



CENTRO DE INVESTIGACIONES BIOLÓGICAS  
DEL NOROESTE, S.C.

---

---

Programa de Estudios de Posgrado

GENÓMICA DE LA ESPERMATOGÉNESIS EN LA  
ALMEJA MANO DE LEÓN *Nodipecten subnodosus*

TESIS

Que para obtener el grado de

Doctor en Ciencias

Uso, Manejo y Preservación de los Recursos Naturales  
( Orientación Acuicultura )

Presenta

**RAÚL ANTONIO LLERA HERRERA**

La Paz, Baja California Sur, Julio de 2013

## ACTA DE LIBERACION DE TESIS

En la Ciudad de La Paz, B. C. S., siendo las 12 horas del día 28 del Mes de Junio del 2013, se procedió por los abajo firmantes, miembros de la Comisión Revisora de Tesis avalada por la Dirección de Estudios de Posgrado del Centro de Investigaciones Biológicas del Noroeste, S. C., a liberar la Tesis de Grado titulada:

**"Genómica de la Espermatogénesis de la Almeja Mano de León *Nodipecten subnodosus*"**

Presentada por el alumno:

**Raúl Antonio Llera Herrera**

Aspirante al Grado de DOCTOR EN CIENCIAS EN EL USO, MANEJO Y PRESERVACION DE LOS RECURSOS NATURALES CON ORIENTACION EN ACUICULTURA

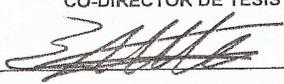
Después de intercambiar opiniones los miembros de la Comisión manifestaron su **APROBACION DE LA TESIS**, en virtud de que satisface los requisitos señalados por las disposiciones reglamentarias vigentes.

### LA COMISION REVISORA

  
DRA. ANA MARÍA IBARRA HUMPHRIES  
CO-DIRECTOR DE TESIS

  
DRA. SILVIA ALEJANDRA GARCÍA GASCA  
CO-DIRECTOR DE TESIS

  
DRA. NORMA YOLANDA HERMÁNDEZ SAAVEDRA  
CO-TUTOR

  
DR. ILIE SAVA RACOTTA DIMITROV  
CO-TUTOR

  
DR. ARNAUD HUVET  
CO-TUTOR

  
DRA. ELISA SERRIERE ZARAGOZA,  
DIRECTORA DE ESTUDIOS DE POSGRADO

**COMITÉ TUTORIAL**

Dra. Ana María Ibarra Humphries  
Dra. Silvia Alejandra García Gasca  
Dra. Norma Yolanda Hernández Saavedra  
Dr. Ilie Sava Racotta Dimitrov  
Dr. Arnaud Huvet

**COMITÉ EVALUADOR DE TESIS**

Dra. Ana María Ibarra Humphries  
Dra. Silvia Alejandra García Gasca  
Dra. Norma Yolanda Hernández Saavedra  
Dr. Ilie Sava Racotta Dimitrov  
Dr. Arnaud Huvet

**JURADO DE EXAMEN**

Dra. Ana María Ibarra Humphries  
Dra. Silvia Alejandra García Gasca  
Dra. Norma Yolanda Hernández Saavedra  
Dr. Ilie Sava Racotta Dimitrov  
Dr. Pedro Cruz Hernández  
Dr. Dariel Tovar Ramírez (suplente)

## RESUMEN

A pesar de los grandes avances en las tecnologías de secuenciación, la información genómica disponible para especies marinas no-modelo de interés ecológico, evolutivo y económico es aún escasa. El objetivo de este trabajo fue identificar genes expresados durante la espermatogénesis en la almeja hermafrodita funcional *Nodipecten subnodosus* (Mollusca: Bivalvia: Pectinidae), con el propósito de obtener un panel de genes que pudieran a futuro permitir el estudio de transcripción diferencial de genes entre almejas diploides y triploides en el contexto del arresto meiótico y esterilidad reproductiva.

Dado que la principal meta era aislar genes involucrados en meiosis y otros procesos relacionados a la maduración testicular, se generaron genotecas de hibridación sustractiva de testículo vs gónada inactiva. Dos aproximaciones de secuenciación fueron empleadas, obteniéndose 325 y 177 ESTs mediante secuenciación capilar (Sanger) y mediante el uso de pirosecuenciación (454-FLX) se maximizó la cantidad de ESTs resultando en 34,276 lecturas. Un total de 1,153 genes de la genoteca de testículo tuvieron un hit por blastx y una anotación ontológica, incluyendo genes de meiosis, espermatogénesis, diferenciación sexual y elementos transponibles. Algunos de los genes de meiosis identificados poseen función en el apareamiento de cromosomas (*scp2*, *scp3*), recombinación y reparación de ADN (*dmc1*, *rad51*, *ccnblip1/hei10*) y en *checkpoints* de meiosis (*rad1*, *hormad1*, *dtl/cdt2*). Los análisis de expresión para genes de meiosis en distintos estadios gametogénicos en ambas regiones sexuales de la gónada confirmaron que la expresión fue específica o incrementada hacia el testículo en maduración. Los genes de espermatogénesis incluyeron algunos reconocidos como testis-específicos (*kelch-10*, *shippo1*, *adad1*), y algunos de ellos se conoce que están relacionados a esterilidad. Los genes de diferenciación sexual incluyen a uno de los genes más conservados en el comienzo de la cascada de determinación sexual (*dmrt1*). La identificación de transcritos codificantes para elementos transponibles, transcriptasas reversas, y transposasas de la genoteca de testículo en maduración son evidencia de que la transposición es un evento activo durante la espermatogénesis en *N. subnodosus*. En relación a la genoteca de gónada inactiva, se identificaron 833 transcritos con anotación funcional relacionada a la maquinaria de transcripción y traducción de proteínas, y más importante, al mantenimiento y control de la línea germinal (*notch homolog 2*, *btg member 2*, *btf3 homolog 4*, *hes-1*, *Tcf21*).

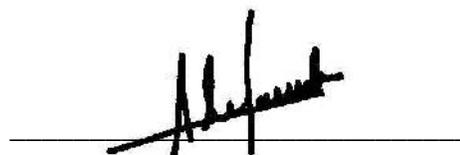
El segundo objetivo fue establecer la mejor estrategia para analizar la expresión por PCR cuantitativa, definiendo el impacto de la elección al azar de genes de referencia en la normalización de la expresión génica y las inferencias derivadas de esto. Para especies no-modelo, con información mínima o inexistente, la cuantificación relativa de la expresión génica requiere de información preliminar que incluya el aislamiento de genes de referencia potenciales y la identificación de aquellos que se expresan establemente bajo las condiciones biológicas de interés. En éste trabajo, secuencias parciales de cuatro genes fueron aisladas de ADNc de tejido gonadal de *N. subnodosus* para ser evaluados como genes de referencia: *18S-rRNA* (*18S*), *riboproteína l8* (*rp-l8*), *actina-* (*act-*), y el *factor de*

*elongación 1* (*ef-1*), con una quinta secuencia codificante para una *alfa-tubulina* (*tub-*) encontrada en las genotecas SSH que también se agregó como un gen potencial de referencia. Los análisis de estabilidad de estos cinco genes usando geNorm y NormFinder indicaron que *18S* junto con *rp-18* fueron los genes más estables para la normalización de expresión génica durante el desarrollo gonadal en ambas partes sexuales. El gen menos estable fue *tub-*, que mostró una expresión sesgada entre regiones sexuales de la gónada, por lo que fue analizado como un gen blanco. La expresión relativa, estimada mediante la normalización con la combinación de *18S* y *rp-18* como genes de referencia, indicaron que en el avance de el desarrollo gonadal, el gen *tub-* fue sobreexpresado en la región masculina, indicando que se trata de una isoforma testículo-específica. Análisis posteriores de su expresión en distintos tejidos indicaron que *tub-* es expresado en forma elevada y específicamente en gónada masculina, aunque la expresión en tejido de músculo aductor fue también observada pero en niveles significativamente menores que en la región testicular. Las tubulinas han sido extensivamente usadas como genes de referencia, pero en este trabajo se ha demostrado que algunas tubulinas pueden no ser adecuadas como genes de referencia. Otro gen aislado de forma dirigida, un miembro de la familia de las DEAD-box ARN helicasas (*ddx*) no mostró cambios en su expresión durante el desarrollo de la gónada o entre regiones sexuales, y fue empleado para analizar su expresión y discutir las distintas inferencias estadísticas que resultan de el uso arbitrario de genes ‘elegidos al azar’ durante la normalización de la expresión génica.

Palabras clave: Espermatogenesis, genómica funcional, meiosis.



Dra. Ana María Ibarra Humphries



Dra. Silvia Alejandra García Gasca

Co-directoras

## ABSTRACT

Despite the great advances in sequencing technologies, genomic and transcriptomic information for marine non-model species with ecological, evolutionary, and economical interest is still scarce. This work was aimed to identify genes expressed during spermatogenesis in the functional hermaphrodite scallop *Nodipecten subnodosus* (Mollusca: Bivalvia: Pectinidae), with the purpose of obtaining a panel of genes that would allow in the future for the study of differentially transcribed genes between diploid and triploid scallops in the context of meiotic arrest and reproductive sterility. Because the principal aim was to isolate genes involved in meiosis and other testis maturation-related processes, suppressive subtractive hybridization libraries of testis vs. inactive gonad were generated. Two sequencing approaches were used, obtaining 352 and 177 ESTs by Sanger sequencing, and by using pyrosequencing (454-Roche) we maximized the identified ESTs to 34,276 reads. A total of 1,153 genes from the testis library had a blastx hit and GO annotation, including genes specific for meiosis, spermatogenesis, sex-differentiation, and transposable elements. Some of the identified meiosis genes function in chromosome pairing (*scp2*, *scp3*), recombination and DNA repair (*dmc1*, *rad51*, *ccnblip1/hei10*), and meiotic checkpoints (*rad1*, *hormad1*, *dtl/cdt2*). Gene expression analyses in different gametogenic stages in both sexual regions of the gonad of meiosis genes confirmed that the expression was specific or increased towards the maturing testis. Spermatogenesis genes included known testis-specific ones (*kelch-10*, *shippo1*, *adad1*), with some of these known to be associated to sterility. Sex differentiation genes included one of the most conserved genes at the bottom of the sex-determination cascade (*dmrt1*). Transcript from transposable elements, reverse transcriptase, and transposases in this library evidenced that transposition is an active process during spermatogenesis in *N. subnodosus*. In relation to the inactive library, we identified 833 transcripts with functional annotation related to activation of the transcription and translation machinery, but more importantly, to germline control and maintenance (*notch homolog 2*, *btg member 2*, *btf3 homolog 4*, *hes-1*, *Tcf21*).

The second goal of this work was to establish the best strategy for analyzing gene expression by quantitative PCR, defining the impact that reference genes randomly chosen have on normalization of gene expression and the inferences derived from that. For non-model species as many used for aquaculture, with minimal or no genomic information, relative quantification of gene expression studies requires preliminary research including the isolation of potential reference genes and the identification of those stably expressed under the biological conditions of interest. In this work four partial gene sequences were isolated from gonad tissue cDNA in the functional hermaphrodite scallop *Nodipecten subnodosus* to be evaluated as reference genes: *18S-rRNA* (*18S*), *riboprotein 18* (*rp-18*), *actin-* (*act-*), and *elongation factor 1* (*ef-1*), with a fifth partial sequence from an *alpha-tubulin* (*tub-*) found in the libraries also included as a potential reference gene. Stability analysis for the five putative reference genes with geNorm and NormFinder indicated that *18S* together with *rp-18* were the most stable genes for normalization of gene

expression during gonad development in both, male and female sexual regions of the hermaphrodite *N. subnodosus*. The least stable gene was *tub-*, showing a biased expression profile between sexual regions of the gonad; therefore this gene was analyzed thereafter as a target gene. Relative expression, estimated by normalization with the combination of *18S* and *rp-18* as reference genes, indicated that as gonad development advanced, *tub-* in the male region was up-regulated pointing toward this gene being a testis-specific *-tubulin* isotype. Further analyzes of gene expression among tissues indicated that *tub-* is specifically and highly expressed in the male gonad, although expression in adductor muscle was also observed but at significantly lower levels. Tubulins have been extensively used as reference genes, but in this work it was shown that some tubulins might not be suitable as reference genes. Another gene found in the libraries, a putative member of the DEAD-box RNA helicase family (*ddx*), showed no changes in expression during gonad development or between sexual regions, and was chosen to analyze gene expression and discuss the different statistical inferences resulting from the arbitrary use of 'randomly-chosen' reference genes when normalizing gene expression.

Keywords: spermatogenesis, functional genomics, meiosis.

## AGRADECIMIENTOS

Agradezco al Centro de Investigaciones Biológicas del Noroeste, y en especial a todo el personal de la Dirección de Estudios de Posgrado, dirigido por la Dra. Elisa Serviere Zaragoza, por el apoyo incondicional para el desarrollo de mi trabajo doctoral, así como por atender las solicitudes de apoyos complementarios para talleres, cursos y congresos. Mi agradecimiento al Consejo Nacional de Ciencia y Tecnología, por la beca de doctorado con registro 204031, así como por la Beca Mixta de Movilidad Nacional otorgada durante la convocatoria 2012. Agradezco también a los siguientes proyectos que financiaron y/o facilitaron la realización de éste trabajo:

- Proyecto SEP-CONACyT 166573 (2012-2015): ‘Genómica del molusco pectínido *Nodipecten subnodosus* en condición diploide y triploide: Mecanismos genéticos de determinación sexual, del control de la fidelidad de la meiosis y del crecimiento muscular’.
- Proyecto SAGARPA-CONACyT 126427 (2010-2012): ‘Aplicación de la genómica funcional como estrategia para la Mejora continua de la industria del camarón – Módulo Reproducción’.

Agradezco ampliamente a mis co-directoras: la Dra. Ana María Ibarra Humphries, por todo el apoyo, atención incondicional y por brindarme siempre todas las oportunidades y recursos necesarios para mi formación y para el desarrollo de este trabajo; y a la Dra. Alejandra García Gasca, por las oportunas recomendaciones en el desarrollo de mi tesis y para mi desempeño profesional. A mis tutores: Dres. Ilie Racotta Dimitrov, Norma Hernández Saavedra y Arnaud Huvet, por estar atentos de mis avances y por las recomendaciones dirigidas en cada reunión semestral.

A Susana Ávila, que como responsable de laboratorio de Genética Acuícola, me permitió desarrollar la mayor parte del trabajo contenido en éste documento, y siempre estuvo al pendiente de que todo salga en tiempo y forma, y me dio muchas lecciones de administración y prevención a la hora de trabajar en un laboratorio. También agradezco ampliamente a José Luís Ramírez Arce y su equipo de trabajo, quienes facilitaron la obtención de material biológico, apoyando salidas de campo, muestreos y brindando lecciones en el mantenimiento de organismos. Definitivamente el apoyo de ambas partes fue muy profesional e importante. También agradezco a mis compañeros de laboratorio, por el apoyo en el trabajo diario y por la retroalimentación constante, así como a los Dres. Pedro Cruz Hernández y Ricardo Pérez Enríquez por sus consejos y recomendaciones.

A Cristina Escobedo Fregoso, muchas gracias por la paciencia y por ayudarme con todo el ánimo posible para que este trabajo comience, continúe, y culmine de la mejor forma.

Al Laboratorio de Histología: M.C. Carmen Rodríguez Jaramillo y a Eulalia Meza Chávez por el apoyo en el procesamiento histológico de muestras.

A la Dra. Gracia Gómez Anduro, Julio Hernández González, así como a todos los estudiantes del laboratorio de Biología Molecular de plantas, que me facilitaron el espacio, materiales y reactivos requeridos para obtener las genotecas sustractivas.

A los Dres. Ricardo Vázquez Juárez, Eduardo Romero Vivas y Fernando Von Borstel Luna, por el apoyo en el trabajo realizado en el servidor de bioinformática del CIBNOR, y al Dr. Cei Abreu Goodger del LANGEBIO, por su asesoría en el procesamiento y análisis informático de datos de secuenciación masiva.

A Marcela Vélez Alavez, por su compañía y apoyo durante una de las etapas más difíciles de este trabajo.

Y finalmente, quisiera agradecer a mi familia y amigos cercanos, por el ánimo siempre ofrecido, y por compartir conmigo el enorme gusto que significó realizar éste trabajo.

## ÍNDICE DE CONTENIDO

<b>COMITÉ TUTORIAL</b> .....	<b>I</b>
<b>RESUMEN</b> .....	<b>II</b>
<b>ABSTRACT</b> .....	<b>IV</b>
<b>AGRADECIMIENTOS</b> .....	<b>VI</b>
<b>ÍNDICE DE CONTENIDO</b> .....	<b>VIII</b>
<b>LISTADO DE FIGURAS</b> .....	<b>XI</b>
<b>LISTADO DE TABLAS</b> .....	<b>XIII</b>
<b>1 INTRODUCCIÓN</b> .....	<b>1</b>
<b>1.1 Triploidía en la almeja mano de león y otras especies</b> .....	<b>2</b>
<b>2 ANTECEDENTES</b> .....	<b>4</b>
<b>2.1 Biología reproductiva</b> .....	<b>4</b>
<b>2.2 Regulación de la meiosis durante la gametogénesis</b> .....	<b>5</b>
<b>2.3 Efectos de la triploidía sobre la gametogénesis</b> .....	<b>7</b>
2.3.1 Posibles causas de la esterilidad o de la inhibición de la meiosis en triploides. ....	8
<b>2.4 Conocimiento genómico en el grupo de los moluscos</b> .....	<b>10</b>
<b>3 JUSTIFICACIÓN</b> .....	<b>12</b>
<b>4 OBJETIVO GENERAL</b> .....	<b>15</b>
<b>4.1 Objetivos particulares</b> .....	<b>15</b>
<b>5 HIPÓTESIS</b> .....	<b>16</b>
<b>6 MATERIALES Y MÉTODOS</b> .....	<b>17</b>
<b>6.1 Material biológico</b> .....	<b>17</b>
6.1.1 Utilizado en la generación de genotecas sustractivas y caracterización de transcritos de genes de interés .....	17
6.1.2 Utilizado en el análisis de validación de genes de referencia y de expresión por qPCR.....	18
<b>6.2 Extracción de ARN</b> .....	<b>21</b>
<b>6.3 Generación de genotecas por hibridación sustractiva</b> .....	<b>22</b>
6.3.1 Secuenciación de genotecas obtenidas por hibridación sustractiva.....	24
6.3.2 Análisis bioinformáticos de secuencias obtenidos en genotecas .....	27
<b>6.4 Aislamiento <i>de novo</i> de transcritos de genes</b> .....	<b>28</b>
6.4.1 Síntesis de ADNc para obtención de transcritos parciales.....	29
6.4.2 Síntesis de ADNc para amplificación de transcritos por RACE.....	29
6.4.3 Amplificación por PCR de punto final para obtención de secuencias parciales de genes .....	30
6.4.4 Amplificación por RACE-PCR de extremos 3' y 5' para transcritos de interés ..	31
6.4.5 Análisis filogenéticos y estructurales de genes ' <i>full-length</i> ' .....	34
<b>6.5 Análisis de expresión génica</b> .....	<b>35</b>

6.5.1	Síntesis de ADNc para la validación de genes de referencia .....	35
6.5.2	Cuantificación de expresión de genes de interés .....	36
6.5.3	Análisis de estabilidad de genes de referencia entre estadios de gametogénesis .....	37
6.5.4	Análisis de expresión de genes <i>ddx</i> , <i>vit</i> y <i>tub-a</i> durante la validación de genes de referencia .....	38
6.5.5	Determinación de expresión relativa mediante cálculo de cantidad relativa (RQ) y análisis estadísticos. ....	39
<b>7</b>	<b>RESULTADOS.....</b>	<b>40</b>
<b>7.1</b>	<b>Generación de información genómica asociada a la reproducción en <i>N. subnodosus</i> .....</b>	<b>40</b>
7.1.1	Identificación de genes expresados diferencialmente en la gametogénesis mediante generación y secuenciación de genotecas de hibridación sustractiva ..	40
7.1.2	Análisis de secuencias de genotecas de hibridación sustractiva. ....	43
7.1.3	Identificación de microsatélites en fragmentos de genes aislados en genotecas de hibridación sustractiva .....	50
<b>7.2</b>	<b>Caracterización de transcritos de genes selectos por RACE .....</b>	<b>51</b>
7.2.1	Transcrito de <i>mre11</i> .....	51
7.2.2	Caracterización de transcritos de <i>dmc1</i> y <i>rad51</i> .....	51
7.2.3	Caracterización de transcrito de <i>pch2</i> .....	57
7.2.4	Caracterización de transcrito de <i>scp3</i> .....	57
<b>7.3</b>	<b>Definición de estabilidad de genes de referencia para la normalización de la expresión génica .....</b>	<b>60</b>
7.3.1	Secuencias parciales de genes de referencia. ....	60
7.3.2	Estabilidad de genes de referencia.....	62
7.3.3	Análisis de expresión de genes <i>ddx</i> , <i>vit</i> y <i>tub-a</i> durante la validación de genes de referencia .....	67
<b>7.4</b>	<b>Ubicuidad de expresión de <i>tub-</i> en distintos tejidos.....</b>	<b>70</b>
<b>7.5</b>	<b>Análisis de expresión de genes asociados a la meiosis .....</b>	<b>71</b>
<b>8</b>	<b>DISCUSIÓN .....</b>	<b>74</b>
<b>8.1</b>	<b>Información transcriptómica asociada a la gametogénesis y obtenida <i>de novo</i> usando como estrategia genotecas sustractivas y secuenciación por 454-FLX ....</b>	<b>74</b>
8.1.1	Descubrimiento, caracterización y expresión de genes selectos e involucrados en la meiosis, recombinación y reparación de ADN durante la gametogénesis .....	78
8.1.2	Descubrimiento de genes de diferenciación sexual y genes específicos de la espermatogénesis .....	84
8.1.3	Descubrimiento de elementos transponibles y enzimas relacionadas en la genoteca de región testicular en maduración .....	86
8.1.4	Genes descubiertos en la genoteca de gónada inactiva están involucrados en el control y mantenimiento de la línea germinal .....	88
<b>8.2</b>	<b>Caracterización de secuencias completas de genes en meiosis por RACE .....</b>	<b>91</b>
<b>8.3</b>	<b>Elección de genes de referencia para la normalización de la expresión génica en gónada de <i>N. subnodosus</i> .....</b>	<b>92</b>
8.3.1	El transcrito del gen ribosomal <i>18S</i> en <i>N. subnodosus</i> sí es amplificado utilizando oligo-dT .....	92

8.3.2	Ventajas y desventajas de los programas GeNorm y NormFinder en la estimación de la estabilidad de genes de referencia.....	94
8.3.3	Descubrimiento de un presunto gen de referencia, <i>tub-</i> , siendo testis- y músculo-específico .....	96
8.3.4	Las tubulinas tejido- o testículo-específicas .....	97
8.3.5	Consecuencias del uso de genes de referencia arbitrariamente elegidos en la normalización de un gen de interés - ejemplo aprendido en el análisis de la expresión de <i>ddx</i> .....	99
<b>9</b>	<b>CONCLUSIONES .....</b>	<b>100</b>
<b>10</b>	<b>LITERATURA CITADA.....</b>	<b>102</b>
	<b>ANEXO I. ....</b>	<b>116</b>
	<b>ANEXO II.....</b>	<b>128</b>

## LISTADO DE FIGURAS

Fig. 1. Eventos de cambios morfológicos propios de la profase 1 .....	6
Fig. 2. Morfología de complejo sinaptonémico (SC) en diploides y triploides.....	10
Fig. 3. Anatomía de la almeja mano de león <i>N. subnodosus</i> .....	18
Fig. 4. Micrografías de los estadios reproductivos evaluados en éste trabajo .....	20
Fig. 5. Esquema de desarrollo de genotecas de hibridación sustractiva .....	23
Fig. 6. Diseño de oligonucleótidos 454-nested 1 y 454-nested 2R, empleados para la adición de adaptadores LIB-L en los extremos de los fragmentos de PCR mediante la reacción anidada.....	26
Fig. 7. Esquema de amplificación de extremos 5' y 3' por RACE en ADNc SMART .....	33
Fig. 8. Verificación de la adición de adaptadores en ADNc realizada en ADNc de músculo esquelético (reacción control) .....	41
Fig. 9. Reacción de amplificación primaria (carriles A) y anidada (carriles B) con oligonucleótidos dirigidos a los adaptadores .....	41
Fig. 10. Amplificación de clonas seleccionadas al azar.....	42
Fig. 11. Prueba de amplificación con oligonucleótidos anidados del sistema Clontech y de oligonucleótidos anidados con adaptadores Lib- .....	43
Fig. 12. Análisis de enriquecimiento de términos ontológicos para genotecas sustraídas recíprocamente entre gónada en maduración testicular y gónada inactiva, sobrerrepresentados en la gónada de testículo .....	46
Fig. 13. Análisis de enriquecimiento de términos ontológicos para genotecas sustraídas recíprocamente entre gónada en maduración testicular y gónada inactiva, sobrerrepresentados en la gónada inactiva.....	47
Fig. 14. Alineamiento múltiple de las secuencias de ortólogos de RAD51A, RAD51C, DMC1 y XXCR3 de especies de invertebrados y vertebrados .....	53
Fig. 15. Árbol filogenético de secuencias de proteínas RAD51A, RAD51C, DMC1 y XXCR3, de especies de invertebrados y vertebrados .....	54
Fig. 16. Alineamientos estructurales del modelo proteico de RAD51 en <i>N. subnodosus</i> contra modelos de RAD51A y RAD51C de <i>Xenopus tropicalis</i> . .....	56
Fig. 17. Alineamiento múltiple de las secuencias de ortólogos de PCH2/TRIP13 de especies de invertebrados y vertebrados .....	58
Fig. 18. Árbol filogenético de secuencias de proteínas PCH2 de especies de invertebrados y TRIP13 de vertebrados .....	59
Fig. 19. Estabilidad de genes de referencia entre regiones gonadales y estadios .....	64
Fig. 20. Valores de estabilidad para los genes de referencia obtenidos por geNorm en cada set de ADNc .....	65
Fig. 21. Valores de estabilidad para los genes de referencia obtenidos por NormFinder en cada set de ADNc .....	66
Fig. 22. Variación relativa de los valores de amplificación ( <i>Ct</i> ) de los genes de referencia entre muestras de gónada inactiva, y en testículo y ovario en estadios de desarrollo temprano y avanzado .....	67
Fig. 23. Expresión relativa de <i>vit</i> , <i>tub-a</i> y <i>ddx</i> en ambas regiones sexuales de la gónada, normalizada con los genes 18S y rp-18 .....	69

- Fig. 24. Expresión relativa de *ddx* estimada después de normalizar con genes individuales (*18S*, *rp-l8*, *efl-1*, o *act-1*), con la combinación de los dos más estables (*18S/rp-l8*), o con la combinación de los cuatro genes de referencia ..... 70
- Fig. 25 Expresión relativa de *tub-a* en diferentes tejidos/órganos de *N. subnodosus*, normalizada con la expresión de los genes de referencia a *ef-1a* y *rp-l8*. ..... 71
- Fig. 26. Expresión relativa de genes seleccionados de la genoteca de testículo en maduración, en distintas condiciones gametogénicas ..... 73

### LISTADO DE TABLAS

Tabla I.	Genes asociados a la meiosis, expresados en la genoteca de región testicular de la gónada en maduración.....	48
Tabla II.	Genes asociados a espermatogénesis y diferenciación sexual expresados en la genoteca de región testicular de la gónada en maduración.....	49
Tabla III.	Genes de elementos transponibles expresados en la genoteca de región testicular de la gónada en maduración.....	50
Tabla IV.	Secuencias de primers para la amplificación por qPCR de genes de referencia y genes de interés.....	61
Tabla V.	Resultados obtenidos en los alineamientos por blastx de la traducción conceptual de las secuencias parciales obtenidas por secuenciación dirigida.....	61
Tabla VI.	Estadísticas descriptivas de los valores de amplificación del set de genes de referencia sobre ADNc sintetizado con oligo-dT y hexanucleótidos aleatorios.....	63
Tabla VII.	Secuencias de primers para la amplificación por qPCR de genes de interés en meiosis.....	72

## 1 INTRODUCCIÓN

La almeja mano de león (*Nodipecten subnodosus*) es una especie de molusco bivalvo pectínido que habita en la costa Oeste del Pacífico. Se le ha adscrito una distribución desde la península de Baja California hasta Perú (Keen, 1971), aunque otros estudios indican que su distribución se limita a la frontera sur del Golfo de California (Smith, 1991), con otra especie, *Nodipecten arthriticus*, estando presente hacia el Pacífico Sur. Recientemente Petersen *et al.* (2010) definieron la estructura poblacional de *N. subnodosus*, encontrando una diferenciación genética entre la almeja distribuida en el Pacífico y en el Golfo de California. Adicionalmente encuentran dos clados a nivel mitocondrial, uno de los cuales se encontró en mucha menor frecuencia en el Golfo de California que en el Pacífico, y aunque una diferenciación a nivel de subespecie no logró hacerse, tal trabajo sugiere la necesidad a futuro de profundizar en estudios para establecer la posición taxonómica del género *Nodipecten*.

Esta especie es explotada bajo el esquema de captura controlada de poblaciones naturales, y no bajo el esquema de cultivo acuícola de organismos producidos en laboratorio a partir de reproductores obtenidos de poblaciones naturales. En un contexto productivo, *N. subnodosus* presenta un alto valor comercial a nivel regional debido al gran tamaño que alcanza su músculo aductor, y es capturado principalmente en la costa del Pacífico de la península de Baja California, siendo la población de la laguna Ojo de Liebre la única capaz de sostener una pesquería constante. La única otra población que se localiza dentro del Golfo de California, desde Bahía de Los Ángeles hasta Bahía de La Paz (Petersen *et al.*, 2010), es demasiado pequeña para sostener una pesquería (Massó-Rojas, 2000, citado en Ramírez-Arce, 2009). Recientemente, la captura de organismos con fines de explotación pesquera se ha visto comprometida por mortalidades (Villalba-Vega, 2011). Dada la importancia económica y la comprometida situación de las poblaciones naturales para explotación, el cultivo de la especie en sitios específicos, bajo condiciones sanitarias adecuadas, podría resultar a futuro en una verdadera alternativa para el aprovechamiento de ésta especie en ambientes donde su cultivo es propicio (Racotta *et al.*, 2003, 2008; Ramírez-Arce, 2009). Para esto, la generación de conocimiento asociado con la

reproducción para la mejora en la producción larvaria es de suma importancia, así como también es importante el desarrollo y optimización de biotecnologías que mejoren las capacidades de cultivo de la especie.

Uno de los problemas principales en la producción en laboratorio de la almeja mano de león resulta del uso de un reducido número de reproductores (Petersen *et al.*, 2008; Petersen *et al.*, 2010), generándose larva con alta consanguineidad, aunado al hecho de que ésta es una especie hermafrodita funcional, que en condiciones normales produce ovocitos y esperma simultáneamente (Rupp *et al.*, 2011). La reducida variabilidad genética en larva de laboratorio puede a su vez generar un impacto genético sobre las poblaciones naturales cuando estas cohabitan y se reproducen en conjunto. Una posible solución a este problema es el desarrollo de cultivos de organismos triploides o estériles, como medida de confinamiento.

La inducción a la triploidía ha sido extensamente aplicada en moluscos bivalvos (Hulata, 2001; Dunham, 2004). En el caso de la almeja mano de león, resulta en una reducción casi total en el desarrollo de la gónada, aunque no necesariamente en una mejora en crecimiento (Maldonado-Amparo *et al.*, 2004; Ramírez-Arce, 2009). Esto es, bajo ciertas condiciones ambientales, principalmente donde la disponibilidad de alimento es limitada (Ramírez-Arce, 2009), la energía normalmente dedicada a la gametogénesis puede redirigirse al crecimiento del músculo aductor en triploides (Racotta *et al.*, 2008), confirmando esto una ventaja sobre los diploides

### 1.1 Triploidía en la almeja mano de león y otras especies

La condición de triploidía se define como la ocurrencia de tres juegos de cromosomas homólogos en el genoma de un organismo, y representa una condición genética de interés en la acuicultura de moluscos bivalvos, ya que generalmente se asocia a un mayor crecimiento del organismo (Piferrer *et al.*, 2009). Particularmente en bivalvos pectínidos (almejas, escalopas), se ha documentado mayor crecimiento del músculo aductor (Tabarini, 1984; Ruiz-Verdugo *et al.*, 2000, 2001; Yang *et al.*, 2000). Esta condición de poliploidía es manipulable en laboratorio, induciéndose en bivalvos generalmente mediante el bloqueo de

la extrusión del segundo cuerpo polar por agentes físicos o químicos (Scarpa *et al.*, 1994; Piferrer *et al.*, 2009), aunque tecnologías alternativas de producción de triploides “naturales” han sido desarrolladas mediante apareamiento entre individuos diploides y tetraploides, éstos últimos siendo obtenidos por bloqueo de la extrusión del cuerpo polar I en ovocitos de individuos triploides (Guo *et al.*, 1996; Eudeline *et al.*, 2000; Maldonado *et al.*, 2003; McCombie *et al.*, 2005). Sin embargo, la producción de triploides naturales es un reto para aquellas especies de bivalvos que presentan alto grado de esterilidad asociado a la condición triploide.

Uno de los métodos extensivamente empleados para la producción de triploides naturales se basa en la consecuente producción de tetraploides a partir de triploides para ser usados como progenitores en cruces de organismos tetraploides con diploides que resultan un alto porcentaje de prole en condición triploide (Guo *et al.*, 1996; McCombie *et al.*, 2005). Este procedimiento se ha documentado como posible en la almeja catarina *Argopecten ventricosus* (Maldonado *et al.*, 2003), ya que cierta proporción de la población de triploides es capaz de generar gametos. Sin embargo, la esterilidad asociada a la triploidía de la almeja mano de león es prácticamente total (Maldonado-Amparo *et al.*, 2004; Ramírez-Arce, 2009), lo que imposibilita la generación de tetraploides hasta el momento. Estudios realizados en almeja catarina han sugerido que la esterilidad de triploides, cuando ocurre, resulta de un arresto meiótico por mecanismos génicos de revisión de la fidelidad en meiosis, o *checkpoints* (Maldonado-Amparo e Ibarra 2002a y b; Maldonado-Amparo *et al.*, 2004). En el caso de la almeja mano de león, se ha observado que la esterilidad derivada en condición triploide ocurre en ambas regiones sexuales de la gónada, por lo que ésta especie resulta entonces un modelo interesante para estudiar los mecanismos genéticos que pudieran estar determinando dicho arresto, ofreciendo la posibilidad de explorar a futuro otras alternativas para la inducción a la esterilidad además de la triploidía, como sería la inhibición de genes específicos y requeridos para completar la meiosis.

## 2 ANTECEDENTES

### 2.1 Biología reproductiva

La reproducción de *N. subnodosus* es discontinua, típicamente anual, con un periodo de maduración gamética durante la temporada cálida (Racotta *et al.*, 2003; Arellano-Martínez *et al.*, 2004, 2011; Maldonado-Amparo *et al.*, 2004; Ramírez-Arce, 2009) y con una talla de primera madurez a los 5 cm de longitud de concha. A partir de esta talla, ambas regiones sexuales de la gónada se desarrollan sincrónicamente, y el desove puede presentar un corto período de retraso entre cada tipo de gameto, generalmente iniciando con la liberación de esperma, evitando la autofecundación (Arellano-Martínez *et al.*, 2011; Rupp *et al.*, 2011). La gametogénesis ocurre en una estructura funcional de la gónada, llamada acino, de una forma similar a la observada en otras especies de pectínidos, en la que los gametos se producen a partir de células germinales primordiales localizadas en el margen interno del acino. En la parte testicular, las células espermáticas maduran a partir de las células germinales en un proceso radial, y de forma contraria al ovario, en la cual los ovocitos se encuentran arrestados en metafase-I desde el inicio de la vitelogénesis y hasta la fertilización externa, en la región testicular todos los estadios meióticos se notan representados en el acino en maduración (Beninger y LePenec, 2006). Su ciclo reproductivo presenta períodos de formación de gametos, un desove generalmente total, fertilización de ovocitos en la columna de agua, desarrollo larvario planctónico, fijación a sustrato bentónico mediante un 'byssum' temporal y crecimiento libre en fondos marinos hasta su madurez reproductiva (Rupp *et al.*, 2011). En cuanto a la estrategia reproductiva, ésta es variable, ya que es de tipo oportunista cuando los organismos crecen en áreas con alta disponibilidad de alimento continuamente (Racotta *et al.*, 2003), pero de tipo conservativa cuando el alimento en el hábitat está restringido a periodos específicos (Racotta *et al.*, 2008; Ramírez-Arce, 2009). Se ha reconocido que durante la actividad gametogénica existe movilización y metabolismo de compuestos bioquímicos, como son algunas clases de lípidos, ácidos grasos insaturados y glucógeno (Racotta *et al.*, 2003, 2008; Arellano-Martínez *et al.*, 2004; Palacios *et al.*, 2005, 2007).

En forma general, en los moluscos el conocimiento de la modulación de la gametogénesis por mecanismos endógeno de tipo endócrino, o más específicos como la diferenciación y mantenimiento de tejido germinal y la diferenciación de regiones sexuales, es prácticamente inexistente (Cunningham *et al.*, 2006; Saavedra y Bacherre, 2006; Boutet *et al.*, 2008; Tanguy *et al.*, 2008; Fleury *et al.*, 2009). Dentro de lo que se conoce de la regulación endógena de la gametogénesis de bivalvos, se sabe que existe una asociación entre los niveles gonadales de monoaminas (norepinefrina, dopamina, serotonina) y el desarrollo de la gónada, con importantes diferencias en función a la parte sexual de la gónada (López-Sánchez *et al.*, 2009), y también se ha demostrado la ocurrencia de hormonas peptídicas similares a la hormona liberadora de la gonadotropina (GnRH) en extractos de ganglios neuronales, así como el efecto de la GnRH de mamíferos en la proliferación de células germinales (Croll y Wang, 2007; Nakamura *et al.*, 2007; Ketata *et al.*, 2008). Más recientemente se identificaron y reportaron las secuencias de transcritos correspondientes a GnRH en el ostión japonés *Crassostrea gigas* (Bigot *et al.*, 2012) y en la almeja *Patinopecten yessoensis* (Treen *et al.*, 2012). Sin embargo, hoy en día el desarrollo de nuevas herramientas de análisis genómico y molecular, potencializan el estudio de aspectos particulares en este sentido (Piferrer *et al.*, 2009). Por ejemplo, Dheilly *et al.* (2012) identificaron recientemente, a través del uso de microarreglos, marcadores potenciales de diferenciación sexual y genes expresados diferencialmente entre espermatozoides y ovocitos en el ostión *Crassostrea gigas*, una especie protándrica con alternancia de sexos.

## 2.2 Regulación de la meiosis durante la gametogénesis

La mayoría de la información existente sobre los mecanismos citogenéticos y moleculares que ocurren en la meiosis se han generado en modelos de relativa simplicidad en cuanto a su manipulación en laboratorio, como lo son la levadura *Saccharomyces cerevisiae* (Roeder, 1997) y el nematodo *Caenorhabditis elegans* (Dernburg *et al.*, 1998; Bhalla y Dernburg, 2005). Sin embargo, en prácticamente todo eucariota, la meiosis involucra dos divisiones celulares secuenciales (meiosis I y II) sin un período de interfase entre ellas. La primera división nuclear incluye la segregación de cromosomas homólogos ya replicados, y

durante la segunda división, las cromátidas hermanas derivadas del único ciclo de replicación del ADN (interfase) son separadas o segregadas, resultando en la formación de cuatro núcleos haploides, seguido de la citocinesis (división celular) que conforma cuatro células haploides. Al igual que en la mitosis, tanto la meiosis I como la meiosis II se subdividen en cuatro estadios generales: profase, metafase, anafase y telofase. Adicionalmente, la profase de la meiosis I (profase I) se subdivide en 5 períodos exclusivos (leptoteno, cigoteno, paquiteno, diploteno y diacinesis), determinados en función al tipo y avance de interacción entre cromosomas homólogos previo a su separación. Es en la profase I donde ocurren los eventos de recombinación de cromosomas homólogos, durante la formación de una estrecha unión física entre homólogos llamada complejo sinaptonémico (CS). La ocurrencia del complejo sinaptonémico se basa en la formación de puentes proteicos entre los bivalentes (par de homólogos), que permite el intercambio recombinatorio mediante un complejo proceso de ruptura de dobles hebras, intercambio, reunión y reparación de ADN (Wyman *et al.*, 2004; Wilkins y Holliday, 2009) (Fig. 1).

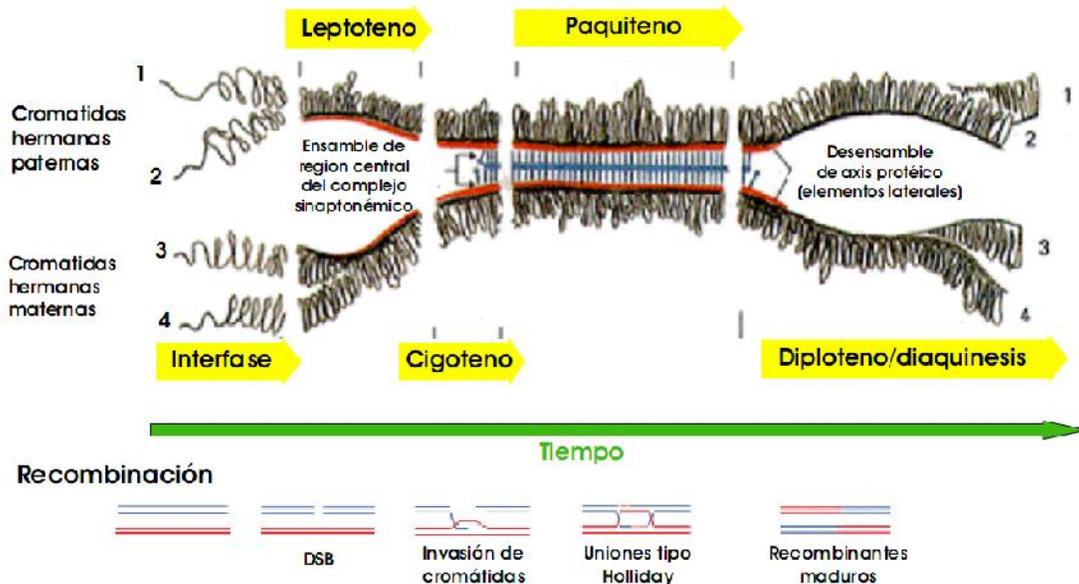


Fig. 1. Eventos de cambios morfológicos propios de la profase I. Durante la fase de paquiteno, los pares de cromosomas homólogos se encuentran completamente cohesionados, y la recombinación ocurre mediante la formación de uniones tipo Holliday. DSB: *Double Strand Breaks*. Figura modificada de Alberts *et al.* (2008).

### 2.3 Efectos de la triploidía sobre la gametogénesis

En pectínidos, como en otras especies de moluscos, la triploidía se asocia a un severo retardo en la reproducción, y a una esterilidad parcial o total. Se propone que la esterilidad asociada a la triploidía propicia una menor movilización de energía hacia la producción de gónada, favoreciendo un balance fisiológico adecuado para el crecimiento del organismo (Tabarini, 1984; Allen y Downing, 1990; Ruiz-Verdugo *et al.*, 2001), aunque los factores ambientales y propios de la biología reproductiva de las diferentes especies no permiten hacer generalidades al respecto (Maldonado-Amparo *et al.*, 2004; Racotta *et al.*, 2008, Ramírez-Arce, 2009). En la almeja mano de león se observa una esterilidad en triploides prácticamente total (Maldonado-Amparo *et al.*, 2004; Ramírez, 2009). Por otro lado, en la almeja catarina (*Argopecten ventricosus*), también hermafrodita funcional, se ha notado que la regionalización de la gónada masculina y femenina se pierde en los triploides (Ruiz-Verdugo *et al.*, 2000) y la esterilidad no es total, produciéndose preferentemente ovocitos sobre espermatozoides inclusive en la región masculina de la gónada de algunos individuos triploides (Maldonado-Amparo y Ibarra, 2002b).

Adicionalmente, los organismos tetraploides de *Argopecten ventricosus* muestran un cambio en la regionalización sexual-gonadal, reduciendo la parte masculina de la gónada más de un 80% (A.M. Ibarra, comunicación personal). Contrariamente, en especies de bivalvos hermafroditas secuenciales, como el ostión japonés, se han documentado tasas de machos/hembras similares entre diploides, triploides y tetraploides, aunque se observa además en triploides un incremento en el número de hermafroditas simultáneos (con gametos tanto femeninos como masculinos) (Allen y Downing, 1990).

