danio*, con un mRNA del gen *egf-cfc* de *C. fornicata* en el que se sustituyó L56 por una treonina, sí que rescató el fenotipo silvestre (89,7% de fenotipos silvestres) (tabla 16; figura 51). En algunos casos se detectaron rescates parciales que denominamos fenotipo "Minion", en el que observamos una importante reducción de la hipertrofia cardiaca, una recuperación aparentemente completa del eje antero-posterior, un desarrollo casi completo de la cabeza y la aparición de dos ojos, que pese a estar más juntos que en el fenotipo silvestre, tienen dos lentes independientes. En estos fenotipos Minion la bilateralidad se rescata de forma muy notable (figura 51).

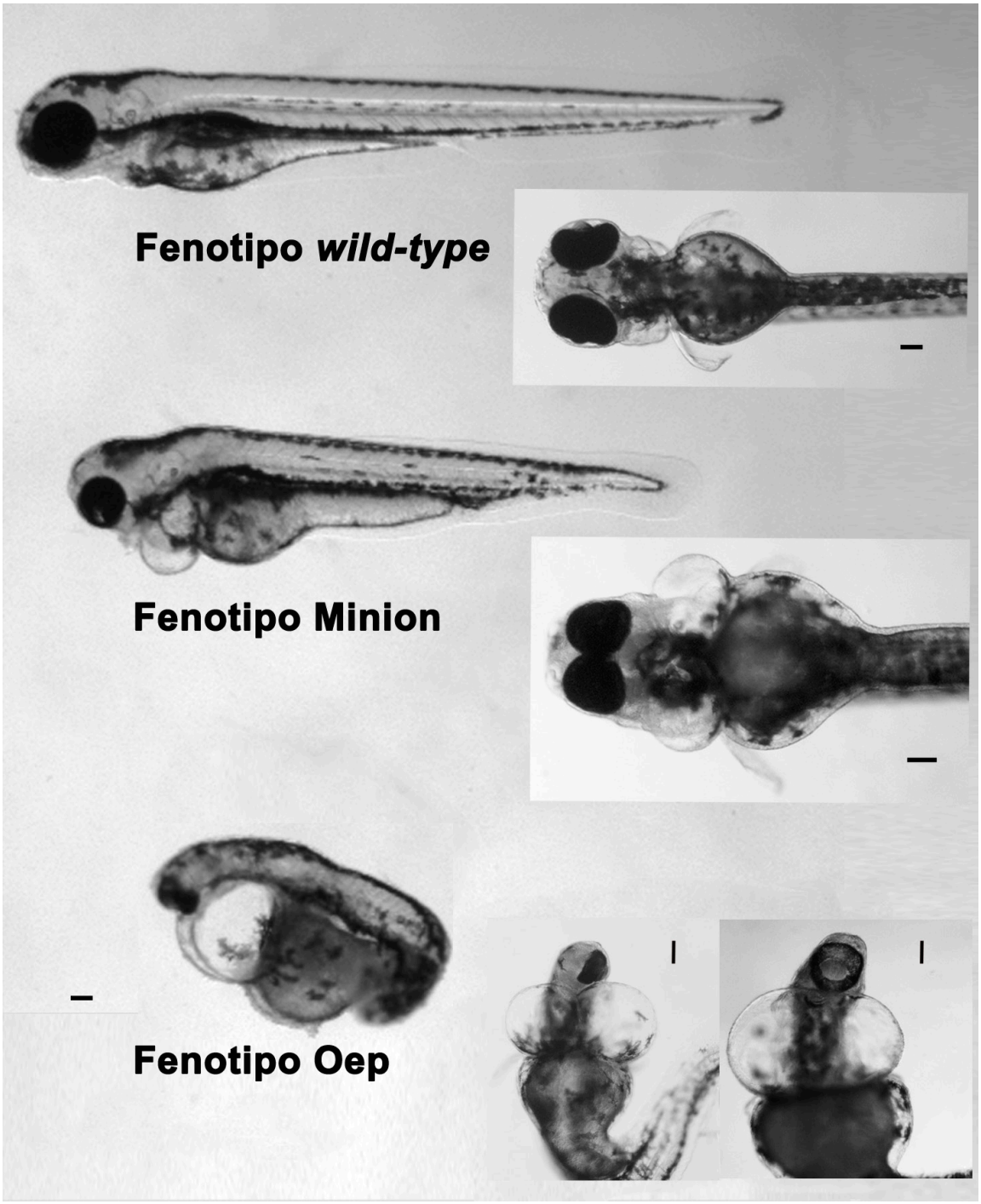


Figura 51. Morfología del pez *D. rerio*. El panel superior muestra un fenotipo silvestre en vista lateral y ventral (recuadro de la derecha). El panel medio muestra un fenotipo Minion en vista lateral y ventral (recuadro de la derecha). Este fenotipo Minion se obtiene después de la inyección de un zigoto de *D. rerio* con la forma del gen *egf-cfc* de *C. fornicata* con una sustitución de la leucina 56 por treonina. El panel inferior muestra un fenotipo Oep clásico en vista lateral y ventral (recuadro de la derecha). Los dos ejemplares en vista ventral muestran dos fenotipos Oep clásicos con un único ojo centralizado (derecha) o desplazado por la errática formación del eje antero-posterior (izquierda). La escala corresponde a 100 μm.

Discusión

El establecimiento de asimetrías morfológicas con respecto al eje de simetría bilateral durante el desarrollo de los organismos bilaterales constituye uno de los ejemplos más interesantes de formación de patrón a diferentes escalas ya que señales a nivel molecular se integran y amplifican en cascadas génicas de activación asimétrica que finalmente conducen a la formación de un patrón asimétrico a nivel del organismo. Por tanto, para entender este proceso se precisa trabajar a varias escalas temporales y espaciales. Consecuentemente, en esta tesis se ha trabajado a diferentes niveles. Por un lado, se ha abordado el estudio de un gen (*inversin*) íntimamente relacionado con el citoesqueleto celular y con la ruptura de la simetría durante los primeros estadios del desarrollo de vertebrados. La mutación en este gen en vertebrados ocasiona 1) una activación de la cascada de señalización que controla las asimetrías (ruta de Nodal) en el lado contrario al silvestre y 2) fenotipos con asimetrías revertidas (*situs inversus*). Por otro lado, se ha trabajado más en detalle en la ruta de señalización Nodal profundizando en el conocimiento del ligando Nodal y del correceptor EGF-CFC, agentes necesarios ambos para la activación de la ruta en vertebrados.

Un funcionamiento deficiente tanto de la proteína Inversin como de los componentes de la ruta de Nodal ocasiona graves malformaciones y defectos en el desarrollo de vertebrados, lo que ha ocasionado que sea en estos organismos donde principalmente se hayan llevado a cabo las investigaciones realizadas hasta el momento. Sin embargo, en esta tesis se ha abordado el estudio de estos genes en otros organismos bilaterales, principalmente en gasterópodos. Los gasterópodos se caracterizan por tener unas asimetrías bilaterales muy bien definidas durante el desarrollo que persisten en las formas adultas. Asimetrías concretas y sus imágenes especulares (*situs inversus*) pueden encontrarse de manera natural, lo que les hace ser unos interesantes modelos para el abordaje de este tipo de estudios. La perspectiva comparada que ha guiado esta tesis ha permitido establecer aquí hipótesis evolutivas sobre estas rutas de señalización e inferir el origen de estos genes y su potencial papel en la formación de patrones asimétricos en los metazoos. A continuación se discuten los principales resultados obtenidos de los estudios aquí realizados.

5.1. La proteína Inversin en Metazoa

Inversin se identificó como la proteína responsable de causar *situs inversus* en ratones mutantes para el gen que la codifica, así como ocasionar malformaciones en los riñones, que no completan el desarrollo durante la embriogénesis y sufren la formación de multitud de quistes

(Mochizuki *et al*., 1998; Morgan *et al*., 1998). Desde su descubrimiento, esta proteína se ha caracterizado también en *Xenopus* (Yasuhiko *et al*., 2001; 2006) y pez cebra (Otto *et al*., 2003; Simons *et al*., 2005), por lo que su presencia en vertebrados quedó confirmada. Sin embargo, se desconocía la existencia de esta proteína en ningún otro grupo no vertebrado. En esta tesis se han aportado evidencias que sustentan la presencia del gen *inversin* en los genomas de representantes de todos los principales grupos de metazoos (figura 18; apéndice 7). Además, se ha podido constatar la existencia de un ortólogo en el grupo hermano de los metazoos, los coanoflagelados. Estos datos apoyan un origen antiguo de este gen que se remonta a un periodo anterior a la aparición de los organismos pluricelulares. Sería necesario determinar la presencia de este gen en otros grupos de eucariotas como los amebozoos, rizarios, excavados y otros organismos unicelulares para determinar el momento concreto en el que se originó.

Estudios detallados en vertebrados y en cultivos celulares han demostrado que esta proteína está asociada al citoesqueleto y posee un papel relevante en el mantenimiento de la integridad de la arquitectura celular (Nürnberger *et al.*, 2004). Además, también se ha visto asociada a otros orgánulos subcelulares como cilios, huso mitótico, cuerpos medios del huso o *midbodies*, centriolos y centrosomas. Las células eucariotas se caracterizan, entre otras cosas, precisamente por tener una compleja estructura citoesquelética que les diferencia de los "procariotas". Es posible por tanto que esta estrecha relación entre Inversin y el citoesqueleto se remonte a los orígenes de las células eucariotas y que ésta haya sido una proteína necesaria para su correcto funcionamiento. Para testar esta hipótesis serán necesarios estudios que profundicen en el papel de esta proteína en la conformación del citoesqueleto en los organismos unicelulares.

En algunos grupos como los ctenóforos, artrópodos y nematodos, este gen parece haberse eliminado del genoma. Probablemente, otra/s proteína/s puede/n haber sustituido funcionalmente a Inversin en la historia evolutiva de estos grupos o incluso otros mecanismos diferentes que no impliquen a Inversin pueden ser responsables de mantener la integridad de su citoesqueleto celular y del resto de funciones de esta proteína.

Excepto coanoflagelados y poríferos que cuentan con dos parálogos, todas las especies aquí revisadas tienen un sólo ortólogo en sus genomas (figura 18). Sin embargo, se han descrito diferentes isoformas en *Mus* generadas por procesamiento alternativo (Nürnberger *et al.*, 2002; Yasuhiko *et al*., 2001, 2006). Estas isoformas tienen diferentes localizaciones subcelulares, interaccionan con diferentes proteínas y presuntamente tienen funciones diferentes (Nürnberger *et al.*, 2002). En esta tesis se ha constatado la presencia de al menos dos isoformas

en el gasterópodo *C. fornicata* (figura 19). Estas dos isoformas están activas en momentos y células diferentes durante el desarrollo embrionario de esta especie (figuras 23, 24 y 52), lo que indica que se comportan de manera independiente y que tienen potencialmente funciones distintas. Las diferencias más significativas en cuanto al patrón de activación espacio temporal entre las dos isoformas son patentes en los estadios anteriores a la gastrulación. En concreto, la isoforma larga tiene un origen materno, con lo que el mRNA se detecta desde el comienzo del desarrollo (figura 23). Sin embargo, la forma corta se activa en los embriones de 37 células (figura 24), cuando presumiblemente comienza la transcripción de RNAs no maternos.

En las primeras divisiones, la isoforma larga está en todos los blastómeros pero posteriormente se localiza exclusivamente en los macrómeros (figura 23). Por otra parte, la isoforma corta muestra un patrón mucho más dinámico y extendido (figura 24). Si bien cuando comienza la transcripción el mRNA está localizado en los blastómeros de los cuadrantes A, B y C siendo mayoritario en los cuadrantes A y B, inmediatamente después, la expresión va cambiando desapareciendo en los micrómeros del cuadrante A y activándose mayoritariamente en los del cuadrante C. El cuadrante C es responsable de la formación del lado derecho en embriones dextrógiros, mientras que A, B y D darán lugar a la zona izquierda, ventral y dorsal, respectivamente. Este patrón se mantiene hasta el comienzo de la gastrulación. Estos datos indican que la isoforma corta de *inversin* está principalmente activa en el lado derecho de la blástula y está completamente ausente del cuadrante D.

Cuando comienza la gastrulación, la isoforma larga se localiza en los macrómeros con la particularidad de que su expresión se ve reducida en el macrómero C (figura 23). Estos datos indican que el lado derecho de la blástula tiene unos niveles de mRNA de la isoforma larga de *inversin* inferiores al resto de cuadrantes del embrión durante la gastrulación. Según avanza la gastrulación, si bien ambas formas están activas exclusivamente en el blastoporo, la expresión sigue siendo asimétrica localizándose la forma larga en los cuadrantes A, B y D, mientras que la forma corta se detecta en los cuadrantes B y C (figuras 23 y 24). El blastoporo en los gasterópodos constituye un centro de señalización importante para la organización del embrión. Por tanto, la activación de este gen de manera restringida en determinados cuadrantes del blastoporo podría implicar diferencias estructurales o de señalización entre diferentes dominios de la gástrula.

Durante la organogénesis y el estadio de larva pre-veliger, la isoforma larga se expresa en el blastoporo y en células sensoriales muy concretas en la parte cefálica (figura 23). Sin embargo, ambas isoformas se expresan de manera asimétrica en el ectodermo del lado derecho (figuras 23 y 24).

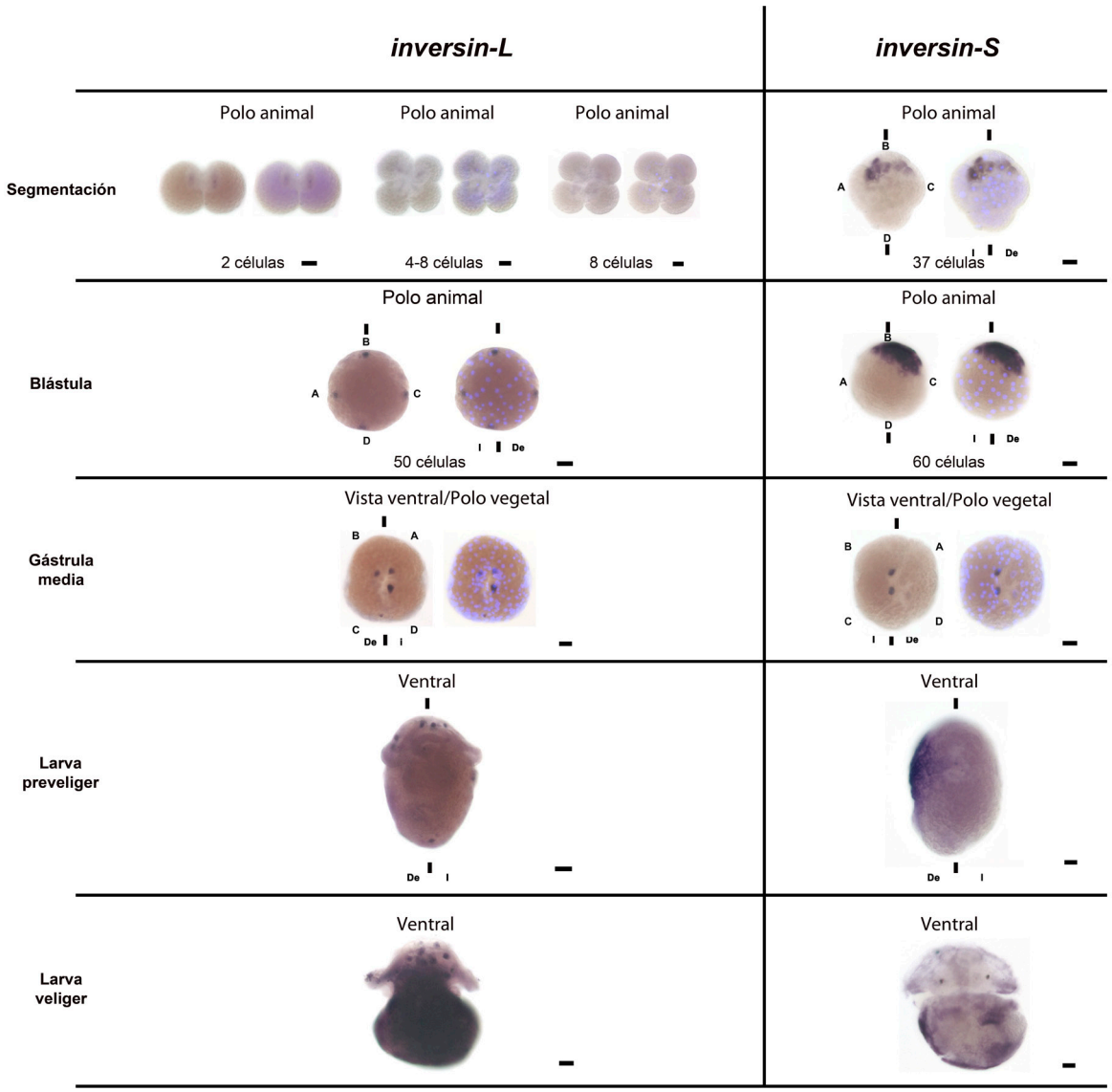


Figura 52. Las isoformas de *inversin* de *C. fornicata* se expresan de forma diferencial tanto espacial como temporalmente. Mientras que la isoforma larga es de origen materno y es detectable ya en las primeras divisiones, la isoforma corta sólo se detecta con la transcripción zigótica. En el estadio de blástula, la isoforma larga se expresa en los cuatro macrómeros (A, B, C y D), mientras que la isoforma corta se expresa en micrómeros de forma polarizada en los cuadrantes B y C principalmente; en este estadio las isoformas larga y corta se expresan al mismo tiempo en el macrómero B. En estadio de gástrula, las isoformas también colocalizan en el macrómero B, pero se diferencian en que el RNAm de *inversin-L* también se detecta en los macrómeros A y D, y el RNAm de la isoforma corta en C. En el estadio de larva preveliger, la isoforma larga se expresa en células muy discretas de la zona anterior, mientras que la corta se expresa de forma asimétrica en la zona anterior derecha. En las larvas veliger, la isoforma larga se expresa de forma discreta en la cabeza y de forma generalizada en el resto del cuerpo, mientras que la isoforma corta se expresa en zonas más discretas y de forma asimétrica en la masa visceral. La escala corresponde a 30 µm.

En resumen, los datos aquí presentados indican que existen diferencias significativas en la activación y regulación de las dos isoformas de mRNA de *inversin* entre el lado derecho e izquierdo del gasterópodo dextrógiro *C. fornicata*. Nunca antes se habían descrito diferencias en la activación del gen *inversin* con respecto al eje de simetría bilateral. Hasta el momento sólo se conocen datos de expresión en *Mus* y *Xenopus* (Mochizuki *et al*., 1998; Yasuhiko *et al*., 2006). En ambos casos se ha descrito una activación muy temprana de los mensajeros, materna en

Xenopus, y exactamente igual en ambos lados del cuerpo durante la embriogénesis. En base a estos datos, estos autores propusieron que Inversin es una proteína intracelular que controla el establecimiento de las AID a través de la regulación de la polaridad celular y no a través de una distribución asimétrica en el embrión. Si bien esto puede ser cierto para vertebrados, no parece ser el caso en gasterópodos puesto que hay una marcada diferenciación en los dominios de expresión de este gen con respecto al eje de simetría bilateral. Particularmente informativos parecen 1) la ausencia de expresión de la forma corta en el cuadrante D (el cuadrante D sirve de "organizador" y centro de señalización en las primeras etapas del desarrollo de gasterópodos), y 2) las diferencias en la expresión en diferentes cuadrantes del blastoporo que constituye otro centro de señalización relevante durante la gastrulación de gasterópodos.

En el presente trabajo sólo se ha detectado un mensajero del gen *inversin* en la especie de gasterópodo levógiro *B. glabrata*. Sería interesante determinar el patrón de activación de este gen en esta especie levógira para comprobar si el mRNA es de origen materno, si su expresión es asimétrica o centrada en algún cuadrante concreto y si termina siendo generalizada en la larva. Estos datos ayudarán a determinar, por ejemplo, si el gen *inversin* sigue algún patrón que pueda correlacionarse con la quiralidad de los gasterópodos.

Los resultados presentados en esta tesis muestran que la proteína Inversin no sólo tiene un origen antiguo en la historia evolutiva de los eucariotas, sino que su estructura molecular también parece haberse conservado estable a lo largo del tiempo. Los dominios presentes en esta proteína se encuentran consistentemente en los ortólogos de todos los representantes aquí estudiados (figura 18; apéndice 7). Estudios realizados principalmente en cultivos celulares han demostrado la interacción y potencial funcionalidad de algunos de estos dominios. Por ejemplo, se ha descrito la interacción de la proteína Inversin con Apc2, una subunidad del complejo promotor de la anafase (APC), complejo responsable de la ubiquitinación de los reguladores celulares de las transiciones mitóticas (Morgan *et al.*, 2002). Los dominios de destrucción D-Box presentes en Inversin son necesarios para la interacción Inversin-APC2 (Morgan *et al.*, 2002). Igualmente, los dominios D-Box se han demostrado necesarios para la degradación por vía del proteasoma de proteínas como Dishevelled (Simons *et al*., 2005; Lienkamp *et al*., 2012). Por su parte, el dominio Nineína parece dirigir la proteína hacia los cilios (Shiba *et al*., 2008). Finalmente, se ha visto una interacción de Inversin con Tubulina polimérica, seguramente a través de los dominios ankirina; puesto que se sabe que la Tubulina interacciona con repeticiones ankirina (Nürnberger *et al.*, 2004). Si bien todos estos estudios intentan aclarar la posible función de los

dominios de esta proteína durante la vida celular, hasta el momento se desconoce el mecanismo exacto por el cual esta proteína regula el establecimiento de asimetrías durante el desarrollo de vertebrados.

Uno de los dominios de la proteína Inversin implicado en generar un correcto plan corporal en *Xenopus* es el dominio IQ, un dominio de unión a la proteína Calmodulina dependiente de calcio (Yasuhiko *et al*., 2001, 2006). La ausencia de calcio ocasiona una conformación en la proteína Calmodulina que le permite unirse a Inversin a través del dominio IQ. Por otra parte, la presencia de calcio provoca un cambio conformacional en la Calmodulina que impide la unión de estas dos proteínas. Experimentos realizados en *Xenopus* han demostrado que la unión de Calmodulina a los dominios IQ es crítica para regular la actividad biológica de la proteína Inversin. La inyección en los blastómeros de embriones de *Xenopus* de mRNA de *inversin* de *Mus* o de *Xenopus* con los dominios IQ mutados ocasiona 1) una modificación en la localización de la activación de la ruta de Nodal, que en vez de activarse en el lado izquierdo como en los tipos silvestres se localiza en el lado derecho, de manera bilateral o se pierde en ambos lados, y 2) fenotipos de asimetrías azarosas o revertidas con respecto a la silvestre (Yasuhiko *et al*., 2001; 2006). En resumen, estos estudios muestran que los dominios IQ de Inversin son críticos para el correcto establecimiento de AID en *Xenopus.*

El estudio de los dominios aquí realizado muestra que todos los ortólogos de Inversin de eucariotas cuentan con dos dominios IQ excepto el ortólogo de la única especie levógira presente en este análisis, *B. glabrata*, que carece del dominio IQ2 (figura 21). Sería muy interesante determinar la secuencia de la proteína Inversin en otros organismos levógiros para comprobar si esta ausencia es una característica propia de este tipo de organización quiral o por el contrario responde a una particularidad de *Biomphalaria*.

Por otra parte, los experimentos funcionales realizados en este trabajo muestran que los dominios IQ de la proteína Inversin de gasterópodos pueden tener un papel en el establecimiento de AID en este grupo. Las inyecciones de mRNA mutantes para los dos dominios IQ de la isoforma larga de *inversin* de *C. fornicata* en huevos fecundados de esta especie tienen como resultado la formación de juveniles con fenotipos completamente simétricos (tabla 10). También las inyecciones de mRNA de *B. glabrata* de *inversin* silvestre (que carece de dominio IQ2) en huevos fecundados de *C. fornicata* ocasionan la aparición de fenotipos completamente simétricos en los juveniles (figura 29; tabla 10). Sin embargo, el porcentaje de individuos simétricos obtenidos en estos experimentos es muy reducido (1,6%; 4,5%). Esto podría deberse a la presencia de dos isoformas del gen *inversin* en la especie *C. fornicata*. Es posible que cuando se inyecte mRNA

mutante para los dos dominios IQ de la isoforma larga de *inversin* de *C. fornicata* en huevos fecundados de esta especie, el efecto se vea parcialmente compensado por la actuación de la isoforma corta. Por otra parte, el fenotipo causado por la inyección de mRNA de *B. glabrata* de *inversin* silvestre en huevos fecundados de *C. fornicata* pudiera igualmente ser compensado por la presencia de las dos isoformas propias de *C. fornicata*. Por tanto, para comprobar la hipótesis que se plantea referente a la compensación entre isoformas, sería necesario eliminar ambas isoformas de *inversin* en *C. fornicata* y evaluar su efecto en la aparición de simetrías.

Un aspecto a destacar de los experimentos funcionales de los dominios IQ llevados a cabo en *Xenopus* es la relevancia de la elección del blastómero inyectado con las formas mutantes (Yasuhiko *et al*., 2001, 2006). Solamente las inyecciones de formas mutantes en el blastómero derecho ocasionan AID revertidas, mientras que las inyecciones en el blastómero izquierdo no producen ningún efecto (Yasuhiko *et al*., 2001, 2006). Esto ha llevado a algunos autores a plantear la posibilidad de que haya diferencias en los niveles de concentración de calcio intracelular entre diferentes blastómeros, que podrían afectar de manera desigual a la actividad biológica de la proteína Inversin (Yasuhiko *et al.*, 2001, 2006). Las inyecciones realizadas en el transcurso de esta tesis se han llevado a cabo en embriones de una célula y por tanto, no se han registrado datos comparativos entre blastómeros. En las primeras divisiones de los embriones de *C. fornicata*, el mRNA de la isoforma larga es de origen materno y está presente en todas las células, por lo que sería interesante comprobar los niveles de concentración de calcio intracelular entre diferentes blastómeros en esos instantes. Si hubiera diferencias en los niveles de calcio intracelular entre diferentes blastómeros, la actividad de Inversin podría verse afectada en esos momentos iniciales. Sin embargo, en etapas posteriores *inversin* en gasterópodos sí que se expresa de manera asimétrica entre blastómeros (a diferencia de lo que ocurre en *Xenopus*). Quizá en esta etapa posterior, donde ya hay diferencias de partida en la distribución del mRNA, la relevancia de la concentración de calcio sea menor. En cualquier caso, se requieren más experimentos para testar estas hipótesis.

Varias evidencias aquí presentadas muestran que la proteína Inversin es una proteína relevante para el estudio de AID en gasterópodos y en otros metazoos. Además, hay diferentes datos que apuntan a la necesidad de seguir profundizando en el conocimiento de esta proteína. A continuación se enumeran algunos de ellos:

- El hecho de que una de las isoformas de *inversin* en *C. fornicata* sea materna puede indicar un papel relevante de este gen durante las primeras divisiones. La tercera división en

caracoles ha resultado ser crítica para el establecimiento de la quiralidad en los adultos. En concreto, el control genético sobre la deformación en la estructura celular (que deja de ser circular) y la orientación de los husos en los cuatro macrómeros durante su división a ocho células, es el que determina la configuración quiral de los blastómeros y por consiguiente, la de los individuos adultos (figura 10). El hecho de que Inversin module la red de Actina durante la mitosis y que su falta de función ocasione una reorganización del citoesqueleto de Actina que haga que se pierda la circularidad celular y que haya modificaciones en la localización de los husos (Werner *et al*., 2013) abre la atractiva posibilidad de que Inversin pueda estar relacionada con este evento en la tercera división de los gasterópodos.

- En relación con el punto anterior, la importancia de la tercera división en embriones de gasterópodos para determinar la quiralidad lleva a preguntarse si hay diferencias en la expresión del gen *inversin* o en la localización subcelular de la proteína entre formas levógiras y dextrógiras durante la tercera división.

- Los datos presentados en esta tesis aportan la primera evidencia de que *inversin* está activo de manera asimétrica en embriones en desarrollo. La expresión está localizada en el lado derecho en las larvas de la especie dextrógira *C. fornicata*. Sería interesante comprobar si también hay una expresión asimétrica en las larvas de la especie levógira *B. glabrata*, y en caso afirmativo, si se corresponde con el lado opuesto, el izquierdo.

- En vertebrados, *inversin* se expresa antes de que se active la ruta de señalización de Nodal y la falta de función de Inversin ocasiona cambios en la localización de la activación de la ruta (Lowe *et al*., 1996; Meno *et al*., 1996; Yasuhiko *et al*., 2001, 2006). En gasterópodos, también *inversin* se expresa antes de que se active la ruta de señalización de Nodal y por tanto, sería relevante comprobar si la localización de la activación de la ruta también es dependiente de Inversin.

- La especie levógira *B. glabrata* ha resultado ser hasta el momento la única especie con un ortólogo de *inversin* que carece de dominio IQ2. Sería interesante profundizar en el conocimiento de este hecho, comprobando su potencial relevancia para el establecimiento de la quiralidad y de AID.

5.2. El ligando Nodal en *C. fornicata*

El ligando Nodal es miembro de la familia de proteínas TGF-β, con un papel fundamental en el establecimiento de la comunicación celular para regular la diferenciación celular, el desarrollo y el crecimiento de los tejidos. Nodal puede actuar como un morfógeno y la correcta activación, tanto suya como de la cascada génica que de él depende, es necesaria en momentos clave durante el desarrollo para el establecimiento de los ejes corporales, la gastrulación y la diferenciación neuronal (Shen, 2007; Schier, 2009). Debido a su carácter multifuncional y las graves consecuencias que su falta de función ocasiona en el desarrollo de vertebrados, este gen ha sido estudiado en detalle durante las últimas décadas principalmente en este grupo animal.

En 2009 Grande y Patel documentaron la presencia de este gen en gasterópodos y demostraron su implicación en el establecimiento de las AID. En concreto, en ese estudio se analizaron dos especies de gasterópodos: *Lottia gigantea* (especie dextrógira) y *B. glabrata* (levógira). El análisis comparado de los resultados de este trabajo con la información existente en deuteróstomos resultó de utilidad para entender cuán conservada está la ruta de Nodal en organismos bilaterales. Sin embargo, aún se desconocen cuáles son los mecanismos que regulan la ruta en gasterópodos así como los efectos ocasionados por la perturbación del funcionamiento del ligando Nodal durante su desarrollo. Con la intención de responder a esta pregunta, en esta tesis se han realizado diferentes aproximaciones utilizando la especie *C. fornicata* como sistema modelo. Hasta el momento, no ha sido posible desarrollar una técnica para la modificación de la expresión de genes específicos en las especies de gasterópodos *L. gigantea* y *B. glabrata* y precisamente por esa razón, los estudios aquí realizados se han centrado en la especie *C. fornicata*, en la cual es posible realizar inyecciones de diferentes reactivos en huevos fecundados, modificar la expresión de genes específicos y seguir sus efectos durante el desarrollo embrionario.

En primer lugar los resultados aquí obtenidos indican la presencia de dos parálogos del gen *nodal* en la especie *C. fornicata* (figura 30; apéndice 13). Hasta el momento, todas las especies de gasterópodos analizadas han resultado tener una sola copia de este gen, a excepción de *Patella vulgata* que cuenta con dos parálogos (figura 30; apéndice 13). Sin embargo, basándonos en nuestros análisis de ortología, estas duplicaciones en *Patella* y *Crepidula* parecen proceder de eventos de duplicación independientes. Un mayor muestreo de este gen en otros gasterópodos servirá para corroborar o refutar esta hipótesis.

Los eventos de duplicación de genes tienen gran relevancia por la posible subfuncionalización de los parálogos o incluso la adquisición de nuevas funciones por parte de uno de ellos. El análisis del patrón de activación de los dos parálogos de *nodal* presentes en *Crepidula* muestra que ambas copias se activan en momentos y lugares diferentes del embrión, lo que apoyaría una independencia funcional entre ambas copias (figuras 33 y 34). La expresión de *nodal1* comienza de manera tardía durante la gastrulación y se concentra principalmente a ambos lados del estomodeo y en la parte anterior, siempre de manera simétrica (figura 33). En las larvas se mantienen estos dominios y además se detecta expresión simétrica en la zona de la concha (figura 33). Sólo el parálogo *nodal2* se activa de manera asimétrica desde etapas muy tempranas (unas 40 células) hasta el estadio larvario (figura 34). El análisis de la expresión de los dos parálogos descritos en *Patella* también muestra una independencia de activación temporal y espacial entre ambas copias (Kenny, 2014). Solamente uno de los parálogos descritos en *Patella* muestra expresión asimétrica, mientras que el otro parálogo tiene un patrón de activación único, con expresión materna que desaparece en el estadio de 37 células y vuelve a aparecer simétricamente en la larva (Kenny, 2014). Así, tanto *Patella* como *Crepidula* poseen cada una un parálogo cuya expresión es asimétrica y además cada una posee asimismo otro parálogo con patrones de expresión no congruentes entre sí. Estos datos apoyarían también la hipótesis anteriormente planteada sobre la existencia de dos eventos de duplicación independientes de *nodal* en gasterópodos: por un lado en la historia evolutiva de *Crepidula* y por otro lado en la línea de *Patella*. En resumen, los datos aquí presentados indican que *nodal2* es el gen con un papel más relevante en el establecimiento de AID en *Crepidula*.

La expresión de *nodal2* en *Crepidula* es asimétrica desde el comienzo, lo cual concuerda con resultados previos obtenidos en otros gasterópodos (Grande & Patel, 2009) y braquiópodos (Grande *et al*., 2014; Martín-Durán *et al.*, 2017). Sin embargo, en vertebrados y otros deuteróstomos, la expresión de *nodal* comienza siendo simétrica para poco después, gracias a su propia autorregulación, ser asimétrica. Por tanto, parece haber claras diferencias en cuanto al control del comienzo de la expresión entre diferentes grupos de organismos bilaterales, reforzando la idea de que diferentes mecanismos regulan estas rutas de señalización asimétricas. Otro aspecto a destacar es que la expresión de *nodal2* en *Crepidula* se restringe a una sola capa germinal, el ectodermo, al igual que ocurre en otros gasterópodos (Grande & Patel, 2009). Los dominios de expresión de *nodal* en Bilateria parecen haber divergido enormemente, puesto que *nodal* se expresa en el mesodermo en braquiópodos (Grande *et al*., 2014; Martín-Durán *et al*., 2017), en el endomesodermo en vertebrados (Hamada *et al*., 2002) y en diferentes

combinaciones de las tres capas germinales en otros deuteróstomos (Duboc *et al*., 2010; Ohguro *et al.*, 2011; Bessodes *et al*., 2012; Morokuma *et al*., 2002; Yu *et al*., 2002; Wlizla, 2011). Estos datos sugieren que la diversificación de los dominios de expresión de *nodal* es congruente con la divergencia de sus funciones en organismos bilaterales, aunque en gasterópodos parece haber una consistencia entre especies en cuanto a su activación espacial y temporal. De hecho, en *Crepidula* la expresión es asimétrica, pero concentrada en el lado derecho, de acuerdo con lo descrito para otros gasterópodos dextrógiros (Grande & Patel, 2009).

En esta tesis se han realizado por primera vez experimentos de manipulación de la expresión del gen *nodal* en organismos no deuteróstomos. En concreto, se ha forzado la sobreexpresión del gen *nodal* desde el comienzo del desarrollo embrionario en el gasterópodo *Crepidula*. Como consecuencia de esta sobreexpresión, los juveniles resultantes muestran patrones de simetría tanto externa, para los que no se detecta ningún crecimiento diferencial de la concha, como interna, con simetría en la masa visceral y gónadas (figura 37; tabla12). La expresión asimétrica de *nodal* en vertebrados se restringe al lado izquierdo y la manipulación y consiguiente activación de la ruta en el lado derecho conlleva la aparición de morfologías típicamente de "tipo izquierda" en el lado derecho (isomerismo de lado izquierdo) (Hamada *et al.*, 2002). Del mismo modo, en equinodermos la expresión normal de *nodal* se reduce al lado derecho y la expresión ectópica en el lado izquierdo conlleva un desarrollo "tipo derecha" en el lado izquierdo (isomerismo de lado derecho)(Duboc *et al*., 2005). Por tanto, estos datos indican que la activación de *nodal* ocasiona una modificación del estado basal del desarrollo; en otras palabras, que *nodal* activamente modifica la simetría que existiría por defecto. Consistentemente, los resultados obtenidos en esta tesis en gasterópodos muestran que una activación ectópica de la ruta de Nodal en el lado donde normalmente no estaría activa ocasiona que ambos lados se comporten de la misma manera y que se pierda la diferenciación bilateral de estructuras. Se ha propuesto que la actividad asimétrica de la ruta de Nodal durante el desarrollo de los gasterópodos conlleva una formación de células productoras de concha diferencial a ambos lados del embrión, lo que ocasionaría diferentes ritmos de producción de concha a ambos lados y finalmente un enrollamiento de la concha debido a este crecimiento desigual (Grande & Patel, 2009). Por tanto, una compensación y unificación de la expresión de *nodal* a ambos lados del embrión en desarrollo conllevaría un equivalente ritmo de producción de concha a ambos lados y finalmente a la formación de una concha completamente simétrica, lo que es consistente con los resultados aquí presentados.

