



FACULTAD DE CIENCIAS DEL MAR Y AMBIENTALES

**LOCALIZACIÓN DE miRNAs EN EL GENOMA DEL LENGUADO
SENEGALÉS (*Solea senegalensis*, Kaup 1858)**

Daniel Ramírez Torres

Tutor/es: Laureana Rebordinos González

Alberto Arias Pérez

Puerto Real a 19 de septiembre de 2017

TRABAJO FIN DE MÁSTER (PERFIL INVESTIGADOR)
MÁSTER ACUICULTURA Y PESCA

**LOCALIZACIÓN DE miRNAs EN EL GENOMA DEL
LENGUADO SENEGALÉS (*Solea senegalensis*, Kaup 1858).**

Memoria presentada por Daniel Ramírez Torres para la obtención del
título de Máster en Acuicultura y Pesca
(Perfil Investigador)

Firma del alumno

Fdo.: Daniel Ramírez Torres

Puerto Real a 19 de septiembre de 2017

TRABAJO FIN DE MÁSTER (PERFIL INVESTIGADOR)
MÁSTER ACUICULTURA Y PESCA

Laureana Rebordinos González, Doctora en Ciencias Biológicas, y Alberto Arias Pérez, Doctor en Biología, como tutores del Trabajo Fin de Máster titulado “LOCALIZACIÓN DE miRNAs EN EL GENOMA DEL LENGUADO SENEGALÉS (*Solea senegalensis*, Kaup 1858)”, realizada por Daniel Ramírez Torres, hacen constar que la presente memoria se ha llevado a cabo bajo su propia dirección en las dependencias del Departamento de Biomedicina, Biotecnología y Salud Pública y autorizan su presentación y defensa.

Para que así conste firman el presente informe en Puerto Real a 19 de septiembre de 2017.

Firma del/los Director/es

Fdo.: Laureana Rebordinos González

Fdo.: Alberto Arias Pérez

TRABAJO FIN DE MÁSTER (PERFIL INVESTIGADOR)
MÁSTER ACUICULTURA Y PESCA

Resumen

Actualmente existe un creciente interés en diversificar el cultivo de peces y su comercialización. Es por esto que en la región sur atlántica española gran parte de los esfuerzos y proyectos realizados se enfocan en nuevas especies de peces comerciales, entre las que se incluye el lenguado senegalés. Hoy en día se desarrollan numerosos proyectos de investigación con el objetivo de conseguir llevar a cabo la producción de *Solea senegalensis*, la cual está limitada dada la dificultad para conseguir la reproducción de esta especie en cautividad. Las especies como el lenguado senegalés junto al lenguado común, el rodaballo o el fletan muestran un gran interés en acuicultura debido a sus características; presentando un alto precio en mercado y acrecentando el interés en su cultivo. En la actualidad, teniendo en cuenta que el cultivo del lenguado senegalés no se ha conseguido de forma integral y dado que el control de su reproducción no es tan conocido como el de otras especies cultivadas es de interés la realización de un estudio que permita ampliar los conocimientos sobre la naturaleza de esta especie con el fin de aportar información que pueda contribuir al desarrollo de la reproducción y producción de esta especie. El presente proyecto bioinformático pretende un estudio de *Solea senegalensis* desde un punto de vista genético con el objetivo principal profundizar en la caracterización del genoma del lenguado senegalés a través del estudio de datos obtenidos de genotecas BAC y de microRNAs, los cuales toman parte en la regulación de la expresión génica y cuyo estudio ayudaría a conocer mejor los mecanismos que regulan determinados aspectos de interés como son el crecimiento o la reproducción, procesos fundamentales en acuicultura. Tras la creación de una base de datos para *S. senegalensis* y la realización de comparaciones entre grupos de secuencias con herramientas Blast[®], recurriendo a información sobre secuencias miRNA maduras ya descritas, se obtuvo como resultado un total de 260 secuencias con similitud a miRNAs identificados previamente en otras especies, llevando a cabo un estudio de la relación de estas secuencias con los miRNAs con los que presentaron homología, describiendo las posibles funciones de los mismos.

Palabras Clave

Solea senegalensis, acuicultura, genoma, microRNA, bioinformática, genética.

Abstract

Currently there is a growing interest in diversifying fish aquaculture and its commercialization. As far as the Spanish Atlantic region is concerned, much of the effort and projects are devoted to new species of commercial fishes, including *Solea senegalensis*. Nowadays, research projects are being carried out with the aim of achieve the production of senegalese sole, which is limited due to the difficulty of reproduction in captivity. Species such as senegalese sole along with common sole, turbot or halibut show great interest in aquaculture due to its characteristics, presenting a high price in the market and increasing the interest in its cultivation. Nowadays, considering that the senegalese sole has not been fully cultivated since the control of its reproduction is not as well known as that of other cultivated species, it is of interest to carry out a study in order to increase the knowledge about the nature of this species that can contribute to the development of its reproduction and production. The present bioinformatic project aims to a study of *Solea senegalensis* from a genetic point of view with the main objective to deepen the characterization of the senegalese sole genome through the study of BAC and microRNAs libraries data, which take part in the gene expression regulation and its study to better understand the mechanisms that regulate the parameters of interest such as growth or reproduction, fundamental processes in aquaculture. Following the creation of a *S. Senegalensis* database and comparisons between groups of sequences using Blast® tools, using information on mature miRNA sequences already described, a total of 260 sequences with similarity to identified miRNAs previously in other species were obtained, carrying out a study of the relation of these sequences with miRNAs which homology was presented, describing their possible functions.

Keywords

Solea senegalensis, aquaculture, genome, microRNA, bioinformatics, genetics.

ÍNDICE

1. Introducción	2
1.1. Taxonomía	2
1.2. Biología de <i>Solea senegalensis</i>	3
1.3. Interés socio-económico y producción	7
1.4. Estudio de microRNAs	10
1.5. Genoteca BAC de <i>Solea senegalensis</i>	12
1.6. Bases de datos genómicas	13
2. Antecedentes y estado actual del tema	14
3. Justificación y objetivos	15
4. Material y métodos	16
4.1. Base de datos de <i>Solea senegalensis</i>	16
4.2. Obtención de secuencias de microRNA	17
4.3. Alineamiento de secuencias e identificación de miRNAs	18
4.4. Análisis de estructuras secundarias	19
4.5. Estudio de la distribución de secuencias	21
5. Resultados y discusión	22
5.1. MiRNAs descritos en peces	23
5.2. MiRNAs descritos en otras especies	24
5.3. Estructuras secundarias	28
5.4. Distribución de secuencias	29
6. Conclusión	32
7. Bibliografía	33

1. Introducción

Contexto biológico y socio-económico de *Solea senegalensis*

1.1. Taxonomía

El lenguado senegalés (*Solea senegalensis* Kaup, 1858), también conocido como lenguado del sur, es un pez plano de cuerpo ovalado y alargado común en las aguas del Mediterráneo y el Atlántico Sur (Vega y otros, 2002). Se trata de un pez óseo de la clase de los teleósteos (Figura 1) perteneciente al orden de los pleuronectiformes, dentro del cual corresponde al género *Solea* de la familia *Soleidae*. La gran mayoría de especies pertenecientes a este orden habitan en ambientes marinos aunque existen algunas especies de agua dulce (Arellano y Sarasquete, 2006).

Reino: Animalia

Filo: Chordata

Subfilo: Vertebrata

Superclase: Osteichthyes

Infraclase: Teleostei

Superorden: Acanthopterygii

Orden: Pleuronectiformes

Suborden: Pleuronectoidei

Familia: Soleidae

Género: *Solea*

Especie: *Solea senegalensis*

Figura 1. Taxonomía de *Solea senegalensis*

Los pleuronectiformes, también conocidos como peces planos, presentan simetría bilateral únicamente en las primeras fases de vida; los adultos no presentan esta simetría ya que estas especies sufren una importante metamorfosis que resulta en la migración de los órganos y uno de los ojos a uno de los flancos del cuerpo (Figura 2). Esta metamorfosis se acompaña de una serie de drásticos cambios como la remodelación del cráneo, la boca y de los orificios operculares así como de las aletas. Se produce una pérdida de la pigmentación

del flanco izquierdo (lado ciego), el cual queda orientado al bentos, mientras que la cara ocular adquiere diferentes tonalidades de marrón, lo que proporciona una gran capacidad de mimetismo en fondos arenosos (Arellano y Sarasquete, 2006; Arjona y otros, 2008). Esta serie de modificaciones supone un cambio tanto conductual como de los hábitos de vida de los individuos.

En el caso de *Solea senegalensis*, la metamorfosis se experimenta después de los primeros 10-12 días de vida; Tras este proceso de aproximadamente una semana de duración, el cuerpo permaneceligeramente redondeado en el flanco donde se localizan los ojos (flanco derecho) y plano en el lado ciego (Rodríguez y Peleteiro, 2014). Los ojos, pequeños y próximos, pueden sobresalir por encima de la superficie del cuerpo, lo que permite a los individuos ver cuando están enterrados en el sustrato (FAO, 2017). Esta especie se caracteriza porque las membranas interradiales de la aleta pectoral del lado ocular son de color negro, un carácter que lo diferencia del lenguado común (*Solea solea*, Linnaeus 1758) el cual presenta en dicha aleta una mancha negra en su mitad posterior (Cañavate, 2005).



Figura 2. Detalle en la aleta pectoral de un espécimen de *Solea senegalensis*. **Fuente:** FAO, 2017

1.2. Biología de *Solea senegalensis*

El lenguado senegalés se trata de una especie marina bentónica que, en el caso de los individuos adultos, suele encontrarse a profundidades de entre 12 y 65 metros mientras que los juveniles se mantienen en profundidades más rasas y en estuarios (Desoutter, 1990).

Habita generalmente en fondos arenosos o fangosos en zonas costeras y en lagos salobres. Poseen pocos y rudimentarios dientes localizados en el lado ciego de la boca, en ambas mandíbulas. Presenta hábitos nocturnos y se alimenta básicamente de invertebrados bentónicos, como larvas de poliquetos y moluscos bivalvos así como de pequeños crustáceos incluso pequeños peces (Desoutter, 1990). Pueden alcanzar los 60 centímetros de longitud en el Atlántico y 45 centímetros en el Mediterráneo así como un peso de 3 kilogramos, siendo las hembras de mayor tamaño que los machos (Rodríguez y Peleteiro, 2014).

Dado que se trata de un pez euritermo y eurihalino, es frecuente su presencia en estuarios y lagunas salobres que comunican con el mar (Imsland y otros, 2003). Su distribución en el Atlántico abarca desde el Golfo de Vizcaya hasta las costas de Senegal como límite sur y las Islas Canarias como límite occidental (Figura 3). En el Mediterráneo Occidental su presencia es menos frecuente (Férrandez y Rodríguez, 2003).



Figura 3. Mapa de distribución del lenguado senegalés. **Fuente:** Férrandez y Rodríguez, 2003

En cuanto a los aspectos reproductivos, cabe destacar que el lenguado senegalés es una especie gonocórica que presenta fecundación externa. Los individuos de diferente sexo no presentan dimorfismo sexual aparente por lo que su sexado no es fácil y puede ser necesario el uso de sistemas de ecografía para la identificación de ovarios o testículos (Rodríguez y Peleteiro, 2014). La diferencia entre sexos solo es evidente en las épocas de puesta, en las cuales las hembras maduras muestran un notable desarrollo gonadal que permite

diferenciarlas externamente de los machos (Anguis y Cañavate, 2005). El macho es más difícil de identificar dado que los testículos generalmente son muy pequeños y, aunque pueden ser fluyentes a lo largo de todo el año, producen un escaso volumen de eyaculado por lo que la madurez no es evidente como en el caso de la hembra.

La madurez se alcanza entre los dos y tres años en el caso de los machos, y entre los tres y cuatro en las hembras (Rodríguez, 1984), cuando los individuos han alcanzado un tamaño de aproximadamente de 30 centímetros de longitud (Rodríguez y Peleteiro, 2014). Las puestas, influenciadas por la temperatura, en el medio natural se realizan durante los meses de marzo a junio a temperaturas entre 15 y 22 °C (Anguis y Cañavate, 2005). Los huevos de *S. senegalensis* (Figura 4) poseen un tamaño de 1 milímetro de diámetro aproximadamente con numerosos aglomerados de glóbulos lipídicos en el interior (Anguis y Cañavate, 2005). Son huevos pelágicos no adhesivos que no reciben cuidados parentales y de los que, pasadas las 48 horas de la puesta, eclosionan larvas pelágicas de simetría bilateral y hábitos diurnos. Durante esta fase las larvas permanecen a lo largo de la columna de agua entre los 10 y 15 metros de profundidad hasta alcanzar una longitud en torno a 2-4 milímetros, después de los 10-12 días de vida. Es entonces cuando estas larvas, en zonas cercanas a la costa, experimentan el proceso de metamorfosis. Este proceso durará una semana al final de la cual se habrá producido el desplazamiento del ojo izquierdo al costado derecho y es entonces cuando la larva, con una longitud entorno a 7,3 milímetros (Dinis y otros, 1999), adquiere un comportamiento nocturno y bentónico que conservará toda su vida (Rodríguez y Peleteiro, 2014).



Figura 4. A la izquierda, espécimen hembra de *Solea senegalensis* en el centro IFAPA de *El Toruño* en el que se aprecia el desarrollo gonadal **Fuente:** Mañanós, 2015. A la derecha, huevos maduros de *Solea senegalensis*.

Fuente: Anguis y Cañavate, 2005.