Estas observaciones sugieren que en algunas especies de moluscos los mecanismos de determinación sexual temprana o de la línea germinal que establecen el tipo de acino a formar durante la proliferación de células primordiales, se ve afectado por la ploidía. Observaciones relacionadas al efecto de la triploidía no solo en la diferenciación sexual sino en el progreso meiótico no se restringen a moluscos, por ejemplo, en algunas especies de peces se ha documentado una respuesta diferencial en arresto gametogénico entre hembras y machos por consecuencia de la triploidía (Benfey *et al.*, 1989; Hussain *et al.*,

1995; Felip *et al.*, 1999). Lo anterior resulta particularmente interesante, ya que el estudio de los efectos reproductivos de la triploidía (arresto reproductivo diferencial entre regiones masculina y femenina) en pectínidos hermafroditas funcionales permitiría aportar al conocimiento de los mecanismos de determinación sexual durante la diferenciación de la línea germinal, así como a posibles mecanismos diferenciales del control de la fidelidad en la meiosis asociados con la recombinación.

### 2.3.1 Posibles causas de la esterilidad o de la inhibición de la meiosis en triploides.

La causa de la esterilidad observada en triploides puede ser un resultado de procesos que operan a distintos niveles. Desde el punto de vista cromosómico, se ha sugerido que los cromosomas sexuales, cuando existen, pueden determinar una causa para las diferencias observadas entre machos y hembras triploides en cuanto a su capacidad de producir gametos (McDermott y Noor, 2010), similar a lo que se ha propuesto en cuanto a la esterilidad de híbridos (Naisbit *et al.*, 2002; Li *et al.*, 2009). Sin embargo, hasta el momento no se han descrito cromosomas sexuales en bivalvos, aunque sí en gasterópodos (Thiriot-Quévieux, 2003). Adicionalmente, para hermafroditas funcionales como las almejas, la condición de cromosomas sexuales heterogaméticos no es algo esperado, incluso si los cromosomas sexuales están presentes. Una segunda causa de la esterilidad, al menos en plantas, se ha asociado a la desregulación dependiente de metilación de genes o secuencias genómicas como los elementos transponibles (TEs) (Verhoeven *et al.*, 2010). En metazoarios, a pesar de la existencia de mecanismos de metilación de DNA y formación de heterocromatina que previene la propagación y transposición de TEs, se conoce que la desregulación de los TEs ocurre durante un período de la espermatogénesis (Zamudio y Bourc'his, 2010). Un mecanismo adicional se sabe que regula a los TEs: ARN pequeños conocidos como PIWI-interacting RNAs (piRNAs) (Bao y Yan, 2012). Una tercera posibilidad de la esterilidad en los triploides es un arresto de la gametogénesis inducido por *checkpoints* meióticos. La esterilidad de moluscos triploides ha sido asociada a un *checkpoint* en paquiteno después de descubrir que el arresto ocurre en los estadios de transición cigoteno-paquiteno de la profase I en la almeja *A. ventricosus* (Maldonado-Amparo e Ibarra 2002a; 2002b).

En forma general los *checkpoints* meióticos operan como mecanismos moleculares de revisión de la fidelidad en el proceso meiótico mediante la activación de genes que pausan el progreso meiótico durante la replicación de ADN, recombinación, o sinapsis de cromosomas cuando se detectan conflictos en dichos procesos durante la meiosis temprana (Murakami y Nurse, 1999; Roeder y Bailis, 2000; MacQueen y Hochwagen, 2011) para evitar la formación de meiocitos aneuploides. Los mecanismos de *checkpoint* han sido detectados como diferenciales entre sexos, así como entre especies (Sun y Kim, 2012), y algunas especies son más tolerantes a la aneuploidía. Por ejemplo, el ostión *Crassostrea gigas* es sólo parcialmente estéril cuando es triploide, produciendo gametos, aunque el crecimiento de organismos aneuploides puede resultar afectado (Thiriot-Quievreux *et al.*, 1988; Zourous *et al.*, 1996; Leitao *et al.*, 2001).

El *checkpoint* de paquiteno opera durante la interacción de cromosomas homólogos en la profase de meiosis I, y tiene como función evitar la progresión a diploteno-diacinesis (posterior a paquiteno) hasta que los eventos de recombinación de cromátidas, que incluye la reparación de los rompimientos de doble cadena, hayan ocurrido de forma completa (Dernburg *et al.*, 1998; Roeder y Bailis, 2000; Meier y Gartner, 2006). Su activación evita la producción de gametos con errores en la recombinación, como translocaciones no recíprocas con la pérdida de fragmentos cromosómicos o gametos con un número de cromosomas desbalanceados (Roeder y Bailis, 2000). Un *checkpoint* generalmente involucra respuestas de arresto en ciclo celular, modificaciones en las histonas, cambios en el transcriptoma y modificaciones post-traduccionales de proteínas de reparación de ADN, y limita, aunque de forma reversible, la progresión del ciclo celular (Clémenson y Marsolier-Kergoat, 2009). El *checkpoint* de paquiteno involucra la expresión de genes (y la correspondiente síntesis-activación) de proteínas específicas de meiosis (Bailis *et al.*, 2000). Se han descrito al menos dos mecanismos particulares de *checkpoint* en paquiteno, uno de ellos dirigido a monitorear intermediarios de recombinación meiótica incompleta, y otro dedicado a revisar la adecuada sinapsis entre homólogos (Dernburg *et al.*, 1998; San-Segundo y Roeder, 1999, 2000; Bailis *et al.*, 2000; Roeder y Bailis, 2000; Bhalla y Dernburg, 2005). En este sentido, y puesto que en la meiosis de triploides ocurre un conflicto meiótico por interacción desbalanceada de cromátidas homólogas durante la

sinapsis y recombinación con la formación de trivalentes (Ruiz-Verdugo *et al.*, 2000; Cunado *et al.*, 2002) (Fig. 2), se hipotetiza que esta interacción irregular activa algún mecanismo de arresto, probablemente mediado por el *checkpoint* en paquiteno (Maldonado-Amparo e Ibarra, 2002a y 2002b; Li *et al.* 2009). La activación de este *checkpoint* evitaría la segregación de trivalentes, la cual resultaría en la generación de intermedios meióticos aneuploides, o con cargas cromosómicas desbalanceadas (Roeder 1997).

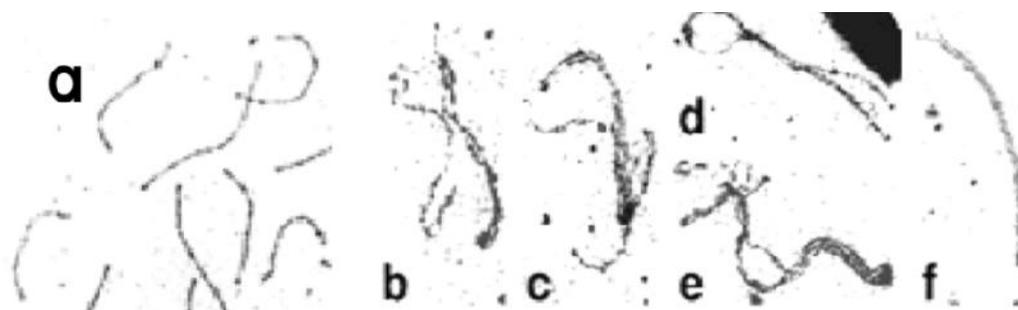


Fig. 2. Morfología de complejo sinaptonémico (SC) en diploides y triploides. En gametos de diploides, se observan apareamientos de cromosomas perfectos (a); en triploides se observa un gradiente de morfologías que va desde trivalentes incompletos de homólogos (b, c, d y e) así como trivalentes perfectos (f) (Cunado *et al.*, 2001; Cunado *et al.*, 2002).

#### 2.4 Conocimiento genómico en el grupo de los moluscos

El grupo de los moluscos ha sido considerado como limitado en cuanto a la información genómica disponible (Song *et al.*, 2006; Boutet *et al.*, 2008). Con el reciente desarrollo de tecnologías de secuenciación masiva, esta tendencia ha sido contrarrestada, y actualmente se cuenta con el genoma de un ostreido, el ostión japonés (*Crassostrea gigas*) (Zhang *et al.*, 2012), y adicionalmente, se ha generado un genoma parcial de la lapa de California *Lottia gigantea* (The Joint Genome Institute - Department of Energy of US), así como de la ostra perlera *Pinctada fucata* (Takeuchi *et al.*, 2012).

El análisis de expresión génica para el entendimiento de los mecanismos genético-moleculares asociados a la gametogénesis, basado en el descubrimiento de genes con

niveles de expresión génica diferencial en un contexto comparativo, p. ej., entre diferentes estadios de desarrollo gametogénico, y su posterior análisis confirmatorio por metodologías de análisis cuantitativos, es una estrategia para ser aplicada en organismos no modelo, incluyendo a *Nodipecten subnodosus*, para los cuales la información genómica de secuencias de genes reportadas en bases de datos es limitada (Saavedra y Bachere, 2006, Fleury *et al.*, 2009). Esta estrategia involucra la generación de genotecas por hibridación sustractiva (SSH) (Diatchenko *et al.*, 1996) para el descubrimiento de genes nuevos cuya expresión resulta modulada diferencialmente en un tejido/órgano entre condiciones de desarrollo, y es considerada como la más adecuada cuando el objetivo es que los genes descubiertos tengan una relevancia en el contexto biológico/fisiológico específico, ya que maximiza la identificación de genes diferencialmente expresados entre condiciones de desarrollo contrastantes, además de que provee con un valioso cuerpo de información transcriptómica primaria para otros trabajos (Saavedra y Bachere, 2006). En cuanto a genotecas SSH testículo-específicas, éstas se han desarrollado para un número de organismos y se han basado en la técnica de secuenciación capilar (Tureci *et al.*, 1998; Wang y Chapin, 2000; Liang *et al.*, 2004; Pinto *et al.*, 2006; Chiba *et al.*, 2007; Leelatanawit *et al.*, 2008; Redon *et al.*, 2010). Adicionalmente, el uso de plataformas de secuenciación masiva (454-Roche) permite maximizar la capacidad de descubrimiento de genes, aunque su uso en combinación con genotecas SSH es incipiente; únicamente se conocen tres trabajos que incorporan esta sofisticación en cuanto a la plataforma de secuenciación: uno que incorpora 454-FLX para secuenciar genotecas de una especie de dinoflagelado (Toulza *et al.*, 2010) y dos trabajos aplicados en planta de soya (Rodrigues *et al.*, 2012; Barros de Carvalho, 2013).

### 3 JUSTIFICACIÓN

La investigación en bivalvos pectínidos ha enfocado esfuerzos en el desarrollo de métodos controlados para la producción continua y suficiente de larvas (Arellano-Martínez *et al.*, 2011a, 2011b; Mazón-Suastegui *et al.*, 2011; Freitas-Valbuena *et al.*, 2011), posibilitando así la aplicación de otras tecnologías encaminadas al control de diversidad y manipulación genética, así como la producción de poliploides (Petersen *et al.*, 2007, 2008; Piferrer *et al.*, 2009; Ibarra *et al.*, 2011). Sin embargo, uno de los aspectos fundamentales para el dominio de la tecnología de producción de larvas para acuicultura para el cual se conoce poco es el entendimiento de los mecanismos genéticos involucrados en la regulación de la gametogénesis.

Para lograr conocer los mecanismos de arresto meiótico en la gametogénesis de almejas triploides, es necesario generar en un principio la información genómica que, a través de su reconocimiento funcional en condiciones de diploides, permita posteriormente avanzar hacia conocer su papel en la esterilidad de triploides. Para otras especies de moluscos, las genotecas de ADNc a partir de tejido gonadal generadas a la fecha han resultado limitadas. Por ejemplo, para la almeja *Argopecten purpuratus*, Boutet *et al.* (2008) desarrollaron genotecas de cDNA de tipo *full-length* sin normalización, y obtuvieron, a través de secuenciación capilar (Sanger), 418 secuencias únicas de transcritos de genes expresados en la gónada de individuos en condición de gametogénesis inactiva y en madurez. La genoteca se desarrolló en pool, lo que impide reconocer si existe una relación de la expresión de los genes identificados con la condición de gametogénesis. Además, para únicamente 80 genes se logró una anotación parcial por análisis de alineamiento local tipo blastx contra bases de datos de proteínas (NCBI-Genbank, nr database). Por otra parte, Ciocan *et al.* (2010) secuenciaron genotecas de cDNA sustractivas (SSH) de tejido gonadal de mejillón (*Mytilus edulis*) en maduración temprana vs tardía, así como en organismos expuestos a estradiol vs organismos no expuestos; los autores reportaron 47 secuencias con similitud significativa a alguna proteína de bases de datos pública. Más recientemente, Dheilly *et al.* (2012) reportaron un análisis transcriptómico basado en microarreglos específicos de la especie de ostión *Crassostrea gigas*, donde identifican 372 y 731

transcritos con expresión diferencialmente modulada durante el transcurso de la espermatogénesis y la ovogénesis, respectivamente, en contraste a otros tejidos somáticos. Lo anterior sugiere que en moluscos bivalvos, los intentos de identificar genes específicos de la gónada han sido limitados, y que existe además la necesidad y oportunidad de incrementar el conocimiento genómico de la gametogénesis en especies de moluscos.

En un contexto transcriptómico, la parte testicular de la gónada resulta ideal para el descubrimiento de genes involucrados en el proceso completo de la gametogénesis ya que en ésta ocurre una meiosis continua y completa, donde se espera que se expresen genes como aquellos que participan en la mitosis de las células germinales, meiosis y recombinación, así como genes del período de post-meiosis y diferenciación. Adicionalmente, genes de determinación/diferenciación sexual (Small *et al.*, 2005; Xia *et al.*, 2007) y elementos transponibles (Zamudio y Burchis, 2010) serán enriquecidos o expresados específicamente durante la diferenciación y desarrollo testicular. Es importante mencionar que a pesar del condensamiento cromosómico que ocurre durante el proceso meiótico de la espermatogénesis, es conocido que la actividad transcripcional es activa y la expresión de los genes puede regularse a nivel transcripcional y post-transcripcional por múltiples mecanismos (Gan *et al.*, 2013). El descubrimiento de genes reproductivos en una especie de pectínido como *N. subnodosus* ofrecerá la oportunidad de desarrollar futuros estudios sobre la genética funcional de la condición de hermafroditismo, desde el mantenimiento de la diferenciación sexual, así como de la diferenciación gamética durante la meiosis.

Una vez identificados o conocidos los genes de interés, el análisis de la expresión génica específica permite la identificación de mecanismos funcionales de procesos celulares, como los que operan durante la espermatogénesis. Como un primer paso para la cuantificación de la expresión génica, es necesario identificar genes expresados en forma estable durante la diferenciación y desarrollo gonadal, particularmente en casos como el de *N. subnodosus*, una especie hermafrodita funcional en la cual ocurren cambios morfológicos y bioquímicos diferentes entre las regiones sexuales de su ovotestis (ovario y testículo). Esto es, el desarrollo gonádico modulará la actividad transcripcional, incluyendo

genes potenciales de referencia, en una forma sesgada hacia el sexo desarrollándose (Lu y Wu, 2011). Esto ha sido encontrado no solamente comparando machos y hembras en especies dioicas, sino también en el caso de plantas y animales hermafroditas funcionales (Jiang *et al.*, 2001; Silveira *et al.*, 2009).

La metodología más robusta utilizada para el análisis de expresión génica es la conocida como PCR en tiempo real y su detección por fluorescencia (o qPCR), en la cual la correcta cuantificación relativa de la expresión de genes de interés depende de múltiples factores metodológicos durante el proceso mismo de generar el ADNc, así como de contar con genes de referencia internos (generalmente genes constitutivos) cuya expresión sea constante entre diferentes condiciones experimentales o entre diferentes tejidos para la normalización del nivel de transcritos del gen de interés (Thellin *et al.*, 1999; Bustin, 2002; Bustin y Nolan, 2004; Bustin *et al.*, 2009). Los genes de tipo constitutivo son aquellos requeridos para el mantenimiento de una función celular básica y de la homeostasis celular, y aunque algunos de estos genes son expresados a niveles relativamente constantes, otros pueden variar su nivel de expresión dependiendo de las condiciones experimentales a las cuales los organismos son sometidos. Debido a esta variación dependiente de las condiciones experimentales, los genes potenciales de referencia deben de ser evaluados para su estabilidad antes de ser utilizados en la normalización de la expresión de genes de interés. Para éste propósito, diversos modelos han sido desarrollados, como los algoritmos de geNorm (Vandesompele *et al.*, 2002), NormFinder (Andersen *et al.*, 2004), y BestKeeper (Pfaffl *et al.*, 2004). Estos algoritmos permiten categorizar a los genes con base en la variación de su expresión, sin embargo lo hacen bajo distintas estrategias, lo cual puede llevar a distintas conclusiones en cuanto a la estabilidad de los genes potenciales de referencia siendo evaluados, por lo que es deseable utilizar varios algoritmos buscando encontrar resultados convergentes, así como también es importante comprender las causas de resultados no convergentes dados los algoritmos siendo utilizados.

## 4 OBJETIVO GENERAL

Generar información genómica asociada a la espermatogénesis, meiosis y mantenimiento de línea germinal en *Nodipecten subnodosus*, validando la expresión de genes de meiosis durante la gametogénesis en organismos diploides con la finalidad de contar con una base molecular para el futuro estudio del arresto meiotico o esterilidad en condición triploide.

### 4.1 Objetivos particulares

- Aislar y anotar un alto número de transcritos expresados a través de la producción de genotecas sustractivas (SSH) y secuenciación masiva (pirosecuenciación) que permita la identificación de genes específicos y/o asociados a la meiosis durante la maduración testicular y el mantenimiento de la línea germinal en reposo reproductivo en diploides de *Nodipecten subnodosus*.
- Establecer la mejor estrategia para el análisis de expresión génica por PCR cuantitativo, definiendo el impacto sobre la normalización de la expresión génica al utilizar genes de referencia "al azar" *versus* genes validados como estables para condiciones específicas.
- Analizar la expresión génica en ovario y testículo de genes identificados con fines de conocer sus patrones de expresión durante la meiosis.
- Caracterizar la secuencia completa de transcritos de genes selectos relacionados al proceso meiótico, y establecer su relación filogenética con sus homólogos/ortólogos en otras especies.

## 5 HIPÓTESIS

1. Dado que la recombinación existe en la región testicular de la almeja *Nodipecten subnodosus*, entonces múltiples genes involucrados en el proceso de meiosis pueden ser aislados a través de secuenciación masiva de genotecas de hibridación sustractiva (SSH) obtenidas en la región testicular en espermatogénesis activa.
2. Los transcritos de genes aislados en la región testicular meióticamente activa e inactiva de *N. subnodosus* podrán ser identificados por su función en la meiosis y/o mantenimiento de la línea germinal tanto por homología con sus ortólogos en otras especies como por su función ontológica inferida.
3. La expresión de genes aislados y anotados como participantes en la meiosis de *N. subnodosus* será específica para el proceso de gametogénesis tanto en la región ovárica como en la testicular.

## 6 MATERIALES Y MÉTODOS

### 6.1 Material biológico

#### 6.1.1 Utilizado en la generación de genotecas sustractivas y caracterización de transcritos de genes de interés

Como material biológico para la generación de genotecas sustractivas, así como para la amplificación de transcritos de interés por PCR degenerativa y RACE-PCR, se utilizó una población experimental de *N. subnodosus* producida con reproductores de Bahía de Loreto, Baja California Sur, y mantenida en el mismo sitio. Ésta población fue generada en el Laboratorio de Genética y Mejoramiento Animal Acuícola (LGyMAA), a partir de la fertilización de ovocitos de una hembra con una mezcla de espermatozoides de tres machos, crecidas hasta su fase de semilla en el laboratorio, y llevadas a campo en bolsas de malla fina para su posterior liberación a canastas ostrícolas en suspensión. Las almejas fueron colectadas en el mes de Julio de 2009, bajo una estrategia selectiva: un grupo de organismos poseían gónadas incipientes, sin coloración en la región femenina, característica de la condición de inactividad gametogénica, y otro grupo fue seleccionado por poseer gónadas de color naranja intenso propio de una condición de avance gametogénico. Todos los organismos se encontraban en un intervalo de 5 a 6 cm de longitud total. Los organismos seleccionados fueron transportados vivos al LGyMAA para un período de depuración en agua de mar filtrada a 21° C por dos días. Posteriormente, los animales fueron sacrificados, y una porción media de gónada incluyendo ambas regiones gonadales fue fijada en solución de Davidson (4% Formaldehído, 10% etanol, en agua de mar filtrada) para su posterior procesamiento histológico que incluyó la inclusión en parafina, corte a 5 micras y tinción con hematoxilina-eosina. Adicionalmente, se disectaron porciones de gónada de la región femenina y masculina de los organismos con gónadas en gametogénesis activa, como se muestra en la figura 3. Para las gónadas inactivas, se disectaron fragmentos de las mismas regiones evitando el tubo intestinal; las regiones son evidenciadas por la transparencia del tejido gonadal al poseer acinos vacíos. Los tejidos

destinados para aislamiento de ARN total fueron preservados en solución fijadora para ácidos nucleicos (sulfato de amonio a una proporción de 70 g en 100 ml de citrato de sodio 25mM, EDTA 10 mM; pH 5.2) en una proporción de 1:10 (tejido:solución), mantenidos a 4° C durante la noche, y almacenados indefinidamente a -20° C.

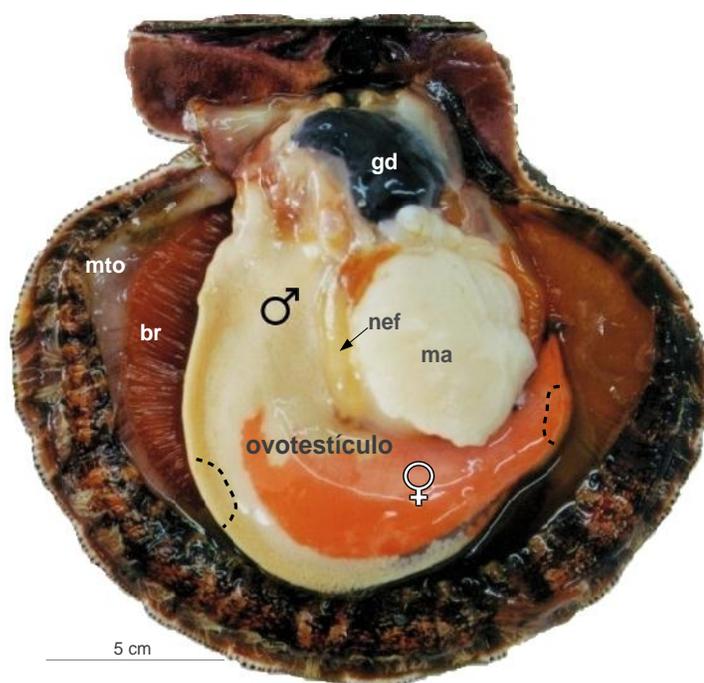


Fig. 3. Anatomía de la almeja mano de león *N. subnodosus* en gametogénesis activa, donde se muestra las porciones ovárica (♀) y testicular (♂) de la gónada, que fueron disectadas para la generación de genotecas de hibridación sustractiva. La región delimitada por líneas paralelas muestra la sección de la gónada tomada para la inspección microscópica de la condición gametogénica por observación histológica. Otros órganos internos: músculo aductor (ma), manto (mto), branquias (br), nefridios (nef), glándula digestiva (dg).

#### 6.1.2 Utilizado en el análisis de validación de genes de referencia y de expresión por qPCR.

Para los análisis de validación de genes de referencia y de expresión génica en distintos estadios reproductivos, se colectaron organismos a partir de una población experimental producida con reproductores de laguna Ojo de Liebre en el LGyMAA y mantenida en la laguna costera de Rancho Bueno, Bahía Magdalena, situada en la costa del Pacífico de la

península de Baja California Sur. Los organismos fueron colectados en Mayo de 2010, transportados a las instalaciones del LGyMAA, mantenidos durante 48 h en agua de mar a 21 °C sin alimentación y posteriormente sacrificados para obtener tejidos de ambas regiones gonadales para histología y para aislamiento de ARN total, como se menciona en la sección 4.1.1. Adicionalmente y con el fin de establecer la ubicuidad de uno de los genes originalmente seleccionados como posible gen de referencia muestras de tejidos de musculo aductor, glándula digestiva, manto, branquias y nefridios fueron preservados en solución fijadora para ácidos nucleicos. Después de la examinación histológica, y basado en la escala de desarrollo gametogénico descrito por Ruiz-Verdugo *et al.* (2001) (Fig. 4) se seleccionaron aquellas muestras representativas de estadios inactivos, y de desarrollo temprano y tardío de la parte ovárica y su correspondiente porción testicular.

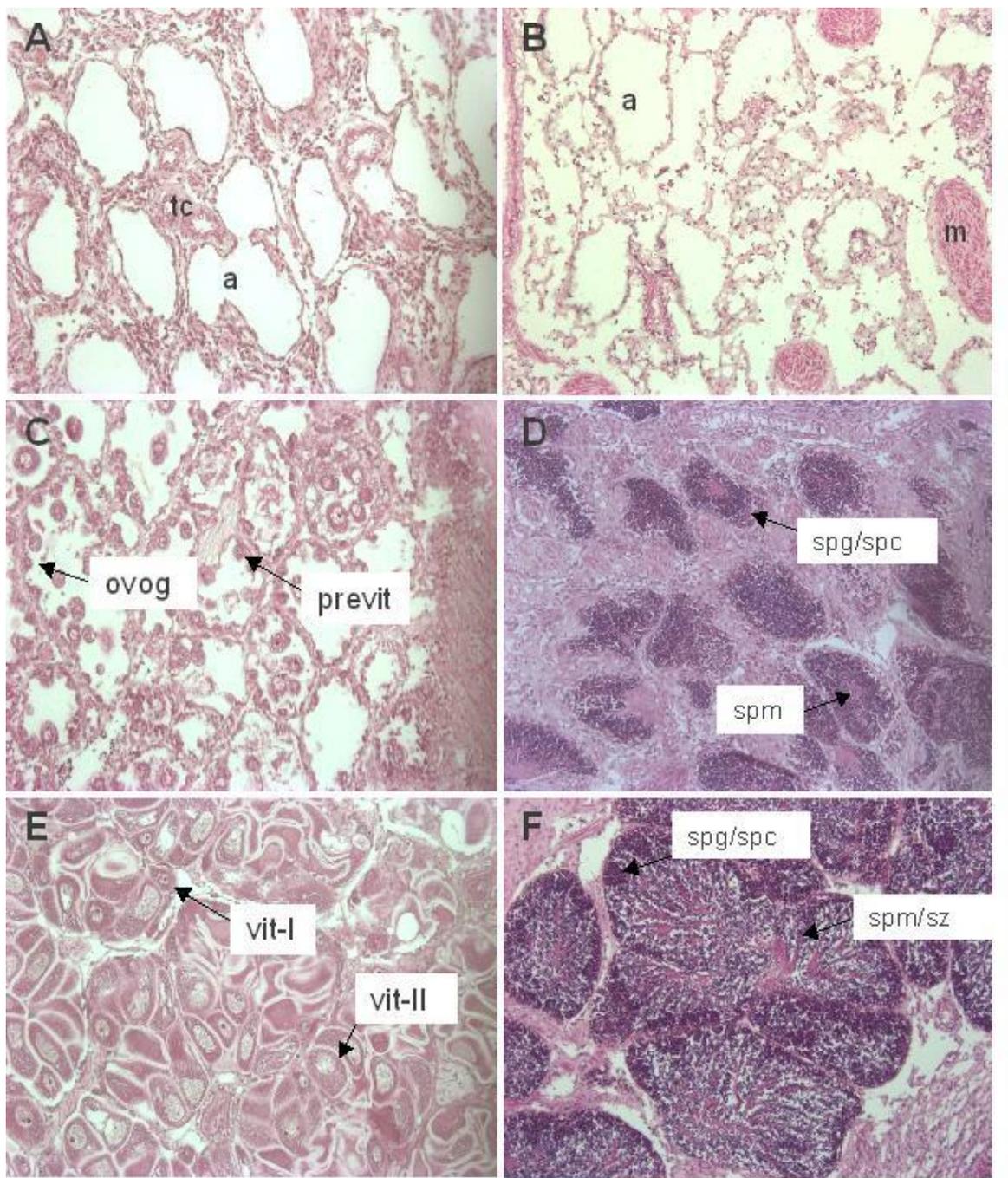


Fig. 4. Micrografías de los estadios reproductivos evaluados en éste trabajo: estadio inactivo (indiferenciado) de pre-gametogénesis (A) y post-desove (B); maduración temprana del ovario (C) y de la región testicular (D); maduración tardía del ovario (E) y de la porción testicular (F). Los tejidos o tipos celulares son denotados como a continuación: tejido conectivo (ct), acinos vacíos (a), fibras musculares (m), ovogonias (ovog), ovocitos previtelogénicos (previt), ovocitos previtelogénicos primarios (vit-I), ovocitos previtelogénicos secundarios (vit-II), espermatogonias (spg), espermatocitos (spc), espermátidas (spm), espermatozoos (spz).

## 6.2 Extracción de ARN

El ARN total fue aislado de muestras individuales mediante el procedimiento de extracción con solución trifásica basada en fenol-guanidina (Scarpa *et al.*, 1994), empleando un reactivo comercial (TriPure, Roche), con menores modificaciones como se describe a continuación. Fragmentos de tejido de 50 mg fueron retiradas de la solución fijadora de ácidos nucleicos retirando el exceso con papel secante, e inmediatamente sumergidas en 1 ml de TriPure, y homogenizados con perlas de vidrio de 435-600  $\mu\text{m}$  (SIGMA) en un equipo FastPrep (MP Biomedicals) empleando dos o tres pulsos de 30 s a 6.0 m/s, con intervalos de 5 minutos de reposo en hielo, hasta obtener homogenizados de tejido totalmente disgregados. Posteriormente, los tubos fueron centrifugados a 12,000 g por 5 min, el sobrenadante recuperado, y se agregó 200  $\mu\text{l}$  de cloroformo, y agitados manualmente 15 s. Los tubos se dejaron reposar por 15 minutos, para posteriormente centrifugar por 15 min a 12,000 g. Se recuperó la fase superior (acuosa) en tubos nuevos, y se agregaron 500  $\mu\text{l}$  de isopropanol dejando reposar por 30 minutos a  $-20^{\circ}\text{C}$ , luego centrifugados por 10 min a 12,000 g y fue descartando el sobrenadante. El precipitado obtenido fue lavado agregando 800  $\mu\text{l}$  de etanol 75%, para después centrifugar a 7,600 g y retirar el sobrenadante. El ARN total fue resuspendido en 100  $\mu\text{l}$  de agua, para luego agregar 10  $\mu\text{l}$  de LiCl 4M y dejar reposando por 2 hr a  $-80^{\circ}\text{C}$ . Posteriormente, los tubos fueron centrifugados a 12,000 g por 20 min, lavados de nuevo con etanol (75%) y centrifugados para retirar éste, y posteriormente secados al aire por 10 min, para posteriormente resuspender el pellet en 50  $\mu\text{l}$  de agua. La concentración del ARN fue estimada por espectrofotometría en UV utilizando un equipo NanoDrop 1000 (Thermo).

Para el ARN destinado para la generación de genotecas de hibridación sustractiva y para síntesis de ADNc para amplificación de transcritos completos por RACE, se siguió el mismo procedimiento de aislamiento de ARN total, pero empleando muestras en pool de gónada en cada condición reproductiva: inactividad (post-desove), madurez tardía de región testicular, y madurez tardía de región ovárica (Fig. 4). Asimismo, los tejidos fueron homogenizados utilizando un homogeneizador mecánico (Polytron, Kinematica) en tubos falcon de 15 ml de volumen y empleando 10 ml de reactivo TriPure. La fracción de ARN

mensajero (ARNm) fue purificada empleando el sistema de purificación con oligo-dT-celulosa Poly-A-Pure (Ambion) a partir de 1-2 mg de ARN total recuperado por muestra previamente resuspendido en 4 ml de NaCl. La solución de ARN total de cada muestra se incubó a 55 °C en presencia de un cartucho de oligo-dT-celulosa, para luego realizar lavados en el conjugado ARNm/oligo-dT-celulosa y así eliminar la fracción no poliadenilada del ARN total. El ARN mensajero (fracción poliadenilada) fue eluido del oligo-dT-celulosa con 100 µl de agua grado biología molecular, y el rendimiento de ARNm fue determinado por espectrofotometría UV, empleando el instrumento NanoDrop 1000 (Thermo).

### 6.3 Generación de genotecas por hibridación sustractiva

Se generaron genotecas recíprocas de ADNc de tipo SSH (*suppression subtractive hybridization*) (Diatchenko *et al.*, 1996), utilizando el kit comercial *PCR-Select® cDNA subtraction* (Clontech), de acuerdo al protocolo proporcionado, y utilizando 2 µg de ARNm por muestra (gónada en maduración de la región testicular; gónada inactiva). A partir del ARNm, se sintetizó ADNc de doble cadena de cada grupo (ADNc-dc) empleando la enzima transcriptasa reversa *SMARTScribe* (Clontech), y se realizó una digestión completa con la enzima de restricción *RsaI* para la posterior ligación de los adaptadores en ambos sentidos del ADNc-dc. Puesto que la eficiente ligación de adaptadores es un paso crucial para el correcto desarrollo de genotecas, éste paso fue corroborado mediante reacciones de PCR con oligonucleótidos dirigidos a cada adaptador y oligonucleótidos específicos para una reacción control, con las instrucciones proveídas en el kit. El ADNc-dc ligado a adaptadores (denominado por el protocolo como *tester*) fue doblemente hibridado de forma recíproca con ADNc-dc de una condición biológica contrastante (*driver*) para luego ser amplificado por reacciones de PCR anidadas; este paso es llamado PCR sustractiva. Por ejemplo, para obtener únicamente fragmentos de ADNc cuya abundancia es relativamente mayor en la región testicular en gametogénesis activa en comparación a la gónada inactiva, el ADNc-dc *tester* corresponde al obtenido a partir de la gónada madura y el *driver* corresponde al obtenido de la gónada inactiva, y la PCR sustractiva resultante generará únicamente fragmentos de ADNc de transcritos que son en

principio únicamente expresados o más abundantes en la gónada testicular (Fig. 5). Un diseño recíproco (usando ADNc-dc de gónada inactiva como *tester* y de gónada testicular en gametogénesis como *driver*) permitiría entonces obtener fragmentos de ADNc correspondientes a transcritos más abundantes en gónada inactiva.

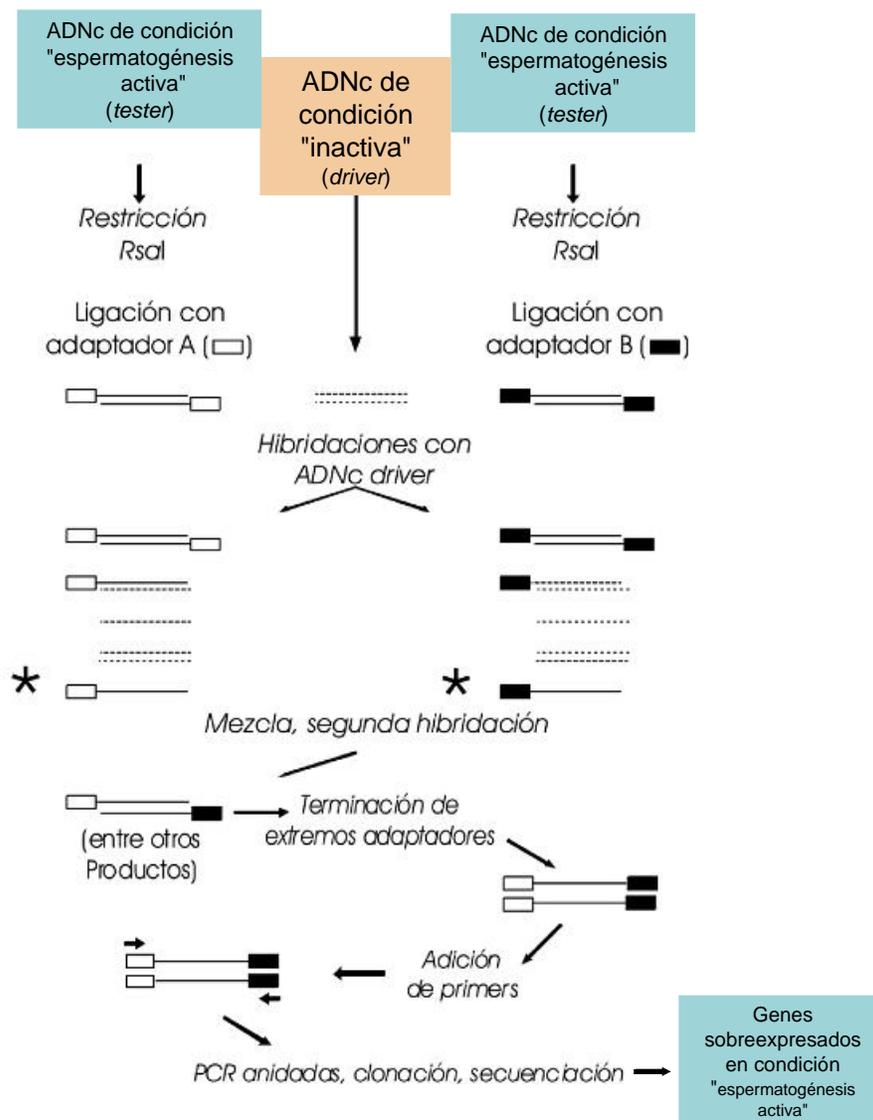


Fig. 5. Esquema de desarrollo de genotecas de hibridación sustractiva. El ADNc de ambas condiciones es sometido a una digestión con endonucleasa *RsaI*, y el ADNc de las gónadas en espermatogénesis activa (*tester*) es ligado con adaptadores A, mientras que el ADNc de la gónada inactiva (*driver*) es ligado a adaptadores B. Ambos ADNc son sometidos a doble hibridación, y los fragmentos (denotados con \*) que no hibridan con ningún fragmento de *driver* por ser específicos de la condición *tester* son amplificados exponencialmente empleando oligonucleótidos dirigidos a los adaptadores A y B para ser secuenciados individualmente. Las secuencias obtenidas representan entonces a los genes específicos de la condición *tester* (espermatogénesis avanzada)

### 6.3.1 Secuenciación de genotecas obtenidas por hibridación sustractiva

Las genotecas obtenidas por hibridación sustractiva fueron secuenciadas mediante dos estrategias: secuenciación por electroforesis capilar (método Sanger) y mediante secuenciación masiva en plataforma 454-FLX de Roche (método 454-FLX), como se describe a continuación.

#### 6.3.1.1 Secuenciación capilar - método Sanger

Para la secuenciación por electroforesis capilar - método Sanger - los fragmentos producidos por PCR sustractiva anidada fueron ligados a vectores de clonación pCR-4.0 (Invitrogen) para su clonación en bacterias competentes DH-5 por transformación por choque térmico (Sambrook y Russell, 2001). El producto de la transformación fue crecido en placas Petri con agar LB (agar - bacto triptona - extracto de levadura; ampicilina 50 µg/ml como medio de selección) e incubadas a 37 °C por ~18 h. Posteriormente, las colonias identificadas como transformantes fueron levantadas con un palillo estéril, y depositadas en placas de 96 pozos con una capa de agar LB-ampicilina para ser enviadas a una compañía de secuenciación (Genewiz, NJ), donde a partir de colonias bacterianas se realizó una amplificación por PCR de ADN plasmídico para su secuenciación por química BigDye Terminator (Applied Biosystems). Antes de enviar a secuenciar las genotecas, se realizó una confirmación de que las clonas seleccionadas aleatoriamente presentaban insertos en el rango de tamaño esperado; de 200 a 1500 pb, como señala el manual de producción de genotecas *PCR-Select® cDNA subtraction kit*, mediante la estrategia de *colony-PCR*. Para la *colony-PCR*, una pequeña muestra de cada colonia seleccionada al azar fue tomada con un palillo, fue resuspendida en 10 µl de agua, hervida a 100° C por 10 min, y 1 µl de la suspensión obtenida fue posteriormente usada como molde para PCR, amplificando con oligonucleótidos universales M13 (sentido: 5'-GTT TTC CCA GTC ACG AC-3'; antisentido: 5'-CAG GAA ACA GCT ATG AC-3'), los cuales hibridan en las regiones flanqueantes a la región de inserción del vector de clonación. Mediante electroforesis en agarosa, se comprobó los tamaños de los insertos contenidos en las clonas ocurrían en un rango de tamaño esperado.

### 6.3.1.2 Secuenciación masiva - método 454

Con el objetivo de incrementar el número de secuencias de los transcritos obtenidos con las genotecas de hibridación sustractiva, se evaluó la secuenciación masiva de las genotecas de hibridación sustractiva generadas. Una de las tecnologías de secuenciación de última generación más populares es la ofrecida por la compañía Roche-Life, que mediante el instrumento FLX genome sequencer, permite la secuenciación de miles de fragmentos de ~400 pb simultáneamente a partir de ADN genómico o de genotecas de ADNc *full-length*.

Ésta tecnología se logró adaptar exitosamente para la secuenciación de varias genotecas sustractivas en una sola celda, mediante modificaciones de las técnicas de secuenciación orientadas a amplicones de PCR (*GS FLX Amplicon DNA Library Preparation Method Manual*; Roche, 2007). Para esto, se sustituyeron los oligonucleótidos anidados del sistema Clontech-PCRselect(R) por un set de oligonucleótidos (454-nested 1; 454-nested 2R) que incluían la secuencia de los adaptadores Lib-L/A y Lib-L/B en el extremo 5' y que son requeridos para la PCR en emulsión (emPCR) realizada durante el proceso de secuenciación; un identificador MID (*Multiplex Identifier*) de 10 bases nucleotídicas único para la librería específica en el oligonucleótido 454-nested 1, y las secuencias de los oligonucleótidos anidados del sistema Clontech en el extremo 3' (Fig. 6).

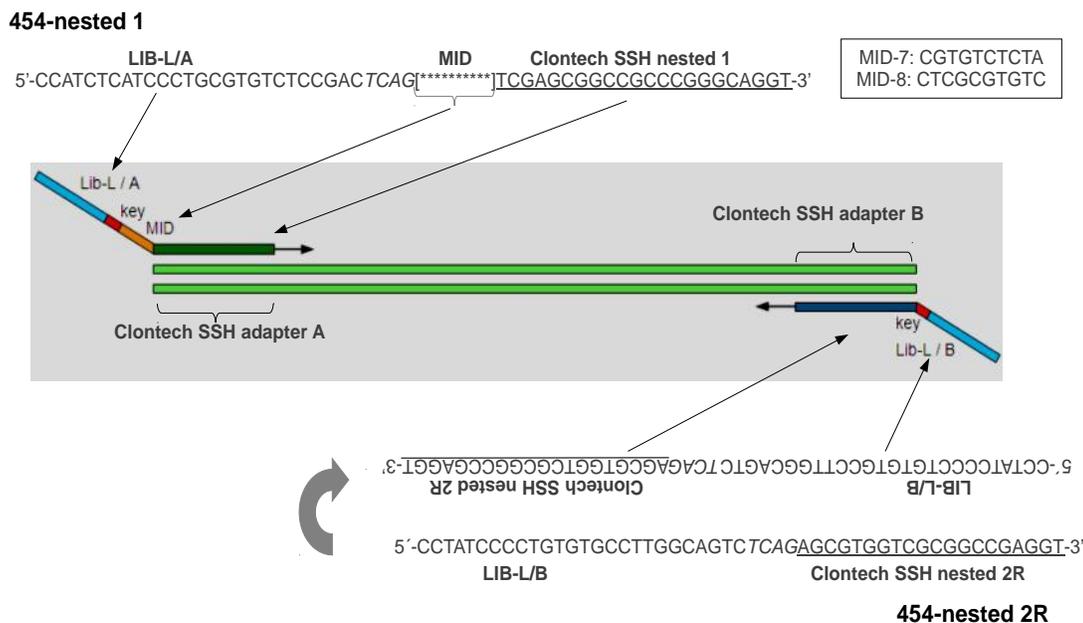


Fig. 6. Diseño de oligonucleótidos 454-nested 1 y 454-nested 2R, empleados para la adición de adaptadores LIB-L en los extremos de los fragmentos de PCR mediante la reacción anidada. Se muestran las secuencias correspondientes a los oligonucleótidos anidados (*nested*) 1 y 2R, originales del sistema PCR-select SSH (Clontech), así como las secuencias de los adaptadores LIB-L A y B requeridos para la secuenciación por el sistema Roche-454 FLX. La secuencia identificadora de cada genoteca se marca como la secuencia MID en el oligonucleótido 454-nested 1.