Con el fin de inhibir la actuación de la proteína Nodal en gasterópodos, en esta tesis se ha generado un mutante dominante negativo para el gen *nodal* a través de la modificación del sitio de corte para el procesamiento del péptido. Los ligandos tipo TGF-β son sintetizados como pro-péptidos y posteriormente deben ser procesados mediante corte (en un sitio con dominio RXXR), dejando una parte pro-dominio en la parte amino-terminal y otra parte que constituye la forma madura del péptido en el carboxilo-terminal. El corte y la dimerización de los ligandos son necesarios para que éstos puedan unirse a los correspondientes receptores (Shen, 2007). Los mutantes para los que se ha modificado el sitio de corte en ligandos TGF-β han resultado ser de utilidad en experimentos de pérdida de función, puesto que su presencia ocasiona una competencia con las formas silvestres que bloquea la activación de díMeros funcionales (Hawley *et al*., 1995; Joseph & Melton, 1998). Sin embargo, los resultados obtenidos en estudios con mutantes para *nodal* realizados en *Xenopus* precisamente contra el sitio de corte, han sembrado dudas con respecto a su utilidad para inhibir la función de Nodal en este grupo. Los resultados en *Xenopus* prueban que incluso los péptidos de Nodal sin procesar pueden generar cierta señalización (Eimon & Harland, 2002). En nuestros análisis, no se ha detectado ningún fenotipo diferente al silvestre en embriones inyectados con formas dominante negativo para el gen *nodal* con los sitios de corte modificados (tabla 12); el desarrollo de estos mutantes es completamente normal. Estos resultados sugieren que este mutante de *nodal* no es capaz de modificar la señalización de la ruta en gasterópodos al igual que ocurre en *Xenopus*. Por tanto, esta diferencia de comportamiento de Nodal con respecto a otros ligandos TGF-β que sí son inhibidos mediante esta técnica parece ser una particularidad de este ligando más que del grupo animal del que se trate.

Otra aproximación para interferir y bloquear la ruta de Nodal durante el desarrollo de *Crepidula* llevada a cabo en esta tesis se basó en la utilización de la proteína Activin A recombinante humana. Activin es un ligando TGF-β estrechamente relacionado con Nodal y que es capaz de sustituirlo funcionalmente *in vitro*. Activin está implicada en diversos procesos biológicos incluyendo inflamación, fibrosis, función del sistema reproductor y hematopoyesis y también tiene un papel relevante durante el desarrollo embrionario en morfogénesis y desarrollo del sistema nervioso de vertebrados (Namwanje & Brown, 2016). Tanto Nodal como Activin señalizan por el mismo tipo de receptor tipo I (TGF-β R1 también llamado receptor Alk4-5-7) y por el mismo tipo de receptor tipo II (ActivinR2). Cuando un ligando TGF-β activo se une a los dos receptores, el receptor I es fosforilado y activado (Massagué, 1998). La activación del receptor permite la endocitosis del complejo y una vez en el citoplasma el receptor I fosforila a

la proteína Smad2 para continuar la señalización de la ruta (revisado en Shen, 2007; Massagué, 2012; Xu *et al*., 2012). Aunque Nodal y Activin señalicen a través de los mismos receptores, los mecanismos que median en la unión del ligando y el receptor, así como el uso de proteínas correceptoras, varían significativamente entre ambos ligandos. Activin se une al receptor tipo II con gran afinidad (Attisano *et al*., 1992) para posteriormente reclutar al receptor tipo I y formar el complejo activo (Carcamo *et al*., 1994; Dijke *et al*., 1994). Por el contrario, Nodal carece de esa afinidad por el receptor tipo II o el de tipo I y requiere del co-receptor EGF-CFC para que esa unión se produzca correctamente (Schier, 2003; Shen, 2007).

Hasta el momento no se ha identificado ningún ortólogo de *activin* para organismos espiralados sino varios parálogos tipo *activin,* producto de diferentes eventos de duplicación y divergencia del gen ancestral (Kenny *et al*., 2014). Se han analizado algunos de estos parálogos y se ha comprobado que no están activos durante el desarrollo de los gasterópodos (Grande & Patel, 2009). Por tanto, las evidencias presentadas hasta el momento apoyan la ausencia de un papel de los parálogos de la proteína Activin durante el desarrollo embrionario de los gasterópodos. En estas condiciones, se espera que el uso de la proteína Activin A recombinante humana en los experimentos realizados en esta tesis durante el crecimiento de embriones de

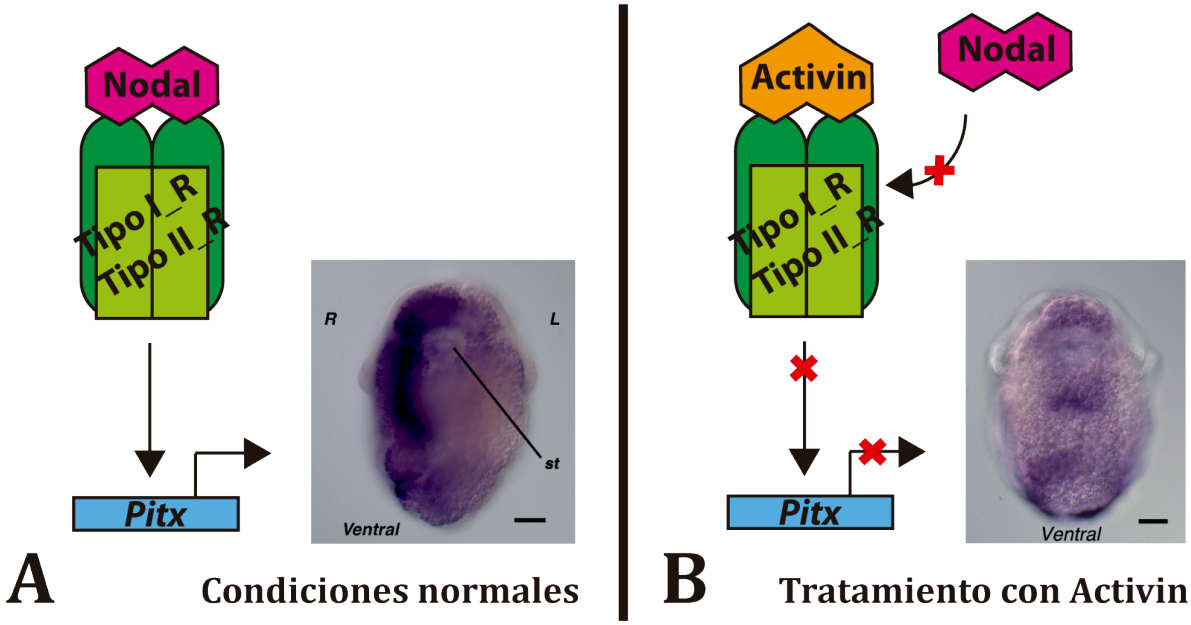


Figura 53. La adición de Activin provoca la inactivación de la ruta de Nodal presumiblemente por competencia por los receptores con el ligando Nodal. A) La unión del ligando Nodal al clúster de receptores provoca una cascada de señalización que desemboca en la activación de la expresión del gen *Pitx*. En las larvas preveliger, *Pitx* se expresa de manera simétrica en la región anterior y asimétrica en el lado derecho; B) la presencia de Activin evitaría, por competencia, la unión de Nodal con sus receptores, lo que conllevaría una inactivación de *Pitx* exclusivamente en su dominio de expresión asimétrico (el único dominio dependiente de la ruta de Nodal), mientras que sus dominios de activación simétricos se verían inalterados. La escala corresponde a 30 µm.

Crepidula interfiera exclusivamente con la ruta de Nodal, puesto que ambos ligandos, Nodal y Activin A recombinante, deben competir por los mismos receptores.

El uso de la proteína Activin A recombinante humana durante 2 días en embriones de *Crepidula* de poco más de 30 horas desde la fertilización ocasiona un 80% de juveniles con fenotipo simétrico (figura 38; tabla 13). La activación de *nodal* en *Crepidula* comienza en torno a las 30 horas después de la fertilización. Por tanto, la presencia inducida del ligando Activin durante el periodo de actuación de la proteína Nodal en *Crepidula* ocasiona fenotipos simétricos, presuntamente como consecuencia de la competencia entre los ligandos Activin y Nodal (figura 53).

Con el fin de comprobar que la ruta de Nodal en *Crepidula* estaba siendo bloqueada por la presencia de la proteína Activin, se analizó el patrón de activación de un gen diana de la ruta de Nodal, *Pitx*, entre los embriones tratados con Activin y los embriones control. El gen *Pitx* es una diana del gen *nodal* en gasterópodos al igual que ocurre en otros organismos bilaterales (Grande & Patel, 2009; Hamada *et al*., 2002; Duboc *et al*., 2010; Ohguro *et al*., 2011; Bessodes *et al*., 2012; Morokuma *et al*., 2002; Yu *et al*., 2002; Wlizla, 2011; Yoshida & Saiga, 2008). *Pitx* se expresa de manera asimétrica en *Crepidula* un poco después de la activación de *nodal* (Figura 39), lo que concuerda con datos previos obtenidos en otros gasterópodos (Grande & Patel, 2009). Además de la expresión asimétrica, *Pitx* también tiene dominios de activación simétricos en *Crepidula* (Figura 39), en otros gasterópodos y en vertebrados, lo que ha llevado a sugerir la presencia de zonas potenciadoras o *enhancers* independientes que regulan cada uno de estos dominios (Christiaen *et al*., 2005, 2007; Grande & Patel, 2009). A diferencia de lo observado en los embriones control, en los embriones tratados con Activin A recombinante humana no se detectó ninguna activación asimétrica de *Pitx* (figura 40). Sin embargo, el dominio de activación simétrico permaneció igual al de los embriones control (figura 40). Estos datos indican que el tratamiento con la proteína Activin A recombinante humana es eficaz para bloquear la ruta de activación de Nodal en *Crepidula* y que dicho bloqueo conlleva la pérdida de AID.

Los resultados aquí expuestos muestran que la proteína Nodal es crítica para establecer AID en gasterópodos. Sin embargo, a pesar de los avances realizados aún quedan muchas preguntas abiertas, como por ejemplo:

- ¿Cómo se inicia la activación asimétrica de *nodal* en gasterópodos? Como se ha indicado anteriormente, en deuteróstomos la activación de *nodal* comienza siendo simétrica y es precisamente su autorregulación la que desencadena su activación asimétrica. Sin

embargo, la activación en gasterópodos es asimétrica desde el primer momento y se desconoce qué factor da origen a ese patrón.

- ¿Qué genes son regulados por Nodal? El único gen diana de Nodal que se ha descrito en gasterópodos es *Pitx*, pero se desconoce qué otros genes pueden estar implicados en la organogénesis y la diferenciación asimétrica.

- Cada vez hay más evidencias que apoyan la conexión que hay entre la ruta de señalización de los TGF-β con otras rutas, como por ejemplo Wnt y Notch (Luo *et al*., 2017). Ya que estas rutas están profundamente implicadas en la definición de ejes corporales en organismos bilaterales, sería interesante comprobar si hay algún tipo de conexión entre todas ellas necesaria para el correcto establecimiento de AID.

5.3. La proteína EGF-CFC en Gastropoda

La familia de los EGF-CFC comprende un grupo de pequeñas glicoproteínas extracelulares asociadas a la membrana que se caracterizan por poseer tres dominios conservados: un dominio tipo EGF en el extremo amino-terminal y un domino CFC rico en cisteínas, seguido de una zona hidrofóbica en su extremo carboxilo-terminal, el dominio GPI. Estudios genéticos detallados demuestran que estas proteínas tienen gran variedad de funciones. Sin embargo, en esta tesis se abordaron en detalle debido a la existencia de una fuerte relación entre las proteínas de la familia EGF-CFC y la ruta de Nodal (Gritsman *et al*., 1999; revisado en Gray & Vale, 2012). Varias evidencias demuestran que Nodal interactúa con las proteínas de la familia EGF-CFC para unirse a sus receptores y provocar así la transducción de la señal (Blanchet *et al*., 2008a, 2008b; Kelber *et al*., 2008; Wechselberger *et al*., 2005).

Recientes estudios centrados en comprender el origen y la evolución molecular de estas proteínas identificaron varios miembros de esta familia exclusivamente en vertebrados y en otros deuteróstomos, lo que llevó a algunos autores a concluir que esta familia de proteínas se originó en el antecesor de los deuteróstomos (Ravisankar *et al*., 2011). Sin embargo, los datos presentados en esta tesis apoyan la existencia de proteínas EGF-CFC en gasterópodos, anélidos y braquiópodos (figura 41). Además, durante el desarrollo de esta tesis se realizaron búsquedas activas en los genomas secuenciados de varias especies de cnidarios y ctenóforos, así como en los de otros bilaterados como *Xenoturbella*, Acoelomorpha, Ecdysozoa y Platyhelminthes, pero

hasta el momento no se ha podido identificar ningún ortólogo en estos grupos. A la vista de estos datos se puede afirmar que la familia de proteínas EGF-CFC no es una novedad evolutiva de los deuteróstomos, sino que al menos ya el antecesor común de deuteróstomos, ecdisozoos y espiralados presentaba algún miembro de esta familia.

El análisis filogenético para reconstruir la historia evolutiva de esta familia de proteínas muestra la presencia de varios genes parálogos dentro de líneas evolutivas específicas, cuyas relaciones son difíciles de establecer (figura 41). Por otro lado, a excepción de *Patella*, todos los organismos no vertebrados aquí señalados parecen tener un solo ortólogo. Estos resultados pueden explicarse en base a, al menos, dos hipótesis plausibles aunque excluyentes entre sí: 1) se han producido varios eventos recientes de duplicación y divergencia de estas proteínas en líneas evolutivas concretas (por ejemplo en mamíferos, *Xenopus* y *Patella*), y 2) se ha producido una duplicación del gen ancestral en el antecesor de todos los deuteróstomos y posteriormente una de las copias se ha perdido en algunos grupos como en los peces, *Gallus* o marsupiales. Entender e inferir la historia evolutiva de esta familia de proteínas requerirá un mayor muestreo de estos genes en más grupos de metazoos. Cabe destacar, por el interés de resultados presentados en esta tesis y comentados más adelante, que tanto *C. fornicata* como el pez cebra *D. rerio* cuentan con un solo ortólogo (en el pez cebra este gen es conocido como *one-eyed pinhead*, *oep*).

Los genes *egf-cfc* han sido caracterizados en detalle en *Mus* (Ding *et al*., 1998), *Homo* (Bramford *et al*., 2000), *Xenopus* (Dorey & Hill, 2006), *Gallus* (Chu *et al*., 2005), *Danio* (Zhang *et al.*, 1998; Gritsman *et al*., 1999) y en hemicordados (Röttinger *et al.*, 2015). En mamíferos hay dos genes de esta familia, *cripto* y *cryptic,* que difieren tanto en sus dominios de activación durante el desarrollo como en las funciones que desempeñan. Los análisis de la activación espacio-temporal de *cripto* en *Mus* muestran que este gen está activo en el epiblasto justo antes de la gastrulación (Ding *et al*., 1998). Durante la gastrulación, *cripto* se expresa en la línea primitiva, células del mesodermo, el nodo y el endomesodermo axial (Chu *et al*., 2005; Ding *et al*., 1998). Finalmente, *cripto* se expresa en diferentes zonas del corazón en desarrollo y posteriormente desaparece (Johnson *et al*., 1994). Se ha demostrado a partir de estudios genéticos que *cripto* se requiere para la formación de las capas germinales y el correcto establecimiento del eje anterior-posterior (Ding *et al*., 1998; Yan *et al*., 1999). Por su parte, *cryptic* se expresa de manera simétrica en el mesodermo lateral , el nodo, notocorda y la placa del suelo (Shen *et al.*, 1997) y controla la determinación de AID (Gaio *et al*., 1999). En concreto, los mutantes de *cryptic* muestran una organización azarosa de la región abdominal, isomerismo del pulmón derecho, heterotaxia vascular y defectos en la disposición del corazón (Gaio *et al*., 1999, Yan

et al., 1999). Cabe destacar que los mutantes de *cryptic* mantienen la expresión asimétrica de *nodal* en el nodo, aunque se pierde su expresión en un momento posterior en el mesodermo lateral. Este hecho sugiere que existe una activación de la ruta en la región del nodo en ausencia de Cryptic, pero que este gen es necesario para la correcta activación de la ruta de Nodal en etapas posteriores. Por otra parte, en el pez cebra, *oep* es de origen materno y su expresión zigótica comienza durante la gastrulación; en etapas posteriores, se expresa en el prosencéfalo, mesodermo lateral y notocorda. Los mutantes zigóticos para *oep* muestran básicamente una suma de los fenotipos observados en los mutantes de *cripto* y *cryptic* en *Mus*: defectos en el establecimiento de los ejes izquierda-derecha y antero-posterior, deficiencia en los movimientos propios de la gastrulación, ciclopia (la presencia de un único ojo), pérdida de la placa precordal, del neuroectodermo ventral y del mesodermo de la cabeza y del tronco (Schier *et al*., 1997; Strähle *et al*., 1997; Gritsman *et al*., 1999; Zhang *et al*., 1998). En resumen, en vertebrados, los genes *egf-cfc* se expresan principalmente en el punto de gastrulación y formación de capas germinales, y están profundamente involucrados en la especificación del mesodermo y en el establecimiento de los ejes corporales.

En esta tesis se ha analizado el patrón de activación del ortólogo de *egf-cfc* en la especie *C. fornicata* con el fin de identificar la ventana espacio-temporal en la que este gen está activo. Se detecta mRNA desde las primeras divisiones durante los estadios de blástula en todos los blastómeros, y posteriormente esta expresión se incrementa y restringe al labio dorsal durante la gastrulación. En las larvas, la señal se mantiene en la región del labio dorsal (ahora ya estomodeo), pero asimismo aparece en la región que comprende el velo, siempre de forma simétrica (figura 45). Además, se observan señales muy intensas, pero poco consistentes entre individuos, en varias células discretas a ambos lados del cuerpo, en la zona posterior de la larva y en la zona dorsal cerca de la concha (Figura 45). Como se ha indicado anteriormente, el labio dorsal es un centro de señalización de gran relevancia durante la gastrulación de los gasterópodos. Por tanto, estos resultados muestran que el gen *egf-cfc* pudiera tener un papel destacado durante la gastrulación y la especificación de las capas germinales en gasterópodos, al igual que ocurre en vertebrados. Sin embargo, cabe destacar que cuando se realizaron los controles negativos (utilizando sondas en el sentido de transcripción del RNA o sonda sentido) se obtuvo un patrón muy parecido al observado en los experimentos (utilizando sondas diseñadas en el sentido complementario al de transcripción del RNA o sonda antisentido) (Figura 45). Con el fin de comprobar este resultado, se sintetizaron varias sondas usando diferentes zonas del gen *egf-cfc* y se repitieron las reacciones variando las condiciones de hibridación y los tiempos de reacción.

En todos los casos en que se obtuvo señal, el resultado siempre fue el descrito anteriormente, confirmando la detección de señales recurrentes de las sondas sentido y antisentido. Una comparación exhaustiva de ambos patrones muestra que las señales obtenidas con la sonda sentido tienen dominios un poco más extendidos que los obtenidos con la sonda antisentido (Figura 45). Estos resultados indican que los transcritos sentido y antisentido del gen *egf-cfc* están activos durante todo el desarrollo embrionario de *C. fornicata*.

Los transcritos antisentido naturales (del inglés, *natural antisense transcripts*, NATs) son moléculas de RNA que complementan al mRNA sentido, que pueden encontrarse de manera endógena y que están activos tanto en eucariotas como en "procariotas" (Jin *et al*., 2008). Se ha demostrado que hasta el 72% de los transcritos tienen un RNA antisentido en los transcriptomas de ratón y humanos (Beiter *et al*., 2009; Werner *et al*., 2009). Los NATs controlan diferentes procesos biológicos como por ejemplo la estabilidad del mRNA, el procesamiento alternativo, la regulación de la traducción y la modificación de la cromatina, entre otros (revisado en Su *et al*., 2010 y Sun *et al.*, 2017). Los NATs regulan al transcrito sentido tanto mediante su unión e interacción directa como mediante la codificación de proteínas que se coexpresan con el RNA sentido y controlan su traducción (Dolnick, 1997). Debido a la complementariedad de las bases, los NATs son considerados unos reguladores de gran especificidad y se ha observado que son capaces tanto de estimular como de inhibir la expresión de los transcritos sentido (Katayama *et al*., 2005; Beiter *et al*., 2009). Asimismo, se ha comprobado que estos reguladores están activos e influyen en el desarrollo embrionario de algunas especies de vertebrados (Shi *et al*., 2017; Piatek *et al*., 2017). En base a los resultados aquí obtenidos, cabe la posibilidad de que el transcrito sentido del gen *egf-cfc* en *C. fornicata* esté siendo regulado por el transcrito antisentido. Se ha estudiado la posibilidad de que el RNA antisentido del gen *egf-cfc* en *Crepidula* pudiera codificar para alguna proteína, primeramente traduciendo el transcrito y posteriormente buscando posibles candidatos en la base de datos del NCBI, pero no se ha detectado ninguno. Esto podría indicar que el RNA antisentido regula al transcrito sentido a través de una interacción directa. Sin embargo, esta hipótesis deberá ser testada con diferentes aproximaciones en el futuro para comprobar su veracidad.

Además del papel fisiológico de los genes *egf-cfc* durante el desarrollo embrionario, estos genes también se han visto implicados en procesos tumorales durante la vida de vertebrados adultos cuando se activan de manera inapropiada (revisado en Gray & Vale, 2012). Por ejemplo, el gen *cripto* se expresa a altos niveles en diferentes tejidos tumorales en humanos a diferencia de lo que ocurre en los tejidos normales. Estos resultados lo han situado como

una diana terapéutica potencial para el control del crecimiento tumoral (revisado en Bianco & Salomon, 2010). Una de las estrategias terapéuticas utilizadas ha sido precisamente el uso de oligonucleóticos antisentido que han sido capaces de afectar la expresión de *cripto* de manera muy eficaz en diferentes tejidos cancerígenos en humanos (Ciardiello *et al*., 1994; De Luca *et al*., 1999; De Luca *et al*., 2000). Por tanto, estos datos indican que el RNA antisentido es capaz de regular al transcrito sentido de genes *egf-cfc* en consonancia con lo comentado anteriormente.

Otra aproximación terapéutica utilizada para neutralizar la acción de estas proteínas EGF-CFC ha sido generar diversos anticuerpos específicos contra ellas en diferentes especies de vertebrados. En concreto, se ha descrito un anticuerpo contra Cripto en *Mus* que es capaz de reconocer la proteína y que colocaliza exactamente con el mRNA (Minchiotti *et al*., 2000). Con el fin de detectar de manera concreta la presencia de la proteína EGF-CFC en *C. fornicata* se diseñó un anticuerpo específico durante el desarrollo de esta tesis, ya que los anticuerpos comerciales disponibles no mostraron ninguna tinción en esta especie de gasterópodo en los experimentos aquí realizados. Los resultados de la tinción de embriones completos con el anticuerpo específico de EGF-CFC de *C. fornicata* muestran que la proteína EGF-CFC se localiza en las membranas de las células del labio dorsal durante la gastrulación. Posteriormente, en las larvas, la señal se mantiene en el estomodeo, aunque se extiende por la parte anterior y a ambos lados de manera simétrica (figura 47). Estos datos refuerzan la hipótesis anteriormente planteada sobre la relevancia de la proteína EGF-CFC en la gastrulación de los gasterópodos.

Las proteínas EGF-CFC son moduladores multifuncionales debido principalmente a su estructura molecular y su composición aminoacídica. Estas proteínas tienen dominios con distintos sitios de reconocimiento que permiten la interacción con diferentes ligandos y receptores (Minchioti *et al*., 2002; Strizzi *et al*., 2005; revisado en Klauzinska *et al*., 2012). Las proteínas EGF-CFC son muy versátiles, siendo por ejemplo capaces de bloquear la señalización de ligandos TGF-β como Activin (Kelber *et al*., 2008; Gray *et al*., 2003; Gray *et al*., 2006) o potenciar la señalización de otros como BMP-4 (Aykul *et al*., 2017) o Nodal (Gritsman *et al*., 1999; Salomon *et al*., 1999; Reissmann *et al*., 2001; Minchiotti *et al*., 2001; Yan *et al*., 2002). Estos resultados aparentemente contradictorios se deben a que las proteínas EGF-CFC pueden actuar no sólo como proteínas de membrana sino también como proteínas solubles si se produce un corte en su dominio GPI (Watanabe *et al*., 2007, 2008). Como proteínas solubles son capaces de inhibir rutas de señalización como las activadas por BMP-4 o Activin mediante el bloqueo de la unión ligando-receptor (Aykul *et al*., 2017), y como proteínas de membrana pueden favorecer la localización de los ligandos en vesículas endocíticas enriquecidas en receptores y potenciar

por tanto la señalización (Blanchet *et al*., 2008a, 2008b; Aykul *et al*., 2017). En resumen, tanto estudios bioquímicos como genéticos han destacado la complejidad funcional de esta aún enigmática familia de pequeñas proteínas. Los resultados de esta tesis muestran que el dominio GPI está presente en todos los ortólogos encontrados (apéndice 10) con lo que estas proteínas podrían potencialmente funcionar como proteínas de membrana o como proteínas solubles.

En lo que respecta a la ruta de Nodal, una caracterización bioquímica *in vitro* mostró que el dominio CFC media en la interacción proteína-proteína de Cripto con el Receptor I de la ruta TGF-β (Alk4) (Calvanese *et al*., 2009). En concreto, se ha descrito que la histidina de la posición 120 de Cripto (H120) y el triptófano de la posición 123 de Cripto (W123), ambas en humanos, son los responsables del reconocimiento en la interacción con el receptor Alk4 (Foley *et al*., 2003; Marasco *et al*., 2006). Los análisis realizados en esta tesis muestran que estas dos posiciones están conservadas en todos los ortólogos EGF-CFC encontrados (excepto W123 para el anélido *Spirobranchus*) y por tanto, todos ellos podrían unirse potencialmente al Receptor I de la ruta TGF-β (figura 54).

Por otra parte, Cripto se une a Nodal gracias al dominio tipo EGF (Bianco *et al*., 2002; Yan *et al*., 2002; Yeo & Withman, 2001). En concreto, dentro del dominio tipo EGF se han identificado tres aminoácidos críticos para la correcta interacción de Cripto con Nodal basados en experimentos de rescate con mutantes: la glicina de la posición 87 (G87), la treonina de la posición 88 (T88) y la fenilalanina de la posición 94 (F94) de Cripto en humanos (Minchiotti *et al*., 2001; Shaolin *et al*., 2007) (figura 54). La sustitución de estos aminoácidos por otros conlleva una completa falta de función de la proteína Cripto. Modelos tridimensionales han confirmado que estos tres aminoácidos están expuestos en el mismo lado de la proteína y que forman una superficie que podría servir de punto de interacción con Nodal. Los resultados de esta tesis muestran que tanto G87 como F94 están conservadas en todos los ortólogos aquí identificados (figura 54). Sin embargo, T88 solamente está conservada en deuteróstomos, mientras que los moluscos poseen una leucina (L) o una arginina (R), los anélidos una metionina (M) o una isoleucina (I), y el braquiópodo una valina (V) (figura 54). En un primer momento se pensó que la fucosilación de la T88 era suficiente y necesaria para la correcta unión a Nodal. Sin embargo, estudios posteriores han demostrado que es la presencia del aminoácido treonina, y no la fucosilación, lo que hace que la unión sea posible (Shaolin *et al*., 2007). De hecho, experimentos de mutagénesis dirigida hacia la T88 para cambiarla por serina, también susceptible de fucosilación, mostraron que la serina no era capaz de mantener la señalización de Cripto/Nodal (Shaolin *et al*., 2007). Por tanto, parece que la fucosilación de T88 no es ni

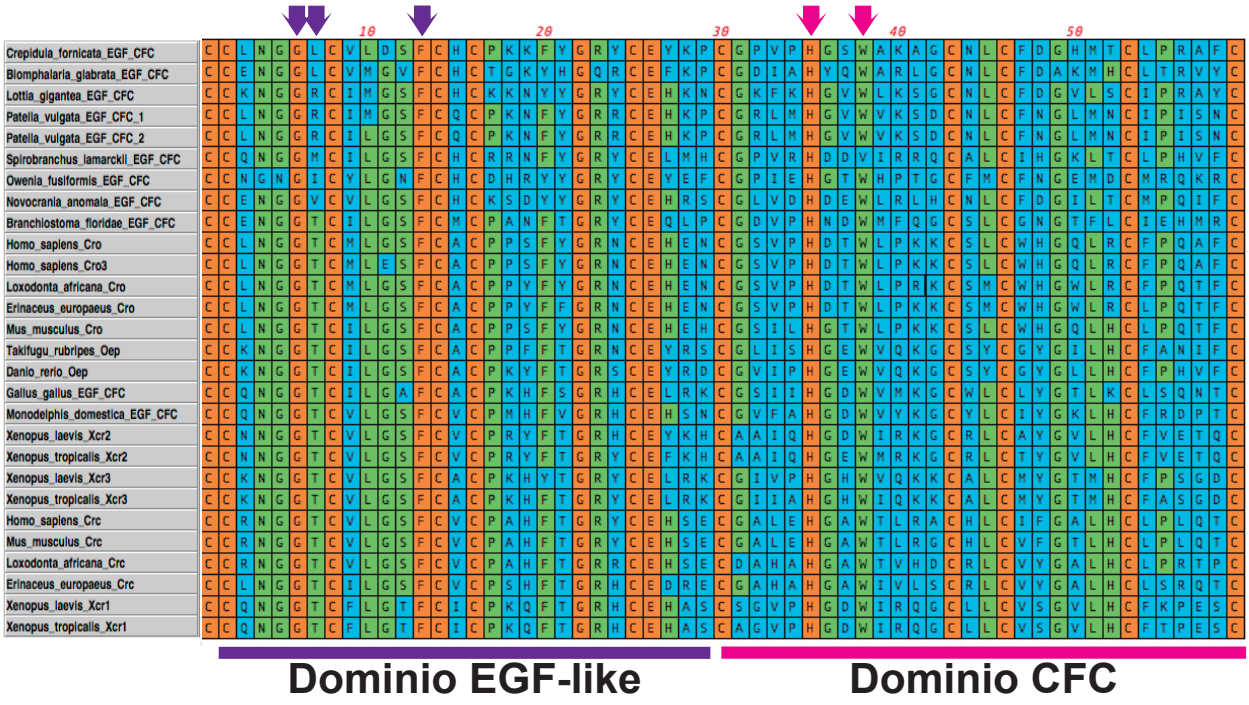


Figura 54. Posiciones relevantes para la unión del co-receptor EGF-CFC al ligando Nodal y al receptor tipo I ALK4 en vertebrados. Dentro del dominio EGF se han identificado tres aminoácidos críticos para la correcta interacción de EGF-CFC con Nodal: la glicina de la posición 87 (G87), la treonina de la posición 88 (T88) y la fenilalanina de la posición 94 (F94) de Cripto en humanos; estas posiciones están marcadas en la figura con las flechas moradas. Nótese que mientras G87 y F94 están conservadas en todos los ortólogos aquí identificados, T88 no está conservada en organismos no deuteróstomos. Por otra parte, se ha descrito que la histidina de la posición 120 (H120) y el triptófano de la posición 123 (W123) de Cripto en humanos son los responsables del reconocimiento en la interacción con el receptor Alk4; estas posiciones están marcadas en la figura con las flechas rosas. Nótese que estas dos posiciones están conservadas en todos los ortólogos EGF-CFC encontrados (excepto W123 para el anélido *Spirobranchus*).

suficiente ni necesaria para la funcionalidad de Cripto en la ruta de Nodal, sino que más bien puede facilitar la estabilidad de la unión de ambas proteínas o no tener ninguna consecuencia funcional (Shaolin *et al*., 2007). Otros cambios de la T88 por alanina, ácido aspártico, ácido glutámico, histidina, lisina, asparagina, glutamina, valina o tirosina, también llevaron a formas defectivas de Cripto (Schiffer *et al*., 2001; Yan *et al*., 2002 Shaolin *et al*., 2007). Por tanto, todos estos experimentos apuntan a que la T88 es crítica para la correcta interacción de Nodal con Cripto, y su sustitución por otros aminoácidos no puede conseguir la activación de la ruta. El hecho de que todos los ortólogos de bilaterados no deuteróstomos identificados en esta tesis carezcan de una treonina en esa posición abre la interesante posibilidad de que Cripto no pueda unirse a Nodal y activar su señalización en estos organismos.

Con el fin de comprobar esta hipótesis, se realizaron diferentes experimentos de mutagénesis dirigida en el ortólogo de *egf-cfc* de *C. fornicata*. En la posición homóloga a la T88 en humanos, *C. fornicata* tiene una leucina (L56). Por una parte, se inyectó una forma mutante de

Crepidula en la que se sustituyó L56 por una alanina en embriones de *C. fornicata*. Los embriones resultantes se desarrollaron de manera normal y no se detectó ninguna modificación, lo que indica que la sustitución de esa posición no modifica la funcionalidad de EFG-CFC (tabla 15). Por otra parte, la inyección de la forma silvestre de *C. fornicata*, con L56, en un fondo mutante *oep* para el pez cebra no fue capaz de rescatar el fenotipo silvestre (tabla 16; figura 51). Por tanto, estos experimentos muestran que la forma *egf-cfc* de *C. fornicata* con L56 no es capaz de activar la ruta de señalización de Nodal en el pez cebra. Sin embargo, la inyección de una forma mutante de *C. fornicata* en la que se sustituyó la L56 por una treonina en el mutante *oep* sí fue capaz de rescatar los fenotipos silvestres (figura 51; tabla 16). En algunos casos se obtuvieron individuos con un rescate parcial del fenotipo silvestre; esto es, con una reducción de la hipertrofia cardiaca, la recuperación aparentemente completa del eje anterior-posterior, un desarrollo casi completo de la cabeza y la generación de dos ojos (figura 51; tabla 16). Estos datos indican que, a pesar de todas las diferencias que se detectan al comparar las secuencias de aminoácidos de EFG-CFC entre el pez cebra y *C. fornicata*, el cambio a una treonina en la posición clave L56 de *C. fornicata* permite la interacción de EFG-CFC con el ortólogo de Nodal en el pez cebra y por tanto la activación de la ruta de señalización. Estas evidencias corroboran que la presencia de una treonina en esa posición es necesaria para la activación de la ruta y la correcta formación del plan corporal. Por tanto, estos datos refuerzan la idea de que EFG-CFC no es un elemento necesario para que Nodal se una a su receptor en los gasterópodos y probablemente tampoco lo sea en otros organismos no deuteróstomos, aunque este hecho deberá ser testado en futuros experimentos.

Varias evidencias han demostrado que Nodal tiene actividad de manera independiente de Cripto en vertebrados. En primer lugar, el hecho de que los fenotipos observados en los mutantes de *nodal* sean mucho más severos que los de los mutantes de *cripto* indica que hay señalización de Nodal que no requiere de esta proteína (Brennan *et al*., 2001; Conlon *et al*., 1994; Ding *et al*., 1998; Liguori *et al*., 2008). Además, se ha visto que una señalización de Nodal independiente de Cripto es capaz de formar correctamente el eje anterior-posterior en ratones, así como de formar tanto el mesoendodermo anterior como el neuroectodermo posterior (Liguori *et al*., 2008). Por tanto, algunos autores han propuesto la existencia de dos rutas para el morfógeno Nodal: una Cripto-dependiente y otra Cripto-independiente, siendo la última activa sólo hasta que comienza la expresión de un regulador negativo denominado *cerberus* (Liguori *et al*., 2008). El regulador Cerberus parece ser una novedad evolutiva de los deuteróstomos, puesto que no se ha encontrado hasta el momento en ningún otro grupo (incluido gasterópodos y en particular

Crepidula) (Kenny *et al*., 2014). Por tanto, es posible que en gasterópodos y otros bilaterados no deuteróstomos Nodal señalice exclusivamente de manera EGF-CFC independiente.