En el lenguado senegalés existe un proceso de cortejo relacionado con el desove en el cual se pueden diferenciar tres etapas (Figura 5); inicialmente el macho descansa sobre la hembra durante un determinado periodo de tiempo para a continuación, en una segunda fase, ambos nadar de forma acoplada hacia la superficie. En una tercera y última etapa, la pareja mantiene los conductos genitales estrechamente unidos y se liberan los gametos (Carazo y otros, 2016). La ausencia de un proceso de cortejo sexual podría ser la razón del fracaso del proceso reproductivo en cautividad (Duncan y otros, 2008) dada la importancia de la proximidad de los gonoporos entre ambos sexos a la hora de la liberación de los gametos para que se lleve a cabo la fertilización de los ovocitos (Baynes y otros, 1994; Carazo y otros, 2016). Superar el problema de la pérdida del instinto del cortejo por parte de los machos es uno de los grandes retos que se plantean en la actualidad y existe gran interés en saber a qué se debe esta dificultad así como a la búsqueda de alternativas como posibles protocolos de fecundación artificial (Rodríguez y Peleteiro, 2014). Es por esto que la reproducción del lenguado senegalés depende en gran medida de los stocks procedentes del medio natural, los cuales realizan la puesta en cautividad sin ninguna dificultad (Anguis y Cañavate, 2005; Rodríguez y Peleteiro, 2014).

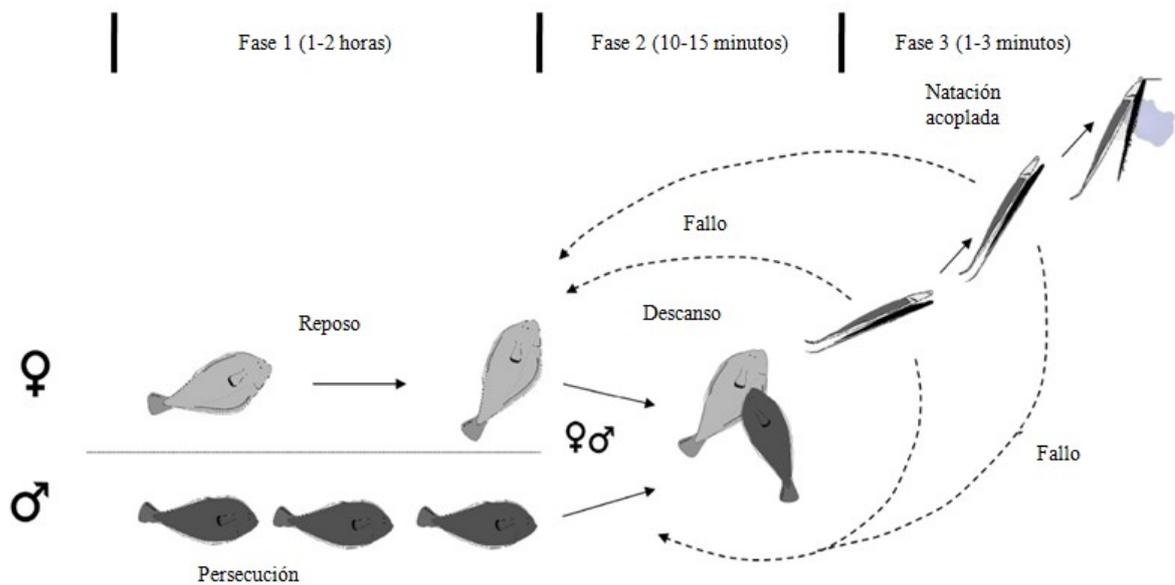


Figura 5. Esquema representativo del proceso de cortejo y desove en *Solea senegalensis*.

Fuente: Carazo y otros, 2016.

En el ámbito genético sobre *S. senegalensis* cabe señalar la realización por parte de diversos autores de estudios tanto a nivel genómico (García y otros, 2013; Benzekri y otros, 2014) como a nivel citogenético y de identificación de especies dentro del género *Solea* (Manchado y otros, 2006; Cross y otros, 2006). En cuanto al estado de las poblaciones de *S. senegalensis* en la naturaleza, cabe destacar que actualmente se carece de la información necesaria para la evaluación del estado de esta especie según los criterios de la Unión Internacional para la Conservación de la Naturaleza (UICN) y hoy en día no hay medidas de conservación. La información sobre el estado, la estructura genética y conectividad entre poblaciones naturales es esencial para comprender el potencial evolutivo, la salud y la sostenibilidad de la especie pero actualmente este conocimiento es reducido. Dentro del ámbito genético caben destacar los trabajos de evaluación y de análisis mediante el uso de marcadores genéticos realizados por Díaz-Ferguson y otros (2007, 2012) sobre poblaciones naturales en la costa sur atlántica española, entre Lisboa y el Golfo de Cádiz. Estos estudios determinaron diferencias entre las poblaciones de Lisboa con el resto de poblaciones (Díaz-Ferguson y otros, 2007), sugiriendo la consideración de tres grupos diferentes (Lisboa, Portimao y el resto de poblaciones). En el estudio realizado por estos autores en 2012, donde se analizaron dos poblaciones portuguesas y seis españolas localizadas en el golfo de Cádiz, se mostró una baja diferenciación genética global y ausencia de una estructura genética para el lenguado senegalés. Sería recomendable tener en cuenta estos aspectos a la hora de establecer planes conservación para la especie (Díaz-Ferguson y otros, 2007, 2012).

1.3. Interés socio-económico y producción.

En lo que respecta a la producción, actualmente existe un creciente interés en diversificar el cultivo de peces y su comercialización, por lo que en la región sur atlántica española, en la cual la especie más extendida y de mayor interés económico es la dorada (*Sparus aurata*), gran parte de los esfuerzos y proyectos realizados se enfocan en diferentes especies de peces comerciales como la hurta (*Pagrus auriga*), el pargo (*Pagrus pagrus*) o la seriola (*Seriola dumerili*) y entre las que se incluye el lenguado senegalés (Arellano y Sarasquete, 2006). Hoy en día se desarrollan numerosos proyectos de investigación con el objetivo de conseguir llevar a cabo la producción de *S. senegalensis*, la cual ha estado limitada por las incidencias con patologías oportunistas debidas al escaso conocimiento científico sobre la fisiología de esta especie, las bajas tasas de crecimiento por el desconocimiento en las

necesidades nutricionales y de acondicionamiento, ya resueltas con el uso de dietas adecuadas y correctas densidades de cultivo, (Cañavate, 2005) y principalmente por la dificultad para conseguir la reproducción en cautividad (Anguis y Cañavate, 2005).

Diversas especies de pleuronectiformes como el lenguado senegalés, el lenguado común (*Solea solea*), el rodaballo (*Psetta máxima*) o el fletan (*Hippoglossus hippoglossus*) son de gran interés para la acuicultura debido a sus características. Su carne es blanca y magra, con un sabor fino y sus espinas se retiran fácilmente, siendo un aporte importante de selenio, fósforo y yodo (Fundación Española de la Nutrición, 2013). Es por estos aspectos que son especies que presentan un alto precio en el mercado, de ahí el interés en su cultivo. La mayoría del lenguado senegalés cultivado es consumido actualmente en los países productores (Tabla 1) y dada la demanda, aún con el aumento de la producción, podrían mantenerse los precios en el mercado ya que se trata de una especie muy cotizada (Olmedo y otros, 2005).

Tabla 1. Producción acuícola del lenguado senegalés en Europa de 2003 a 2015.

Fuentes: Ministerio de Agricultura y Pesca, Alimentación y Medio Ambiente, 2015. FAO, 2015.

País	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
España	52	75	60	80	60	55	180	204	110	194	440	805	582
Francia	-	-	-	-	-	-	-	142	200	220	-	260	260
Portugal	4	4	11	9	8	13	14	14	50	100	154	-	-
Italia	-	-	-	-	-	19	14	14	10	-	-	-	-
Islandia	-	-	-	-	-	-	-	-	-	-	-	-	290
Total	56	79	71	89	68	87	208	374	370	514	594	1065	842

España fue el primer país productor de esta especie durante los primeros años y su producción, localizada principalmente en Galicia, Andalucía y Canarias, ha ido aumentando con el paso de los años, superando en 2013 las 400 toneladas de producción, llegando a un máximo de 805 y convirtiéndose en el mayor productor en la Unión Europea (FAO, 2013). Posteriormente el cultivo se extendió a Portugal, Italia y Francia, que actualmente es el segundo país productor tras España. Italia, sin embargo, tras un período de producción entre los años 2008 y 2011 abandonó su producción. En el futuro se espera que la producción siga aumentando con su desarrollo en países productores como Islandia y la preparación de

nuevas instalaciones previstas en España (Arellano y Sarasquete, 2006; Rodríguez y Peleteiro, 2014).

En lo que respecta a la pesquería de esta especie, según la FAO, su pesca se realiza dentro de las zonas 27 (Atlántico Noreste), 34 (Atlántico Centro-Este) y 47 (Atlántico Sureste), fundamentalmente con artes de trasmallo y arrastre de fondo (Figura 6).

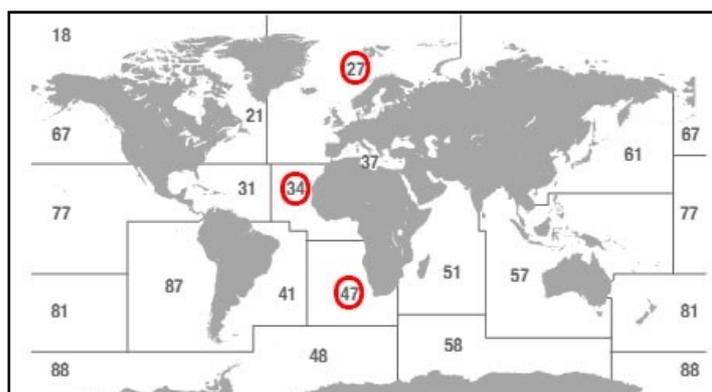


Figura 6. Zonas de pesca de *Solea senegalensis*. Fuente: FAO, 2017.

En cuanto a la pesca dentro de la región andaluza y según los datos de captura registrados (Tabla 2) se aprecia que, durante la última década, existe una tendencia creciente de capturas en peso así como del precio que esta especie alcanza en mercado. Destaca una caída de los precios en el año 2013 seguida de nuevo por un aumento, siendo su precio comparable en los últimos años con el de especies como la dorada o la lubina, los cuales se encuentran en torno a los 9 – 13 euros por kilo (IDAPES, 2015). En el último año se aprecia un considerable aumento tanto de las capturas en peso como del precio de venta alcanzado en lonja.

Tabla 2. Captura en peso del lenguado senegalés en Andalucía y precio medio de primera venta en lonja.

Fuente: IDAPES, 2015.

Año	Peso (kg)	Valor (euros)	Precio medio (euros/kg)
2010	14.741,40	177.125	12,02
2011	14.701,58	186.818,15	12,71
2012	15.968,10	208.704,05	13,07
2013	16.600,40	174.096,31	10,49
2014	12.250,65	151.266,34	12,35
2015	13.490,11	173.538,61	12,86
2016	18.503,36	261.328,68	14,12

1.4. Estudio de microRNAs

Los microRNAs o miRNAs son pequeñas moléculas de ARN no codificantes, de entre 20 y 25 nucleótidos de longitud que intervienen en la regulación de la expresión génica (Skok y otros, 2013). Inicialmente descubiertos por Howard Robert Horvitz (Ferguson y otros 1987, Chang y Mendell, 2007) y descritos por primera vez por Lee y otros autores (2013), estas moléculas presentan una secuencia complementaria a ARN mensajero que, en un proceso postranscripcional, hibridan con estos uniéndose a determinadas regiones diana, impidiendo su traducción a proteínas en un proceso que se conoce como silenciamiento génico y jugando un importante papel en la expresión génica (Skok y otros, 2013; Paczynska y otros, 2015). Los miRNAs se conservan en numerosas especies; se ha demostrado su presencia en multitud de organismos tanto vegetales (Park y otros, 2002) como animales (Campos y otros; 2014) poniéndose de manifiesto su conservación (Shi y otros, 2012) y su papel dentro de multitud de procesos fisiológicos incluidos la diferenciación de células madre y procesos de hematopoyesis así como de respuesta inmune. La mutación de proteínas involucradas en los procesos de biogénesis de miRNAs ha demostrado perjudicar al desarrollo animal (Campos y otros; 2014), un aspecto que destaca la importancia de estas moléculas.

Los miRNAs se transcriben a partir de genes del ADN contenido en el núcleo de las células, pero estos no son traducidos a proteínas. La ARN Polimerasa II origina un microRNA primario, también conocido como pri-miRNA (Figura 7), que consta de dos cadenas complementarias donde una de ellas incluye el microRNA (Paczynska y otros, 2015) y que hibridan entre si formando una estructura en forma de bucle también denominada en inglés como *hairpin* o *loop-structure*. Esta estructura de doble cadena es reconocida por la proteína *DGCR8* la cual, asociada con la enzima *DROSHA*, cortan el pri-miRNA para originar un microRNA precursor, el cual es transportado al citoplasma por la molécula *EXPORTIN 5* a través de los poros nucleares y donde se producirá el silenciado génico (Huang y otros, 2015). En el citoplasma, este precursor es reconocido por la enzima *DICER* la cual vuelve a cortar la molécula de microRNA. Acto seguido, el complejo *AGO2* interviene en la separación de las cadenas complementarias, las cuales son liberadas. Una vez que la cadena se encuentra libre constituye lo que se conoce como microRNA maduro que junto a *AGO2* y una serie de proteínas complementarias forman el complejo RISC (*RNA Induce Silencing*

Complex), que permitirá llevar a cabo el silenciado génico (Park y otros, 2002; Paczynska y otros, 2015).

El ARN mensajero de los genes diana sobre los que actúa el complejo RISC es complementario a la secuencia del microRNA, lo que permite la unión entre ambas. Tras la unión, existen dos formas en las que el complejo RISC puede inactivar la expresión del gen: Cortando el ARN mensajero, el cual será eliminado posteriormente por la célula, o simplemente inhibiendo la traducción evitando la unión de las subunidades del ribosoma (Huang y otros, 2015). En ambos casos se evitará que se produzca la traducción de proteínas.