La amplificación de la genoteca se realizó bajo el protocolo de amplificación anidada del protocolo original de librerías SSH de Clontech. De ésta forma, fue posible utilizar éste tipo de tecnología para la secuenciación de distintas genotecas de forma simultánea. Las genotecas entonces obtenidas para genes específicos de gónada masculina (M) y de indiferenciada-sustraídas con gónada masculina ( $I_M$ ) fueron preparadas desde la PCR anidada usando los oligonucleótidos modificados, además de marcadas diferencialmente con secuencias MID únicas mediante el programa *sffinfo* y *sfffile* incluido en el software Newbler 2.5.3 (Roche)

Las genotecas fueron enviadas y procesadas en el laboratorio de secuenciación GenoSeq-UCLA, aplicando un esfuerzo de secuenciación exploratorio (1/16 del total de muestras incluidas en la corrida de secuenciación estándar). Las secuencias obtenidas fueron entregadas en archivos de secuenciación 454 formato \*.sff.

### 6.3.2 Análisis bioinformáticos de secuencias obtenidos en genotecas

Las secuencias de cada genoteca (testis e inactiva) obtenidas por 454-FLX en formato \*.sff fueron separadas con base en las secuencias MID de los adaptadores LIB-L/A y LIB-L/B (ver sección anterior), y posteriormente convertidas a archivos \*.fastq con *score* de calidad Phred en escala Sanger (Cock *et al.*, 2010). Se utilizaron los programas *sfffile* y *sffinfo* (sff tools, Roche) para su procesamiento y limpieza de adaptadores con las herramientas de los programas Newbler 2.5 y Trimmomatic (Loshe *et al.*, 2012).

Los archivos de secuencias obtenidas por secuenciación capilar tipo Sanger fueron transformadas del formato \*.ab1 a archivos de secuencia en fasta (\*.seq y \*.qual) usando el programa Phred (Ewing *et al.*, 1998) y procesadas para la remoción de vectores y adaptadores con el software SnoWhite 1.1.4 (Barker *et al.*, 2010) implementado en Perl y Cutadapt (Martin, 2011), para su posterior conversión a formato \*.fastq.

Para ambos sets de lecturas por cada plataforma de secuenciación, se eliminaron las secuencias con un índice de calidad en escala Phred menor a 20 y menores a 50 pb. Ambos tipos de lecturas fueron importados al software Newbler 2.5 para realizar un ensamble de lecturas híbrido. Se obtuvieron secuencias ensambladas (*contigs*) con traslapes de al menos 40 bases y 90% de identidad, y se separaron de las secuencias que no fue posible ensamblar (*singlets*) mediante un programa tipo script implementado en Perl (desarrollado por Tim Booth, NERC *Environmental Bioinformatics Centre*).

#### 6.3.2.1 Anotación de secuencias obtenidas en genotecas

Tanto las lecturas obtenidas por secuenciación Sanger como las obtenidas del ensamblaje de lecturas de 454-FLX, que incluye las secuencias tanto tipo *contigs* como *singlets*, fueron transferidas a un archivo multifasta el cual fue sometido a un proceso de anotación funcional empleando la plataforma bioinformática Blast2GO<sup>®</sup> (Gotz *et al.*, 2008), configurando los parámetros de BlastX con un valor  $E < 10^{-6}$  contra la base de datos de proteínas *nr* de NCBI. Los *contigs* construidos con lecturas específicas de cada genoteca de hibridación sustractiva (gónada de región testicular en maduración; gónada inactiva) fueron identificados mediante la exploración y análisis de los archivos de ensamble que incluyen

la información relacional de lecturas que conforman a cada *contig*, para posteriormente realizar un análisis de enriquecimiento en cada genoteca por la prueba exacta de Fisher (opción de análisis en el software Blast2GO<sup>®</sup>), generando tablas de categorías funcionales que se encontraron más representadas en cada genoteca.

#### 6.3.2.2 Identificación de locus de microsatélites en fragmentos de transcritos de genes aislados en genotecas de hibridación sustractiva

Las lecturas individuales obtenidas por ambas plataformas de secuenciación fueron exploradas *in silico* con el programa QDD (Megléczy *et al.*, 2010), el cual es implementado en lenguaje Perl y utiliza ClustalW2 (Larkin *et al.*, 2007) y blast+ (Camacho *et al.*, 2009) para alinear las lecturas flanqueantes de las repeticiones, y así reducir la redundancia de potenciales microsatélites. También se incorporó Primer3 v1.1.4 (Untergasser *et al.*, 2007) para diseñar el mejor par de oligonucleótidos para cada locus microsatélite en las secuencias que flanquean al mismo. El número mínimo de repeticiones permitido fue de cinco para los motifs di-, tri-, tetra- y pentanucleótidos. Éste programa también fue empleado para determinar aquellos microsatélites que eran polimórficos, basados en las variaciones observadas en la extensión del motif entre las lecturas individuales que conservan las mismas regiones flanqueantes. Los microsatélites fueron posteriormente explorados con blastn contra la base de datos nr-GenBank con fines de establecer su similitud con genes y secuencias de microsatélites ya reportadas.

#### 6.4 Aislamiento *de novo* de transcritos de genes

Para la evaluación y estandarización de los genes de referencia a utilizar en la normalización de la expresión génica, se realizó la amplificación parcial por PCR de punto final de transcritos identificados como potenciales genes de referencia (*18S-RNA*, *riboproteína 'rp-l8'*, *alfa-tubulina 'tub-a'*, *beta-actina 'act-'*, y *factor de elongación 1-alfa 'ef-1a'*) así como de dos genes selectos a utilizar como genes a normalizar durante la validación de estabilidad (*DEAD-box helicase 'ddx'* y *vitelogenina 'vtg'*).

Adicionalmente, con el fin de obtener las secuencias completas de algunos de los

transcritos de genes descubiertos en las genotecas así como de otros genes conocidos por su función en la meiosis, se utilizó la técnica de RACE (*Rapid Amplification of cDNA Ends*). Esta estrategia fue usada para amplificar los transcritos completos de los genes de meiosis *synaptonemal complex protein 3 (scp3)*, *disrupted meiotic cDNA (dmc1)*, *RAD51 homolog (rad51)*, aislados en la genoteca de hibridación sustractiva de gónada-región testicular en maduración, así como de los genes *meiotic recombination homolog 11 (mre11)* y *pachytene checkpoint protein 2 (pch2)* cuyos fragmentos parciales fueron primero aislados por PCR degenerativa.

#### 6.4.1 Síntesis de ADNc para obtención de transcritos parciales

Antes de la síntesis de ADNc, el ARN total aislado fue cuantificado por espectrofotometría UV utilizando un instrumento NanoDrop-1000 (Thermo), y 2 µg de ARN fueron resuspendidos en 10 µl de mezcla de reacción con 1 U de DNAsa RQ1 (Promega) y fueron incubadas a 37 °C por 30 mins, seguido de una incubación a 65 °C por 10 min para inactivar la DNAsa. Una vez realizado el tratamiento con DNAsa, la síntesis de ADNc fue realizada a partir de 500 ng de ARN total tratado con DNAsa, y agregando 500 ng de cada primer de transcripción reversa (oligo-dT), e incubando ésta mezcla a 70 °C por 5 min, 25°C por 5 min, y seguido inmediatamente de una incubación en hielo. Posteriormente, la mezcla enzimática para transcripción reversa fue agregada obteniendo una reacción con 1 µl de transcriptasa reversa, 0.5 mM de dNTP's, 40 U de RNAsin, MgCl<sub>2</sub> 5.6 mM y DTT 10 mM, en un volumen final de 20 µl. La transcripción reversa se llevó a cabo a 42° C durante 60 min, incubándose posteriormente a 75 °C por 5 min para inactivar a la transcriptasa reversa. El ADNc generado fue almacenado a -20 °C hasta su utilización en la obtención de transcritos parciales por amplificación por PCR de punto final.

#### 6.4.2 Síntesis de ADNc para amplificación de transcritos por RACE

La síntesis de ADNc para la secuenciación completa de transcritos de interés se realizó por medio de la técnica RACE (*Rapid Amplification of cDNA Ends*), descrita por Zhu *et al.* (2001), y que permite la amplificación independiente de los extremos 3' y 5' a partir de un primer fragmento intermedio de ADNc de secuencia conocida (Fig. 7). Para la síntesis de

ADNc y la amplificación de los extremos 3' y 5' en los transcritos de interés, se empleó 1 µg de ARN mensajero (fracción poliadenilada) purificado como se indica en la sección 5.2, y se siguió el protocolo proporcionado en el kit comercial *SMARTer RACE cDNA Amplification Kit* (Clontech). En resumen, el ADNc para la amplificación por RACE de los extremos 3' fue sintetizado con la enzima *SmartScribe* (Clontech), y se empleó un oligo-dT (5'-AAG CAG TGG TAT CAA CGC AGA GTA CT<sub>(30)</sub>N<sub>-1</sub> N-3'; N = A, C, G, o T; N<sub>-1</sub> = A, G, o C) con la secuencia del adaptador SMART de Clontech, mientras que el ADNc para la amplificación de los extremos 5' fue generado con un oligo-dT<sub>(15)</sub> y un adaptador SMARTer IIA (5'-AAG CAG TGG TAT CAA CGC AGA GTA CXX XXX -3' (las X muestran una secuencia desconocida protegida por propiedad intelectual de Clontech) que resulta ligado en el extremo 5' del ADNc por la actividad deoxinucleotidil transferasa intrínseca de la transcriptasa reversa *SmartScribe*. Ambas reacciones se realizaron en un volumen final de 20 µl, y empleando una concentración de 10 mM de dNTPs, 40 U de RNAsin y 10 mM de DTT para la reacción de transcripción reversa.

#### 6.4.3 Amplificación por PCR de punto final para obtención de secuencias parciales de genes

La estrategia empleada para obtener fragmentos parciales específicos de secuencias de ADNc para los genes *pch2*, *mrel1*, *vlg*, *ddx*, así como para los potenciales genes de referencia *18S*, *rp-18*, *tub-a*, *act-*, *ef-1a*, se basó en un aislamiento por PCR empleando oligonucleótidos degenerativos. Para el diseño de oligonucleótidos, se realizaron alineamientos de secuencias completas de nucleótidos/proteínas descargadas de la base de datos en GenBank en diversas especies de invertebrados y vertebrados, buscando identificar regiones conservadas de al menos 7-8 aminoácidos. Las regiones conservadas en aminoácidos fueron posteriormente realineadas incluyendo las secuencias nucleotídicas para la elección manual de oligonucleótidos con múltiples variaciones de bases en determinadas posiciones, empleando la nomenclatura IUPAC, y comprobando que las diferencias en temperatura de disociación para cada par de oligonucleótidos, así como la estabilidad de dímeros fuera mínima, utilizando el calculador de termodinámica para oligonucleótidos de Sigma-Aldrich (<http://www.sigma-geosys.com/calc/DNACalc.asp>).

Las reacciones de amplificación por deg-PCR se prepararon con las siguientes concentraciones de componentes: 10 mM de dNTP, 1 U *Taq* (Platinum *Taq*, Invitrogen), 1.5 a 2.5 mM de MgCl<sub>2</sub>, y agregando de 10 a 20 pmol de oligonucleótidos degenerativos o específicos y 1 µl de ADNc; todo en reacciones de 50 µl. Las condiciones de amplificación incluyeron un paso de disociación de 94 °C por 15 s, seguido de la hibridación de los oligonucleótidos a 4°C por debajo de su T<sup>°m</sup> (temperatura de disociación media, calculada en [www.oligocalculator.com](http://www.oligocalculator.com)) durante 30 s, y su extensión a 72 °C por 1 min/Kb. Después de 30-35 ciclos de amplificación por PCR, los productos fueron separados en geles de agarosa al 1%, y los productos de tamaño esperado fueron aislados y purificados de la banda del gel de agarosa (PCR gel purification kit, Qiagen), insertados en el vector pCR2.1 vector (Invitrogen) o pGEM-T (Promega), y clonados en *E. coli* DH5-. La amplificación se realizó a partir de una poza de ADNc sintetizado como se describe en la sección 6.4.1, a partir de muestras de gónada en condiciones de inmadurez y madurez (incluyendo regiones tanto femenina como masculina para la condición de madurez) hasta obtener fragmentos de amplificación únicos de tamaño esperado. En algunos casos, cuando se observó más de un fragmento, se purificaron las bandas obtenidas directamente del gel de agarosa, y se clonaron independientemente. Para el aislamiento de ADN plasmídico, 1 ml de cultivo bacteriano fue lisado en condiciones alcalinas y posteriormente extraído con fenol-cloroformo (Sambrook y Russell, 2001). El ADN plasmídico fue enviado a secuenciar por electroforesis capilar a la compañía Genewiz (NJ) o Macrogen (Corea). Las secuencias obtenidas fueron editadas mediante la herramienta de limpieza de vector VecScreen (<http://www.ncbi.nlm.nih.gov/VecScreen/VecScreen.html>), y se corroboró su identidad mediante un análisis de alineamiento por blastx contra la base de datos no redundante de GenBank-NCBI, registrando los mejores hits en similitud.

#### 6.4.4 Amplificación por RACE-PCR de extremos 3' y 5' para transcritos de interés

La técnica de *Rapid Amplification of cDNA Ends* (RACE) fue empleada para amplificar de forma específica los extremos 5' y 3' de una secuencia parcial de transcritos para los que únicamente se cuenta con una secuencia intermedia, a partir de la cual se diseñan oligonucleótidos que se orientan al sentido de cada extremo del transcrito. En la figura 7 se

presenta el esquema del uso de oligonucleótidos gen-específicos para la amplificación de extremos por RACE-PCR. Los oligonucleótidos específicos para algunos de los genes seleccionados fueron diseñados a partir de la secuencia específica de fragmentos parciales obtenidos en las genotecas de hibridación sustractiva; para el caso de los transcritos aislados por PCR degenerativa, se empleó la secuencia así obtenida. Los oligonucleótidos obtenidos a partir de secuencias específicas se diseñaron utilizando el software Primer3 v0.4 (<http://frodo.wi.mit.edu/primer3/>), Las amplificaciones de los fragmentos extremos 5' y 3' fueron realizadas empleando la enzima polimerasa incluida en el sistema *Advantage cDNA amplification* (Clontech), también siguiendo la recomendación del sistema de amplificación RACE. Las reacciones fueron preparadas en volumen de 20 µl, empleando 0.4 µl de enzima polimerasa, así como 10 pmol de oligonucleótidos de amplificación específicos y 10 pmol µl de la mezcla de oligonucleótidos UPM dirigidos a la secuencia de adaptador SMARTer agregado durante la síntesis del cDNA (5'-CTA ATA CGA CTC ACT ATA GGG C-3 y 5'-CTA ATA CGA CTC ACT ATA GGG CAA GCA GTG GTA TCA ACG CAG AGT-3'; ambos en proporción 1:0.2). Las reacciones se condujeron con un programa de amplificación propuesto por el fabricante del kit (desnaturalización a 94 °C x 30 S, alineamiento de oligonucleótidos a 55-60 °C por 30 S y una extensión a 72 °C por 2 min, amplificando a 25-35 ciclos). Los productos fueron analizados en electroforesis de agarosa para la revisión de la amplificación de fragmentos de tamaño esperado, y cuando la amplificación era tenue, se agregaron 3-5 ciclos adicionales. Los fragmentos amplificados fueron aislados a partir de bandas cortadas en el gel de agarosa, eluidas con un sistema de adsorción en sílica (PCR gel purification & Clean-Up Kit; MO BIO, CA), y posteriormente fueron clonados con el sistema TOPO/TA en células competentes DH5 como fue descrito anteriormente para la posterior purificación de ADN plasmídico y su secuenciación capilar por el método de Sanger (Genewiz, NJ).

## Amplificaciones PCR-RACE

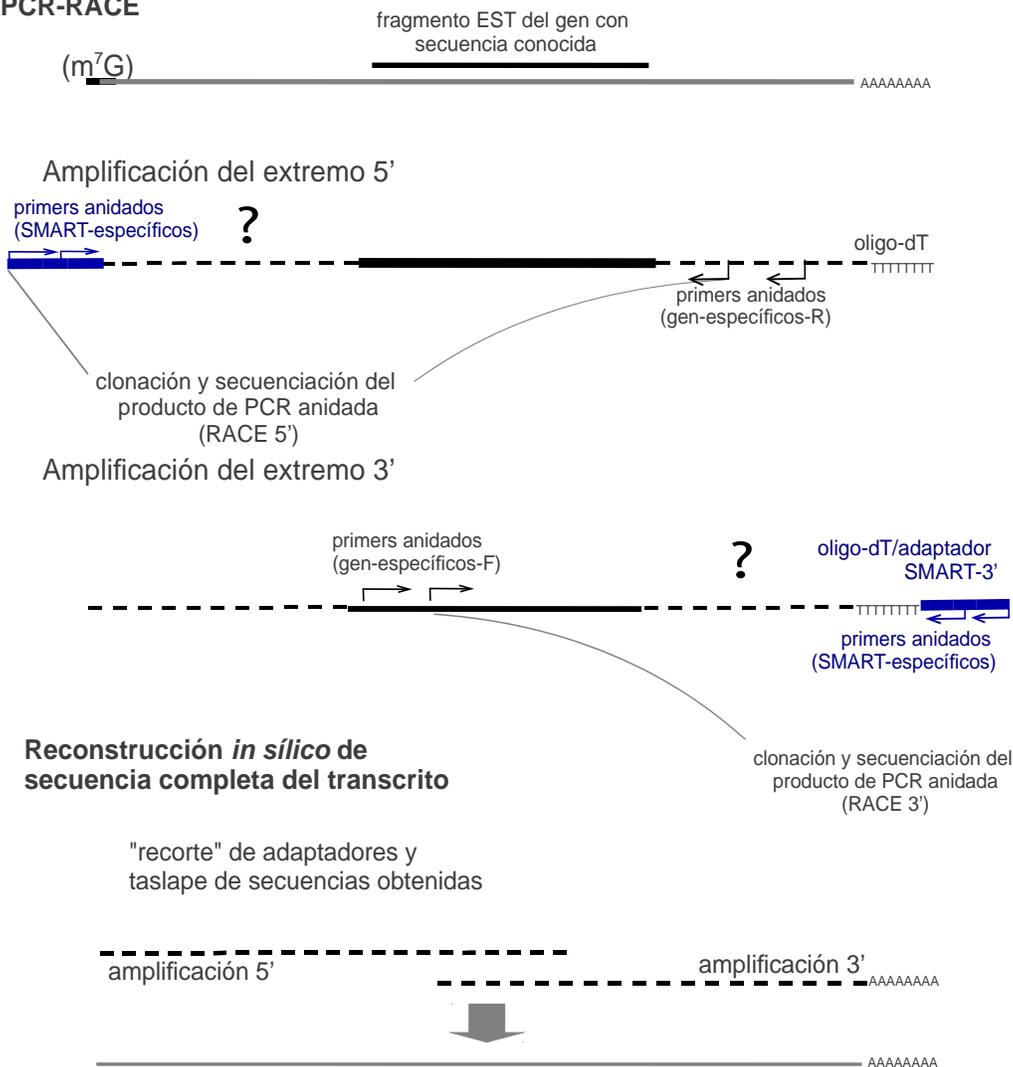


Fig. 7. Esquema de amplificación de extremos 5' y 3' por RACE en ADNc SMART. La amplificación de cada extremo se basa en una reacción de PCR con oligonucleótidos específicos del gen y oligonucleótidos dirigidos al adaptador SMART que previamente fue incorporado al ADNc durante la transcripción reversa (no mostrado en el esquema). El ensamble de la secuencia completa se basa en el traslape de ambas regiones (5' y 3') en la región intermedia conocida. Cada secuencia RACE es amplificada por PCR anidadas (dobles flechas) para asegurar la especificidad de las reacciones PCR.

Las secuencias obtenidas fueron editadas, ensambladas y verificadas con los cromatogramas correspondientes para obtener una secuencia única, nombrada generalmente como *full-length cDNA* o ADNc de secuencia completa, la cual incluye ambos extremos no traducidos (UTRs) y/o desconocidos dentro del marco de lectura, así como el fragmento intermedio originalmente conocido. Posteriormente se condujeron análisis bioinformáticos en la secuencia única obtenida para determinar la región correspondiente al marco de lectura de traducción proteica y su traducción conceptual en aminoácidos (ORFinder, NCBI; <http://www.ncbi.nlm.nih.gov/gorf/gorf.html>), para posteriores análisis filogenéticos.

#### 6.4.5 Análisis filogenéticos y estructurales de genes '*full-length*'

Para corroborar la correcta identificación de los transcritos de interés (*pch2*, *rad51* y *dmrt1*) y determinar la posición filogenética que guardan con respecto a otros grupos filogenéticos, se obtuvieron secuencias completas de proteínas ortólogas identificadas en diversos grupos de metazoarios a partir de las bases de datos Entrez-GenBank de NCBI. Las secuencias de aminoácidos de ortólogos, incluyendo la traducción conceptual de cada transcrito obtenido en *N. subnodosus*, fueron alineadas mediante el programa MUSCLE ver. 3.8 (Edgar, 2004) en configuración por defecto. Después de los alineamientos, las regiones ambiguas (con gaps o pobremente alineadas) fueron removidas mediante la construcción de bloques de alineamientos mediante el software GBlocks v.0.91b (10 posiciones como tamaño mínimo de cada block posterior a la limpieza de gaps, máximo 8 posiciones no conservadas permitidas dentro de cada bloque) (Castresana, 2000). Para las secuencias de PCH2, se generaron árboles a partir de los alineamientos se generaron mediante distintas aproximaciones. se empleó el software PhyML ver.3.0 con la función aLRT (*approximate likelihood ratio test for branches*) activada, e incluyendo una matriz de sustitución de aminoácidos tipo WAG, derivada de casi 4000 secuencias agrupadas en 182 familias proteicas y diseñada para analizar proteínas globulares bajo métodos empíricos de máxima verosimilitud (Whelan y Goldman, 2001). Los árboles obtenidos fueron analizados estadísticamente mediante un método aproximativo basado en máxima verosimilitud (Anisimova y Gascuel, 2006). El análisis se condujo en la plataforma Phylogeny.fr

(Dereeper *et al.*, 2008), la cual alberga las herramientas de análisis anteriormente mencionadas en un servidor web (<http://www.phylogeny.fr/version2.cgi/index.cgi>). Para los árboles filogenéticos de RAD51/DMC1 se generaron árboles por máxima verosimilitud, empleando el software MEGA5 (Tamura *et al.*, 2011), con las matrices de sustitución de aminoácidos determinadas como más apropiadas con la función *MODELTEST* (Posada y Crandall, 1998) incorporado en MEGA5, y con un soporte estadístico de *bootstrap* con 1000 iteraciones.

La secuencia de aminoácidos de RAD51 fue contrastada mediante comparación estructural en modelos proteicos generados *in silico*, empleando el servidor web Swiss-Model (Arnold *et al.*, 2006). Los modelos obtenidos para la secuencia de *N. subnodosus*, así como de las referencias descargadas de GenBank, fueron exportados en formato \*.PDB y alineados mediante el servidor web FATCAT (Ye y Godzik, 2004).

## 6.5 Análisis de expresión génica.

### 6.5.1 Síntesis de ADNc para la validación de genes de referencia.

A partir de ARN total de muestras de gónada individuales, se sintetizó ADNc mediante dos estrategias de *priming*, empleando: a) oligonucleótidos hexanucleótidos aleatorios y b) oligo-dT como iniciadores de transcripción reversa, con el objetivo de definir el método de *priming* más adecuado para el uso de genes de referencia potencialmente no poliadenilados (p.ej., fragmento de la subunidad *18S* del ARN ribosomal). Se siguió el mismo procedimiento que el descrito en la sección 6.4.1, pero en este caso se sintetizaron dos ADNc, agregando 500 ng de cada primer de transcripción reversa (oligo-dT o hexanucleótidos aleatorios) en reacciones independientes preparadas de forma simultánea.

### 6.5.2 Síntesis de ADNc para análisis de expresión de genes de meiosis.

El resto de análisis de expresión de genes de interés fue realizado únicamente en ADNc sintetizado con oligo-dT, y siguiendo el mismo procedimiento que el descrito en la sección 6.4.1. El ADNc así generado fue almacenado a -20 °C hasta su análisis por qPCR.

### 6.5.2 Cuantificación de expresión de genes de interés

La cuantificación de expresión se realizó mediante la síntesis de ADNc (*ImProm II reverse transcription kit*, Promega, EU), con un tratamiento posterior con DNAsa I (Promega) y la amplificación mediante PCR en tiempo real (qPCR) de fragmentos específicos de cada mensajero a analizar, empleando el sistema de detección fluorescente con EvaGreen en una mezcla de reacción de PCR (Hernández-Arteaga y López-Revilla, 2008; Llera-Herrera *et al.*, 2012). La reacción de PCR se realizó por duplicado o triplicado en cada muestra, y adicionalmente se analizó de forma paralela los genes de expresión constitutiva como referencia de normalización entre reacciones. Las reacciones de amplificación se realizaron por triplicado o duplicado, en placas de 96 pozos con cubiertas adheribles plásticas, empleando el instrumento CFX96 (Bio-Rad). La mezcla contuvo 2.5 mM de MgCl<sub>2</sub>, 200 μM de cada deoxinucleótido, 0.46 μM de cada primer (sentido y antisentido), 0.45 U de *Taq* polimerasa GoTaq (Promega) y compuesto intercalante fluorescente EvaGreen (Biotium) a una concentración final 1X, todo en un volumen final de 15 μl. En cada análisis de expresión, se preparó mezcla de reacción suficiente para todas las reacciones, evitando así que exista variación técnica en los resultados por efecto de la variación en concentración y calidad de los componentes, y en un inicio se comparó el desempeño de la mezcla realizada contra un sistema comercial de qPCR (SsoFast EvaGreen qPCR SuperMix 2x; Bio-Rad), obteniendo resultados casi idénticos cuando se comparó la amplificación de diluciones seriales (1:5) idénticas en términos de pendiente e intersección en  $y$  (donde  $x$  es el  $C_t$  registrado, y  $y$  es el logaritmo de la concentración) de las curvas estándar generadas (mix comercial: pendiente = -3.482, intersección= 9.208; mezcla preparada en laboratorio: pendiente= -3.505; intersección= 9.288). Las condiciones de amplificación para la qPCR fueron en todos los casos: 94 °C 3 min (desnaturalización inicial; 40 ciclos de: 94 °C (10 seg), 60°C (15 seg) y 72 °C (30 seg), adquiriendo la fluorescencia a 79 °C (2 seg). Al final del programa de qPCR, una disociación de 60 a 94 °C fue llevada a cabo a un rampeo térmico de 1 °C/seg, para obtener una curva de disociación de los amplicones generados y así inspeccionar la amplificación específica de cada reacción.

La estimación de la eficiencia de amplificación se realizó para cada gen, mediante el cálculo de la pendiente de la amplificación sobre diluciones seriadas (1:5), comenzando a partir de 0.5  $\mu$ l de un pool de cDNA de distintos estadios/sexos. Para aquellos genes de expresión baja ( $Ct \sim 28$  o mayor), se empleó una dilución 1:10,000 de un producto de PCR a 30 ciclos. Los valores de eficiencia fueron obtenidos a partir de la pendiente obtenida en la función de la concentración (log) sobre  $Ct$  (lineal), usando la ecuación  $E = (10^{(-1/pendiente)} - 1)$ , y empleados para la corrección de eficiencia de cada gen en los posteriores cálculos de expresión relativa como sugieren (Pfaffl, 2001).

Los valores de expresión relativa para cada gen *de interés* fueron determinados mediante el cálculo de expresión relativa (RE), con base en la comparación de su amplificación contra un factor de normalización que integra la amplificación de los genes constitutivos para cada muestra, como se detalla más adelante.

### 6.5.3 Análisis de estabilidad de genes de referencia entre estadios de gametogénesis.

Las secuencias parciales de genes obtenidos en *N. subnodosus* con fines de utilizarlas como genes de referencia, fueron incluídos los genes *18S*, *rpl8*, *ef1*, y *act*. Adicionalmente se incluyó un fragmento de un transcrito de un gen *alfa-tubulina* aislado en las genotecas, *tub-a*. A partir de esas secuencias parciales se diseñaron oligonucleótidos apropiados para su amplificación por qPCR (amplicones de 150-200 pb;  $T^{\circ}m \sim 60^{\circ}C$ ), y se analizó la estabilidad de estos genes en un contexto de desarrollo gametogénico en ambas regiones de la gónada. La estabilidad se analizó en muestras de gónada en estadio de indiferenciación ( $n = 6$ ), estadio temprano de espermatogénesis y ovogénesis ( $n = 6$  por cada región sexual) y estadio avanzado de espermatogénesis y ovogénesis ( $n = 6$  por cada región sexual). Los análisis de estabilidad se realizaron mediante los macros en Office Excel (Microsoft) de análisis de GeNorm (Pfaffl *et al.*, 2004) y NormFinder (Vandesompele *et al.*, 2002), proporcionados por los autores. Debido a que en estos análisis se encontró que el gen *alfa-tubulina* presentaba la menor estabilidad de expresión entre los 5 genes de referencia originalmente seleccionados, este gen fue considerado en adelante como un gen de interés (blanco) para evaluar su expresión relativa. Para evaluar el empleo de los genes de referencia finalmente seleccionados, se analizó la expresión de los genes *ddx*, *tub-a* y *vit* en

contexto gametogénico como se detalla más adelante.

#### 6.5.4 Análisis de expresión de genes *ddx*, *vit* y *tub-a* durante la validación de genes de referencia.

Se condujo un análisis de expresión de los genes *ddx*, *vit* y *tub-a* (este último originalmente seleccionado como potencial gen de referencia) utilizando ADNc sintetizado con oligo-dT. Las secuencias parciales de éstos genes de interés fueron aislados mediante PCR degenerativa en una estrategia similar a la empleada para el set de genes de referencia (sección 6.4.3). La expresión relativa fue calculada con base en distintas combinaciones de genes de referencia como factores de normalización.

##### 6.5.4.1 Ubicuidad de expresión de *tub-a*.

Se re-analizó la expresión de *tub-a* en ambas regiones sexuales en maduración activa, además de tejidos somáticos de músculo aductor, branquias, manto, glándula digestiva y nefridios (n=6 muestras por tejido). De éstas muestras, se obtuvo ARN total y se sintetizó ADNc mediante *priming* con oligo-dT. Las reacciones de qPCR se realizaron por duplicado, incluyendo todas las muestras en una placa de reacción para cada gen, evitando variaciones entre reacciones. Se analizó todo el set de genes de referencia antes empleado en la sección correspondiente a la selección de genes de referencia entre estadios de gametogénesis mediante los algoritmos de geNorm y NormFinder, y el mejor par de genes de referencia fue empleado para la normalización de la expresión relativa, empleando el software qBasePlus.

##### 6.5.4.2 Análisis de expresión de genes de meiosis y diferenciación sexual.

Se realizó la cuantificación relativa de la expresión de genes identificados en el proceso meiotico: *scp3*, *dmc1*, *rad51*, *mre11*, *horma*, *e3-ccnb1ip1*, *p33-ringo* y *dmrt-1*. Los análisis se realizaron mediante qPCR en ADNc sintetizado con oligo-dT de muestras de gónadas indiferenciadas, así como en regiones sexuales de gónadas en estadio de maduración temprana y tardía. Para cada estadio, seis muestras individuales fueron analizadas en duplicado. Los valores de expresión relativa fueron normalizados con la media geométrica

de la expresión de los genes de referencia *rp-18* y *18S*, y analizados mediante la comparación de la expresión relativa (RE).

#### 6.5.5 Determinación de expresión relativa mediante cálculo de cantidad relativa (RQ) y análisis estadísticos.

La expresión génica relativa fue calculada a partir de las cantidades relativas (RQ) de cada gen (de referencia o de interés), y fueron estimadas con la ecuación:

$$RQ = (1+E)(Ct \text{ promedio} - Ct) \quad (\text{ecuación 1})$$

Los valores de expresión relativa (RE) fueron calculados entonces a partir de la proporción de la cantidad relativa (RQ) de cada muestra individual con la ecuación:

$$RE = RQ_i / RQ_{nf} \quad (\text{ecuación 2})$$

Donde  $i$  = gen de interés y  $nf$  = factor de normalización = media geométrica de los genes de referencia determinado para cada caso). Los valores de Ct mayores a 35 fueron considerados como amplificación indetectable. Los análisis estadísticos se realizaron mediante ANOVA de una vía, después de transformar a logaritmo natural. Las medias fueron comparadas mediante la prueba *post-hoc* de Fisher, con el software Statistica versión 7 (StatSoft). Los resultados se presentan una vez re-transformados desde los valores de logaritmo natural a escala lineal, y las diferencias significativas son presentadas con un valor de significancia de  $P < 0.05$ ).

## 7 RESULTADOS

### 7.1 Generación de información genómica asociada a la reproducción en *N. subnodosus*.

#### 7.1.1 Identificación de genes expresados diferencialmente en la gametogénesis mediante generación y secuenciación de genotecas de hibridación sustractiva.

Se desarrollaron dos genotecas de genes específicos o expresados diferencialmente entre dos condiciones gonadales: estadio de inactividad (gónada entera) y estadio de maduración avanzada (región sexual testicular), y los fragmentos obtenidos fueron secuenciados mediante dos plataformas. Durante el desarrollo de las genotecas, diversos controles de reacción fueron considerados para confirmar que el flujo de reacciones y procedimientos ocurrieron adecuadamente. El primer paso de revisión fue la síntesis de ADNc doble cadena fragmentado con la enzima de restricción *RsaI*, y al cual se le agregaron adaptadores en ambos extremos. Para esto, se realizó una PCR con combinaciones de primers para actina humana (incluidos en el kit de generación de genotecas) y primers dirigidos a los adaptadores SMART, en una genoteca sustractiva "control" también incluidos en el kit de generación de genotecas (ARNm de músculo esquelético humano contra el mismo ARNm adicionado con 0.2% de fago X174 - digerido con *HaeIII* -) realizada estrictamente en paralelo a las de las genotecas de gónada (Fig. 8). Una vez revisado el hecho de que las genotecas fueron realizadas adecuadamente, se revisó que los fragmentos de ADN obtenidos en las genotecas de hibridación sustractiva se ubicaron en el rango de 200 a 1500 pb (Fig. 9), lo que posibilita su análisis mediante clonación y secuenciación Sanger, y sugiere que el procedimiento de construcción de genotecas se completó adecuadamente. Como confirmación de la adecuada clonación de los fragmentos de las genotecas en el vector pGEM en medio selectivo con X-gal, se realizó una revisión de la ocurrencia de insertos en 100 clonas seleccionadas al azar, de las cuales se determinó que el 96% de las clonas seleccionadas tenían un fragmento de ADN inserto (Fig. 10)

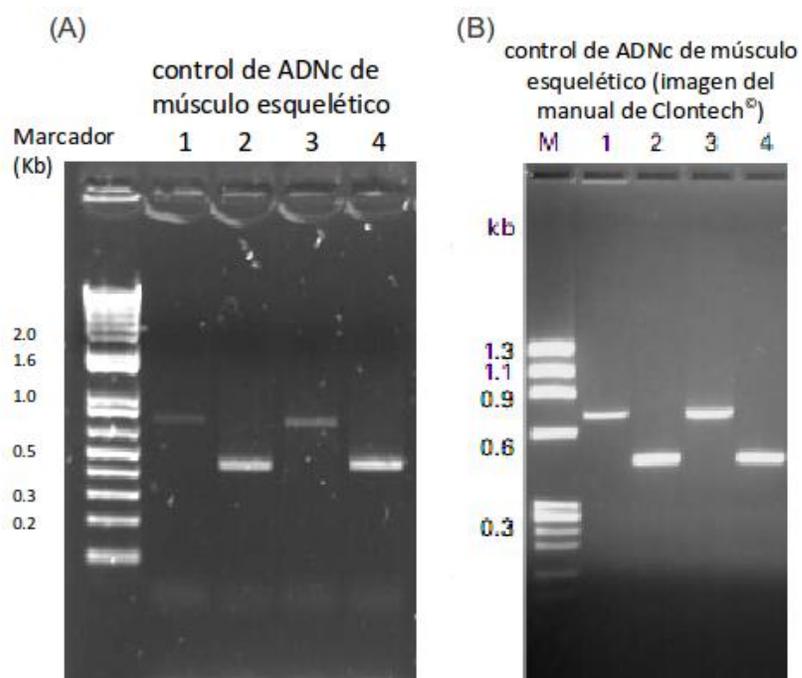


Fig. 8. Verificación de la adición de adaptadores en ADNc realizada en ADNc de músculo esquelético (reacción control), y amplificando mediante PCR con combinaciones de oligonucleótidos para el gen *-actina* de humano y oligonucleótidos dirigidos a los adaptadores A y B en extremo 5' (primer actina-sentido + primer adapter; carriles 1 y 3) y 3' (carriles 2 y 4) (imagen A). Para efectos comparativos, se muestra el resultado esperado proporcionado por el manual de kit de genotecas SSH de Clontech (imagen B).

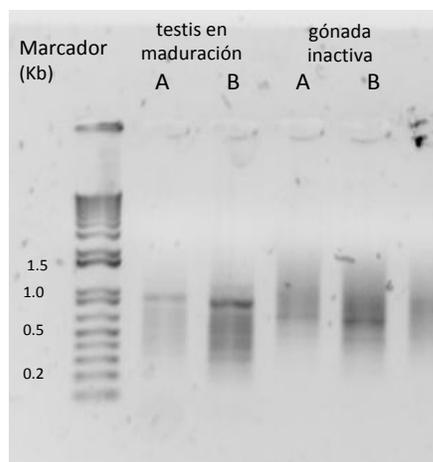


Fig. 9. Reacción de amplificación primaria (carriles A) y anidada (carriles B) con oligonucleótidos dirigidos a los adaptadores en productos de ADNc obtenidos después de la doble-hibridación sustractiva, representando aquellos fragmentos de transcritos expresados diferencialmente.

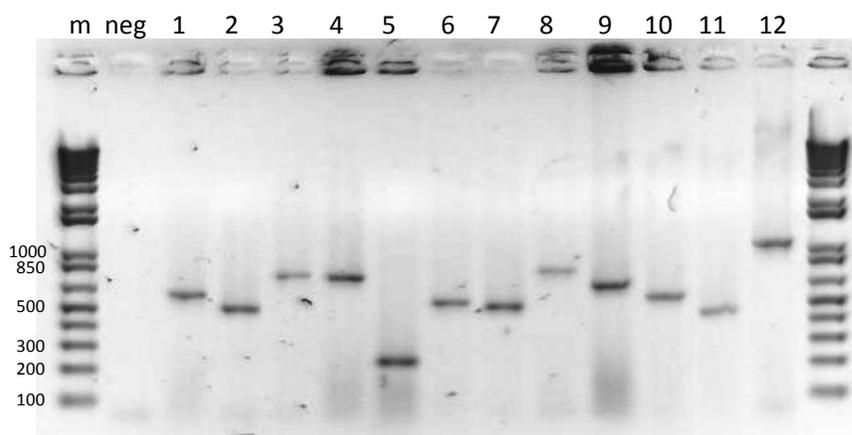


Fig. 10. Amplificación de clonas seleccionadas al azar de las regiones insertadas oligonucleótidos M13, universales en el vector) de fragmentos obtenidos en la genoteca de hibridación sustractiva de gónada de región testicular en maduración - sustraída con gónada inactiva. El carril 5 muestra un producto de amplificación de una clona sin inserto (184 pb)

Para la preparación de las genotecas para su secuenciación en la plataforma 454, se monitoreó mediante electroforesis en agarosa la amplificación con oligonucleótidos convencionales anidados dirigidos a los adaptadores de genotecas SSH, en paralelo a la amplificación con oligonucleótidos modificados que incluyen las secuencias adaptadores para secuenciación por 454-FLX (Lib-L) en el extremo 5'. Los resultados indican que la adición de oligonucleótidos modificados para la adición de adaptadores Lib-L, no afecta el patrón electroforético de los fragmentos amplificados que conforman la genoteca, y de hecho, la intensidad de la amplificación incrementó en comparación a la reacción que emplea oligonucleótidos convencionales anidados (Fig. 11). Este resultado sugiere que las genotecas se encontraban listas para ser secuenciadas en la plataforma 454-FLX mediante el protocolo de secuenciación de amplicones.

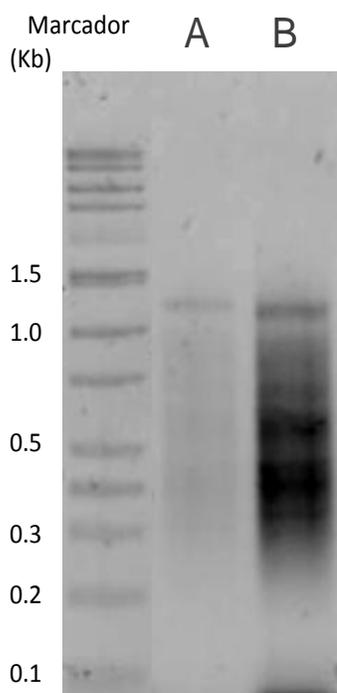


Fig. 11. Prueba de amplificación con oligonucleótidos anidados del sistema Clontech (carril 1) y de oligonucleótidos anidados modificados con las secuencias de adaptadores Lib-L requeridos para la secuenciación en sistema 454-FLX.

#### 7.1.2 Análisis de secuencias de genotecas de hibridación sustractiva.

Combinando las secuencias obtenidas mediante los métodos de electroforesis capilar (Sanger) y secuenciación masiva (454-FLX), se obtuvieron un total de 34,276 lecturas válidas mayores 50 pb y un valor de calidad Phred mayor a Q20. El número de secuencias obtenidas por Sanger fue 356 y 187 de las genotecas de región testicular en maduración e inactiva, respectivamente. Mediante 454-FLX, se obtuvieron 19,876 y 13,857 lecturas, respectivamente. Las lecturas obtenidas por secuenciación masiva fueron depositadas en el *European Nucleotide Archive* del servidor EMBL-EBI, bajo la clave de acceso ERP002162; y las secuencias obtenidas mediante secuenciación capilar (Sanger) fueron depositadas en la base de datos dbEST de NCBI (accesos JZ199499 - JZ200027). Posteriormente, con el software Newbler 2.5 (Roche) se realizó un ensamble híbrido con las lecturas obtenidas en ambas plataformas, el cual resultó en la generación de 3,727

contigs y 8,279 singletons. El tamaño modal (N50) de los isotigs generados fue de 429 pb.