Esta hipótesis lleva a preguntarse por el rol que desempeña EGF-CFC en gasterópodos. Para intentar resolver esta pregunta, se realizaron experimentos de sobreexpresión y de falta de función (con la deleción completa del dominio EGF) de EGF-CFC en *C. fornicata*. La sobreexpresión de EGF-CFC en *C. fornicata* no produjo ninguna modificación visible en el fenotipo de los juveniles resultantes (tabla 15). Sin embargo, en los experimentos de falta de función se obtuvieron fenotipos de embriones radializados, en los que no puede distinguirse ningún eje corporal ni ningún destino celular y en los que no se produce la gastrulación (figura 49; tabla 15). Este fenotipo recuerda mucho al obtenido en la inhibición de la ruta MAPK (*mitogen activated protein kinase*) en *C. fornicata* (Henry & Perry, 2008). Esto, junto con el hecho de que Cripto activa la ruta MAPK en vertebrados (De Santis, 1997; Kannan *et al*., 1997; Ebert *et al*., 1999; Bianco *et al*., 2004) puede indicar que la ruta MAPK está siendo regulada de alguna manera por EGF-CFC en *C. fornicata*. Esta hipótesis deberá ser comprobada por futuros experimentos.

En resumen, los resultados aquí expuestos muestran que la proteína EGF-CFC puede no ser un elemento necesario para la activación de la ruta de Nodal en gasterópodos. Por tanto, EGF-CFC puede que no esté implicado directamente en el establecimiento de las AID controladas por Nodal en gasterópodos. La ausencia de la T88 en todos los ortólogos aquí descritos para organismos no deuteróstomos podría implicar que la ruta de Nodal comienza a actuar de una manera EGF-CFC dependiente en el antecesor de todos los vertebrados (y probablemente de todos los deuteróstomos). Futuras investigaciones funcionales en bilaterados no deuteróstomos servirán para testar esta hipótesis.

Los resultados aquí presentados abren una interesante línea de investigación que deja algunas preguntas abiertas, como por ejemplo:

- ¿Existe alguna relación directa entre la ruta MAPK y la proteína EGF-CFC en gasterópodos? La activación de MAPK por parte de Cripto en vertebrados ocurre a través de la unión de Cripto a la proteína Glipicano-1; esta unión activa una tirosín-kinasa citoplasmática, c-Src, que a su vez activa la ruta MAPK (Bianco *et al*., 2003). Sería necesario identificar ortólogos de estos componentes en los gasterópodos y describir su activación, funcionalidad y su relación con EGF-CFC.

- Además de la ruta MAPK, las proteínas EGF-CFC también intervienen en las rutas de Wnt y Notch en vertebrados (revisado en Klauzinska *et al*., 2014). Recientemente se ha demostrado que EGF-CFC regula la ruta de Wnt mediante su unión con los co-receptores LRP5 y LRP6; esta unión facilita a su vez la unión de Wnt3a con su receptor que produce la estabilización de β–Catenina, lo que eleva su actividad transcripcional (Nagaoka *et al*., 2013). Por otra parte, EGF-CFC es capaz de potenciar la ruta de Notch, favoreciendo el procesamiento extracelular del precursor de Notch-1 a través de una proteasa Furin-l (Watanabe *et al*., 2009). Sería interesante dilucidar si EGF-CFC es realmente un componente implicado en estas rutas de señalización en gasterópodos.

- El hecho de que la estructura molecular de las proteínas EGF-CFC esté muy conservada entre diferentes bilaterados puede indicar que estas proteínas tienen la posibilidad de actuar tanto de una manera autónoma en la célula (como proteínas de membrana) como de manera no autónoma (al ser proteínas solubles cuando proteoliza el sitio de unión del dominio GPI). Cabe la posibilidad por tanto de que las proteínas EGF-CFC en organismos no deuteróstomos también puedan actuar como proteínas solubles regulando la actividad de otros genes, una posibilidad merecedora de atención.

5.4. Generación de herramientas y puesta a punto de diferentes metodologías para el estudio de organismos no modelo como los gasterópodos

A lo largo del desarrollo de esta tesis, una parte muy importante del esfuerzo se invirtió en la generación de recursos y a la puesta a punto de diferentes técnicas para el abordaje de las cuestiones planteadas.

En primer lugar, se han generado tres bases de datos transcriptómicas para las especies de gasterópodos *C. fornicata* y *B. glabrata* (apartado 3.5.). El desarrollo de estas bases de datos supuso un punto de inflexión durante el trabajo aquí realizado, puesto que permitió la búsqueda de genes de interés de manera rápida y efectiva. Antes de la existencia de estas bases de datos, la búsqueda de cada gen debía realizarse primero, a través del diseño de cebadores degenerados en base a las secuencias de ortólogos disponibles para otras especies. Las amplificaciones por PCR realizadas con cebadores degenerados son un reto, ya que las condiciones para que dichas amplificaciones sean positivas deben afinarse muy ajustadamente. En segundo lugar se

requieren PCRs de tipo RACE para obtener los fragmentos completos de los genes estudiados. De nuevo, este tipo de PCRs constituye un paso delicado, puesto que la existencia de mRNA incompletos o una baja eficiencia en la retrotranscripción original del cDNA molde, entre otros factores, pueden dificultar la amplificación de los extremos 5´y 3´ UTR. En tercer lugar, tras la secuenciación completa del gen deseado, deben diseñarse cebadores específicos en una zona adecuada para conseguir la amplificación completa de la zona codificante que puedan utilizarse en estudios por ejemplo, de hibridaciones *in situ* o construcciones con formas silvestres y/o mutantes. Estos tres pasos de gran dificultad se reducen a uno solo desde el momento en que se dispone de una base de datos transcriptómica, ya que los dos primeros pasos no son necesarios cuando las búsquedas bioinformáticas son capaces de encontrar el gen en cuestión. Estos recursos genómicos generados ya han sido publicados en revistas de reconocido prestigio internacional (apéndices 11 y 12) y están disponibles para otros investigadores. En resumen, las bases de datos transcriptómicas generadas para las especies estudiadas han sido la base de muchos experimentos y constituyen una pieza fundamental en la planificación de futuros estudios tanto de biología del desarrollo como de otras disciplinas y suponen por tanto, una herramienta de gran utilidad para el estudio de estas especies.

Otro punto a destacar lo componen los anticuerpos específicos diseñados y sintetizados durante el desarrollo de esta tesis para las especies de gasterópodos *C. fornicata* y *B. glabrata*. Los anticuerpos comerciales diseñados para vertebrados de las proteínas aquí estudiadas no dieron resultados positivos cuando se usaron en embriones de *Crepidula* o *Biomphalaria*. Dada la relevancia que tiene la localización de la proteína en este tipo de estudios, se consideró necesario realizar un esfuerzo para la síntesis de anticuerpos específicos (apartado 3.12.). El trabajo se centró en el diseño, síntesis y comprobación de la especificidad de los anticuerpos sintetizados para las proteínas estudiadas (tabla 2). Durante el desarrollo de esta tesis se ha comenzado esta tarea aunque aún queda mucho trabajo que realizar en cuanto a la comprobación de todos los sueros generados (únicamente se ha testado el suero 3 extraído de los conejos A y B después de ser inoculados con los péptidos específicos para cada anticuerpo; aún quedan por comprobar los sueros 1, 2, 4, 5 y 6). Los datos obtenidos hasta el momento indican que los anticuerpos sintetizados probablemente reconocen en su mayoría las proteínas para las que fueron diseñados. La proteína Inversin (proveniente de las dos isoformas) y la proteína Nodal2 de *C. fornicata*, así como la proteína Inversin y la proteína EGF-CFC de *B. glabrata* pudieran estar siendo reconocidas específicamente por los anticuerpos sintetizados (usando técnicas de

western blot) (figuras 25, 26, 35 y 46). Por otra parte, la proteína EGF-CFC de *C. fornicata* pudiera estar siendo reconocida de manera específica por los anticuerpos sintetizados en reacciones de inmunohistoquímica sobre individuos completos (figura 47).

Además del testado de los sueros que aún no se han probado, queda por realizar una segunda comprobación de especificidad que incluye la utilización de los péptidos específicos diseñados para cada gen. Estos péptidos ya han sido sintetizados y purificados de manera individual. En el futuro, estos péptidos específicos individualizados se incubarán con las muestras de suero con el anticuerpo. Como resultado de esta incubación, se espera que los péptidos se unan a los anticuerpos de manera específica y por tanto, que bloqueen los puntos de unión. Seguidamente se incubará esta mezcla con las muestras de proteína total proveniente de los embriones de la especie de gasterópodo deseada. Esta incubación no debería dar lugar a la aparición de ninguna banda en los *western blot*, ya que al estar bloqueados los anticuerpos, éstos no serían capaces de unirse a las proteínas presentes en el embrión. De esta manera, se estará comprobando si los anticuerpos reconocen de manera específica a las proteínas endógenas.

Otro de los grandes esfuerzos realizados en esta tesis ha sido el diseño y puesta a punto de experimentos para el estudio funcional de los genes de interés. En concreto, se ha utilizado la técnica CRISPR, una técnica novedosa en el campo que ha demostrado ser de gran utilidad para el silenciamiento de genes diana (apartado 3.18.). Sin embargo, los datos aquí obtenidos parecen indicar que no se ha producido un correcto silenciamiento de los genes deseados (tablas 8 y 11). Se necesitan más experimentos para intentar detectar el problema en el desarrollo de esta técnica. Particularmente necesario sería comprobar que realmente se ha conseguido eliminar del genoma el gen diana en los zigotos inyectados con el molde para el RNA guía y la nucleasa Cas9. Esto podría comprobarse, por ejemplo, mediante PCR para amplificar el gen diana en las larvas que fueron inyectadas en estadio zigoto. Hasta el momento este tipo de experimentos no se ha realizado, aunque se realizarán en un futuro cercano.

Por otra parte, también se han utilizado cebadores *morpholino antisense* con el fin de impedir la transcripción de los genes diana (apartado 3.19.). Los datos aquí obtenidos parecen indicar que no se ha impedido la transcripción de los genes deseados, puesto que en todos los experimentos se obtenían individuos con un fenotipo silvestre (tablas 9 y 14). Es necesario un mayor esfuerzo experimental para optimizar esta técnica. En concreto, una vez que los protocolos de inmuno-detección con cada uno de los anticuerpos específicos se hayan optimizado, éstos podrán usarse para detectar si la proteína está siendo sintetizada tanto en los

embriones control como en aquéllos inyectados con el morfolino. Este abordaje experimental será llevado a cabo también en un futuro próximo.

Finalmente cabe destacar que la generación de mutaciones dirigidas sobre el molde silvestre de los genes estudiados sí ha resultado ser una técnica con resultados prometedores (apartado 3.16.). La especificidad de esta técnica es extremadamente elevada; en algunas ocasiones se han generado mutantes en los que sólo había un cambio de un único aminoácido en la secuencia de aminoácidos. Se han obtenido fenotipos diferentes a los silvestres en los juveniles provenientes de zigotos inyectados con formas mutantes de manera consistente y robusta en diferentes ensayos. Asimismo, los experimentos de rescate llevados a cabo han conseguido obtener el fenotipo silvestre. Estos resultados sugieren que la función de los genes estudiados ha sido modificada de manera específica (tabla 10, 12, 15 y 16). Estos prometedores resultados abren una interesantísima vía para el abordaje de estudios funcionales en gasterópodos.

Conclusiones

1) El gen *inversin* tiene un origen antiguo en la historia evolutiva de los animales. Existen ortólogos de este gen en coanoflagelados y en los principales grupos de metazoos (Porifera, Cnidaria y Bilateria).

2) La estructura molecular de las proteínas Inversin identificadas en estos grupos está muy conservada. En todas las proteínas ortólogas pueden distinguirse 16 repeticiones ankirina y un número variable de dominios de destrucción D-Box y de localización nuclear NLS. Además, todas estas proteínas ortólogas presentan dos dominios IQ excepto la de la especie levógira de gasterópodo *B. glabrata*, que carece del dominio IQ2.

3) El gasterópodo dextrógiro *C. fornicata* tiene dos isoformas (*inversin-L* e *inversin-S*), generadas por procesamiento alternativo del gen *inversin*. Estas dos isoformas se expresan en diferentes momentos y dominios durante el desarrollo embrionario de esta especie, aunque ambas presentan patrones de activación asimétricos con respecto al eje de simetría bilateral.

4) La proteína Inversin en el gasterópodo dextrógiro *C. fornicata* se localiza en el blastoporo durante la gastrulación y la organogénesis.

5) La sobre-expresión de la isoforma *inversin-L* en embriones de *C. fornicata* no tiene ningún efecto fenotípico apreciable durante el desarrollo. Sin embargo, la presencia de una forma dominante negativa de la isoforma *inversin-L* (que carece de ambos dominios IQ) durante el desarrollo embrionario de *C. fornicata* ocasiona la aparición de fenotipos simétricos, aunque con una frecuencia muy baja. Igualmente, la presencia del transcrito silvestre del gen *inversin* de la especie levógira *B. glabrata* (que carece del dominio IQ2) durante el desarrollo embrionario de la especie *C. fornicata* ocasiona la aparición de fenotipos simétricos, aunque también con una frecuencia muy baja.

6) El gasterópodo *C. fornicata* tiene dos parálogos del gen *nodal* (*nodal1* y *nodal2*) que se expresan en diferentes momentos y dominios durante el desarrollo embrionario de esta especie; sólo *nodal2* presenta un patrón de activación asimétrico en el lado derecho del embrión.

7) El gen *nodal2* regula el establecimiento de asimetrías izquierda-derecha (AID) durante el desarrollo embrionario de la especie *C. fornicata*. La sobre-expresión del gen *nodal2* en embriones de *C. fornicata* ocasiona la aparición de fenotipos simétricos.

8) La presencia de la proteína recombinante Activin durante el desarrollo embrionario de la especie *C. fornicata* ocasiona la inactivación de la ruta de Nodal y la aparición de fenotipos simétricos.

9) La familia de proteínas EGF-CFC es muy antigua y estaba presente al menos en el antecesor común de deuteróstomos, ecdisozoos y espiralados.

10) Los transcritos sentido y antisentido del gen *egf-cfc* están activos durante todo el desarrollo embrionario de la especie de gasterópodo *C. fornicata*, mientras que la proteína EGF-CFC se localiza en las membranas de las células del estomodeo durante la gastrulación y organogénesis.

11) Las proteínas EFG-CFC no parecen ser un elemento necesario para que el ligando Nodal se una a su receptor en los gasterópodos y probablemente tampoco en otros organismos no deuteróstomos. Esto se debe a la ausencia de una treonina (T88 en humanos) en la posición homóloga en la proteína EFG-CFC de gasterópodos y organismos no deuteróstomos. Por tanto, las proteínas EFG-CFC no parecen tener un papel en el establecimiento de AID en gasterópodos.

12) La presencia de una forma dominante negativa del gen *egf-cfc* (que carece del dominio EGF) durante el desarrollo embrionario de *C. fornicata* ocasiona la aparición de embriones con fenotipos radializados, en los que no se ha producido gastrulación, definición de ejes ni diferenciación celular.

1) The *inversin* gene has an ancient origin in the evolutionary history of animals. Orthologs exist for this gene in choanoflagelates and in the main groups of Metazoa (Porifera, Cnidaria and Bilateria).

2) The molecular structure of the identified Inversin proteins in these groups is very conserved. All the orthologs show 16 ankyrin repetitions and a variable number of D-Box and NLS domains. Also, all these orthologs show two IQ domains, except the sinistral gastropod species *B. glabrata*, which lacks the IQ2 domain.

3) The dextral gastropod, *C. fornicata*, has two isoforms (*inversin-L* and *inversin-S*), generated by alternative splicing. These two isoforms are expressed in different moments and domains during the embryonic development in this species, although both show asymmetrical expression patterns with respect to the bilateral axis.

4) Inversin protein in the dextral gastropod *C. fornicata* localizes on the blastopore during gastrulation and organogenesis.

5) The over-expression of *inversin-L* isoform in *Crepidula* embryos doesn't have any appreciable phenotypic effect during development. However, the presence of a dominant negative form for *inversin-L* isoform (lacking both IQ domains) during *Crepidula* embryonic development promotes the symmetric phenotype, but at very low frequency. Likewise, the presence of the *inversin* wild-type transcript for the sinistral species, *B. glabrata* (lacking the IQ2 domain) during *Crepidula* development causes the symmetric phenotype, also at very low frequency.

6) The gastropod *C. fornicata* has two paralogs for *nodal* gene (*nodal1* and *nodal2*) which are expressed at different times and regions during the embryonic development of this species; only *nodal2* shows an asymmetrical pattern on the right side of the embryo.

7) *nodal2* gene regulates the establishment of left/right asymmetries (LRA) during the embryonic development of *C. fornicata.* The over-expression of *nodal2* in *Crepidula* embryos causes symmetric phenotypes.

8) The presence of recombinant Activin protein during *Crepidula* development causes the inactivation of the Nodal pathway and symmetric phenotypes.

9) The protein family EGF-CFC is very ancient and was present, at least, in the common ancestor of Deuterostomia, Ecdysozoa and Spiralia.

10) Both sense and antisense *egf-cfc* transcripts are active through the development of the gastropod *C. fornicata*, while the EGF-CFC protein localizes on the dorsal lip cell membranes in the stomadeum during gastrulation and organogenesis.

11) The EGF-CFC proteins don't seem to be a necessary element for the Nodal ligand to join its receptor in gastropods and probably in other non-deuterostome organisms. This is due to the absence of a threonine (T88 in humans) at the homologous position in the protein EGF-CFC in gastropods and non-deuterostome organisms. Hence, the EGF-CFC proteins do not seem to play a role in the establishment of LRA in gastropods.

12) The presence of a dominant negative form for *egf-cfc* gene (which lacks the EGF domain) during the development of *C. fornicata* leads to radialized embryos, in which there is no gastrulation, axis definition or cellular differentiation.

Bibliografía

Abascal F, Zardoya R, Posada D. 2005. ProtTest: selection of best-fit models of protein evolution. Bioinformatics; 21(9):2104-2105

Adema CM, Hillier LW, Jones CS, Loker ES, Knight M, Minx P, Oliveira G, Raghavan N, Shedlock A, do Amaral LR, et al. 2017 Whole genome analysis of a schistosomiasis-transmitting freshwater snail. Nat Commun; 8:15451

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol; 215(3):403-410

Attisano L, Wrana JL, Cheifetz S, Massague J. 1992. Novel activin receptors: distinct genes and alternative mRNA splicing generate a repertoire of serine/threonine kinase receptors. Cell; 68:97–108

Aw S, Adams DS, Qiu D, Levin M. 2008. H,K-ATPase protein localization and Kir4.1 function reveal concordance of three axes during early determination of left-right asymmetry. Mech Dev; 125(3-4):353-372

Aw S, Levin M. 2009. Is left-right asymmetry a form of planar cell polarity?. Development; 136(3):355-366

Aykul S, Parenti A, Chu KY, Reske J, Floer M, Ralston A, Martínez-Hackert E. 2017. Biochemical and cellular analysis reveals ligand binding specificities, a molecular basis for ligand recognition, and membrane association-dependent activities of Cripto-1 and Cryptic. J Biol Chem; 292(10):4138-4151

Beiter T, Reich E, Williams RW, Simon P. 2009. Antisense transcription: a critical look in both directions. Cell Mol Life Sci; 66(1):94–112

Bennett PM. 2004. Genome plasticity: insertion sequence elements, transposons and integrons, and DNA rearrangement. Methods Mol Biol; 266:71-113

Bergmann DC, Lee M, Robertson B, Tsou MF, Rose LS, Wood WB. 2003. Embryonic handedness choice in C. elegans involves the Galpha protein GPA-16. Development; 130(23):5731-5340

Bessodes N, Haillot E, Duboc V, Röttinger E, Lahaye F, Lepage T. 2012. Reciprocal signaling between the ectoderm and a mesendodermal left-right organizer directs left-right determination in the sea urchin embryo. PLoS Genet; 8(12):e1003121

Bianco C, Adkins HB, Wechselberger C, Seno M. Normanno N, De Luca A, Sun Y, Khan N, Kenney N, Ebert A, Williams KP, Sanicola M, Salomon DS. 2002. Cripto-1 activates nodal- and ALK4-dependent and -independent signaling pathways in mammary epithelial Cells. Mol Cell Biol; 22:2586–2597

Bianco C, Strizzi L, Rehman A, Normanno N, Wechselberger C, Sun Y, Khan N, Hirota M, Adkins H, Williams K et al. 2003. A Nodal- and ALK4-independent signaling pathway activated by Cripto-1 through Glypican-1 and c-Src. Cancer Res; 63(6):1192-1197

Bianco C, Normanno N, Salomon DS, Ciardiello F. 2004. Role of the cripto (EGF-CFC) family in embryogenesis and cáncer. Growth Factors; 22(3):133-139

Bianco C, Salomon DS. 2010. Targeting the embryonic gene Cripto-1 in cáncer and beyond. Expert Opin Ther Pat; 20(12):1739-1749

Blanchet MH, Le Good JA, Oorschot V, Baflast S, Minchiotti G, Klumperman J, Constam DB. 2008. Cripto localizes Nodal at the limiting membrane of early endosomes. Sci Signal; 1(45):ra13

Blanchet MH, Le Good JA, Mesnard D, Oorschot V, Baflast S, Minchiotti G, Klumperman J, Constam DB. 2008. Cripto recruits Furin and PACE4 and controls Nodal trafficking during proteolytic maturation. EMBO J; 27(19):2580-2591

Boorman CJ, Shimeld SM. 2002. Pitx homeobox genes in Ciona and Amphioxus show left-right asymmetry is a conserved chordate character and define the ascidian adenohypophysis. Evol Dev; 4(5):354-365

Bramford RN, Roessler E, Burdine RD, Splakoglu U, De la Cruz J, Splitt M, Goodship JA, Towbin J, Bowers P, Ferrero GB, et al. 2000. Loss-of-function mutations in the EGF-CFC gene CFC1 are associated with human left-right laterality defects. Nat Ganet; 26(3):365-369

Brennan J, Lu CC, Norris DP, Rodriguez TA, Beddington RS, Robertson EJ. 2001. Nodal signalling in the epiblast patterns the early mouse embryo. Nature; 411(6840):965-969

Camey T & Verdonk NH. 1969. The Early Development of the Snail Biomphalaria Glabrata (Say) and the Origin of the Head Organs. Netherlands J Zool; 20(1):93-121

Calvanese L, Saporito A, Oliva R, D' Auria G, Pedone C, Paolillo L, Ruvo M, Marasco D, Falcigno L. 2009. Structural insights into the interaction between the Cripto CFC domain and the ALK4 receptor. J Pept Sci; 15(3):175-183

Calvanese L, Marasco D, Doti N, Saporito A, D'Auria G, Paolillo L, Ruvo M, Falcigno L. 2010. Structural investigations on the Nodal-Cripto binding: a theoretical and experimental approach. Biopolymers; 93(11):1011-1021

Cannon JT, Vellutini BC, Smith J 3rd, Ronquist F, Jondelius U, Hejnol A. 2016. Xenacoelomorpha is the sister group to Nephrozoa. Nature; 530(7588):89-93

Capdevilla J, Vogan KJ, Tabin CJ, Izpisúa-Belmonte JC. 2000. Mechanisms of left/right determination in vertebrates. Cell; 101:9–21

Carcamo J, Weis FM, Ventura F, Wieser R, Wrana JL, Attisano L, Massague, J. 1994. Mol Cell Biol; 14:3810–3821

Carr M, Leadbeater BS, Hassan R, Nelson M, Baldauf SL. 2008. Molecular phylogeny of choanoflagellates, the sister group to Metazoa. PNAS; 105(43):16641-16646

Carroll SB, Gates J, Keys DN, Paddock SW, Panganiban GE, Selegue JE, Williams JA. 1994. Pattern formation and eyespot determination in butterfly wings. Science; 265(5168):109-114

Chang C, Hsieh YW, Lesch BJ, Bargmann CI, Chuang CF. 2011. Microtubule-based localization of a synaptic calcium-sugnaling complex is required for left-right neuronal asymmetry in C. elegans. Development; 138(16):3509-3518

Cheatle Jarvela AM, Pick L. 2016. Evo-Devo: Discovery of Diverse Mechanisms Regulating Development. Curr Top Dev Biol; 117:253-274

Chin D, Means AR. 2000. Calmodulin: a prototypical calcium sensor. Trends Cell Biol; 10:322-328

Christiaen L, Bourrat F, Joly JS. 2005. A modular cis-regulatory system controls isoform-specific pitx expression in ascidian stomodaeum. Dev Biol; 277(2):557-566

Christiaen L, Jaszczyszyn Y, Kerfant M, Kano S, Thermes V, Joly JS. 2007. Evolutionary modification of mouth position

in deuterostomes. Semin Cell Dev Biol; 18(4):502-511

Chu J, Ding J, Jeays-Ward K, Price SM, Placzek M, Shen MM. 2005. Non-cell-autonomous role for Cripto in axial midline formation during vertebrate embryogenesis. Development; 132(24):5539-5551

Ciardiello F, Tortora G, Bianco C, Selvam MP, Basolo F, Fontanini G, Pacifico F, Normanno N, Brandt R, Persico MG, et al. 1994. Inhibition of Cripto expression and tumorigenicity in human colon cancer cells by antisense RNA and oligodeoxynucleotides. Oncogene; 9(1):291-298

Conklin EG. 1897. The embryology of Crepidula, a contribution to the cell lineage and early development of some marine gasteropods. J Morph; 13(1):1-226

Conlon FL, Lyons KM, Takaesu N, Barth KS, Kispert A, Herrmann B, Robertson EJ. 1994. A primary requirement for nodal in the formation and maintenance of the primitive streak in the mouse. Development; 120(7):1919-1928

Crivici A, Ikura M. 1995. Molecular and structural basis of target recognition by calmodulin. Annu Rev Biophys Biomol Struct; 24:85-116

Davison A, McDowell GS, Holden JM, Johnson HF, Koutsovoulos GD, Liu MM, Hulpiau P, Van Roy F, Wade CM, Beneriee R, et al. 2016. Formin is associated with left-right asymmetry in the pond snail and the frog. Curr Biol, 26(5):654-660

De Alburquerque LP, Pontual EV, Santana GM, Silva LR, Aquiar Jdos S, Coelho LC, Rêgo MJ, Pitta MG, da Silva TG, Melo AM, Napoleao TH, Paiva PM. 2014. Toxic effects of Microgramma vacciniifolia rhizome lectin on Artemiasalina, human cells, and the schistosomiasis vector Biomphalaria glabrata. Acta Trop; 138:23-27

De Luca A, Cassamassimi A, Selvam MP, Losito S, Ciardiello F, Agrawai S, Salomon DS, Normanno N. 1999. EGF-related peptides are involved in the proliferation and survival of MDA-MB-468 human breast carcinoma cells. Int J Cancer; 80(4):589-594

De Luca A, Arra C, D'Antonio A, Casamassimi A, Losito S, Ferraro P, Ciardiello F, Salomon DS, Normanno N. 2000. Simultaneous blockage of different EGF-like growth factors results in efficient growth inhibition of human colon carcinoma xenografts. Oncogene; 19(51):5863-5871

De Santis ML, Kannan S, Smith GH, Seno M, Bianco C, Kim N, Martinez-Lacaci I, Wallace-Jones B, Salomon DS. 1997. Cripto-1 inhibits beta-casein expression in mammary epithelial cells through a p21ras-and phosphatidylinositol 3'-kinase-dependent pathway. Cell Growth Differ; 8(12):1257-1266

Delsuc F, Brinkmann H, Philippe H. 2005. Phylogenomics and the reconstruction of the tree of life. Nat Rev Genet; 6(5):361–375

Dictus WJ, Damen P. 1997. Cell-lineage and clonal-contribution map of the trochophore larva of Patella vulgata (Mollusca). Mech Dev; 62:213-226

Dijke PT, Yamashita H, Ichijo H, Franzen P, Laiho M, Miyazono K, Heldin CH. 1994. Characterization of type I receptors for transforming growth factor-beta and activin. Science; 264(5155):101–104

Ding J, Yang L, Yan YT, Chen A, Desai N, Wynshaw-Boris A, Shen MM. 1998. Cripto is required for correct orientation of the anterior-posterior axis in the mouse embryo. Nature; 395(6703):702-707

Dolnick BJ. 1997. Naturally occurring antisense RNA. Pharmacol Ther; 75(3):179-184

Dorey K & Hill CS. 2006. A novel Cripto-related protein reveals an essential role for EGF-CFCs in Nodal signalling in Xenopus embryos. Dev Biol; 292(2):303-316

Driever W, Nüsslein-Volhard C. 1989. The bicoid protein is a positive regulator of hunchback transcription in the early Drosophila embryo. Nature; 337(6203):138-143

Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian Phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol; 29(8):1969–1973

Drummond AJ & Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol; 7:214

Duboc V, Röttinger E, Lapraz F, Besnardeau L, Lepage T. 2005. Left-right asymmetry in the sea urchin embryo is regulated by nodal signaling on the right side. Dev Cell; 9(1):147-58

Duboc V, Lapraz F, Saudemont A, Bessodes N, Mekpoh F, Haillot E, Quirin M, Lepage T. 2010. Nodal and BMP2/4 pattern the mesoderm and endoderm during development of the sea urchin embryo. Developmen; 137(2):223-235

Ebert AD, Wechselberger C, Frank S, Wallace-Jones B, Seno M, Martínez-Lacaci I, Bianco C, De Santis M, Weitzel HK, Salomon DS. 1999. Cripto-1 induces phosphatidylinositol 3'-kinase-dependent phosphorylation of AKT and glycogen synthase kinase 3beta in human cervical carcinoma cells. Cancer Res; 59(18):4502-4505

Eimon PM, Harland RM. 2002. Effects of heterodimerization and proteolytic processing on Derrière and Nodal activity: implications for mesoderm induction in Xenopus. Development; 129-3089-3103

Eley LT, Tumpenny L, Yates LM, Craighead AS, Morgan D, Whistler C, Goodship JA, Strachan T. 2004. A perspective on inversin. Cell Biol Intern; 28(2):119-124

Foley SF, Van Vlijmen HW, Boynton RE, Adkins HB, Cheung AE, Singh J, Sanicola M, Young CN, Wen D. 2003. The CRIPTO/FRL-1/CRYPTIC (CFC) domain of human Cripto. Functional and structural insights through disulfide structure analysis. Eur J Biochem; 270(17):3610-3608

Freeman G, Lundelius JW. 1992. Evolutionary implications of the mode of D quadrant specification in coelomates with spiral cleavage. J Evol Biol; 5(2):205-247

Gaio U, Schweickert A, Fischer A, Garratt AN, Müller T, Azcelik C, Lankes W, Strehle M, Britsch S, Blum M, Birchmeier C. 1999. A role of the cryptic gene in the correct establishment of the left-right axis. Curr Biol; 9(22):1339-1342

Galant R, Carrol SB. 2002. Evolution of a transcriptional repression domain in an insect Hox protein. Nature; 415(6874):910-913

Grande C, Patel NH. 2009. Nodal signalling is involved in left-right asymmetry in snails. Nature; 457(7232):1007-1011

Grande C. 2010. Left-right asymmetries in Spiralia. Integr Comp Biol; 50(5):744-55

Grande C, Martín-Durán JM, Kenny NJ, Truchado-García M, Hejnol A. 2014. Evolution, divergence and loss of the Nodal signalling pathway: new data and a synthesis across the Bilateria. Int J Dev Biol; 58(6-8):521-32

Gray PC, Harrison CA, Vale W. 2003. Cripto forms a complex with activin and type II activin receptors and an block activin signaling. PNAS; 100(9):5193-5198

Gray PC, Shani G, Aung K, Kelber J, Vale W. 2006. Cripto binds transforming growth factor beta (TGF-beta) and inhibits TGF-beta signaling. Mol Cell Biol; 26(24):9268-9278

Gray PC & Vale W. 2012. Cripto/GRP78 modulation of the TGF-β pathway in development and oncogenesis. FEBS Lett; 586(14):1836-1845

Gritsman K, Zhang J, Cheng S, Heckscher E, Talbot WS, Schier AF. 1999. The EGF-CFC protein one-eyed pinhead is essential for nodal signaling. Cell; 97(1):121-132

Hamada H, Mano C, Watanabe D, Saijoh Y. 2002. Establishment of vertebrate left-right asymmetry. Nat Rev Genet; 3(2):103-113

Harrison CA, Gray PC, Vale WW, Robertson DM. 2005. Antagonists of activin signaling: mechanisms and potential biological applications. T Endo Met; 16(2):73-78

Hawley SH, Wunnenberg-Stapleton K, Hashimoto C, Laurent MN, Watabe T, Blumberg BW, Cho KW. 1995. Disruption of BMP signals in embryonic Xenopus ectoderm leads to direct neural induction. Genes Dev; 9:2923-2935

Hejnol A. 2010. A Twist in Time — The Evolution of Spiral Cleavage in the Light of Animal Phylogeny. Integr Comp Biol; 50(5): 695–706

Henry JJ, Perry KJ. 2008. MAPK activation and the specification of the D quadrant in the gastropod mollusc, Crepidula fornicata. Dev Biol; 313(1):181-195

Henry JJ, Collin R, Perry KJ. 2010. The Slipper Snail, Crepidula, an emerging lophotrochozoan model system. Biol Bull; 218:211-229

Henry JQ. 2014. Spiralian model systems. Int. J. Dev. Biol; 58:389-401

Hozumi S, Maeda Rm Taniguchi K, Kanai M, Shirakabe, Sasamura T, Spéder P, Noselli S, Aigaki T, Murakami R, Matsuno K. 2006. An unconventional myosin in Drosophila reverses the default handedness in visceral organs. Nature; 440(7085):798-802

Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics; 17(8):754-755

Hurst LD & Smith NGC. 1998. The evolution of concerted evolution. Proc R Soc Lond Ser B; 265: 121–127

Inaki M, Liu J, Matsuno K. 2016. Cell chirality: its origin and roles in left-right asymmetric development. Philos Trans R Soc Lond B Biol Sci; 371(710)

Jao LE, Wente SR, Chen W. 2013. Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. PNAS; 110(34):13904-13909

Johnson SE, Rothstein JL, Knowles BB. 1994. Expression of epidermal growth factor family gene members in early mouse development. Dev Dyn;201(3):216-226

Joseph EM, Melton DA. 1998. Mutant Vg1 ligands disrupt endoderm and mesoderm formation in Xenopus embryos. Development; 125:2677-2685

Jurado LA, Chockalingam PS, Jarrett HW. 1999. Apocalmodulin. Physiol Rev; 79:661-682

Kannan S, De Santis M, Lohmeyer M, Riese DJ 2nd, Smith GH, Hymes N, Seno M, Brandt R, Bianco C, Persico G et al. 1997. Cripto enhances the tyrosine phosphorylation of Shc and activates mitogen-activated protein kinase (MAPK) in mammary epithelial cells. J Biol Chem; 272(6):3330-3335

Katayama S, Tomaru Y, Kasukawa T, Waki K, Nakanishi M, Nakamura M, et al. 2005. Antisense transcription in the mammalian transcriptome. Science; 309(5740):1564–1566

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C. 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics; 28(12):1647–1649

Kelber JA, Shani G, Booker EC, Vale WW, Gray PC. 2008. Cripto is a noncompetitive activin antagonist that forms analogous signaling complexes with activin and nodal. J Biol Chem; 283(8):4490-4500

Kenny NJ, 2014. The Evolution and Development of Left / Right Asymmetry in the Lophotrochozoa. Tesis doctoral. St. Cross College; University of Oxford, Inglaterra

Kenny NJ, Namigai EK, Dearden PK, Hui JH, Grande C, Shimeld SM. 2014. The Lophotrochozoan TGF-β signalling cassette – diversification and conservation in a key signalling pathway. Int J Dev Biol; 58(6-8):533-549

Kenny NJ, Truchado-García M, Grande C. 2016. Deep, multi-stage transcriptome of the schistosomiasis vector Biomphalaria glabrata provides platform for understanding molluscan disease-related pathways. BMC Infect Dis; 16(1):618

Keyte AL, Alonzo-Johnsen M, Hutson MR. 2014. Evolutionary and developmental origins of the cardiac neural crest: building a divided outflow tract. Birth Defects Res C Embryo Today; 102(3):309-323

Klauzinska M, Castro NP, Rangel MC, Spike BT, Gray PC, Bertolette D, Cuttitta F, Salomon D. 2014. The multifaceted role of the embryonic gene Cripto-1 in cancer, stem cells and epithelial-mesenchymal transition. Semin Cancer Biol; 29:51-58

Knytl M, Smolík O, Kubícková S, Tlapáková T, Evans BJ, Krylov V. 2017. Chromosome divergence during evolution of the tetraploid clawed frogs, Xenopus mellotropicalis and Xenopus epitropicalis as revealed by Zoo-FISH. PLoS One;12(5): e0177087

Kruithof-de Julio M, Álvarez MJ, Galli A, Chu J, Price SM, Califano A, Shen MM. 2011. Regulation of extra-embryonic endoderm stem cell differentiation by Nodal and Cripto signaling. Development; 138(18):3885-3895

Kuroda R, Endo B, Abe M, Shimizu M. 2009. Chiral blastomere arrangement dictates zygotic left-right asymmetry pathway in snails. Nature; 462(7274):790-794

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, López R et al. 2007. Clustal W and Clustal X version 2.0. Bioinformatics; 23(21):2947-2948

Levin M, Pagan S, Roberts DJ, Cooke J, Kuehn ME, Tabin CJ. 1997. Left/right patterning signals and the independent regulation of different aspects of situs in the chick embryo. Dev Biol; 189(1):57-67

Lienkamp S, Ganner A, Boehlke C, Schmidt T, Arnold SJ, Schäfer T, Romaker D, Schuler J, Hoff S, Powelske C. 2010. Inversin relays Frizzled-8 signals to promote proximal pronephros development. PNAS; 107(47):20388-93

Lienkamp S, Ganner A, Walz G. 2012. Inversin, Wnt signaling and primary cilia. Differentiation 83(2):549-555

Liguori GL, Borges AC, D'Andrea D, Liguoro A, Gonçalves L, Salgueiro AM, Persico MG, Belo JA. 2008. Cripto-independent Nodal signaling promotes positioning of the A-P axis in the early mouse embryo. Dev Biol; 315(2):280-289

Lobikin M, Wangc G, Xuc J, Hsiehd YW, Chuangd CF, Lemirea JM, Levin M. 2012. Early, nonciliary role for microtubule proteins in left–right patterning is conserved across kingdoms. PNAS; 109(31):12586-12591

Long M, Betrán E, Thornton K, Wang W. 2003. The origin of new genes: glimpses from the young and old. Nat

Lowe LA, Supp DM, Sampath K, Yokoyama T, Wright CV, Potter SS, Overbeek P, Kuehn MR. 1996. Conserved left-right asymmetry of nodal expression and alterations in murine situs inversus. Nature; 381(6578):158-161

Luo K. 2017. Signaling cross-talk between TGF-β/Smad and other signaling pathways. Cold Spring Harb Perspect Biol; 9(1):a022137

Lyons DC, Perry KJ, Lesoway MP, Henry JQ. 2012. Cleavage pattern and fate map of the mesentoblast, 4d, in the gastropod Crepidula: a hallmark of spiralian development. Evodevo; 3(1):21

Lyons DC, Henry JQ. 2014. Ins and outs of Spiralian gastrulation. Int J Dev Biol; 58(6-8):413-28

Lyons DC, Perry KJ, Henry JQ. 2015. Spiralian gastrulation: germ layer formation, morphogenesis, and fate of the blastopore in the slipper snail Crepidula fornicata. EvoDevo; 6:24

Maddison WP, Maddison DR. 2011. Mesquite: a modular system for evolutionary analysis. Version 2.75. http://mesquiteproject.org

Maldonado M. 2012. Metazoos. En: El árbol de la vida: sistemática y evolución de los seres vivos. P Vargas y R Zardoya (Eds.). Págs: 159-177. ISBN: 978-84-615-9740-6

Mallo M, Wellik DM, Deschamps J. 2010. Hox genes and regional patterning of the vertebrate body plan. Dev Biol; 344(1):7-15

Marasco D, Saporito A, Ponticelli S, Chambery A, De Falco S, Pedone C, Minchiotti G, Ruvo M. 2006. Chemical synthesis of mouse cripto CFC variants. Proteins; 64(3):779-788

Martín-Durán JM, Vellutini BC, Hejnol A. 2016. Embryonic chirality and the evolution of spiralian left-right asymmetries. Philos Trans R Soc Lond B Biol Sci; 317(1710)

Martín-Durán JM, Ryan JF, Vellutini BC, Pang K, Hejnol A. 2017. Increased taxon sampling reveals thousands of hidden orthologs in flatworms. Genome Res;27:1-10.