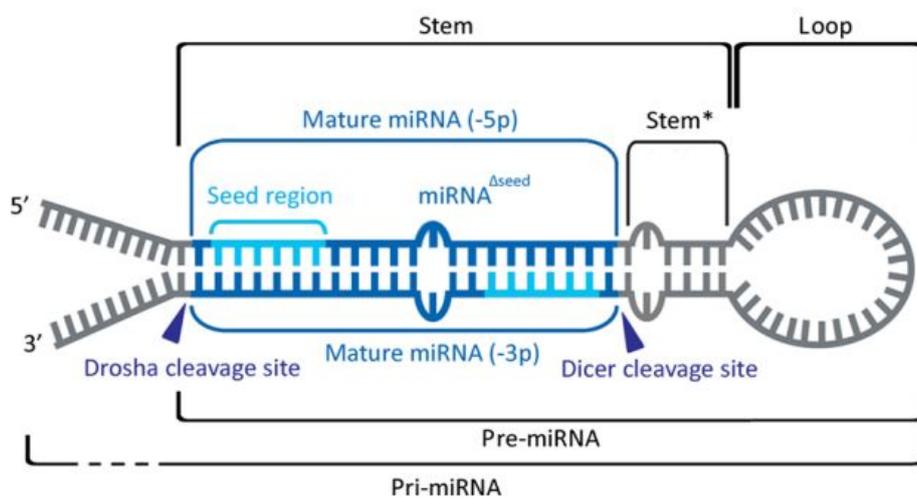


Figura 7. Representación de la estructura del microRNA. **Fuente:** Skok y otros, 2013

Teniendo en cuenta que un solo miRNA puede actuar sobre miles de ARN mensajeros controlando así múltiples genes, del mismo modo que un gen puede estar controlado por múltiples miRNAs, varios estudios presentan la relación de estas moléculas con la inhibición de la expresión génica e incluso se ha demostrado su intervención en procesos fisiológicos como el desarrollo muscular (Campos y otros; 2013 y 2014), la respuesta inmunológica (Shi y otros, 2012) e incluso el desarrollo de cánceres (Lu y otros; 2005).

En el caso del estudio genético de *S. senegalensis*, el trabajo realizado por Campos y otros (2014) relaciona los cambios en la tasa de crecimiento durante la ontogenia con la temperatura, siendo el crecimiento durante las fases embrionarias mayor a temperaturas de entre 18 – 21°C que a temperaturas menores (15°C). Esta diferencia viene acompañada de una diferenciación en la expresión de genes relacionados con el crecimiento muscular, la

cual parece estar regulada por miRNAs, aunque su papel en los procesos de desarrollo y crecimiento aún ha de ser investigado.

La relación de los miRNAs con aspectos como el crecimiento o en la respuesta inmunológica supone que estos adquieran un importante papel en cualquier estudio genético en acuicultura al tratarse de aspectos esenciales para la producción. El caso del lenguado senegalés, el estudio de estas moléculas podría dar respuesta a los problemas existentes que dificultan su cultivo.

1.5. Genoteca BAC de *Solea senegalensis*

Una genoteca consiste en una colección de fragmentos de ADN de longitud variable que, en conjunto, contienen toda la información del genoma de una especie en forma de fragmentos aleatorios. Existen diversos tipos dependiendo del material genético utilizado y suponen una herramienta imprescindible para cualquier estudio dentro del campo de la genética. Concretamente en este estudio se hizo uso de datos procedentes de una genoteca BAC de *S. senegalensis* elaborada a partir de larvas y proporcionados por el área de genética dentro del departamento de Biomedicina, Biotecnología y Salud Pública de la Universidad de Cádiz.

Las genotecas BAC (*Bacterial Artificial Chromosome*), descritas inicialmente por Shizuya y otros (1992), suponen un importante sistema de clonación bacteriana derivado de plásmidos de bacterias *Escherichia coli* para el estudio de genomas con capacidad de generar fragmentos de entre 100 y 300 kilobases con un alto grado de estabilidad (Shizuka y otros, 1992). Dentro de su estructura (Figura 8), es posible diferenciar tres regiones principales: Un origen de replicación, un gen de resistencia a antibiótico y una región donde se realiza la inserción del fragmento de ADN exógeno y que recibe el nombre de *polylinker*. Estos fragmentos de ADN se insertan en vectores sintéticos que se introducen en poblaciones bacterianas en las que las bacterias replican el fragmento de ADN permitiendo que este se pueda amplificar, aislar y manipular (Shizuya y otros, 1992). Este material genético generado puede ser almacenado facilitando la creación de genotecas BAC que presentan ventajas respecto a otras debido a la disminución del quimerismo y su facilidad a la hora de aislar clones (Shizuka y otros, 1992; Amemiya y otros, 1999).

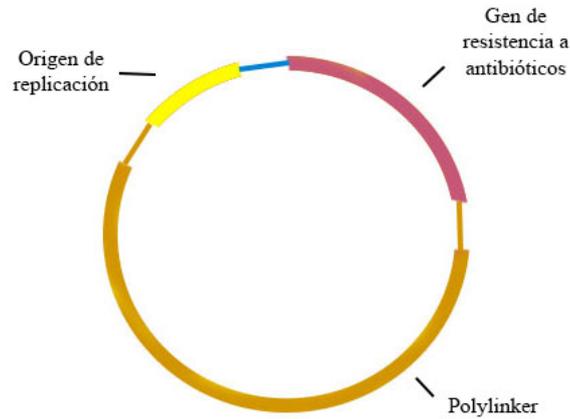


Figura 8. Esquema representativo de la estructura de un BAC

1.6. Bases de datos genómicas

Las bases de datos genómicas consisten en sistemas de almacenamiento que incluyen secuencias de nucleótidos o proteínas, pudiendo también incluir información de las mismas. Estas bases permiten acceder a secuencias de genes disponibles de cualquier especie de la cual se encuentre información dentro de la base de datos. Destacan principalmente la base estadounidense GenBank, la europea EMBL (European Molecular Biology Laboratory) y la japonesa DDBJ (DNA DataBank of Japan). Estas tres organizaciones intercambian datos entre sí y en su conjunto recogen los genomas de una gran cantidad de especies. La información es de libre acceso y los datos almacenados están disponibles para su descarga en formato *fasta*, un tipo de formato de texto comúnmente utilizado en informática para la representación de secuencias de nucleótidos haciendo uso de códigos de una única letra y permitiendo la adición de información adicional sobre las secuencias.

Dada la naturaleza del presente estudio es de especial interés la base de datos miRBase, una base de datos disponible en línea (www.mirbase.org) gestionada por el laboratorio Griffiths-Jones en la Facultad de Ciencias de la Naturaleza en la Universidad de Manchester con fondos del BBSRC (Biotechnology and Biological Sciences Research Council) que, del mismo modo que las bases de datos genómicas, recoge las secuencias de microRNA publicadas y proporciona información sobre la localización y la secuencia de las mismas.

2. Antecedentes y estado actual del tema

En la década de los 80, dada la abundancia de las capturas y a las ventajas de cultivo frente a otras especies, se inicia en España el cultivo del lenguado senegalés. Sin embargo, destacan desde las fases tempranas los problemas encontrados para conseguir la reproducción de esta especie en cautividad dada la existencia de persistentes patologías y a la incapacidad de los machos de la F1 para aparearse, que tiene como consecuencia la necesidad de un constante aporte de individuos salvajes para completar el ciclo y mantener la producción. Estos aspectos originan un creciente interés en la especie que ha llevado a su estudio desde diferentes puntos de vista, entre los que se incluye el estudio genómico que permita avanzar en el conocimiento del genoma de *S. senegalensis* y conocer los mecanismos que regulan la expresión genética en la especie. A día de hoy, junto a los problemas del cortejo, destacan los problemas de disfunción reproductiva asociados a la producción en cautividad como principal impedimento para la producción. En el caso de las hembras, el desove impredecible se acompaña de una baja calidad de los huevos; en los machos disminuye la producción de esperma así como su calidad (Carazo y otros, 2008 y 2009).

Aun siendo un tema de estudio reciente, los miRNAs adquieren un papel importante en los estudios genéticos dado a que por su alto grado de conservación es posible localizarlos en el genoma de casi cualquier especie viendo así su relación con genes de interés. Como se comenta anteriormente, los miRNAs son pequeñas secuencias de ARN que intervienen en la regulación de la expresión génica y cuyo estudio ayudaría a conocer mejor los mecanismos relacionados con aspectos de interés como el crecimiento, la reproducción o la respuesta inmunológica, procesos fundamentales en acuicultura. Según el registro disponible en la base de datos miRBase, hasta ahora únicamente se han identificado miRNAs en 9 especies de peces: *Cyprinus carpio*, *Danio rerio*, *Fugu rubripes*, *Hippoglossus hippoglossus*, *Ictalurus punctatus*, *Oryzias latipes*, *Paralichthys olivaceus*, *Salmo salar* y *Tetraodon nigroviridis* (Huang y otros, 2015). Estos estudios se complementan con los estudios realizados que han tratado los miRNAs y su posible relación con aspectos como el crecimiento, como es el caso de Campos y otros autores (2013 y 2014) comentado anteriormente, y que contribuyen a un mayor conocimiento del genoma de esta especie. Otros estudios también han tratado la relación de los miRNAs con la maduración sexual en otras especies dentro del orden de los

pleuronectiformes (Bizuayehu y otros, 2012) o incluso su posible relación con el cáncer y la generación de tumores (Fornari y otros, 2010).

3. Justificación y objetivos

Teniendo en cuenta que el cultivo del lenguado senegalés no se ha conseguido de forma integral y dado que el control de su reproducción no es tan conocido como el de otras especies de interés comercial cultivadas dentro del territorio español como *Sparus aurata* (Linnaeus, 1758), es de interés ampliar los conocimientos sobre la naturaleza del lenguado senegalés con el fin de aportar información que pueda contribuir al desarrollo de la reproducción y producción de esta especie con gran potencial económico.

El proyecto que aquí se presenta busca contribuir al estudio de esta especie de interés en acuicultura para alcanzar un mayor conocimiento de su genoma. En los últimos años se han estudiado múltiples aspectos de la biología de *S. senegalensis*, incluidos los genéticos. Como se ha comentado anteriormente, actualmente los miRNAs adquieren un papel importante dado a que por su alto grado de conservación es posible localizarlos en el genoma de casi cualquier especie viendo así su relación con genes de gran interés en dicho ámbito. Esto hace posible conocer mejor los mecanismos que regulan determinados aspectos como son el crecimiento, la reproducción o la respuesta inmunológica, contribuyendo a la producción de animales sanos de alta calidad y seguridad para el consumidor. El presente proyecto bioinformático pretende un estudio de *S. senegalensis* desde un punto de vista genético con el objetivo principal de profundizar en la caracterización del genoma del lenguado senegalés a través del estudio de datos obtenidos de genotecas BAC y de microRNAs. Para la consecución de este objetivo principal se plantean una serie objetivos parciales:

- 1. Crear una base de datos a partir de secuencias de BACs de *Solea senegalensis*.
- 2. Rastrear las secuencias de microRNAs extraídas de una base de datos genómica en secuencias de BACs de *S. senegalensis*.
- 3. Analizar las estructuras secundarias de los microRNAs en BACs y compararlas con las obtenidas en otros organismos ya descritos.
- 4. Localizar y describir microRNAs en los mismos BACs.

4. Material y métodos

4.1. Base de datos de *Solea senegalensis*

Para el estudio de las secuencias de ADN de *S. senegalensis*, obtenidas mediante secuenciación a partir de una genoteca BAC y con el objetivo de identificar secuencias de microRNAs, el tratamiento de estas secuencias se llevó a cabo con la herramienta de edición BioEdit Sequence Alignment Editor[®] versión 7.2.5, un editor de secuencias de nucleótidos (Hall, 1999) y BLAST[®] (Basic Local Alignment Search Tool) versión 2.6.0, una aplicación que permite la realización de comparaciones de secuencias de nucleótidos o proteínas con bases de datos genómicas, calculando la significación estadística de las coincidencias permitiendo inferir las relaciones funcionales y evolutivas entre las secuencias así como ayudar a identificar a los miembros de las familias de genes. Con BioEdit[®], se concatenaron secuencias de la genoteca BAC de *S. senegalensis* distribuidas en 8 lotes, dando lugar a un único archivo en formato *fasta*, del mismo modo que se encuentran los datos disponibles en las bases de datos genómicas.

El material con el que se trabajó y con el que se llevaron a cabo los análisis consistió en 2.960 contigs de 77 BACs de *S. senegalensis*. Los contigs presentaron una longitud media de 4.126 nucleótidos de longitud, con un valor de desviación típica de 7.426. En total, el número de nucleótidos que se manejó fue de 12.651.172, aunque teniendo en cuenta que algunos de los BACs fueron enviados a secuenciar más de una vez, omitiendo los BACs secuenciados duplicados, quedarían un total de 11.348.681 nucleótidos. Entre los BACs secuenciados más de una vez se encuentran 05K05 en los lotes 1 y 2, 10K23 en los lotes 3 y 6, 11O20 en los lotes 3, 4 y 6, 16E16 en los lotes 3 y 5, 19J21 en los lotes 3 y 6, 21O23 en los lotes 3 y 5, 28F09 en los lotes 4 y 5 y por último, 63A03 y 63A07 en los lotes 2 y 5. En el caso de estos BACs duplicados finalmente se utilizaron únicamente los últimos resultados, de forma que para el caso del BAC 05K05 presente en los lotes 1 y 2, por ejemplo, únicamente se utilizó en el lote 2.