Como parte del proceso de anotación funcional, se encontraron 2,037 transcritos con al menos un hit significativo por blastx; y de éstos, 51% de los mejores hits fueron contra proteínas de *C. gigas*, una especie cuyo genoma fue secuenciado y anotado recientemente. En relación a cada genoteca específica, 1,153 y 833 transcritos de las genotecas de la región testicular de la gónada en maduración y de la gónada inactiva respectivamente, presentaron un hit significativo en contra de la base de datos *nr*-GenBank ([archivo suplementario 1](#)), y de éstas cifras, 768 (67%) y 697 (84%) obtuvieron al menos un término ontológico (GO) que fue asociado de forma significativa mediante la plataforma de anotación blast2GO ([archivo suplementario 2](#))

La eficiencia de la sustracción durante la generación de las genotecas se observó cuando solamente 91 isotigs (de un total de 3,727) fueron construidos con lecturas de ambas genotecas; y de éstos, 48 tuvieron al menos un hit de blastx y al menos un GO asociado (3 tuvieron un hit por blastx pero no anotación GO). Entre éstos transcritos construidos con lecturas de ambas genotecas, se encontraron aquellos codificantes para proteínas ribosomales (n = 13), así como transcritos anotados como *cytochrome syntase* (n = 4), *ATP synthase* (n = 3) y *NADH dehydrogenase* (n = 2), entre otros. Los restantes 43 transcritos no obtuvieron un hit con blastx. Una prueba posterior de la eficiencia de sustracción durante la generación de las genotecas para la identificación de genes específicos de las condiciones biológicas en comparación se obtuvo mediante un análisis de enriquecimiento de términos ontológicos. Una comparación entre isotigs y singletons de cada genoteca mostraron los términos que están enriquecidos específicamente entre las genotecas de gónada inactiva y la región testicular en maduración. Los términos ontológicos obtenidos mediante blast2GO de *Meiosis* (GO:0007126), *Meiosis-I* (GO:0007127), *meiotic cell cycle* (GO:0051321) y *M phase of meiotic cell cycle* (GO:0051327) fueron frecuentemente asociados a genes caracterizados en la genoteca de la gónada masculina en maduración; adicionalmente, diversos términos asociados a espermatogénesis también se encontraron enriquecidos en esta genoteca (Fig. 12). Entre los genes identificados, se incluyen aquellos involucrados en la fidelidad de la replicación de ADN, sinapsis de cromosomas,

recombinación y *checkpoint* en meiosis (Tabla I), así como genes identificados dentro de la espermatogénesis (Tabla II) codificantes para antígenos de espermatozoides, cinasas testículo-específicas, así como proteínas constituyentes del flagelo. Adicionalmente a genes de meiosis y espermatogénesis, se identificaron genes con el término '*sex*' en su descripción (Tabla II) como *dmrt1* (*doublesex and mab-3 related transcription factor 1*), *dual specificity protein kinase clk2*, y *kelch-10 like*. Finalmente, se identificaron transcritos anotados con el prefijo '*transpos*' y '*retrotranspos*' para elementos transponibles y retrotransposones (Tabla III), incluyendo entre otros, *piggyBac transposable element-derived protein 3 like*, *retrotransposon transcriptase*, *RNA-directed DNA polymerase* derivadas de dos clases de elementos transponibles, *BS* y *jockey-like*, y una *retrovirus-related POL polyprotein* de un *type-I retrotransposable element R2*.

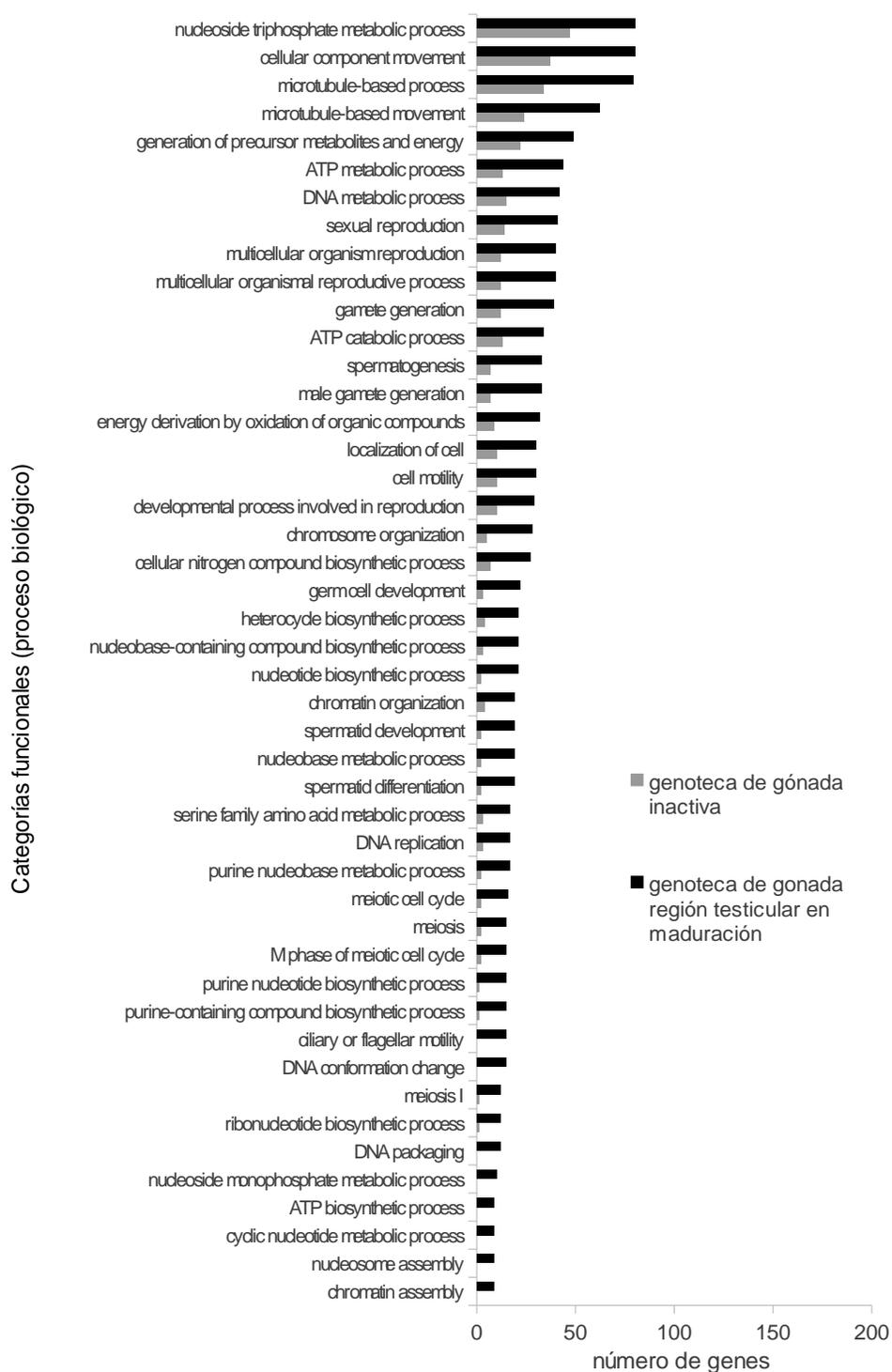


Fig. 12. Análisis de enriquecimiento de términos ontológicos para genotecas sustraídas recíprocamente entre gónada en maduración testicular y gónada inactiva, sobrerrepresentados en la porción testicular en maduración.

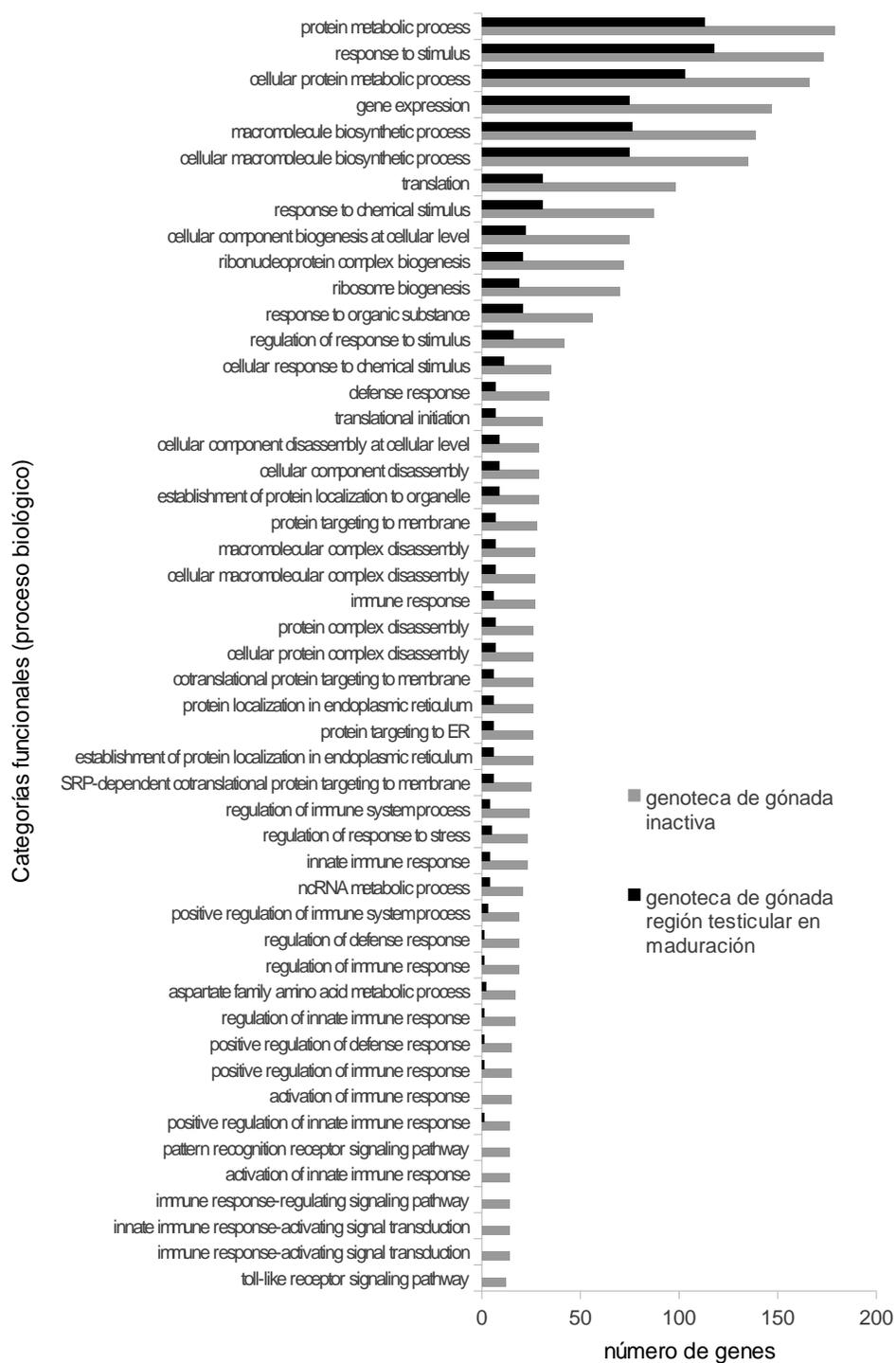


Fig. 13. Análisis de enriquecimiento de términos ontológicos para genotecas sustraídas recíprocamente entre gónada en maduración testicular y gónada inactiva, sobrerrepresentados en la gónada inactiva.

Tabla I. Genes asociados a la meiosis, expresados en la genoteca de región testicular de la gónada en maduración, con los términos ontológicos de *meiosis/meiotic; recombination, repair y checkpoint*.

Descripción de la anotación	valor E	mejor hit (acceso GenBank)
ATP-dependent DNA helicase Q1	1.27 e-45	EKC18968
ATP-dependent DNA helicase Q4	1.14 e-21	EKC33210
brain tumor protein	2.77 e-15	EGI57493
cell cycle checkpoint protein RAD1	5.58 e-22	BAD86790
cyclin dependent kinase 1	2.10 e-10	ADZ23998
denticleless protein homolog	2.99 e-36	EKC37484
citrate synthase	2.53 e-49	EDS33990.1
DNA ligase 1	1.09 e-16	XP_003127315
	4.63 e-47	EKC32253
	7.63 e-40	XP_002160108
	4.03 e-34	EFN72821
DNA polymerase delta subunit 2	3.29 e-60	EKC39895
DNA repair protein RAD51 homolog 3	1.06 e-62	EKC25786
DNA replication licensing factor MCM6	2.91 e-15	BAG61268
dual specificity protein kinase TTK	1.59 e38	EKC17444
e3 ubiquitin-protein ligase CCNB1IP1	9.79 e-19	XP_780975
	1.28 e-36	XP_002588845
histone 3	2.92 e-57	CAJ81662
histone 3.2-like	2.48 e-18	EJT46200
HORMA domain-containing protein 1-like	8.50 e-85	XP_002168952
	3.82 e-14	XP_003727321
importin subunit alpha-2	6.81 e-20	EKC18904
integrase family protein	4.29 e-22	XP_003390809
	3.38 e-14	EKC22946
	3.17 e-8	EKC28263
meiotic recombination protein DMC1/LIM15 homolog	1.89 e-94	XP_003221056
	5.91 e-45	XP_002831176
	5.66 e-8	XP_001367929
p33 RINGO	4.49 e-48	EKC24735
RuvB-like protein 2	3.26 e-63	EKC28232
serine threonine-protein kinase PLK1	1.18 e-48	EKC29729
synaptonemal complex central element protein 2	4.81 e-11	XP_002591412
synaptonemal complex protein 3	1.32 e-45	EKC40091
TFIIH basal transcription factor complex helicase subunit	1.84 e-11	AEO89688
ubiquitin-conjugating enzyme e2 c isoform 2	7.40 e-6	NP_861515
Werner syndrome ATP-dependent helicase	2.10 e-11	XP_003389810
zinc finger and BTB domain-containing protein 40	1.48 e-13	XP_002735651

Tabla II. Genes expresados en la genoteca de región testicular de la gónada en maduración, extraídos mediante la búsqueda de términos ontológicos (GOs) con los prefijos *sperm/spermatogenesis/spermatid* y *sex*.

Descripción de la anotación	valor E	mejor hit (acceso GenBank)
adenosine deaminase domain-containing protein 1-like isoform 2	1.14 e-10	EKC39025
C-Myc binding protein	7.84 e-44	XP_002604299
coiled-coil domain-containing protein 135	1.57 e-23	EKC19048
DnaJ subfamily member 13	5.88 e-51	EKC20056
doublesex and mab-3 related transcription factor 1	4.69 e-22	CAC42778
dual specificity protein kinase	1.86 e-7	XP_002432301
dynein heavy chain	5.06 e-90	EKC20777
growth arrest-specific protein 8	6.30 e-64	EKC37009
H1 histone member oocyte-specific-like	3.75 e-21	P22974
Kelch-like protein 10	1.92 e-57	EKC42522
	2.72 e-32	EFN76175
motile sperm domain-containing protein 2	1.65 e-26	EKC28345
nucleoside diphosphate kinase homolog 5-like	8.83 e-38	EKC43147
outer dense fiber protein 3-like protein 2-like	1.25 e-16	XP_002732801
parkin coregulated gene / protein homolog	2.68 e-57	EKC36429
	8.50 e-54	XP_003215805
peptidyl-prolyl cis-trans isomerase FKBP4	2.31 e-44	EKC41424
predicted protein (DM domain)	7.00 e-6	XP_001632590
probable ubiquitin carboxyl-terminal hydrolase FAF-X	4.12 e-11	XP_001627156
ropporin-1-like protein	2.73 e-126	EKC31617
sperm flagellar protein 2	1.07 e-7	XP_001623400
sperm phosphodiesterase 5	1.81 e-34	EKC21406
sperm-associated antigen 16 protein	2.32 e-81	EKC18115
	1.16 e-14	XP_002591791
sperm-associated antigen 17	2.31 e-54	EKC26001
spermatogenesis associated 4	1.66 e-83	EKC20206
spermatogenesis-associated protein 1	5.44 e-16	XP_002733667
tektin 2	7.57 e-73	EKC18085
	1.06 e-14	XP_002591791
testis-specific serine threonine-protein kinase 1	5.53 e-31	XP_002738751
testis-specific serine threonine-protein kinase 2	1.52 e-17	EHA99087
testis-specific serine threonine-protein kinase 3-like	9.62 e-24	EKC40940

Tabla III. Genes expresados en la genoteca de región testicular de la gónada en maduración, extraídos mediante la búsqueda de términos ontológicos (GOs) para *transposable*, *transposase*, *retrotransposon*, *retrotransposable* y *DNA-polymerase*.

Descripción de la anotación	valor E	mejor hit (acceso GenBank)
DNA-mediated transposase	1.14 e-13	ABH09251
piggyBac transposable element-derived protein 3-like	4.71 e-6	XP_003706354
tyrosine recombinase-like	6.04 e-9	DAA01995
pol-like polymerase (25 transcripts)	1.27 e-30	XP_002160589
retrotransposon transcriptase	4.82 e-10	EFN57005
retrotransposon unclassified	5.10 e-25	BAC82624
retrotransposon-like family member ( <i>retr-1</i> )-like	7.34 e-41	XP_003730886
retrovirus-related Pol polyprotein from type-1 retrotransposable element R2	2.51 e-40	XP_002127101
RNA-directed DNA polymerase from transposon Bs (13 transcripts)	2.89 e-21	XP_002127101
RNA-directed DNA polymerase from mobile element <i>jockey</i> -like (20 transcripts)	7.50 e-27	XP003727887
transposable element TCB2 transposase	1.28 e-5	ABV31711

En contraste a la gónada de la región testicular, en la genoteca de gónada inactiva diversos términos GO fueron asociados a la maquinaria de traducción proteica y la biosíntesis de ribosomas, así como a respuesta a estrés e inmune (Fig. 13). Adicionalmente, se identificaron otros genes involucrados en la diferenciación sexual y en el control y mantenimiento de la línea germinal como *notch homolog 2*, *btg member 2*, *pumilio domain-containing protein*, *cleavage and polyadenylation specificity factor subunit 5*, *nuclear receptor subfamily 0 group b member 1*, *transcription factor btf3 homolog 4*, *ccr4-not transcription complex subunit 10* y *transcription factor hes-1*, entre otros genes (Archivos suplementarios 1 y 2, en formato electrónico anexo).

### 7.1.3 Identificación de microsatélites en fragmentos de genes aislados en genotecas de hibridación sustractiva

Se identificaron 507 secuencias con microsatélites potenciales, y para 310 de éstas se logró un adecuado diseño de oligonucleótidos para su amplificación por PCR. Del total, 272 microsatélites puros y 38 compuestos fueron encontrados. Las secuencias microsatélites, con oligonucleótidos diseñados, y con el mejor hit contra secuencias nucleotídicas en

GenBank, son presentadas en el [archivo suplementario 3](#). Un microsatélite encontrado en una secuencia singleton (ERR224474.16682; SRA-NCBI) fué idéntico al microsatélite NsubA1B12 reportado para la misma especie (acceso GenBank FJ986341.1).

## 7.2 Caracterización de transcritos de genes selectos por RACE.

### 7.2.1 Transcrito de *mre11*.

La secuenciación de los extremos de ADNc 5' y 3' mediante RACE del gen *Mre11* permitió reconocer un ensamble de 1588 nucleótidos, que traducido en su marco de lectura +2, muestra una secuencia de 498 aminoácidos. En éste marco de lectura no fue posible identificar un codón de paro de traducción, pero sí una poliadenilación potencialmente prematura que comienza en la posición 1549 de la secuencia nucleotídica (Anexo I). El análisis de ésta secuencia en bases de datos proteicos mediante el programa blastp (secuencias proteicas *nr*) resultó en una estrecha similitud con la proteína MRE11 de *C. gigas* (score=837; valor E = 0) (Anexo I). y mediante un análisis de búsqueda de familias proteicas (Pfam 24.0; <http://pfam.sanger.ac.uk/>) dentro de la secuencia completa de la proteína se identificaron dos familias PfamA, ambas relacionadas con la arquitectura propia de proteínas del tipo MRE11: la familia de las fosfoesterasas similares a la calcineurina (acceso Pfam PF00149.21, score = 90.8; valor E =  $5.6^{-26}$ ) y la familia de enlace a DNA tipo *Mre11* (acceso Pfam PF04152; score = 259.9; valor E =  $1.1^{-77}$ ).

### 7.2.2 Caracterización de transcritos de *dmc1* y *rad51*

La secuencia completa del transcrito codificante para *dmc1* fue obtenida mediante la amplificación de extremos 3' y 5' por la técnica de RACE. Se obtuvo un ensamble de secuencia de 1274 nucleótidos, con un marco de lectura (+1) de 341 aminoácidos en la que puede identificarse una región correspondiente la familia proteica RAD51 (clave Pfam PF08423; score = 404.2, valor E =  $1.41^{-121}$ ) (Fig. 10). La secuencia del cDNA de *dmc1*, así como su traducción conceptual, fue depositada en GenBank, y está disponible bajo el acceso KC119408.

Por otra parte, la secuencia completa del transcrito de *rad51*, también obtenida por PCR-

RACE, posee una región codificante de 1861 nucleótidos (acceso GenBank KC119409). El marco de lectura inferido por su traducción conceptual muestra una secuencia de 360 aminoácidos, con un dominio conservado para la familia de recombinasas RAD51. Debido a que éste dominio también se encuentra presente en la secuencia de *dmc1*, se realizó un árbol filogenético de secuencias utilizando homólogos tanto de la subfamilia DMC1 como de RAD51 y XXRC3, y se utilizó el homólogo ancestral de REC-A como grupo externo. Las claves de las secuencias empleadas se muestran en el alineamiento múltiple realizado para la construcción del árbol, y fue recortado para eliminar las partes de las secuencias altamente divergentes, incluyendo la parte correspondiente a las región comprendida entre las posiciones 55 a 263 de la referencia de REC-A de *Escherichia coli* (ZP\_06991525.1) (Fig. 14). El árbol filogenético realizado muestra nombres estandarizados de los ortólogos de Rad51-C (RAD51-3) y RAD51-A (RAD51; RAD51-1), y separa claramente a las secuencias en grupos de RAD51 (A, C), XXRC3 y DMC1. La secuencia de RAD51 aislada en *N. subnodosus* se agrupa dentro del clado correspondiente a RAD51C (Fig. 15).



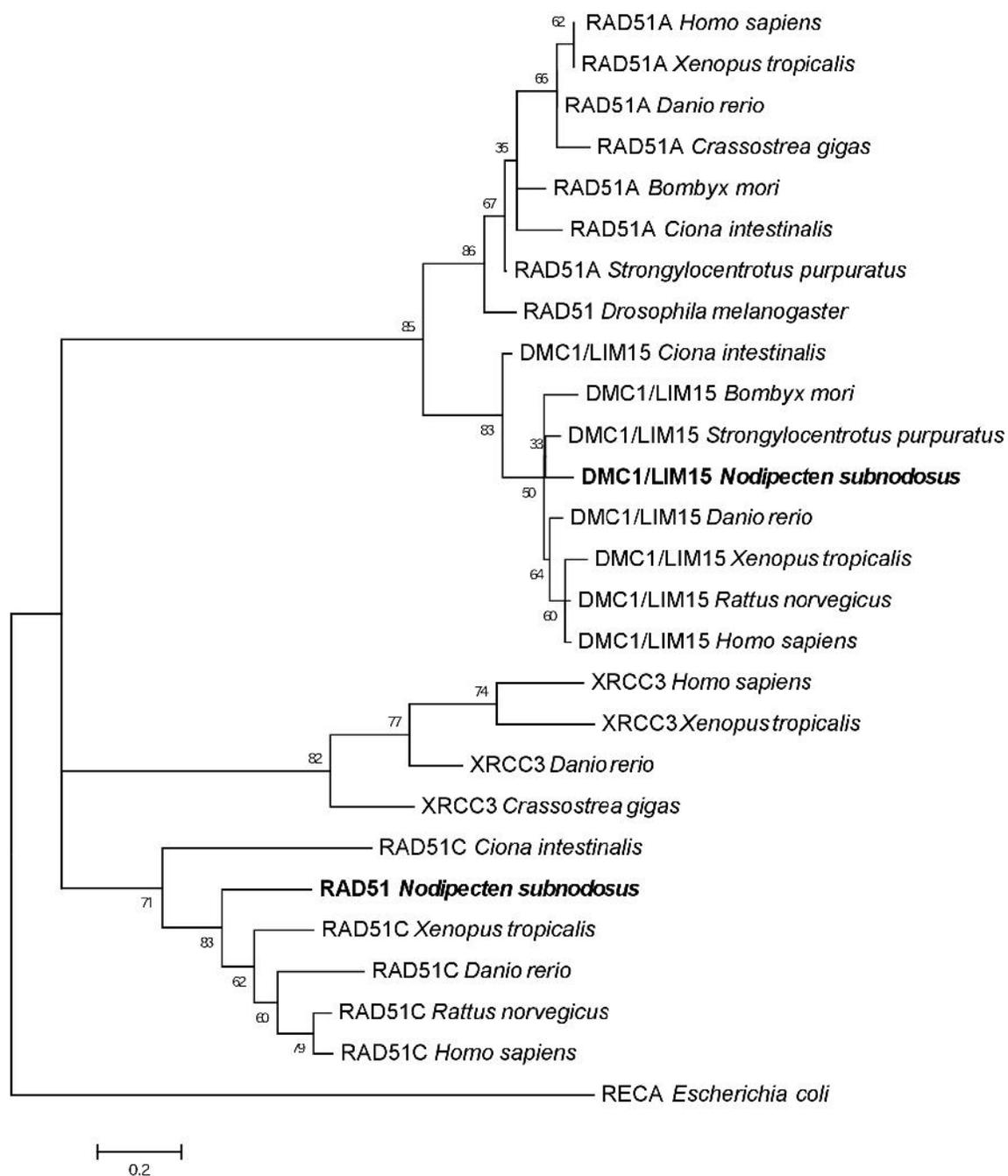


Fig. 15. Árbol filogenético de secuencias de proteínas RAD51A, RAD51C, DMC1 y XRCC3, de especies de invertebrados y vertebrados, incluyendo las caracterizadas en el presente estudio para *N. subnodosus*. La secuencia de *recA* de *Escherichia coli* fue empleada como grupo externo. Los valores de los nodos representan el soporte de bootstrap del clado en porcentaje.

Como soporte adicional de asignación a subfamilias de RAD51, se realizó una comparación estructural del modelo proteico de la secuencia en aminoácidos del transcrito obtenido en *N. subnodosus* contra ortólogos de RAD51A (acceso GenBank NP\_001016393.1) y RAD51C (acceso GenBank NP\_001016923.1) de *Xenopus tropicalis*, mediante generación de modelos estructurales en Swiss-Model (Arnold *et al.*, 2006), y un posterior análisis de comparación estructural por superposición: *Flexible structure AlignmentT by Chaining AFPs (Aligned Fragment Pairs) with Twists* - FATCAT - (Ye y Godzik, 2004). De éste análisis comparativo, se infiere que la secuencia obtenida en la almeja es más similar a una RAD51C, a partir de que se obtuvieron scores de alineamientos estructurales de 666.4 y 725.2 contra las secuencias de RAD51A y RAD51C de *X. tropicalis*, respectivamente, ambos significativos ( $1.9^{-13}$  y  $4.0^{-13}$ , respectivamente) (Fig. 16).

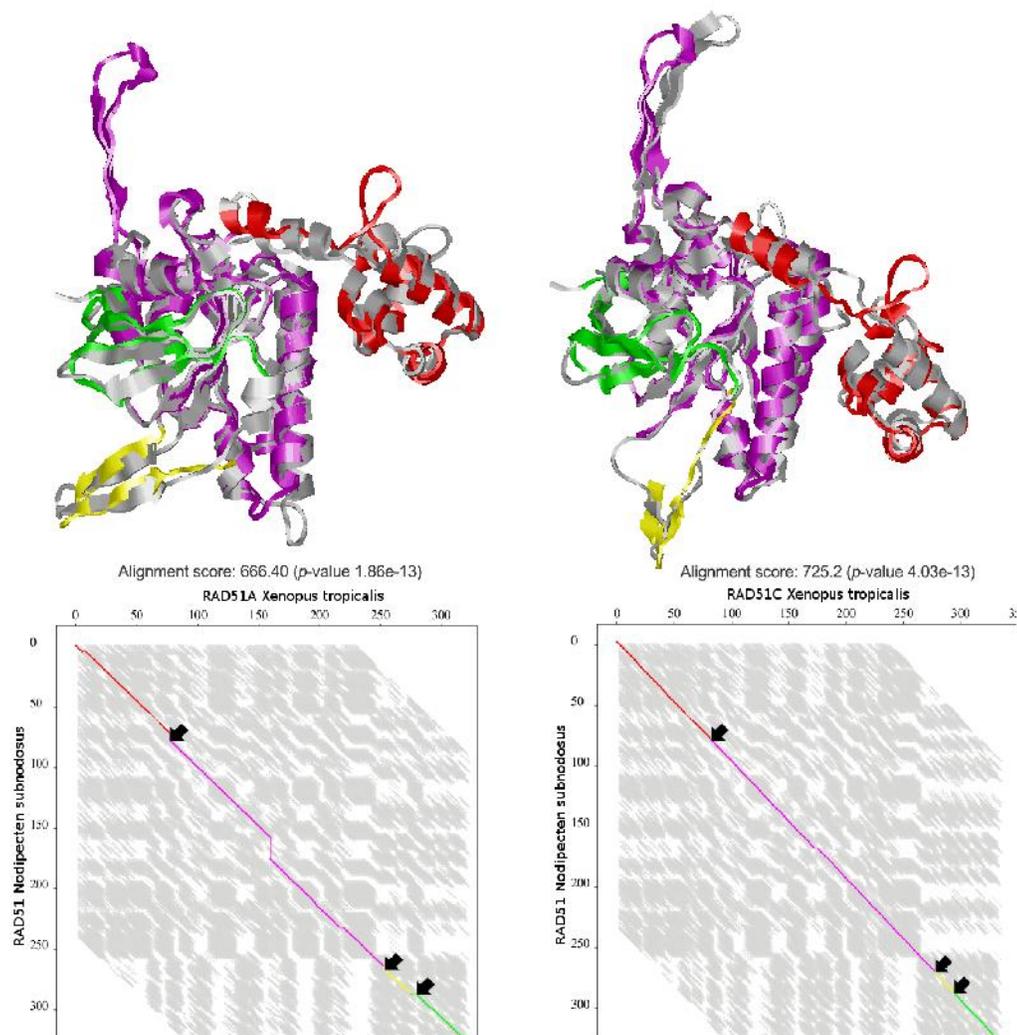


Fig. 16. Alineamientos estructurales del modelo proteico de RAD51 en *N. subnodosus* contra modelos de RAD51A y RAD51C de *Xenopus tropicalis*.

### 7.2.3 Caracterización de transcrito de *pch2*

Mediante la amplificación por RACE del homólogo *pch2* en *N. subnodosus* se obtuvo un transcrito de 1861 pb (acceso GenBank KC904225), con una secuencia codificante con similitud significativa a proteínas *pch2* en invertebrados y con homólogos en vertebrados nombrados como *thyroid receptor interacting protein 13*. Los mejores hits contra proteínas de bases de datos públicas se encontraron contra diversas secuencias de homólogos de PCH2/TRIP13 de invertebrados y vertebrados (valores de  $E = 0$ ). Cuando se realizó una búsqueda por blastx en las proteínas anotadas del genoma del ostión *C. gigas*, se obtuvo un hit significativo contra *Nuclear valosin-containing-like protein* (score = 71.2, valor  $E = 6^{-13}$ ). Un árbol filogenético a partir de un bloque conservado de secuencias de ortólogos de invertebrados y vertebrados (PCH2; TRIP13) (Fig. 17) permite ubicar al homólogo identificado en *N. subnodosus* entre secuencias homólogas de invertebrados, las cuales se disponen en un clado basal en comparación al clado de homólogos PCH2/TRIP13 propios de vertebrados (Fig. 18)

### 7.2.4 Caracterización de transcrito de *scp3*.

El transcrito correspondiente a *scp3* aislado mediante RACE presentó 984 pb, y se detectó un marco de lectura de 241 aminoácidos, los cuales presentan una similitud de 69% con la proteína anotada como *synaptonemal complex protein 3* en *C. gigas*, con un score = 297 y un valor  $E = 2^{-98}$ . Dentro de la secuencia, se identificó un dominio conservado Cor1/Xlr/Xmr (pFAM04803), con una similitud de 1.81<sup>-30</sup>; la secuencia del transcrito fue depositada en GenBank-NCBI bajo el acceso KC119411.

## MUSCLE (3.8) multiple sequence alignment

```

Toxoplasma_gondii_ME49{237837989      SSQLEINAHSLFSRWFSSESGKLVLMKMTTIKDLLEDESCFVCLIDEVESLSTARRAAM
Trichoplax_adhaerens{195996749}      YQQLIEINSHSLFSKWFSESGKLVKMFQKIQELVDDSDAL-----VESLTAARSAAL
Oreochromis_niloticus{348534013}     YGQFVEINSHSLFSKWFSESGKLVTKMFQKIQQLIIDDKDALVFVLIDEVESLTAARNACQ
Danio_riero{41053680}                 HSQFVEINSHSLFSKWFSESGKLVTKMFQKIQELIIDDKDALVFVLIDEVESLTAARSAAQ
Saccoglossus_kowalevskii{2912434}    YGQLIEINSHSLFSKWFSESGKLVKMFQKIQEFTIDQDALICVLIDEVESLTAARKSSM
Amphimedon_queenslandica{3403677}    YGQQLVEINSHSLFSKWFSESGKLVQKMFSKIHTLVEDDNALICILIDEVESLTVARNAAM
Branchiostoma_floridae{260819230}    YGQFIEINSHSLFSKWFSESGKLVLMKMFQKIQELIIDSDALVCLIDEVESLTAARKASL
Nodipecten_subnodosus                 YGQLIEINSHSLFSKWFSESGKLVKMFQKIQELIIDKKEALVFVLIDEVESLTAARKSAM
Nematostella_vectensis{156353324}    YGQLIEINSHSLFSKWFSESGKLVKMFQKIQELIIDKKEALVCLIDEVESLTAARKAAM
Strongylocentrotus_purpuratus{72}     YGQQLVEINSHSLFSRWFSSESGKLVKMFQKIQELVSDPDLSLVCILIDEVESLTAARKSAL
Strongylocentrotus_purpuratus{11}    YGQQLVEINSHSLFSRWFSSESGKLVKMFQKIQELVSDPDLSLVCILIDEVESLTAARKSAL
Xenopus_tropicalis{45361301}         YGQQLVEINSHSLFSKWFSESGKLVTKMFQKIHELINDKEALVFVLIDEVESLTAARKASR
Xenopus_laevis{148229431}            YGQQLVEINSHSLFSKWFSESGKLVTKMFQKIHDLINDKEALVFVLIDEVESLTAARKASR
Ornithorhynchus_anatinus{3453265}    YGQLIEINSHSLFSKWFSESGKLVTKMFQKIQELIIDKKNALVFVLIDEVESLTAARNACR
Monodelphis_domestica{126320866}     YGQLIEINSHSLFSKWFSESGKLVTKMFQKIQELIIDKKEALVFVLIDEVESLTAARNACR
Cricetulus_griseus{354506641}        YGQLIEINSHSLFSKWFSESGKLVTRMFQKIQDLIIDKKEALVFVLIDEVESLTAARNACR
Oryctolagus_cuniculus{291390587}     YGQLIEINSHSLFSKWFSESGKLVTRMLQKIQDLIIDKKEALVFVLIDEVESLTAARSACR
Oryctolagus_cuniculus{291413971}     YGQLIEINSHSLFSKWFSESGKLVTRMFQKIQDLIIDKKEALVFVLIDEVESLTAARSACR
Loxodonta_africana{344308311}        YGQLIEINSHSLFSKWFSESGKLVTKMFQKIQDFIIDKDALVFVLIDEVESLTAARNACR
Equus_caballus{338718796}            YGQLIEINSHSLFSKWFSESGKLVTKMFQKIQDLIIDKDALVFVLIDEVESLTAARNACR
Mus_musculus{110625724}              YGQLIEINSHSLFSKWFSESGKLVTKMFQKIQDLIIDKKEALVFVLIDEVESLTAARNACR
Rattus_norvegicus{58865438}          YGQLIEINSHSLFSKWFSESGKLVTKMFQKIQDLIIDKKEALVFVLIDEVESLTAARNACR
Cavia_porcellus{348552660}           YGQLIEINSHSLFSKWFSESGKLVTRMFQKIQDLIIDKKEALVFVLIDEVESLTAARNACR
Ailuropoda_melanoleuca{301782535}    YGQQLVEINSHSLFSKWFSESGKLVTKMFQKIQDLIIDKDALVFVLIDEVESLTAARNACR
Callithrix_jacchus{296194996}        YGQLIEINSHSLFSKWFSESGKLVNMFQKIQDLIIDKDALVFVLIDEVESLTAARNACR
Pan_troglodytes{332864074}           YGQLIEINSHSLFSTWFSSESGKLVTKMFQKIQDLIIDKDALVFVLIDEVESLTAARNACR
Nomascus_leucogenys{332228131}       YGQLIEINSHSLFSKWFSESGKLVTKMFQKIQDLIIDKDALVFVLIDEVESLTAARNACR
Pongo_abelii{297674865}              YGQLIEINSHSLFSKWFSESGKLVTKMFQKIQDLIIDKDALVFVLIDEVESLTAARNACR
Macaca_mulatta{109076630}            YGQLIEINSHSLFSKWFSESGKLVTKMFQKIQDLIIDKDALVFVLIDEVESLTAARNACR
Pan_troglodytes{332820827}           YGQLIEINSHSLFSKWFSESGKLVTKMFQKIQDLIIDKDALVFVLIDEVESLTAARNACR
Homo_sapiens{11321607}               YGQLIEINSHSLFSKWFSESGKLVTKMFQKIQDLIIDKDALVFVLIDEVESLTAARNACR
Homo_sapiens{261337161}              YGQLIEINSHSLFSKWFSESGKLVTKMFQKIQDLIIDKDALVFVLIDEVESLTAARNACR
Canis_lupus_familiaris{74003129}    YGQLIEINSHSLFSKWFSESGKLVTKMFQKIQDLIIDKDALVFVLIDEVESLTAARNACR
Sus_scrofa{305855152}                YGQLIEINSHSLFSKWFSESGKLVTRMFQMIQDLIIDKDALVFVLIDEVESLTAARNACR
Bos_taurus{297468101}                YGQQLVEINSHSLFSKWFSESGKLVTRMFQKIQDLIIDKDALVFVLIDEVESLTAARNACR
Anolis_carolinensis{327275055}       YGQLIEINSHSLFSKWFSESGKLVTKMFQKIQALIIDKDALVFVLIDEVESLTAARSAVR
Taeniopygia_guttata{224045739}       YGQLIEINSHSLFSKWFSESGKLVTKMFQKIQELVDDPDALVFVLIDEVESLTAARSAFR
Gallus_gallus{118086321}             YGQLIEINSHSLFSKWFSESGKLVTKMFQKIQELIIDKDALVFVLIDEVESLTAARSAFK
Meleagris_gallopavo{326916921}      YGQLIEINSHSLFSKWFSESGKLVTKMFQKIQELIIDRDALVFVLIDEVESLTAARSAFK
.*:***:***** ***** .*:  *:  :::  ...      ****:  .* :

```

Fig. 17. Alineamiento múltiple de las secuencias de ortólogos de PCH2/TRIP13 de especies de invertebrados y vertebrados, incluyendo la caracterizada en el presente estudio para *N. subnodosus*. Se muestra la región conservada del alineamiento correspondiente al rango de posición 209 a 271 de la secuencia de *N. subnodosus*.

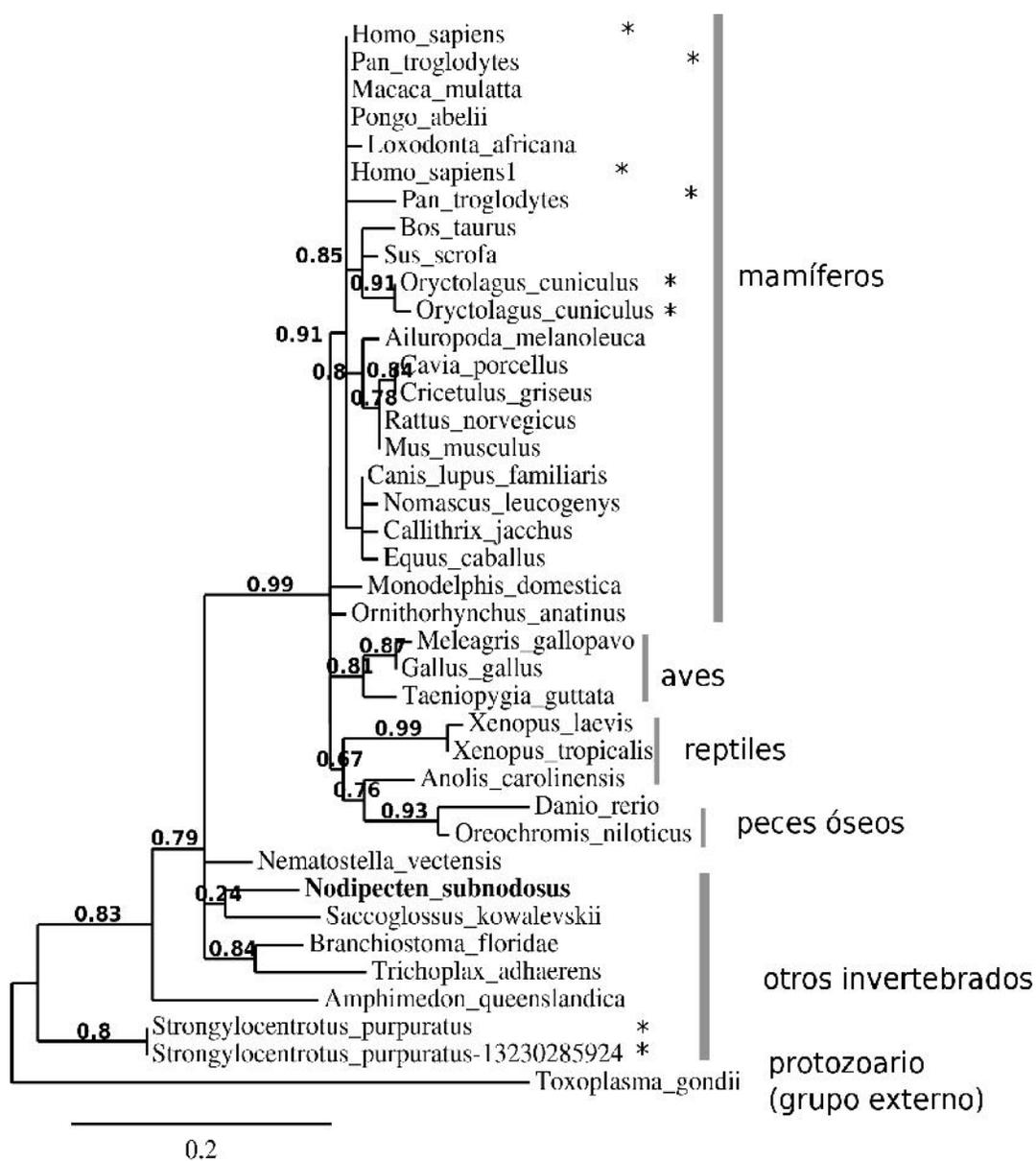


Fig. 18. Árbol filogenético de secuencias de proteínas PCH2 de especies de invertebrados y TRIP13 de vertebrados, incluyendo la secuencia caracterizada en el presente estudio para *N subnodosus*. Los valores de los nodos representan el soporte de bootstrap del clado en proporción de 0 a 1.

### 7.3 Definición de estabilidad de genes de referencia para la normalización de la expresión génica

#### 7.3.1 Secuencias parciales de genes de referencia.

Mediante la estrategia a clonación parcial dirigida de genes por PCR degenerativa (deg-PCR), se logró la obtención de secuencias parciales de transcritos para ser utilizados como potenciales genes de referencia. Un fragmento de ARN ribosomal *18S* de 168 pb fue obtenido empleando oligonucleótidos específicos en regiones conservadas, y mediante alineamientos múltiples contra fragmentos *18S* de otros pectínidos [*Pecten jacobaeus* (AY070112.1), *Pecten maximus* (L49053.1), *Chlamys farreri* (AF526252.1) y *Mimachlamys varia* (DQ279939.1)] se identificó una similitud del 100% entre las posiciones 913-1030 de las secuencias completas del *18S-rRNA* de secuencias relacionadas. Por otra parte, el fragmento de *act-* obtenido fue de 206 pb, y presentó una similitud nucleotídica alrededor al 90% con transcritos de *act-* de otras especies de bivalvos [*Chlamys farreri* (AY335441.2), *Mizuhopecten yessoensis* (DQ7878P), *Placopecten magellanicus* (U55046.1) y *C. gigas* (AB071191.1)] (anexo 1). Del fragmento obtenido para *act-*, se diseñaron oligonucleótidos específicos para su cuantificación por qPCR. Las claves de acceso GenBank de las secuencias depositadas, así como los oligonucleótidos diseñados para amplificación por qPCR de los genes de referencia se presentan en la Tabla IV.

Adicionalmente, se obtuvieron secuencias parciales para los genes constitutivos *factor de elongación 1-alfa* (*efl-*; acceso GenBank JN034909.1), *proteína ribosomal L8* (*rpl8*; acceso GenBank JN034908.1), *tubulina alfa* (*tub-*; acceso GenBank JN034910.1), así como de genes de interés en el contexto gametogénico para validar los análisis de estabilidad de los genes de referencia: *vitelogenina* (*vit*; acceso GenBank JN034911.1) y una helicasa de la familia *DEAD-box* (*ddx*; acceso GenBank JN034912.1). Los mejores hits contra proteínas de referencia de GenBank mediante alineamientos por blastx se muestran en la Tabla V). Un fragmento de 781 pb del transcrito *efl-* se obtuvo por amplificación PCR empleando oligonucleótidos degenerados, y fue confirmado como el transcrito esperado cuando el resultado de Blastx mostró un hit de 83% con *Placopecten magellanicus*.