Martindale MQ, Doe CQ, Morrill JB. 1985. The role of animal-vegetal interaction with respect to the determination of dorsoventral polarity in the equal-cleaving spiralian,Lymnaea palustris. Wilhelm Roux's Archiv; 194(5): 28-295

Martindale MQ. 1986. The organizing role of the D quadrant in an equal-cleaving spiralian, Lymnaea palustris as studied by UV laser deletions of macromeres at intervals between third and fourth quartet formation. Inter J Invert Repro Dev; 9:229–242

Massagué J. 1998. TGF-beta signal transduction. Annu Rev Biochem; 67:753-91

Massagué J. 2012. TGF-β signalling in context. Nat Rev Mol Cell Biol; 13(10):616-630

McDowell G, Rajadurai S, Levin M. 2016. From cytoskeletal dynamics to organ asymmetry: a nonlinear, regulative pathway underlies left-right patterning. Philos Trans R Soc Lond B Biol Sci; 371(1710):pii20150409

McGrath J, Somlo S, Makova S, Tian X, Brueckner M. 2003. Two populations of node monocilia initiate left–right asymmetry in the mouse. Cell; 114(1):61-73

Meno C, Saijoh Y, Fujii H, Ikeda M, Yokoyama T, Yokoyama M, Toyoda Y, Hamada H. 1996. Left-right asymmetric expression of the TGF beta-family member lefty in mouse embryos. Nature; 381(6578):151-155

Meyers EN, Martin GR. 1999. Differences in left-right axis pathways in mouse and chick: functions of FGF8 and SHH. Science; 285(5426):403-406

Minchiotti G, Parisi S, Lihuori G, Signore M, Lania G, Adamson ED, Lago CT, Persico MG. 2000. Membrane-anchorage of Cripto protein by glycosylphosphatidylinositol and its distribution during early mouse development. Mech Dev; 90(2):133-142

Minchiotti G, Manco G, Parisi S, Lago CT, Rosa F, Persico MG. 2001. Structure-function analysis of the EGF-CFC family member Cripto identifies residues essential for nodal signalling. Development; 128(22):4501–4510

Minchiotti G, Parisi S, Liguori GL, D'Andrea D, Persico MG. 2002. Role of the EGF-CFC gene cripto in cell differentiation and embryo development. Gene; 287(1-2):33-37

Mochizuki T, Saijoh Y, Tsuchiva K, Shirayoshi Y, Takai S, Taya C, Yonekawa H, Yamada K, Nihei H, Nakatsuji N, Overbeek PA, Hamada H, Yokoyama T. 1998. Cloning of inv, a gene that controls left/right asymmetry and kidney development. Nature; 395(6698):177-181

Monteiro A. 2015. Origin, development, and evolution of butterfly eyespots. Annu rev Entomol; 60:253-271

Morgan D, Turnpenny L, Goodship J, Dai W, Majumder K, Matthews L, Gardner A, Schuster G, Vien L, Harrison W, et al. 1998. Inversin, a novel gene in the vertebrate left-right axis pathway, is partially deleted in the inv mouse. Nat Genet; 20(2):149-156

Morgan D, Goodship J, Essner JJ, Vogan KJ, Turnpenny L, Yost HJ, Tabin CJ, Strachan T. 2002. The left-right determinant inversin has highly conserved ankyrin repeat and IQ domains and interacts with calmodulin. Hum gen 110(4): 377-384

Morokuma J, Ueno M, Kawanishi H, Saiga H, Nishida H. 2002. HrNodal, the ascidian nodal-related gene, is expressed in the left side of the epidermis, and lies upstream of HrPitx. Dev Genes Evol; 212(9):439-446

Müller GB. 2007. Evo–devo: extending the evolutionary synthesis. Nat Rev Gen; 8(12):943-949

Naganathan SR, Fürthauer S, Nishikawa M, Jülicher F, Grill SW. 2014. Active torque generation by the actinomyosin cell cortex drives left-right symmetry breaking. Elife; (3):e04165

Naganathan SR, Middelkoop TC, Fürthauer S, Grill SW. 2016. Actomyosin-driven left-right asymmetry: from molecular torques to chiral self organization. Curr Opin Cell Biol; 38:24-30

Nagaoka T, Karasawa H, Turbyville T, Rangel MC, Castro NP, Gonzales M, Baker A, Seno M, Lockett S, Greer YE, Rubin JS, Salomon DS; Bianco C. 2013. Cripto-1 enhances the canonical Wnt/β-catenin signaling pathway by binding to LRP5 and LRP6 co-receptors. Cell Signal; 25(1):178-189

Namwanje M, Brown CW. 2016. Activins and Inhibins: Roles in Development, Physiology, and Disease. Cold Spring Harb Perspect Biol; 8(7):pii:a021881

Nishide K, Mugitani M, Kumano G, Nishida H. 2012. Neurula rotation determines left-right asymmetry in ascidian tadpole larvae. Development; 139(8):1467-1475

Nogi T, Yuan YE, Sorocco D, Perez-Tomas R, Levin M. 2005. Eye regeneration assay reveals an invariant functional left-right asymmetry in the early bilaterian, Dugesia japonica. Laterality; 10(3):193-205

Nonaka S, Tanaka Y, Okada Y, Takeda S, Harada A, Kanai Y, Kido M, Hirokawa N. 1998. Randomization of left-right asymmetry due to loss of nodal cilia generating leftward flow of extraembryonic fluid in mice lacking KIF3B motor protein. Cell; 95(6):829-837

Nosenko T, Schreiber F, Adamska M, Adamski M, Eitel M, Hammel J, Maldonado M, Müller WEG, Nickel M, Schierwater B, Vacelet J, Wiens M, Wörheide G. 2013. Deep metazoan phylogeny: When different genes tell different stories. Mol Phy and Evol; 67(1):223-233

Nürnberger J, Bacallao RL, Phillips CL. 2002. Inversin forms a complex with catenins and N-cadherins in polarized epithelial cells. Mol Biol Cell; 13(9):3096-3106

Nürnberger J, Kribben A, Opazo-Sáez A, Heusch G, Philipp T, Phillips CL. 2004. The Invs gene encodes a microtubule-associated protein. J Am Soc Nephrol; 15(7):1700-1710

Nürnberger J, Kavapurackal R, Zhang SJ, Opazo-Saez A, Heusch G, Phillipp T, Pietruck F, Kribben A. 2006. Differential tissue distribution of the Invs gene product inversin. Cell Tissue Res, 323(1): 147-155

Ohguro Y, Takata H, Kominami T. 2011. Involvement of Delta and Nodal signals in the specification process of five types of secondary mesenchyme cells in embryo of the sea urchin, Hemicentrotus pulcherrimus. Dev Growth Differ; 53(1):110-123

Ohno S. 1970. Evolution by gene duplication. Libro ISBN: 978-3-642-86661-6

Okada Y, Nonaka S, Tanaka Y, Saijoh Y, Hamada H, Hirokawa N. 1999. Abnormal nodal flow precedes situs inversus in iv and inv mice. Mol Cell; 4(4):459-468

Olson CM. 1994. Understanding and Evaluating a Meta-analysis. Acad Emerg Med; 1(4):392-398

Otto EA, Schermer B, Obara T, O'Toole JF, Hiller KS, Mueller AM, Ruf RG, Hoefele J, Beekmann F, Landau D, et al. 2003. Mutations in INVS encoding inversin cause nephronophthisis type 2, linking renal cystic disease to the function of primary cilia and left-right axis determination. Nature genetics; 34(4):413-420

Perry KJ & Henry JQ. 2015. CRISPR/Cas9-mediated genome modification in the mollusc, Crepidula fornicata. Genesis; 53(2):237-244

Perry KJ, Lyons DC, Truchado-Garcia M, Fischer AH, Helfrich LW, Johansson KB, Diamond JC, Grande C, Henry JQ. 2015. Deployment of regulatory genes during gastrulation and germ layer specification in a model

spiralian mollusc Crepidula. Dev Dyn; 244(10):1215-1248

Philippe H, Brinkmann H, Copley RR, Moroz LL, Nakano H, Poustka AJ, Wallberg A, Peterson KJ, Telford MJ. 2011. Acoelomorph flatworms are deuterostomes related to Xenoturbella. Nature; 470(7333):255-258

Piatek MJ, Henderson V, Fearn A, Chaudhry B, Werner A. 2017. Ectopically expressed Slc34a2a sense-antisense transcripts cause a cerebellar phenotype in zebrafish embryos depending on RNA complementarity and Dicer. PLoS One; 12(5):e0178219

Pick L. 2016. Hox genes, evo-devo, and the case of the ftz gene. Chromosome; 125(3):535-551

Qiu D, Cheng SM, Wozniak L, McSweeney M, Perrone E, Levin M. 2005. Localization and loss-of-function implicates ciliary proteins in early, cytoplasmic roles in left-right asymmetry. Dev Dyn; 234(1):176-189

Rankin CT, Bunton T, Lawler AM, Lee SJ. 2000. Regulation of left-right patterning in mice by growth/differentiation factor-1. Nat Genet; 24(3):262-265

Ravisankar V, Singh TP, Manoj N. 2011. Molecular evolution of the EGF-CFC protein family. Gene; 482(1-2):43-50

Raya A, Kawakami Y, Rodríguez-Esteban C,Ibañes M, Rasskin-Gutman D, Rodríguez-León J, Büscher D, Feijó JA, Izpisúa Belmonte JC.. 2004. Notch activity acts as a sensor for extracellular calcium during vertebrate left–right determination. Nature; 427:121-128

Raya A, Izpisúa-Belmonte JC. 2006. Left-right asymmetry in the vertebrate embryo: from early information to higher-level integration. Nat Rev Genet; 7(4):283-293

Reissmann E, Jörnvall H, Blokzijl A, Andersson O, Chang C, Minchiotti G, Persico MG, Ibáñez CF, Brivanlou AH. 2001. The orphan receptor ALK7 and the Activin receptor ALK4 mediate signaling by Nodal proteins during vertebrate development. Genes Dev; 15(15):2010-2022

Rincón-Limas DE, Lu CH, Canal I, Calleja M, Rodríguez-Esteban C, Izpisúa-Belmonte JC, Botas J. 1999. Conservation of the expression and function of apterous orthologs in Drosophila and mammals. PNAS; 96(5):2165-2170

Romano V, Raimondo D, Calvanese L, D'Auria G, Tramontano A, Falcigno L. 2012. Toward a better understanding of the interaction between TGF-β family members and their ALK receptors. J Mol Model; 18(8):3617-3625

Ronshaugen M, McGinnis N, McGinnis W. 2002. Hox protein mutation and macroevolution of the insect body plan. Nature; 415 (6874):914-917

Rosenberg MI, Lynch JA, Desplan C. 2009. Heads and tails: evolution of antero-posterior patterning in insects. Biochim Biophys Acta; 1789(4):333-342

Röttinger E, DuBuc TQ, Amiel AR, Martindale MQ. 2015. Nodal signaling is required for mesodermal and ventral but not for dorsal fates in the indirect developing hemichordate, Ptychodera flava. Biol Open; 4(7):830-842

Rutenberg J, Cheng SM, Levin M. 2002. Early embryonic expression of ion channels and pumps in chick and Xenopus development. Dev Dyn; 225(4):469-484

Ryan JF, Pang K, Schnitzler CE, Nguyen AD, Moreland RT, Simmons DK, Koch BJ, Francis WR, Havlak P; NISC Comparative Sequencing Program, et al. 2013. The genome of the ctenophore Mnemiopsis leidyu and its

implications for cell type evolution. Science; 342(6164):1242592

Saijoh Y, Oki S, Ohishi S, Hamada H. 2003. Left-right patterning of the mouse lateral plate requires nodal produced in the node. Dev Biol; 256(1):160-172

Salomon DS, Bianco C, De Santis M. 1999. Cripto: a novel epidermal growth factor (EGF)-related peptide in mammary gland development and neoplasia. Bioessays; 21(1):61-70

Sanetra M, Begemann G, Becker MB, Meyer A. 2005. Conservation and co-option in developmental programmes: the importance of homology relationships. Front Zool; 2:15

Schier AF, Neuhauss SC, Helde KA, Talbot WS, Driever W. 1997. The one-eyed pinhead gene functions in mesoderm and endoderm formation in zebrafish and interacts with no tail. Development; 124(2):327-342

Schier AF. 2003. Nodal signaling in vertebrate development. Annu Rev Cell Dev Biol; 19:589–621

Schier AF. 2009. Nodal morphogens. Cold Spring Harb Perspect Biol; 1(5):a003459

Schiffer SG, Foley S, Kaffashan A, Hronowski X, Zichittella AE, Yeo CY, Miatkowski K, Adkins HB, Damon B, Whitman M, Salomon D, Sanicola M, Williams KP. 2001. Fucosylation of Cripto is required for its ability to facilitate nodal signaling. J Biol Chem; 276:37769–37778

Schierwater B, Eitel M, Jakob W, Osigus HJ, Hadrys H, Dellaporta SL, Kolokotronis SO, Desalle R. 2009. Concatenated analysis sheds light on early metazoan evolution and fuelsa modern “urmetazoon” hypothesis. PLoS Biol; 7(1):e20

Schumacher JA, Hsieh YW, Chen S, Pirri JK, Alkema MJ, Li WH, Chang C, Chuang CF. 2012. Intercellular calcium signalling in a gap junction-coupled cell network establishes asymmetric neuronal fates in C. elegans. Development; 139(22):4191-4201

Shaolin S, Ge C, Luo Y, Hou X, Haltiwanger RS, Stanley P. 2007. The threonine that carries fucose, but not fucose, is required for Cripto to facilitate nodal signaling. J Biol Chem; 282:20133–20141

Shen MM, Wang H, Leder P. 1997. Adifferential display strategy identifies Cryptic, a novel EGF-related gene expressed in the axial and lateral mesoderm during mouse gastrulation. Development;124(2):429-442

Shen MM. 2007. Nodal signaling: developmental roles and regulation. Development; 134(6):1023–1034

Shen MM, Schier AF. 2000. The EGF-CFC gene family in vertebrate development. Trends Genet; 16(7):303-309

Shen XX, Hittinger CT, Rokas A. 2017. Contentious relationships in phylogenomic studies can be driven by a handful of genes. Nat Eco Evo; (1)

Shi S, Ge C, Luo Y, Hou X, Haltiwanger RS, Stanley P. 2007. The threonine that carries fucose, but not fucose, is required for Cripto to facilitate Nodal signaling. J Biol Chem; 282(28):20133-20141

Shi C, Zhang L, Qin C. 2017. Long non-coding RNAs in brain development, synaptic biology, and Alzheimer’s disease. Brain Res Bull; 132:160-169

Shiba D, Yamaoka Y, Hagiwara H, Takamatsu T; Hamada H, Yokoyama T. 2008. Localization of Inv in a distinctive intraciliar compartment requires the C-terminal ninein-homolog-containing region. J Cell Sci; 122(1):44-

Shibazaki Y, Shimizu M, Kuroda R. 2004. Body handedness is directed by genetically determined cytoskeletal dynamics in the early embryo. Curr Biol; 14(16):1462-1467

Shimadera S, Iwai N, Deguchi E, Kimura O, Fumino S, Yokoyama T. 2007. The inv mouse as an experimental model for biliary atresia. J Pediatr Surg; 42(9):1555-1560

Shimeld SM, Levin M. 2006. Evidence for the regulation of left-right asymmetry in Ciona intestinalis by ion flux. Dev Dyn; 235(6): 1543-1553

Shinohara K, Hamada H. 2017. Cilia in left-right symmetry breaking. Cold Spring Harb Perspect Biol. pii:a028282

Shiratori H, Hamada H. 2006. The left-right axis in the mouse: from origin to morphology. Development; 133:2095-2104

Simons M, Gloy J, Ganner A, Bullerkotte A, Bashkurov M, Krönig C, Schermer B, Benzing T, Cabello OA, Jenny A, et al. 2005. Inversin, the gene product mutated in nephronophthisis type II, functions as a molecular switch between Wnt signaling pathways. Nat Genet; 37(5):537-543

Spéder P, Adám G, Noselli S. 2006. Type ID unconventional myosin controls left-right asymmetry in Drosophila. Nature, 440(7085):803-807

Spéder P, Noselli S. 2007. Left-right asymmetry: class I myosins show the direction. Curr Opin Cell Biol. 19(1):82-87

Spéder P, Petzoldt A, Suzanne M, Noselli S. 2007. Strategies to establish left/right asymmetry in vertebrates and invertebrates. Curr Opin Genet Dev; 17(4):351-358

Srivastava M, Begovic E, Chapman J, Putnam NH, Hellsten U, Kawashima T, Kuo A, Mitros T, Salamov A, Carpenter ML, et al. 2008. The Trichoplax genome and the nature of placozoans. Nature; 454(7207):955-960

Strähle U, Jesuthasan S, Blader P, García-Villalba P, Hatta K, Ingham PW. 1997. one-eyed pinhead is required for development of the ventral midline of the zebrafish (Danio rerio) neural tube. Genes Funct; 1(2):131-148

Su WY, Xiong H, Fang JY. 2010. Natural antisense transcripts regulate gene expresión in an epigenetic manner. Biochem Biophys Res Commun; 396(2):177-181

Sun Y, Li D, Zhang R, Peng S, Zhang G, Yang T, Qian A. 2017. Strategies to identify natural antisense transcripts. Biochimie; 132:131-151

Takeda S, Yonekawa Y, Tanaka Y, Okada Y, Nonaka S, Hirokawa N. 1999. Left-right asymmetry and kinesin superfamily protein KIF3A: a new insights in determination of laterality and mesoderm induction by kif3A-/- mice analysis. J Cell Biol; 145(4):825-836

Tee YH, Shemesh T, Thiagarajan V, Hariadi RF, Anderson KL, Page C, Volkmann N, Hanein D, Sivaramakrishnan S, Kozlov MM, Bershadsky AD. 2015. Cellular chirality arising from the self-organization of the actin cytoskeleton. Nat Cell Biol; 17(4):445-457

Telford MJ, Budd GE, Philippe H. 2015. Phylogenomic insights into animal evolution. Curr Biol; 25(19):876-887

Tisler M, Wetzel F, Mantino S, Kremnyov S, Thumberger T, Schweickert A, Blum M, Vick P. 2016. Cilia are required for asymmetric nodal induction in the sea urchin embryo. BMC Dev Biol; 16(1):28

True JR, Carroll SB. 2002. Gene co-option in physiological and morphological evolution. Annu Rev Cell Dev Biol; 18:53-80

Ulloa L, Creemers JW, Roy S, Liu S, Mason J, Tabibzadeh S. 2001. Lefty proteins exhibit unique processing and activate the MAPK pathway. J Biol Chem; 276:21387-21396

Unolt M, Putotto C, Silvestri LM, Marino D, Scarabotti A, Massaccesi V, Caiaro A, Versacci P, Marino B. 2013. Transposition of Great Arteries: New Insights into the Pathogenesis. Front Pediatr; 1:11

Van den Biggelaar JA, Guerrier P. 1979. Dorsoventral polarity and mesentoblast determination as concomitant results of cellular interactions in the mollusk Patella vulgata. Dev Biol; 68(2):462-471

Verdonk NH, Van den Biggelaar JAM. 1983. Early development and the formation of the germ layers. En: The Mollusca. NH Verdonk, JAM Van den Biggelaar, y AS Tompa (Eds.). Págs: 91-122. Academic Press, New York.

Watanabe D, Saijoh Y, Nonaka S, Sasaki G, Ikawa Y, Yokoyama T, Hamada H. 2003. The left-right determinant Inversin in a component a node monocilia and other 9+0 cilia. Development; 130(9):1725-1734

Watanabe K, Hamada S, Bianco C, Mancino M, Nagaoka T, Gonzales M, Bailly V, Strizzi L, Salomon DS. 2007. Requirement of glycosylphosphatidylinositol anchor of Cripto-1 for trans activity as a Nodal co-receptor. J Biol Chem; 282(49):35772-3586

Watanabe K, Nagaoka T, Strizzi L, Mancino M, Gonzales M, Bianco C, Salomon DS. 2008. Characterization of the glycosylphosphatidylinositol-anchor signal sequence of human Cryptic with a hydrophilic extension. Biochim Biophys Acta; 1778(12):2671-2681

Watanabe K, Nagaoka T, Lee JM, Bianco C, Gonzales M, Castro N, Rangel MC, Sakamoto K, Sun Y, Callahan R, salomon DS. 2009. Enhancement of Notch receptor maturation and signaling sensitivity by Cripto-1. J Cell Biol; 187(3):343-353

Walentek P, Beyer T, Thumberger T, Schweickert A, Blum M. 2012. ATP4a is required for Wnt-dependent Foxj1 expression and leftward flow in Xenopus left-right development. Cell Rep; 1(5):516-527

Wechselberger C, Strizzi L, Kenney N, Hirota M, Sun Y, Ebert A, Orozco O, Bianco C, Khan NI, Wallace-Jones B, et al. 2005. Human Cripto-1 overexpression in the mouse mammary gland results in the development of hyperplasia and adenocarcinoma. Oncogene; 24(25):4094-4105

Werner A, Swan D. 2010. What are natural antisense transcripts good for? Biochem Soc Trans; 38(4):1144-1149

Werner ME, Ward HH, Phillips CL, Miller C, Gattone VH, Bacallao RL. 2013. Inversin modulates the cortical actin network during mitosis. Am J Physiol Cell Physiol; 305(1):36-47

Wlizla M. 2011. Evolution of nodal signaling in deuterostomes: insights from Saccoglossus kowalevskii. Doctoral dissertation. University of Chicago, Chicago, EEUU. UMI Number: 3487671

Xu J, Keymeulen V, Wakida NM, Carlton P, Berns MW, Bourne HR. 2007. Polarity reveals intrinsic cell chirality. PNAS; 104(22):9296-9300

Xu P, Liu J, Derynck R. 2012. Post-translational regulation of TGF-β receptor and Smad signaling. FEBS Lett;

586(14):1871-1884

Yan YT, Gritsman K, Ding J, Burdine RD, Corrales JD, Price SM, Talbot WS, Schier AF, Shen MM. 1999. Conserved requirement for EGF-CFC genes in vertebrate left-right axis formation. Genes Dev; 13(19):2527-2537

Yan YT, Liu JJ, Luo YEC, Haltiwanger RS, Abate-Shen C, Shen MM. 2002. Dual roles of Cripto as a ligand and coreceptor in the nodal signaling pathway. Mol Cell Biol; 22:4439–4449

Yasuhiko Y, Imai F, Ookubo K, Takakuwa Y, Shiokawa K, Yakoyama, T. 2001. Calmodulin binds to inv protein: implication for the regulation of inv function. Dev Growth Differ; 43:671-681

Yasuhiko Y, Shiokawa K, Mochizuki T, Asashima M, Yokoyama T. 2006. Isolation and characterization of Xenopus laevis homologs of the mouse inv gene and functional analysis of the conserved calmodulin binding sites. Cell Research; 16:337-346

Yeo C, Whitman M. 2001. Nodal signals to Smads through Cripto-dependent and Cripto-independent mechanisms. Mol Cell; 7(5):949-957

Yokoyama T, Copeland NG, Jenkins NA, Montgomery CA, Elder FF, Overbeek PA. 1993. Reversal of left-right asymmetry: a situs inversus mutation. Science; 260(5108):679-682

Yoshida K, Saiga H. 2008. Left-right asymmetric expression of Pitx is regulated by the asymmetric Nodal signaling through an intronic enhancer in Ciona intestinalis. Dev Genes Evol; 218(7):353-360

Yu JK, Holland LZ, Holland ND. 2002. An amphioxus nodal gene (AmphiNodal) with early symmetrical expression in the organizer and mesoderm and later asymmetrical expression associated with left-right axis formation. Evol Dev; 4(6):418-425

Zhang J, Talbot WS, Schier AF. 1998. Positional cloning identifies zebrafish one-eyed pinhead as a permissive EGF-related ligand required during gastrulation. Cell; 92(2):241-251

Zhang J. 2003. Evolution by gene duplication: an update. Trends Eco Evo; 18(6): 292-298

Zecca M, Basler K, Struhl G. 1995. Sequential organizing activities of engrailed, hedgehog and decapentaplegic in the Drosophila wing. Development; 121(8):2265-2278

Apéndices

Apéndice 1: Créditos de las imágenes utilizadas en esta tesis

https://en.wikipedia.org/wiki/Hexactinellid
http://www.eldiario.es/politica/carabela-portuguesa-aguaviva-temporales-Canarias_0_229227159.html
https://www.vistaalmar.es/images/ampliadas2/medusa-peine-ctenoforo.jpg
https://es.pinterest.com/pin/476185360583792575/
https://es.pinterest.com/pin/208361920234098321/
https://es.pinterest.com/pin/4644405837971435/
http://serc.carleton.edu/microbelife/topics/tardigrade/index.html
http://knowledgebase.lookseek.com/Starfish.html
http://www.underwaterkwaj.com/shell/cone/Conus-ammiralis.html
https://www.vistaalmar.es/ciencia-tecnologia/biologia/5346-cuando-la-sepia-contiene-la-respiracion-se-vuelve-casi-invisible-a-los-tiburones.html
http://www.thebiologistapprentice.com/blog/urocordados
http://beautifulnow.is/bnow/new-beautiful-purple-life-purple-sea-urchin-super-urchins-purple-tomato-indigo-rose-sugary-sky
https://www.vistaalmar.es/especies-marinas/general/4821-conoce-las-estrellas-reales-de-buscando-a-nemo.html
http://losdaimonions.foroactivo.com/t1022-rana-arboricola-de-ojos-rojos
https://es.pinterest.com/explore/azul-ruso/
http://alfonsosegura.com/fotografia/index.php/Colibries/Colibr-34

Apéndice 2: Bases de datos consultadas en las búsquedas de ortólogos en distintos Metazoa y en Coanoflagelata

Grupo	Phyla	Taxa	Especie	Fuente genómica
Coanoflagelado	Choanoflagelata		*Salpingoeca rosetta*	GenBank
Metazoos no bilaterales	Placozoa		*Trichoplax adherens*	http://genome.jgi.doe.gov/Triad1/Triad1.home.html
	Porifera		*Oscarella lobularis*	GenBank
			Amphimedon queenslandica	GenBank
	Cnidaria		*Oopsacas minuta*	GenBank
			Aurelia aurita	Genoma de la especie cortesía del Dr. Gold* (Caltech, , CA)
			Exaiptasia pallida	GenBank
	Ctenophora		*Mnemiopsis leidyi*	GenBank https://kona.nhgri.nih.gov/mnemiopsis
Bilateria	Xenacelomorpha	Xenoturbella	*Xenoturbella bocki*	GenBank; accession number: SRX1343818
		Acelomorpha	*Isodiametra pulchra*	GenBank; accession number: SRX1343817
			Convoluriloba macropiga	GenBank; accession number: SRX1343815
	Ecdysozoa	Priapulida	*Halicryptus spinulosus*	GenBank; accession number: SRX1343820
			Priapulus caudatus	GenBank; accession number: SRX507009
		Nematoda	*Caenorhabditis elegans*	GenBank
		Arthropoda	*Drosophila melanogaster*	GenBank http://flybase.org/
	Spiralia	Gastrotricha	*Lepidodermella squamata*	GenBank; accession number: SRX1000997
		Bryozoa	*Membranipora membranacea*	GenBank; accession number: SRX1121923
		Brachiopoda	*Novocrania anomala*	GenBank; accession number: SRX1343816
		Mollusca	*Lottia gigantea*	http://genome.jgi-psf.org/Lotgi1/Lotgi1.home.html
			Biomphalaria glabrata	Transcriptoma realizado y publicado durante esta tesis (apéndice 11)
			Crepidula fornicata	Dos transcriptomas realizados durante esta tesis (apéndice 10) Transcriptoma cortesía del Dr. Smith (MBL, Woods-Hole, MAS)
		Annelida	*Owenia fusiformis*	Transcriptoma de la especie cortesía del Dr. Harald Hausen
			Capitella teleta	http://genome.jgi.doe.gov/Capca1/Capca1.home.html
			Dinophilus gyrociliatus	GenBank; accession number: SRX2030658
		Nemertea	*Lineus ruber*	Transcriptoma cortesía del Dr. Hejnol (Sars Institute, Noruega)
	Deuterostomia	Hemichordata	*Saccoglossus kowalevskii*	GenBank
		Vertebrata	*Danio rerio*	GenBank https://zfin.org/
			Xenopus laevis	GenBank http://www.xenbase.org/entry/
			Gallus gallus	GenBank
			Ratus norvegicus	GenBank
			Mus musculus	GenBank

*Gold DA, Katsuki T, Yan X, Li Y, Regulski M, Ibberson D, Holstein TW, Steele RE, Jacobs DK, and Greenspan RJ. The genome of the jellyfish *Aurelia* and the evolution of animal complexity. En revisión.