A las secuencias se le otorgó una nomenclatura conveniente que permitiera su identificación, diferenciando dentro del mismo los siguientes campos, expuestos con el siguiente ejemplo:

01 - 56H24 - 07 - 014
1 2 3 4

1. Número del cromosoma; indicando "NA" en aquellos casos en los que se desconoce y "99" si hay hibridación con varios cromosomas.
2. Coordenada del BAC secuenciado en las genotecas disponibles en el laboratorio; indicando en primer lugar la placa, seguida de la coordenada dentro de esta placa. Cabe destacar que algunos BACs fueron enviados a secuenciar más de una vez.
3. Número del lote en que se envió a secuenciar el BAC.
4. Número de secuencia dentro de las obtenidas para el BAC en el lote determinado.

La creación de este único archivo *fastay* su integración con la herramienta *makeblastdb*, dentro de la aplicación BLAST[®], permitieron realizar las comparaciones de secuencias en busca de homologías.

4.2. Obtención de secuencias de microRNA

La base de datos miRBase recoge las secuencias de microRNA publicadas y proporciona información sobre la localización y la secuencia de las mismas. Cada entrada dentro de la base representa un fragmento de microRNA que origina la estructura *hairpin*; tanto las secuencias *hairpin* como las secuencias maduras, las cuales hacen referencia a las secuencias de microRNAs maduros, están disponibles para su descarga en formato *fasta*. De esta forma, es posible obtener datos de secuencias de miRNAs para la realización de análisis en busca de homologías existentes entre las secuencias de *S. senegalensis* con secuencias ya descritas de miRNAs disponibles en miRBase, donde se proporciona información de la especie o las especies donde han sido descubiertas y se hace referencia a estudios y autores.

En el presente proyecto, se hizo uso de las secuencias maduras disponibles en miRBase y descargadas en marzo de 2017. Existen un total de 35.829 secuencias maduras, de las cuales fueron utilizadas 23.572 después de un proceso de unificación de secuencias, con *FASTA*

Unique Sequences versión 1.0. Esta herramienta, disponible dentro de la web del Centro Nacional para la Información Biotecnológica (NCBI; www.ncbi.nlm.nih.gov), permite la unificación de secuencias repetidas o que presentan gran similitud entre ellas, con el objetivo de obtener secuencias únicas y evitar resultados duplicados. Junto a estas, se añadieron al análisis 48 secuencias maduras de miRNAs ya descritas previamente en *S. senegalensis* por Campos y otros (2014).

4.3. Alineamiento de secuencias e identificación de miRNAS

El programa BLAST[®] (Basic Local Alignment Search Tool) desarrollado por el National Center for Biotechnology Information permitió comparar secuencias en los BACs del lenguado senegalés, contenidas en la base de datos creada, con las secuencias de miRNAs maduros obtenidas en miRBase. Dentro de las opciones ofrecidas por la aplicación *blastn*, dentro del programa BLAST[®], se diferencian 3 tipos de análisis: *megablast*, para comparaciones entre secuencias muy similares en casos en los que existan, por ejemplo, errores de secuenciación, *blastn*, en aquellos casos de comparaciones entre especies, y *blastn-short*, optimizado para secuencias de menos de 30 nucleótidos. Dadas las características de los datos con los que se trabajó en el presente ensayo, y teniendo en cuenta que la longitud de las secuencias miRNAs maduras se encuentra en torno a 20 - 25 pares de bases, se trabajó con análisis tipo *blastn* y *blastn-short*.

Para el análisis fue necesario tener en cuenta una serie de parámetros dentro de los que proporciona BLAST[®], estableciendo unos valores predeterminados que permitieran garantizar la fiabilidad del análisis. Entre estos parámetros destacan:

- ***e-value***: Parámetro que describe el número de emparejamientos aleatorios esperados cuando se hace una búsqueda en una base de un tamaño particular (Blast.ncbi.nlm.nih.gov, 2017). A menor valor, mayor será la similitud entre las secuencias comparadas
- ***frames***: Orientación de la secuencia examinada y la secuencia de la base de datos.
- ***gapopen***: *Indels* (inserciones y/o deleciones).
- ***mismatch***: Emparejamientos incorrectos.
- ***pident***: Porcentaje mínimo de identidad entre miRNA y secuencia en el BAC.
- ***qcovhsp***: Cobertura de la secuencia examinada por segmento.
- ***qlen***: Longitud de la secuencia examinada.

La elección de intervalos de valores para algunos de estos parámetros permitió realizar un filtrado de secuencias, obteniendo aquellos casos que mostrasen mayor fiabilidad. Se tomaron como referencia los parámetros descritos por Huang y otros (2015) con pequeñas modificaciones. Los intervalos establecidos fueron, según el parámetro:

- ***e-value***: Menor a 10^{-6} .
- ***gapopen***: Igual a 0. No se permitió la presencia de *indels*.
- ***mismatch***: Menor o igual a 4. Se permitió un máximo de 4 emparejamientos incorrectos.
- ***pident***: Mayor o igual al 75%.
- ***qcovhsp***: Igual a 100. Seleccionando así alineamientos donde el miRNA ha alineado en toda su longitud.

Tras el análisis, BLAST® permite obtener la posición de las secuencias miRNA potenciales en el BAC donde se encuentren homologías, permitiendo así aislar el miRNA y llevar a cabo su extracción mediante la aplicación *blastdbcmd*, dentro de este mismo paquete informático. Para la utilización de esta aplicación fue necesario especificar el comienzo y el final de las secuencias seleccionadas tras el filtrado. En este caso, la extracción realizada fue mayor a la longitud de las secuencias seleccionadas, extrayéndose una región adicional entorno al sitio donde se encontró similitud incluyendo 100 pares de bases en sentido 5' y otras 100 en sentido 3'. Esto permitió no solo la obtención de las secuencias miRNA, sino también el estudio de las estructuras secundarias de las mismas al obtener una región mayor a la longitud de las secuencias maduras, un aspecto necesario para el estudio de las secuencias secundarias, posteriormente realizado.

4.4. Análisis de estructuras secundarias

Tras obtener secuencias con similitud a miRNAs conocidos, se estudió si formaban las estructuras secundarias típicas de estas moléculas, lo cual apoyaría que se trata de secuencias activas. Se hizo uso de la aplicación MIRENA versión 2.0, la cual consiste en una herramienta que permite la identificación de pre-miRNAs con alta sensibilidad y especificidad detectando nuevos miRNAs por homología de miRNAs conocidos y con la que fue posible obtener el valor de la energía libre de las secuencias. MIRENA es un algoritmo de

búsqueda diseñado para la predicción de miRNAs y que permite confirmar las secuencias pre-miRNA (Carbone y Mathelier, 2010). Esta aplicación calcula, dada una determinada secuencia pre-miRNA, la energía libre mínima de equilibrio de la estructura secundaria de dicha secuencia, también llamada *adjusted minimum folding energy* (AMFE), como el cociente (Ecuación 1) entre la denominada como *minimum folding energy* o “energía mínima de pliegue” y la longitud en nucleótidos (n) de la secuencia pre-miRNA, multiplicado por 100 (Carbone y Mathelier, 2010).

$$AMFE = \frac{MFE}{l(n)} * 100$$

Ecuación 1. Calculo de energía libre de equilibrio en MIRENA.

El índice de energía libre mínima o *minimun free energy index* se calcula como el cociente entre la energía libre mínima de equilibrio y la composición en porcentaje de guanina y uracilo en la secuencia pre-miRNA (Ecuación 2).

$$\frac{AMFE}{G - U\%}$$

Ecuación 2. Calculo del índice de energía libre en MIRENA.

La estructura secundaria adoptada por un pre-miRNA ha de satisfacer, según las reglas establecidas por MIRENA, 5 criterios comentados brevemente a continuación (Carbone y Mathelier, 2010):

1. Una secuencia madura no puede plegarse sobre si misma, es necesaria la existencia de una secuencia pre-miRNA.
2. El cociente entre la longitud de la secuencia madura del miRNA y la longitud de la secuencia pre-miRNA (en nucleótidos), ha de encontrarse entre los valores 0,83 y 1,17 (Ecuación 3).
3. El porcentaje de nucleótidos no coincidentes en la secuencia del pre-miRNA ha de ser menor o igual al 26% de la longitud de la secuencia.
4. El valor de la energía libre de equilibrio o *adjusted minimum folding energy* a de ser menor o igual a -32.

5. La secuencia pre-miRNA ha de tener un índice de energía libre mínima menor a -0,85, un criterio que permite diferenciar secuencias pre-miRNAs de secuencias de ARN con alta especificad.

Con estos criterios, MIRENA permitió la identificación de pre-miRNAs, obteniendo el valor de energía libre y el estudio de las estructuras secundarias.

Posteriormente se hizo uso de la aplicaicón RNAfold. Dentro del paquete The ViennaRNA Packageversión 2.3.5 (Lorenz y otros, 2011), RNAfold es una herramienta quepermitepredecir estructuras de secuencias de ARN o de ADN de cadena simple mediante la minimización de la energía libre de la estructura. El programa puede obtener la estructura secundaria que tendría la secuencia, eligiendo entre las estructuras que presentan una menor energía de liberación (y con ella mayor estabilidad) y permitiendo confirmar si se tratan de secuencias activas al presentar la estructura en forma de bucle, característica de los pre-miRNAs. Estos forman una estructura secundaria con un bucle al final en la cual el "tallo" contiene la secuencia miRNA madura (Figuras 7 y 9).

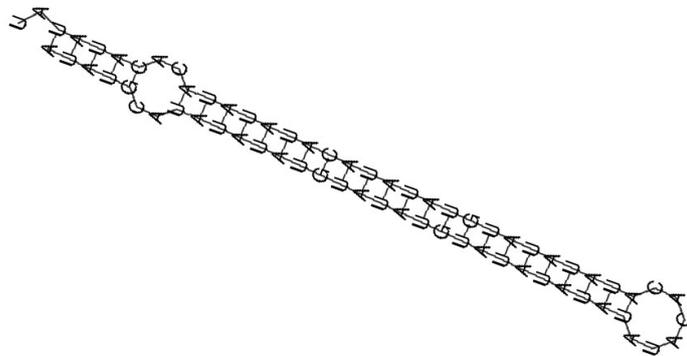


Figura 9. Estructura en forma de "horquilla" generada en RNAfold.

4.5. Estudio de la distribución de secuencias

Para conocer la distribución de las secuencias identificadas estas se clasificaron haciendo uso de la nomenclatura utilizada, agrupando las secuencias en cromosomas y BACs, permitiendo conocer el número de secuencias identificadas en cada categoría. Para estudiar si había diferencias significativas en cuanto al número de miRNAs entre cromosomas se utilizó una prueba X^2 de bondad de ajuste realizada con Excel. Se hizo uso de los datos

observados, y con el objetivo de obtener los valores esperados, se asumió igual frecuencia de miRNAs para todos los cromosomas, una vez ponderado por el tamaño de los mismos. Esta frecuencia esperada se obtuvo como la división del número total de miRNAs entre la suma del tamaño de las secuencias obtenidas para todos los cromosomas. Para obtener los valores esperados se multiplicó esa frecuencia por el tamaño total de secuencia obtenido de cada cromosoma. Dado que no todas las frecuencias esperadas eran mayores que 5, para evitar reducir la exactitud de los resultados se agruparon los cromosomas con menor frecuencia en la categoría “otros”.

5. Resultados y discusión

Se realizaron un total de 560.248 comparaciones, aunque tras la realización de una serie de filtrados adecuados estableciendo una serie de valores convenientes que garantizaran la fiabilidad del análisis y el estudio de las estructuras secundarias, se descartaron gran parte de las comparaciones seleccionándose como resultado final un total de 408 secuencias. Para una presentación más simplificada de los resultados, estas secuencias se agruparon según su similitud dando lugar como resultado definitivo 260 secuencias de *S. senegalensis* similares a miRNAs maduros previamente descritos de los cuales 12 fueron en peces. Estos resultados se presentan en los Anexos 1 y 2, detallados a continuación. Junto a las secuencias identificadas, en estos anexos se muestran los nombres otorgados a dichas secuencias, el miRNA con el que se presenta homología, el extremo de la cadena más próximo a la secuencia (3' o 5'), el número de desemparejamientos (*mismatch*), la cadena (Pos; positiva o Neg; negativa), la longitud en pares de bases del precursor, el porcentaje de la composición de adenina y uracilo con respecto al promotor de la secuencia madura así como la energía libre de la estructura secundaria. Se incluyen en la columna final miRNAs similares, mapeados aproximadamente en la misma posición y que podrían asociarse a las secuencias maduras identificadas en *S. senegalensis* dada su similitud y sus características. Estas secuencias presentan un e-valor bajo y un número reducido de desemparejamientos, sin embargo, con el objetivo de obtener una perspectiva simplificada de los resultados, se presentan asociados a la secuencia principal la cual presentaba los valores más bajos de estos parámetros.

El primer indicio de la multitud de miRNAs diferentes existentes comenzó con la realización de estudios de secuenciación en pequeños mamíferos y otro tipo de organismos como moscas o gusanos (Lagos-Quintana y otros, 2001). Actualmente se han identificado cientos de estas secuencias en diferentes organismos y la dificultad de su estudio se debe a que la expresión de algunos miRNAs puede limitarse no solo a un tipo celular sino a una etapa determinada del ciclo de vida del organismo estudiado, dificultando la detección de estos (Houbaviyy otros, 2003). A continuación, se trata la posible relación de las secuencias obtenidas como resultado de los análisis realizados en este ensayo con miRNAs ya descritos en estudios previos por parte de otros autores, realizando una recopilación bibliográfica de las posibles funciones de estas secuencias.