Tabla IV. Secuencias de primers para la amplificación por qPCR de genes de referencia y genes de interés. Las temperaturas de disociación ( $T^{\circ}m$ ) fueron determinadas con el calculador de oligos de SIGMA (<http://www.sigma-genosys.com/calc/DNACalc.asp>).

Gen	acceso GenBank	orientación	secuencia de primers para qPCR	$T^{\circ}m$	eficiencia (0-1)
18S	n.a.	sentido	5'-GAAATTCCTGGATCGCCGTA-3'	63.7	93%
		antisentido	5'-GCCGAGTCATTGAAGCAACT-3'	64.3	
ef-1	JN034909	sentido	5'-GTGCCAGTGGGTAGGGTAGA-3'	63.9	82%
		antisentido	5'-CTCCAGCAACGTTTCCTCTC-3'	63.9	
act-	n.a.	sentido	5'-CAGAGCAAGAGAGGTATCCTCACC-3'	66.2	93%
		antisentido	5'-GTTGAAGGTCTCGAACATGATCTG-3'	65.8	
rp-18	JN034908	sentido	5'-CGTCATGGATACATCAAGGGT-3'	63.5	96%
		antisentido	5'-CAAACAGTCCAGTGATACATGCC-3'	64.3	
tub-	JN034910	sentido	5'-TTCAGCCTGATGGACAAATG-3'	63.6	94%
		antisentido	5'-TCTGGATGAAACAGCTGACG-3'	64.1	
ddx	JN034911	sentido	5'-CCAATCTTGCCCTCGTTCAAT-3'	63.8	93%
		antisentido	5'-GCTTTAATCCTGGCCCCTAC-3'	63.4	
Vit	JN034912	sentido	5'-CTGGCCAAAGCGAGTTCTAC-3'	63.7	88%
		antisentido	5'-TCGTTGAACATGGTGTTCGT-3'	63.6	

n.a. indica que la secuencia no fue sometida a la base de datos pública de GenBank, pero se presenta en la sección de Anexos II

Tabla V. Resultados obtenidos en los alineamientos por blastx de la traducción conceptual de las secuencias parciales obtenidas por secuenciación dirigida contra bases de datos de proteínas de referencia (*refseq*; NCBI-GenBank).

secuencias parciales en <i>N. subnodosus</i> (acceso GenBank)	descripción de mejor hit	especie	Score	valor E	similitud
rp-18 JN034908	putative ribosomal protein L (AAN05596)	<i>Eurythoe complanata</i>	317	1 <sup>-114</sup>	99 %
ef-1 JN034909	elongation factor 1 alpha (AFS50747.1)	<i>Placopecten magellanicus</i>	452	5 <sup>-158</sup>	89 %
tub- JN034910	alpha tubulin (XP_002580034.1)	<i>Schistosoma mansoni</i>	222	1 <sup>-56</sup>	100 %
ddx JN034911	pl10 (ABE27760.1)	<i>Azumapecten farreri</i>	210	1 <sup>-74</sup>	97 %
vit JN034912	vitellogenin (CAQ06469.1)	<i>Pecten maximus</i>	193	5 <sup>-59</sup>	99 %

La secuencia obtenida para *rp-l8* fue confirmada mediante la misma herramienta, presentando una identidad de 99% con la secuencia del bivalvo *Eurythoe complanata*. La traducción conceptual de *vit* presentó una identidad significativa con la secuencia de vitelogenina de *Pecten maximus* (92% de identidad; valor E de  $1^{-67}$ ), incluyendo un dominio conservado, vitellogenin\_N (PF01347.16), identificado en la base de datos Pfam (Finn *et al.*, 2010).

### 7.3.2 Estabilidad de genes de referencia

#### 7.3.2.1 Análisis de abundancia relativa de los genes de referencia en función al métodos de síntesis de ADNc (oligo-dT vs. hexanucleótidos aleatorios)

La estadística descriptiva contrastando la expresión relativa de todos los genes de referencia en muestras de ADNc sintetizadas con oligo-dT y hexámeros aleatorios se muestran en la tabla VI. Los Cts promedios de todos los genes traducidos (*rp-l8*, *ef-1*, *act-*, *tub-*), fue siempre menor cuando el ADNc es sintetizado con hexanucleótidos aleatorios en comparación al ADNc sintetizado con oligo-dT. Cuando la diferencia de Cts fueron transformadas a abundancia relativa (RQ; sección de métodos) ponderado por la eficiencia de amplificación, se denota que para éstos transcritos derivados de ARNm se observan relaciones de sobreabundancia que van de 1.7 a 3.1 veces en ADNc sintetizado con hexanucleótidos aleatorios en contraste al sintetizado con oligo-dT. Para el fragmento *18S*, ésta relación es de 12.9 veces (Tabla VI).

Cuando los genes transcritos en reverso a ADNc mediante ambos métodos (oligo-dT y hexanucleótidos aleatorios) y analizados por qPCR fueron agrupados y analizados por geNorm, el gen *18S* fue denotado como el más estable independientemente del método de priming empleado en las muestras en el cual fue analizado (Fig. 19a). Los valores de estabilidad *M* estimados para *18* (como una estabilidad pareada entre las muestras obtenidas mediante síntesis de ADNc con oligo-dT y hexanucleótidos aleatorios) fueron las menores (menor valor = mayor estabilidad) ( $M = 0.68$ ). *18S* fue denotado como el segundo gen más estable, resultando en  $M = 0.86$  y  $0.96$  cuando las muestras fueron sintetizadas con hexanuucleótidos aleatorios y oligo-dT, respectivamente. De una forma similar,

NormFinder indicó que los genes más estables fueron *18S* y *rp-18* (Fig. 19b).

Tabla VI. Estadísticas descriptivas de los valores de amplificación del set de genes de referencia sobre ADNc sintetizado con oligo-dT y hexanucleótidos aleatorios.

Gen	<i>18S</i>		<i>rp-18</i>		<i>act-</i>		<i>ef-1</i>		<i>tub-</i>	
	oligo-dT	hexa	oligo-dT	hexa	oligo-dT	hexa	oligo-dT	hexa	oligo-dT	hexa
Tipo de ADNc										
n	28		28		28		28		28	
media geom [Ct]	19.9	16.0	27.0	26.3	27.9	27.3	26.3	25.4	29.0	27.3
min [Ct]	18.4	14.8	25.0	24.3	23.5	22.9	23.9	23.2	21.6	20.6
max [Ct]	23.8	18.0	30.8	30.2	34.6	38.7	34.0	31.4	36.1	32.5
desv. est. [ $\pm$ Ct]	0.9	0.7	1.2	1.4	2.8	2.7	2.2	1.7	3.5	2.9
CV [% Ct]	4.5	4.2	4.4	5.3	9.8	9.8	8.3	6.8	11.9	10.6
$C_t$	3.9		0.8		0.6		0.9		1.7	
relación de abundancia relativa (oligo-dT / hexa)	-12.9		-1.7		-1.5		-1.9		-3.1	

Mediante el uso de geNorm, *18S* y *rp-18* fueron los genes más estables en las muestras de ADNc sintetizadas con oligo-dT; en contraste, para las muestras de ADNc sintetizadas con hexanucleótidos aleatorios, *rpl8* y *ef-1a* fueron los genes más estables, seguidas por *18S* (Fig. 20). El gen más inestable detectado por geNorm fue *tub-a*. De hecho, la desviación estándar y coeficiente de variación (CF) fue menor para *18S* (CV de 4.5 y 4.2) y *rp-18* (CV 4.4 y 5.3), que para el resto de los genes de referencia potenciales. El gen con el mayor CV fue *tub-a* (CV de 11.9 y 10.6) (Tabla VI).

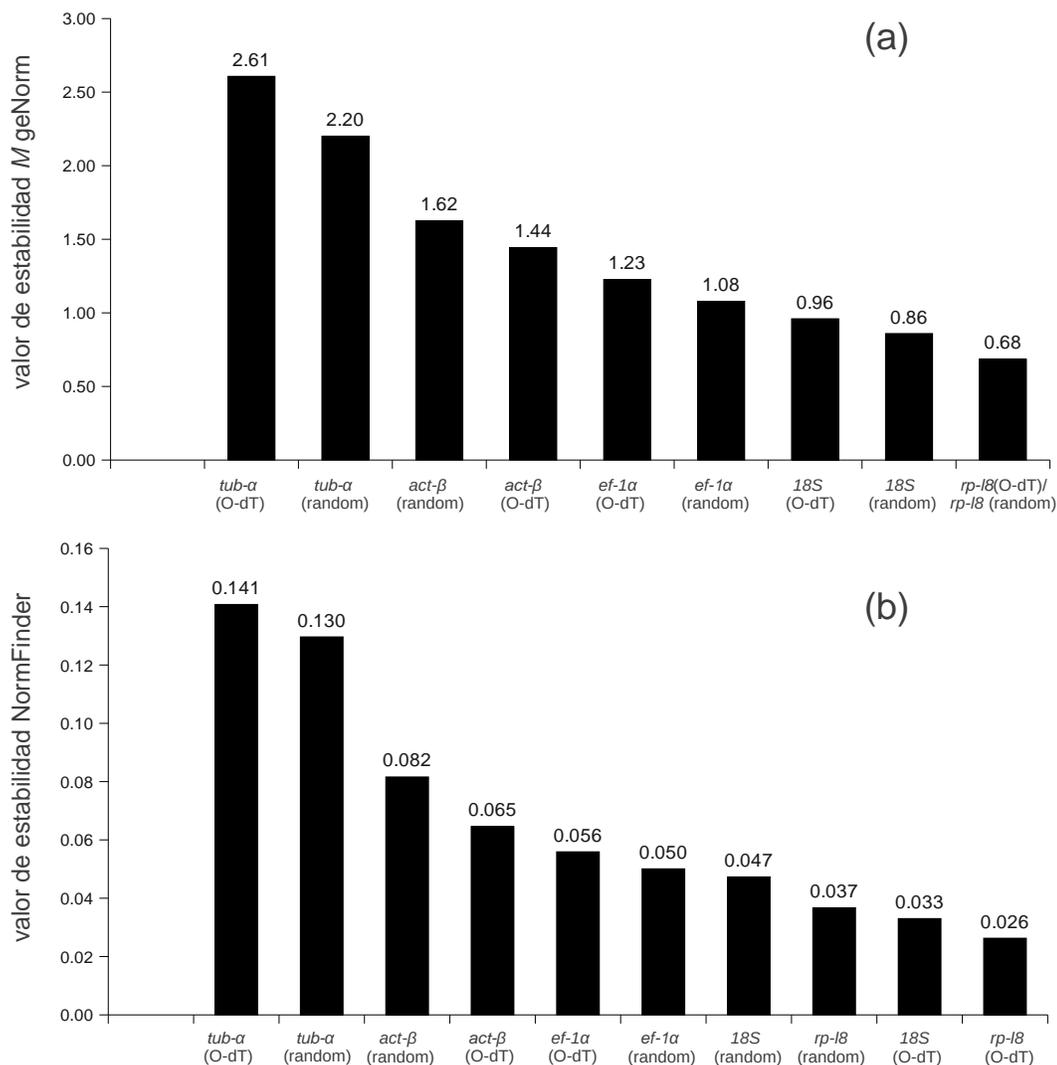


Fig. 19. Estabilidad de genes de referencia entre regiones gonadales y estadios ( $n = 30$ ). se muestran los resultados por geNorm (a) y NormFinder (b) para cada gen y método de síntesis de ADNc en función al primer empleado para la transcripción reversa: oligo-dT, hexanucleótidos aleatorios (random).

Los valores de estabilidad obtenidos por NormFinder para el mismo set de genes de referencia también fueron registrados y analizados de forma separada para las muestras de ADNc sintetizadas con oligo-dT y hexanucleótidos aleatorios (Fig. 21). Tanto *18S* como *rp-18* fueron encontrados como el par de genes más estables en ambos sets de muestras de ADNc (valor de estabilidad de 0.032). Sin embargo, el gen más estable en ADNc/oligo-dT

fue *18S* (valor de estabilidad de 0.031), y en ADNc/hexanucleótidos aleatorios fue *rp-18* (valor de estabilidad de 0.037). El gen menos estable, sin considerar el primer de síntesis de ADNc, fue de nuevo *tub-a*, con valores de estabilidad de 0.133 y 0.138 en ADNc sintetizado con oligo-dT y hexanucleótidos aleatorios, respectivamente.

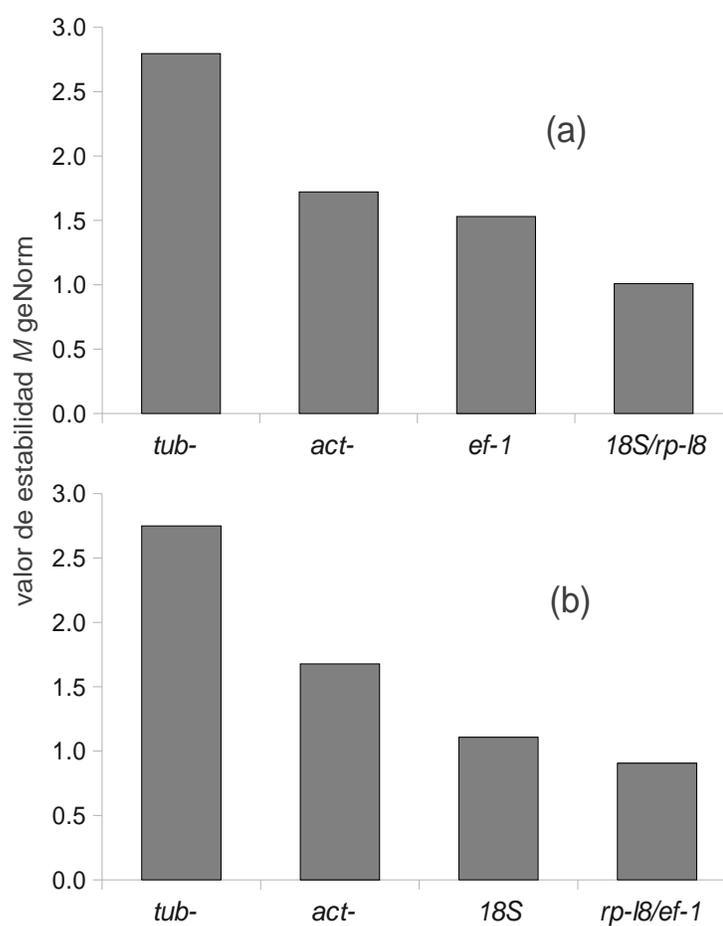


Fig. 20. Valores de estabilidad para los genes de referencia obtenidos por geNorm en cada set de ADNc sintetizado con oligo-dT (a) y hexanucleótidos aleatorios (b).

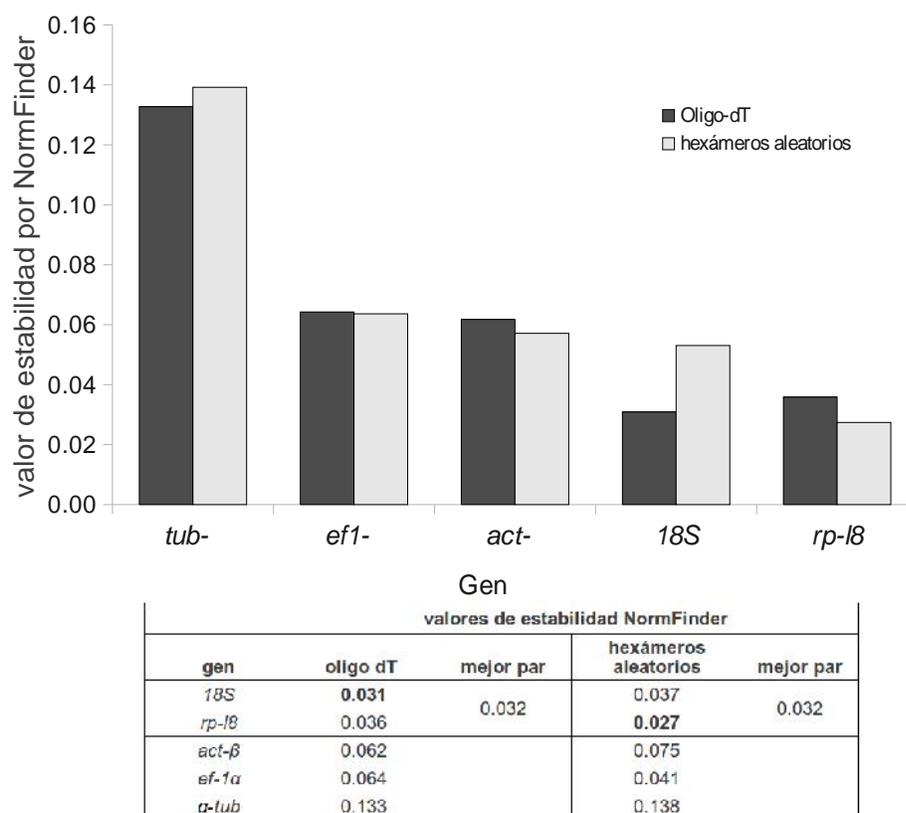


Fig. 21. Valores de estabilidad para los genes de referencia obtenidos por NormFinder en cada set de ADNc sintetizado con oligo-dT (a) y hexanucleótidos aleatorios (b).

La fuente de variación biológica estimada para la expresión de todos los genes de referencia, basada en la magnitud de la variación entre cada condición gonadal analizada, y por cada método de síntesis de ADNc, se presenta en la figura 22. Se puede notar que *tub-a* muestra el mayor nivel de variación entre-grupos estimada en Ln (*Ct*) (menor *Ct* = mayor expresión), indicando que la expresión de este gen se encuentra sesgada hacia la región testicular. Para los otros genes, incluyendo *18S*, la tendencia de la variación es positiva hacia la región testicular, pero con una variación notablemente menor que la registrada para *tub-a*. Cuando se comparan los patrones de variación entre métodos de priming de ADNc, la variación entre-grupos es relativamente similar.

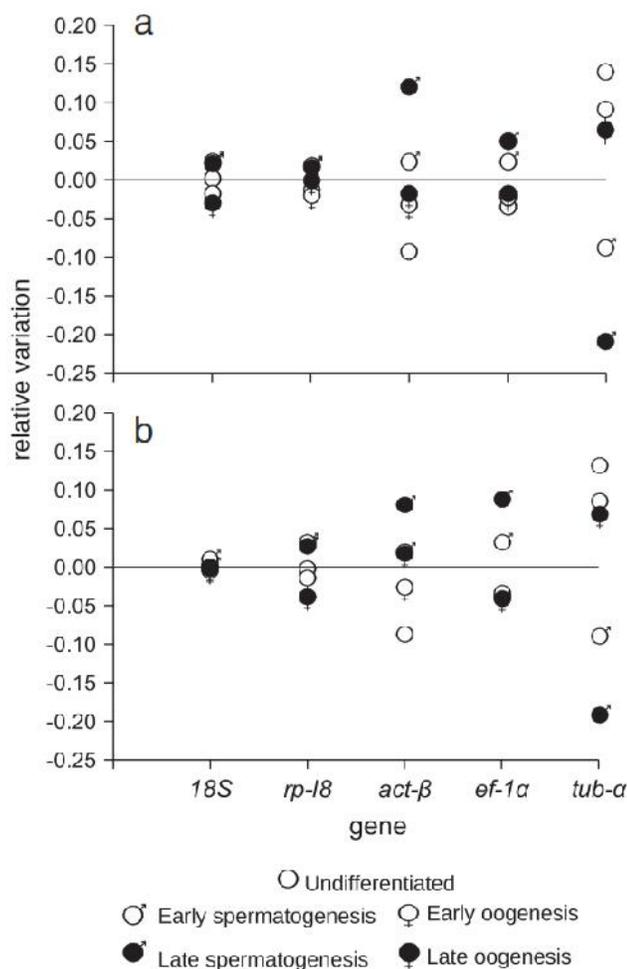


Fig. 22. Variación relativa de los valores de amplificación ( $C_t$ ) de los genes de referencia entre muestras de gónada inactiva, y en testículo y ovario en estadios de desarrollo temprano y avanzado en ADNc preparado con oligo-dT (a) o hexanucleótidos aleatorios (b).

### 7.3.3 Análisis de expresión de genes *ddx*, *vit* y *tub-a* durante la validación de genes de referencia.

Eligiendo los dos genes más estables para normalización en tejido gonadal (*rp-18* y *18S*), se analizó la expresión relativa de tres genes blanco. La expresión de los genes *tub-a* y *vit* mostraron una clara asociación hacia la región ovárica (Fig. 23). *tub-a* mostró una sobreexpresión de 50 y 330 veces en los estadios temprano y avanzado de maduración testicular en comparación a los estadios correspondientes de la región ovárica, y de 130 y 600 veces a la expresión detectada en la gónada inactiva. Aunque la expresión de *tub-a* fue

aparentemente mayor (4.6 veces) en el estadio de madurez tardía vs el estadio de maduración temprana, ésta diferencia no fue estadísticamente significativa. Un patrón similar se encontró en la parte ovárica para *vit*, que mostró una expresión incrementada desde el estadio de ovogénesis temprana hacia la avanzada (tardía), registrando niveles de expresión de ~1000 veces en el estadio de ovogénesis avanzada en comparación a la gónada inactiva.

El efecto de utilizar genes de referencia elegidos al azar para la cuantificación de la expresión relativa fue analizado usando a *ddx* como gen de interés. Se usaron uno, dos, o todos los potenciales genes de referencia, pero no *tub-a* para normalizar en esta comparación. Como se muestra en la figura 23, no se encontraron diferencias significativas en la expresión relativa entre regiones sexuales o estadios de desarrollo de las gónadas cuando se usaron los dos genes más estables para éste caso (*18S + rp-18*), ni se encontraron diferencias cuando se usó a cada uno de estos genes de referencia por separado (Fig. 24). Sin embargo, cuando se normalizó con *act-*, tanto la región testicular como ovárica en los estadios de maduración temprana y avanzada muestran una aparente sobreexpresión de *ddx* cuando se contrasta con la gónada indiferenciada. Cuando se emplea únicamente a *ef1-* como gen de normalización, observamos una aparente sobreexpresión de *ddx* en la parte testicular en madurez avanzada. Cuando se utilizó la media geométrica de todos los genes de referencia (*rp-18+18S+ef1- +act-*), la interpretación de las diferencias de expresión es como a continuación: gónada indiferenciada  $\leq$  gónada femenina (sin importar el estadio) = región testicular en estadio temprano de desarrollo < región testicular en estadio avanzado de desarrollo (Fig. 24).

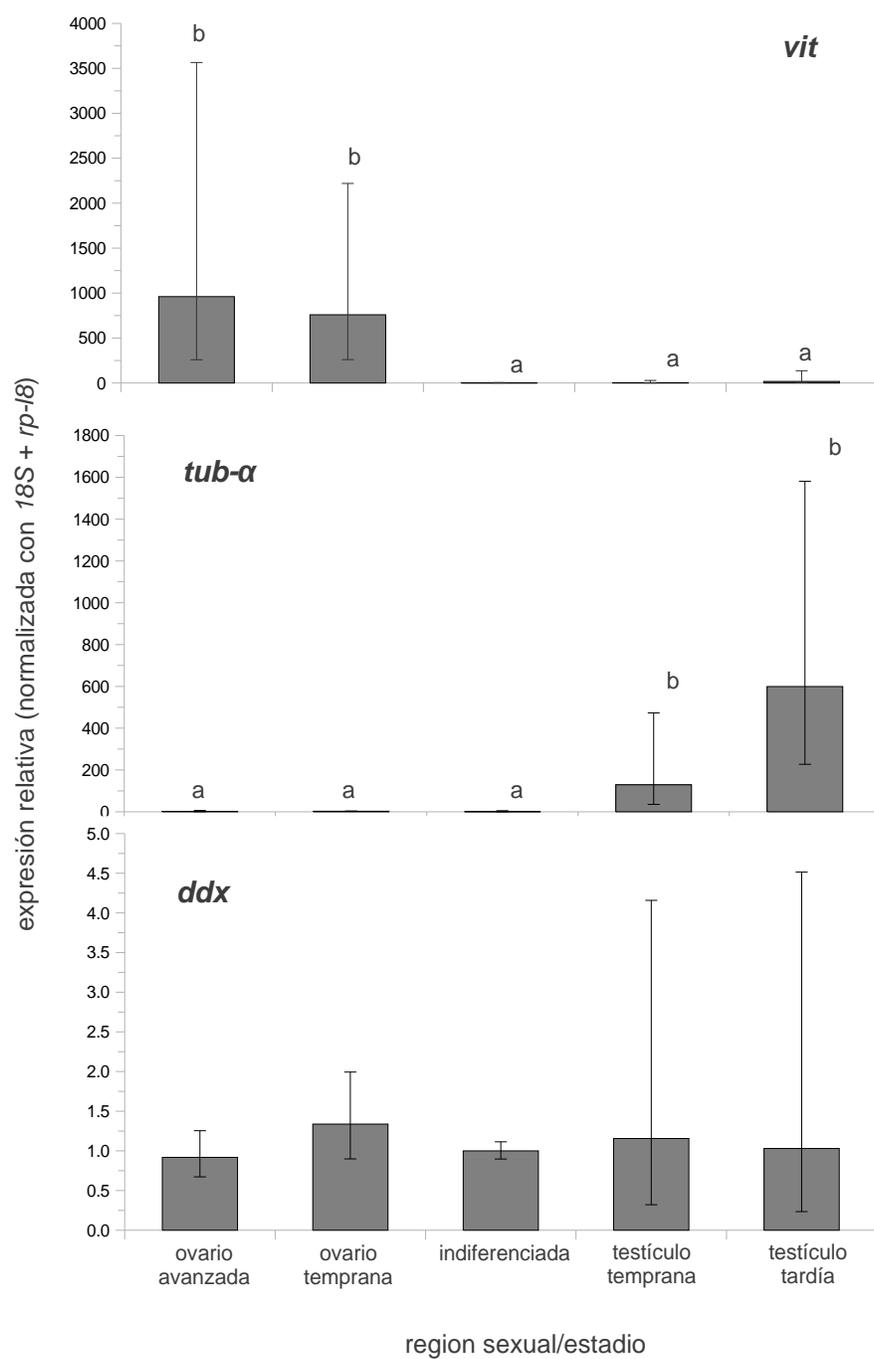


Fig. 23. Expresión relativa de *vit*, *tub-a* y *ddx* en ambas regiones sexuales de la gónada, normalizada con los genes 18S y rp-18 (n=6 para cada sexo/estadio gametogénico). Diferentes letras indican diferencias significativas ( $p < 0.05$ ); y las barras de dispersión representan el intervalo de confianza al 95%.

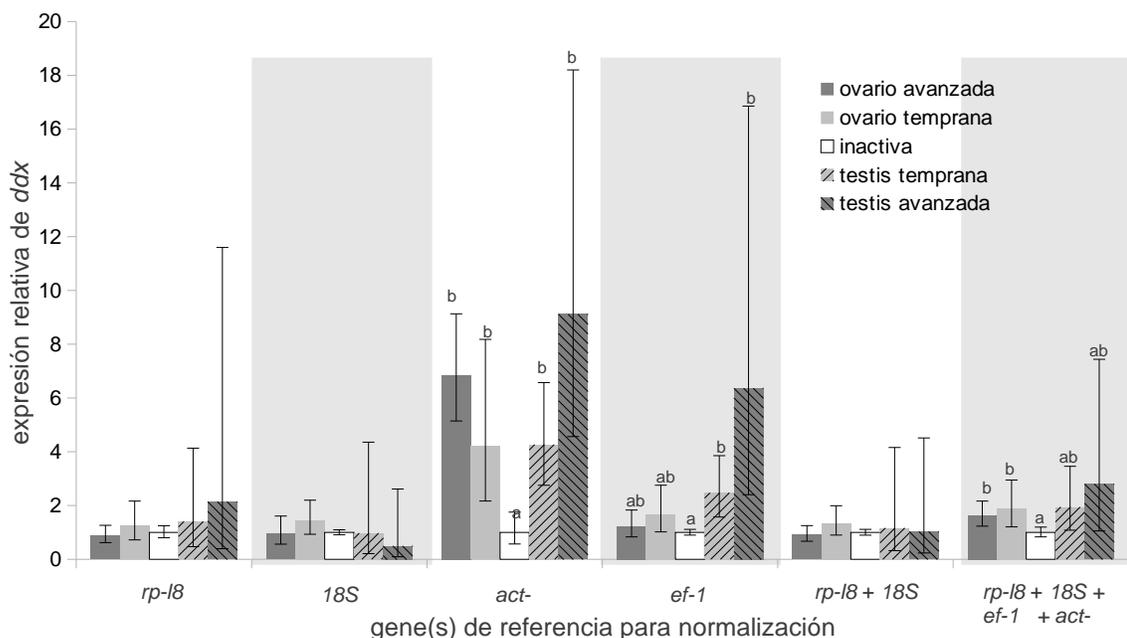


Fig. 24. Expresión relativa de *ddx* estimada después de normalizar con genes individuales (*18S*, *rp-18*, *ef-1*, o *act-*), con la combinación de los dos más estables (*18S/rp-18*), o con la combinación de los cuatro genes de referencia ( $n = 6$  para cada sexo/estadio de gametogénesis). Diferentes letras indican diferencias significativas ( $p < 0.05$ ); y las barras de dispersión representan el intervalo de confianza al 95%.

#### 7.4 Ubicuidad de expresión de *tub-* en distintos tejidos.

Ambos análisis de estabilidad, empleando geNorm y NormFinder, indicaron que los dos genes más estables para la normalización de expresión entre distintos órganos/tejidos fueron *ef-1* y *rp-18*, por lo que fueron entonces empleados para la normalización de la expresión de *tub-* en distintos tejidos. Éste gen mostró una elevada expresión en testículo, y en menor grado, en músculo aductor; el resto de los tejidos analizados mostraron un nivel de expresión no detectable o despreciable. La posibilidad de que se estuvieran obteniendo distintos amplicones en los dos tejidos analizados fue descartada por la secuenciación de los productos obtenidos y por el análisis de disociación de los amplicones en el paso final de la qPCR; ambos resultados confirmaron amplicones idénticos (Fig. 25).

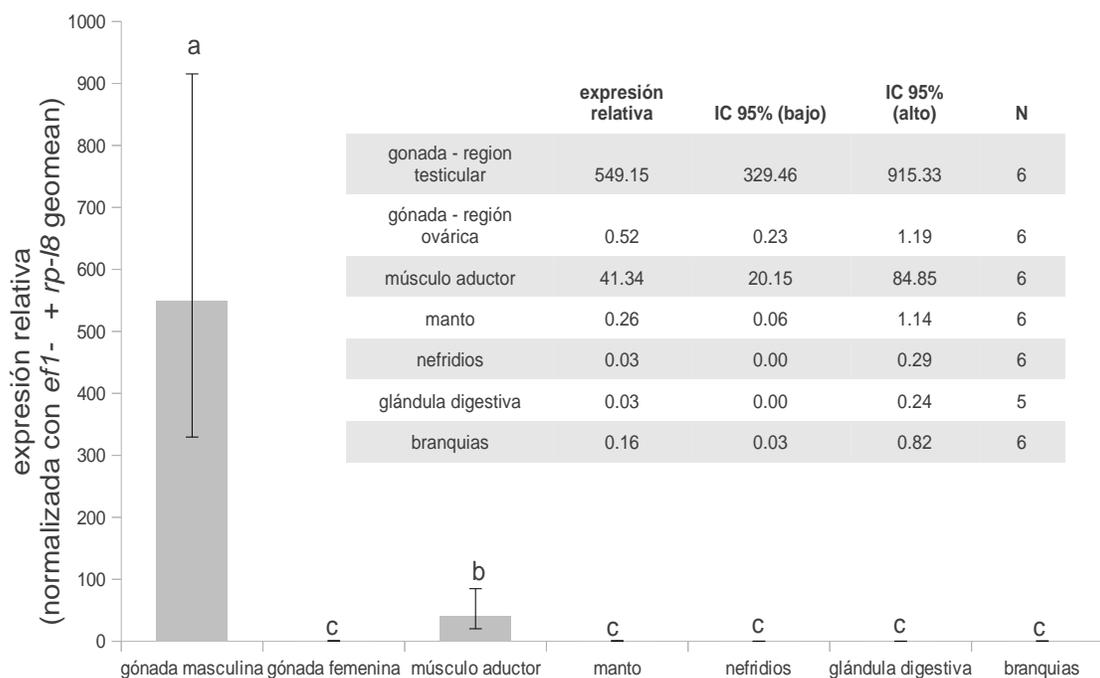


Fig. 25 Expresión relativa de *tub-a* en diferentes tejidos/órganos de *N. subnodosus*, normalizada con la expresión de los genes de referencia a *ef-1a* y *rp-18*. Distintas letras indican diferencias significativas ( $p < 0.05$ ); las barras de dispersión muestran el intervalo de confianza al 95%.

### 7.5 Análisis de expresión de genes asociados a la meiosis

Se evaluó la expresión relativa por qPCR de siete genes involucrados en el proceso de meiosis, así como de un gen que participa potencialmente en la diferenciación sexual de la gónada, comparando la expresión entre estadios de desarrollo y regiones sexuales de la gónada de *N. subnodosus*. Los accesos de las secuencias correspondientes a los genes analizados, así como los oligonucleótidos diseñados, se muestran en la tabla VII.

Cinco de éstos genes mostraron un incremento en el perfil de expresión desde el estadio inactivo (ambas regiones sexuales de la gónada indiferenciada) hacia los estadios de maduración; de estos, cuatro genes mostraron un incremento de su expresión en la región testicular: *Horma domain-containing protein 1 (horma)*, *synaptonemal complex protein 3 (scp3)*, *disrupted meiotic cDNA (dmc1)* y *meiotic recombination element (mre11)*. Por otro

lado, *doublesex and mab-3 related transcription factor 1 (dmrt1)* únicamente se encontró expresado en la región masculina de la gónada, tanto en estadio temprano como avanzado. Aunque algo de expresión para *e3 ubiquitin ligase ccnb1lp1 (e3-ccnb1lp1)* fue detectada en los estadios de maduración ovárica, no se observaron diferencias significativas de éstos estadios con la gónada inactiva, mientras que en la región testicular la expresión de éste gen fue significativamente mayor que en el resto de las condiciones, y mostró una tendencia a incrementar desde el estadio temprano al avanzado de maduración. El único gen que no mostró una expresión significativamente mayor en el estadio de madurez avanzada tanto de la región ovárica como testicular fue *p33-RINGO*. La expresión de éste gen sólo fue observada en la gónada en maduración temprana (ambas regiones sexuales).

Tabla VII. Secuencias de primers para la amplificación por qPCR de genes de interés en meiosis. Las temperaturas de disociación ( $T^{\circ}m$ ) fueron determinadas con el calculador de oligos de SIGMA (<http://www.sigma-geosys.com/calc/DNACalc.asp>). n.a. indica que la secuencia no fue sometida a la base de datos pública de GenBank.

Gen	acceso GenBank	orientación	secuencia de primers para qPCR	$T^{\circ}m$	eficiencia (0-1)
<i>dmcl</i>	KC119408	sentido	5'-AGGTGCCAATGGTTATACCG-3'	63.4	0.96
		antisentido	5'-TGAAGACACCTGGCTCCTC-3'	63.1	
<i>rad51</i>	KC119410	sentido	5'-CTTCACCCATTCGCACTTTT-3'	63.8	0.98
		antisentido	5'-TAGCGTGGCCTTTCACCTTTC-3'	64.0	
<i>sep3</i>	KC119411	sentido	5'-CGCAGTTTTCAACCAGTGGGAAG-3'	70.4	0.97
		antisentido	5'-TTGTGGCAGGTCCCAAGATCC-3'	70.4	
<i>p33-RNGO</i>	isotig01997 <sup>na</sup>	sentido	5'-ATGGGCTCTTGGTCACAAATGG-3'	68.9	0.94
		antisentido	5'-CCATTATCTCTTCACAGCATCGCT-3'	67.2	
<i>e3-ccnb1lp1</i>	G29YTMO03G46S6 (SRA-NCBI)	sentido	5'-TTTGCTTGGGTGACATCTTG-3'	63.5	1.0
		antisentido	5'-ACTGCTCATTTGGCTGAAGA-3'	62.5	
<i>hormad</i>	isotig00711 <sup>na</sup>	sentido	5'-TTTCAAGCCAGCACAGAGTG-3'	64.2	0.98
		antisentido	5'-TTCCTGCACAGGACTCTCAA-3'	63.7	
<i>dmrt</i>	G29YTMO03GFPXJ (SRA-NCBI)	sentido	5'-GGAAGAAGTGCACGTGCG-3'	65.8	0.96
		antisentido	5'-CGTAATGTTGTCCAGGGCT-3'	62.5	
<i>mre11</i>	n.a.	sentido	5'-GCGAGACCGTTAAGAAGCAC-3'	63.8	0.94
		antisentido	5'-CAATGCTGTGGTCTTGAGGA-3'	64.0	

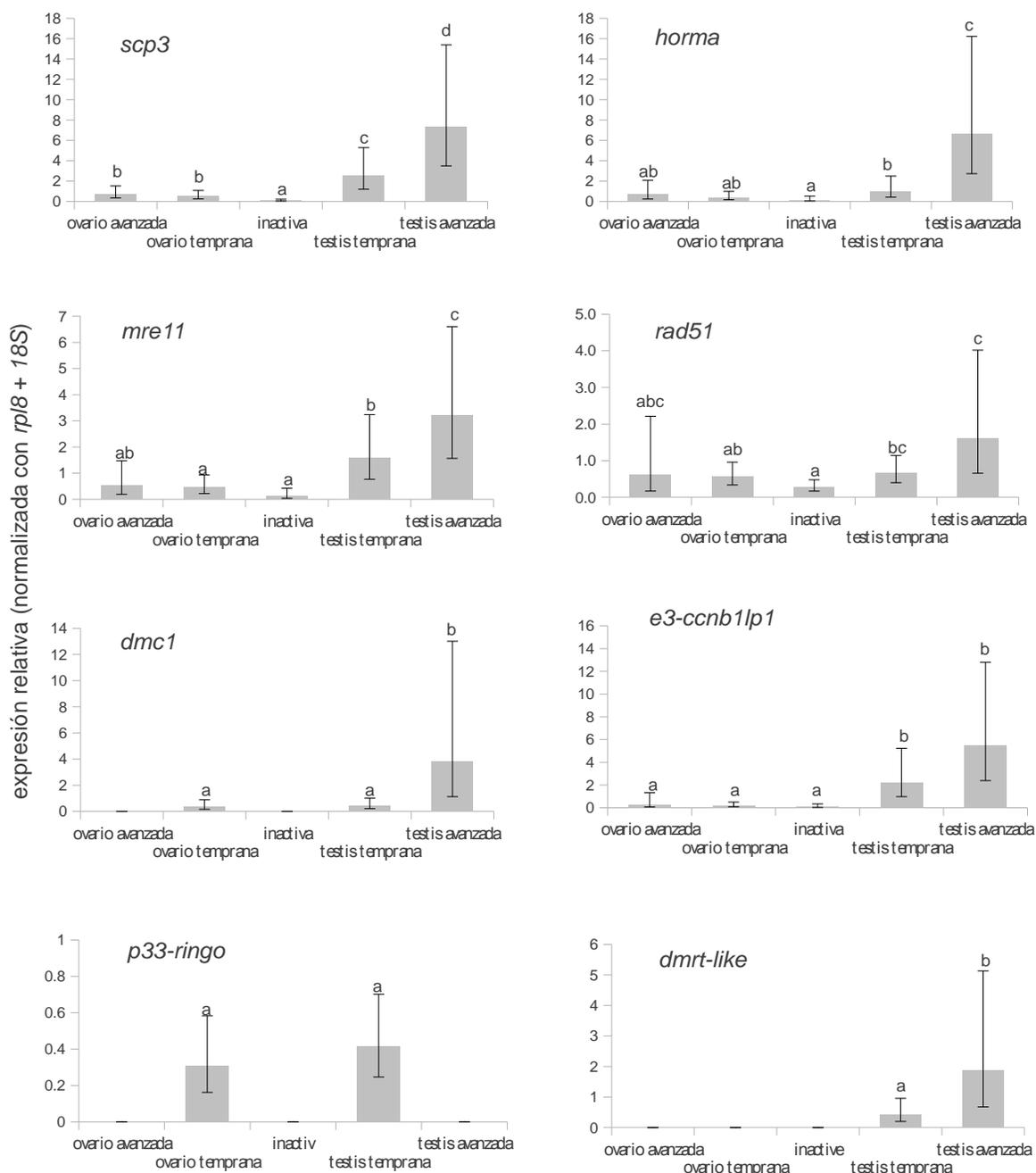


Fig. 26. Expresión relativa de genes seleccionados de la genoteca de testículo en maduración, en distintas condiciones gametogénicas (n = 6 para cada estadio): ovogénesis temprana y tardía en la región ovárica, y estadio temprano y tardío en la región testicular. Distintas letras indican diferencias significativas ( $p < 0.05$ ; prueba post-hoc de Fisher para diferencias entre grupos). Las barras de dispersión muestran el intervalo de confianza al 95%.

## 8 DISCUSIÓN

8.1 Información transcriptómica asociada a la gametogénesis y obtenida *de novo* usando como estrategia genotecas sustractivas y secuenciación por 454-FLX.

La generación de genotecas por hibridación sustractiva y su secuenciación ha sido una técnica ampliamente utilizada para la obtención de información transcriptómica de genes que se expresan específicamente bajo condiciones biológicas contrastantes, incluyendo el testículo *vs.* ovario, o testículo en diferentes etapas de desarrollo en los vertebrados e invertebrados marinos (Vandesompele *et al.*, 2002). Esta metodología se basa en secuenciación *de novo*, por lo que es especialmente apta para los organismos no-modelo que carecen de información genómica, que además de ser un método sensible para discriminar genes diferencialmente expresados, permite el aislamiento de transcritos raros o que se expresan en un número bajo de copias (Andersen *et al.*, 2004). Sin embargo, una desventaja en su aplicación mediante la estrategia de secuenciación clásica (por ej. por metodología de Sanger) puede ser que el esfuerzo y costo de secuenciación requeridos para obtener suficiente información de genes involucrados en un proceso biológico en cada condición comparativa específica puede ser alto, ya que se requiere secuenciar individualmente miles de clonas. En otros moluscos bivalvos un esfuerzo considerable de secuenciación usando genotecas SSH para generar información transcriptómica *de novo* se realizó inicialmente para el ostión *C. gigas*, con la secuenciación de 8.064 clonas de 6 bibliotecas recíprocamente sustraídas de los diferentes órganos (Fleury *et al.*, 2009).

En el presente trabajo se demostró que mediante la implementación de secuenciación masiva, y la combinación de ambas estrategias [(1) genotecas de ADNc por hibridación sustractiva, y (2) pirosecuenciación por la plataforma 454-Roche], es posible obtener una gran cantidad de información *de novo* para transcritos expresados diferencialmente entre la región testicular en maduración *vs* gónada inactiva de la almeja mano de león. De hecho, en este trabajo se obtuvo con éxito en el orden de 10 a 100 veces más secuencias que lo obtenido en trabajos anteriores (Zhang *et al.*, 2007a; Leelatanawit *et al.*, 2008; Fleury *et al.*, 2009; Redon *et al.*, 2010; Sun *et al.*, 2010). Sorprendentemente, y a pesar del impacto

positivo que las plataformas de secuenciación masiva han tenido en el desarrollo de metodologías para el descubrimiento y la caracterización de genes, hasta la fecha sólo dos trabajos reportan el uso de plataformas de secuenciación masiva para las genotecas SSH, uno con un dinoflagelado tóxico y empleando la plataforma 454-Roche, similar al presente trabajo (Toulza *et al.*, 2010), y el otro con soya, empleando la plataforma de secuenciación Illumina (Rodrigues *et al.*, 2012). Los resultados obtenidos indican que la estrategia de combinar genotecas de ADNc del tipo SSH con la secuenciación masiva aumenta significativamente la cantidad de información transcriptómica para genes muy raros en los estudios comparativos de los organismos marinos no modelo.