Apéndice 3: Cebadores (primers) utilizados en esta tesis

NOMBRE	SECUENCIA	NOMBRE	SECUENCIA
Cripto-Crepi-F	GCGCTACTGTGAGTATGAGCTGTC	**InvCrepiBamH1-F**	GGGGGATCCATGGCAGCCACGGCCAATCCCGAT
Cripto-Crepi-R	GCATCTTCGGCGTCGTATTTACTG	**InvCrepiXho1-R**	CCCCTCGAGAGCTCTCCTAAGATTGTAGGAAAA
Cripto-Crepi-F-IN	GTATGAGCTGTCAGAGAAGCCA	**InvCrepi_1QAA_F**	GCCCTGACAGCTGCGCTGGCCTGGAGGAAGTTC
Cripto-Crepi-R-IN	TCGTATTTACTGGTGGTGGAGCCC	**InvCrepi_1QAA_R**	GGCAATTTGTCTTTCCCACGGCAGCC
Criptocrepi-F-2	GGGCGAAGGCAGGATGTAACTT	**INVs_crepi_F**	GGGCAACAAAAATCAGAACAACAACAAGTA
Criptocrepi-F-2-IN	ATGTAACTTTGTGTCGCTGCTTTG	**INVs_crepi_R**	TTCACCTGCTCTTTTTCTTGCAGCTTCTTTCTTCAGC
Criptocrepi-R-2	TGGCGAAGATGACCATTTCGTCT	**NodalCrepi_F2**	GTGGGTCATAATAAACACAGGTCT
Criptocrepi-R-2-IN	ACCATTTCGTCTTGGTAGTCCCAGA	**NodalCrepi_F2_in**	AAACACAGGTCTCGTGACAAGAGA
CriptoCrepi-F3	GAAGTTTTACGGGCGCTACTGTGA	**NodalCrepi_R2**	CAGCTATCATATTTTCGTGGTGGC
CriptoCrepi-F3-in	GCGCTACTGTGAGTATGAGCTGTC	**NodalCrepi_R2_in**	TTTCGTGGTGGCGAACCACAATCT
CriptoCrepi-R3	TGACACGAGTGTCCATGAGGTCCA	**Nodal2Crepi_F3**	ACTTGACCGTGGTAGGTCAAGACT
CriptoCrepi-R3-in	CCATGAGGTCCAGAGTATCGTAAC	**Nodal2Crepi_F3_in**	TAGGTCAAGACTCAGATGTCCAGT
CriptoCrepi-R4	TTCTATTATGTATACGCATATTCC	**Nodal2Crepi_R3**	CTACTGGTCCTAGAGGTAGTAGAT
CriptoCrepi-R4-in	TACGCATATTCCACTTCCAGTTTT	**Nodal2Crepi_R3_in**	GAGGTAGTAGATGCTCTCTTGTCA
CriptoCrepi-R5	AGTTATGTGCATAGGCAGCATGGT	**Nd2_Crepi_DCS_F**	GCCCCCAAGCACGAAAAATGCCAGATGTAC
CriptoCrepi-R5-in	ATCTTTAAGTGTAGTTATGTGCAT	**Nd2_Crepi_DCS_R**	ACGTTTAGATTCCCTACTGGTCCTAGAGG
CriptoCrepi-R6	CGTCACATCCATCGAAGGCACGTG	**Nod2Crepi_F4**	CACAAGAGGTGATGTCTCACAGAG
CriptoCrepi-R6-in	CGAAGGCACGTGGCAGGCAGGTCA	**Nod2Crepi_F4in**	TGTCTCACAGAGAGGTCTCTAGGT
CriptoCrepi-F4	GACGTGTCGTCTGTCAGGGGTCTG	**Nod2Crepi_R4**	GTGCACACATGGACGACAATCACC
CriptoCrepi-F4-in	GTCAGGGGTCTGATCACGCTGGTC	**Nod2Crepi_R4in**	ACGACAATCACCTACACCCACACT
CriptoCrepi-F5	GCGAGAGACTGTTGACAGTTATAA	**Nod2_Crepi_EcoI_F**	GGGGAATTCATGCCATTATTTTTCGGCAGCC
CriptoCrepi-F5-in	TGACAGTTATAAAGGATTTCTGTA	**Nod2_Crepi_SalI_R**	CCCGTCGACCCTACACCCACACTCATCAGC
CriptoCrepiBamH1-F	GGGGGATCCATGGACGTGTCGTCTGTCAGGGGT	**Cripto-Biom-F**	ATGTGGCTAGTAATGATAAGACTC
CriptoCrepiXhoI-R	CCCCTCGAGTGACACGAGTGTCCATGAGGTCCA	**Cripto-Biom-R**	GAGCACTCAATCACAAAAGTGAA
CR_L56T_F	GTCGTGCTGCCTGAACGGGGGGCACGTGTGTGCTGG	**Cripto-Biom-F-IN**	ATGATAAGACTCGCCAAGTCGGAC
CR_L56T_R	TTGGAGGGTCTAGGGTTCAAGTCTCCCAGCCTG	**Cripto-Biom-R-IN**	CACAAAAGTGAATGGTAACACTTG
CR_L56A_F	GTCGTGCTGCCTGAACGGGGGGCGCGTGTGTGCTGG	**Biom-INV-R**	CCGTTACCTCTAGCTGGGCCCCAT
CR_L56A_R	TTGGAGGGTCTAGGGTTCAAGTCTCCCAGCGTCTTC	**Biom-INV-R-in**	GCTGGGCCCCATGGTCCAGCAAGA
CR_S49A_F	CTTGAACCCTAGACCCTCCAAGGCGTGCTGCCTGA	**Biom-INV-R-2**	CAGTACATGAGTGGACTACGTCC
CR_S49A_R	TCTCCCAGCGGTCTTCTTCGACTTCGTCTGTGGGC	**Biom-INV-R-2-in**	GGACTACGTCCAAACTGATCACC
CR_S61A_F	CGGGGGCCTGTGTGCTGGACGCCTTCTGCCATT	**Biom-INV-F**	TGTACCGATGAAGCGGATCTAGAA
CR_S61A_R	TTCAGGCAGCACGACTTGGAGGGTCTAGGGTTC	**Biom-INV-F-in**	GAAGCGGATCTAGAAGGACGTACT
CR_DEGF_F	AAGCCATGTGGTCCTGTGCCTCATGGCAGCTGGGC	**InvBiom-F2**	GAGTCAACCTAGCAGTAATGGGA
CR_DEGF_R	GGGTTTGGGGTTCAAGTCTCCCAGCGTCTTCTTCG	**InvBiom-F2-in**	CAGTAATGGGAACTTAATGTTAG
Inv-Crepi-F	GGGCTGCACACAAGGGTAACTACC	**InvBiom-R3**	ACAGCAGCTGCATGGATCTGGGT
Inv-Crepi-R	TCTACCAGTAGCTTCACAGCATCC	**InvBiom-R3-in**	TGGATCTGGGTCACTACTACTTCTAA
Inv-Crepi-F-IN	AGGGTAACTACCGCCTTATGAAGC	**INVBiom_F3**	CCAGAAGATCCAGTCAGCTTTCCG
Inv-Crepi-R-IN	TTCACAGCATCCAGATGGCCCTTA	**INVBiom_F3_in**	GTCAGCTTTCCGAGGTTACAGAGT

Inv-Crepi-F-2-IN	GGATCAAGTCAAGCATACGCCCTT	**INVBiom_R4_in**	TGACAGTTTGATTAACGCCTAGTA
InvCrepiF3	CACGCAGGATAAAGAGGGCATGGC	**InvBiomF4**	GTAACTAATTGACTTAGGGGTTAT
InvCrepiF3-in	AGAGGGCATGGCTTGCCTCCACTG	**InvBiomF4-in**	CTTAGGGGTTATTTTGACTTGCAT
InvCrepi-R2	TGAGATAGTTGAAGTCGCCGCATC	**InvBiomR5**	GAACTTAATAGTTTAGCTCAGTCA
InvCrepi-R2-in	AGTCGCCGCATCCTGACTTCATTC	**InvBiomR5-in**	TTAGCTCAGTCACAATGCATTATC
InvCrepi_F4	AGAGTGTCCAAGGAATCAAGGAGT	**InvBiomBamH1-F**	GGGGGATCCATGGTTTCCCATTACTTGTGTCAT
InvCrepi-F4-in	GAATCAAGGAGTGTCCAAGGTCGT	**InvBiomXhoI-R**	CCCCTCGAGACACAGTTTATCACTTGCAATGAC
InvCrepi-R3	AGCTCTCCTAAGATTGTAGGAAAAT	**InvBiomWP_F**	TGCAGTAGCTGATGGAAATGAG
InvCrepi-R3-in	ATTGTAGGAAAATTGCGAATTCGC	**Oep_ZF_F**	GGCCAGCGGAATGACGAGTCAAC
InvCrepiF5	GCAGACCGTTGATCTAGCCAACTA	**Oep_ZF_R**	AAACTCATTTACAGCAGGCGGTG
InvCrepiF5-in	TCTAGCCAACTACGGAGGTTAAAG	**Oep_ZF_BI_F**	GGGGGATCCATGACGAGTCAACTGTTCGGGT
InvCrepiR5	CTTTGGCACAATACTGATCAACAA	**Oep_ZF_SI_R**	CCCGTCGACCAGCAGGCGGTGTAAAATAAAAGTG
InvCrepiR5-in	ACTGATCAACAAAGATGCAGGCAC	**sgRNA_R1**	AAAAGCACCGACTCGGTGCC
InvLCrepiWP_F	GCAGGCACAGACAAAGCAAGCATC	**Sp6_tr**	TGATTTAGGTGACACTATAG
InvCrepi_wp2_R	CATTCTGAGAAGATTCTGTCACC		

Apéndice 4: Diseño de los anticuerpos específicos para las especies *C. fornicata* y *B. glabrata*

A continuación se facilitan las secuencias aminoacídicas de las proteínas Inversin, Cripto y Nodal2 para las especies *C. fornicata* y *B. glabrata*, indicando las regiones elegidas para la generación de anticuerpos. El tipo de anticuerpo obtenido es de tipo policlonal, a través de la inoculación en conejo de péptidos sintéticos. En cada secuencia se aplica el siguiente código de color: Regiones seleccionadas para la generación de anticuerpos, zona de procesamiento alternativo, **zonas potenciales de unión a proteínas del complejo MHCII.**

Inversin *C. fornicata*

MAATANPDGPSASARQDSPQVPFASGTPRMENITQLHAAAVNGEKQA**LNRLIAGCGAGEMD**
VGDQFGRNPLMYCVLADRLDCAEL**LLKAGSQ**VNFRDKGGRTGLHWAAHKGNYRLMKLLL**SRG**
ADCREKDNEGQTPLHLCTRHKLPKCMAIL**LRQLLPGEIDNQDNNKRTALHWASSYG**NLEHV
KMLIKPDSNIGIPDIEGKTPLHWAASSRDPEAVQCVRLILETTPSVINWQDYEGRTALHLAVADG
NTSIVDLLTSVEKCNVSALDNVFRTPLHWAAVLGHSAIVGQLLQSNADSTSSDSNGATPLHYAAQ
NNFDETVNVFLSRKNTDEADLEGRTAFMWAAGKGADNVLRVFFHHGVDVQQEDKTGGTAWC
QDQWPRIKSSIRPCFGRVRWGTLTLHKPSSIMEHGVDVLDEDGRSPLHWAALGGHAYICQILIKY
GIDPNIRDHSGRSPLHCAAYGGYVNCMSALMEHQADPNTQDKEGMACLHWACSKGHLDAVKLL
VENDAFPNHMEFREERYTPLDYALMGEHHDVAQYMIEQGALSITGIQDIAATKIQSAFRGYRVRK
TFIERKKLLMKHEMLKKEAARKRAGEG**KKSEEQLRKESEDLAMEPLVAPLHHQLPQKTQTPS**
PPHQKSPSREQQQSPTKQLETKNPETESQKKQKETVQAQTKQASEQPQTKEQVSEPTQTK
QEETEQSEIAQQTTEEPHITQQESEQTQTTHEEQKTEQSQTVQNDAQQKHAEQKASEVTPL
NSEKSEEVQIQQQEAEHRQAKEQELKQPDTKQPEKAHQQTDLQEAEQTLTGQQKSEQQQV
QQEKTQQLKDSQEISSTSDTQQQAPASPVSSFGDKADPPKDAHVRHRTDSDKQKHAGRVS
KESRSVQGRKKKDKGPSAEEKVLQTKSDKANVKKGEKKRKGDIQKKQHAASVIQRAWRK
HRRSRQRHRQQPMVTESSQNEVRMRRLQLSQAAWERQIAALTIQLAWRKFYRRKL**LKSLNPN**
RRQLRMWDPDIIALKQQALIAEIYSEQLMVPFWHPTLKPSVRPFWFRLVPSPAAVSFNFAVDHY
HPLK**LWQGMRPHSWTGPVNGTVDLDFVPKSASSHRPERSANSQFSYNLRRA**

- Anticuerpos sintetizados:
  - CGPSASARQDSPQVPFASGT-amide
  - CGAGEMDVGDQFGRNPL-amide
  - Acetyl-LRQLLPGEIDNQDNNKRTAC-amide

Nodal2 *C. fornicata*

MPLFSAAIRAQGKCYFSAQCVGGEWLLHWLVLTMLLWTSSCPVMAI**TLPVALPPSHPNYYLTV**
VGQDSDVQSSTTPA**QTHNSNHVPVDVQGLGIREDQVMYHFYKNGRL**GSGYYRKASTVFMVR
LFNDLIRG**LAPVPSARSVSLGRNKDASDEDHPFTDTIRSFSA**GEVNFSKNKNATIHFFVPFLPP
NERLRSAEIRF**LRRPGARKKKRRNRYKLS**ITVLRGGRVVEKVL**LRER**PVTQQDNDVFDITRVL
DSWINSHHGNVSLKVRIPRNLENSILLSNASLNSTSLVALYLEDREFLNNMYESYTTDDLR**GNRS**
DARASSPGGGEVGHNNHSSRDKSAS**TTSSTSSESNRRNNSNNNKN**RWYRAPKHENCQMYDF
NVDFDFIGWGQWIIHPKNFNASFCYGECPSPVDSKYKPTNHAMLQTLMSQKKRRLAPPTCCVPI
NLNPLSMLYFEFDEIVVRHHENMIADECGCSW

- Anticuerpos sintetizados:
  - CPVTQQDNDVFDITRVLDS-amide
  - GGEVGHNNHSSRDKSAS-amide

EGF-CFC *C. fornicata*

MDVSSVRGLITLVIQLHTIAVFS**QAVSSQSPQTKSKKTLGD**LNPKPSKSCCLNGGLCVLGSFCHCP
KKF**YGRYCEYE**LSEKPCGPVPHGSWAKA**GCNLCRCFDG**HMTCLPR**AFDGCDDKPL**TGEQPQ
NIWDYQDEMV**IFANQ**GSTTSKYDAEDADYYDYYDDNGGGDGGG**GGVRHSAGARLVP**GLCCVA
VWSMLCVLVTILWTSWTLVS

- Anticuerpos sintetizados:
  HZ-Ahx-LSEKPCGPVPHGSWAKA-amide
  C-Ahx-TGEQPQNIWDYQDEMV-amide

Inversin *B. glabrata*

MVSHYLCHCCVSARF**LGL**YTEAQSETPRTPFTSKPEG**QRLEVVTQIHA**AAVNGEKQTLLRLIA**A
SASNGDLDVGDQFGRSPLMYC**VLGDRLECAEVL**LKTGAQVNLKDKGGRTALHWAAHKG**NL
RFMKLL**LSRGADCREKDNEGQTALHLC**TRHKLPKCMALL**LRQLTPGEIDDQDKNKRTALHW
AASY**GNLEHVKM**LIKQDSNIGIPDTEGKTPLHWAASTRDREAVSC**VKLILET**TPSVINWQDY
EGRTALHL**AVADGNESVVNVLTTVKNCNVSALDNVFRTPLHWAAVLGHSKIVSILLDSGADSSSS
DSNGATALHYAAQNNFSDTVSVFLKHNCTDEADLEGRTAFMWAAGKGADDVIRVFLEHKVDIH
QEDKTGGTSLHAAALSGHSSTVKLLLDHGAQLEVTDQAKHTPLFRACEMGHMDVVQTLIDCHA
RVDVLDQYGRSPLHWAALGGHAYICQLLIKYGVNPDVRDHSARAPLHCAAFGGFV**NCMSVLME
HEADPNSQDKEGLTCLHWACSKGH**LDAVKLLIEYNAYPNHMAFTEDRYTPLDYALMGEHHEV
AQYMIEQGALSITGIQDIAAQKIQSAFRGYRVRKAFIERKKLFMKHEKLKKEAARKRAGEEGKRT
MEKQ**TS**FGKSISFDKQTSFDKQT**SLDKQTSLDRQTPFDSQT**LAQADRLTVSHTTTSRNAEKLP
ENSSKTLQAECNHLHNLE**IASGHMTSPANIQSNIKNHTADSENYDQNSNTRNINEKRN**ADINS
PKQTQESQNDKQP**MKTELKVHKPRLRASQRRRLLQ**ADQNKGSLLGVNQTVIASDKLC

- Anticuerpos sintetizados:
  Acetyl-YTEAQSETPRTPFTSKPEGC-amide
  CFGKSISFDKQTSFDKQT-amide
  CADINSPKQTQESQNDKQP-amide

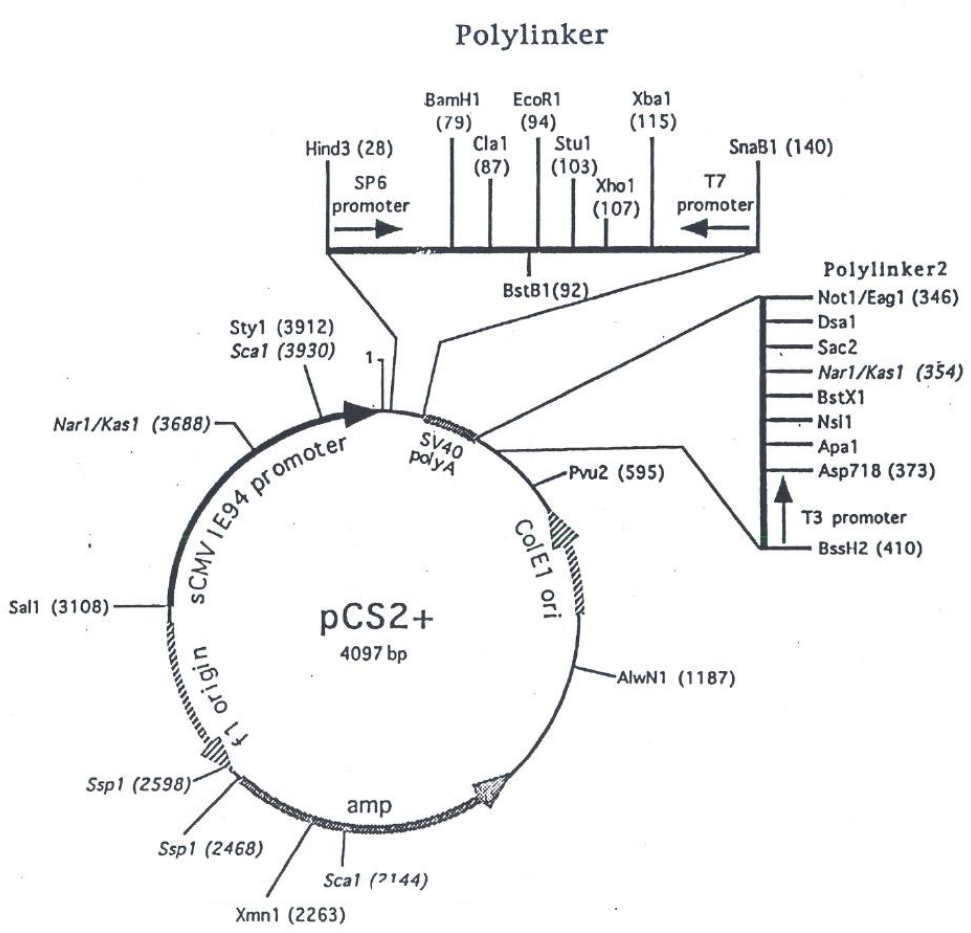
EGF-CFC *B. glabrata*

MANLFIMWLVM**IRLAKSDQGNYTDLHESITGNNVVSHPSL**LLSMLTPT**TETSKHQEEGQSSE
VKQKCCENGGLCVMGVFCHCTGK**YHGQRCEFERKPCGDIAHYQWARLGCNLCRCFDAKMHCL
TRVYEG**CDDKPPKTDVDT**SDYLNDREIFPADSDSDY**DNYGSGDYSYNEDDRVQRSGSSKL**IT
HLLTMLLVIVTVMFLIGSPWQVLPFTFVIECS

- Anticuerpos sintetizados:
  Acetyl-TETSKHQEEGQSSEVKQ-Ahx-C-amide
  CSDYLNDREIFPADSDSDY-amide

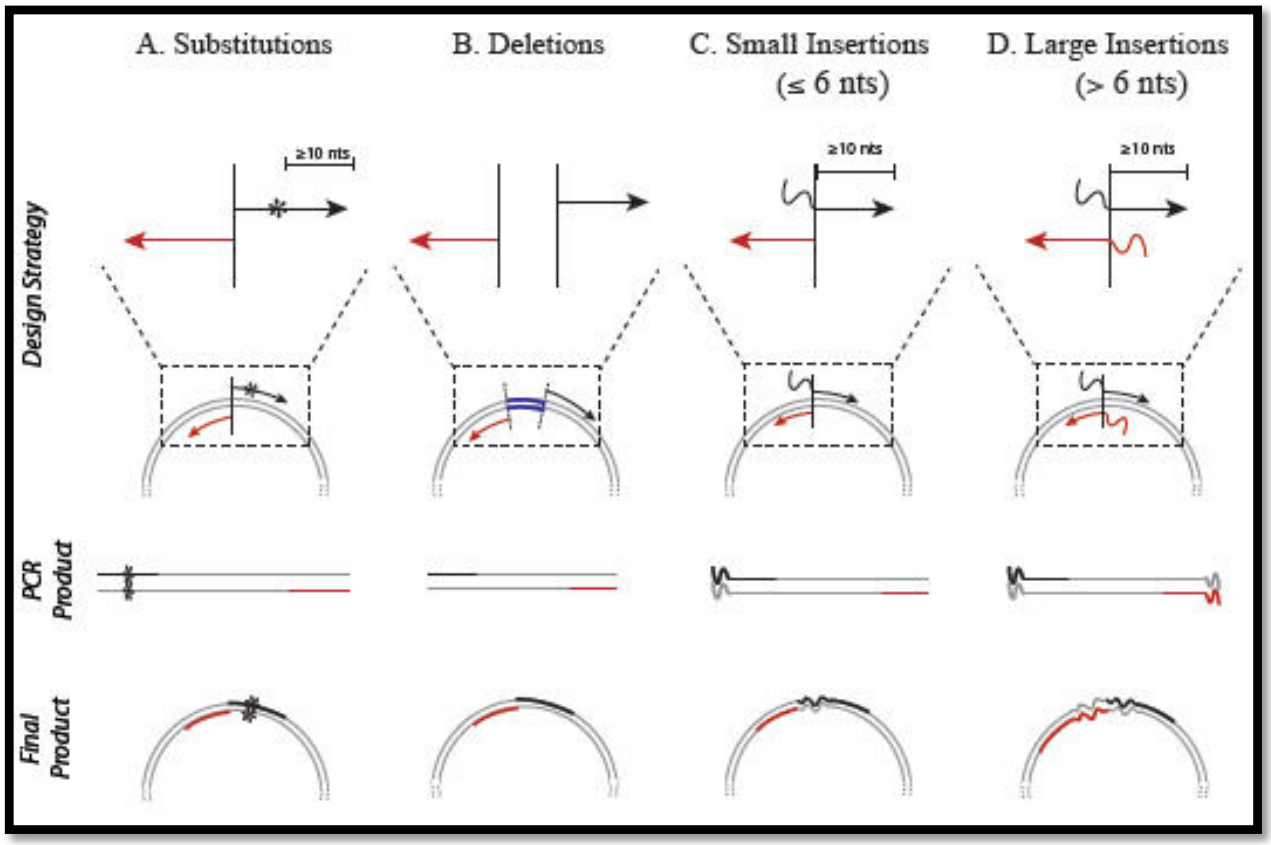
Apéndice 5: Vector utilizado en las construcciones para la generación de mRNA

Los sitios de clonaje utilizados fueron BamHI, EcoRI y XhoI, de acuerdo a las necesidades de cada caso. Se tuvo en cuenta para algunos clonajes que los extremos generados por las enzimas XhoI y SalI son cohesivos entre sí (isosquizómeros).

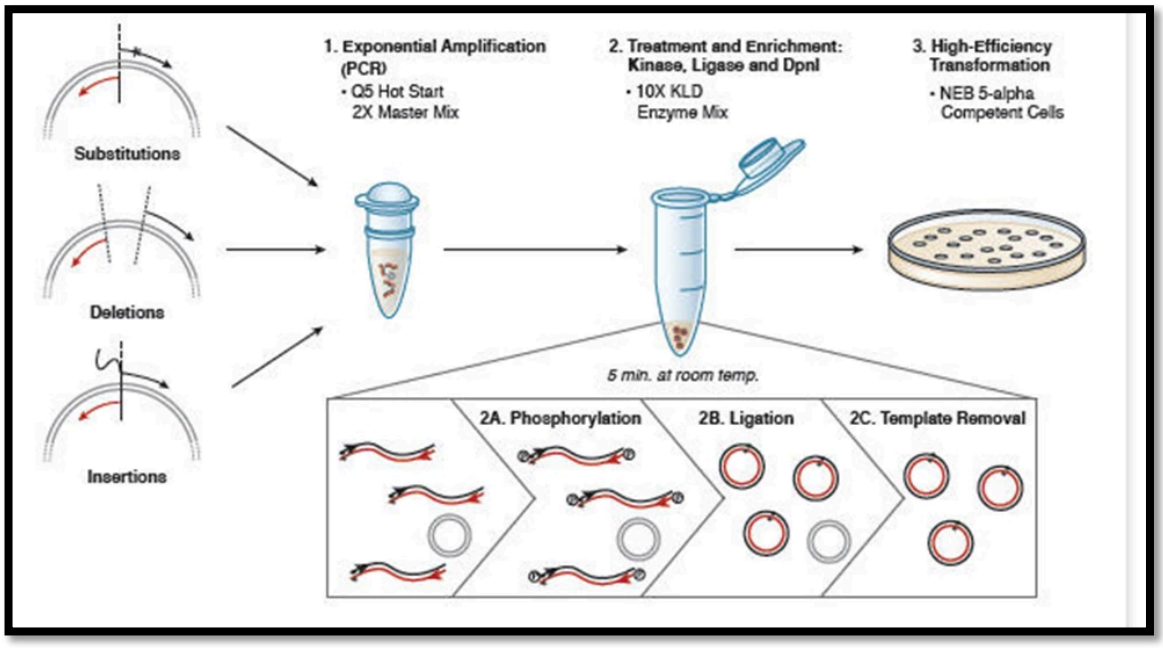


Apéndice 6: Estrategia seguida para la generación de construcciones mutantes

El kit indicado en materiales y métodos (3.16.) sugiere las siguientes estrategias en el diseño de cebadores, de acuerdo al tipo de mutación que vaya a generarse:



El kit proporciona un buffer que fosforila los extremos 5' de los productos de PCR obtenidos, recirculariza por ligación el vector y elimina el molde (wt). En esta tesis, se realizaron estos mismos pasos pero con productos adquiridos de forma independiente, como se indica en el apartado 3.16. de materiales y métodos. Así, aunque se suele aconsejar adquirir cebadores comerciales ya fosforilados y purificados por HPLC, en esta tesis se diseñaron de forma estándar y la fosforilación se realizó en el laboratorio con enzima PNK. Tras la PCR se añadió enzima dpnI de digestión rápida para eliminar el molde y tras limpiar el producto de PCR por columna, se llevó a cabo la ligación a través de la ligasa T4 (ligasa rápida de la casa New England, BioLabs). El producto obtenido puede utilizarse directamente para el protocolo de transformación, para el cual se utilizaron las bacterias NEB turbo de la misma casa comercial.



(Figuras obtenidas de https://www.neb.com/products/e0554-q5-site-directed-mutagenesis-kit#pd-description)

Apéndice 7: Secuencias aminoacídicas de la proteína Inversin identificadas en distintas especies

A continuación se facilitan las secuencias aminoacídicas de la proteína Inversin en las distintas especies en las que se ha identificado. Así mismo, se indican los dominios (potenciales) encontrados en cada una de ellas, de acuerdo al siguiente código de color: Dominios ankirina, dominios IQ, dominios de localización nuclear de cualquiera de los tipos (canónico monopartido, bipartido y no canónico) y dominios de destrucción D-box. En negrita se señala la secuencia delecionada por procesamiento alternativo en Inversin de *C. fornicata*. Todos los dominios se obtuvieron a partir de secuencias consenso proporcionadas en otras publicaciones (ver tabla 7). En cursiva se indican otros potenciales dominios obtenidos en el análisis *in silico*, que en principio se han descartado en base al alineamiento con el resto de secuencias.

> *Salpingoeca_rosetta*_A [hypothetical protein PTSG_12654; XP 004991715.1]
MSLDLAVPEARFDRAPSATSTRLGAADSLVNDIAGMTDTDLFKSISDVTYQ**EQHQVSFM**
HLTAMTNNLNDLQGLINQHPDYVNAVDSLDRTPLYFAILANQPKTCKLLIDAGALV
DVQDADGRTPGHWATFHGCPECLDVLLAAGATPDVRDKDGRTMLHWATAQKQI
KCLQHLLKQQDLDLDTRDSEEMTPLHWACFHNHAKHVQLLLNAKADITAVDMEG
KTALHWTSGNKDAAVVKVLLASKPELANQGDNEGRTPLHLCVGDASKAIAQAILKS
AKTDPTLVDHTGRTALHWAVALANDTMIGVLAKHRDSVRVQDEQGATPVHYAAQ
LDSEPCTKALLKSMPKDMVDVEDREQRTALFWAIAKDHLSTAELLLKAGANPNHV
DSTGRTILHISALSGKLGAVKLLTKYKADKEIQDSAQQTPLFIACENGFADLVSYLLD
NKANDAIIDAEGRNIVHAAAISGNKNVLAECLQLRPHDMNSVDERGETPVHYAAYF
GQFDCVQELLSSGAEPNCFDLEGVTPLHWACSQGFPDVVRELLNYNAYPNFTDRSE
DQLTPLDYALSGGFQECVDLLVEARGATGDELRAFAANT**IQAAFRVFQAKRIL**QNLR
KRKQSNAAIVIQAAARGFLERNRYKSMREDPSQRHQQQEPANKDDEERKRKEEEERRR
RQQEEEERRERE**RKAKAEQEEKLRKKK**QKQEQERR*RREEARKKKEAEERKKR*EQEDL
KRSQSLRLQEEEE**RRLDEKRRAEEEERRKR**ERELKQLREERQSSQRLMTALRGQQARQ
SVVRREQERLQLLRQKIRAATAIQQWWRVCKARQRQLEGYPTGLTPVESPNATRRK**KK**
PRSRTTSPQGGKGRSPRRQQRRANG**RSLLTHTRE**EDRKRLVAALT**IQLWWRRYLARK**
YAKRRQRTRVQLHGPAVSNRQQAKRLVTVYNTTGKAPSSTRVRQFQPKPQRAVRTHLL
PEPPSPVDLSWASAASTYFKPPNPGRRRPRGLHATRRRQRQPAYANMQDPEHMLTLPII
HGTQGSRKNDARLNQAVSLPPV

>*Salpingoeca_rosetta*_B [hypothetical protein PTSG_05580; XP 004993448.1]
MSEELPGLDPSSVLTPEEAAED**INPLHYAAAMGDKQSLQGILRMATADQVNMLDVYG**
RTALVYAVVSDKFTCAEMLIRKGISVDTPDGEGRTALHWACFHDKPKMVKLLLAN
GADTFIQDNDGRTCAHLACSAKSVKSFMAVVKARGADHDIDARDNWGMACIHWA
AYGGSVDMLKHILLKLGGLFLLRDQEGKSPVHWAAGSGNLSCLQFLITFWNSGKNR
HPDEVDVDMDVVNTVDLEGRAPLHIAVASGNEKVVNYLASLNTISIMQQDSLGRTS
LHWAAHANHAKAATILLRRGGTRQFFTADANGASPLHYAANEGNVDVVDALLAFD
DVHDVPDTYGRTALMVAAITGNIYIAQSIILSRKCDVLAKDHAGNTALHAAAYYGH
HDIIAMLIHAGAPAHDPDFTDKTALFHATEAGHAECVRVLMENGADLSFRDEEGRS
LLHWAAVSGHTDVVRVLLSSPEIDVDAKDTSLRTPLHNATFIGNAEIVGQLLDAGAN
VDAQDAMGIAPLHWAASRELVDVSKLLISRGATVNVSEYHEERLTPLDYANMTSNS
ELVKLLRKHNALTIEEIRDQAALQ**LQAWWKGHQARITL**LSLFKKHLGVEESKAAIAPK
KDFDGQRPVMKSRSS**KKAK**PLSPSAIVPKHDHLEPLRYSSTPSVTKEPLPQKRSGVTLPE

IKGASPSPSKASLPMPSSSSHHSTSSLSGATTTTSSTDRRARRRGKTKSNVVPLPSINRTP
GMEAQPLSSVGNNNKSSPTGRRGSVSKLTSRKQQRAAHKPGIPSLPIRNETQAQVQRER
SRVKDIRAKVMAARVIQRAFRRWLFHPERQVPSPVHSPQRPRVFEGTVGFSVDENDKR
QQIAALTIQLAWRQYIGRRAEREARQTSLNQSLARQMARQRQQKAVYDSTVPVVQWK
PTLAPIVRPLEFRSLPSPAVTSFTMALNTYTQPFADIQ*RHRLQRAEE*MLRLRQEDAERRL
MQTEERM**RSVLSLTAQ**RPSTALLYPVIESDHEDAVAYEQFQHDLHQSQQQQQQQQQQ
QQQQRRTRVDPHEIVRQYQAKRGIVA

>*Oscarella_lobularis*_1 [AMQ67871.1]
MDPRGGLQPSYATSGISFEPSSQPALPGVAEFYVSSRELLPSSSFVQPP**DTVHASPLHGV**
AVNGERDELLNLISSNAYDLDCGDQFGRTPLVYSVLGDRLDCAEILLKAGAQVNQAD
MDGRTALHWAAYQDNYRCLRMLLMKGANWKAKDNEGRTALHMATGHTSCKCLR
ELLKKMKQGSREVNDSDNEKMTALHWSAYNNNPEAIRFLIKAGADIVLPDVEGKTA
LHWTAGNNDGASVRAILEISPAVINLQDNEGRAALHLAVAEGNENTVGALTSVPGV
KCDVTSRDHMFRTPLHWAAVLGHTHIVGLLLERNADFSSTDGNGATALHYAAQNN
FGETVATMLSFPQVKDLHDNDGQTALMWAAGKGSYEVVRTMMERKLDVHACDN
QGGTALHAAAYAGHANIARVLLQHGAPVNTLDSRNHTPLFRACEMGHTEGVVTLIQ
GGAQVDIVDVDGRTPLHWAALGGHAYICLSLIQNGSTVDVKDKEGRSPLECASYGGF
VNCMAVLLENGCDVNTRDNEGITSIHWASASGHLEAIRLLFEYGANPNFMEVDGDR
LTPLDYAIINDHQEAAQCLIEQGALSITGIQDLAATA**IQSAWRGHVVRKAI**AK*KKEK*R
EAPSASSGGRPVSVGSRGDDE**KRKR**EESARKKAFERRAALERELADAERLDDKERQTEL
KRRQRQLEQEERLRERQQGSPRRYQSGTLKAHSIQDQQTLAKKE**RQRLMNIRQ**KTRAA
LVIQLAWRKYMRQ**KRRR**AALASKRKPSQMGSVEWRRQIAALT**IQLAWRQYMRRKLL**
QLYS**KRQK**ILHEWSPSVLAAKQRALVEKIYSQEVRIVQYRPPSPKPMVRPAYFRYIPSQA
ALFFNFAVDQYNPMTARLGMTRTFDMKTGSTRVRPFSLRRGTGGWKKEEPFRKGQSA
AFGGEPRSLASSVLF

>*Oscarella_lobularis*_2 [partial; AMQ6787.1]
[30]EKAYNMDEFLEPDKG**KLEGLHKAAAYGDDVAIMNIMKIEGQRHHHIMYTGLDQ**
LDAYGRSPLMYAVLSESLACMEVLLEFGALREQIDVNGRSALHYAAFYGRSRALQYL
LDKAANVTARDAFGRTALHWAAMINNPKCLSLLIQYATGHHLGLDWTDQEQVIPL
HVAAQFNRERNVALLMKAGARCDCVDMTGQTALHYCLGNTNVDCLNTICSGHPAIL
NVKDGSGATLLHWASGHGQAHIVAALVSFDGCDIESRDQWSCTPAHYACLYGNANC
LRILLAHDAHDYATDINGLTCLHYAVQENHLDCIHLLVSSPVASNLPDSEGHRPLTW
AAALGNHNAIKSLLQNSIFVSAIDQEEEEGRTALHVASLNGHALCVKALLENGSHIDA
LDAYGCSSLHLAASQSHINCIQELLSYGADVFACDANGQSVLHYAVHSRSVAAVELIL
QTNVNVDGVDSINSSPLHYATTENLTDIVTCLVRYGADPALLGVDGLSSLHIAVKNG
SDASLVTALLSGAPNPSVWNEDGETPLDFCYKFGQLQLVDTLINAGAMKFTDIQSFA
ATA**IQAAYRGWKARREL**KGEAKLARAALKIQSYYRGYRGR**KRYR**EVVRRDKAARVVQ
RNYRNFVLRKERKRELLRLRRKTEVTVLISEFHERLLIDRARVKAKPIVDRKGGEDSDKE
KAVSNQERLRMEMGLRIREIEAERQ**KKWREFLEEEETRRK**NKAIEKDVAYWNGIIRES
KERAHRAINDRVNRVKACQRAACIIQRWWRACTKGRNDRRSSRVIRIRGLMIGRRRAA
KV**IQRAWRDKKKRDAG**REKSRKQARAFNEETAPYIRPTLVQGKPRHKLGPLRIGSGVQ
DRSKTSYQAKKITSPTNTINKRGATVEFRPWTVLPLAKSLSASKRKEKVTLPRIDSNSTA
LVAQK

>*Amphimedon_queenslandica* [XP_019851880.1]
MADQMLRDGEAGE**EAIITTELNSAAYSGDKDLLLALLKEHDIDECDEQGRTPLMYAS**
MVNQYDIVDLLLKRKASSTIQDTLGQTPLHLAASNNAFECLKLLARSATELNIKDMD

GRTPLHLSSAKSNLRCLDFLIRKLRSSELNEGDNEKMTALHWCAYNNNSEGVRKLL
QSGADIVLTDIDGKTGLHWTASNEDDSTIKVLLDKAPTAINLKDKDERTVLHLSVA
AGNATVIMTLLNSSSVRCDVMAVDSDFRTPLHWAAVLGLSQIVGILMENGADPTAV
DATGATPLRYAAMKDHEDTVEVLLSFNSDDIPDNEGRTALMVAAEKGYYNIVKKMI
ERKVEINAQDNQGATALHLAAYSGHSDICQLFIQHNADINATDINGHSPLFRACECG
HNPVVLTLIQHKALVDLIDNEGRTCLHWAASGGHEFIVTTLIHSGLPVDILDNENRSS
LHCAAYNGSVGCCSELINANAFINQTDKDGVTPLHWACAGGSSDCVQFLLSKGANPN
AMDSSEQLTPLDYAILENHQELAQALIINGALTIASIQELAAIM**IQKVVRGYLARKRF**K
LLHAEKVKGEVKVEETGGTETEATDAVSSRPDSSHRRSLELQVAESERQAKEWALRSPS
TRALQSQLARKDRQRMTLHRKKTHAALVIQYAWRKY**KRRKHYLEIEDAKRK**ARQRLG
LDHWRKEIAALV**IQLAWRQYQRKQVL**KLAN**KRRR**ILHQWTPSVLAAKQRLLVEKVYG
QELKTSQYHPPKPRPMVRPAYLQLVASPAATSFNFAVNQYQTSFSPQSRVF