5.1. MiRNAs descritos en peces

Entre las secuencias resultantes, se presentan 12 casos con similitud a miRNAs ya descritos en peces (Anexo 1), donde los miRNAs hhi-miR-728 y ipu-miR-7562 se presentan dos veces, siendo por tanto 10 los miRNAs referenciados: ccr-miR-135c, ccr-miR-153b, ccr-miR-460-3p, ccr-miR-460-5p, dre-miR-153b-5, dre-miR-457b-3p, hhi-miR-430b, hhi-miR-728, ipu-miR-7562 y ssa-miR-430a-5p. Cada miRNA presenta en su nombre un prefijo que hace referencia a la especie en la que fue identificado. A continuación se detalla la información acerca de estas secuencias haciendo referencia a sus correspondientes estudios.

Los miRNAs ccr-miR-135c, ccr-miR-153b, ccr-miR-460-3p y ccr-miR-460-5p fueron identificados en la carpa común o europea (*Cyprinus carpio*) en un estudio de identificación de miRNAs y su relación con el desarrollo esquelético-muscular realizado por Yan y otros (2012). En el pez zebra (*Danio rerio*), los miRNAs dre-miR-153b-5p y dre-miR-457b-3p han sido descritos por Chen y otros (2005), los cuales estudian su relación con el desarrollo embrionario de esta especie, concretamente en el período de blástula. Los miRNAs hhi-miR-430b y hhi-miR-728 fueron identificados en el fletán (*Hippoglossus hippoglossus*) en el estudio realizado por Bizuayehu y otros (2012 y 2013) donde se mostró su expresión durante el desarrollo temprano de los individuos así como en el desarrollo sexual. El miRNA ipu-miR-7562 fue identificado en el pez gato americano (*Ictalurus punctatus*) junto a otros miRNAs (ipu-miR-129b, ipu-miR-7553,...) estudiados por Xu y otros (2013) los cuales presentaron expresión en muestras de tejido de cerebro, corazón, músculo e hígado. La expresión

múltiple en estos tejidos podría indicar que estos estén asociados con funciones fundamentales como el metabolismo (Xu y otros, 2013). Por último, en el salmón común o del Atlántico (*Salmo salar*) se identificó el miRNA ssa-miR-430a-5p, en el estudio realizado por Andreassen y otros (2012) expresado en muestras de tejido hepático, aunque sin identificar su función.

5.2. MiRNAs descritos en otras especies

De las secuencias identificadas, 248 no mostraron similitud con miRNA descritos en peces. Dado el elevado número de resultados obtenidos, se calculó la frecuencia de aparición de estas secuencias con respecto al número de secuencias totales, las cuales se presentan en la Tabla 3, con el objetivo de tratar en este ensayo aquellos miRNAs con los que de forma más frecuente se ha presentado similitud.

Tabla 3. Frecuencia de aparición de los miRNAs maduros (N=260). Secuencias con frecuencia igual a uno agrupados en la categoría otros.

miRNA homólogo	n	n/N
dme-miR-4979-3p	58	0,223
mmu-miR-466a-5p	20	0,077
mmu-miR-467g	13	0,05
mmu-miR-467e-3p	11	0,042
mmu-miR-466h-3p	8	0,031
mmu-miR-669n	8	0,031
rno-miR-466b-2-3p	7	0,027
efu-miR-8915	6	0,023
mmu-miR-466p-5p	6	0,023
bta-miR-669	5	0,019
gma-miR1533	5	0,019
rno-miR-466b-3p	5	0,019
dme-miR-9388-5p	4	0,015
hsa-miR-3960	4	0,015
miR-467g	4	0,015
mmu-miR-669f-5p	4	0,015
cin-miR-4194-3p	3	0,012
mmu-miR-3082-5p	3	0,012
hhi-miR-728	2	0,008
hsa-miR-620	2	0,008
ipu-miR-7562	2	0,008
miR-5112	2	0,008
mmu-miR-466f	2	0,008
mmu-miR-467c-3p	2	0,008
mmu-miR-669b-3p	2	0,008

mmu-miR-669d-5p	2	0,008
rno-miR-3584-5p	2	0,008
Otros*	68	0,272

*: aca-let-7e-3p, aca-miR-153-5p, ath-miR169a-3p, ath-miR414, ath-miR854a, bbe-miR-4857-3p, bdi-miR530a, bfl-miR-4875a-5p, bmo-miR-2807b, bta-miR-2293, bta-miR-2917, bta-miR-568, bta-miR-6522, ccr-miR-135c, ccr-miR-153b, ccr-miR-460-3p, ccr-miR-460-5p, cel-miR-253-5p, cfa-miR-1840, cin-miR-3598-3p, chi-miR-1814, dme-miR-4968-3p, dme-miR-9388-3p, dre-miR-153b-5p, dre-miR-457b-3p, dsi-miR-2580, eca-miR-8995, efu-let-7f, efu-miR-9318a, gga-let-7j-3p, gga-miR-1617, gga-miR-1648-5p, gga-miR-2127, gga-miR-7462-5p, gga-miR-7466-3p, hhi-miR-430b, hsa-miR-1322, hsa-miR-1976, hsa-miR-4270, hsa-miR-4293, hsa-miR-4303, hsa-miR-4507, hsa-miR-548v, hsa-miR-6089, hsa-miR-8079, lja-miR7535, mdo-miR-7399-5p, mdo-miR-885, mmu-miR-1895, mmu-miR-2137, mmu-miR-297a-5p, mmu-miR-3077-3p, mmu-miR-466b-5p, mmu-miR-466d-3p, mmu-miR-466d-5p, mmu-miR-466f-5p, mmu-miR-467h, mmu-miR-669l-3p, oan-miR-1346, oan-miR-218-3p, ppy-miR-1322, prd-let-7-5p, prd-miR-7966a-3p, rgl-miR7797b, sbi-miR5386, sly-miR5303, ssa-miR-430a-5p, stu-miR7997a.

El miRNA dme-miR-4979-3p, el cual presenta la mayor frecuencia de aparición, se identificó en el estudio realizado en la mosca de la fruta (*Drosophila melanogaster*) realizado por Berezikov y otros (2011) y cuya expresión se dio, aunque con poca frecuencia y junto a otros miRNAs, en las muestras analizadas en diferentes etapas desde estados larvarios a individuos adultos. El segundo miRNA más frecuente, dme-miR-9388-5p, se identificó en esta especie por Mohammed y otros (2013) aunque sin determinar su función.

Los miRNAs mmu-miR-466a-5p, mmu-miR-467g, mmu-miR-467e-3p, mmu-miR-466h-3p, mmu-miR-669n, mmu-miR-466p-5p, mmu-miR-669f-5p, mmu-miR-3082-5p, mmu-miR-466f, mmu-miR-467c-3p, mmu-miR-669b-3p y mmu-miR-669d-5p, presentados por frecuencia de aparición, fueron identificados en el ratón común (*Mus musculus*). El miRNA mmu-miR-466a-5p fue identificado en el estudio realizado por Zuoren y otros (2005) donde se analizaron muestras de tejido testicular tanto en individuos de entre 15 y 20 días como en adultos y donde tuvo lugar la expresión del mismo, sugiriendo su presencia tanto en células germinales como en células somáticas (Zouren y otros, 2005). Por otro lado, los miRNAs mmu-miR-466f y mmu-miR-467g fueron identificados por Calabrese y otros (2007) en muestras de tejido ovárico. Nuevamente en un estudio de la especie, realizado por Chiang y otro autores (2010) donde se trataron tejidos de cerebro, ovario y testículo, se identificó la expresión de los miRNAs mmu-miR-466p-5p, mmu-miR-467e-3p, mmu-miR-467c-3p, mmu-miR-669f-5p, mmu-miR-3082-5p, mmu-miR-669d-5p y mmu-miR-669b-3p, siendo este último identificado anteriormente por Mineno y otros (2006) en su estudio realizado en embriones de ratón. El miRNA mmu-miR-466h-3p fue también identificado en este caso en tejido cerebral en el estudio realizado por Landgraf y otros (2007). Kuchenbauer y otros

(2008) identificaron el miRNA mmu-miR-669n en su estudio donde se trató médula ósea de esta especie aunque no fue validado experimentalmente.

En la rata parda (*Rattus norvegicus*) fueron identificados los miRNAs rno-miR-466b-2-3p, rno-miR-466b-3p y rno-miR-3584-5p; en los dos primeros casos en el estudio realizado por Landgraf y otros (2007) estos miRNAs se identificaron, al igual que en el caso del ratón, en muestras tratadas de tejido cerebral, aunque no se relacionó su presencia con ninguna función. Por otro lado, en el caso de rno-miR-3584-5p, Linsen y otros (2010) identificaron la presencia de este miRNA nuevamente en tejido cerebral aunque tampoco se especificó su posible función. Posteriormente en el estudio realizado por Lu y otros (2015) se identificó y sugirió su posible relación con los mecanismos de acción de anestesia general en el sistema nervioso central en mamíferos (Luy otros, 2015), aunque aún no se conoce con certeza su función en el cerebro.

El miRNA efu-miR-8915 identificado en el murciélago moreno (*Eptesicus fuscus*) por Platt y otros (2014) en un estudio donde se trató el genoma de esta especie junto al del perro y el caballo, aunque no se determinó experimentalmente su posible función. Del mismo modo, en el caso de la vaca (*Bos taurus*), en un estudio del ovario bovino por parte de Hossain y otros (2009) el miRNA bta-miR-669 fue identificado, nuevamente sin determinar su posible función. Ocurre lo mismo en el caso de gma-miR1533, identificado por Subramanian y otros (2008) en su estudio realizado en brotes de soja (*Glycine max*).

Se presentan dos casos de miRNAs estudiados previamente en seres humanos, hsa-miR-3960 y hsa-miR-620, mostrando este último una menor frecuencia de aparición. Respecto al estudio de miRNAs en el ser humano, cabe destacar que existen disponibles en miRBase 2588 secuencias maduras, de las cuales únicamente 296 han sido verificadas experimentalmente. Teniendo en cuenta que en 2006 el número de secuencias maduras presentes era de 321, siendo 234 de ellas verificadas (Cummins y otros, 2006), se pone en manifiesto la dificultad de demostrar experimentalmente la función de estas secuencias dado el elevado número existente y a que su expresión puede estar limitada a un determinado tipo celular o periodo de vida. En el estudio realizado por Hu y otros (2011) se identificó el miRNA hsa-miR-3960, el cual desempeña un papel regulador en la diferenciación osteoblástica y mantiene relación con el miRNA miR-2861 (Hu y otros, 2011)

descrito anteriormente por Li y otros (2009). Estos autores sugieren la intervención de este miRNA en los procesos de diferenciación osteoblástica y formación ósea (Li y otros, 2009; Hu y otros, 2011). En este estudio realizado en osteoblastos en *M. musculus* por Li y otros (2009) se identificó que el miRNA miR-2861 juega un papel regulador en la diferenciación osteoblástica reprimiendo la expresión de la histona deacetilasa 5 (HDAC5), implicada en la regulación de la transcripción (Li y otros, 2009). Es importante destacar que se encontró que miR-2861 se conserva en seres humanos, y una mutación homocigótica en el premiRNA pre-miR-2861 que bloquea la expresión de miR-2861 es causante de osteoporosis en adolescentes. Del mismo modo, la no expresión de este miRNA en el caso de ratón tiene como consecuencia la inhibición de la formación ósea así como la disminución de la masa ósea (Li y otros, 2009). Agrupado en el mismo loci se encuentra hsa-miR-3960, identificado en el ser humano y en el ratón (Hu y otros, 2011), el cual también interviene en la diferenciación osteoblástica manteniendo una relación de retroalimentación con miR-2861 (Hu y otros, 2011). Por último, en un estudio realizado sobre el cáncer colorrectal por Cummins y otros (2006), donde se trataron muestras de mucosa colónica, se identificó hsa-miR-620 como un posible nuevo miRNA, aunque no se especificó su posible función (Cummins y otros, 2006).

El miRNA miR-467g, comentado anterior mente (Calabrese y otros, 2007), junto a miR-5112, descrito como un posible nuevo miRNA por Spierings y otros (2011), fueron ambos identificados en el ratón aunque se presentan por separado estos dos casos dado a que se tratan de secuencias identificadas en el estudio realizado por Campos y otros (2014) añadidas anteriormente a las secuencias maduras analizadas, identificadas en dicho ensayo en el lenguado senegalés junto a otros 48 secuencias de miRNAs maduros que nunca antes habían sido descritos en peces. En este estudio se identificó miR-467g, con un bajo número de lecturas durante el desarrollo mientras que miR-5112 presentó variaciones de expresión durante el mismo. Teniendo en cuenta los estudios previos realizados en ratón donde se sugiere la relación de estos miRNAs con el crecimiento muscular y óseo (Calabrese y otros, 2007), es posible que jueguen un papel en el desarrollo de *S. senegalensis* (Campos y otros, 2014).

Por último, entre los miRNAs que presentaron menor frecuencia se incluyeron miR-4194-3p y ipu-miR-7562, este último ya descrito en el apartado anterior en el pez gato mientras

que cin-miR-4194-3p se identificó en el urocordado *Ciona intestinalis*. En el estudio realizado por Hendrix y otros (2010) se llevó a cabo una identificación de secuencias miRNAs secuenciando el genoma de esta especie, tratando huevos no fertilizados así como individuos adultos, considerando así diferentes etapas de desarrollo y donde se sugirió cin-miR-4194-3p como posible miRNA aunque no fue verificado experimentalmente.

En la categoría “Otros” se recogen aquellas secuencias con similitud a miRNAs que presentaron una baja frecuencia de aparición. Estos miRNAs han sido identificados tanto en especies de mamíferos y peces, incluyendo algunas de las descritas previamente, como en aves (gga-miR-1648-5p, *Gallus gallus*), reptiles (aca-let-7e-3p, *Anolis carolinensis*), insectos (bmo-miR-2807b, *Bombyx mori*) y plantas (ath-miR169a-3p, *Arabidopsis thaliana*).