Aún más, y si bien el objetivo principal de este trabajo no incluyó el identificar microsatélites -repeticiones de secuencias simples (SSRs), un producto de la secuenciación masiva fue el poder identificar la presencia de este tipo de locus en algunos transcritos de las librerías. Las SSR identificados en ESTs tienen la ventaja de un mayor grado de conservación entre especies que las SSR genómicas (también llamadas SSRs anónimas), y por lo tanto son más fáciles de emplear simultáneamente entre especies cercanas para estudios de genómica comparativa estructural (Wren *et al.*, 2000; Li *et al.*, 2004; Ellis y Burke, 2007). De hecho, dos de los EST-SSR identificados en este trabajo tuvieron un hit (blastn) significativo con dos microsatélites descritos para otras especies de pectínidos, el microsatélite B119 de *Argopecten purpuratus* (Genbank JN799265.1) y la clona microsatélite CFFD148 de *Chlamys farreri* (Genbank EF148946.1) ([archivo suplementario 3](#)). Un tercer microsatélite identificado en este trabajo resultó ser idéntico a NsubA1B12 reportado en el grupo de ligamiento Nsub1 por Petersen *et al.* (2012). Aunque se ha documentado que los microsatélites en genes codificantes presentan la mayor densidad de repeticiones de trinucleótidos en la región 5'-UTR (Lawson y Zhang, 2008), los microsatélites polimórficos localizados en genes codificantes son de interés en un contexto fisiológico por considerarse como no-neutrales para la selección, ya que la expresión de los transcritos que las contienen puede verse alterada por cambios en el número de repeticiones, afectando potencialmente la región promotora, la estabilidad del UTR 3', así como el marco de lectura de la traducción (Wren *et al.*, 2000; Li *et al.*, 2004; Ellis y Burke, 2007). Interesantemente, para algunos de los microsatélites en los cuales se obtuvo una

anotación de la secuencia flanqueante, se encontró que la repetición se localizaba en el 3'-UTR del ARNm, como una repetición tri-TAT en la región flanqueante a una secuencia similar a la *troponin-I* del ADNc de *Patinopecten yessoensis* (GenBank AB008006.1). En cuanto a microsatélites localizados en el marco de lectura de algunos genes, se ha documentado que las expansiones o contracciones de los SSR localizados en la región codificadora puede dar lugar a cambios de marco de lectura, codones de paro prematuros, y cambios de secuencias de aminoácidos, aun cuando la repetición más común en la región codificante del ARNm es de tipo trinucleótido (o sus múltiplos). Es importante mencionar que los microsatélites identificados en este estudio contribuyen con el panel de microsatélites polimórficos generado inicialmente por Ibarra *et al.* (2006) y Petersen *et al.* (2009) para la especie sujeto de estudio, y podría ayudar a refinar el mapa genético en el que se obtuvo un mapa consenso con 22 grupos de ligamiento (Petersen *et al.*, 2012), cuando el número haploide de cromosomas es de 19 (Ibarra *et al.*, 2011).

Regresando al transcriptoma y a las genotecas obtenidas, el ensamble 'híbrido' realizado, el cual combina todas las lecturas obtenidas de ambas genotecas, pero mantiene la información del origen de las lecturas individuales que conforman cada isotig (= contig) o singleton, permitió 'dividir' o separar los transcritos del ensamble final en conjuntos de transcritos representados en cada genoteca, así como los transcritos que son construidos con lecturas de ambas genotecas. Este análisis indicó que el número de isotigs que se conformaron con lecturas procedentes de las dos genotecas comprendía solamente el 2.6% de un total de 3,727, lo cual es evidencia de la adecuada "sustracción" durante la generación de las genotecas SSH, eliminando aquellos transcritos que son comunes entre ambas condiciones. Algunos trabajos (Cao *et al.*, 2004; Bui *et al.*, 2005; Fleury *et al.*, 2009; Toulza *et al.*, 2010; Rodrigues *et al.*, 2012) han sugerido que durante la generación de genotecas SSH ocurre un enriquecimiento de transcritos raros, pero que aquellos transcritos muy abundantes en ambas condiciones pueden escapar de la normalización sustractiva. Estos transcritos abundantes pueden incluir transcritos codificantes para proteínas ribosomales -componentes de los ribosomas-, que poseen una alta tasa de transcripción y que también están implicadas en muchas funciones extra-ribosomales no relacionadas a la propia maquinaria de traducción (Warner y McIntosh, 2009). Del mismo modo, algunos

transcritos como la *cytochrome C oxidase*, *ATP synthase* y *NADH dehydrogenase*, que funcionan en las mitocondrias y participan en las demandas energéticas basales, son de esperarse que se expresen tanto en la región testicular en maduración como en la gónada inactiva. Sin embargo, al menos dos transcritos interesantes se encontraron conformados por lecturas derivadas de ambas genotecas. Uno de ellos codifica para una proteína homóloga de MAGO NASHI, que en *Drosophila melanogaster* está involucrado en el establecimiento de la delimitación del eje anteroposterior en el embrión a través de una correcta localización de ARNm *oskar* hacia el polo posterior (Boswell *et al.*, 1991). En *Caenorhabditis elegans*, también hermafrodita como la especie de éste trabajo, éste gen es necesario para la morfogénesis embrionaria como en *Drosophila*, pero interesantemente también es requerido para la regulación de la transición de esperma a ovocitos en hermafroditas adultos. Esto fue demostrado por primera vez por estudios de silenciamiento por iARN (interferencia de ARN) en el que los nematodos a los que se les silenció *mago nashi* mostraron masculinización de la línea germinal (Li *et al.*, 2000). Estos autores propusieron que MAGO NASHI inhibe la expresión de genes masculinizantes como *fog*, *fem*, y *gld* en la línea germinal, lo que permite que la ovogénesis progrese. El producto de éste gen, junto con Y14, forma parte del complejo EJC (*splicing-dependent exon-exon junction complex*) relacionado al sistema de regulación de ARNm llamado *nonsense-mediated mRNA decay*, pero su modo de acción específico en la promoción de la transición de esperma a ovocitos sigue siendo desconocido (Kawano *et al.*, 2004). Independientemente de su modo de acción, el involucramiento de MAGO NASHI en el *splicing* de genes específicos ha sido más recientemente demostrado también en *Drosophila* fuera de la línea germinal (Roignant y Treisman, 2010). Otro transcrito de interés en el contexto de línea germinal, que fue identificado como expresado en ambas genotecas, y anotado como *proteína yellow*, forma parte de la familia de proteínas YELLOW, hasta ahora sólo identificadas en artrópodos, hongos y algunas bacterias (Ferguson *et al.*, 2011). Se sabe que las proteínas YELLOW están implicadas en el comportamiento de apareamiento y en mecanismos de melanización, y han sido también reportadas como sobreexpresadas por el mecanismo denominado *amplificación génica* en las células del folículo del ovario de *Drosophila*, donde es requerido para la formación de

la cubierta del ovocito (Claycomb *et al.*, 2004). Adicionalmente, algunas proteínas YELLOW han sido encontradas en asociación con la proteína específica de macho FRUITLESS en las neuronas del sistema nervioso central (SNC), pero también están presentes en el SNC de hembras (Radovic *et al.*, 2002). Considerando que la función específica de éste gen en la gónada de la almeja es desconocida, el hallazgo de este transcrito tanto en la gónada inactiva como en la región testicular en maduración no es exclusivo de la almeja, ya que todas las proteínas YELLOW se han descubierto como transcritos en múltiples tejidos en *Bombyx mori*, incluyendo testículo y ovario (Xia *et al.*, 2006). La función tanto de MAGO NASHI como de la proteína YELLOW en la almeja tendrá que ser estudiada a futuro con fines de establecer su asociación con el control de la línea germinal.

#### 8.1.1 Descubrimiento, caracterización y expresión de genes selectos e involucrados en la meiosis, recombinación y reparación de ADN durante la gametogénesis.

La recombinación meiótica es un proceso por el cual el cruzamiento genético o intercambio de los cromosomas homólogos se produce durante la gametogénesis. En algunas especies, la recombinación depende del sexo, por ejemplo, la recombinación es ausente en machos de *D. melanogaster*; cuando la recombinación no se produce, no se detecta la transcripción de algunos genes que codifican para funciones específicas en la recombinación (McKee y Ren, 1996; McKee *et al.*, 2012). En la almeja mano de león, se sabe que ocurre la recombinación meiótica durante la producción de gametos masculinos, ya que se determinó recientemente que las tasas de recombinación de marcadores microsatélites cuando se analizan a partir de ambos sexos fueron similares (Petersen *et al.*, 2012). Los genes que participan en la recombinación meiótica identificados en las genotecas SSH representan una corroboración funcional de que la recombinación ocurre durante la meiosis masculina en esta especie. Por ejemplo, algunos de los genes implicados en la recombinación y formación del complejo sinaptonémico (SC) se encuentran en la genoteca de testículo de maduración. El complejo sinaptonémico es una estructura específica de meiosis que consta de dos elementos paralelos axiales (AEs), cada uno asociado a un par de cromátidas hermanas, y un filamento axial situado entre los cromosomas homólogos en sinapsis. Dos

genes que codifican para proteínas constituyentes del complejo sinaptonémico, SCP2 y SCP3, se encontraron expresados en la genoteca de testículo de maduración. Las proteínas SCP2 y SCP3 son dos de las tres proteínas que forman los filamentos axiales del complejo sinaptonémico en la meiosis-I, y son esenciales para el apareamiento de cromosomas, sabiéndose esto debido a que las mutaciones de los genes codificando estas proteínas y que causan su disfunción o no-transcripción en ratones muestran consecuencias fenotípicas como la ausencia de AEs y de sinapsis entre pares de cromosomas homólogos (Pelttari *et al.*, 2001). Otra proteína que participa en la formación de SC es HORMA-domain containing protein 1 (HORMAD1). Esta proteína se co-localiza con las SCP en el elemento axial del SC, y es considerada como un componente estructural central en el complejo sinaptonémico ya que mutaciones en éste gen resultan en la alteración de la localización de SCP2 y SCP3 durante la recombinación temprana, perdiéndose la estructura tripartita del SC (axial y centrales), y dando lugar a la formación de meiocitos aneuploides (Shin *et al.*, 2010). Los resultados de la expresión cuantitativa de *scp3* y *horma*, mostrando un incremento de éstas desde el estadio de inactividad (gónada indiferenciada) hasta el estadio de maduración tardía de la región testicular, indican que la formación de complejos sinaptonémicos, y por lo tanto la meiosis I, se estaba desarrollando hasta lo que en este trabajo se denominó ‘maduración tardía’. De hecho, la diferencia principal entre los estadios tempranos y tardíos de la maduración testicular fue el tamaño de los acinos: mientras que en el estadio de madurez temprana testicular los acinos eran pequeños y se encontraron rodeados de tejido de soporte y conectivo, en el estadio de madurez tardía de la región testicular los acinos eran de mayor tamaño, y se encontraron llenos de células en todas las etapas de espermatogénesis, incluyendo espermatoцитos en estadio inicial (Fig. 4). La expresión de estos genes durante la meiosis I es interesante, ya que la diferenciación morfológica de los espermatoцитos primarios y secundarios no se ha observado en esta especie de almeja. Estos genes podrían servir como futuros marcadores de esta etapa meiótica y esto debe ser probado, de manera similar a lo realizado por Bannister *et al.* (2007).

Además de la cohesión adecuada de los cromosomas homólogos durante la meiosis en la formación de SC, el ADN debe ser reparado después de las rupturas de doble hebra

necesarias para que ocurra la recombinación. En este trabajo el tercer gen cuya expresión aumentó durante la maduración de la región testicular fue *dmc1*, que codifica para uno de los dos recombinasas que participan en el paso crucial de la reparación de las rupturas de doble-hebra del ADN que ocurre durante la recombinación, DMC1 y RAD51. Aunque estas proteínas comparten un alto nivel de secuencia y similitud estructural, se sabe que tienen diferencias bioquímicas y funcionales. El gen *dmc1* se expresa exclusivamente en la meiosis y su producto proteico es requerido para el adecuado apareamiento de pares de cromosomas homólogos y la formación del SC (Bishop *et al.*, 1992). Por otro lado, RAD51, el homólogo eucariota de la enzima bacteriana RecA (Tarsounas *et al.*, 1999), se expresa tanto durante la mitosis como en la meiosis. DMC1 tiene mayor afinidad para la unión al ADN de cadena sencilla y RAD51 forma filamentos nucleoprotéicos helicoidales que promueven la invasión de hebras durante el entrecruzamiento de ADN para la formación de estructuras heterodúplex de ADN (Li y Ma, 2006). Estas dos proteínas co-localizan en sitios de recombinación en una proporción equimolar durante el subestadio meiotico de paquiteno, y debido a que ambos tienen capacidades de interacción ADN-proteína y proteína-proteína equivalentes, se sabe que forman un complejo funcional con actividad recombinasa (Tarsounas *et al.*, 1999). En la almeja mano de león, la expresión de *dmc1* se limitó a aquellas fases de la maduración ovotesticular donde se espera que ocurra la meiosis I, lo que confirma su papel de recombinación meiótico-específica en esta especie. En contraste, la expresión de *rad51* estuvo presente en todas las etapas de desarrollo de las gónadas, incluyendo la gónada inactiva, lo que confirma que este gen es necesario en otros procesos de reparación del ADN, además de la meiosis. En este sentido, se sugiere llevar a cabo análisis posteriores para establecer su posible uso como marcadores transcripcionales de arresto meiótico durante el paquiteno de meiocitos de almejas triploides. Ambos de estos genes fueron caracterizados para su transcrito completo, y el análisis filogenético y estructural los coloca en clados separados entre sí, asociando RAD51A a DMC1 más estrechamente que a RAD51C. La confirmación de la identificación del transcrito de *rad51* obtenido en *Nodipecten subnodosus* como codificando para un transcrito de la subfamilia RAD51C se realizó por un análisis de superposición de estructuras proteicas. Es importante mencionar que se carece de

información relacionada a la distinción funcional de las isoformas A y C de RAD51 en el contexto de recombinación meiótica.

Otro gen anotado como un gen asociado a la meiosis y que se encontró sobreexpresado en la región de testículo fue *ubiquitin ligase e3 ccnblip1* (*e3-ccnblip1*; también conocido como *Human Enhancer of Invasion 10* o *Hei10*), que codifica para una proteína RING (*Really Interesting New Gene*) con un dominio ‘de dedo de zinc’, ha sido reconocido como esencial para la fertilidad en ratones ya que participa en la resolución de intermediarios de recombinación mediante la formación/maduración del quiasma (Strong y Schimenti, 2010). Este componente de la maquinaria de ubiquitinación también participa en la maduración meiótica en la frontera profase-metafase I, regulando el equilibrio en la acumulación y degradación de la ciclina B (Ward *et al.*, 2007). Un ortólogo se ha descrito recientemente en el arroz, y se propone que presenta homología con la ZIP3 de levaduras y con ZHP 3 de *Caenorhabditis elegans*, con función en la formación de *crossovers* normales durante la recombinación meiótica (Wang *et al.*, 2012). En contraste con *horma* y *scp3*, la expresión de *e3-ccnblip1* no aumentó durante la maduración testicular, lo que podría ser un indicador de que su función principal ocurre durante la meiosis temprana similarmente a lo observado en ratones. Además, el patrón de expresión que se observó para *e3-ccnblip1* fue específico para la región testicular, con la expresión en tejido de la región ovárica sin diferencias significativas con la expresión estimada en la gónada inactiva, lo que sugiere que la regulación de este gen es testículo-específica en *N. subnodosus*.

Entre los genes anotados con participación en la meiosis y seleccionados para ser analizados en cuanto a su expresión entre los estadios de la gametogénesis y región sexual, *p33-ringo* fue el único con una expresión específica en la maduración temprana de ambas regiones sexuales. La proteína codificada por este gen, también conocido como *speedy*, primero fue implicada como un iniciador de la maduración de ovocitos mediante la inducción de la transición G2-M en los ovocitos de *Xenopus* (Ferby *et al.*, 1999, citado por Yamashita, 2000). Posteriormente se demostró que induce a la progresión de G1-S cuando se sobreexpresa en diferentes tipos celulares de humanos, y más recientemente ha sido asociada con la promoción de tumores, posiblemente debido a previene o limita la

activación de *checkpoints* (McAndrew *et al.*, 2009). El hecho de que no se detectó expresión en la región testicular en maduración puede ser el resultado de que este gen resulta expresado en niveles muy bajos y no es detectable por qPCR, incluso si la sustracción tuvo éxito en el aislamiento de dos lecturas idénticas que comprenden el isotig p33-ringo. Además, la nula expresión en la región ovárica en maduración indica que este gen posiblemente no está implicado en la maduración de ovocitos como en *Xenopus*. Finalmente, es importante mencionar que un estudio reciente de la familia de proteínas Speedy/Ringo en vertebrados ha encontrado que esta familia de genes se expresa predominantemente en el testículo, aunque niveles menores de expresión fueron detectados también en otros tejidos (Chauhan *et al.*, 2012). La caracterización del transcrito de este gen, aunado con experimentos de interferencia de la transcripción, serán en conjunto importante para entender la función del mismo durante la meiosis inicial en la almeja, especialmente cuando los triploides sufren un arresto meiótico.

Considerando que varias otras proteínas que están implicadas en la sinapsis y recombinación cromosómica se expresaron durante la maduración testicular, es importante mencionar también otros genes que están potencialmente implicados en la sincronización y la revisión de estos acontecimientos en un contexto de ciclo celular. En particular, los denominados genes de '*checkpoint*' (punto de control) (también incluidos en la Tabla I), son de interés para ser analizados a futuro en el contexto de triploides inducidos por la esterilidad. En forma general un *checkpoint* del ciclo celular funciona para mantener la integridad del genoma, garantizando la secuencia de acontecimientos que se produce durante el proceso de la división celular y la segregación cromosómica (Bailis *et al.*, 2000). En diversas especies tales como *S. cerevisiae*, *C. elegans* y *D. melanogaster*, las células meióticas con defectos en la recombinación y sinapsis detonan un retraso o arresto en la etapa de paquiteno de la profase I (*checkpoint* en paquiteno) (Li *et al.*, 2009). Uno de los genes de *checkpoint* meiótico descubiertos en las genotecas es *HORMA domain-containing protein 1* (*horma*), anteriormente discutido como un gen participante en la formación de SC, pero que se reconoce por su participación en la dirección simultánea de varios procesos meióticos. En *C. elegans*, *HTP-1*, el homólogo del *HORMA domain-containing protein 1*, participa en el control de la meiosis mediante la inhibición del ensamble de SC entre

cromosomas no homólogos (Martinez-Perez y Villeneuve, 2005), induciendo una señal de "espera" en la progresión de la sinapsis hasta que la homología de los cromosomas haya sido verificada (MacQueen y Hochwagen, 2011). Otro gen de *checkpoint* de meiosis identificado en la genoteca de la región testicular fue *denticleless protein homolog* (*denticleless E3 ubiquitin protein ligase homolog*), o *Cdt2* en la levadura, se sabe que participa en el control de re-replicación durante la fase S mediante la promoción de la degradación de una proteína necesaria para la replicación del complejo pre-replicación: CDT1 (chromatin licensing and DNA replication factor). CDT1 también es reconocida como un componente esencial del *checkpoint* que opera en la transición G2/M de vertebrados (Sansam *et al.*, 2006), de acuerdo con la observación de una sobreexpresión de éste gen en las gónadas en desarrollo de *Drosophila* (Sloan *et al.*, 2012).

Así como algunas de las proteínas mencionadas antes con una doble función en diferentes etapas de la meiosis, también se ha reconocido que existe una doble función de ciertas proteínas que participan en la revisión o reparación del daño en ADN inducido por mitosis y meiosis, induciendo mecanismos de *checkpoint* (Freire *et al.*, 1998). Al menos dos transcritos de este tipo fueron identificados en la genoteca de la región testicular de la gónada en maduración. Uno de ellos es *Rad1*, que codifica una proteína de *checkpoint* implicada en el arresto celular en G2 hasta que se completa la reparación de ADN (Rowley *et al.*, 1992). En el humano, RAD1 es abundante en testículo, y se asocia con los cromosomas tanto en sinapsis como en asinapsis durante la profase meiótica I de la espermatogénesis (Freire *et al.*, 1998). Otro gen de *checkpoint* identificado como expresado en la región testicular en maduración fue *dual specificity protein kinase TTK*, la cual se sabe que activa por fosforilación a p53 cuando el huso mitótico se interrumpe después de un prolongado período en mitosis sin división celular. En respuesta al estrés, p53 modula la expresión de genes blanco como *p21*, que son necesarios para el arresto del ciclo celular (Huang *et al.*, 2009).

Adicionalmente a genes tipo '*checkpoint*' descubiertos en la genoteca de la región testicular, un gen aislado *de novo* fue *pch2* (*pachytene checkpoint protein 2*) conocido por su función específica en el control de la fidelidad de la meiosis. Este gen fue originalmente

identificado en levaduras (San Segundo y Roeder 1999) pero que ha sido estudiado en organismos modelo como *C. elegans* y *D. melanogaster* (Bhalla y Dernburg 2005; Joyce y McKim, 2009). Se sabe que la proteína codificada por el mismo actúa en el monitoreo de los intermediarios de la recombinación y sinapsis, induciendo un arresto o retraso meiótico hasta que ocurra la reparación completa de los rompimientos de doble cadena entre cromátidas de cromosomas homólogos por efecto de recombinación. También se sabe que en mutantes disfuncionales para este gen, no ocurre un retraso o arresto meiótico a pesar de existir rompimientos de la doble cadena del DNA no reparados o resueltos (Joyce y McKim, 2009).

#### 8.1.2 Descubrimiento de genes de diferenciación sexual y genes específicos de la espermatogénesis.

En los organismos con una determinación cromosómica del sexo, una vez que el destino sexual de la gónada se ha establecido, genes específicos de diferenciación sexual son transcritos para mantener el sexo de la gónada (Lalli, 2003). En hermafroditas funcionales como la almeja mano de león, los genes de diferenciación sexual se espera que conduzcan a la formación de las regiones sexuales separadas, y que presenten un perfil de regulación específica para cada región sexual. La evidencia de esto es el descubrimiento de varios genes en la genoteca de región testicular cuya anotación funcional sugiere que se encuentran implicados en la diferenciación sexual hacia macho. Uno de esos genes fue anotado como *dmrt1* (*doublesex and mab-3 related transcription factor 1*), que se sabe que regula la diferenciación sexual masculina desde nematodos hasta los mamíferos. El homólogo de *dmrt* en *Drosophila*, *dsx* o *doublesex*, participa en la parte inferior de la cascada de la diferenciación sexual, siendo uno de los genes mejor caracterizados dentro de aquellos participantes en el mecanismo de diferenciación sexual hasta el momento; este gen se encuentra finamente regulado por splicing alternativo y sexo-específico (Shukla y Nagaraju, 2010). En *C. elegans*, el homólogo de *dmrt1*, *mab-3*, es requerido también para la diferenciación de órganos sexuales y somáticos tanto en machos como en hermafroditas, y su *splicing* sexo-específico se regula por TRA-1 (Yi y Ross, 2000). El análisis de expresión de *dmrt* muestra un patrón de expresión diferencial entre partes sexuales de la

gónada, lo que justifica investigación futura en elucidar el rol funcional que posee este gen en la diferenciación regional de la gónada hermafrodita de *N. subnodosus*.

Otro gen descubierto en la genoteca y anotado con una función en la diferenciación sexual fue *dual specificity protein kinase, clk-2*, cuya proteína CLK2 se reconoce que determina el uso alternativo de los exones 2 y 3 en el homólogo humano de TRA-2 (TRA-2-BETA 1), una proteína SR - nombradas así por sus característicos residuos de serina (S) y arginina (R) - que regula el *splicing* alternativo de transcritos (Stoilov *et al.*, 2004). En *Drosophila*, las mutaciones en el homólogo de *clk* conducen a la disrupción del *splicing* sexo-específico del ARN mensajero prematuro de *doublesex*, lo cual se sabe que depende de TRA y TRA-2 (Du *et al.*, 1998). Finalmente, un tercer gen anotado como implicado en la diferenciación sexual fue *fkbp52 (peptidyl-prolyl cis-trans isomerase)*, que codifica para una co-chaperona de HSP90, y se sabe que es crítico para el desarrollo adecuado y la diferenciación de los órganos reproductivos masculinos en mamíferos (Neuer *et al.*, 2000).

Adicionalmente a los genes antes mencionados como involucrados directamente en la diferenciación sexual, genes cuya expresión está limitada a tejido testicular fueron también identificados en este trabajo. Estudios previos han establecido que en el nematodo hermafrodita *C. elegans*, los perfiles transcripcionales de la gónada testicular y el tejido somático son diferentes (Wang *et al.*, 2009). Hoy en día se sabe que la expresión específica de genes en testículo pueden ser regulada ya sea por promotores específicos reconocidos por factores de transcripción expresados en tejido testicular (Kleene, 2001; Chang *et al.*, 2011), o por un *splicing* sexo-específico (Telonis-Scott *et al.*, 2009). El conjunto de genes descubiertos en *N. subnodosus* cuya anotación funcional los supone como testículo-específicos (Tabla II) incluye algunos de los genes que se expresan específicamente en testículo en otras especies. Por ejemplo, mediante hibridación *in situ*, Yan *et al.* (2004) determinaron que *Kelch-like 10* se encuentra expresado en las espermatidas de ratón; estos autores también demostraron mediante análisis realizados en base de datos UniGene que la mayoría de las ESTs reportados (45 de 46 ESTs) de este gen se obtuvieron a partir de bibliotecas de testículo. Otros transcritos que codifican para proteínas específicas de la espermatogénesis y que fueron encontradas en la genoteca de la región testicular en este

trabajo fueron algunas proteínas estructurales de células de esperma, anotadas como *sperm antigen 16* y *17*, o 'antígenos de esperma' porque fueron descubiertos inicialmente como proteínas inmunorreactivas en experimentos de proteínas diana para la anticoncepción en seres humanos (Domagala y Kurpisz, 2004). Por ejemplo, el gen *sperm antigen 16*, expresado en la fase meiótica, se localiza en el aparato central de la cauda de los espermatozoides y es esencial para la motilidad flagelar (Zhang *et al.*, 2007b). Otro constituyente de la cauda de los espermatozoides es el transcrito identificado como *outer dense fiber protein 3-like*, que sólo se expresa en células haploides y es también conocido como SHIPPO1, codifica para una proteína específica de testículo localizado en los flagelos de espermátidas elongadas y a lo largo de toda la longitud de los espermatozoides maduros (Zhang *et al.*, 2007b). Otro transcrito encontrado en la genoteca de testes fue la *adenosine deaminase domain-containing protein 1-like isoform 2 (adad1)*, el cual codifica para una proteína reconocida por su interacción con ARN de doble hebra. Este gen es también específico de testículo, y en ratones su expresión se limita a las células en espermátocitos en etapa de paquiteno hasta la etapa de espermátida (Schumacher *et al.*, 1995). Las mutaciones en este gen son conocidas por dar lugar a espermatozoides anormales e inducir esterilidad masculina (Connolly *et al.*, 2005). Las aplicaciones potenciales de estos genes específicos de la espermatogénesis en la inducción a una esterilidad en almeja serán presentadas más adelante, en relación a otro gen descubierto en las genotecas, *alfa-tubulina*.

### 8.1.3 Descubrimiento de elementos transponibles y enzimas relacionadas en la genoteca de región testicular en maduración

Los elementos transponibles (TEs, por sus siglas en inglés) son elementos móviles que representan una gran parte del genoma en la mayoría de los eucariotas, incluyendo al del grupo de los moluscos. Por ejemplo, en el genoma publicado recientemente para el ostión *C. gigas*, se estimó que los elementos transponibles en repeticiones invertidas (MITes) comprenden 8,82% del genoma, y los fragmentos de genes derivados de transposasas y transcriptasas se detectaron no sólo en el genoma, sino también como transcritos expresados (Zhang *et al.*, 2012). Se sabe que los elementos transponibles se expresan

generalmente en las células germinales y sus precursores embrionarios para asegurar su herencia (Zamudio y Bourc'his, 2010). Sin embargo, la relajación del control epigenético sobre la expresión de TEs, junto con su integración por de-metilación del genoma, puede conducir a una incorporación potencialmente disruptiva de TEs en el genoma (Bao y Yan, 2012), lo cual puede ser importante en el contexto de esterilidad observada en triploides. El hallazgo de diversas transposasas y polimerasas, tales como *DNA-mediated transposase*, *retrotransposon transcriptase* and *TCB2 transposases* es evidencia de que, no sólo la transcripción de transposones, sino también de algunas de las enzimas necesarias para su retrotransposición y su integración en el genoma, son expresados activamente durante la espermatogénesis de la almeja. Entre esos TEs, se encontraron transcritos de *piggyBac transposable element-derived protein 3-like*, un transposón pequeño del tipo de repeticiones invertidas de ADN, el cual fue aislado originalmente a partir de un genoma de la polilla (Sarkar *et al.*, 2003), pero cuya presencia no se había reportado anteriormente en otros moluscos, se utiliza en el presente como un vector para la transgénesis en insectos y protozoarios parásitos (Su *et al.*, 2012).

Otros TEs fueron encontrados en la región testicular de la gónada en maduración, *retrotransposon-like retr-1*, así como un segundo retrotransposon no clasificado. También fueron encontrados varios transcritos anotados como transcriptasas reversas, como la *reverse transcriptase of the Jockey-like element* y del *transposon Bs*. Estas enzimas poseen la actividad transcriptasa para la polimerización de ADN dirigida por ARN, y requerida para la reinserción de retrotransposones en el genoma (Ivanov *et al.*, 1991). Sorprendentemente, aun con la cercanía evolutiva con el ostión *C. gigas* y dado que una gran proporción del genoma de éste posee TEs como se mencionó antes, ninguno de los transcritos identificados como TEs en la almeja presentaron un "mejor hit" en blastn con algún TE del ostión. Lo anterior se observó a pesar de que 11 proteínas asociadas con TEs fueron predecidas a partir del genoma del ostión *C. gigas*, y dos de ellos se encontraron también en la almeja: una proteína anotada como *RNA-directed DNA polymerase from Bs* y otra anotada como *transposase Tcb2*.

En un futuro cercano resultará importante evaluar si algunos de los transcritos de elementos

transponibles o sus enzimas, recientemente descubiertos en la almeja, presentan una alteración en su regulación durante la gametogénesis en triploides, de tal forma que su posible falta de regulación pueda explicar el arresto meiótico en triploides.

#### 8.1.4 Genes descubiertos en la genoteca de gónada inactiva están involucrados en el control y mantenimiento de la línea germinal.

Con respecto a la genoteca de gónada inactiva, es importante denotar que las muestras utilizadas para la preparación de la genoteca se colectaron en Junio, cuando la temperatura aumenta y la concentración de clorofila está disminuyendo en el ambiente de cultivo (Ramírez-Arce, 2009). En la fecha de muestreo, el 50% de la población de almejas estaba en fase reproductiva inactiva, mientras que el otro 50% se encontraba en alguna fase de maduración o maduros, aunque los testículos de algunas almejas hermafroditas (10%) ya habían liberado gametos (JL Ramírez Arce, datos no publicados). Las observaciones histológicas de las almejas muestreadas indican que algunos organismos inactivos podrían, en algunos casos, superar un primer proceso gametogénico incompleto, con evidencia de reabsorción de las células gaméticas residuales dada por la presencia de hemocitos en todo el tejido gonadal, tanto en el interior como rodeando los acinos rotos. Esta condición reproductiva puede explicar la ocurrencia de muchos de los transcritos que se encontraron en las gónadas inactivas. Los análisis de enriquecimiento de términos ontológicos mostraron las categorías funcionales que son reguladas diferencialmente entre las condiciones de reproducción (gónadas inactivas *vs* testículo en maduración). En contraste con la genoteca de región testicular en maduración, en la cual genes con anotación funcional que incluye términos ontológicos relacionados a la meiosis y el desarrollo espermático fueron notablemente sobrerrepresentados, en la biblioteca inactiva se mostraron enriquecidos aquellos términos ontológicos relacionados a funciones de respuesta inmune y defensa, metabolismo de las proteínas y de la biogénesis de ribonucleoproteínas, lo que sugiere que el proceso de síntesis/degradación de proteínas es activo durante la regeneración de los acinos en los animales después del desove.

Adicionalmente y como una sorpresa importante, algunos genes que se descubrieron en la gónada inactiva fueron indicativos de la ocurrencia de procesos tales como el control y

mantenimiento de la línea germinal. Esto es, el mantenimiento de la línea germinal después de la primera maduración y desove se esperaría que ocurra durante esta condición de inactividad gametogénica ya que una segunda maduración se producirá aproximadamente un año después de la primera (ver Racotta *et al.*, 2003; Maldonado-Amparo *et al.*, 2004). Dado lo anterior, es claro que la identidad sexual de las células germinales masculinas y femeninas tiene que ser mantenida y especificada cuando inicie un nuevo ciclo gametogénico.

Entre los genes que se encontraron en la gónada inactiva se identificó a *notch homolog 2*, un gen que codifica para la proteína del tipo de receptor NOTCH, que se sabe es responsable de mantener un equilibrio entre la proliferación celular y la apoptosis. Las proteínas NOTCH son activadas por la interacción con ligandos específicos, y posteriormente son procesadas por proteólisis específica que resulta en dos fragmentos, dando como resultado un fragmento activo intracelular que activa la transcripción de genes blanco de NOTCH en el núcleo (Wang *et al.*, 2010). En *C. elegans*, la proteína GLP-1 es uno de dos receptores NOTCH en esa especie, la cual, junto con las proteínas FBF (*fem-3 binding factors*), funciona promoviendo las divisiones mitóticas continuas en la línea germinal, previniendo simultáneamente su entrada a la meiosis (Kimble y Crittenden, 2007). La proteínas FBF 1 y 2 también controlan a las células germinales mediante la represión de la traducción de dos reguladores de la meiosis (Crittenden *et al.*, 2002). Estas proteínas son homólogos de las proteínas PUMILIO o PUF, las cuales se caracterizan por ser proteínas de unión al ARN y que inhiben o regulan la traducción del ARNm por su interacción con elementos reguladores en el 3'-UTR. Interesantemente, un segundo transcrito encontrado en la gónada inactiva fue anotado como *pumilio domain-containing protein*. Si bien no se conoce si la línea germinal de la almeja utiliza un mecanismo de represión similar al del nematodo hermafrodita, la identificación de varios genes homólogos a genes involucrados en el control de la diferenciación de la línea germinal permite proponer que tal probabilidad existe y debe ser investigada a futuro. Por ejemplo, entre otros transcritos anotados e identificados como participantes en el control de la línea germinal, *ccr4-not transcription complex subunit 10* es un gen conocido por ser parte de los complejos *ARNm deadenylase* dirigidos a las proteínas reclutadas por PUF, dando como

resultado la represión de la traducción o la inestabilidad del mensajero ya que promueve el acortamiento de las señales de poliadenilación. También se ha descubierto que PUF forma un complejo de represión traduccional junto con EF1 y AGO (Friend y Campbell, 2012). Otro transcrito interesante encontrado en la gónada inactiva fue *btg member 2*, una proteína Tob/BTG que en vertebrados trabaja en el control del crecimiento y la diferenciación, y en *C. elegans* codifica para la proteína específica de línea germinal FOG-3 (cuyo nombre deriva del fenotipo de los mutantes: *Feminization Of Germline*) que regula el que las células germinales se diferencien hacia esperma (Lee *et al.*, 2011). Además, se encontró también en la gónada inactiva un *transcription factor homolog bt3 4 (bicaudal-C* en *Drosophila*), que codifica una proteína de unión a ARN implicada en el control de línea germinal, que regula la longitud de poliadenilación de ARNm específicos de la línea germinal mediante el reclutamiento del producto de *ccr4-not deadenylase* (Chicoine *et al.*, 2007).

Otros dos transcritos encontrados y anotados como factores de transcripción se encuentran involucrados en vías de diferenciación sexual en la línea germinal: el factor de transcripción *hes-1* y el factor de transcripción 21 o *tcf21*. Este último, *tcf21*, es un gen que en ratas ha sido identificado como un elemento regulado directamente por SRY, el gen de determinación testicular que promueve la diferenciación de las células de Sertoli en ratas (Bhandari *et al.*, 2011). El otro factor de transcripción, *hes-1 (hairy and enhancer of split-1* en *Drosophila*), está implicado en la regulación negativa de la transcripción, y se ha encontrado que tiene una expresión estadio-dependiente también en células de Sertoli (Hasegawa *et al.*, 2012). Otro transcrito encontrado en la genoteca de gónada inactiva fue *nuclear receptor subfamily 0 group b member 1*, también conocido como *dax-1* por las iniciales de *dosage-sensitive sex reversal, adrenal hypoplasia critical region on the X chromosome, gene 1*, cuyas mutaciones causan una hipoplasia suprarrenal en el hombre. Este gen es conocido por ser un regulador negativo de la producción de hormonas esteroideas, que actúa por medio de la represión transcripcional de genes implicados en la vía esteroidogénica (Lalli y Sassone-Corsi, 2003), y también por antagonizar e inhibir la función del receptor de andrógenos (AR), posiblemente mediante la modulación de la transcripción de genes AR-dependientes (Holter *et al.*, 2002). Tomados en conjunto, estos

resultados muestran que el mantenimiento de la línea germinal a través de la transcripción de los genes encontrados en la almeja, es un proceso esencial, incluso durante el período de reposo de la gónada después del desove. La importancia del descubrimiento de estos genes está en su posible aplicación sobre el control de la entrada a la meiosis en organismos tanto diploides como triploides. En primera instancia se podría hipotetizar que su posible sobreexpresión, por ejemplo en triploides, resulta en el arresto del proceso de gametogénesis comúnmente visto en organismos de este tipo para esta especie. En segunda instancia se puede hipotetizar que la inducción de la sobreexpresión de algunos de estos genes en diploides resultará en organismos estériles.

## 8.2 Caracterización de secuencias completas de genes en meiosis por RACE

**RAD51:** El análisis filogenético de las secuencias obtenidas en éste estudio se separaron en los clados correspondientes a cada subfamilia DMC1 y RAD51, y se ajustaron a la topología del árbol presentada por (Lin *et al.*, 2006). Sin embargo, puesto que el análisis filogenético incluye un bloque conservado de secuencia donde los miembros de RAD51A y C presentan una similitud muy elevada, se realizó una confirmación estructural por superposición de modelos que confirma que la secuencia del transcrito aislada en este trabajo corresponde a la familia de RAD51C.

La secuencia de PCH2 inferida de la traducción conceptual del transcrito obtenido en éste trabajo también fue analizado en un contexto filogenético, en el cual se observa una separación de las secuencias en clados con concordancia filogenética, separando a los invertebrados del resto de secuencias ortólogas propias de vertebrados; incluso dentro del clado de secuencias de vertebrados, se observa una separación clara entre aquellas secuencias de peces, reptiles, aves y mamíferos. También es interesante mencionar que mediante análisis informáticos de alineamientos locales (blastn, blastx tblastn) no se encontraron ortólogos de éste gen en el genoma completo de *C. gigas*, o en la base nr de GenBank acotada a especies de moluscos; sin embargo, sí se identificaron transcritos de genes en otras especies de cnidarios, anélidos, urocordados (incluidas en el análisis filogenético), además de las secuencias analizadas en *Caenorhabditis elegans* (San

Segundo y Roeder, 1999) y en *D. melanogaster* (Joyce y McKim, 2009), lo que indica que éste gen pueda haberse perdido durante la evolución en algunas especies de moluscos bivalvos, pero no en *N. subnodosus*.

### 8.3 Elección de genes de referencia para la normalización de la expresión génica en gónada de *N. subnodosus*.

En este trabajo se propuso establecer la estabilidad de la expresión para diversos genes potencialmente a utilizar como genes de referencia en la normalización de la expresión génica entre regiones sexuales de la gónada en diferentes estadios de la gametogénesis. Adicionalmente se evaluó la pertinencia de utilizar dos métodos de transcripción reversa del ARN a ADNc para la obtención de transcritos de genes ribosomales presuntamente no poliadenilados (p.ej. 18S-RNA), pero que potencialmente pueden ser utilizados como genes de referencia.

#### 8.3.1 El transcrito del gen ribosomal *18S* en *N. subnodosus* sí es amplificado utilizando oligo-dT.

A pesar de la suposición de que la síntesis de ADNc en presencia de oligo-dT excluye al ARN ribosomal en el proceso de transcripción reversa, ampliamente mencionado en protocolos convencionales de RT-PCR (O'Connell, 2002) fue demostrada como incorrecta para el caso de la almeja mano de león. La transcripción reversa empleando hexanucleótidos aleatorios en lugar de oligo-dT se recomienda cuando los transcritos de los genes blanco por amplificar mediante PCR son no poliadenilados, tal como el ARNm de histonas, ARNm procariota, o ARN ribosomal como *18S* o *28S*. En este estudio se demostró, mediante la comparación de la estabilidad relativa del transcrito de este gen (entre el resto de genes de referencia) en pools de ADNc cebado por oligo-dT y hexanucleótidos aleatorios, que el ARN ribosomal *18S* puede ser utilizado como un gen de referencia en *N. subnodosus* cuando el ADNc es sintetizado con cebadores oligo-dT. También se observó que la abundancia relativa del transcrito de *18S*, aunque inferior en ADNc sintetizado con oligo-dT en comparación al ADNc sintetizado con hexanucleótidos aleatorios, fue mayor que cualquiera de los genes de ARNm analizados. Puesto que una de

las mejores características de un gen de referencia es una alta abundancia (que generalmente resulta en una mínima variación en las réplicas de qPCR), estos resultados apoyan que el fragmento *18S* en ADNc oligo-dT es un gen potencial de referencia para la normalización de la expresión por qPCR en la almeja *N. subnodosus*. De manera similar a este resultado, otros autores han encontrado que los genes del ARN ribosomales pueden ser retrotranscritos utilizando oligo-dT. Por ejemplo, Xue *et al.* (2010) evaluaron la amplificación de un fragmento del ARN ribosomal 28S en ADNc sintetizado utilizando oligo-dT o en combinación de oligo-dT y oligos-antisentido para 28S en líneas celulares de insecto en cultivo *in vitro*. Los autores observaron diferencias de 6 a 13 *Ct*s (*Ct* = ciclo al cual la amplificación es detectada sobre un umbral de fluorescencia) en la amplificación de 28S, con valores de *Ct* más altos (mayor *Ct* = menor abundancia del blanco de amplificación) cuando se amplifica el ADNc retrotranscrito sólo con oligo-dT. Incluso con esta menor abundancia de ARNr 28S en ADNc sintetizado únicamente con oligo-dT, se observó que este gen era el más estable entre un conjunto de cuatro genes cuando las líneas celulares fueron infectadas por virus. Zhu y Altmann (2005), también observaron que el *18S* es un transcrito que puede ser detectado por qPCR en oligo-dT/ADNc utilizando diversos tejidos de rata, incluso después de encontrar que en oligo-dT/ADNc había una cantidad 100 veces menor de éste transcrito cuando se compara con ADNc retrotranscrito con oligonucleótidos antisentido para 18S y cebadores oligo-dT. En nuestro estudio, la diferencia en la cantidad del retrotranscrito obtenido entre oligo-dT y cebadores aleatorios fue sólo de 13 veces, coincidente con los resultados reportados por Xue *et al.* (2010) y Zhu y Altmann (2005) que también encontraron que la transcripción reversa de ARN ribosomal es menos abundante en ADNc sintetizado con oligo-dT que con hexanucleótidos aleatorios. La diferencia podría ser debida a una mayor abundancia de *18S* poliadenilado en *N. subnodosus* que en células de insecto y de rata. Entre las posibles razones por las cuales el ARN ribosomal puede ser un blanco de amplificación con oligo-dT/ADNc podrían estar las siguientes: (1) la aparición de poliadenilación verdadera en el extremo 3' por su transcripción vía ARN polimerasa I o III, o (2) por la presencia de sitios ricos en adeninas (poli-A's) en regiones intermedias del transcrito ribosomal, o (3) por auto-priming del propio ARN ribosomal, resultado de una estructura terciaria compleja durante la

transcripción reversa, y que resulte en la detección por PCR, incluso sin necesidad de añadir ningún tipo de cebador en la reacción RT como fue demostrado por (Tuiskunen *et al.*, 2010). En el primer caso, se ha documentado en la levadura *Saccharomyces cerevisiae* que los niveles bajos pero significativos de ARN ribosomal son de hecho poliadenilados en el extremo 3' por la acción de la *Pap1p* poli-A polimerasa, probablemente como parte de un mecanismo de control de biogénesis y reciclaje de ARN (Kuai *et al.*, 2004). En el humano se ha reconocido también la existencia de fragmentos de ARN ricos en regiones poliadeniladas, no sólo en el extremo 3', sino también en regiones intermedias de los transcritos de ARN ribosomal (Slomovic *et al.*, 2006, 2010). Aún más, una exploración de genotecas de ADNc realizado por Gonzalez y Sylvester (1997) mostró que la presencia de ARN ribosomal es un fenómeno común, incluso después de realizada la selección de ARN poliadenilado y síntesis de ADNc usando oligo-dT como cebador para la retrotranscripción, lo que sugiere que la poliadenilación de regiones intermedias del ARN ribosomal es un proceso generalizado. Más recientemente Zhuang *et al.* (2013) documentaron que la poliadenilación de *18S* es un fenómeno común en diversas especies de algas, probablemente como un mecanismo regulatorio transcripcional que ocurre en el núcleo. Tomando en conjunto todos estos argumentos, es posible concluir que el uso del ARN ribosomal *18S* como un gen de referencia potencial cuando se amplifica en oligo-dT-ADNc es posible, y no debe ser descartado *de facto*, sino hasta realizar análisis de estabilidad de expresión en comparación a otros genes de referencia en las muestras a analizar.