>*Aurelia_aurita* [transcriptoma no publicado; referencia en apéndice 2]
MPRTTLNRVEEEKGEEAKVISPEDGAN**LSPLHLAAIMGNKSKMAALLKEHKDDIDKG**
DKIGRTPLIYSILGDRYECAEALIKAGADINQTDVDERTALHWAAYQGNQKFVKLLL
DKGAQQGSQDKEGRTPLHLSLSHDNIKVMSRLLKDTDVDCRDGTGMTPLHFSVAY
ERVEHTKMLLHAKASPDCQDNDKKTVLHWAKSNKDTTIVSKLLDYERVMLILNSK
DSEGRTALHLMVAVDNRKLVRYLTSREDCLLSEQDHMQRTPLHLAAILGHTEIASV
LLARGCEPLKPDNNGVRPLHYAAQNNCGKTVDAMLQHDNVTDVPDDDGRTALM
WAAAKGANEALSVLIQHKQDILASDRSTATALHMACEGGHLQCVSLLLQNQADVNA
LDKQKYTPLFCACQMGQRDVVETLLLHGATVDLEDKEGRTPLHWAAVGGRTDICG
MLLDHGLDISRRDKLGRPALHYAAYGGFYSCMSLLLEFGADVDTRDKTGVTALHW
ACASGSLDAVKLLFRYNAFPNFVEMEGDKLTPLDYAIIREHQDVAQFMIEQGAFTIA
GIQDEAATV**IQATWRGYRLRKYF**QEQKELLILHERKRRETKKSYLTDNEELEFENDQD
FPEEGVEFSNSLNLLNGRDFADYFGTPASRLDEEVVNDKDNDGSSRDGSRLVKNEARDI
ENETLNIENEASDFENEAPNLGYEPQNDADSGQHELLDERRSVIGYDNEDVASDRGGSS
LGGSFILRENLEESQKDISTDQPQRHKSFEATEVSRKISNGNLKEDFVPKLTNPITKNDSG
KDCSRDHVQSRSERIFGLKSDELDSGSNTDESAGSGEN**KRLR**ERVNSIEGKDHSDISLVP
AQITSDVKHTTVAHDKRENVDFGETNFTKEQRPRVGSGRAMQLVEEAGQNRPNTSEKN
SYMQT**KRHK**ITKLERSARIDRSPSQSRMSVNDYNSESASDLKPWQIY**KRER**YRKNLIRR
KMESAVIIQRAFRTHILRHKPRFAEQMNEEEREKEAKAKEVVQDIAALV**IQLHWRKYL**
KRKLVQERRNESRRKDSADLRLLAARERKPKVILHPRSPPESTRRKSHHASKQRLKTAD
VNTTRTTPEQLELPNINN**RKVRIDETYSVKKRR**DPLNEVQPKSRPKTVGAVSKRGSRS
GQKISPPPTTKSHRNSDAHSKVASNAPQGSSTAFKMYMPQNWSSNLQKFPLRY

>*Exaiptasia_pallida* [KXJ12652.1]
MSQLSIRPISLHNSLSHSEQDEDNLSHLDDFDNPEG**LSNTSTLLHAAAVNGNKRMLQK**
LLQEVNSSDEINRGDQFGRTPLVYSVLGDRLDCAKLLILYGASVSMSDFDGRTPLHW
AAYQGSQKCVKLLLARGADWKAKDTKGRTPLHFAASHQSSKCLAVLLKKAPKGDV
HVDDPDNDSMTALHWGASYGNTEHVRMLLKAKASVKLLDLESKTPLHWATTNRN
PSVVKLILEQTTEIINWKDVEGRTALHLAVARQNEDVVKSLLVVEDCNPSLQDNMG
RTPLHWATVLENSRIVKLLLRRGADYKVADGNGSTVLHYAAQGDFAQTVEVLLSSC
DIKDTPDDKGRTALMWAAGKGASSSLDAMLRHHLDVHETDNAGRTALHIAAYTG
DVNCVEMLLKYGASVNALDQLNHTPLFRACEMGYADVVNLLKDNKAMLDLEDQD
GRSPLHWAALNGHVHICFILVTNNAQVNRGDRLGCTPLHCAAYGGFVNCMSLLIDN
GADVNKLDNEGTSAFHWSCSSGHLEAVKLLFRKKAFPNYTKLNEERLTPLDFAIIGD
HQDVAQYMIEQGGLTFRAIQDVAATT**IQSCWRGYRVRKPF**QK**RKTLYMRHE**QLK**R**
DMKRRNSAAKEEELKSHTVEAAMSASVEPQVPLWGVGRRQALCHSTSGGLFQLDLDL

QPEGKAHDYSDDDHDSYETVLDGETSDGEYDFLSTILEASFPSVNGSVFETMADRDGTIK
QNNRISLAAETIAENEDEKHEQESQVMEESQNSSSLDNRISNNPPDITSAERSNTSCSED
RELDDCTDFTPPPNVASPARSNTSCSEDREQDGRTDFIPPSNVATADRSNAFYSEDGELD
DCTDFIPPSNDASTERTNTSSRSEDRRLDGRTDFIPLRENKILLHSQTSENARKERKRRE
LIRAKMQAATTIQRWYRRNRISKQHEYXVAELQNDHEELIKQIAALT**IQLWWRKYLCR**
HYHHDNHEITSGHEFSRSETSKENYSQFKTSRTRKKDSQVRGFVQNSWKGSRSNRKNV
SSVTSIGSQRRSTLSSVSGRRSTRPSLSQKRVTNNKIRPSVGRMTKTGWGQDFDVNFSRK
LSKNSWKS

>*Oopsacas_minuta* [AMQ67864.1]
MSFTATNSSHRSISNLFSLEQN**TSPLHDAVISQDLPSLRQLLTSNSYDIDVLDECGRTP**
LIYSIFVGSFVCFELLLQSGSNIYILDFDSKSPLHWACQLGRLRIVKYLISKGCDIYQKD
SDCRTALHHATVQHLPKLLSMILKNYDQSLIDQPDSLGMSALHWSAYYNNSEACRM
LIRNRSIPSSVDNEGRTPIHLCAGNDDCSSIKVFLEFAGQTILNLQDSQGCTLLHHIVAI
QNVSLLSTITNPQNSSWLMNLNLDSKDLQHRTPLHWAFALGSSVIAGILISRGADW
KVIDDNGYLPLHYAIERNFIDCIDCYFTFDFTDDLEDNNGRTCLMLSCQLGHHESTR
AILERRISGVNGFSDKNGMSALHFASIGGNINCLKLLLQYGALVDMQDNHLRTSLSYA
CEYGYAFIVEYLINSGGLIDLPSIDGRLPIHYAALTNSCSTIAFLISKGNQLNTPDQMGR
IPLHYACTLDNVELITMLGFAQINSQDLSGFTPLHWAVMHENIHTIKILVENDAKINV
MDYSMQRRTPLDLSLFIEDREIYSFLRYLGCVTGIEIQTLAVIE**IQRFWRHISRQRLKK**
NLRITSPEKPVQKSRVQPSLRSIDSSLTNDVHKTNCDEKYFPNLEIHPQTTSQTGISISKSV
DLQRVTLFHRKTEAALIIQLFWRKYLRNKQLKSLHKVEHTTKQSERLFEYFNRKREMAA
LT**IQLAWRQFLRKKLL**NIQII**KRKK**LFQWSPGCLAIRQRLLIEQLYSQQFHPFYYTPSSIS
RADRPSFIKYIPSPAALSFNFAINQYCRVTDCHF

>*Xenoturbella_bocki* [SRX1343818]
MENDSTNTMVSTKNSITNIQADDMSQDPILSPTSLITAGSSSSALSARSNSSS**SLSPVLMA**
CVEGQKHLLQRYIIDDPSCIEMRDHCGRTPLIFCVLADRVDCAKVLLDYNADPSAMD
SGRRTALHWAAHKDNYQFVKLLVSWGVDCCGKDAEGLTPIHLATHHQSTKCLAFL
LKQLLPGEVDIQDDNKRTPLHWSASFGNSEHVKLLLKQDSNVGLPDSEGKTPLHW
AVNGKGAGAITTAELLLDVAPSVINWQDYEGRMALHLAISEGNADMVDVLTCWDS
CKMDCLDNMFRTPLHWAAVLGRAEIIKMLLERNADWSCSDANGATPLHYAAQKN
HFEAVEALLSHRGVTDEPDIEGRTAVMWAAGRGADDVIEVMLAFDHDIHATDKNG
GTALHAACISGKPSSVKILIQRNADIDATDMMKHTPLFRACELGHVEIANVLIEQGAS
LSLLDEAGRTPVHWAAYSGHINVCLLLLNAGMDPNTRDFSGRSPLHCAAYGGHTDC
MELLIDNGAHINVQDKEGMTALHWCCSSGYLDAVTLLFRYLAYPNHIEFTEDRYTP
LDYALLGAHQQVSQFMIEHGALSIIGIQDIAASR**IQAFYRGHRVRKSF**RE**RKLLLMKH**
EELRTKRRSEVLKRVDNSQLGSQIVPPANIITKKLSASKSNSLTSGSHDHMSSSSLASVQG
TLYSTLSNEKESLVSSSLDLSDMIFPIRDKSVADTVAEERQRLSEFRKKSEAASRIQKSWR
DYTFHSKKLVKTFREILRTKRFLSREQQLRREMAALT**IQLAWRRYYREKI**QKSLYPN**RK**
MLHHWSEPVRVSKQSFLVEQVYGERLEVCDWTIPRNPKAARPSYFKYIPSPAALSYNFA
ADQYHPVMARQGIAGMNTFTQSSSRVQIPQLKQSSNWPQTDAGVTLYADELLNLPMTP
SMSSDRFSYNLR

>*Isodiametra_pulchra* [SRX1343817]
MGDMSGSWLGRHSLTQLAIPAKT**QHVTGALHVAAINGDKPAIVKLISSSHDARATAD**
SPDSFGRTPLMYAVLGGRLDSAELLLKAGCNINAKDNARRTALHWAAHKENHKFF
KLLFNRGCDFRAKDMEGLTAFHLCMKHRTPKSASLLIKALNANEIEEQDNFKRTAL
HIAAMYGSHEQVQLLLSRDANVGIPDGAGKTPLHWAINSKNQYSGQIVQQLIDNMP

SVINWQDHDGRTSLHLAVAEGSYNLVKTLISNESCKLACLDNNFRTPLHWAAVLGH
TQIVELLLNKGCDWTCSDANGATPLHYAAQKNHSETVKAFLQFPELTDVADIEGRT
ALMWAAGSGSTETLSTFLTFSPDLTQTDKNGATALHAASASGSASCVKVLLEAKCAI
SADNMKLTPLARAAEMGHVPVAQLLIKFGAKVTHQDQQGRFPLHWAALGGSPKM
CELLLQSHNDANCSDYFGRTPLQYAAQKGFNNCMSKLIDFLADPNLQDAQGRTALH
WSCSLGHLDAAKLLIKSGSFPNHMEFTEERYTPLDYALLNNHQEVVQYLLEHGALSI
NGIRDLAAAK**IQAAFKGLRERRSF**NER**RTLLLKHDQ**LRKKRNSHLGDLQPLIAERIHTL
AEERSVPPFKMPFSDESPSNQEDNPPLEQDKSPPEPKDQTPPEQVKMSDEDKPDSSSTA
ASTSKSSSTASKTSSLKSSSSKSSSTSESEKEGKFNIAEPTKDMENQKRRSRKVAHFEDQT
ERKNDAVSVSSQHSNMTGEANQGEASGASRPSSSMSSSSSTSGSASLSSRISGKPLGLAVS
QSTPKPPKAEPSESEEAEAQRAVSPIGVRNATESVVPFDITADESRRRSSSVSSYTPRRQL
DFTNTNFSLSDEDVTAVQSSDDSCALAGGVAGSPMMDGSFDEVMERQSEFRRLRVAEQ
RRRAEIHRRNHAAAVIQKAWRAYGFKS**KKFYKAFKEVIHKRR**TLKQEQTLKKEVAALT
IQLAWRRYYRRKILRTSNPYKMLVNKYSVKVRLQKQNFLESRIYSEKESLRHSPHYNPL
LHTLQQPGPGGLPVNHFYGVTPPKPKRPFYFRSVISQSAADYNFALNNYHPTRYAYGGV
APKRIHSGKRYSYNMRA

>*Convolutriloba_macropiga* [SRX1343815]
MDNHNVRNSRSVSSRRPDHRDANTRNYTSSHHHFSSNQRSNRSSSSMAFESSSARFSSS
AGIGLSDRSQIGSPTASFAAPSLSRLTIPYVN**PHLHTNALHLACVNGDKQLLIKLVTSSD**
ASGLNLHTSSRSVDRGNGSRASMRSMNINTYFTGAIPDVDTPDNIGRTPLMYCVLG
NKFDCAEYLIKSTRANVNAADNLGRTALHWAAHKEHHKMLKLLISKGAADCRAKD
RDGLTPFHLCMKNRSSKSTALIMKNLNANDLEEQDNYKRTALHLAAMYGSYDQVQ
MLLQRDANVGIPDGEGKTPLHWAINSKHQHSCSIVQCLVDHMPSVVNWQDHDGR
TSLHLAVAEGSYNLVKTLIATDGCKLSCLDNQFRTPLHWAAVMDSPASAAAQIVTLL
LHKDCDWTCSDSNGATPLHYAVQKNNAEIVRAFLEFPELVDQADIEGRSAFMWAA
GAGALESLDLMVDHGADLSQSEKNGATALHAAATSGNVDCVKLLLRVGCKMTADN
LKLTPLMRACELGHVAIAQLLIRHGAKITQQDREGRYPLHWACFGGNLRMCELLVV
HGNVNCSDYMGRTPLQYAAQKGFNNCISLLIQHGADPNLQDAQGRTALHWSCSLG
HLDVVRLLLKWEAFPNYMEFSEERYTPLDYALLNSHHEVAQYMIENAGALSINGIRD
IAAAK**IQAAFKGARERRSF**VER**RTLLLKHD**KIRKNKSSSTGQQQHNNSAYSSLQNSEAE
VLPLVRETFYSAPEDYSPNKSNMEAVDYSNPETQPQNEPRTSEAEQPVFSEAQNDVSEIS
TKESKGSLTGNSAAAKDKSKSNSSRSSSSRSSKSGHKIELAQFKNESQTDINQERPISKSSE
IDQDGAGLSGSKQENVGTPNESTLSAKFENSMVGGQDPVKHSESGSLKIEWDKTTDMPY
KQDDSKAEEVEQPAGIKYDDSSVHSLNISESQEKSNDYESQGALPDTEHQTSEPKVTENL
EGANDEEWEESRTLTRVKDETDVPPLSDQQNFADFFMNDAESEQGQNLQDKDTTEEP
EKVTMRSEISSKTQKSSMRGTIKDSENFYPHSRSERTTPGHTSRAISETSSIRGQVRDEQD
NSISEDRNQVAQSTPKPTHTQDVNEKGKRKVKSSNRERAQKVADSSSSTSAESHRSGSS
MSIKSEDIVQRQEQGSEYGTGVEQEPETGEAEREKSITSEYRAMRAESVLPFNVTADEES
QKSDARSQKSSKHSDTAGALESRGPEKTATTDKNEGDEVTGELCSSTGGLGAGVQKDSV
SAMLERQDHFRRIRLEEQQRRLEIQRRNNAAKLIQKAWRTYGFKSKKFYKAFKEVFLLR
RSIGQEQSLKQEVAALT**IQLAWRRYYRRKVL**ASNNPYRTMLNKFSRDVRNAKQNFVET
RIYTRQQRAHQFYGMRVERPKRPYYANRIPYQSTLDYNFALQQYKPILYSQFETVAPRRE
RSARRFSYNLRT

>*Saccoglossus_kowalevskii* [low quality protein; XP_006819100.1]
MSKSQREQRKMQTRQASRTIPSFEVPGAEGNIGSDTSLKR**DVAPTSQIHAAAVNGDKSI**
LQKLLSANPHQIDGQDQFGRTPLMFAVLADRLECTEILLKAGANVDAKDSGGRTAI
HWAAHKGHFKCLKLLISKGANCKEKDSEGQTALHLSTRHKNTKCLALLMKQLHVD

LGEVDEQDSAKRTALHWSASYGNEEAVRMLIKQDSNIGIPDTEGKTPLHWAATAG
QDSSAVNTVKLLLESAPSVINWQDYEGRTALHLTVADGNEPIVGALTSLEKCNVTAL
DNMFRTPLHWAAVLGHTKIVQLLLNRKADYASTDSNGATPMHYAAQNNYAETVD
AFLSRENVTDEPDLEGRTALMWAAGKGADDVIRTILKRNSDINATDKTGGTGKHK
NAISSNNKSFSLLLXHGAHVNACDMMKHTPLFRACEMGHLEVVKTLMDGGAKVNI
ADQDGRTPLHWAALGGHAVICETLMKHGISVDVRDHVGRTPLQCAAYGGYINCMS
LLMENGADPNLQDHEGMTALHWACSSGCLDAIKLLFEYKAFPNHMEFNEDRFTPL
DYALLNDHHDVAQYMIEQGALSITGIRDLVASKIQAHYRGYLVRKTFLERKKLLMKH
EQLRKDAAKKKEQETKKRQLEYRKREEQRRQRLHHKQGNESQNSTDSEQDKINNQH
RQHPVKETSVTSSDTEESVSSSRSDKYSVSSDRYSDNFEDEEMMERDREKEERKKNKSI
VNKERNRLALWRKKNEAAAVIQRAWKTYLFRRSSMSRAASEVARNYRRLKGEQDYRR
QIAACVIQLTWRKYCRRKLEKSLGKNKKILYPWSPPVLAAKQRHLVEQIYSEVLIAREW
YPELPRVIRPDYMKFVPSAAALSYNFAVDQYHPMLARRGVGKAENVPPWSPRLRPSSM
RRDTYGWTPQSGELFDEDFRHMSLANSREASADHFSYNLHR

>*Danio_rerio* [NP_694502.1]
MAMLLPQNPSQVHAAAVNGDKNTLHKLITESALRDSEDQFGRTPLMYCVLADRLDC
AEVLLKAGAGINKTDHSQRTALHLAAQKGNVRFMKLLLSRHADWRLKDLEEMTPL
HLASRHSSSKPLSLLLKHMAPGEVDTQDRNKQTALHWSAFYNHPEHVKLLIKHDS
NIGIPDSEGKIPLHWAAHNKHPNATRTVRCILEAAPTESLLNWQDYEGRTPLHFAV
ADGNEAVVEVLTSYEGCSVTAYDNLFRTPLHWAALLGHAKIVHLLLERNKSGMIPS
DSQGATPLHYGAQSNFADTVAVFLKHHSVRDEPDLEGRTAFMWAAGKGSNDVIKI
MLDLKKDIDINMTDKYGGTALHAAALSGHVSTVRLLLEQGGMVDPLDVMKHTPLF
RACEMGHRDVILTLIKGGARVDLVDIDGHSALHWAALGGNAEVCEVLMENGISPNL
QDQAGRTPLQCAAYAGYINCMALLIQHDADPNIQDKEGRTALHWSCNNGYLDAVK
LLLGCGAFPNHMEHTEERYTPLDYALLGEHQELTQFLLEHGALSIAAIQDIAASSIQA
LYKGYKVRRAFRERKKLLMRHEQLRKDAAKKREEERRREAEQQLSFAEAGQKQRVL
LAAVGVEKLSLDEAEQRVKDSVAAKGHKHKKSSSAHNSQSRREKPSRAERRTREPETSD
APSIRSLPPVTPMSTKKCPVTREEVCVRETFGPDTGISLNCGSANERRSPAGSSRPGSAKP
VHTGTHVAPGHMETTPTPKPRPSTTGAHSKNASQDTPQHNETQTTSKGNKPISEHKPT
GSQPSNNTSVTRQKEKRQEKETHREKDKRSRTEGDKQTVREKQKGTGIERAKERLMGR
TRKKLAEKEKEKKKDGTCSKNQAAVVIQRAWRRSCVRGRIRKVLCRSLKGVESAEATA
LLIQLLWEWPVLHDHTHRKPSDVQAPPTRIAGKKSSVLQNIYGAAPSKRGTSLRAAALK
TQSQSQSQVLLDLSLRTHKQLSAVECVNLVDSVSQAKQFSYHLRPSSGASQSSQN

>*Xenopus_laevis*_A [NP_001079230.1]
MSSPPQGSSLASPVQAAAVTGDKTTLLKLIASSPEVIDQEDQLGRTPLMYSVLGDRRS
CAEALLKHGAQVNHPDRSGRTALHLAAQTGNHRLLKLLLSRKADCTHRDLRDITAV
HLSTRHQDTRCLALILKYTPPGQVDAQDQRKQTALHWSAYYNRPRHVRLLVRHGS
NIGIPDTEGKIPLHWAAGHKDPEAALTVRCLLEAAPTESLLNWQDYEGRTPLHLAV
GDGNQEVVRLLTSYRGCNVAPYDNLFRTPLHWAALLGYTPIAHLLLETNNSPNIPSD
SQGATPLHYAAQGNCPDTVRVLLSHISVRDEADLEGRTAFMWAAGKGSDEVVRTM
LELDPELEVNRTDKYGGTALHAASLSGQITTVRILLENRVQVDAVDVMKHTALFRA
CEMGHREVISTLIKGGAKVHLVDKDGRSPLHWAALGGNANVCQILIENNINPDAQD
YEGRTPLQCAAYGGYIGCMEVLMENKADPNIQDKNGRTALHWSCNNGYLDAVKLL
LGYSAFPNQMENTEERYTPLDYALLGGHQEVIQFMLEHGALSIAAIQDIAASKIQAVY
KGHKVRRAFQERKNLLMKHEQLRKGAAAKKREGENRQKGKVGQTEGKQKDENHVM
RQDKSNEHIQNEVMREWYGEETGRAEDRKEEHQEENQNIEPKQLKHSKHMEQNSKSIA
KNQKRAGHIQSSPIEHVHTNSIQTRMSPSRTSISHSSPLGNETPKNMYWDDNPTQNNTQ

PRRTSRPQIESPNIIVHRIEDLVQKESRRKSHREERKGSHRQRASSHHRLQASERETAGSV
IHGEVEFKKKETKKGRRTAAGTSKIRASGEAGRLSQSEKEFSSTGIQGRVDCITSPESCET
PSRVCRERKMISAKSGQRPLTETQSPEKACQGSSALKPSLTSHTKQTAIASKCLDSTPSYI
GFGEAIKPLTPMGILREGSFSSKWQNIDIELIPLEARLQLVEKEKARKQLFQRKKHAATV
IQKAWRTYCIRKSSRKTRHSHLRNNPRAMV

>*Xenopus_laevis*_B [NP_001083066.1]
MSNPPQGSSLASPIQAAAVTGDKTTLLRLIASSPQVIDQEDQLGRTPLMYSVLGDRRS
CAEALLKHGAKVNRPDRSGRTALHLAAQTGNHRLLKLLLSRKADCTHRDLCDITAL
HLSTRHQDTQCLVLLLKYTPPGQVDAQDQRKQTALHWSAYYNRPQHVRLLVRHG
SNIGIPDTEGKIPLHWAAGHKDPEAALTVRCLFEAAPTESLLNWQDYEGRTPLHLA
VGDGNQEVVRLLTSYRGCNVAPYDNLFRTPLHWAALLGHTPIAHLLLERNNSPNIPS
DSQGATPLHYAAQGNCPDTVRVLLSHPSVRDEADLEGRTALMWAAGKGSDEVVRT
MLELNPKLEVNRTDKYGGTALHAASLSGQITTVRILLENRAQADAVDVMKHTPLFR
ACEMGHREVIATLIKGGAKVHLVDKDGRSPLHWAALGGNANVCQILIENNINPDAQ
DYEGRTPLQCAAYGGYIGCMEVLMENKADPNIQDKNGRTALHWSCNNGYLDAVKL
LLGYNAFPNQMENTEERYTPLDYALLGGHQEVIQFMLEHGALSIAAIQDIAAFKIQAV
YKGHKVRRAFQERKNLLMKHEQLRKGAAAKKREGENRQKVKVGQTKGKQKDADSM
ERQNKSNEQIIKNEVVHEWQGEASGNAEDRKGKHREENLETNHLQHSKHMAKNQRIT
AQIQSSPSEHEHTNSIQIRTSPSGTSNTQSSPLGNEIPKNMYWDDNPSQKHTQTRRTSRH
QMESPDVVVHRIEDLIQKESRRKSHREERKGSHRQRQSSDYRLHTSEKEASDSAIHREEE
GKKKETKKGRRTVAVTPKIEACCKGGCGKLSQSEKVSLGIGIQGRVDCITSPEPCETPSRV
CRERKTISAKTGQRPLTETHKPPGKACRSSSALKSSLPSHIKQTAIDSKCLDSTLSYIGFGE
AIKPLFPMGILREGSFFSKWQNIDIELIPVQARLQLVEREKARKQLFQRKNHAATVIQKA
WRTYWVRKSSCKTRHSRSQNNPPAMV

>*Gallus_gallus* [NP_989882.1]
MIIPCKSMNISANCLFSGSSLASEVHAAAVNGDKSTLLKLIAGNSELKDKEDQFGRTPL
MYCVLADRVDCAEALLKAGADVNRADRSRRTALHLAAQKGNYRFMKLLLARRGN
WMQKDLEGMTPLHLTTRHKSPKCLALLLKHMAPGEVDTQDRNKQTALHWSAYY
NNPEHVKLLIKHDSNIGIPDIEGKIPLHWAANNKDPSAIHTVKCILEAAPTESLLNW
QDYEGRTPLHFAVADGNVAVVDVLTSYEGCNVTSYDNLFRTPLHWAALLGHAQIV
HLLLERNKFGTIPSDSQGATPLHYAAQSNFAETVEVFLKHPSVKDDSDLEGRTSFM
WAAGKGSDNVIRTMLDLKLDIDINMTDKYAGTALHAAALSGHVSTVKLLLERNAQ
VDALDVMKHTPLFRACEMGHKEVIQTLIKGGARVDLVDQDGHSPLHWAALGGNA
DVCQILIENKINPNVQDYAGRTPLQCAAYGGYINCMVVLLENNADPNIQDKEGRTAL
HWLCNNGYLDAIKLLLGFDAFPNHMENSEERYTPLDYALLGEHHEVIQFMLEHGAL
SIAAIQDIAAFKIQAVYKGYKVRKAFQERKNLLMKHEQLRKDAAAKKREEESKRKEA
SLQKGMQNMEQNKFQVQLSTAVREKTASTLQLSNKQTDLQNKRPLSVSASQIQLGRNS
RGSPKACRSKGSPKESCLSSELQSEGHNIRQELLRKHIKSKPSCVHFHCGKVKEVTKVEA
KHQVAAATELNGEKHKEHAVEANGTSAHGNRRHASACGTAGAGEKTRDQSLSSSGNRG
HCEGTSVVVCNVSCAGGIARNSKRCEAVPKSKRHQQKSRHKEVNYERCSPAGSSRPGSA
KVVFVNTRNATVCAIEHANNVGNHELAKKTSPLLSTETESTGTGPRNPAACSALDDSLN
LEKTGEVGSRSAGDQLCSVAWQSTNIELIPLEIRMQIIEKERTRKELFRKKNYAATVIQR
TWRSYRLRQELSQLLSAKRQRKEDEDKWRQETAAFLIQVAWKKQLNHSPQKSVPSCK
SLKSVNKTSSAIKTSKQSILKQIYGRSQEGRVYQPARPPSKLKLSDVQLVSANNLQYVNLL
ENVGKSKQFSYNMRPSTAAKSKSTRLEH

>*Rattus_norvegicus* [NP_001101402.1]
MNISENVLSTGSSL**ASQVHAAAVNGDKGALQRLIVGNSALKDKEDQFGRTPLMYCVL**
ADRLDCADALLKAGADVNKTDHSRRTALHLAAQKGNYRFMKLLLTRRANWMQK
DLEEMTPLHLSTRHRSPKCLALLLKFMAPGEVDTQDKNKQTALHWSAYYNNPEHA
KLLIKHDSNIGIPDVEGKIPLHWAANHKDASAVHTVRCILDAAPTESLLNWQDYEG
RTPLHFAVADGNLTVVDVLTSYESCNITSYDNLFRTPLHWAALLGHAQIVHLLLERN
KSGTVPSDSQGATPLHYAAQSNFAETVKVFLTHPSVRDDSDLEGRTSFMWAAGKGS
DDVLRTMLSLKSDIDINMADKYGGTALHAAALSGHVSTVKLLLDNDAQVDATDVM
KHTPLFRACEMGHRDVIQTLIKGGARVDLVDQDGHSLLHWAALGGNADVCQILIEN
KINPNVQDYAGRTPLQCAAYGGYINCMAVLMENNADPNIQDKEGRTALHWSCNNG
YLDAIKLLLDFAAFPNQMENNEERYTPLDYALLGERHEVIQFMLEHGALSIAAIQDIA
AFK**IQAVYKGYKVRKAF**RD**RKNLLMKHE**QL**RKDAAAKKREEENKRK**EAEQKGQPD
TDPPRPQVLPCPPSPQNEPSRQSAAPSKQPPASHTVQSPDPEQGRPAGQCPGRASHRDC
SSDLQGTASRKPSESPREHCRGASAGVHPRSCEDGDSYRHQGTSSAEKCRGETSGEHQR
CEQGPSRARQPVSTGSAGPAEKGEDSSPAITSASQRDRPRKPS**KRQDGAPRSRGASQKR**
RTHQLRDRCSPAGSSRPGSAKAEVARADQSSVHHDTPRNKVTQDKLVGGVSSDVPLSTE
ELGSGCQKLYEDLCASPENNVAHGPRPGQSVNIHLLPVEQRLLIIQRERS**RKELFRRKN**K
AAAV**IQRAWRSYQLRKHL**SRLLHLKQLGARDVLRCTQLCAALLLQVWRKELELKFPKS
ISVSRTSKSPSKGMSAAKSARHSVLRQIYGCSQEGKVHHPIRSPKVPPVLHLS**SVNSLQSI**
HLDSSGRSKKFSYNLQPASQSKNKPKL

>*Mus_musculus** [CAA09388.1]
MNISEDVLSTGSSL**ASQVHAAAINGDKGALQKLIVGNSALRDKEDRFGRTPLMYCVL**
ADRVDCADALLKAGADVNKTDHSRRTALHLAAQKGNYRFMKLLLTRRANWMQK
DLEEMTPLHLSTRHRSPKCLALLLKFMAPGEVDTQDKNKQTALHWSAYYNNPEHA
KLLIKHDSNIGIPDVEGKIPLHWAANHKDPSAVHTVRCILDAAPTESLLNWQDYEG
RTPLHFAVADGNLTVVDVLTSYESCNITSYDNLFRTPLHWAALLGHAQIVHLLLERN
KSGTIPSDSQGATPLHYAAQSNFAETVKVFLQHPSVKDDSDQEGRTSFMWAAGKG
NDDVLRTMLSLKSDIDINMSDKYGGTALHAAALSGHVSTVKLLLDNDAQVDATDV
MKHTPLFRACEMGHRDVIQTLIKGGARVDLVDQDGHSLLHWAALGGNADVCQILIE
NKINPNVQDYAGRTPIQSAAYGGNINCMAVLMENNADPNIQDKEG*RTALHWSCN*NG
YLDAIKLLLDFAAFPNQMENNEERYTPLDYALLGERHEVIQFMLEHGALSIAAIQDIA
AFK**IQAVYKGYKVRKAF**RD**RKNLLMKHE**QL**RKDAAAKKREEENKRK**EAEQQKGQL
DTDPPRSHCSSSAPVLPCPPSPQNEASKQDATPSKQPPASHTVQSPDPEHSRPPGRCPGR
ASQGDSSIDLQGTASRKPSETPIEHCRGPSACVHPRSWEGGNSSKNQGTSSVEKRRGETN
G*KHRR*CEEGPSSARQPLCTGSGRPAEKGEDSSPAVASASQQDHP*RKPN***KRQDRAARPR**
GASQKRRTHQLRDRCSPAGSSRPGSAKGEVACADQSSLHRHTPRSKVTQDKLIGGVSSG
LPLSTEASRSGCKQLYEDICASPETGVAHGPPPGQCMNIHLLPVEQRLLIIQRERS**RKELF**
RRKNKAAAV**IQRAWRSYQLRKHL**SRLLHLKQLGAREVLRCTQVCTALLLQVWRKELE
LKFPKSISVSRTSKSPSKGSSATKYARHSVLRQIYGCSQEGKGHHPIKSSKAPAVLHLS**SVN**
SLQSIHLDNSGRSKKFSYNLQPSSQSKNKPKL

>*Halicryptus_spinulosus* [SRX1343820]
MAVCISCCLNSKRNNSHLQLFGSRDGDVLSLTSLPHTAAADCSQRIMSTDSQSA**DVFYIR**
QTHAAAINGEKSLLQKLTAGNAGLIDAGDQFGRTALMFAVLVDRTDSAEWLLKHG
ANIKAKDKGGRTALHWAAHKGSYKCVKLLLSKGSDWSDKDQEGQTALHLSTRHKS
PWCLTLILKQLSAGEVDNQDNNKRTALHWSAAYGNLEHMKLLLRHDSNIGIPDIEG
KTPLHWAASCRGVDAIVCCKLLLAIAPTVINWQDYDGRTALHLAVADGNEPVVAAL

TSVDKCTVGALDNMFRTPLHWAAVLGHAAIVKLLLESGADYASSDANGATPLHYA
AQNNHAETVKVFLSRDGVSDVPDLEGRSALMWAAGKGANDVIRTMMRFNCDIYAT
DKNGGTGERATPISATCIAERAQKVERSVEKLPAGSTCVTFCLAALHAAALSGHAAT
VLLLQEFGADVNATDNMKHSPVFRACEMGNTDVVRALIEGGASLDLFDQDGRSPLH
WAALGGHAYICQTLGQYGVDPDVRDDSGRSPLQCAAYGGFVNCMSVLVENQADVD
LQDNEGMTALHWACVSGHLDAVKLLLECRAFPNHVKVTEDRFTALDYALLGEHHD
VAQYLVEQGGLSINAIQDIAAAK**IQAVFRGYRVRKTF**LKRKSLLMKHEKLRKDAAMLV
KNAVLNRVEHVLFRILLIAVRLGVRCFCLRHINMLFEKSGGDAQENAGETFACRGRH**RR**
RKEQKQEQQRRERRRREDEEAFARE…

>*Priapulus_caudatus* [SRX507009]
…GGARLDGSDQDGRSPLHWATLGGHAYICLTLAAYAVDPDLRDHSGRSPLQCAAY
GGFVNCMSVLIENQADLDLQDNEGMTALHWACVSGHLDAVKLLLECGAFPNHVKF
TEDRFSALDYALMGEHQEVAQYLIEQGGLSINAIQDIAVAK**IQAVFRGYRVRKTF**LK**R**
KTLLMKHEKLRKDAAM**KRAK**EELQKKTQEQATGIFFDSLPTNETNKNQNNNNNEAS
AKIHRRLEGTDIALV…

>*Lepidodermella_squamata* [SRX1000997]**
…MITSQHLAAAVCSTGSGNLKSPGTIINDPAGHSPIFYAIHCGNEKAVRSLIQLPHLIH
QADSHGHTALAIAVGRGELKVVQDLISAGADICSKDVNGVTPLHLATKSTPEIMKAI
AAKVNVHNVDITDSSGRTPLHWAAAQARENYCRILLQAGANTAKLDQNRRTPLH
WLSVSENFETSSSVATLFLRTYPDLVRWQDVRGATPFHLAACIGNATLLETLLEYAD
KDTINRADNLFRTPLHYACSIGRAKIVELLLESGAKIDCADCEGATAAHLAAYWNHT
SCLALLIERGTPSLISSLVDYNGQSPLHWAVSKRNFDAARCLCSAMASPVSSLLLIAAA
NGDDAMVDLLLTNGVRVSDSLDPKTSSTALIAAAQHGHVTVLMTLVYRGKANPRQ
TDLRGRTALHWAAMSGSAESCKFLLEVGCPYEAADYLGCNALHCAAFSAAECVKVL
LPKSSPEVKDRYGFTPLHWACLGRNAGSVKLLIAAGAKVNALGITRDRHTPLDCAL
RRNATEVAEILVGAGGMTAHGVKEYSATK**IQSLWRSYRVRKVV**KTLKRHELRLAKG
QHYLKNLKEKPQPDSPFQDEVVDAVYYGMGKSRLEAGSSREDNATSDHSFFDEEQADR
SNTDQNGEVAPPLVSSLNTNGYGDSEYRERRGSTHSVRFANVESSSNMQDRHHNDDRR
ESSQQLSDQMANRFKQLRIGENQSQDQPSTSNNDGSMSMATVVSVVVSA**KRNR**RYNGL
VINDDDFDDDDEDDMTPEELQQKYYRMTEEQIILKESEAA**RKLRVDEMKRVEAVRRR**
FQAASV**IQRSWRRYKRNK**KNTQTQKLTARQMRLRAGEQAWKRQMAAITLQNAWRK
HHRNEVLKSIELHPDLKQVQSPLSSMMRSTRQAVLVSSVYRKSVPLSLLHPPNRPKTGV
AERASAPKPTVKPLSDLSKPRPAAISFNFAHRQYTSQTYKSNH