5.3. Estructuras secundarias

Las estructuras secundarias de las secuencias identificadas se generaron, como ya se ha comentado anteriormente, con RNAfold. Para confirmar si se tratan de secuencias activas se comprobó manualmente que las estructuras resultantes presentaban la ya mencionada estructura característica de los pre-miRNAs, diferenciando en ellas un bucle final así como identificando la secuencia madura en el “tallo” de la estructura (Figura 9). Dada la cantidad de secuencias obtenidas, los resultados presentados se limitan a las secuencias anteriormente presentadas en el Anexo 1, es decir, aquellas con homología de miRNAs descritos previamente en peces, las cuales se presentan en el Anexo 3.

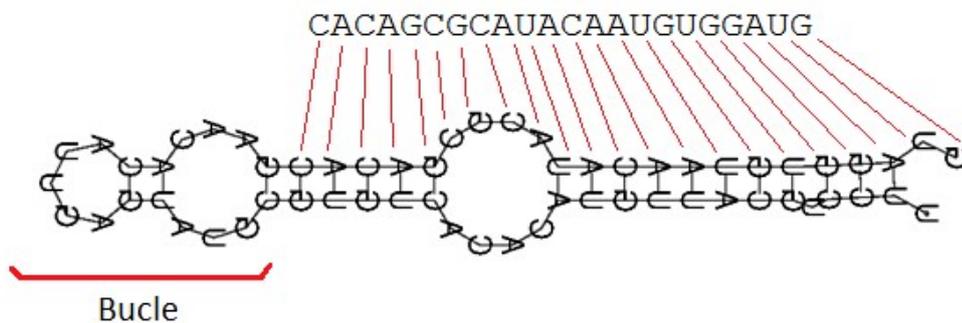


Figura 9. Estructura generada en RNAfold en la que se identificó la secuencia 01-48P07-07-005 con similitud al miRNAccr-miR-460-3p y el bucle característico de un pre-miRNA.

5.4. Distribución de las secuencias

A continuación se presenta la distribución resultante de las secuencias identificadas con similitud a miRNAs según la distribución en cromosomas (Figura 10) y BACs (Tabla 4). Cabe señalar que los cromosomas 5 y 14 presentan un valor de 0 dado que, de los BACs estudiados hasta ahora, ninguno corresponde a dichos cromosomas. Debido al elevado número de categorías en el caso de los BACs esta información se presenta en forma de tabla ordenados según su frecuencia. Adicionalmente se incluye en el Anexo 4 la longitud en nucleótidos por cromosoma y BAC.

Figura 10. Distribución del número de secuencias identificadas con similitud a miRNAs en cromosomas. Se representa la frecuencia absoluta por cromosoma, denominando como “NA” los casos correspondientes a BAC cuyo cromosoma es desconocido y “99” para BAC que hibridan con varios cromosomas.

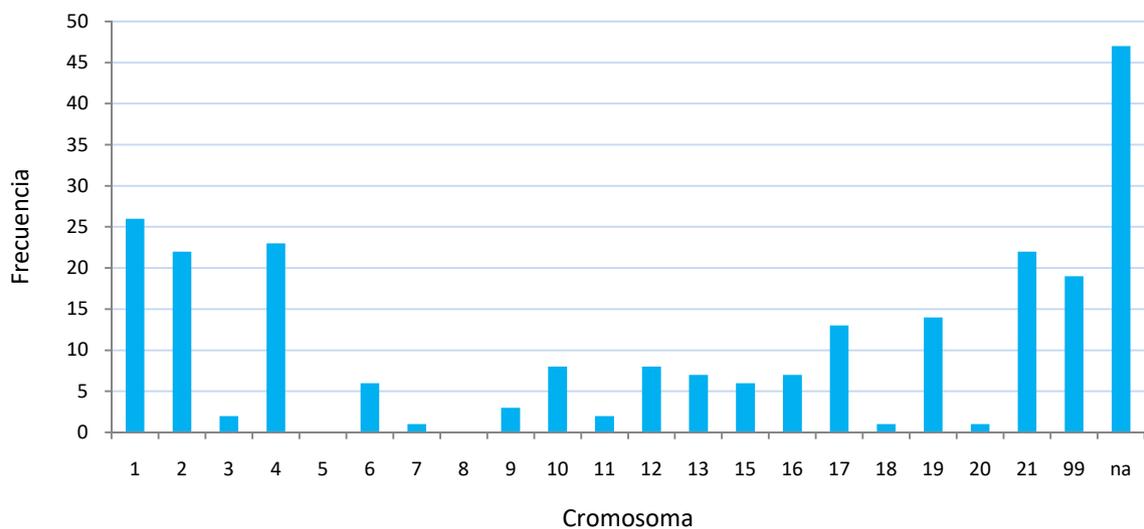


Tabla 4. Distribución del número de secuencias identificadas con similitud a miRNAs en cromosomas. Se representa la frecuencia absoluta por BAC.

BAC	N	BAC	N	BAC	N	BAC	N
30H22	14	36E03	6	01C02	3	36J02	2
0000	11	08O07	5	04D15	3	38N10	2
46C05	9	12K16	5	06P22	3	48K07	2
05K05	8	30J04	5	13F02	3	56H24	2
63A03	8	35D17	5	19J21	3	16E16	1
15I19	8	50K03	5	20D18	3	03C15	1
31C01	8	08A23	4	21O23	3	04M14	1
63A07	7	10L10	4	32B08	3	19H19	1
29D04	7	12D22	4	46P22	3	31F01	1

46B02	7	12K06	4	11O20	2	36H02	1
52G10	7	31O01	4	09B02	2	36M02	1
71N11	7	53D20	4	09J04	2	4E10	1
09N08	6	48P07	4	12N15	2	52C17	1
12D24	6	48P07	4	13F04	2	52E 07	1
31N01	6	10K23	3	31B01	2		

En el caso de los cromosomas, han sido 1, 2, 4 y 21 los que han presentado mayor número de secuencias con similitud a miRNAs, presentando valores próximos a la categoría “99” (Hibridación entre cromosomas) y por debajo de de “NA” (Cromosoma desconocido). En cuanto a los BACs, han sido 30H22, 0000 y 46C05 los que muestran el mayor número de secuencias, siendo 0000 el resultado de secuenciar un conjunto de BACs. Los BACs 05K05, 63A03, 15I19 y 31C01 son los siguientes en encabezar la lista. De cara a los resultados, se observa que las secuencias identificadas en los cromosomas 1, 2 y 4 presentan similitud a miRNAs cuya función, sugerida en la bibliografía comentada anteriormente, está relacionada con el desarrollo y el crecimiento como es el caso de los miRNAs descritos en la carpa europea y su relación con el desarrollo esquelético-muscular (Yan y otros, 2012) o en el caso del ratón nuevamente con relación en el crecimiento muscular y óseo (Calabrese y otros, 2007). Cabe destacar que una gran parte de las secuencias identificadas en el cromosoma 1 se encuentran dentro del BAC 05K05, uno de los BACs que se encuentra entre los que encabezan la lista.

En el BAC 30H22, donde se encuentra el mayor número de secuencias identificadas, se han encontrado homologías con varios de los miRNAs mencionados previamente y que presentaron las mayores frecuencias de aparición (Tabla 3), como dme-miR-4979-3p, mmu-miR-669n, mmu-miR-466h-3p, rno-miR-466b-3p o rno-miR-466b-3p entre ellos. La mayor parte de las secuencias con similitud a estos miRNAs, identificados en muestras de tejido cerebral (Landgraf y otros, 2007) y estudios de desarrollo larvario (Berezikov y otros, 2011), se encuentran localizan en el cromosoma 21 aunque de momento se desconoce la función de estos.

Por otro lado, en el cromosoma 6, presentando una frecuencia menor, se han identificado secuencias con similitud a miRNAs identificados en etapas tempranas de los individuos y aspectos sexuales, como es el caso del estudio por Mineno y otros (2006) realizado en

embriones de ratón o el fletán en el estudio realizado por (Bizuayehu y otros, 2012 y 2013) donde se mostró su expresión durante el desarrollo temprano de los individuos así como en el desarrollo sexual. Teniendo en cuenta estos aspectos, sería de interés la realización de estudios de expresión en individuos jóvenes de *S. senegalensis* que permitieran identificar la expresión de estas secuencias en muestras de tejido muscular, óseo o gonadal.

Tras el cálculo del estadístico χ^2 (Tabla 5) el análisis de la frecuencia de miRNAs entre cromosomas mostró diferencias significativas ($\chi^2 = 39,95$, $gl = 12$, $p\text{-valor} = 7,3 \cdot 10^{-5}$), lo que indicaría que la distribución de los miRNAs no es aleatoria. En futuros estudios sería de interés estudiar si los miRNAs identificados se ubican cerca de su gen diana. De forma adicional, se presenta en el anexo 5 la relación de BACs y el gen de interés a buscar.

Tabla 5. Frecuencias observadas y esperadas por cromosoma.

Cromosoma	Frecuencia observada	Frecuencia esperada
1	26	37.8
2	22	20.9
4	23	24.0
6	6	8.5
9	3	5.3
10	8	5.5
13	7	5.4
16	7	8.6
17	13	6.9
19	14	7.4
20	1	7.5
21	22	9.9
Otros*	20	24.5

* Cromosomas 3, 5, 7, 8, 11, 12, 14, 15 y 18.

6. Conclusiones

El objetivo del presente estudio ha sido identificar secuencias de microRNA potenciales, analizando una base de datos creada para *S. senegalensis*, buscando en la misma secuencias con similitud a miRNAs maduros identificadas en otras especies, estudiando posteriormente su estructura secundaria para confirmar su posible funcionalidad.

Dados los resultados, se puede concluir:

- Se analizaron 2.960 contigs de 71 BACs, repartidos en 8 lotes. Un total de 12.651.172 nucleótidos.
- Se hizo uso de las secuencias maduras disponibles en miRBase, manejando un total de 23.572 secuencias.
- Se realizó un total de 560.248 comparaciones, obteniendo como resultado final 260 secuencias con similitud a miRNAs maduros, 12 de los cuales han sido previamente identificados en peces.
- Las secuencias identificadas presentan la estructura secundaria típica de las secuencias pre-miRNA, un aspecto que apoya su posible funcionalidad.
- En un futuro sería de interés la realización de estudios de expresión, con el objetivo de identificar los miRNAs aquí presentados.

Bibliografía

- Amemiya, C.T., Zhong, T.P., Silverman, G.A., Fishman, M.C., y Zon, L.I. (1999). Zebrafish YAC, BAC, and PAC genomic libraries. *Methods Cell Biol.*, 60. 235 – 258.
- Andreassen, R., Worren, M. M., & Høyheim, B. (2013). Discovery and characterization of miRNA genes in atlantic salmon (*Salmo salar*) by use of a deep sequencing approach. *BMC Genomics*, 14, 482.
- Anguis, V. y Cañavate, J.P. (2005). Spawing of captive Senegal sole (*Solea senegalensis*) under a naturally fluctuating temperature regime. *Aquaculture*, 243. Centro de Investigación y Formación Acuícola y Pesquera “El Toruño”, IFAPA, Junta de Andalucía. 133 – 145.
- Arellano, J.M., y Sarasquete, C. (2006). *Biología de Solea senegalensis*. Atlas histológico del lenguado senegalés. Consejo superior de investigaciones científicas (CSIC). Madrid. 21–25.
- Arjona, F. J., Vargas, L., Martín, M. P., Flik, G., Mancera, J. M. y Klaren, P. H. M. (2008). The involvement of thyroid hormones and cortisol in the osmotic acclimation of *Solea senegalensis*. *General and comparative endocrinology* 155(3). 796-803.
- Baynes S.M., Beard, T.B., Hallam, J.D. y Howell, B.R.(1994). A description of spawning behaviour of captive Dover sole, *Solea solea* (L). *Neth. J. SeaRes.*, 32. 271 – 275.
- Benzekri, H., Armesto, P., Cousin, X., Rovira, M., Crespo, D., Merlo, M.A., Mazurais, D., Bautista, R., Guerrero-Fernández, D., Fernández-Pozo, N., Ponce, M., Infante, C., Zambonino, J.L., Nidelet, S., Gut, M., Rebordinos, L., Planas, J.V., Bégout, ML., Claros, M.G. y Manchado, M. (2014). De novo assembly, characterization and functional annotation of Senegalese sole (*Solea senegalensis*) and Common sole (*Solea solea*) transcriptomes: Integration in a database and design of a microarray. *BMC Genomics*, 952.
- Berezikov, E., Robine, N., Samsonova, A., Westholm, J. O., Naqvi, A., Hung, J.H., Lai, E. C. (2011). Deep annotation of *Drosophila melanogaster* microRNAs yields insights into their processing, modification, and emergence. *Genome Research*, 21(2), 203–215.
- Bizuayehu, T.T., Lanes, C.F., Furmanek, T., Karlsten, B.O., Fernandes, J.M., Johansen, S.D. y Babiak, I. (2012). Differential expression patterns of conserved miRNAs and isomiRs during Atlantic halibut development. *BMC Genomics*. 13:11.
- Bizuayehu, T. T., Fernandes, J. M. O., Johansen, S. D., & Babiak, I. (2013). Characterization of Novel Precursor miRNAs Using Next Generation Sequencing and Prediction of miRNA Targets in Atlantic Halibut. *PLoS ONE*, 8(4).
- Blast.ncbi.nlm.nih.gov. (2017). *BLAST Frequently Asked questions*. [En línea] Disponible en:https://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE_TYPE=BlastDocs&DOC_TYPE=FAQ#expect (Citado el 29 de agosto de 2017).
- Calabrese, J. M., Seila, A. C., Yeo, G. W., & Sharp, P. A. (2007). RNA sequence analysis defines Dicer’s role in mouse embryonic stem cells. *Proceedings of the National Academy of Sciences of the United States of America*, 104(46), 18097–18102.
- Campos, C., Valente, L.M., Conceicao, L.E., Engrola, S., Sousa, V., Rocha, E. y Fernandes, J.M. (2013). Incubation temperature induces changes in muscle cellularity and gene expression in Senegalese sole (*Solea senegalensis*). *Gene* 2013, 516(2). 209–217.