### 8.3.2 Ventajas y desventajas de los programas GeNorm y NormFinder en la estimación de la estabilidad de genes de referencia.

En las especies con poca o ninguna información genómica disponible, un análisis de expresión génica requiere del aislamiento y la secuenciación dirigida (al menos parcial) de la secuencia del ADNc de posibles genes de referencia, tomando en consideración los ya referidos en la literatura como potencialmente estables en condiciones biológicas similares, pero no considerarlos como estables *de facto*. En este trabajo se pre-seleccionó un conjunto de posibles genes de referencia junto con *18S*, incluyendo cuatro genes codificantes para proteínas, y se evaluó cuáles eran los más estables para la normalización de la expresión de

los genes 'blanco'. Para la estabilidad de la expresión génica se compararon dos programas para este análisis, 'geNorm' y 'NormFinder', ya que se ha demostrado que en los pequeños conjuntos de genes de referencia, especialmente cuando dos genes verdaderamente inestables están en gran medida correlacionados en su expresión, la selección de los mejores genes de referencia por el método de GeNorm podría resultar sesgada (De Santis *et al.*, 2010), con un alto riesgo de realizar una interpretación errónea de la verdadera expresión relativa de los genes de interés.

En una primera instancia de análisis, y debido a la sensibilidad antes mencionada de GeNorm para detectar genes correlacionados por su expresión, se utilizó este programa para comprobar si la expresión de *18S* y los otros genes eran equivalentes cuando se analizan simultáneamente los dos conjuntos de ADNc (oligo-dT y hexanucleótidos aleatorios); es decir, para probar la hipótesis de que un gen particular debería presentar una estabilidad similar sin importar el método de *priming* utilizado para la transcripción reversa y la condición biológica en la que se evaluaron. Se esperaba que cada uno de los genes siendo evaluados pero transcritos con ambos métodos fueran clasificados cercanamente en cuanto a su estabilidad y esto fue lo observado, al encontrar que la estabilidad de cada uno de los genes ocurrió en forma de agrupación por el método de *priming*. Esto confirmó que la estabilidad relativa de *18S* no se ve afectada cuando se amplifica en ADNc obtenido con oligo-dT.

Sin embargo, en una segunda instancia de análisis en donde se analizó la estabilidad relativa de cada gen dentro de cada conjunto de ADNc en forma independiente en función de la estrategia de *priming*, el análisis de geNorm detectó diferentes conjuntos de genes como los más estables para cada método de *priming*: a *18S* y *rp-18* como los genes más estables cuando se analizó en el ADNc obtenido con oligo-dT, y a *efl-* y *rp-18* para el ADNc obtenido con hexanucleótidos aleatorios. Esto podría ser debido a que el algoritmo de GeNorm requiere un mayor conjunto de genes de referencia (al menos ocho) para determinar con precisión el conjunto de genes más estable (Vandesompele *et al.*, 2002). Si es así, y solamente se evalúa un número reducido de genes de referencia, un resultado así podría ser tomado como evidencia de que *efl-* es suficientemente estable para ser incluido

como una parte de un conjunto de genes de referencia, cuando en realidad no lo es, como fue demostrado en el análisis por medio de NormFinder. Otro problema se deriva porque GeNorm no detecta la variación entre-grupos como una fuente de sesgo para análisis de cuantificación relativa, ocultando diferencias cuando los genes están en realidad expresados diferencialmente. La condición de inestabilidad de *efl*- como un gen de referencia para las condiciones específicas evaluadas en este trabajo fue evidenciada también al contrastar los coeficientes de variación (CV) para la expresión relativa de cada gen, observándose que el CV para *efl*- fue mayor que 5%, mientras que los CVs de *18S* y *rp-18* eran menos o igual a 5%, así como al analizar los genes de referencia con el segundo programa, NormFinder, el cual sí considera la variación entre-grupos. De Santis *et al.* (2010) y Zhong *et al.* (2008) observaron que incluso cuando un par de genes de referencia se expresan diferencialmente entre los grupos experimentales, GeNorm erróneamente podría seleccionar como "estable" cuando su expresión únicamente se correlaciona, porque, como se ha indicado antes, este programa realmente elige los genes expresados de manera 'más similar' y no necesariamente 'más estable', en algunos casos sin tener en cuenta las posibles diferencias en expresión entre los grupos experimentales.

Finalmente, e independientemente del programa de análisis utilizado, los análisis de estabilidad de los genes de referencia pre-seleccionados permitieron establecer que el gen con la menor estabilidad de los cinco originalmente incluidos fue el de la *alfa-tubulina*.

### 8.3.3 Descubrimiento de un presunto gen de referencia, *tub*- , siendo testis- y músculo-específico.

Una vez definidos los genes de referencia más estables para el análisis de la expresión de genes blanco o de interés, se evaluó la expresión relativa de tres genes, *vitelogenina*, *ddx* y el encontrado como más inestable de los cinco originales de referencia: *tub*- , La expresión de *vit* ocurrió durante la ovogénesis como era de esperarse, debido a que éste gen codifica una lipoproteína expresada en las células del folículo del ovario, y se acumula en los ovocitos en crecimiento (Osada *et al.*, 2004; Llera-Herrera *et al.*, 2012b). Aunque la expresión de *vit* fue 1.2 veces mayor durante la ovogénesis tardía en contraste con la ovogénesis temprana no fue significativamente diferente. Por otro lado, cuando *tub*- se

analizó como un gen de interés (y no como un gen de referencia potencial) se observó que se expresa sólo en la región de la gónada masculina durante la espermatogénesis temprana y tardía, de manera que muestra una tendencia a aumentar a medida que avanzó la espermatogénesis. Las proteínas TUBULIN- y - son proteínas estructurales de los microtúbulos (filamentos de citoesqueleto) y sus propiedades moleculares son diferentes dependiendo de las modificaciones post-traduccionales que ocurren en las regiones C-terminal (Kierszenbaum, 2002; Hammond *et al.*, 2008). Existen al menos dos tubulinas específicas de testículo que participan en la estructura del axonema en flagelos (Nielsen *et al.*, 2010). La expresión no detectada en la gónada indiferenciada o en la región femenina, así como una elevada expresión en la región masculina, en la que divisiones celulares se están produciendo continuamente (espermatogónias a espermatocitos y espermátidas) para culminar con la formación del axonema en los espermatozoides maduros, indican que éste transcrito tiene una expresión testículo-específica. El análisis de expresión posterior en un contexto somático (tejidos de músculo aductor, manto, nefridios, glándula digestiva y branquias) confirmaron lo último, aunque cierta expresión - menor a la registrada en testículo - fue observada en músculo aductor.

#### 8.3.4 Las tubulinas tejido- o testículo-específicas

Las tubulinas componen una familia de proteínas eucariotas estructurales que conforman a los microtúbulos. Todas las células eucariotas requieren microtúbulos, que son componentes del citoesqueleto que participan en la división celular, forma, motilidad y tráfico intracelular (Nielsen *et al.*, 2010). Particularmente en células espermáticas, isotipos específicos de las tubulinas de la subfamilia fueron inicialmente descritos por Kempfues *et al.* (1980). Posteriormente, Theurkauf y Baum (1986), describieron cuatro -tubulinas en *D. melanogaster*, una de las cuales fue testículo específica ( 3). En el ratón, Villasante *et al.* (1983) reportaron que los dímeros formados por las tubulinas 3/7 y 3 eran expresados sólo en testículo maduro, conformando la mayor parte de la estructura de los microtúbulos en las células germinales masculinas, pero que un isotipo, la *tub- 4* se expresa exclusivamente en músculo. Posteriormente, Lewis y Cowan (1988), reportaron una nueva *tubulin-* (*tub- 8*) en humanos y en ratón, que resulta expresada exclusivamente en

musculo esquelético y en testículo. En el artrópodo *Bombyx mori*, Kawasaki *et al.* (2003) encontraron también que la *tub- 3* era testículo-específica, y posteriores análisis filogenéticos confirmaron que éste isotipo ocurre en algunos linajes de artrópodos, y se relaciona a la familia del isotipo *tub- 1* (Nielsen *et al.*, 2010). En todos los casos, los análisis estructurales y por mutación realizados en los isotipos testículo-específicos resultaron en ensamblajes de axonemas disfuncionales y en genotipos estériles para el sexo masculino, confirmando el rol fundamental de la mayoría de las isoformas testículo-específicas para la adecuada función del esperma. Resulta interesante que la expresión tejido-específica de las tubulinas es mediada por sus regiones promotoras (Bo y Wensink, 1989), lo que ha sido de interés por su aplicación en la tecnología de transgénicos para el control biológico de plagas y para la inducción dirigida de la expresión tejido-específica de transgenes utilizando promotores nativos como los de las tubulinas y . Hasta el momento de este trabajo no existía información acerca de una tubulina testículo-específica en moluscos. Para el mejillón *Mytilus edulis*, dos *tub-* se encuentran reportadas en las bases de datos de GenBank: la primera (acceso GenBank ABA03050.1), fue recientemente utilizada por Ciocan *et al.* (2011) como un gen de referencia para la normalización de la expresión génica; los mismos autores reportan una segunda *tub-* aislada de testículo (acceso GenBank AEM36066.1), pero no se analizó su perfil de expresión para identificar especificidad en algún tejido en particular, y por lo tanto no se confirmó o descartó su naturaleza de testículo-específica. Sin embargo, esa secuencia parcial alinea con una similitud del 100% con la aislada en éste trabajo para *N. subnodosus* (GenBank AEN71147.1), mientras que no se produjo un alineamiento significativo contra la primer secuencia aislada de *M. edulis*. También recientemente, en un estudio que empleó una plataforma de microarreglos para definir genes expresados establemente entre distintos tejidos somáticos y germinales del ostión del Pacífico *C. gigas* para que sean empleados como genes de referencia candidatos, Dheilly *et al.* (2011) identificaron que una *tub-* resulta expresada establemente. Lo anterior indica la necesidad de realizar análisis de estabilidad de expresión para genes antes de que sean considerados como potenciales genes de referencia. Finalmente, se recomienda realizar investigaciones orientadas al desarrollo de posibles aplicaciones biotecnológicas de la especificidad tisular de la expresión de la

*tub-a* identificada en éste trabajo.

### 8.3.5 Consecuencias del uso de genes de referencia arbitrariamente elegidos en la normalización de un gen de interés - ejemplo aprendido en el análisis de la expresión de *ddx*.

El análisis de la expresión de un miembro de la familia *ddx* resultó ser un ejemplo interesante de lo que puede suceder cuando la normalización de la expresión génica se lleva a cabo con genes de referencia seleccionados arbitrariamente o basándose únicamente en la literatura disponible. Cuando se simuló lo que ocurriría cuando se utilizan diferentes genes o combinaciones de genes, se encontraron múltiples posibilidades para la interpretación de los resultados, pero la sobreexpresión de este gen durante la espermatogénesis masculina tardía fue la interpretación más frecuente. Sin embargo, cuando la expresión relativa de este gen se evaluó usando los genes encontrados como más estables (*rp-18* y *18S*), se hizo patente que no hay diferencias entre las partes sexuales de la gónada y tampoco entre etapas de desarrollo. *ddx* codifica para un miembro no identificado de la familia helicasa DEAD-box que incluye las subfamilias VASA y PL10 de las helicasas del ARN ATP-dependientes. Aunque *vasa* se expresa específicamente en las células germinales, y tiene un papel funcional en la formación de la línea germinal en el ostión del Pacífico *C. gigas* (Fabioux *et al.*, 2004a, 2004b), el gen *ddx* parcial aislado en el presente trabajo tuvo un mejor hit (empleando blastx contra bases de datos de proteínas de referencia de GenBank) contra la subfamilia PL10 de las helicasas DEAD-box. En el peneido *Marsupenaeus japonicus*, un gen *pl-10* se ha encontrado altamente expresado durante el desarrollo embrionario y larvario, mostrando una expresión relativamente reducida pero mantenida en las gónadas femenina y masculina adultas (Sellars *et al.*, 2007), similarmente a lo encontrado para *ddx* en este trabajo. Su función en peneidos como moluscos es sin embargo desconocida, y se requiere de realizar investigación adicional a futuro con el fin de evaluar su papel funcional y molecular.

## 9 Conclusiones

Los resultados obtenidos han permitido tener un avance significativo en cuanto al conocimiento genómico asociado a la reproducción en la almeja *Nodipecten subnodosus*, identificando un alto número de genes involucrados en la gametogénesis y meiosis que, a través del estudio de su expresión en condiciones comparativas de gametogénesis de diploides y triploides, se espera que contribuirá de forma importante al conocimiento de su función en el contexto de arresto meiótico de los organismos triploides, ya sea como genes marcadores o incluso como posibles determinantes del arresto reproductivo. Muchos de los genes identificados en éste trabajo prometen ser los primeros en ser descubiertos en la gametogénesis de cualquier especie de molusco bivalvo, y en el caso particular de *pch2*, representa el primer registro de éste ortólogo en cualquier especie de bivalvo, nunca antes reportado siquiera en el genoma del ostión *C. gigas*, recientemente secuenciado. El descubrimiento de genes con función en la meiosis, recombinación y reparación del ADN también permite la identificación de marcadores transcripcionales para la meiosis y la espermatogénesis, proporcionando un panel de genes para llevar a cabo caracterizaciones funcionales mediante metodologías de genética inversa como la conocida por interferencia de ARN, una herramienta ya utilizada en los bivalvos marinos durante el proceso de gametogénico. Adicionalmente, la combinación de la estrategia de generación de genotecas por hibridación sustractiva y metodologías de secuenciación masiva como la que ofrece la plataforma 454-Roche, posibilita la re-secuenciación de genotecas ya desarrolladas por la comunidad científica, con costos significativamente menores, y aumentando el potencial de descubrimiento de nuevos genes específicos expresados diferencialmente entre distintas condiciones biológicas.

Adicionalmente, la anotación de los genes obtenidos por secuenciación, permitió identificar un gen potencialmente involucrado con el mecanismo de decisión en la diferenciación sexual (*dmrt-1*), el cual ha sido extensivamente documentado en *D. melanogaster*. Se sabe que éste gen posee una capacidad de *splicing* alternativo (Burtis y Bake, 1988), determinado por dosaje génico mediado por *Tra* (Hoshijima *et al.*, 1991). La posible participación de este gen como parte de una cascada de determinación sexual entre

regiones gonadales, abre la posibilidad de profundizar en el conocimiento de cómo esta especie hermafrodita funcional es capaz de generar gametos simultáneamente en el mismo órgano. La caracterización del gen (transcrito, organización genómica), su regulación y función por estrategias de interferencia del ARN ya se encuentra en proceso. Por otro lado, la anotación de los genes aislados en la genoteca de gónada inactiva permitió identificar varios genes potencialmente involucrados en el control y mantenimiento de la línea germinal.

El establecimiento de los mejores genes de referencia en el contexto gamético/reproductivo permite una robusta interpretación de la expresión génica, y también identifica un transcrito específico de testículo/músculo (*tub-*), lo que posibilita que posteriores estudios puedan desarrollarse con métodos de análisis de expresión optimizados para la correcta interpretación de la expresión diferencial. En este estudio, se demostró que el gen *18S-rARN* de *N. subnodosus* es eficientemente retrotranscrito utilizando oligo-dT como estrategia de cebado. Para las condiciones específicas evaluadas en este estudio, se identificó a *18S* como uno de los genes de referencia más estables para los estudios entre diferentes etapas gametogénicas y regiones sexuales en el tejido gonadal, que junto con la proteína ribosomal *rp-18*, proporcionan la mejor condición de normalización para los análisis de expresión de genes diana, permitiendo una interpretación más robusta de los resultados.

## 10 Literatura citada

- Allen Jr, S.K. S.L. Downing. 1990. Performance of triploid Pacific oysters, *Crassostrea gigas*: gametogenesis. *Can J Fish Aquat Sci.* 47(6):1213–1222.
- Andersen, C.L., J.L. Jensen, T.F. Ørntoft. 2004. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res.* 64(15):5245–50.
- Anisimova, M., O. Gascuel 2006. Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. *System Biol.* 55(4):539–52.
- Arellano-Martínez, M., B.P. Ceballos-Vázquez, C. Lodeiros-Seijo, L.F. Freitas-Valbuena, E. Palacios, I. Racotta, M.E. Valdez-Ramírez N. García, F.A. García-Domínguez (2011). Reproducción. En: Maeda-Martínez A.N., C. Lodeiros-Seijo (eds.) *Biología y cultivo de los moluscos pectínidos del género Nodipecten*. Limusa - Centro de Investigaciones Biológicas del Noroeste, S.C., 99-134p.
- Arellano-Martínez, M., B.P. Ceballos-Vázquez, C. Ruíz-Verdugo, E. Pérez de León, R. Cervantes-Duarte, P.M. Domínguez-Valdéz. 2011. Growth and reproduction of the lion's paw scallop *Nodipecten subnodosus* in a suspended culture system at Guerrero Negro lagoon, Baja California Sur, Mexico. *Aquac Res.* 42(4):571–582.
- Arellano-Martínez, M., B.P. Ceballos-Vazquez, M. Villalejo-Fuerte, F. Garcia-Dominguez, J.F. Elorduy-Garay, A. Esliman-Salgado, I.S. Racotta. 2004. Reproduction of the lion's paw scallop *Nodipecten subnodosus* Sowerby, 1835 (Bivalvia: Pectinidae) from Laguna Ojo de Liebre, BCS, México. *J Shellfish Res.* 23(3):723–730.
- Arellano-Martínez, M., I.S. Racotta, B.P. Ceballos-Vázquez, J.F. Elorduy-Garay. 2004. Biochemical composition, reproductive activity, and food availability of the lion's paw scallop *Nodipecten subnodosus* in the Laguna Ojo de Liebre, B.C.S., Mexico. *J Shellfish Res.* 23:15–23.
- Bailis, J.M., A. V Smith, S. Roeder. 2000. Bypass of a meiotic checkpoint by overproduction of meiotic chromosomal proteins. *Mol Cell Biol.* 20(13):4838–4848.
- Bannister, L.A., R.J. Pezza, J.R. Donaldson, D.G. de Rooij, K.J. Schimenti, R.D. Camerini-Otero, J.C. Schimenti. 2007. A dominant, recombination-defective allele of Dmc1 causing male-specific sterility. *PLoS Biol.* 5(5):e105.
- Bao, J.,W. Yan. 2012. Male germline control of transposable elements. *Biol Reprod.* 86(5):162, 1–14.
- Barker, M.S., K.M. Dlugosch, L. Dinh, R.S. Challa, N.C. Kane, M.G. King, L.H. Rieseberg. 2010. EvoPipes.net: Bioinformatic Tools for Ecological and Evolutionary Genomics. *Evol Bioinformatics.* 6:143–9.
- Barros de Carvalho, G. A., J.S.S. Batista, F.C. Marcelino-Guimarães, L.C. Do Nascimento, M. Hungria. 2013. Transcriptional analysis of genes involved in nodulation in soybean roots inoculated with *Bradyrhizobium japonicum* strain CPAC 15. *BMC Genomics*, 14:153.
- Benfey, T.J., H.M. Dye, I.I. Solar, E.M. Donaldson. 1989. The growth and reproductive endocrinology of adult triploid Pacific salmonids. *Fish Physiol Biochem.* 6(2):113.

- Beninger, P.G., M. Le Pennec. 2006. Structure and function in scallops. En: *Scallops Biology, ecology and aquaculture*. Shumway, S.E. y J.G. Parsons (editors) Elsevier. Pp. 123-211.
- Bhalla, N., A.F. Dernburg. 2005. A conserved checkpoint monitors meiotic chromosome synapsis in *Caenorhabditis elegans*. *Science*. 310(5754):1683–6.
- Bhandari, R.K., I. Sadler-Riggleman, T.M. Clement, M.K. Skinner. 2011. Basic helix-loop-helix transcription factor TCF21 is a downstream target of the male sex determining gene SRY. *PLoS One*. 6(5):e19935.
- Bigot, L., C. Zatylny-Gaudin, F. Rodet, B. Bernay, P. Boudry, P. Favrel 2012. Characterization of GnRH-related peptides from the Pacific oyster *Crassostrea gigas*. *Peptides*. 34:303-310.
- Bishop, D.K., D. Park, L. Xu, N. Kleckner 1992. DMC1: a meiosis-specific yeast homolog of *E. coli* recA required for recombination, synaptonemal complex formation, and cell cycle progression. *Cell*. 69(3):439–56.
- Bo, J., P. Wensink 1989. The promoter region of the *Drosophila* alpha 2-tubulin gene directs testicular and neural specific expression. *Development*. 106(3):581-587.
- Boutet, I., D. Moraga, L. Marinovic, J. Obrequé, P. Chavez-Crooker. 2008. Characterization of reproduction-specific genes in a marine bivalve mollusc: Influence of maturation stage and sex on mRNA expression. *Gene*. 407:130–138.
- Bustin S.A., V. Benes, J.A. Garson, J. Hellemans, J. Huggett, *et al.* (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clinical Chem* 55: 611–622.
- Bustin, S.A. (2002). Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. *J Mol Endocrinol*. 29(1):23–39.
- Camacho, C., G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer, T.L. Madden. 2009. BLAST+: architecture and applications. *BMC Bioinformatics*. 10:421.
- Chang, P.L., J.P. Dunham, S. V Nuzhdin, M.N. Arbeitman 2011. Somatic sex-specific transcriptome differences in *Drosophila* revealed by whole transcriptome sequencing. *BMC Genomics*. 12(1):364.
- Chauhan, S., X. Zheng, Y.Y. Tan, B.-H. Tay, S. Lim, B. Venkatesh, P. Kaldis. 2012. Evolution of the CDK-activator Speedy/RINGO in vertebrates. *Cell Mol Life Sci*. 69(22):3835–50.
- Cheung-Flynn, J., V. Prapapanich, M.B. Cox, D.L. Riggs, C. Suarez-Quian, D.F. Smith. 2005. Physiological role for the cochaperone FKBP52 in androgen receptor signaling. *Mol Endocrinol*. 19(6):1654–66.
- Chiba, T., J. Yao, Y. Higami, I. Shimokawa, M. Hosokawa, K. Higuchi. 2007. Identification of differentially expressed genes in senescence-accelerated mouse testes by suppression subtractive hybridization analysis. *Mammalian Genome*. 18(2):105–12.
- Cicoine, J., P. Benoit, C. Gamberi, M. Paliouras, M. Simonelig, P. Lasko. 2007. Bicaudal-C recruits CCR4-NOT deadenylase to target mRNAs and regulates oogenesis, cytoskeletal organization, and its own expression. *Devel Cell*. 13(5):691–704.
- Chomczynski, P., N. Sacchi. 1987. Single-Step Method of RNA Isolation by Acid Guanidinium Extraction. *Anal Biochem*. 162:156–159.
- Ciocan, C.M., E. Cubero-Leon, C. Minier, J.M. Rotchell. 2011. Identification of

- reproduction-specific genes associated with maturation and estrogen exposure in a marine bivalve *Mytilus edulis*. PLoS One. 6(7):e22326.
- Claycomb, J.M., M. Benasutti, G. Bosco, D.D. Fenger, T.L. Orr-Weaver. 2004. Gene amplification as a developmental strategy: isolation of two developmental amplicons in *Drosophila*. Dev Cell. 6(1):145–55.
- Clémenson, C. M.-C. Marsolier-Kergoat 2009. DNA damage checkpoint inactivation: adaptation and recovery. DNA Repair. 8(9):1101–9.
- Connolly, C., A. Dearth, R. Braun 2005. Disruption of murine *Tenr* results in teratospermia and male infertility. Dev Biol. 278:13-21.
- Crittenden, S.L., D.S. Bernstein, J.L. Bachorik, B.E. Thompson, M. Gallegos, A.G. Petcherski, G. Moulder, R. Barstead, M. Wickens, J. Kimble. 2002. A conserved RNA-binding protein controls germline stem cells in *Caenorhabditis elegans*. Nature. 417(6889):660–3.
- Croll, R.P., C. Wang 2007. Possible roles of sex steroids in the control of reproduction in bivalve molluscs. Aquaculture. 272(1-4):76.
- Cuñado, N., J. Terrones, L. Sánchez, P. Martínez, J.L. Santos. 2002. Sex-dependent synaptic behaviour in triploid turbot, *Scophthalmus maximus* (Pisces, Scophthalmidae). Heredity. 89(6):460–4.
- Cunningham, C., J.-I. Hikima, M. Jenny, R. Chapman, G.-C. Fang, C. Saski, M.L. Lundqvist, R.A. Wing, P.M. Cupit, P.S. Gross, G. Warr, J.P. Tomkins. 2006. New resources for marine genomics: Bacterial Artificial Chromosome libraries for the Eastern and Pacific oysters (*Crassostrea virginica* and *C. gigas*). Mar Biotechnol. 8(5):521–533.
- Dereeper, a, V. Guignon, G. Blanc, S. Audic, S. Buffet, F. Chevenet, J.-F. Dufayard, S. Guindon, V. Lefort, M. Lescot, J.-M. Claverie, O. Gascuel. 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucl Acids Res. 36:W465–9.
- Dernburg, A.F., K. McDonald, G. Moulder, R. Barstead, M. Dresser, A.M. Villeneuve. 1998. Meiotic recombination in *C. elegans* initiates by a conserved mechanism and is dispensable for homologous chromosome synapsis. Cell. 94(3):387–398.
- Dheilly, N.M., C. Lelong, A. Huvet, P. Favrel. 2011. Development of a Pacific oyster (*Crassostrea gigas*) 31,918-feature microarray: Identification of reference genes and tissue-enriched expression patterns. BMC Genomics. 12(1):468.
- Dheilly, N.M., C. Lelong, A. Huvet, K. Kellner, M.-P. Dubos, G. Riviere, P. Boudry, P. Favrel. 2012. Gametogenesis in the Pacific oyster *Crassostrea gigas*: a microarrays-based analysis identifies sex and stage specific genes. PLOS One. 7(5):e36353.
- Diatchenko, L., Y-F Lau, A.P. Campbell, A. Chenchik, et al., 1996. Suppression subtractive hybridization: A method for generating differentially regulated or tissue-specific cDNA probes and libraries. PNAS. 93:6025-6030.
- Domagala, A. and M. Kurpisz 2004. Identification of sperm immunoreactive antigens for immunocontraceptive purposes: a review. Reproductive biology and endocrinology. 2:11.
- Du, C., M.E. McGuffin, B. Dauwalder, L. Rabinow, W. Mattox 1998. Protein phosphorylation plays an essential role in the regulation of alternative splicing and sex determination in *Drosophila*. Mol Cell. 2(6):741–50.

- Dunham, R.A. 2004. Chapter 4: Polyploidy. en: Aquaculture and Fisheries Biotechnology: Genetic Approaches. CABI publishing. Oxford pp. 35-66.
- Edgar, R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucl Acids Res.* 32(5):1792–1797.
- Ellis, J.R., J.M. Burke. 2007. EST-SSRs as a resource for population genetic analyses. *Heredity.* 99(2):125–32.
- Eudeline, B., S.K. Allen, X. Guo. 2000. Optimization of tetraploid induction in Pacific oysters, *Crassostrea gigas*, using first polar body as a natural indicator. *Aquaculture.* 187(1-2):73.
- Fabioux, C., A. Huvet, C. Lelong, R. Robert, S. Pouvreau, J.Y. Daniel, C. Minguant, M. Le Pennec 2004a. Oyster vasa-like gene as a marker of the germline cell development in *Crassostrea gigas*. *Biochem Biophys Res Commun.* 320(2):592–598.
- Fabioux, C., S. Pouvreau, F. Le Roux, A. Huvet 2004b. The oyster vasa-like gene: a specific marker of the germline in *Crassostrea gigas*. *Biochem Biophys Res Commun.* 315(4):897–904.
- Felip, A., S. Zanuy, M. Carrillo, F. Piferrer 1999. Growth and gonadal development in triploid sea bass (*Dicentrarchus labrax* L.) during the first two years of age. *Aquaculture.* 173(1-4):389.
- Ferguson, L.C., J. Green, A. Surridge, C.D. Jiggins 2011. Evolution of the insect yellow gene family. *Mol Biol Evol.* 28(1):257–72.
- Fleury, E., A. Huvet, C. Lelong, J. de Lorgeril, V. Boulo, *et al.* 2009. Generation and analysis of a 29,745 unique Expressed Sequence Tags from the Pacific oyster (*Crassostrea gigas*) assembled into a publicly accessible database: the GigasDatabase. *BMC Genomics.* 10:341.
- Freire, R., J.R. Murguia, M. Tarsounas, N.F. Lowndes, P.B. Moens, S.P. Jackson. 1998. Human and mouse homologs of *Schizosaccharomyces pombe* rad1+ and *Saccharomyces cerevisiae* RAD17: linkage to checkpoint control and mammalian meiosis. *Genes Dev.* 12(16):2560–2573.
- Freites-Valbuena L.F., J.M. Mazón-Suastegui, A.N. Maeda-Martínez, *et al.* 2011. Preengorda, cultivo intermedio y engorde de los pectínidos. En: Maeda-Martínez A.N., C. Lodeiros-Seojo (eds.) *Biología y cultivo de los moluscos pectínidos del género Nodipecten*. Limusa - Centro de Investigaciones Biológicas del Noroeste, S.C., 313-350p.
- Friend, K. and Z. Campbell. 2012. A conserved PUF–Ago–eEF1A complex attenuates translation elongation. *Nat Struct Biol* 19(2):176–183.
- Gan, H., T. Cai, X. Lin, Y. Wu, X. Wang. 2013. Integrative proteomic and transcriptomic analyses reveal multiple post-transcriptional regulatory mechanisms of mouse spermatogenesis. *Mol Cell Proteomics.* 12(5) 1144-1157.
- Gonzalez, I.L., J.E. Sylvester. 1997. Incognito rRNA and rDNA in databases and libraries. *Genome Res.* 7(1):65–70.
- Götz, S., J.M. García-Gómez, J. Terol, T.D. Williams, S.H. Nagaraj, M.J. Nueda, M. Robles, M. Talón, J. Dopazo, A. Conesa. 2008. High-throughput functional annotation and data mining with the Blast2GO suite. *Nucl Acids Res.* 36(10):3420–35.
- Guo, X., G.A. DeBrosse, S.K. Allen. 1996. All-triploid Pacific oysters (*Crassostrea gigas*

- Thunberg) produced by mating tetraploids and diploids. *Aquaculture*. 142(3-4):149.
- Hammond, J.W., D. Cai, K.J. Verhey. 2008. Tubulin modifications and their cellular functions. *Curr Opin Cell Biol*. 20(1):71–6.
- Hasegawa, K., Y. Okamura, Y. Saga. 2012. Notch signaling in Sertoli cells regulates cyclical gene expression of *Hes1* but is dispensable for mouse spermatogenesis. *Mol Cell Biol*. 32(1):206–15.
- Holter, E., N. Kotaja, S. Mäkela, L. Strauss, S. Kietz, O.A. Jänne, J.-A. Gustafsson, J.J. Palvimo, E. Treuter. 2002. Inhibition of androgen receptor (AR) function by the reproductive orphan nuclear receptor DAX-1. *Mol Endocrinol*. 16(3):515–28.
- Huang, Y.-F., M.D.-T. Chang, S.-Y. Shieh. 2009. TTK/hMps1 mediates the p53-dependent postmitotic checkpoint by phosphorylating p53 at Thr18. *Mol Cell Biol*. 29(11):2935–44.
- Hussain, M.G., G.P.S. Rao, N.M. Humayun, C.F. Randall, D.J. Penman, D. Kime, N.R. Bromage, J.M. Myers, B.J. McAndrew. 1995. Comparative performance of growth, biochemical composition and endocrine profiles in diploid and triploid tilapia *Oreochromis niloticus* L. *Aquaculture*. 138(1-4):87.
- Ibarra, A., C. Alfonsi, J.E. Perez 2011. Genética y biotecnologías para el mejoramiento del cultivo y la conservación del recurso *Nodipecten* spp. En: Maeda-Martínez A.N., C. Lodeiros-Sejo (eds.) *Biología y cultivo de los moluscos pectínidos del género Nodipecten*. Limusa - Centro de Investigaciones Biológicas del Noroeste, S.C., 135–144p.
- Ibarra, A.M., J.L. Petersen, T.R. Famula, B. May. 2006. Characterization of 35 microsatellite loci in the Pacific lion-paw scallop (*Nodipecten subnodosus*) and their cross-species amplification in four other scallops of the Pectinidae family. *Mol Ecol Notes*. 6(1):153–156.
- Ibarra, A., C. Alfonsi, J.E. Perez. 2011. Genética y biotecnologías para el mejoramiento del cultivo y la conservación del recurso *Nodipecten* spp. En: Maeda-Martínez A.N., C. Lodeiros-Sejo (eds.) *Biología y cultivo de los moluscos pectínidos del género Nodipecten*. Limusa - Centro de Investigaciones Biológicas del Noroeste, S.C.
- Ivanov, V. a, a a Melnikov, a V Siunov, I.I. Fodor, Y. V Ilyin. 1991. Authentic reverse transcriptase is coded by *jockey*, a mobile *Drosophila* element related to mammalian LINEs. *EMBO J*. 10(9):2489–95.
- Jiang, M., J. Ryu, M. Kiraly, K. Duke, V. Reinke, S.K. Kim. 2001. Genome-wide analysis of developmental and sex-regulated gene expression profiles in *Caenorhabditis elegans*. *PNAS* 98(1):218–223.
- Kataoka, N., M.D. Diem, V.N. Kim, J. Yong, G. Dreyfuss. 2001. Magoh, a human homolog of *Drosophila* mago nashi protein, is a component of the splicing-dependent exon-exon junction complex. *EMBO J*. 20(22):6424–33.
- Kawano, T., N. Kataoka, G. Dreyfuss, H. Sakamoto. 2004. Ce-Y14 and MAG-1, components of the exon–exon junction complex, are required for embryogenesis and germline sexual switching in *Caenorhabditis elegans*. *Mechanisms of Development*. 121(1):27–35.
- Keen, M.A. 1971. *Sea shells of tropical west America*. Stanford University Press, Stanford (CA). 374p.

- Kemphues, K.J., E.C. Raff, T.C. Kaufman, R.A. Raff. 1980. Mutation in a testis-specific beta-tubulin in *Drosophila*: analysis of its effects on meiosis and map location of the gene. *Cell*. 21(2):445–51.
- Ketata, I., X. Denier, A. Hamza-Chaffai, C. Minier 2008. Endocrine-related reproductive effects in molluscs. *Comp Biochem Physiol C*. 147(3):261–270.
- Kierszenbaum, A.L. 2002. Sperm axoneme: a tale of tubulin posttranslation diversity. *Mol Reprod Dev*. 62:1–3.
- Kimble, J., S.S.L. Crittenden. 2007. Controls of germline stem cells, entry into meiosis, and the sperm/oocyte decision in *Caenorhabditis elegans*. *Annu Rev Cell Dev Biol*. 23:405–33.
- Kleene, K.C. 2001. A possible meiotic function of the peculiar patterns of gene expression in mammalian spermatogenic cells. *Mech Dev*. 106(1-2):3–23.
- Kuai, L., F. Fang, J.S. Butler, F. Sherman. 2004. Polyadenylation of rRNA in *Saccharomyces cerevisiae*. *PNAS*. 101(23):8581–8586.
- Lalli, E. 2003. Sexy splicing: regulatory interplays governing sex determination from *Drosophila* to mammals. *J Cell Sci*. 116(3):441–445.
- Lalli, E., P. Sassone-Corsi. 2003. DAX-1, an unusual orphan receptor at the crossroads of steroidogenic function and sexual differentiation. *Mol Endocrinol*. 17(8):1445–53.
- Larkin, M.A, G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, J.D. Thompson, T.J. Gibson, D.G. Higgins. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics*. 23(21):2947–8.
- Lawson, M.J. L. Zhang. 2008. Housekeeping and tissue-specific genes differ in simple sequence repeats in the 5'-UTR region. *Gene*. 407(1-2):54–62.
- Lee, M.-H., K.W. Kim, C.T. Morgan, D.E. Morgan, J. Kimble. 2011. Phosphorylation state of a Tob/BTG protein, FOG-3, regulates initiation and maintenance of the *Caenorhabditis elegans* sperm fate program. *PNAS*. 108(22):9125–30.
- Leelatanawit, R., S. Klinbunga, T. Aoki, I. Hirono, R. Valyasevi, P. Menasveta. 2008. Suppression subtractive hybridization (SSH) for isolation and characterization of genes related to testicular development in the giant tiger shrimp *Penaeus monodon*. *BMB Reports*, 41(11):796–802.
- Leitao, A., P. Boudry, C.E. Thiriou-Quievreux. 2001. Negative correlation between aneuploidy and growth in the Pacific oyster, *Crassostrea gigas*: ten years of evidence. *Aquaculture*. 193:39–48.
- Li, W., R. Boswell, W.B. Wood. 2000. mag-1, a homolog of *Drosophila mago nashi*, regulates hermaphrodite germ-line sex determination in *Caenorhabditis elegans*. *Dev Biol*. 218(2):172–82.
- Li, W., H. Ma. 2006. Double-stranded DNA breaks and gene functions in recombination and meiosis. *Cell Res*. 16(5):402–412.
- Li, X.C., B.C. Barringer, D.A. Barbash. 2009. The pachytene checkpoint and its relationship to evolutionary patterns of polyploidization and hybrid sterility. *Heredity*. 102(1):24–30.
- Li, Y.-C., A.B. Korol, T. Fahima, E. Nevo. 2004. Microsatellites within genes: structure, function, and evolution. *Mol Biol Evol*. 21(6):991–1007.
- Liang, G., X. Zhang, L. Wang, Y. Sha, J.C. Zhang, S.Y. Miao, S.D. Zong, *et al.* 2004.

- Identification of differentially expressed genes of primary spermatocyte against round spermatid isolated from human testis using the laser capture microdissection technique. *Cell Res*, 14(6):507–512.
- Lin, Z., H. Kong, M. Nei, and H. Ma. 2006. Origins and evolution of the *recA/RAD51* gene family: evidence for ancient gene duplication and endosymbiotic gene transfer. *PNAS*. 103(27):10328–33.
- Llera-Herrera, R., C. Vázquez-Boucard, A. García-Gasca, A. Huvet. 2012. Co-expression and regulation of ovarian vitellogenins in the Pacific oyster *Crassostrea gigas*. *Aquac Res. en prensa*.
- López-Sánchez, J.A., A.N. Maeda-Martínez, R.P. Croll, H. Acosta-Salmón. 2009. Monoamine fluctuations during the reproductive cycle of the Pacific lion's paw scallop *Nodipecten subnodosus*. *Comp Biochem Physiol A*. 154(3):425–8.
- Lu, X., C.I. Wu. 2011. Sex, sex chromosomes and gene expression. *BMC Biol*. 9:30.
- MacQueen, A.J., A. Hochwagen. 2011. Checkpoint mechanisms: the puppet masters of meiotic prophase. *Trends Cell Biol*. 21(7):393–400.
- Maldonado R., A. Ibarra, J. Ramírez. 2003. Induction to tetraploidy in catarina scallop, *Argopecten ventricosus* (Sowerby II, 1842). *Cienc Mar*. 29(2):229–238.
- Maldonado-Amparo, R., A.M. Ibarra. 2002a. Ultrastructural characteristics of spermatogenesis in diploid and triploid catarina scallop (*Argopecten ventricosus* Sowerby II, 1842). *J Shellfish Res*. 21(1):93–101.
- Maldonado-Amparo, R. and A.M. Ibarra 2002b. Comparative analysis of oocyte type frequencies in diploid and triploid catarina scallop (*Argopecten ventricosus*) as indicators of meiotic failure. *J Shellfish Res*. 21(1):597–604.
- Maldonado-Amparo, R., J.L. Ramirez, S. Avila, A.M. Ibarra. 2004. Triploid lion-paw scallop (*Nodipecten subnodosus* Sowerby); growth, gametogenesis, and gametic cell frequencies when grown at a high food availability site. *Aquaculture*. 235(1-4):185–205.
- Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J*. 17(1):10–12.
- Martinez-Perez, E., A. Villeneuve. 2005. HTP-1-dependent constraints coordinate homolog pairing and synapsis and promote chiasma formation during *C. elegans* meiosis. *Genes Dev*. 2727–2743.
- Mazón-Suastegui, J.M., A.N. Maeda-Martínez, M. Robles-Mungaray, J.P. De La Roche, G.S. Rupp, M. Mendes-De-Bem, L.A. Velasco, L.F. Freitas-Valbuena. 2011. Avances en la producción de juveniles de *Nodipecten* spp. En: Maeda-Martínez A.N. and Lodeiros-Seijo C. (eds.) *Biología y Cultivo de los Moluscos Pectínidos del Género Nodipecten*. Limusa - Centro de Investigaciones Biologicas del Noroeste, S.C.,
- McAndrew, C., R. Gastwirt, D. Donoghue. 2009. The atypical CDK activator *Spy1* regulates the intrinsic DNA damage response and is dependent upon p53 to inhibit apoptosis. *Cell Cycl*. 8(1):66–75.
- McCombie, H., C. Ledu, P. Phelipot, S. Lapegue, P. Boudry, A. Gerard. 2005. A complementary method for production of tetraploid *Crassostrea gigas* using crosses between diploids and tetraploids with cytochalasin B treatments. *Mar Biotechnol*. 7(4):318–330.