>*Membranipora_membranacea* [SRX1121923]
MPIKLFSKRKSSKGKTGKLSRKGSKDDILQEEDPQGASGSSGNYFKYGATGSMLSNSTSS
GSENSAKFGHPACKQCGLLMKDGSHSPTCQLCLRLDPTFGQYLTTGQGLEGGGIGAETQ
VSEIVKSVEILDSNR**AQIMVTQMHAAAVNGDKSQLSKLLLTNNKEVDTGDQFGRTPL**
IFCVLADRVECAEILLKAGCKVDKKDETGRTALLWAAHKGHYRCLKLLLTRGANWK
EKDKEGETALHLSTRHRSSKCLSLLKSHLAPGEVDDQDNNKRTALHWAASYGNLE
HMKLLIKQDSNSGIPDSDGKTPLHWAASSHDENAAACVRLLLETTPFIMNWQDFEG
RTPLHLAVEIGNNSVVSLLTSYGSCNVSALDNMFRTPLHWAAALGHSKIAGMLLDSK
ADFSSADSNGATALHYSAQNNHAKTVEVFLTRASITDEPDLEGRTALMWAASKGA
DEAIATWVRHNPDLRHIDKNGGTGLHAAALSGHASTVRLLLEHGSDIDAQDLLMH
TPLFRACEMGHTDVVQALIDGGAQVEVIDQDGRSPLHWAALGGHAYICQILIKYGV
DPNVRDFSGRTSLQCAAYGGYVNCMSVLIENKANTNTQDKEGMSALHWACSRGHL
DSVKLLVEFRAFPNHMESTEDRYTPLDYALMGEHHEVAQFMIEQGGLSITGIQDIAA

CKVQSAFRGWKVRKAFLERKELLVRHDQIRKDAARKRAGEEYKKKEDVRNKRELG
NRIKQQSSDTAEPVVVERPVSKKKKKDELTEEERILKDAKQKQAIVSSERSRLTYHRRR
HNAALCIQKAWKMFKFRTQGLMKNARDAIKLKRLDAGELEWKKQIAACTIQLAWRKY
FRRKLLRALNKTNNAVMHMWDPEMLAFRQRSLVEHIYSEELYAPHWHPVLTAPIRPY
WARYLPSSAALSFNFAIDQYTQPREIFWDLQQQGHHDMDELDMEIDGDLARLRMTERA
GGSPTHYSYNMNS

>*Novocrania_anomala* [SRX1343816]
MARHGLCLFKSNRVEPDPHPQEDSEMAAVARSRGHLSKVSGPFENSGPIQNMEATTQ
MHAAAVNGDKAALAKLLVSSPKEIDIGDQFGRTPLMFCVLADRLECAELLMKAGAT
VEKKDKGGRNALHWSAHKGNFRCLKLLLSKGANWKEKDSEGQTALHLCTRHKSP
KCLALLLKQLAPGEVDDQDKNKRTALHWSASYGNLEHIKMLIKQDSNIGIPDVEGK
TPLHWAASSRDPAAVSCVKLLLETAPSVINWQDYEGRTALHLAVADGNDDVVDAL
TSLEKCNVSALDNMFRTPLHWAAVLGHSKIVSLLLSKHADYASSDSNGATPLHYAA
QNNFADTVEVFLAEKIKDEPDAEGRKAFMWAAGKGADDVIKTWVKYGVHLQQTD
RNGGTALHAAALSGHASTVKLLLDNKAKINAEDLLKHTPLFRACEMGHTDVVQTLI
DYGAKVEVFDQSGRSPLHWGALGGHAYICQTLIKYGIDPNVKDNSGRTPLQCAAYG
GYVNCISVLIENKANPNTQDQEGMTALHWGCSKGHLDTVKLLVEFCAFPNHMEFTE
DRYTPLDYALMGEHHEVAQFMIEQGGLSINGIQDIAACKIQSAFRGWRVRKTFLERK
KLLMKHEQLRKDAAKKRAEEEYKRKEDLRNKQEVDSRRKMQVEFTDQPVTRSSSRS
KKKKEPQLSEEEIILLKTKQNKELVNKERDRVAEFKKKQDAAKVIQRAWKHYVFRYKG
VMKSAAAAIKSKRLRAGEEEWKQQIAALTIQLAWRKYYRKKLLQSLTPRSKRMLHM
WDPEVVAMKQKMMVEAIYTEKIVANEWHPTLNPPVRPLWFKYIPSPAALSFNFAMEQ
YNPELVEAQAMGGPTDSMEFLAMQDALGHMKLSNSSEARQFSYNMYDY

>*Lottia_gigantea* [protID_159385]
MLTSPIVNSTQDNNLIVNFQSEEEDKDCDYESVKQLHAAAVNGDKHYLARLLTAFHH
YIDYGDEFGRNALIFCILADRYDCAELLLKNGCNINHRDKGGRTPLHWAAHKGNYK
MLKLLLSKGASWYEKDNESQTALHLSTRHKQYKCLSLLLRHLSPAEIDVQDNNKRT
ALHWAASYGNFEHVKMLLRQDSNIGIPDIEGKTPLHWAASSHDRLAVNCVKLILET
TPSVINWQDFEGRTALHLAVADGNESIVNILTSVSKCDVSALDNMFRTPLHWAAVL
GHSKIVRILLEKGADYTNSDFNGATPLHYAAQNNFSETVEIFLSSITLTEDMHDNERR
SAFMWAAGKGADDVINVFFKHGIDLHQVDKNGATALHAAALSGHESTVKLLLVLG
AHIDIEDHVHHTPLFRACEMGHTDVVCCLLDMCSKVDTVDQEGRSTLHWAALGGH
ANICQILINHEVDPNVRDFSGRTPLHCAAYGGYVNCMSVLIEFNADSNAQDSEGMTP
LHWACSKGDLDAVKLLLENNAYPNQMEYTEDRYTPLDYALMGENHDVAQYMIEQ
GALSITGIQDIAAQKIQSAFRGYRVRKTFTERKILLMKHEKLKKEAARKRAVTADDKH
SDYTLQTACDKRAIHIPLVTQTFIVDDIRKVQISQTLDTESSTVHSCPGDPLTEDSSASKK
TAKTKIGRVNQMDRNQAVTIIQRAWQRYRSRKCAKSISSCQVYDLQHQMAALNIQLAW
RKYYRKKQLKTVKPNSQLLQNWHPSFVSDKQQTLSQQMKADHSNVKVTKDFSRKFSY
NLHN

>*Biomphalaria_glabrata****
MVSHYLCHCCVSARFLGLYTEAQSETPRTPFTSKPEGQRLEVVTQIHAAAVNGEKQTL
LRLIAASASNGDLDVGDQFGRSPLMYCVLGDRLECAEVLLKTGAQVNLKDKGGRTA
LHWAAHKGNLRFMKLLLSRGADCREKDNEGQTALHLCTRHKLPKCMALLLRQLTP
GEIDDQDKNKRTALHWAASYGNLEHVKMLIKQDSNIGIPDTEGKTPLHWAASTRD
REAVSCVKLILETTPSVINWQDYEGRTALHLAVADGNESVVNVLTTVKNCNVSALD
NVFRTPLHWAAVLGHSKIVSILLDSGADSSSSDSNGATALHYAAQNNFSDTVSVFLK

HNCTDEADLEGRTAFMWAAGKGADDVIRVFLEHKVDIHQEDKTGGTSLHAAALSG
HSSTVKLLLDHGAQLEVTDQAKHTPLFRACEMGHMDVVQTLIDCHARVDVLDQYG
RSPLHWAALGGHAYICQLLIKYGVNPDVRDHSARAPLHCAAFGGFVNCMSVLMEHE
ADPNSQDKEGLTCLHWACSKGHLDAVKLLIEYNAYPNHMAFTEDRYTPLDYALMG
EHHEVAQYMIEQGALSITGIQDIAAQK**IQSAFRGYRVRKAF**IE**RKKLFMKHE**KLKKEA
ARKRAGEEGKRTMEKQTSFGKSISFDKQTSFDKQTSLDKQTSLDRQTPFDSQTLAQADR
LTVSHTTTSRNAEKLPENSSKTLQAECNHLHNLEIASGHMTSPANIQSNIKNHTADSEN
YDQNSNTRNINEKRNADINSPKQTQESQNDKQPMKTELKVHKPRLRASQR**RRLLQAD**
QNKGSLLGVNQTVIASDKLC

>*Crepidula_fornicata_L*
MAATANPDGQSGSARQDSPQVPFASGTP**RMENITQLHAAAVNGEKQALNRLIAGCGA**
GEMDVGDQFGRNPLMYCVLADRLDCAELLLKAGSQVNFRDKGGRTGLHWAAHKG
NYRLMKLLLSRGADCRERDNEGQTPLHLCTRHKLPKCMAILLRQLLPGEIDNQDNN
KRTALHWASSYGNLEHVKMLIKQDSNIGIPDIEGKTPLHWAASSRDPEAVQCVRLIL
ETTPSVINWQDYEGRTALHLAVADGNTSIVDLLTSVEKCNVSALDNVFRTPLHWAA
VLGHSAIVGQLLQSNADSTSSDSNGATPLHYAAQNNFDETVNVFLSRKNTDEADLEG
RTAFMWAAGKGADNVLRVFFHHGVDVQQEDKTGGTALHAAAFSGHAATVKLLLQ
HGARINATDQVKHTPLFRACEMGHTDVAQALIDNGARVDVLDEDGRSPLHWAALG
GHAYICQILIKYDIDPNIRDHSGRSPLHCAAYGGYVNCMSALMEHQADPNTQDKEG
MACLHWACSKGHLDAVKLLVENDAFPNHMEFREERYTPLDHALMGEHHDVAQYM
IEQGALSITGIQDIAATK**IQSAFRGYRVRKTF**IE**RKKLLMKHE**MLKKEAARKRAGEG**K**
KSEEQLRKESEDLAMEPLVAPLHHQLPQKTQTPSPTHQKSPSRKQQQSPTKQLETK
NPETESQKKQK**ETVQAQTKQASEQPQTKEQVSEPTQTKQEETEQSEIAQQTTEEPH**
ITQQESEQTQTTHEEQKTEQLQTVQNDAQQKHAEQKASEVTPLNSEKSEEVQIQQQ
EAEHRQAKEQELKQPDTKQPEKAHQQTDLQEAEQTLTGQQKSEQQQVQQEKTQQL
KDSQEISSTSDTQQQAPASPVSSFGDKTDSPKDAHVRHRTDSDKQKHAGRVSKESRSVQ
GRK**KKDK**GPSAEEKVLQTKSDKANVKKGE**KKRK**GDIQKRQHAASVIQRAWRKHRRSR
QRHRQQPMVTESSQNEVRMRRLQLSQAAWERQIAALT**IQLAWRKFYRRKLL**KSLNPN
RRQLRMWDPDIIALKQQALIAEVYSEQLMVPFWHPTLKPSVRPFWFRLVPSPAAVSFNF
AVDHYHPLKLWQGMRPHSWTGPVNGTVDLDFVPKSASSHRPERSANSQFSYNLRRA

>*Crepidula_fornicata_S*
MAATANPDGQSGSARQDSPQVPFASGTP**RMENITQLHAAAVNGEKQALNRLIAGCGA**
GEMDVGDQFGRNPLMYCVLADRLDCAELLLKAGSQVNFRDKGGRTGLHWAAHKG
NYRLMKLLLSRGADCRERDNEGQTPLHLCTRHKLPKCMAILLRQLLPGEIDNQDNN
KRTALHWASSYGNLEHVKMLIKQDSNIGIPDIEGKTPLHWAASSRDPEAVQCVRLIL
ETTPSVINWQDYEGRTALHLAVADGNTSIVDLLTSVEKCNVSALDNVFRTPLHWAA
VLGHSAIVGQLLQSNADSTSSDSNGATPLHYAAQNNFDETVNVFLSRKNTDEADLEG
RTAFMWAAGKGADNVLRVFFHHGVDVQQEDKTGGTALHAAAFSGHAATVKLLLQ
HGARINATDQVKHTPLFRACEMGHTDVAQALIDNGARVDVLDEDGRSPLHWAALG
GHAYICQILIKYDIDPNIRDHSGRSPLHCAAYGGYVNCMSALMEHQADPNTQDKEG
MACLHWACSKGHLDAVKLLVENDAFPNHMEFREERYTPLDHALMGEHHDVAQYM
IEQGALSITGIQDIAATK**IQSAFRGYRVRKTF**IE**RKKLLMKHE**MLKKEAARKRAGEGQ
QKSEQQQVQQEKTQQLKDSQEISSTSDTQQQAPASPVSSFGDKTDSPKDAHVRHRTDSD
KQKHAGRVSKESRSVQGRK**KKDK**GPSAEEKVLQTKSDKANVKKGE**KKRK**GDIQKRQH
AASVIQRAWRKHRRSRQRHRQQPMVTESSQNEVRMRRLQLSQAAWERQIAALT**IQLA**
WRKFYRRKLLKSLNPNRRQLRMWDPDIIALKQQALIAEVYSEQLMVPFWHPTLKPSV

RPFWFRLVPSPAAVSFNFAVDHYHPLKLWQGMRPHSWTGPVNGTVDLDFVPKSASSH
RPERSANSQFSYNLRRA

>*Owenia_fusiformis*
MAAETANHAGGGNMAKGMDSNMNLDGGTVN**SNNQITEMHAAAVNGDKPGLAKLLV**
SNAKHLDAGDVFGRTPLMFCVLADRLDCADLLLKAGAQVNKKDKGGRTALHWAS
HKGNFRCMKLLLSKGANWKEKDNEGQTAMHLSTRHKSPKCLALIMKQLEPGEIDD
QDKNKRTALHWSASYGNLEHIRMLIKQNSNIGIPDIEGKTPLHWAASSKDASAVNC
VRLILETAPSVINWQDYEGRTALHLAVADGNEAVVDALTSLAKCNVSALDNMFRTP
LHWAAVLGHSRIVDVLLTKKAEYSSSDSNGATPLHYAAQNNFAATVEVFLSHKAVT
DEPDVEGRTAFMWACGKGANDVIKAFVKSRTDLHLVDKNGGSALHAAALSGHAAS
VKLLLESGSQINTIDLLKHTPLFRACEMGHTDVVHVLIDKGAKVEVTDQDGRSPLH
WAALGGHAYICQTLIKYGIDPNVRDHSGRCPLQCAAYGGFVNCMSVLIESKADPDSQ
DNEGMCALHWAASKGHLDAVKLLIEYKAFPNHMEFTEDRYTPLDYALMGEHTEVS
QYLLEQGALSITGIQNIAASK**IQSSFRGFRVRRTF**LE**RKKLLFKHE**QLRKDAAKKKAEE
EYKKTFDQRNKKEVSERLKVEKLLFVEP**KKTK**EQLAEEERKKLEAQQRAEEAERLKEK
QPKIHRRKKGPQYTEEEMKAMQMKENREIVATERERLNEFRSRQHAAHVIQRSWKLYK
FRTKGILKTASEAIKDKRLELGEDNWKREIAALT**IQLAWRKFYRKKLL**KSLSGKNRKIIH
MWDPNVIAAKQQKLLEQIYTEHIEAPEWHPKLHRPARPSWFKYVPSAAALSYNFAIDQ
YHPLVPRNGTSSHVSRHEDDRLRTSTRGSGGSWSPDNNTEDFQNMKLTDSQNASNYSY
NMYEY

>*Capitella_teleta*
MSLCCVRCRTTSPHDSIVLDAKDVDFDIGEDNFGHDEMEEEDGRKRRLKWGRGSRWSK
RMSTKLKVEGYWAQSTMEELGNGRLFCDSSDYIQGGLVRRSQPTTHQAALYVAHGYIPG
GLTEAQTETFNSGA**HLGGVTSIHAAAVMGDKSALSKLIMTASPCDLGAGDKFGRTPL**
MYCMLGDRIECAELLIKGGADVNKKDIGGRTSLHWAAHKGNLKCLKMLISKGASW
REKDSEGQSVLHLATRHKSPKCLSFLLKLLEVSEVDDQDKNRRTPLHWAASYGNYD
QVKMLIKQDSNIGIPDVEGRTPLHWAASSKDRQAGDCVKLILETTPSVIDWQDYHG
WSALHLAVADGNEAVVQVLVSIENVNVSALDNMFRTPLHWAAVLGHSKISALLLQS
GSDFSSADSNGATPLHYAAQNNFEETVDVFLHHKDMKDQTDIDGRSAFIWACSKGA
VEVMSCFIRHRVDIQQTDKDGGTPLHASALAGHTSSVKLLLDHGAAIHQVDLVKHT
PLFRACEMGHTEVTQALIDAGAKVDIVDQDGRSPLHWAALGGHAYICQVLIKYGVD
PSVRDYSNRTPLQCAAYGGYVNCMSVLLENHADYNTQDNEGMTALHWACSKGHL
DSVKLLIEYHAFPNHMEFTEDRYTSLDYALMGEHIDVAQFMIEQGGLSITGIQDIAAL
K**IQSVFRGYRVRKTF**LD**RKILLMKHE**QL**RKDAAKKRAQEEQIKRR**EMTQESLESDLE
GQGTASSSSTFSGQNRKKKESIAGERAKLQNKQVAQAAAENKQRIEKHRTSDWAARV*I*
*QGAWRQWRSRAGL*REEEALLKEKQWLLQFSALT**IQLAWRRYYRRKL**LRSMSNAKDHR
RKILQSWDPEVVAAKQKTLISQIYTEEK**KRPR**LYAASTRTPYRPPWLKAVPSPAAMSYN
FALHQYHPARRRFRSSWATESSDDLNKMKISDSFQPHHFSYNLYDS

>*Dinophilus_gyrociliatus*
MQPRKKSAKPKLLNDRRSSGTN**YSRLTNLHKYISNGDKESLEKQLEVEGTKNLNTQD**
KTGKSPLLYSILGEYTDCFEVLLNYDPDITIKDSTGKTALHYAVQKGNIAFIEMLLQR
GANIMDKDNEGLNSLLLSSRYSNAKCLNLMLKHLNARDLEDTDKQKKTALHWSCSS
GHLEIVKILINNGMKASTQDITGKTSLHWAAMGKGQFATDCISYLLEVEPSIINVQDY
EGRTSLHQAAAYGDENAVRTLASNEKCSVSCVDSLFRTPLHWASVLGRQEIVKILME
FKIDCGMTDSNGATALHYSAKNNYSGCVKALLGGISQVIRDESDGEGRTALIWASGH
GADDALSLLVEAGCNVHHVDNHGATALHAAAQNGHASTVKLLLAYGSDKNVMDQL

SLTPLFRACEYGHEEVLRILYQSGVDLETRDADGRSALHWASLSGHAHICSMLIKFGL
NPDVYDRSGRTPLQCASYSGHVNCMIILIEYKADPNWQDNEGMTALHWACSKGHL
ESAKILIQFQASLNNMELTEDKYTPLDYALLENHQKVAQFLIEQGALSIRGIQESAATK
IQSVIRGLAVRRELEEK**RNLLRKKHQ**KQLEEKRKQSRHIRAEKLPVTKKNSKNSKQIDR
SETNEMILE**RKNSYRRNISREKERKK**LIKAKITAAIVIQRAWRRYKSSGFVRNGNDDMI
RELAALT**IQLAWRKYYRRRVL**SKLNELTS**KKRK**VLQIWDPAVISAKQEALVRHIYGDE
NVTEPDQWFPKLKAINRPLWLKYASSPAALSFNFAVDQYNPYV

>*Lineus_ruber*
MELGGPLHAGG**TEAVTQMHAAAVNGDKSMLAKLIVASSRDIDLGDQFGRTPLMFCV**
LADRPECAQLLLKAGAAVDKNDRGGRTALHWAAHKGNFKLLRLLLSKGALWKEK
DNEGQTPLHLSTRHKSPKCLALLMKHVGIGEIDDQDKNKRTALHWSSSYGNLEHIR
MLIKQDSNIGIPDTEGKTPLHWAASSHDSQAVNCVNLLLETAPSVINWQDYEGRTA
LHLAVADGNEAVCDALTSLEKCSVSALDNMFRTPLHWAAVLGHSNIVNLLLNRNAD
YSSSDSNGATPLHYAAQNNYSETVEVFLKRREIRDEPDVDGRTAFMWAAGKGADD
VIKTWVKYGADIQEVDKNGGTALHAAAVAGHATTVKLLLDAKASIDGTDLMKHTP
LFRACEMGHTEVVQTLIDLGARVDMLDQDGRSPLHWAALGGHAYICQILIKYGVDP
NIRDLSGRTPLQCAAYGGFVNCMSVLIEHQSDPNAQDNEGMTALHWGCSKGHLDA
VKLLLEYRAFPNHMEFTEDRYTPLDYALIGEHHDVAQFMIEQGGLSISGIQEIAAAK**I**
QSYFRGFRVRKTFLE**RKKLLMKHE**QL**RKDAAKKRAEDEYKKK***EDIRNKREVDERRR*
*R*EQDKSQVAQSTASSSTVQEPTTPCTTPRPETASVAGSGTAGSGMVRSSRK**KRTK**ELTE
EEKILIKRQEKKEIVDAERMRQAEFRRKWKAAVVIQRCWRYYLYFKKGLIKKAGDNIRT
KRLAAGEEEWYRQIAALT**IQLAWRKFYRRKLM**RSLTPKSRHIMHMWDPEVIATKQRA
MLRDIYCEELQVPVWHPKIRPPVRPSWFKYIPSPAALSFNFAVDQYHPLVARRGSSRLT
HVNDNSTRVRPFSMKRDSEGWSHFTDEATKELNGVKLSDKTDRSKTWSYNMREF

* Dominios de *Mus musculus* descritos en la literatura, considerados y no considerados en este análisis:

RTALHWSCN: D-Box descrita no considerada.

RKDAAAKKREEENKRK: NLS descritano considerada.

KHRR: NLS descrita, no demostrada funcionalmente, que no consideramos.

*RKPN***KRQDRAARPRGASQK**: NLS descrita que coincide parcialmente con nuestros análisis de consenso. Se consideró el fragmento azul.

RKELFRRKN: D-Box descrita sí considerada y que coincide con nuestros análisis de consenso

** Posible falta de secuencia en la parte N-terminal de la proteína.

*** Posible dominio IQ divergente, no funcional, en *B. glabrata:* IQSNIKNHTADSEN

Apéndice 8: Secuencias nucleotídicas completas de la región codificante de los genes *inversin, nodal1, nodal2* y *egf-cfc* de las especies *C. fornicata* y *B. glabrata*

En las secuencias presentadas a continuación se ha seguido el siguiente código de color: Codones de iniciación y terminación, y **procesamiento alternativo**

>*inversin_Crepidula_fornicata*

ATGGCAGCCACGGCCAATCCCGATGGCCCGTCGGCCTCTGCTCGTCAGGACAGTCCACA
GGTCCCCTTTGCTAGTGGCACCCCCCGCATGGAAAACATTACACAGCTTCATGCTGCCG
CTGTCAACGGGGAAAAACAAGCTCTCAATCGACTCATTGCAGGATGCGGAGCGGGAGA
GATGGACGTGGGGGACCAGTTTGGTCGCAACCCTCTCATGTACTGCGTGCTGGCTGATA
GACTGGACTGTGCTGAGCTGCTGCTGAAGGCTGGCTCCCAGGTCAACTTCAGGGACAAA
GGGGGCCGCACTGGTCTTCACTGGGCTGCACACAAGGGTAACTACCGCCTTATGAAGCT
GCTATTGTCTCGTGGTGCTGATTGTCGGGAGAAAGACAATGAAGGGCAGACCCCTCTCC
ATTTGTGCACACGACACAAACTGCCCAAGTGCATGGCTATACTGCTGCGACAGCTGCTG
CCTGGAGAGATTGACAATCAAGATAATAACAAGCGGACAGCCTTGCACTGGGCGTCATC
CTATGGCAACCTAGAACATGTCAAGATGTTGATAAAGCCGGACTCCAACATAGGAATCC
CTGACATAGAAGGGAAGACTCCGTTACACTGGGCTGCCAGCAGTCGCGACCCTGAAGCT
GTGCAGTGCGTCAGACTCATCCTGGAGACCACGCCCAGTGTCATTAACTGGCAGGACTA
CGAAGGTCGCACTGCCCTTCACTTGGCTGTTGCTGATGGCAACACAAGTATTGTCGATC
TCTTGACATCAGTAGAGAAATGCAATGTTTCTGCTCTAGACAATGTGTTTCGCACACCT
CTTCATTGGGCTGCCGTTTTAGGGCACTCTGCTATCGTGGGCCAGTTACTGCAGAGCAA
TGCAGACTCAACCAGTTCAGACTCAAATGGTGCTACTCCTCTTCATTATGCAGCTCAGA
ATAATTTTGATGAAACGGTGAATGTGTTCTTATCAAGGAAAAATACTGATGAAGCGGA
TTTGGAGGGCCGCACAGCGTTTATGTGGGCAGCGGGCAAAGGTGCTGACAACGTGTTAC
GGGTATTTTTTCATCACGGCGTCGATGTACAGCAGGAGGACAAAACTGGTGGCACGGCA
TGGTGCCAGGATCAATGGCCACGGATCAAGTCAAGCATACGCCCTTGTTTCGGGCGTGT
GAGATGGGGCACACTGACGTTGCACAAGCCCTCATCGATAATGGAGCACGGAGTGGAT
GTTTTGGATGAAGACGGGCGATCCCCTTTACATTGGGCAGCTTTAGGAGGTCATGCTTA
CATCTGCCAGATTTTGATCAAATATGGTATTGATCCAAATATCCGTGACCATTCTGGAC
GTTCTCCTTTGCACTGCGCAGCTTATGGTGGCTATGTGAATTGTATGTCTGCCTTGATG
GAACACCAAGCTGACCCCAACACGCAGGATAAAGAGGGCATGGCTTGCCTCCACTGGGC
ATGTAGTAAGGGCCATCTGGATGCTGTGAAGCTACTGGTAGAAAATGACGCCTTTCCC
AATCACATGGAGTTCAGAGAAGAAAGATACACTCCGTTGGACTATGCCTTGATGGGGG
AGCACCACGATGTGGCGCAATACATGATTGAACAAGGGGCTTTGTCAATCACCGGCATT
CAGGACATTGCTGCCACCAAAATACAGTCAGCATTCCGAGGCTACAGGGTGCGGAAAAC
TTTCATAGAACGTAAAAAACTGTTGATGAAACATGAGATGCTGAAGAAAGAAGCTGCA
AGAAAAAGAGCAGGTGAAGGGAAAAAAAGTGAGGAACAGCTGAGGAAAGAATCAGA
AGACCTGGCGATGGAGCCTCTCGTTGCACCGTTACATCATCAGCTTCCTCAGAAAA
CACAAACTCCATCTCCCCCACACCAGAAATCTCCATCACGAGAACAGCAGCAATCAC
CAACAAAGCAGTTAGAAACAAAGAATCCAGAAACAGAGTCACAGAAAAAGCAAAA
GGAAACAGTGCAGGCACAGACAAAGCAAGCATCAGAACAGCCACAAACAAAGGAGC
AGGTTTCAGAACCCACTCAGACAAAACAAGAAGAAACTGAGCAGTCAGAGATAGCA
CAACAGACAACAGAAGAGCCACATATAACACAACAAGAATCTGAACAGACACAAAC
CACACATGAAGAACAAAAAACAGAGCAGTCACAGACAGTGCAAAATGACGCACAAC
AGAAACATGCAGAGCAAAAGGCATCAGAGGTAACTCCATTAAATTCAGAAAAAAG
TGAAGAGGTACAGATACAGCAACAAGAAGCTGAACATAGGCAAGCAAAGGAACAG
GAATTGAAACAACCAGATACAAAGCAACCAGAGAAAGCACACCAACAAACAGACCT
ACAAGAAGCAGAACAAACACTCACAGGGCAACAAAAATCAGAACAACAACAAGTACA
ACAAGAAAAGACACAACAACTTAAGGACAGTCAGGAAATTTCATCAACATCCGACACA
CAACAGCAAGCACCGGCTTCTCCAGTGTCTTCATTTGGGGATAAGGCAGACCCACCGAA

GGACGCTCATGTTAGACACAGGACAGATTCTGACAAACAGAAACATGCTGGGAGAGTG
TCCAAGGAATCAAGGAGTGTCCAAGGTCGTAAAAAGAAGGACAAAGGGCCAAGTGCGG
AAGAAAAGGTGCTACAGACAAAGTCTGACAAGGCAAATGTGAAAAAGGGGGAGAAAAA
ACGCAAAGGAGACATCCAGAAAAAACAACATGCCGCCTCAGTCATCCAAAGAGCATGGC
GAAAGCACAGACGCTCAAGACAACGACACCGCCAGCAGCCCATGGTGACAGAATCTTCT
CAGAATGAAGTCAGGATGCGGCGACTTCAACTATCTCAGGCTGCGTGGGAAAGACAAA
TTGCCGCCCTGACAATTCAGCTGGCCTGGAGGAAGTTCTATAGACGGAAACTGCTGAAA
TCATTGAATCCGAACCGTAGACAGCTTCGGATGTGGGACCCTGACATTATAGCCCTCAA
GCAGCAGGCTCTGATTGCAGAAATTTATAGTGAACAGCTGATGGTACCCTTCTGGCACC
CCACCTTGAAGCCTTCCGTCCGCCCCTTCTGGTTCCGCCTCGTCCCCTCTCCTGCAGCGG
TGTCATTCAACTTTGCTGTGGACCACTATCATCCGTTGAAGCTGTGGCAAGGCATGCGG
CCACATTCATGGACTGGCCCAGTTAACGGCACTGTCGATTTAGATTTTGTCCCCAAATC
TGCTTCCTCTCACCGCCCAGAACGCTCAGCGAATTCGCAATTTTCCTACAATCTTAGGA
GAGCTTGA

>*inversin_Biomphalaria_glabrata*
ATGGTTTCCCATTACTTGTGTCATTGCTGTGTTTCTGCTCGCTTCTTGGGACTTTATAC
AGAAGCACAGAGTGAAACACCAAGAACACCATTCACATCTAAGCCTGAAGGTCAGAGG
TTAGAAGTAGTGACCCAGATCCATGCAGCTGCTGTGAATGGAGAGAAGCAGACTCTAC
TCAGACTAATTGCAGCTTCAGCTTCCAATGGAGATCTTGATGTTGGTGATCAGTTTGGA
CGTAGTCCACTCATGTACTGTGTACTGGGGGACAGGCTGGAGTGTGCCGAGGTTCTTCT
CAAGACTGGGGCCCAGGTGAACCTGAAAGACAAAGGAGGGAGGACTGCCCTTCACTGG
GCAGCACACAAGGGAAATCTTCGATTTATGAAGCTTTTGCTTTCTCGGGGCGCAGACTG
CCGTGAGAAGGACAACGAAGGTCAGACTGCTCTGCATCTGTGCACGAGACATAAGTTAC
CCAAGTGCATGGCCTTGTTGCTGAGACAACTGACTCCTGGGGAGATAGATGATCAGGAC
AAAAATAAAAGAACTGCTTTACACTGGGCTGCTTCTTATGGCAACTTGGAACATGTCA
AGATGCTTATTAAACAGGATTCCAACATTGGTATACCAGACACAGAAGGCAAGACACC
ATTACACTGGGCTGCAAGTACTAGGGACAGAGAAGCTGTTAGCTGTGTGAAACTGATT
TTGGAAACCACACCTAGTGTTATAAATTGGCAAGACTATGAAGGTCGAACGGCTTTGC
ACCTTGCAGTAGCTGATGGAAATGAGAGTGTTGTTAATGTTCTGACTACTGTCAAAAA
CTGCAATGTCTCCGCTTTAGATAATGTGTTTCGAACTCCGCTTCATTGGGCAGCAGTCT
TAGGTCATTCCAAGATAGTGTCTATTTTATTAGACAGTGGGGCTGACAGTTCAAGTTCT
GACTCCAATGGAGCGACAGCACTGCACTACGCAGCTCAGAATAATTTCAGTGATACTGT
TTCAGTGTTCTTGAAACACAACTGTACCGATGAAGCGGATCTAGAAGGACGTACTGCTT
TCATGTGGGCCGCTGGTAAGGGTGCCGATGATGTCATCCGAGTTTTCCTAGAACACAAA
GTTGACATACACCAAGAGGACAAAACTGGTGGAACATCACTTCATGCTGCAGCATTGTC
TGGTCACAGTTCTACAGTTAAACTCTTGCTGGACCATGGGGCCCAGCTAGAGGTAACGG
ATCAGGCAAAACATACCCCTTTGTTCAGGGCATGTGAGATGGGCCACATGGATGTTGTC
CAGACGCTGATCGACTGCCATGCCAGAGTGGATGTGCTGGACCAGTATGGAAGGTCACC
ATTACACTGGGCAGCCCTAGGTGGACATGCTTACATATGCCAGCTCTTGATAAAGTATG
GTGTGAACCCAGATGTTAGGGACCACTCTGCAAGAGCTCCACTCCACTGTGCTGCATTT
GGTGGCTTTGTCAACTGTATGTCTGTATTGATGGAGCATGAAGCAGATCCCAATTCTCA
GGACAAAGAAGGCTTGACTTGTCTACACTGGGCATGTTCCAAAGGTCATCTTGATGCTG
TAAAACTACTCATTGAATACAATGCTTACCCCAATCATATGGCTTTCACAGAGGACAGG
TACACTCCTCTTGATTATGCATTAATGGGAGAGCACCATGAGGTAGCTCAGTATATGAT
AGAGCAAGGGGCATTATCCATCACAGGGATTCAAGACATTGCTGCCCAGAAGATCCAGT
CAGCTTTCCGAGGTTACAGAGTGAGAAAGGCTTTTATAGAAAGGAAAAAATTATTTAT
GAAACATGAAAAGTTGAAAAAAGAAGCAGCTAGAAAGAGAGCTGGAGAAGAAGGAAA
GAGAACTATGGAGAAACAGACCTCATTCGGGAAATCAATTTCCTTCGATAAACAGACA

TCATTTGATAAACAGACTTCTTTAGACAAGCAGACTTCTTTGGATAGGCAGACGCCATT
CGATTCCCAGACTTTGGCTCAAGCAGACAGACTGACCGTTTCCCACACAACTACAAGCA
GAAATGCTGAAAAACTTCCAGAAAATAGTTCTAAAACTTTACAAGCAGAATGCAACCA
TTTACATAACCTTGAAATTGCAAGTGGTCACATGACTTCCCCAGCTAACATCCAATCCA
ACATAAAAAATCACACAGCTGATTCAGAAAATTATGATCAAAATTCTAACACTAGGAA
TATCAATGAAAAAAGAAATGCAGACATCAATAGTCCAAAACAAACGCAAGAATCTCAA
AATGACAAACAGCCAATGAAAACTGAACTAAAGGTTCATAAGCCCCGCCTTCGTGCAA
GCCAGAGAAGACGTCTGCTACAAGCAGATCAAAACAAAGGATCATTACTAGGCGTTAA
TCAAACTGTCATTGCAAGTGATAAACTGTGTTAA