- Campos, C., Sundaram, A., Valente, L., Conceicao, L., Engrola, S. y Fernández, J. (2014). Thermal plasticity of the miRNA transcriptome during *Senegalese sole* development. *BMC Genomics*.15:525. 4 – 13.
- Cañavate, J.P. (2005). Opciones del lenguado senegalés *Solea senegalensis* Kaup, 1858 para diversificar la acuicultura marina. Instituto Español Oceanográfico. 147 – 154.
- Carazo, I., Mañanos, E., Norambuena, F. y Duncan, N. (2008). Preliminary observations of behaviour of Senegal sole broodstock (*Solea senegalensis*) induced to ovulate and release eggs using different male and female hormone (GnRH α and hCG) treatments. A: Sixth International Symposium on Fish Endocrinology.
- Carazo, I., Chereguini, O., Huntingford, F., Martín, I., Norambuena, F. y Duncan, N. (2009). Observaciones del cortejo de lenguado (*Solea senegalensis*, Kaup 1858) salvaje mantenido en cautividad. A: XII Congreso Nacional de Acuicultura.
- Carazo, I., Chereguini, O., Duncan, N., Huntingford, F. y Martín, I. (2016). Reproductive ethogram and mate selection in captive wild *Senegalese sole* (*Solea senegalensis*). *Spanish Journal of Agricultural Research*, Volume 14, 4, 6 – 13.
- Carbone, A., Mathelier, A. (2010). MIRENA: finding microRNAs with high accuracy and no learning at genome scale and from deep sequencing data. *Bioinformatics*. 26 (18): 2226-2234.
- Chang, T.C. y Mendell, J.T. (2007). MicroRNAs in vertebrate physiology and human disease. *Annu Rev Genomics Hum Genet*, 8. 215 – 239.
- Chen, P.Y., Manninga, H., Slanchev, K., Chien, M., Russo, J.J., Ju, J., Sheridan, R., John, B., Marks, D.S., Gaidatzis, D., Sander, C., Zavolan, M. y Tuschl, T. (2005). The developmental miRNA profiles of zebrafish as determined by small RNA cloning. *Genes Dev* 19. 1288–1293.
- Chiang, H. R., Schoenfeld, L. W., Ruby, J. G., Auyeung, V. C., Spies, N., Baek, D., Johnston, W.K, Russ, C., Shujun, L., Babiarz, J.E., Blelloch, R., Schroth, G.P., Nusbaum, C., Bartel, D.P. (2010). Mammalian microRNAs: experimental evaluation of novel and previously annotated genes. *Genes & Development*, 24(10), 992–1009.
- Cross, I., Merlo, M.A., Manchado, M., Infante, C., Cañavate, J.P. y Rebordinos, L. (2006). Cytogenetic characterization of the sole *Solea senegalensis* (Teleostei: Pleuronectiformes: Soleidae): Ag-NOR, (GATA) n , (TTAGGG) n and ribosomal genes by one-color and two-color FISH. *Genetica*, 128. 1–3.
- Cummins, J. M., He, Y., Leary, R. J., Pagliarini, R., Diaz, L. A., Sjoblom, T., ... Velculescu, V. E. (2006). The colorectal microRNAome. *Proceedings of the National Academy of Sciences of the United States of America*, 103(10), 3687–3692.
- Desoutter, M. (1990). Soleidae. In J.C. Quero, J.C. Hureau, C. Karrer, A. Post y L. Saldanha (eds.) Check-list of the fishes of the eastern tropical Atlantic (CLOFETA). JNICT, Lisbon; SEI, Paris; and UNESCO, Paris. Vol. 2. 1037 – 1049.
- Diaz-Ferguson, E.; Cross, I.; Barrios, M.; Pino, A.; Castro, J.; Bouza, C.; Martínez, P. & Rebordinos, L. (2012). Genetic characterization, based on microsatellite loci, of *Solea senegalensis* (Soleidae, Pleuronectiformes) in Atlantic coast populations of the SW Iberian Peninsula *Ciencias Marinas*, 38, 129-142
- Dinis, M.T., Ribeiro, L., Soares, F. y Sarasquete, C. 1999. A review on the cultivation potential of *Solea senegalensis* in Spain and in Portugal. *Aquaculture* 176, 27–38.

- Dovc, P., Godnic, I., Horvat, S., Kovac, M., Kunej, T., Skok, J. y Zorc, M. (2013). Genome-wide in silico screening for microRNA genetic variability in livestock species. *Animal Genetics*. Stichting International Foundation for Animal Genetics, 44, 669 – 677.
- Duncan N., Carazo I., Norambuena F., Estevez A., Chereguini O. y Mañanos E. (2008). Reproductive behaviour of *Senegal sole (Solea senegalensis)*. *International Sole Workshop*, Faro, Portugal, 11-13.
- FAO (2017). Cultured Aquatic Species Information Programme. *Solea solea*. Text by Colen, R., Ramalho, A., Rocha, F., and Dinis, M.T. FAO; Fisheries and Aquaculture Department [En línea]. Actualizado 18 de Febrero de 2014. (Citado el 15 de enero 2017). Disponible en: http://www.fao.org/fishery/culturedspecies/Solea_spp/en
- Fao.org. (2017). Organización de las Naciones Unidas para la Alimentación y la Agricultura. [En línea] (Citado el 1 de mayo de 2017) Disponible en: <http://www.fao.org/home/es/>
- Fen.org. (2013). Fundación Española de la Nutrición. Orden pleuronectiformes *Solea solea* [En línea]. (Citado el 10 de Marzo de 2017). Disponible en: <http://www.fen.org.es/mercadofen/pdfs/lenguado.pdf>
- Ferguson, E.L., Sternberg, P.W., Horvitz, H.R. (1987). A genetic pathway for the specification of the vulval cell lineages of *Caenorhabditis elegans*. *Nature* 326:259–67
- Fernández, B. y Rodríguez, J.L. (2003). Guía de la Piscicultura Europea. Consellería de Pesca e Asunto Marítimos. Xunta de Galicia. 45 – 46.
- Fishbase.org. (2017). Order Summary for Pleuronectiformes. *Solea senegalensis*. [En línea] (Citado el 5 de marzo de 2017). Disponible en: <http://www.fishbase.org/summary/ordersSummary.php?order=Pleuronectiformes>
- Fornari, F., Milazzo, M., Chieco, P., Negrini, M., Calin, G.A., Grazi, G.L., Pollutri, D., Croce, C.M., Bolondi, L. y Gramantieri, L. (2010). MiR-199a-3p regulates mTOR and c-Met to influence the doxorubicin sensitivity of human hepatocarcinoma cells. *Cancer Res*, 70, 12.
- García, A., Merlo, M.A., Ponce, M., Portela, S., Cross, I., Manchado, M. y Rebordinos, L. (2013). A preliminary genetic map in solea senegalensis (pleuronectiformes, soleidae) using bac-fish and next-generation sequencing. *Cytogenet. Genome Res.*, 141. 2–3.
- Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41: 95-98.
- Hendrix, D., Levine, M., & Shi, W. (2010). miRTRAP, a computational method for the systematic identification of miRNAs from high throughput sequencing data. *Genome Biology*, 11(4), R39.
- Houbaviy, H.B., Murray, M.F., Sharp, P.A. (2003). Embryonic stem cell-specific microRNAs. *Dev Cell* 5: 351–358.
- Hossain, M. M., Ghanem, N., Hoelker, M., Rings, F., Phatsara, C., Tholen, E., Schellander, K., Tesfaye, D. (2009). Identification and characterization of miRNAs expressed in the bovine ovary. *BMC Genomics*, 10, 443.
- Hu, R., Liu, W., Li, H., Yang, L., Chen, C., Xia, Z.-Y., ... Luo, X.-H. (2011). A Runx2/miR-3960/miR-2861 Regulatory Feedback Loop during Mouse Osteoblast Differentiation. *The Journal of Biological Chemistry*, 286(14), 12328–12339.
- Huang, Y., Ren, H., Wang, Y., Yu, X., Zhan, J. y Zou, Q. (2015) Computational identification and characterization of miRNAs and their genes from five cyprinidae fishes. *Saudi Journal of Biological Sciences*. King Saun University. 2 – 10.

- Huang, Y.; Zou, Q.; Ren, H. T. & Sun, X. H. (2015). Prediction and characterization of microRNAs from eleven fish species by computational methods Saudi Journal of Biological Sciences, Elsevier BV, 22, 374-381.
- IDAPES. (2015). Sistema de información andaluz de comercialización y producción pesquera. Consejería de Agricultura, Pesca y Desarrollo Rural. Junta de Andalucía. [En línea] (Citado el 5 de marzo de 2017). Disponible en: www.juntadeandalucia.es/agriculturaypesca/idapes/
- Imsland, A. K., Foss, A., Conceicao, L. E. C., Dinis, M. T., Delbare, D., Schram, E., Kamstra, A., Rema, P. y White, P. (2003). A review of the culture potential of *Solea solea* and *Solea senegalensis*. Review of Fish Biology and Fisheries, 13. 379 – 409.
- IUCN. (2016). The IUCN Red List of Threatened Species. Version 2016-3. Text by Monroe, T., Adeofe, TA, Camara, K., Camara, YH, Cissoko, K., de Moraes, L., Djiman, R., Mbye, E., Sagna, A., Sylla, M. y Tous, P. 2015. *Solea senegalensis*. [En Línea] (Citado el 5 de marzo de 2017). Disponible en: <http://www.iucnredlist.org/details/15622678/0>
- Kuchenbauer, F., Morin, R. D., Argiropoulos, B., Petriv, O. I., Griffith, M., Heuser, M., ... Humphries, R. K. (2008). In-depth characterization of the microRNA transcriptome in a leukemia progression model. Genome Research, 18(11), 1787–1797.
- Lagos-Quintana, M., Rauhut, R., Lendeckel, W., Tuschl, T. (2001). Identification of novel genes coding for small expressed RNAs. Science 294: 853–858.
- Lagos-Quintana, M., Rauhut, R., Yalcin, A., Meyer, J., Lendeckel, W., Tuschl, T. (2002). Identification of tissue-specific microRNAs from mouse. Curr Biol 12: 735–739.
- Landgraf, P., Rusu, M., Sheridan, R., Sewer, A., Iovino, N., Aravin, A., ... Tuschl, T. (2007). A Mammalian microRNA Expression Atlas Based on Small RNA Library Sequencing. Cell, 129(7), 1401–1414.
- Lee, R.C., Feinbaum, R.L., Ambros, V. (1993). The *Caenorhabditis elegans* heterochronic gene encodes small RNAs with antisense complementary to *lin-14*. Cell, 75, 843 – 854.
- Li, H., Xie, H., Liu, W., Hu, R., Huang, B., Tan, Y.F., Liao, E.Y., Xu, K., Sheng, Z.F., Zou, H.D, Xian-Ping, W., Luo, X.H. (2009). A novel microRNA targeting HDAC5 regulates osteoblast differentiation in mice and contributes to primary osteoporosis in humans. The Journal of Clinical Investigation, 119(12), 3666–3677.
- Linsen, S. E., de Wit, E., de Bruijn, E., & Cuppen, E. (2010). Small RNA expression and strain specificity in the rat. BMC Genomics, 11, 249.
- Lorenz, Ronny and Bernhart, Stephan H. and Höner zu Siederdisen, Christian and Tafer, Hakim and Flamm, Christoph and Stadler, Peter F. and Hofacker, Ivo L. ViennaRNA Package 2.0 Algorithms for Molecular Biology, 6:1 26, 2011, doi:10.1186/1748-7188-6-26
- Lu, J., Getz, G., Miska, E.A., Alvarez-Saavedra, E., Lamb, J., Peck, D., Sweet-Cordero, A., Ebert, B.L., Mak, R.H., Ferrando, A.A., Downing, J.R., Jacks, T., Horvitz, H.R., Golub, T.R. (2005). MicroRNA expression profiles classify human cancers. Nature, 435, 7043.
- Lu, Y., Jian, M.-Y., Ouyang, Y.-B., & Han, R.-Q. (2015). Changes in Rat Brain MicroRNA Expression Profiles Following Sevoflurane and Propofol Anesthesia. Chinese Medical Journal, 128(11), 1510–1515.
- Manchado, M., Rebordinos, L., Infante, C. (2006). U1 and U2 small nuclear RNA genetic linkage: A novel molecular tool for identification of six sole species (Soleidae, Pleuronectiformes). J. Agric. Food Chem., 54, 11.