- McDermott, S., M. Noor. 2010. The role of meiotic drive in hybrid male sterility *Phyl Trans Roy Soc B.* 365:1265–1272
- Meglécz, E., C. Costedoat, V. Dubut, A. Gilles, T. Malausa, N. Pech, J.-F. Martin. 2010. QDD: a user-friendly program to select microsatellite markers and design primers from large sequencing projects. *Bioinformatics.* 26(3):403–4.
- Meier, B. A. Gartner. 2006. Meiosis: checking chromosomes pair up properly. *Curr Biol.* (2004):2004–2006.
- Murakami, H., P. Nurse. 1999. Meiotic DNA replication checkpoint control in fission yeast. *Gene Dev* 13:2581–2593.
- Naisbit, R.E., C.D. Jiggins, M. Linares, C. Salazar, J. Mallet. 2002. Hybrid Sterility, Haldane's rule and speciation in *Heliconus cydno* and *H. melpomene*. *Genetics* 1526:1517–1526.
- Nakamura, S., M. Osada, A. Kijima. 2007. Involvement of GnRH neuron in the spermatogonial proliferation of the scallop, *Patinopecten yessoensis*. *Mol Reprod Dev.* 74:108-115.
- Newmark, P. a, S.E. Mohr, L. Gong, R.E. Boswell. 1997. Mago nashi mediates the posterior follicle cell-to-oocyte signal to organize axis formation in *Drosophila*. *Development.* 124(16):3197–3207.
- Nielsen, M.G., S.R. Gadagkar, L. Gutzwiller. 2010. Tubulin evolution in insects: gene duplication and subfunctionalization provide specialized isoforms in a functionally constrained gene family. *BMC Evol Biol.* 10:113.
- Nolan, T., & Bustin, S. 2008. Procedures for Quality Control of RNA Samples for Use in Quantitative Reverse Transcription PCR. *Essentials of nucleic acid analysis: a robust approach.* Royal Society of Chemistry. p. 189.
- O'connell, J. 2002. Chapter 2: The basics of RT-PCR - Some Practical considerations. en: *RT-PCR Prococols.* Humana Press. 19-25.
- Osada, M., M. Harata, M. Kishida, A. Kijima. 2004. Molecular Cloning and Expression Analysis of Vitellogenin in Scallop, *Patinopecten yessoensis* (Bivalvia, Mollusca). *Mol Reprod Dev.* 67:273–281.
- Palacios, E., I.S. Racotta, O. Arjona, Y. Marty, J.R. Le Coz, J. Moal, J.F. Samain 2007. Lipid composition of the pacific lion-paw scallop, *Nodipecten subnodosus*, in relation to gametogenesis: 2. Lipid classes and sterols. *Aquaculture.* 266(1-4):266.
- Palacios, E., I.S. Racotta, E. Kraffe, Y. Marty, J. Moal, J.F. Samain 2005. Lipid composition of the giant lion's-paw scallop (*Nodipecten subnodosus*) in relation to gametogenesis: I. Fatty acids. *Aquaculture.* 250(1-2):270.
- Pelttari, J., M.R. Hoja, L. Yuan, J.G. Liu, E. Brundell, P. Moens, S. Santucci-Darmanin, R. Jessberger, J.L. Barbero, C. Heyting, C. Hoog. 2001. A meiotic chromosomal core consisting of cohesin complex proteins recruits DNA recombination proteins and promotes synapsis in the absence of an axial element in mammalian meiotic cells. *Mol Cell Biol.* 21(16):5667–5677.
- Petersen, J.L., M.R. Baerwald, A.M. Ibarra, B. May 2012. A first-generation linkage map of the Pacific lion-paw scallop (*Nodipecten subnodosus*): Initial evidence of QTL for size traits and markers linked to orange shell color. *Aquaculture.* 350-353:200–209.
- Petersen, J.L., A.M. Ibarra, B. May. 2007. Genetic consequences of a mass spawning event

- in the lion-paw scallop, *Nodipecten subnodosus*. *Aquaculture*. 272(2007):S299–S299.
- Petersen, J.L., A.M. Ibarra, B. May. 2009. Thirty-seven additional microsatellite loci in the Pacific lion-paw scallop (*Nodipecten subnodosus*) and cross-amplification in other pectinids. *Conserv Gen Res*. 1(1):101–105.
- Petersen, J.L., A.M. Ibarra, B. May. 2010. Nuclear and mtDNA lineage diversity in wild and cultured Pacific lion-paw scallop, *Nodipecten subnodosus*(Baja California Peninsula, Mexico). *Marine Biology*. 157(12):2751–2767.
- Petersen, J.L., A.M. Ibarra, J.L. Ramirez, and B. May. 2008. An Induced Mass Spawn of the Hermaphroditic Lion-Paw Scallop, *Nodipecten subnodosus*: Genetic Assignment of Maternal and Paternal Parentage. *J Hered*. 99(4):337–348.
- Pfaffl, M.W. 2001. A new mathematical model for relative quantification in real-time RT–PCR. *Nucl Acids Res*. 29(9):e45.
- Pfaffl, M.W. 2004. Quantification strategies in real-time PCR. En: Bustin S. (eds.) *AZ of quantitative PCR*. International University Line, La Jolla, CA. 87–112p.
- Pfaffl, M.W., A. Tichopad, C. Prgomet, T.P. Neuvians. 2004. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper–Excel-based tool using pairwise correlations. *Biotechnol Lett*. 26(6):509–515.
- Piferrer, F., A. Beaumont, J.C.J. Falguière, M. Flajshans, P. Haffray, L. Colombo, M. Flajshans. 2009. Polyploid fish and shellfish: Production, biology and applications to aquaculture for performance improvement and genetic containment. *Aquaculture*. 293(3-4):125–156.
- Pinto, P.I.S., H.R. Teodósio, M. Galay-Burgos, D.M. Power, G.E. Sweeney, A.V.M. Canário. 2006. Identification of estrogen-responsive genes in the testis of sea bream (*Sparus auratus*) using suppression subtractive hybridization. *Mol Rep Dev*, 73(3):318–29.
- Racotta, I.S., E. Palacios, A.M. Ibarra, J.L. Ramírez, F. Arcos, O. Arjona. 2008. Comparative biochemical composition of ploidy groups of the lion-paw scallop (*Nodipecten subnodosus* Sowerby) supports the physiological hypothesis for the lack of advantage in triploid mollusc's growth in food-rich environments. *Mar Biol*. 153(6):1245–1256.
- Racotta, I.S., J.L. Ramírez, A.M. Ibarra, M.C. Rodríguez-Jaramillo, D. Carreño, E. Palacios. 2003. Growth and gametogenesis in the lion-paw scallop *Nodipecten* (*Lyropecten*) *subnodosus*. *Aquaculture*. 217(1-4):335–349.
- Radovic, A., P.J. Wittkopp, A.D. Long, and M.D. Drapeau 2002. Immunohistochemical colocalization of Yellow and male-specific Fruitless in *Drosophila melanogaster* neuroblasts. *Biochemical and biophysical research communications*. 293(4):1262–4.
- Ramírez-Arce, J. 2009. Evaluacion de la ventaja productiva y grado de esterilidad en triploides de almeja mano de leon *Nodipecten subnodosus*(Sowerby 1835) como una alternativa para el cultivo en el Parque Naciona Bahía de Loreto, Golfo de California. CICIMAR,
- Rodrigues, F. A., J. Marcolino-Gomes, J. De Fátima Corrêa Carvalho, L.C. Do Nascimento, N. Neumaier, J.R.B. Farias, M.F. Carazzolle, *et al.* 2012. Subtractive libraries for prospecting differentially expressed genes in the soybean under water deficit. *Genet*

- Mol Biol. 35(1(suppl)):304–14.
- Roeder, G. J.M. Bailis. 2000. The pachytene checkpoint. Trends Genet. 9525(1999):395–403.
- Roeder, G.S. J.M. Bailis. 1997. Meiotic chromosomes. It takes two to tango. Gene Dev. 11:2600–2621.
- Rowley, R., S. Subramani, P.G. Young. 1992. Checkpoint controls in *Schizosaccharomyces pombe*: RAD1. 1(4):1335–1342.
- Rozen, S. H. Skaletsky. 2000. Primer3 on the WWW for General Users and for Biologist Programmers. Method Mol Biol. 132:365–386.
- Ruiz-Verdugo, C.A., I.S. Racotta, A.M. Ibarra. 2001. Comparative biochemical composition in gonad and adductor muscle of triploid and diploid catarina scallop (*Argopecten ventricosus* Sowerby II, 1842). J Exp Mar Biol Ecol. 259(2):155–170.
- Ruiz-Verdugo, C.A., J.L. Ramírez, S. Allen. 2000. Triploid catarina scallop (*Argopecten ventricosus* Sowerby II, 1842): growth, gametogenesis, and suppression of functional hermaphroditism. Aquaculture. 186(1-2):13–32.
- Rupp, G.S., M.E. Valdéz-Ramírez, M. Leme da Fonseca. 2011. Biología y ecología. En: Maeda-Martínez A.N. and Ceijo-Lodeiros C. (eds.) Biología y cultivo de los moluscos pectínidos del género *Nodipecten*. Limusa - Centro de Investigaciones Biológicas del Noroeste, S.C., 27–54p.
- Saavedra, C., E. Bacherre 2006. Bivalve genomics. Aquaculture. 256:1–14.
- Sambrook, J. and D.W. Russell 2001. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press. Cold Spring Harbor.
- San-Segundo, P.A., G.S. Roeder 1999. Pch2 Links Chromatin Silencing to Meiotic Checkpoint Control. Cell. 97(3):313.
- San-Segundo, P.A., G.S. Roeder 2000. Role for the Silencing Protein Dot1 in Meiotic Checkpoint Control. Mol. Biol. Cell. 11(10):3601–3615.
- Sansam, C.L., J.L. Shepard, K. Lai, A. Ianari, P.S. Danielian, A. Amsterdam, N. Hopkins, J.A. Lees. 2006. DTL/CDT2 is essential for both CDT1 regulation and the early G2/M checkpoint. Gene Dev. 20(22):3117–29.
- Sarkar, A., C. Sim, Y.S. Hong, J.R. Hogan, M.J. Fraser, H.M. Robertson, F.H. Collins. 2003. Molecular evolutionary analysis of the widespread piggyBac transposon family and related “domesticated” sequences. Mol Genet Genomics. 270(2):173–80.
- Scarpa, J., J.E. Toro, K.T. Wada. 1994. Direct comparison of six methods to induce triploidy in bivalves. Aquaculture. 119(2-3):119–133.
- Schramm, G., I. Bruchhaus, T. Roeder 2000. A simple and reliable 5'-RACE approach. Nucl. Acids Res. 28(22):e96.
- Schumacher, J.M., K. Lee, S. Edelhoff, and R.E. Braun. 1995. Distribution of Tenr, an RNA-binding protein, in a lattice-like network within the spermatid nucleus in the mouse. Biol Reprod. 52(6):1274–83.
- Sellars, M.J., R.E. Lyons, P.M. Grewe, T. Vuocolo, L. Leeton, G.J. Coman, B.M. Degan, N.P. Preston. 2007. A PL10 vasa-like gene in the kuruma shrimp, *Marsupenaeus japonicus*, expressed during development and in adult gonad. Mar Biotechnol. 9(3):377–87.
- Shin, Y.-H., Y. Choi, S.U. Erdin, S.A. Yatsenko, M. Kloc, F. Yang, P.J. Wang, M.L.

- Meistrich, A. Rajkovic. 2010. Hormad1 mutation disrupts synaptonemal complex formation, recombination, and chromosome segregation in mammalian meiosis. *PLoS Genetics*. 6(11):e1001190.
- Shukla, J.N., J. Nagaraju 2010. Doublesex: a conserved downstream gene controlled by diverse upstream regulators. *J Genet*. 89(3):341–56.
- Silveira, E.D., M. Alves-Ferreira, L.A. Guimarães, F.R. Da Silva, V.T.D.C. Carneiro. 2009. Selection of reference genes for quantitative real-time PCR expression studies in the apomictic and sexual grass *Brachiaria brizantha*. *BMC Plant Biol*. 9:84
- Sloan, R., C. Swanson, L. Gavilano. 2012. Characterization of null and hypomorphic alleles of the *Drosophila* 1 (2) *dtl/cdt2* gene: Larval lethality and male fertility. *Fly*. 6:173–183.
- Slomovic, S., E. Fremder, R.H.G. Staals, G.J.M. Pruijn, G. Schuster. 2010. Addition of poly(A) and poly(A)-rich tails during RNA degradation in the cytoplasm of human cells. *PNAS*. 107(16):7407–12.
- Slomovic, S., D. Laufer, D. Geiger, G. Schuster. 2006. Polyadenylation of ribosomal RNA in human cells. *Nucl. Acids Res*. 34(10):2966–2975.
- Small, C.M., G.E. Carney, Q. Mo, M. Vannucci, A.G. Jones. 2009. A microarray analysis of sex- and gonad-biased gene expression in the zebrafish: evidence for masculinization of the transcriptome. *BMC Genomics*. 10:579
- Smith, J.T. 1991. Cenozoic giant pectinids from California and the Tertiary Caribbean Province: *Lyropecten*, “*Macrochlamys*”, *Vertipecten*, and *Nodipecten* species. United States Geological Survey Professional Paper. 1391:1–155.
- Song, L., W. Xu, C. Li, H. Li, L. Wu, J. Xiang, X. Guo 2006. Development of expressed sequence tags from the bay scallop, *Argopecten irradians irradians*. *Mar Biotechnol*. 8(2):161–169.
- Stoilov, P., R. Daoud, O. Nayler, S. Stamm 2004. Human tra2-beta1 autoregulates its protein concentration by influencing alternative splicing of its pre-mRNA. *Hum Mol Genet*. 13(5):509–24.
- Strong, E.R., J.C. Schimenti 2010. Evidence Implicating CCNB1IP1, a RING Domain-Containing Protein Required for Meiotic Crossing Over in Mice, as an E3 SUMO Ligase. *Genes*. 1(3):440–451.
- Su, H., X. Liu, W. Yan, T. Shi, X. Zhao, D.P. Blake, F.M. Tomley, X. Suo 2012. piggyBac transposon-mediated transgenesis in the apicomplexan parasite *Eimeria tenella*. *PloS One*. 7(6):e40075.
- Sun, S.C., N.H. Kim. 2012. Spindle assembly checkpoint and its regulators in meiosis. *Hum Rep* 18
- Tabarini, C.L. 1984. Induced triploidy in the bay scallop, *Argopecten irradians*, and its effects on growth and gametogenesis. *Aquaculture*. 42(2):151–160.
- Takeuchi, T. *et al.* 2012. Draft Genome of the Pearl Oyster *Pinctada fucata*: A Platform for Understanding Bivalve Biology. *DNA Res* 19:117–130.
- Talavera, G., J. Castresana. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst Biol*. 56(4):564–77.
- Tanguy, A. *et al.* 2008. Increasing genomic information in bivalves through new EST

- collections in four species: Development of new genetic markers for environmental studies and genome evolution. *Gene*. 408(1-2):27.
- Tarsounas, M., T. Morita, R.E. Pearlman, P.B. Moens. 1999. RAD51 and DMC1 form mixed complexes associated with mouse meiotic chromosome cores and synaptonemal complexes. *J Cell Biol*. 147(2):207–220.
- Telonis-Scott, M., A. Kopp, M.L. Wayne, S. V Nuzhdin, L.M. McIntyre. 2009. Sex-specific splicing in *Drosophila*: widespread occurrence, tissue specificity and evolutionary conservation. *Genetics*. 181(2):421–34.
- Thellin, O., W. Zorzi, B. Lakaye, B. De Borman, B. Coumans, G. Hennen, T. Grisar, *et al.* 1999. Housekeeping genes as internal standards: use and limits. *J Biotechnol*, 75(2-3):291–295.
- Theurkauf, W. H. Baum. 1986. Tissue-specific and constitutive alpha-tubulin genes of *Drosophila melanogaster* code for structurally distinct proteins. *PNAS*. 83(22):8477-8481.
- Thiriou-Quévieux, C. (2003). Advances in chromosomal studies of gastropod molluscs. *J Mollusc Stud*. 69:187–201.
- Thiriou-Quévieux, C., T. Neol, S. Bougrier, S. Dollot. 1988. Relationship between aneuploidy and growth rate in pair mating of the oyster, *Crassostrea gigas*. *Aquaculture*. 75:89–96.
- Toulza, E., M.S. Shin, G. Blanc, S.S. Audic, M. Laabir, Y. Collos, J.-M. Claverie, *et al.* 2010. Gene Expression in Proliferating Cells of the Dinoflagellate *Alexandrium catenella* (Dinophyceae). *Appl Environ Microbiol* 76(13), 4521–4529.
- Treen, N., N. Itoh, H. Miura, I. Kikuchi, T. Udea, K.G. Takahashi, T. Ubuka, K Yamamoto, P.J. Sharp, K Tsutsui, M. Osada. 2012. Mollusc gonadotropin-releasing hormone directly regulates gonadal functions: A primitive endocrine system controlling reproduction. *Gen Comp Endocrinol*. 176(2): 167-172.
- Tuiskunen, A., I. Leparco-Goffart, L. Boubis, V. Monteil, J. Klingström, H.J. Tolou, A. Lundkvist, S. Plumet. 2010. Self-priming of reverse transcriptase impairs strand-specific detection of dengue virus RNA. *J Gen Virol*. 91:1019–27.
- Türeci, O., U. Sahin, C. Zwick, M. Koslowski, G. Seitz, M. Pfreundschuh. 1998. Identification of a meiosis-specific protein as a member of the class of cancer/testis antigens. *PNAS* 95(9):5211–5216
- Vandesompele, J., K.D. Preter, F. Pattyn, B. Poppe, N. V Roy, A.D. Paepe, F. Speleman. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol*. 3(7):research0034.1 – 0034.11.
- Verhoeven, K.J.F., P.J. Van Dijk, A. Biere. 2010. Changes in genomic methylation patterns during the formation of triploid asexual dandelion lineages. *Mol Ecol* 19(2):315–324
- Villalba-Vega, J.N. 2011. Caracterización de bacterias asociadas a las lesiones de la concha de la almeja mano de león *Nodipecten subnodosus* (Sowerby, 1835) cultivada en Bahía Tortugas, Baja California Sur. Tesis de Maestría. Centro de Investigación Científica y de Educación Superior de Ensenada. Ensenada, México.
- Villasante, A., D. Wang, P. Dobner, P. Dolph, S.A. Lewis, N.J. Cowan. 1986. Six mouse alpha-tubulin mRNAs encode five distinct isotypes: testis-specific expression of two

- sister genes. *Mol Cell Biol.* 6(7):2409-2419.
- Wang, K., M. Wang, D. Tang, Y. Shen, C. Miao, Q. Hu, T. Lu, Z. Cheng 2012. The role of rice HEI10 in the formation of meiotic crossovers. *PLoS Genetics.* 8(7):e1002809.
- Wang, W., R.E. Chapin. 2000. Differential gene expression detected by suppression subtractive hybridization in the ethylene glycol monomethyl ether-induced testicular lesion. *Toxicol Sci.* 56(1):165–174
- Wang, X., Y. Zhao, K. Wong, P. Ehlers, Y. Kohara, S. Jones, M. Marra, R. Holt, D. Moerman, D. Hansen. 2009. Identification of genes expressed in the hermaphrodite germ line of *C. elegans* using SAGE. *BMC Genomics.* 10(1):213.
- Wang, Z., Y. Li, D. Kong, A. Ahmad, S. Banerjee, F.H. Sarkar 2010. Crosstalk between miRNA and Notch signaling pathways in tumor development and progression. *Cancer Lett.* 292(2):141–148.
- Ward, J.O., L.G. Reinholdt, W.W. Motley, L.M. Niswander, D.C. Deacon, L.B. Griffin, K.K. Langlais, V.L. Backus, K.J. Schimenti, M.J. O'Brien, J.J. Eppig, J.C. Schimenti. 2007. Mutation in mouse *hei10*, an e3 ubiquitin ligase, disrupts meiotic crossing over. *PLoS Genetics.* 3(8):e139.
- Weaver, T.L.O. 1995. Review Meiosis in *Drosophila*: Seeing is believing. 92:10443–10449.
- Wilkins, A.S., R. Holliday. 2009. The evolution of meiosis from mitosis. *Genetics.* 181(1):3–12.
- Wren, J.D., E. Forgacs, J.W. Fondon, A. Pertsemliadis, S.Y. Cheng, T. Gallardo, R.S. Williams, R. V Shohet, J.D. Minna, y H.R. Garner. 2000. Repeat polymorphisms within gene regions: phenotypic and evolutionary implications. *Amer J Hum Genet.* 67(2):345–56.
- Wyman, C., D. Ristic, R. Kanaar. 2004. Homologous recombination-mediated double-strand break repair. *DNA Repair.* 3(8-9):827–833.
- Xia, A.-H., Q.-X. Zhou, L.-L. Yu, W.-G. Li, Y.-Z. Yi, Y.-Z. Zhang, Z.-F. Zhang 2006. Identification and analysis of YELLOW protein family genes in the silkworm, *Bombyx mori*. *BMC Genomics.* 7:195.
- Xia, W., L. Zhou, B. Yao, C.J. Li, J.F. Gui. 2007. Differential and spermatogenic cell-specific expression of DMRT1 during sex reversal in protogynous hermaphroditic groupers. *Mol Cell Endocrinol,* 263(1-2):156–72
- Xue, J.-L., T.Z. Salem, C.M. Turney, y X.-W. Cheng 2010. Strategy of the use of 28S rRNA as a housekeeping gene in real-time quantitative PCR analysis of gene transcription in insect cells infected by viruses. *J Virol Method.* 163(2):210–215.
- Yamashita, M. 2000. Toward modeling of a general mechanism of MPF formation during oocyte maturation in vertebrates. *Zool Sci.* 17(7):841–851.
- Yan, W., L. Ma, K.H. Burns, M.M. Matzuk 2004. Haploinsufficiency of kelch-like protein homolog 10 causes infertility in male mice. *PNAS.* 101(20):7793–7798.
- Yang, H., F. Zhang, X. Guo. 2000. Triploid and tetraploid Zhikong scallop, *Chlamys farreri* Jones et Preston, produced by inhibiting polar body I. *Mar Biotechnol.* 2:466–475.
- Ye, Y., A. Godzik. 2004. FATCAT: a web server for flexible structure comparison and structure similarity searching. *Nucl Acids Res.* 32:W582–5.
- Yi, W., J. Ross. 2000. Mab-3 is a direct tra-1 target gene regulating diverse aspects of *C.*

- elegans* male sexual development and behavior. *Development*. 4480:4469–4480.
- Zamudio, N. D. Bourc'his 2010. Transposable elements in the mammalian germline: a comfortable niche or a deadly trap?. *Heredity*. 105(1):92–104.
- Zhang, B., M. Gallegos, A. Puoti, E. Durkin. 1997. A conserved RNA-binding protein that regulates sexual fates in the *C. elegans* hermaphrodite germ line. *Nature*. 390:477-484.
- Zhang, G. *et al.* 2012. The oyster genome reveals stress adaptation and complexity of shell formation. *Nature*. 490:49–54.
- Zhang, Z., M.A. Zariwala, M.M. Mahadevan, P. Caballero-Campo, X. Shen, E. Escudier, B. Duriez, A.-M. Bridoux, M. Leigh, G.L. Gerton, M. Kennedy, S. Amselem, M.R. Knowles, J.F. Strauss. 2007. A heterozygous mutation disrupting the SPAG16 gene results in biochemical instability of central apparatus components of the human sperm axoneme. *Biol Reprod*. 77(5):864–71.
- Zhong, Q., Q. Zhang, Z. Wang, J. Qi, Y. Chen, S. Li, Y. Sun, C. Li, X. Lan. 2008. Expression profiling and validation of potential reference genes during *Paralichthys olivaceus* embryogenesis. *Mar Biotechnol*. 10(3):310–318.
- Zhu, L.J. S.W. Altmann 2005. mRNA and 18S-RNA coapplication-reverse transcription for quantitative gene expression analysis. *Anal Biochem* 345(1):102–109.
- Zhuang, Y., H. Zhang, S. Lin 2013. Polyadenylation of 18S rRNA in algae. *J Phycol.* in press.
- Zouros, E., C. Thiriot-Quievreux, G. Kotoulas. 1996. The negative correlation between somatic aneuploidy and growth in the oyster *Crassostrea gigas* and implications for the effects of induced polyploidization. *Genet Res*. 68(2):109–116.

**ANEXO I.**

SECUENCIAS AISLADAS EN ESTE TRABAJO.

Secuencias parciales y alineamientos nucleotídicos obtenidas para el fragmento de *ARN ribosomal 18S* y para el transcrito *act-* .

```
>Nodipecten subnodosus ribosomal RNA, partial sequence (168 pb)
GAAATCTTTGGATCGCCGTAAGACGAACTACTGCGAAAGCATTGGCCAAGAATGTTTCATTAATCAAGATCGAAAG
TCAGAGGTTTCAAGACGATCAGATACCGTCGTAGTTCTGACCATAAACGATGCCAACTAGCGATCCGCCGGAGTTGC
TTCAATGACTCGGC
```

CLUSTAL 2.0.12 multiple sequence alignment

```
Pecten_jacobaeus|AY070112.1      GGTGTTAGAGGTGAAATCTTTGGATCGCCGTAAGACGAACTACTGCGAAA
Pecten_maximus|L49053.1         GGTGTTAGAGGTGAAATCTTTGGATCGCCGTAAGACGAACTACTGCGAAA
Chlamys_farreri|AF526252.1     GGTGTTAGAGGTGAAATCTTTGGATCGCCGTAAGACGAACTACTGCGAAA
-----GAAATCTTTGGATCGCCGTAAGACGAACTACTGCGAAA
Nodipecten_subnodosus          GGTGTTAGAGGTGAAATCTTTGGATCGCCGTAAGACGAACTACTGCGAAA
Mimachlamys_varia|DQ279939.1   GGTGTTAGAGGTGAAATCTTTGGATCGCCGTAAGACGAACTACTGCGAAA
*****
```

```
Pecten_jacobaeus|AY070112.1      GCATTTGCCAAGAATGTTTTTCATTAATCAAGATCGAAAGTCAGAGGTTTCG
Pecten_maximus|L49053.1         GCATTTGCCAAGAATGTTTTTCATTAATCAAGATCGAAAGTCAGAGGTTTCG
Chlamys_farreri|AF526252.1     GCATTTGCCAAGAATGTTTTTCATTAATCAAGATCGAAAGTCAGAGGTTTCG
Nodipecten_subnodosus          GCATTTGCCAAGAATGTTTTTCATTAATCAAGATCGAAAGTCAGAGGTTTCG
Mimachlamys_varia|DQ279939.1   GCATTTGCCAAGAATGTTTTTCATTAATCAAGATCGAAAGTCAGAGGTTTCG
*****
```

```
Pecten_jacobaeus|AY070112.1      AAGACGATCAGATACCGTCGTAGTTCTGACCATAAACGATGCCAACTAGC
Pecten_maximus|L49053.1         AAGACGATCAGATACCGTCGTAGTTCTGACCATAAACGATGCCAACTAGC
Chlamys_farreri|AF526252.1     AAGACGATCAGATACCGTCGTAGTTCTGACCATAAACGATGCCAACTAGC
Nodipecten_subnodosus          AAGACGATCAGATACCGTCGTAGTTCTGACCATAAACGATGCCAACTAGC
Mimachlamys_varia|DQ279939.1   AAGACGATCAGATACCGTCGTAGTTCTGACCATAAACGATGCCAACTAGC
*****
```

```
Pecten_jacobaeus|AY070112.1      GATCCGCCGGAGTTGCTTCAATGACTCGGCGGGCAGCTTACGGGAAACCA
Pecten_maximus|L49053.1         GATCCGCCGGAGTTGCTTCAATGACTCGGCGGGCAGCTTACGGGAAACCA
Chlamys_farreri|AF526252.1     GATCCGCCGGAGTTGCTTCAATGACTCGGCGGGCAGCTTACGGGAAACCA
Nodipecten_subnodosus          GATCCGCCGGAGTTGCTTCAATGACTCGGC-----
Mimachlamys_varia|DQ279939.1   GATCCGCCGGAGTTGCTTCAATGACTCGGCGGGCAGCTTACGGGAAACCA
*****
```

//



Secuencias completas de ADNc traducidas caracterizadas por RACE y depositadas en la  
base de datos pública GenBank - NCBI (fecha de liberación: 01/Dic/2013)

LOCUS seq1 1274 bp mRNA linear INV 15-OCT-2012  
DEFINITION [Nodipecten subnodosus] DMCl/LIM15 homolog.  
ACCESSION KC119408

VERSION

KEYWORDS .

SOURCE Nodipecten subnodosus

ORGANISM Nodipecten subnodosus

Eukaryota; Metazoa; Lophotrochozoa; Mollusca; Bivalvia;  
Pteriormorphia; Pectinoidea; Pectinoidea; Pectinidae; Nodipecten.

REFERENCE 1 (bases 1 to 1274)

AUTHORS Llera-Herrera,R., Garcia-Gasca,A. and Ibarra,A.M.

TITLE Identification and isolation of spermatogenesis genes from the  
scallop Nodipecten subnodosus by suppressive subtractive  
hybridization and 454-FLX sequencing

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1274)

AUTHORS Llera-Herrera,R., Garcia-Gasca,A. and Ibarra,A.M.

TITLE Direct Submission

JOURNAL Submitted (15-OCT-2012) Aquaculture Genetics and Breeding  
Laboratory, Centro de Investigaciones Biologicas del Noroeste,  
S.C., Instituto Politecnico Nacional, No. 195, Playa Palo de Santa  
Rita, La Paz, BCS 23096, Mexico

COMMENT Bankit Comment: ALT EMAIL:aibarra@cibnor.mx.

Bankit Comment: TOTAL # OF SEQS:4.

##Genome-Assembly-Data-START##

Sequencing Technology :: Sanger dideoxy sequencing

Assembly Method :: CAP3 v. 32-bit Linux version

##Genome-Assembly-Data-END##

FEATURES Location/Qualifiers

source

1..1274

/organism="Nodipecten subnodosus"

/mol\_type="mRNA"

/db\_xref="taxon:330909"

mRNA

1..1274

5'UTR

1..54

CDS

55..1080

/codon\_start=1

/translation="MESEDQVVEEVLGNEDDDESLEFQDIDMLQNHGINVADIKKQV  
GICTIKGIQMTTKRCLCNIGISEAKMEKIKEASGKLC DAGFLTALEYS DKR KQV FRI  
STGSQEFDKLIGGGLESMSITEAFGEFRTGKTQLSHTLCVTTQLPGANGYTGKVMFI  
DTENTFRPDRLRAIADRFNLDQSAVLNDNVLYARAYTSEHQFELLDFVAAKFYEPPGVF  
KLLIIDSIMVLFVRVDFSGRGELVDRQQLAQML SRLQKISEEYNVAVFVTNQMTADPG  
ATMSFQADPKPIGGNILAHASTTRLSLRKGRGENRIAKIYDSPDLPENEATYAITAG  
GIADAKE"

3'UTR

1081..1274

BASE COUNT 466 a 212 c 269 g 327 t

//

LOCUS seq2 1417 bp mRNA linear INV 15-OCT-2012

DEFINITION [Nodipecten subnodosus] DNA repair protein RAD51 homolog.

ACCESSION KC1194089

VERSION

KEYWORDS .

SOURCE Nodipecten subnodosus

ORGANISM Nodipecten subnodosus

Eukaryota; Metazoa; Lophotrochozoa; Mollusca; Bivalvia;  
Pteriomorphia; Pectinoidea; Pectinoidea; Pectinidae; Nodipecten.

REFERENCE 1 (bases 1 to 1417)  
AUTHORS Llera-Herrera,R., Garcia-Gasca,A. and Ibarra,A.M.  
TITLE Identification and isolation of spermatogenesis genes from the scallop Nodipecten subnodosus by suppressive subtractive hybridization and 454-FLX sequencing

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1417)  
AUTHORS Llera-Herrera,R., Garcia-Gasca,A. and Ibarra,A.M.  
TITLE Direct Submission  
JOURNAL Submitted (15-OCT-2012) Aquaculture Genetics and Breeding Laboratory, Centro de Investigaciones Biologicas del Noroeste, S.C., Instituto Politecnico Nacional, No. 195, Playa Palo de Santa Rita, La Paz, BCS 23096, Mexico

COMMENT Bankit Comment: ALT EMAIL:aibarra@cibnor.mx.  
Bankit Comment: TOTAL # OF SEQS:4.  
##Genome-Assembly-Data-START##  
Sequencing Technology :: Sanger dideoxy sequencing  
Assembly Method :: CAP3 v. 32-bit Linux version  
##Genome-Assembly-Data-END##

FEATURES Location/Qualifiers  
source 1..1417  
/organism="Nodipecten subnodosus"  
/mol\_type="mRNA"  
/db\_xref="taxon:330909"  
mRNA 1..1417  
5'UTR 1..198  
CDS 199..1281  
/codon\_start=1  
/translation="MQTQRELSSYPIAPVHKTCLVAAGFLIVEDLKGMPSELSKETG  
IGMEESLEILKLVLTGTDGKGDNSGQVDKSALDILKQEQSLPYIVTFSEQLDNMLGGGI  
PLCKITEFCGAPGTGKTQICMQLAVDAQIPECLGGLEAEVVYIDTEGSFIVERLVDIV  
TASVEHCKQIAGGEQDSEDLQGFTPENVLSRIHYRCHDYIELLATIHLPLDFVKHHS  
KVRLVIVDSVAFHFRHDFDDMSLRTRLLTSTAQNLIKLATYQLAVVLTNQMTTKVRM  
GEASQLVPALGESWGHASTIRVILYWEGQKRYAWLYKSPSHREARVPYQVTMGGIRDV  
VTEESNDNDNDINTEENPTKRQRIT"  
3'UTR 1282..1417

BASE COUNT 462 a 283 c 331 g 341 t

//

LOCUS seq3 1440 bp mRNA linear INV 15-OCT-2012  
DEFINITION [Nodipecten subnodosus] DNA repair protein RAD51 homolog, extended 3'-UTR.  
ACCESSION KC119410  
VERSION  
KEYWORDS .  
SOURCE Nodipecten subnodosus  
ORGANISM Nodipecten subnodosus  
Eukaryota; Metazoa; Lophotrochozoa; Mollusca; Bivalvia;  
Pteriomorphia; Pectinoidea; Pectinoidea; Pectinidae; Nodipecten.

REFERENCE 1 (bases 1 to 1440)  
AUTHORS Llera-Herrera,R., Garcia-Gasca,A. and Ibarra,A.M.  
TITLE Identification and isolation of spermatogenesis genes from the scallop Nodipecten subnodosus by suppressive subtractive hybridization and 454-FLX sequencing

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1440)  
AUTHORS Llera-Herrera,R., Garcia-Gasca,A. and Ibarra,A.M.  
TITLE Direct Submission  
JOURNAL Submitted (15-OCT-2012) Aquaculture Genetics and Breeding

Laboratory, Centro de Investigaciones Biologicas del Noroeste,  
S.C., Instituto Politecnico Nacional, No. 195, Playa Palo de Santa  
Rita, La Paz, BCS 23096, Mexico

COMMENT Bankit Comment: ALT EMAIL:aibarra@cibnor.mx.  
Bankit Comment: TOTAL # OF SEQS:4.  
##Genome-Assembly-Data-START##  
Sequencing Technology :: Sanger dideoxy sequencing  
Assembly Method :: CAP3 v. 32-bit Linux version  
##Genome-Assembly-Data-END##

FEATURES Location/Qualifiers  
source 1..1440  
/organism="Nodipecten subnodosus"  
/mol\_type="mRNA"  
/db\_xref="taxon:330909"  
mRNA 1..1440  
5'UTR 1..198  
CDS 199..1281  
/codon\_start=1  
/translation="MQTQRELSSYPIAPVHKTKLVAAGFLIVEDLKGMPSELSKETG  
IGMEESLEILKLVLTGTDGKGDNSGQVDKSAKDILKQEQSLPYIVTFSEQLDNMLGGGI  
PLCKITEFCGAPGTGKTQICMQLAVDAQIPECLGGLEAEVVYIDTEGSFIVERLVDIV  
TASVEHCKQIAGGEQDSEDLQGFTPENVLSRIHYRCHDYIELLATIHLLPDFVKHHS  
KVRLVIVDSVAFHFRHDFDMSLRTRLLTSTAQNLIKLATYQLAVVLTNQMTTKVRM  
GEASQLVPALGESWGHASTIRVILYWEGQKRYAWLYKSPSHREARVPYQVTMGGIRDV  
VTEESNDDNDINTEENPTKRQRIT"  
3'UTR 1282..1440

BASE COUNT 470 a 288 c 334 g 348 t  
//

LOCUS seq4 984 bp mRNA linear INV 15-OCT-2012  
DEFINITION [Nodipecten subnodosus] Synaptonemal complex protein 3.  
ACCESSION KC119411  
VERSION  
KEYWORDS .  
SOURCE Nodipecten subnodosus  
ORGANISM Nodipecten subnodosus  
Eukaryota; Metazoa; Lophotrochozoa; Mollusca; Bivalvia;  
Pteriomorphia; Pectinoidea; Pectinoidea; Pectinidae; Nodipecten.

REFERENCE 1 (bases 1 to 984)  
AUTHORS Llera-Herrera,R., Garcia-Gasca,A. and Ibarra,A.M.  
TITLE Identification and isolation of spermatogenesis genes from the  
scallop Nodipecten subnodosus by suppressive subtractive  
hybridization and 454-FLX sequencing  
JOURNAL Unpublished

REFERENCE 2 (bases 1 to 984)  
AUTHORS Llera-Herrera,R., Garcia-Gasca,A. and Ibarra,A.M.  
TITLE Direct Submission  
JOURNAL Submitted (15-OCT-2012) Aquaculture Genetics and Breeding  
Laboratory, Centro de Investigaciones Biologicas del Noroeste,  
S.C., Instituto Politecnico Nacional, No. 195, Playa Palo de Santa  
Rita, La Paz, BCS 23096, Mexico

COMMENT Bankit Comment: ALT EMAIL:aibarra@cibnor.mx.  
Bankit Comment: TOTAL # OF SEQS:4.  
##Genome-Assembly-Data-START##  
Sequencing Technology :: Sanger dideoxy sequencing  
Assembly Method :: CAP3 v. 32-bit Linux version  
##Genome-Assembly-Data-END##

FEATURES Location/Qualifiers  
source 1..984  
/organism="Nodipecten subnodosus"  
/mol\_type="mRNA"

```

/db_xref="taxon:330909"
mRNA 1..984
5'UTR 1..52
CDS 53..778
/codon_start=1
/translation="MPPARAQKKNNTGSPNQKQQTkskiELDDHDDDDIESREHSL
DSSPARDETPiIPRQGKKRAAEDPGEFNTEMQKMLECFGADITKTLVnKRKRLEQFTQ
NSLKASnKKVEEiWKMQQLERgKLHTEYCRQVNTVFNQWEADLEKTKEQEKLVTFLK
QQQKLFQQARVVQSRLKtIRQLHEQYTKGMEDLGTCHNSQQSNVQSELKKEMTLQK
KILMDTQQQEMANVRKSLQTMLF"
3'UTR 779..984
BASE COUNT 394 a 173 c 203 g 214 t
//
LOCUS seq1 1870 bp mRNA linear INV 14-APR-2013
DEFINITION Pachytene checkpoint protein 2 homolog (pch2), mRNA.
ACCESSION KC904225
VERSION
KEYWORDS .
SOURCE Nodipecten subnodosus
ORGANISM Nodipecten subnodosus
Eukaryota; Metazoa; Lophotrochozoa; Mollusca; Bivalvia;
Pteriomorphia; Pectinoida; Pectinoidea; Pectinidae; Nodipecten.
REFERENCE 1 (bases 1 to 1870)
AUTHORS Llera-Herrera,R., Garcia-Gasca,A. and Ibarra,A.M.
TITLE Identification and isolation of spermatogenesis genes from the
scallop Nodipecten subnodosus by suppressive subtraction
hybridization and 454-FLX sequencing
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1870)
AUTHORS Llera-Herrera,R., Garcia-Gasca,A. and Ibarra,A.M.
TITLE Direct Submission
JOURNAL Submitted (14-APR-2013) Aquaculture Program, Centro de
Investigaciones Biologicas del Noroeste, Mar Bermejo 195, La Paz,
BCS 23090, Mexico
COMMENT Bankit Comment: ALT EMAIL:raul.llera@gmail.com.
Bankit Comment: TOTAL # OF SEQs:1.
FEATURES
source Location/Qualifiers
1..1870
/organism="Nodipecten subnodosus"
/mol_type="mRNA"
/db_xref="taxon:330909"
gene 275..1579
/gene="PCH2"
/gene_synonym="TRIP-13"
CDS 275..1579
/gene="PCH2"
/gene_synonym="TRIP-13"
/codon_start=1
/gene_synonym="TRIP-13"
/translation="MSSITMDQVSELGDGHSdstAISpvyHIEICQSPSSTAQGNIVK
EHVLNLLHRQkVAYGDFVlTEfDDPFLHEHVNSIALCDSdLSGSNRQTIDLCGRELQL
HVFQLQEDGPGIEELDEEDLAAASHWLLPARDLHGmWDSLvFDENVKSQLLNyATTTL
LFSDRQVDSNIITWnKVLLHGPPGTGKtSLCKALAqKLVVRLSDrySYGQLIEINSH
SLFSKWFSESGKLVmKMFQKIqELIDdKEAIVFVLIDEVESLTSARKsAMSGSEPSDA
IRVVNALLTQLDQIKRQPNVMILtTsnVTGAIDLAFVDRADIKQYIGPPSPAAIFKIF
HSCINELMRVDIISPAQQLDLRGLeAMRFmENDATKNSLVLrDIANKSHGLSGRTLR
KLpFIAHAMFVQgASVtLEDFLVALSRGvDRQFQEREDLTKE"
BASE COUNT 550 a 377 c 426 g 517 t

```

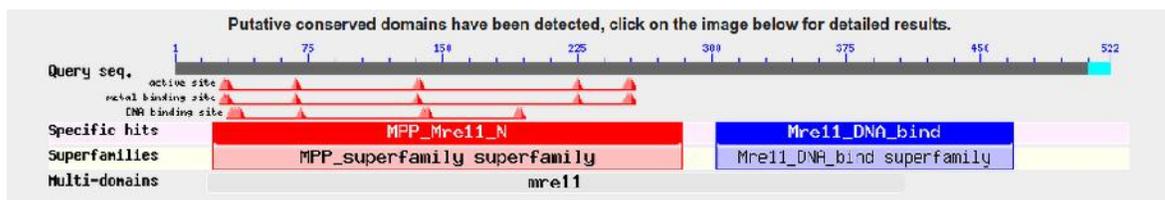
## Secuencia traducida de ADNc del transcrito *mre11* obtenida por RACE, y análisis por blastp contra bases de datos públicas de proteínas y dominios conservados

LOCUS Mre11 1588 bp linear 17-APR-2013  
 DEFINITION Nodipecten subnodosus MRE11, partial  
 ACCESSION Mre11  
 VERSION  
 KEYWORDS .  
 SOURCE Unknown.  
 ORGANISM Nodipecten subnodosus.  
 Unclassified.

**FEATURES** Location/Qualifiers  
 source 1..1588  
 CDS 23..1588  
 /note="predicted coding region"  
 /translation="LTKKQYPSIIMPQTEDESDILKIIIVATDIHLGYAEKDGIRGNDS  
 LVAFEEILENAKNHDVDFILLGGDLFHENKPPRRIMHGICISLLRKYCMGDKPVQFEYL  
 SDQSKDFNHCQFPYLNIEDPNLNISIPVFSIHGNHDDPTGQGNLCTLDILHTSGFVNY  
 FGKHTSLEKIEMSPLLLQRGSTKCLAMYGLGSIRDERLHRMFLKKTVTMLRPRENPDEW  
 FNMFVIHQNRSKHTTTNYIPEQFLDDFDLVIWGHEHECRIEPEWNSVQNFYVSQPGS  
 SVATSLSEGETVKKKHIGLLQIKGKNFKMTKIPLETVRPFYMEDVVLSDTSRNPQDHSI  
 AKKVEAFCSEKVEGLLEKAESEHSGNGKQPDMLIRLRIDYSGGFEPFSAHRFGQKFI  
 DKVANPKDMIQFQRKKVISTKNDKEDKDALGAIKADTLDTSRVEDMVKDYFSSSDCK  
 MQLKLLTEKGMGQAVQEFVDKKEERAISELVKYQLQKTQTYLKQRNANEEIIDAQVTR  
 YKEEKKKKKKKKKK"

BASE COUNT 566 a 273 c 340 g 409 t

### Detección de dominios conservados para la traducción conceptual de *mre11* (*Conserved domains*; NCBI)



## Búsqueda mediante blastp para la traducción conceptual de *mre11* contra base de datos de proteínas (nr-GenBank)

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenPept Graphics Distance tree of results Multiple alignment

Description	Max score	Total score	Query cover	E value	Max ident	Accession
<input type="checkbox"/> Double-strand break repair protein MRE11 [Crassostrea gigas]	837	837	95%	0.0	78%	<a href="#">CNC_9294.1</a>
<input type="checkbox"/> PREDICTED: double-strand break repair protein MRE11A-like [Strongylocentrotus purpuratus]	683	683	95%	0.0	62%	<a href="#">XP_798167.3</a>
<input type="checkbox"/> PREDICTED: double-strand break repair protein MRE11A [Orcinus orca]	682	682	96%	0.0	63%	<a href="#">XP_004265514.1</a>
<input type="checkbox"/> Double-strand break repair protein MRE11, partial [Columba livia]	678	678	95%	0.0	63%	<a href="#">EMC88261.1</a>
<input type="checkbox"/> PREDICTED: double-strand break repair protein MRE11A [Ovis aries]	675	675	95%	0.0	64%	<a href="#">XP_004016004.1</a>
<input type="checkbox"/> PREDICTED: double-strand break repair protein MRE11A isoform 2 [Callithrix jacchus]	676	676	95%	0.0	63%	<a href="#">XP_002754717.1</a>
<input type="checkbox"/> PREDICTED: double-strand break repair protein MRE11-like [Meleagris gallopavo]	674	674	95%	0.0	62%	<a href="#">XP_003203558.1</a>
<input type="checkbox"/> PREDICTED: double-strand break repair protein MRE11A [Callithrix jacchus]	674	674	95%	0.0	63%	<a href="#">XP_003734104.1</a>
<input type="checkbox"/> Double-strand break repair protein MRE11A [Bos grunniens mutus]	674	674	95%	0.0	64%	<a href="#">ELR48187.1</a>
<input type="checkbox"/> PREDICTED: double-strand break repair protein MRE11A [Odobenus rosmarus divergens]	673	673	95%	0.0	63%	<a href="#">XP_004412484.1</a>