>*nodal1_Crepidula_fornicata*
ATGCTGGCCTGCAGTCCTGTAGGGACAATCACAGCTTTCCATCCTTCTGACGTTCTTTC
TGGAAAGCAGCGAGGTCCTGTTCGTTTTGGAGCGCCTCGCCTGTCTTCTGTGGACATGG
AGAAGATTAGCAAGGAGGAGGTGAACTACCATCTGTATCAGAACGGGCGTCTGGATGC
TTGGCACCAGCGTAAAGCCTCCACCGTGTTCCTGGTGGACCTCTTCAACCACCTGGTTC
GGGGTCATCCAGTAGCATCTCACAGACGCCACTCACCAGCCAGGGCACGCATGGGCAGT
AACGCCCATCTCCCTTTTACTGATACGATTCGGAGCTTCTCTGCAAACGAGGTAACATT
CAGCCCCAAGAAGAAGGCCACAGTTGAGTTCCCCATGCCTTTCCTGCCGCCCAACGAAC
GCCTAGGGTCCGCGGAACTACGCTTCCTACGGCGAACGAATACGTCCAAGATACGGAGC
AAGACTCGCTACAAGCTGCCCATGACAGTGATTCGCGGTGGACGTGTAATGAAGAGAA
CGGTGCTGAGGGAACGTGCTGTGCGTCAGAGAGAATACCATGTGTTTGACGTGAGCCG
ATTTCTAGACACCTGGATTAACTCATACCACGGCAACATCACACTTCAAGTGCGCATCC
CACGAAAACTGGAGTCATCTATCCAGCTGAGGAACACTTCTCTCACCTCTTTGATCGCT
CTGTATTTAGAAGACCAGGACTTCCTCAAAAACATGTATGAAAGTTACACTACGAACA
AGTTAAAGCAGCAAACATTGAACTCCCACGAGTATACTCAAGTCAAGAATCTGTCAGA
TCTTCAGATATCTCATACTCGTGAAAAGAGACATATGAGGTCTCAGATACTAGATTCGC
CACAACAGAGAGATACAAGATCTCGAGTTAAAAGAGACTCAAGCCGTCGGAAATCAAG
ACATCGTAAAAATAAACGATGGTATCGTGCCTCAAAACGTGCCAAATGTCGGATGTAT
AATTTTAAGGTGGACTTTGATTTCATCGGATGGGGACAATGGATTATTCAACCTAAAG
AATTCAACTCGCGATTTTGTTACGGAAAATGCCCATCCCCTGTAGAGAGCAAGTACAAA
CCGACCAATCACGCCATGCTGCAGACTCTGATGAGAGAGAAGATGCCCCGCCTGGCACC
CCCCACCTGTTGTGCTCCTGTCAGACTCAAGCCGCTTAGTATGTTGTACTTTGAGTACG
ACGAAATTGTGGTCCGCCACCACGATGACATGATTGCCGATCAGTGTGGGTGTAG

>*nodal2_Crepidula_fornicata*
ATGCCATTATTTTCGGCAGCCATACGTGCGCAAGGAAAGTGTTATTTCAGAGCACAATG
TGTCGGCGGAGAGTGGTTGCTGCACTGGTTGGTGCTGACAATGCTGTTATGGACATCGT
CGTGCCCCGTGATGGCCATTACACTACCCGTGGCTTTGCCCCCTTCCCACCCTAAATATT
ACTTGACCGTGGTAGGTCAAGACTCAGATGTCCAGTCCAGCACGACCCCGGCCCAAACC
CACAACTCTAACCATGTTCCCGTGGACGTTCAGGGTTTGGGGATCCGGGAGGACCAGGT
GATGTACCATTTCTACAAGAACGGCCGGCTGGGTTCAGGCTACTATCGTAAGGCGTCCA
CCGTGTTCATGGTTCGCCTCTTTAACGATCTCATTCGAGGCCTAGCTCCTGTTCCCTCGG
CCCGTAGTGTGTCGCTAGGGCGCAACAAGGACGCTAGCGACGAGGACCACCCTTTCACG
GACACCATTCGGAGCTTTTCTGCTGGAGAGGTAAAATTTTCCAAGAAAAAGAAAGCCA
CAATCCACTTCTTCGTGCCTTTCTTGCCGCCCAACGAACGCCTTCGCTCTGCGGAAATCC
GCTTCTTGCGTCGTCCAGGAGCCCGAAAGAAGAAGCGACGGAAACGCTACAAGCTGAGA
ATCACGGTGTTGCGAGGAGGTCGGGTAGTGGAAAAGGTCCTTCTCCGAGAACGTCCCGT
GACTCAGCAAGACAACGATGTGTTCGACATTACTCGGGTCCTCGATTCCTGGATCAACT
CCCACCACGGAAACGTCTCGCTCAAGGTTCGAATCCCTCGCAAACTAGAGAAATCCATC

TTGCTGAGTAACGCCTCGCTGAAATCTACATCCCTAGTGGCGTTGTATCTAGAAGATCG
GGAGTTCTTGAAAAATATGTACGAGAGTTACACAACAGACGATCTTCGTGGAAACCGT
AGTGATGCTCGAGCCAGAAGCCCTGGGGGTGGGGAGGTGGGTCATAATAAACACAGGT
CTCGTGACAAGAGAGCATCTACTACCTCTAGGACCAGTAGGGAATCTAAACGTCGCAAA
AACAGAAACAATAAAAAGAAACGTTGGTATCGAGCCCCCAAGCACGAAAAATGCCAGA
TGTACGACTTTAAAGTGGATTTTGATTTCATTGGGTGGGGTCAGTGGATAATCCACCCT
AAGAACTTCAACGCGAGGTTTTGTTACGGCGAGTGTCCGTCTCCCGTGGATAGTAAGTA
CAAGCCGACCAACCACGCCATGCTGCAGACTCTGATGAGGCAAAAGAAGCGCCGTCTGG
CACCGCCTACTTGCTGTGTGCCCATCAAACTCAAACCGTTGAGCATGCTGTACTTTGAA
TTTGACGAGATTGTGGTTCGCCACCACGAAAATATGATAGCTGATGAGTGTGGGTGTA
GGTGA

>*egf-cfc_Crepidula_fornicata*
ATGGACGTGTCGTCTGTCAGGGGTCTGATCACGCTGGTCATACAGCTCCACACCATTGC
CGTTTTTTCACAAGCGGTGTCGTCACAAAGCCCACAGACGAAGTCGAAGAAGACGCTGG
GAGACTTGAACCCTAGACCCTCCAAGTCGTGCTGCCTGAACGGGGGCCTGTGTGTGCTG
GACAGCTTCTGCCATTGCCCCAAGAAGTTTTACGGGCGCTACTGTGAGTATGAGCTGTC
AGAGAAGCCATGTGGTCCTGTGCCTCATGGCAGCTGGGCGAAGGCAGGATGTAACTTGT
GTCGCTGCTTTGACGGTCACATGACCTGCCTGCCACGTGCCTTCGATGGATGTGACGAC
AAGCCGCTGACGGAGGAACAACCTCAAAACATCTGGGACTACCAAGACGAAATGGTCA
TCTTCGCCAACCAGGGCTCCACCACCAGTAAATACGACGCCGAAGATGCTGACTACTAC
GACTACTACGACGATAACGGTGGTGGTGATGGTGGTGGTGGTGGCGTTAGACATTCTG
CTGGTGCCAGACTGGTGCCCGGGCTGTGCTGTGTGGCTGTGTGGAGCATGCTGTGTGTG
CTCGTTACGATACTCTGGACCTCATGGACACTCGTGTCATAA

>*egf-cfc_Biomphalaria_glabrata*
ATGGCAAACCTTTTCATCATGTGGCTAGTAATGATAAGACTCGCCAAGTCGGACCAGGG
AAATTATACAGATCTACATGAGTCTATTACCGGCAACAACGTGGTCAGCCACCCATCCC
TTCTACTCTCTATGTTGACACCAACTACTGAGACATCCAAGCATCAAGAAGAGGGACAG
TCTAGCGAAGTCAAGCAGAAGTGTTGTGAAAATGGAGGCCTTTGTGTTATGGGAGTCT
TCTGTCATTGTACAGGAAAATATCATGGCCAAAGATGTGAATTTGAAAGAAAACCCTG
TGGTGATATCGCCCATTACCAGTGGGCCAGATTAGGATGTAACTTGTGTAGATGCTTTG
ATGCTAAGATGCACTGTTTAACAAGAGTCTACGAAGGCTGTGATGACAAGCCACCAAA
AACAGATGTTGATACCTCTGATTACCTGAATGATAGGGAGATATTCCCAGCAGACAGT
GACAGTGATTATGATAATTATGGAAGTGGGGATTATAGTTACAATGAAGATGACAGGG
TTCAAAGATCAGGCTCTTCTAAGCTTATCACCCACTTACTCACAATGCTGTTGGTAATA
GTGACTGTGATGTTTTTGATTGGCTCCCCATGGCAAGTGTTACCATTCACTTTTGTGAT
TGAGTGCTCTTAG

Apéndice 9: Secuencias aminoacídicas de las proteínas de la familia EGF-CFC de las distintas especies analizadas en esta tesis

Entre corchetes se indican los "GenBank; accession number" o el código identificador de la proteína. Para el resto, la secuencia se obtuvo de las bases de datos especificadas en el apéndice 2.

>*Novocrania_anomala*_EGF_CFC
MSPATRLVLVGIFPFLTLEVLAATLPGVPLISRQMRAANPQGITKEVRKNCCENGGVCVL
GSFCHCKSDYYGRYCEHHSKQRSCGLVDHDEWLRLHCNLCQCFDGILTCMPQIFDDQIC
DKNLTATFKINYRKHHKKFETKQNVDLEKMSSLQDLVNYSPDFRFSPVMLMLTCITSVI
MRR

>*Patella_vulgata*_EGF_CFC_1
MDKIITCLMFVAMVIQFVETRQTQDVPRWLQISTKTVKIRYHHGDIIPKKIQQHVEIIRM
FNSSKSSTSCCLNGGRCIMGSFCQCPKNFYGRRCEHKNVQKPCGRLMHGVWVKSDCNLC
FNGLMNCIPISNPGCDAERTKQDQMRDVYIFPPKRNVYRDHVRKEDDLQSAAVQTMNR
LYLLTIVLYILT

>*Patella_vulgata*_EGF_CFC_2
MDQIVVRSLLVSMVIQFAESKHTQDIPRWLQISSKTVKVRHRHGMPKEIQQHVEIIQMF
NSSKSSTSCCLNGGRCILGSFCQCPKNFYGRRCEHKNVEKPCGRLMHGVWVKSDCNLCR
CFNGLMNCIPISNPGCDAERTKQDQMRDVYIFPPKRHVYRHVRKENVPQSAAVQTMNR
LYLLTIVLFVLC

>*Lottia_gigantea*_EGF_CFC [Prot ID 233075 estExt_fgenesh2_pg.C_sca_340150]
MDISYLTRCVLLLVGIILTTEFVESKKETPKWIKRSLKRERIRHHHGGFMKKQSSESFDNP
QAYKATCCKNGGRCIMGSFCHCKKNYYGRYCEHKSDNKNCGKFKHGVWLKSGCNLCRC
FDGVLSCIPRAYSGCDDKPVHIRDKYNPDDEVYVFPTETEDKVEEDSNYYDVTPSDAVNI
SKSWISVLSAVVLSFRIIYR

>*Crepidula_fornicata*_EGF_CFC
MDVSSVRGLITLVIQLHTIAVFSQAVSSQSPQTKSKKTLGDLNPRPSKSCCLNGGLCVLDS
FCHCPKKFYGRYCEYELSEKPCGPVPHGSWAKAGCNLCRCFDGHMTCLPRAFDGCDDK
PLTEEQPQNIWDYQDEMVIFANQGSTTSKYDAEDADYYDYYDDNGGGDGGGGGVRHSA
GARLVPGLCCVAVWSMLCVLVTILWTSWTLVS

>*Owenia_fusiformis*_EGF_CFC
MRTSNMISVLWLVFITGVSGHSIESGKRGQHKNHPEKRDRVNEEITNEISTNPGRKHHN
RANGGTSKKVKITKKSKKDCCNGNGICYLGNFCHCDHRYYGRYCEYYLEHEFCGPEIEHG
TWHPTGCFMCFCFNGEMDCMRQKRFDTCTMEEPFPVFPWEQNSVQDTDTVTHANKK
THQHTKKTNHHNKKTHHHNKETNEYNTGTHHRTHHALFITLCLMVIQLLYIF

>*Biomphalaria_glabrata*_EGF_CFC
TEDVLNRTVDVTKHPRRCDVSKSMANLFIMWLVMIRLAKSDQGNYTDLHESITGNNVV
SHPSLLLSMLTPTTETSKHQEEGQSSEVKQKCCENGGLCVMGVFCHCTGKYHGQRCEFE

RKPCGDIAHYQWARLGCNLCRCFDAKMHCLTRVYEGCDDKPPKTDVDTSDYLNDREIF
PADSDSDYDNYGSGDYSYNEDDRVQRSGSSKLITHLLTMLLVIVTVMFLIGSPWQVLPFT
FVIECS

>*Branchiostoma_floridae*_EGF_CFC [EGF-CFC XM_002207525]
MRALNLPGSHVRAPRQVTLHEHLMKPANHSADCCENGGTCILGSFCMCPANFTGRYCE
QHAVTLPCGDVPHNDWMFQGCSLCRCGNGTFLCIEHMRPNCESERDVVNVVNREGNL
NPLSVLDLFDQSRN

>*Danio_rerio*_Oep [NP571167]
MTSQLFGFLMFAVIICQAVSLESGCEGSECVKVGVSGKPKQYAEFLNKFNEMNTQTPQR
QHRNAEAALPFVGLTGVAKQSRTCCKNGGTCILGSFCACPKYFTGRSCEYDERLRDCGVI
PHGEWVQKGCSYCRCGYGLLHCFPHVFSKDCDVFSKDCDDSQEVRWHRSGSLRTLSSTI
VMFATFILHRLL

>*Takifugu_rubripes*_Oep [ENSTRUP00000013088]
FLDQFSQVSSSNSGDRKHRDASAVLPFVGLTGSKSLSRSCCKNGGTCILGSFCACPPFFTG
RNCEYDQRIRSCGLISHGEWVQKGCSYCRCGYGILHCFANIFHKDCVERRHSHIILLIPEIV
VLRKATLLPTLLMSFVLIICCFLI

>*Xenopus_laevis*_Xcr1 [AAA83569]
MQFLRFLAILIFSAKHFIKHCKGETCMGLNCNDPRLLEAIKSNTINQLLHDTINATHGKSP
PKSTKTLPFLGITDSKKLNRKCCQNGGTCFLGTFCICPKQFTGRHCEHERRPASCSGVPH
GDWIRQGCLLCRCVSGVLHCFKPESEDCDVVHEKNMRSGVPRMQLSLIIYCFLTANLFYH
IVWHLNIGL

>*Xenopus_laevis*_Xcr2 [CAI15753]
MLCRDHGRFICITLALLTFYCCKGLALEAAEYSNKGSSNATKEDLLINHNRTIDTFYHLN
KDNEKRNQHKTEGLVPFIGLTDSSKLSKHCCNNGGTCVLGSFCVCPRYFTGRHCEYDER
AKHCAAKIQHGDWIRKGCRLCRCAYGVLHCFVETQTDCDEVEEEEIHSSAPMEQSSIYLI
AFGILFCIYTLV

>*Xenopus_laevis*_Xcr3 [NP_001089135]
MVWLKRLFYLTVQLSAVIRLGTPKQNGTMINAGSELSEDTYVQKLLHHVDHAGKEKSSV
ETIPFIGLTKNHALDKHCCKNGGTCVLGSFCACPKHYTGRYCELHVHNRKCGIVPHGHW
VQKKCALCRCMYGTMHCFPSGDCDAKDYVEESQIFLSSSPRVRSCLIAWFFLIAASNTLS
H

>*Xenopus_tropicalis*_Xcr1 [ENSXETP00000020142]
FLFYRFLAILVFSAKYFIKHCKGESCVGLYCNDTGLSLAIKSNTIHQILQDTINATHGKSPE
KSAKTLPFLGITDSKKLNKKCCQNGGTCFLGTFCICPKQFTGRHCEHERRPASCAGVPHG
DWIRQGCLLCRCVSGVLHCFTPESEDCGPVHEKNMRSGVPRTQLSLFIYYLFIANLFYHIV
WHLNIGL

>*Xenopus_tropicalis*_Xcr2 [ENSXETP00000020145]
NTEKNFCFKIILKIIVQILKLGSGCNHPHCPPVHTSGAVETHNRTIDTFYNLNKDNEKRN
HHKTEPLVPFVGLTDSSKLSKQCCNNGGTCVLGSFCVCPRYFTGRYCEFDERAKHCAAKI
QHGEWMRKGCRLCRCTYGVLHCFVETQTDCGKLLGQGG

>*Xenopus_tropicalis*_Xcr3 [NP_001078829]
MINTGSELSENTYIQKLLQHVDHNEKETIPFIGLTKNNALDKHCCKNGGTCVLGSFCACP
KHFTGRYCELHVHNRKCGIIAHGHWIQKKCALCRCMYGTMHCFASGDCDAKDYVKDTH
IFLSRSPSVSYCLTSWFLVIAASNTLLH

>*Homo_sapiens*_Crc [NP_115934]
MTWRHHVRLLFTVSLALQIINLGNSYQREKHNGGREEVTKVATQKHRQSPLNWTSSHF
GEVTGSAEGWGPEEPLPYSRAFGEGASARPRCCRNGGTCVLGSFCVCPAHFTGRYCEHD
QRRSECGALEHGAWTLRACHLCRCIFGALHCLPLQTPDRCDPKDFLASHAHGPSAGGAP
SLLLLLPCALLHRLLRPDAPAHPRSLVPSVLQRERRPCGRPGLGHRL

>*Mus_musculus*_Crc [NP_031711]
MTWRHHARFLFTVSLALQIINLGNSYQREKHNGRREEVTKVATQKHQQSPLNWTSSHS
GEVTGSAQGWGPEEPLPYSRAFGEGASARPRCCRNGGTCVLGSFCVCPAHFTGRYCEHD
QRRSECGALEHGAWTLRGCHLCRCVFGTLHCLPLQTPGSCDPKDFLASHAPGRGAGGAP
SLLLLLPCALLHHLLRPDAPAHPRSLVPSVLQRERRPLRKAGTWASPFIFTVVNNRCV

>*Loxodonta_africana*_Crc [ENSLAFP00000005045]
MIRGHHVRLLFMISLALQIIHLGKSYQREKHKGGKEEINNATTQKFQQKTLHWTWTLN
NFSEANGSAQGWRRPPDSPGFRDSVSTQSSCCRNGGTCVLGSFCVCPAHFTGRRCEHDP
KRSECDAHAHGAWTVHDCRLCRCVYGALHCLPRTPGPCDLKDFLTSHSNGLSSQHTLSF
LILLPCFLLCLVLCGHLSGKEHPRSDIRSITEPEEWHRARAIPG

>*Erinaceus_europaeus*_Crc [ENSETEP00000015228]
LMRGHRVRLLFRISMALQIIPMRNSYQREKYEGDKEKISNATTQSHQQKPHNWTFNNFS
DVTEGGGNCDSLPRFWGLHKPSRRPRCCLNGGTCILGSFCVCPSHFTGRHCEDQRRECG
AHAHGAWIVLSCRLCRCVYGALHCLSRQTQGFC

>*Monodelphis_domestica*_EGF_CFC [ENSMODP00000037671]
MAWIYHVRRLFLITLVFQVIHLGNGFNEKEKLKGSAEEPINTTAPKQQPPNQNPTWNST
RDLNRSSESPKQFQSRPFIPFIGFKDGTKLCCQNGGTCVLGSFCVCPMHFVGRHCEHDER
KSNCGVFAHGDWVYKGCYLCRCIYGKLHCFRDPTLNNCDLNEEFVTLQSNGSRLQYTLY
FLLPFSCFLLQWIVLQK

>*Gallus_gallus*_EGF_CFC [NP_990031]
MFWRKHVRILFTVTLIWQAIHLGKGREKEHEKDVKNFNDTAQKQSPKNSVTIIDAFSD
MNQSYQSRKQQNSREFVPFTGITESKNLNRNCCQNGGTCILGAFCACPKHFSGRHCELR

KCGSIIHGDWVMKGCWLCRCLYGTLKCLSQNTQDGCELRREEEIIRLYSNGLRLQQTMS
ALICLLTFLLELCCWQL

>*Mus_musculus*_Cro [AAH52646]
MGYFSSSVVLLVAISSVFEFGPVAGRDLAIRDNSIWDQKEPAVRDRSFQFVPSVGIQNSKS
LNKTCCLNGGTCILGSFCACPPSFYGRNCEHDARKEHCGSILHGTWLPKKCSLCRCWHG
QLHCLPQTFLPGCDGHVMDQDLKASRTLCQTPSVMTTFMLAGACLFLDMKV

>*Loxodonta_africana*_Cro [ENSLAFP00000007663]
MDRRKMECFSSSVIFIMAFSKAFELGLVAGLGHQELASSPQGDADIRSQEEPVVRHQPFQ
FVPFMGIQDSKELNKTCCLNGGTCMLGSFCACPPYFYGRNCEHDVRKENCGSVPHDTW
LPRKCSMCKCWHGWLRCFPQTFLPGCDGRVMDEHLEASRTPELLPSVCTTLMLASLCL
STLLY

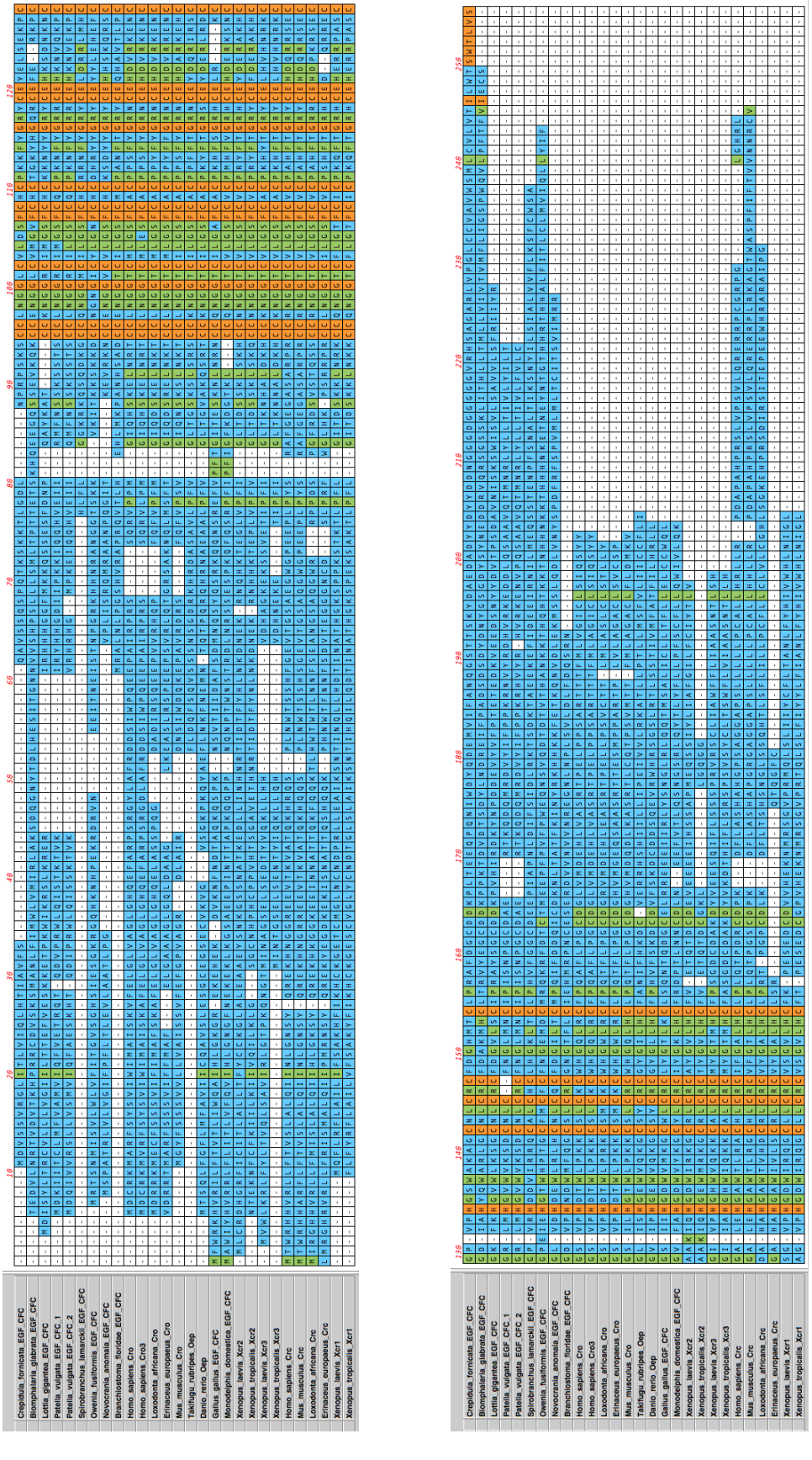
>*Erinaceus_europaeus*_Cro [ENSEEUP00000004109]
VDRRTMARFSSSVIFFMAFELGAAAGLGQLDAAGLKEASLRSPEEPSVRLQTGRSAKNQL
MSFTGIQDSRELNKTCCLNGGTCMLGSFCACPPYFFGRNCEHDVRRENCGSVPHDTWLP
KKCSMCKCWHGWLRCLPQTFLPGCDGHVMGEQSAASRTTESMPPARTPLVLAGACLSV
CSP

>*Homo_sapiens*_Cro [NP_003203]
MDCRKMARFSYSVIWIMAISKVFELGLVAGLGHQEFARPSRGYLAFRDDSIWPQEEPAIR
PRSSQRVPPMGIQHSKELNRTCCLNGGTCMLGSFCACPPSFYGRNCEHDVRKENCGSVP
HDTWLPKKCSLCKCWHGQLRCFPQAFLPGCDGLVMDEHLVASRTPELPPSARTTTFML
VGICLSIQSYY

>*Homo_sapiens*_Cro3 [P51864]
MDCRKMVRFSYSVIWIMAISKAFELGLVAGLGHQEFARPSRGDLAFRDDSIWPQEEPAI
RPRSSQRVLPMGIQHSKELNRTCCLNGGTCMLESFCACPPSFYGRNCEHDVRKENCGSV
PHDTWLPKKCSLCKCWHGQLRCFPQAFLPGCDGLVMDEHLVASRTPELPPSARTTTFM
LAGICLSIQSYY

>*Pomatoceros_lamarckii*_EGF_CFC
HLFLGGKRKQSQDKGCCQNGGMCILGSFCHCRRNFYGRYCELDRRLMHCGPVRHDDVI
RRQCALCHCLHGKLTCLPHVFGGCDEEPIAPNLDIFSIDLRKTSDKTAEEKEFIDEEIRNIL
MEQSSTENPFSNALTIKSNYILSIALVVFLKSFCWSA

Apéndice 10: Alineamiento completo de las secuencias aminoacídicas de las proteínas EGF-CFC identificadas



Agradecimientos

Agradecimientos

Escribir agradecimientos siempre es difícil porque uno no quiere terminar teniendo un discurso infinito como el de Almodóvar, pero tampoco quiere dejar de agradecer a todos aquéllos que de algún modo han contribuido a que su tesis haya sido posible… Y, bueno, ya me conocéis… Haré lo que me parezca jajaja No obstante, como siempre nos terminamos dejando a alguien, pido perdón de antemano.

Primero quiero agradecer a Cristina, mi boss, por la oportunidad que me dio para hacer la tesis con ella; dije "si no es contigo, no hago tesis", y parece que mi insistencia funcionó!! Sé que todos los estudiantes de doctorado dirán lo mismo: "la tesis es como una montaña rusa, lo pasé fatal"… y una vez me dijiste "Esto, es como una carrera de fondo"… Cris, esto ha sido un verdadero Ironman!!! Ha habido muchos obstáculos en el camino y sólo tengo palabras de gratitud hacia ti, por no dejar de confiar en mí, por empujarme a explorar mis límites dándome a la vez una seguridad brutal y reenfocarme cuando me dispersaba, a no tener miedo a equivocarme, por darme la oportunidad de conocer grandes investigadores y darme la máxima formación y experiencia posibles, por darme un apoyo que ya quisieran muchos estudiantes tener por parte de sus directores, por enseñarme tantas lecciones dentro y fuera del labo, por los ratitos de charla y por tus "historias de abuela cebolleta" que a mí me encantan… Has sido más que una directora de tesis porque me has tendido la mano y has tirado de mí cuando parecía que no iba a poder, algo que no es nada fácil de encontrar. No cambies, sigue dando guerra y sigue trasmitiendo esa pasión por lo que haces (gracias a Pepe Bella, que es una gran guía y supo perfectamente con quien juntarme!); eres una gran investigadora con una visión del mundo y la ciencia que se debería oír más… Y esto por resumir un poco; de momento "sólo" gracias por todas las cosas que me has ayudado a alcanzar, entre ellas, este futuro best-seller jeje. Por cercanía, aprovecho aquí para agradecer también a la gente del Museo (Carolina, Lourdes, Rafa, Íñigo y Jorge (gracias por enseñarme a maquetar *in extremis*, eres un crack!!!), Mario…) por su ayuda, apoyo, consejos y su agradable forma de ser. También aprovecho aquí para agradecer a los miembros de mi tribunal el haber aceptado venir a evaluarme a pesar de esta horrible fecha!!! Gracias por el esfuerzo, de verdad, os lo agradezco infinitamente.

Second, thank you very much to a very special lab for me… Prof. Jonathan Henry lab, that greeted me warmly the half part of this thesis (Jon, Kim, Alvin, Paul, Surabhi, Tayyab, Cornelia, Marty, Ashley and Elena). There are so many things to thank and so great moments, that I don't know how I could start! Jon, my mentor, you are one of the most special people I've ever met; a great scientist, enthusiastic and crazy child-mind, always with interesting questions, always creating. Your passion for science is contagious, you are like a Jedi!… and also a very generous and unselfish person. Thanks for always making me feel part of your lab, for helping me like I were one more of your students, for teaching me, for Woods-Hole!, for caring about my thesis so much… To me, you are my second advisor, despite it's not written on any paperwor. Kim… nearly my third advisor!! The most efficient person I've ever met in a lab; you may have secret super powers because it's simply awesome the way you work. You are a beautiful person, so thoughtfull and kind; thanks for being so patient, for teaching me all about microscopy and protocols for Crepidula, and working with the gooey frogs hehe and thanks for supporting me so much. Thanks to Alvin and Paul!! I enjoyed so much sharing lab with you, Zipi & Zape!! Smart people, always interesting and funny conversations going on, also made me feel part of the lab, volleyball on weekends with Paul's team, the "welcome display" on my desk every year, creating words in spanish ("desktopo" is my favourite)… you're great guys, I have really nice memories with both of you and "Bret" hahaha. Thanks to Surabhi and Tayyab!! You are worderfull labmates

too, so helpfull, so nice with me, we laughed sooo much… best wishes for you and your thesis!! We will keep in touch, lab!!

Also thank very much to Jon, Alejandro Sánchez-Alvarado and Richard Berigner for inviting me to the Embryology course at the MBL in Woods-Hole (US), giving me the chance to do crucial experiments for this thesis, and also meet so many brilliant investigators. There are no words for describing the experience. I also want to thank here to all the people from 2016 year, for so many crazy moments!! Also people that I met on 2013, specially to Nathan, who also contributed importantly to this thesis, assembling our super helpful trancriptomes, for example!! Also helped me the first year that I arrrived to US, at Woods-Hole! Thanks a lot.

Aprovecho este momento para agradecer a todos mis amigos de Urbana, Illinois, ahí "in the middle of the corn"! Feli, Bea, Kat, Jin, Karina, Pedro, Camila, Chris Kinson, OneKey!, Dani, Pericles, Piyush… y tantísimos otros que me dejo sin nombrar… Que habéis sido parte de mi vida a rachas durante 4 años, como los amigos del pueblo jajaja pero gracias por hacer que cada viaje para volver se hiciese menos y menos duro, siendo siempre igual de acogedores conmigo. Gracias por todas esas noches salseras, por los interesantes debates, por esas barbacoas con bailoteo incluido, por los Thanksgivings y viajes exprés a Chicago sólo para bailar 3 horas!! En especial, a mi pequeña familia Felipe y Bea, que han sido y son como dos hermanos mayores para mí, a los que quiero mucho; os deseo lo mejor porque no merecéis menos. Beits, gracias por hacerme estas portadas tan chulas!!! Eres una artista!!

También quiero agradecer a mis nuevos compis de labo (Irene, Nuria, Pablo, Fran, Luis, Miguel Ángel, Claudia; y Turi, Claudia, Ana…). Sois geniales, de verdad… sólo empezando a escribir esto me estoy riendo porque sois "la alegría de la huerta" jajaja Y a pesar de haber compartido con vosotros poco tiempo de la tesis, ha sido intenso y os tengo que agradecer el súper acogimiento (conmigo, con Cris e incluso con Mike cuando ha venido)… espero haberos devuelto al menos un poquito durante mi estancia!! Irene y Nuria, gracias por ser tan buenas compañeras; suerte con todo, aunque estoy segura de que os irá bien. A los polluelos, con los que tanto tanto tanto me he reído cuando más tensa estaba, mucha suerte, valéis "un potosí", sois unos espartanos! Seguid con esa ilusión, que os llevará lejos, pero a portarse bien que os pongo un punto negro a todos en menos que canta un gallo!!!! Jejeje

Gracias a gente del CBM que aportó su granito de arena para que esta tesis saliese adelante. Gracias a ese maravilloso 414, el labo de Isabel Guerrero (Isabel, Pedro mi gran compi de bench, Irene, Laura, Gus, Carmen, Adrián, Julia, Nicole…) por los que siento un gran respeto, aprecio y gratitud, por hacerme sentir una más y acogerme cuando lo necesité. Gracias también a Curro, que me ha salvado más de 2, 3 y 4 veces… más paciencia que el santo Job; gracias por tus consejos, tus ánimos de teniente-coronel, por tu apoyo y por tu desesperante sentido del humor!!! Gente con la que compartí parte de la tesis y tengo gratos recuerdos: Javi, Óscar, Cris, "Los Baonza". También quiero agradecer a Flor su maravilloso trabajo con los peces; te deseo mucha mucha suerte en tu nueva etapa, eres una gran investigadora, y es muy fácil y agradable trabajar contigo.

Gracias a todos mis compis y amig@s de Bio… Inés y sus polluelos, Elena, Sandra, JuanPa, Anuska, Marta, Pablo… no he compartido mucho tiempo con vosotros, pero habéis aportado mucho a mi salud mental jajaja sois muy grandes. Aquí hago especial mención al "Equipo Actimel", Inés e Irene; chicas, lo hicimos!!!! Felicidades a las tres!!! También aprovecho aquí a

agradecer a mis amigas (ay Fuckencias mías jaja) Mimi, Almu, Patri, Maite, todas, y recordar a mi amigo Sergio… por nuestras noches de Dr. Why con cerve 0.9% jajaja también a Lago, por vuestra amistad y por escuchar a pesar de no entender mucho, pero decirme "que soy lista" y "que voy a por el Nobel" jajaja Os quiero mucho; gracias por estar siempre tan pendientes y darme ánimos. Gracias a todos mis amig@s con los que comparto mi afición (o más bien vicio) por el baile (Manu, Luigi, Álex, Rober, Luis, Iris, Coral, Esther, JanCarlo…) por el desestrés y la alegría que provocáis! Este verano toca celebrar! Por supuesto, a mis YonKizomberos!! Albir y mis compis de clase; nunca han faltado reconfortantes palabras de ánimo y vuestro cariño… Gracias, de verdad por vuestros empujoncitos y estar pendientes. Gracias a mis Bailongos, l@s chic@s de la clase con Sergio!! Con los que me divierto siempre y rápido me han recargado las pilas para escribir.

Nos acercamos al final, y es momento de agradecer a mi familia, mi padre y mi hermana Ana. Gracias por creer en mí siempre, por vuestro cariño, por apoyarme y animarme a pesar de que mis bichos no sean muy agraciados jajaja; esta tesis también es vuestra. Gracias sobretodo por proveerme de tanto helado de menta con chocolate que casi vaciamos los mercados del país!!! Thanks, Mike, for stay by my side and being my wall, always saving me... yup, you created an injection system just for me, so I could stop crying at Woods-Hole hahaha Thanks for making these thousand kilometers feel way more less; it will be even less soon. Thanks for being "my cheerleader"!! También gracias a Merche por su enorme contribución; tus directrices y tu ayuda han sido bastante críticas para conseguirlo, y por ello, muchas gracias. El último es para la persona a la que está dedicada esta tesis y no llegó a verme empezar, pero plantó la semilla que llevó a que fuera posible… por inculcarme los mejores valores, por fomentar constantemente mi curiosidad por el mundo, por enseñarme a no ser conformista y sí a ser valiente, a que si quieres algo debes luchar como gato panza arriba… y cuya imagen y recuerdo me ha dado fuerza para que esto llegara a buen puerto; gracias a mi madre.