- Manchado, M., Zuasti, E., Cross, I., Merlo, M.A., Infante, C. y Rebordinos, L. (2006). Molecular characterization and chromosomal mapping of the 5S rRNA gene in 39 *Solea senegalensis*: A new linkage to the U1, U2, and U5 small nuclear RNA genes. *Genome*, 49, 1.
- Mapama.gob.es. (2017). Ministerio de Agricultura y Pesca, Alimentación y Medio Ambiente. El mercado del lenguado en España. [En línea] (Citado el 1 de mayo de 2017). Disponible en: <http://www.mapama.gob.es>
- Mineno, J., Okamoto, S., Ando, T., Sato, M., Chono, H., Izu, H., ... Kato, I. (2006). The expression profile of microRNAs in mouse embryos. *Nucleic Acids Research*, 34(6), 1765–1771.
- Mirbase.org. (2017). *miRBase*. [En línea] Disponible en: <http://www.mirbase.org/index.shtml>
- Mohammed, J., Flynt, A. S., Siepel, A., & Lai, E. C. (2013). The impact of age, biogenesis, and genomic clustering on *Drosophila* microRNA evolution. *RNA*, 19(9), 1295–1308.
- Olmedo, M., J.B. Peleteiro, R. Cal y F. Linares. 2005. Crecimiento de juveniles de lenguado (*Solea senegalensis*, Kaup 1858) en Galicia. En: IX Congreso Nacional de Acuicultura (Cádiz, mayo 2003). La acuicultura como actividad económica en las zonas costeras: Libro de Actas (12-16 de mayo, 2003. Cádiz, España): 383-385. Consejería de Agricultura y Pesca, Junta de Andalucía. Sevilla, España.
- Paczynska, P., Grzenmski, A., Szydlowski, M. (2015). Distribution of miRNA genes in the pig genome. *BMC Genet.*, 16(1). 1 – 4.
- Park, W., Li, J., Song, R., Messing, J., Chen, X. (2002). Carpel Factory, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in *Arabidopsis thaliana*. *Curr Biol.*, 12 (17).
- Porta J, 2006. Desarrollo de microsatélites y su aplicación al cultivo del lenguado senegalés (*Solea senegalensis*). Tesis doctoral. Universidad de Málaga. 32.
- Rodríguez, R. B. (1984). Biología y cultivo de *Solea senegalensis* Kaup 1858 en el Golfo de Cádiz. Tesis doctoral. Departamento de Zoología. Universidad de Sevilla.
- Rodríguez, J.L. y Peleteiro, J. B. (2014). Cultivo del lenguado senegalés (*Solea senegalensis*). Cuadernos de acuicultura, 7. Fundación Observatorio Español de Acuicultura. Madrid. 12 – 36.
- Platt, R.N, Vandewege, M.W., Kern, C., Schmidt, C.J., Hoffman, F.G., Ray, D.A. (2014). Large Numbers of Novel miRNAs Originate from DNA Transposons and Are Coincident with a Large Species Radiation in Bats. *Mol Biol Evol* 2014; 31 (6): 1536-1545. doi: 10.1093/molbev/msu112
- Shi, B., Gao, W., Wang, J. (2012) Sequence Fingerprints of MicroRNA Conservation. *Plos One*. 7 (10). Thomas Preiss, The John Curtin School of Medical Research.
- Shizuya, H., Birren, B., Kim, U.J., Mancino, V., Slepak, T., Tachiiri, Y., Simon, M. (1992). Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc. Natl. Acad. Sci.U.S.A.*, 89, 8794–8797.
- Skok, J., Godic, I., Zorc, M., Horvat, S., Dovc, P., Kovac, M. y Kunej, T. (2013). *Animal Genetics*. Stichting International Foundation for Animal Genetics, 44. 669 – 677.
- Spierings, D. C., McGoldrick, D., Hamilton-Easton, A. M., Neale, G., Murchison, E. P., Hannon, G. J., Green, D.R., Withoff, S. (2011). Ordered progression of stage-specific miRNA profiles in the mouse B2 B-cell lineage. *Blood*, 117(20), 5340–5349.

- Subramanian, S., Fu, Y., Sunkar, R., Barbazuk, W. B., Zhu, J.-K., & Yu, O. (2008). Novel and nodulation-regulated microRNAs in soybean roots. *BMC Genomics*, *9*, 160.
- Vega, L., Díaz, E., Cross, I., Rebordinos, L. (2002). Cytogenetic and isozymatic characterization of the sole *Solea senegalensis* Kaup, 1858. *Boletín Instituto Español Oceanografía*, *18*(1–4). 245–250.
- Xu, Z., Chen, J., Li, X., Ge, J., Pan, J., y Xu, X. (2013). Identification and Characterization of MicroRNAs in Channel Catfish (*Ictalurus punctatus*) by Using Solexa Sequencing Technology. *PLoS ONE*, *8*(1).
- Yan, X., Ding, L., Li, Y., Zhang, X., Liang, Y., Sun, X., y Teng, C.-B. (2012). Identification and Profiling of MicroRNAs from Skeletal Muscle of the Common Carp. *PLoS ONE*, *7*(1).
- Zuoren, Yu., Raabe, T., Norman, B.H. (2005) MicroRNA Mirn122a Reduces Expression of the Posttranscriptionally Regulated Germ Cell Transition Protein 2 (Tnp2) Messenger RNA (mRNA) by mRNA Cleavage. *Biol Reprod*; *73* (3): 427-433.

Anexo 1. Secuencias de *Solea senegalensis* con similitud a miRNAs maduros descritos previamente en peces*.

Secuencia madura predicha	Nombre de la secuencia	miRNA homólogo	Extremo	Mismatch	Cadena	LP	A + U	Energía libre	miRNAs similares
CCUGCAUUGUACACACUGUGCG	01-48P07-07-005	ccr-miR-460-5p	5'	0	Pos.	60	55,0	-22,5	ola-miR-460-5p
CACAGCGCAUACAAUGUGGAUG	01-48P07-07-005	ccr-miR-460-3p	3'	0	Pos.	61	54,1	-22,5	aca-miR-460a-3p oan-miR-460-3p ola-miR-460-3p
CACACACACUCACACACACACA	02-04D15-06-001	ipu-miR-7562	3'	3	Pos.	131	45,8	-78,3	
UUGCAUAGUCACAAAAAUGAGC	04-12D24-04-001	ccr-miR-153b	3'	0	Neg.	91	58,2	-39,5	cfa-miR-153 efu-miR-153 oha-miR-153-3p
GUCAUUUUUGUGGUUUGCAGCU	04-12D24-04-001	dre-miR-153b-5p	5'	0	Neg.	91	58,2	-39,5	aca-miR-153-5p mdo-miR-153-1-5p oan-miR-153-1-5p spu-miR-153-5p
AUACUAAGUACACUACGUUUA	06-10K23-03-004	hhi-miR-728	3'	0	Neg.	80	56,3	-36,2	
AUACUAAGUACACUACGUUUA	06-10K23-06-002	hhi-miR-728	3'	0	Pos.	80	56,3	-36,2	
UAUGGCUUUUCUGCUCCGCGUGA	16-09N08-05-001	ccr-miR-135c	5'	4	Pos.	61	47,5	-28,4	cin-miR-135-5p dre-miR-135c
CACACACACUCUCACACACACA	19-31C01-06-008	ipu-miR-7562	3'	4	Pos.	134	48,5	-50,8	
UCCUGUAUUGCUGUCCGGUGU	na-00000-01-685	dre-miR-457b-3p	3'	3	Pos.	62	43,5	-26,8	

UACCCUAACUUUAGCAUUGACU	na-09B02-02-015	ssa-miR-430a-5p	5'	4	Pos.	60	61,7	-20,0
UAAGUGCUUCUCUUUGGGGUUG	na-09B02-02-015	hhi-miR-430b	3'	0	Pos.	78	55,1	-28,3

*Los miRNAs presentados han sido previamente descritos por otros autores en las especies *Cyprinus carpio*, *Danio rerio*, *Hippoglossus hippoglossus*, *Ictalurus punctatus* y *Salmo salar*.

Abreviaturas: LP – longitud del precursor (nt), Energía Libre – Energía libre mínima de la estructura secundaria, A+U – Porcentaje de adenina y uracilo en la longitud de la secuencia del precursor del miRNA maduro.

Anexo 2. Secuencias de *Solea senegalensis* con similitud a miRNAs maduros descritos previamente en otras especies

Secuencia madura predicha	Nombre de la secuencia	miRNA homólogo	Extremo	Mismatch	Cadena	LP	A + U	Energía libre	miRNAs similares
ACGUGUGUGUGUGCGUGCAUGU	01-01C02-04-001	mmu-miR-466f	5'	3	Pos.	71	47,9	-31,5	
AUAUAUGUAUAUGUAUAUAU	01-01C02-04-001	dme-miR-4979-3p	3'	4	Neg.	66	90,9	-27,8	
AUAAAAUAAUAAUAAUGA	01-01C02-04-001	gma-miR1533	3'	2	Pos.	67	94,0	-21,6	
CUUUUCACACACACACCAACA	01-05K05-01-001	cel-miR-253-5p	3'	2	Pos.	108	57,4	-38,9	
AUAUAUGUAUAUGUAUAUAU	01-05K05-01-001	dme-miR-4979-3p	5'	4	Pos.	93	89,2	-32,1	

UAUACAUAUAUACACAUAUAU	01-05K05-01-001	miR-467g	3'	2	Neg.	68	85,3	-24,4	
AUAUAUACACACACAAAUAUAC	01-05K05-01-003	rno-miR-466b-3p	3'	3	Neg.	74	81,1	-28,8	
UAUAUAUACACACACAAAUAU	01-05K05-01-003	mmu-miR-467g	3'	2	Neg.	73	82,2	-28,3	miR-467g *
AUACAUGUGUGUGUAUAUAUAU	01-05K05-01-003	dme-miR-4979-3p	5'	2	Neg.	73	82,2	-28,3	
UAUAUAUACACACACAUGUAU	01-05K05-01-003	mmu-miR-467g	3'	2	Pos.	84	82,1	-35,8	
CGGACGGCCAGCUGCAGGUGGCCG	01-05K05-01-184	efu-miR-8915	3'	2	Neg.	80	32,5	-38,9	
UAUACAUAUAUACACAUAUAU	01-05K05-02-002	miR-467g	3'	2	Neg.	68	85,3	-24,4	
AUAUAUGUAUAUGUAUAUAUAU	01-05K05-02-002	dme-miR-4979-3p	5'	4	Neg.	93	89,2	-32,1	
AUACAUGUGUGUGUAUAUAUAU	01-05K05-02-006	dme-miR-4979-3p	5'	2	Pos.	73	82,2	-28,3	
UAUAUAUACACACACAUGUAU	01-05K05-02-006	mmu-miR-467g	3'	2	Neg.	84	82,1	-35,8	miR-467g *
UAUAUAUACACACACAAAUAU	01-05K05-02-006	mmu-miR-467g	3'	2	Pos.	73	82,2	-28,3	
CUGAUCCCUGCUUCUUCGCAG	01-05K05-02-162	mmu-miR-3077-3p	5'	1	Neg.	60	45,0	-23,4	
GGUGAUUAUCCUGUUGACCUG	01-05K05-02-180	eca-miR-8995	3'	1	Neg.	62	51,6	-23,4	
CGGACGGCCAGCUGCAGGUGGCCG	01-05K05-02-317	efu-miR-8915	3'	2	Pos.	80	32,5	-38,9	

UAUGUGUGUGUGUAUAUAUAUA	01-10L10-03-001	mmu-miR-466a-5p	5'	4	Neg.	63	87,3	-27,0	
AUGUAUGUGUGUAUGUAUAUGU	01-10L10-03-001	mmu-miR-297a-5p	5'	3	Neg.	186	84,4	-99,0	mmu-miR-297b-5p mmu-miR-297c-5p
AUACAUAUAUAUGUAUAUAUAU	01-10L10-03-001	dme-miR-4979-3p	5'	4	Pos.	60	91,7	-24,7	
AUAUAUAUGUAUGUAUAUAUAU	01-10L10-03-001	dme-miR-4979-3p	5'	4	Neg.	60	91,7	-28,3	
CGGACGGCCAGCUGCAGGUGGCG	01-12D22-02-001	efu-miR-8915	3'	2	Neg.	81	32,1	-40,2	
CGGACGGCCAGCUGCAGGUGGCG	01-12D22-02-001	efu-miR-8915	3'	2	Neg.	81	32,1	-40,2	
CGGACGGCCAGCUGCAGGUGGCG	01-12D22-02-002	efu-miR-8915	3'	2	Neg.	81	33,3	-40,2	
CGGACGGCCAGCUGCAGGUGGCG	01-12D22-02-002	efu-miR-8915	3'	2	Pos.	81	32,1	-40,2	
CAUAUACAUAUAUAUAUAUAU	01-48P07-07-003	mmu-miR-669b-3p	5'	4	Pos.	61	91,8	-20,7	mmu-miR-467e-3p mmu-miR-669l-3p
UAUGUGUGUGUGUGUGUAUAUA	01-48P07-07-005	mmu-miR-466a-5p	5'	4	Pos.	65	83,1	-31,3	
UAUGUGUGUGUAUAUAUAUAUA	01-52C17-07-001	mmu-miR-466a-5p	5'	3	Pos.	71	85,9	-24,6	mmu-miR-466p-5p
UAUGUGUGUGUGUAUAUAUAUA	01-56H24-07-001	mmu-miR-466a-5p	3'	4	Neg.	97	88,7	-48,7	
AUUUGUGUGUGUAUAUAUGU	01-56H24-07-006	mmu-miR-669n	5'	3	Neg.	83	90,4	-36,2	

CCUCUUCUCCUCCUUGCUGU	02-04D15-06-001	hsa-miR-1976	5'	3	Neg.	119	50,4	-43,7	
GAGGAGAAGAGGGAGGAGGAG	02-04D15-06-001	ath-miR854a	3'	3	Pos.	127	49,6	-45,8	
AUUUGUGUGUGUGUGUGUGU	02-06P22-04-006	mmu-miR-669n	3'	2	Pos.	183	48,1	-70,2	
UACACACACACACACACAC	02-06P22-04-009	mmu-miR-466h-3p	5'	2	Neg.	95	51,6	-51,4	
AUACACACACACACAUACACA	02-06P22-04-017	rno-miR-466b-2-3p	5'	2	Neg.	117	57,3	-40,9	
UCACGUGACGGGAAUCGGCG	02-46C05-07-001	cfa-miR-1840	5'	2	Neg.	153	42,5	-61,7	
UAUAUAUAUACACACAUAUAU	02-46C05-07-002	mmu-miR-467g	3'	2	Neg.	97	88,7	-38,8	miR-467g *
AUAUAUGUGUGUAUAUAUAU	02-46C05-07-002	dme-miR-4979-3p	5'	1	Pos.	60	86,7	-27,3	
UAUAUAUAUACACACAUAUAU	02-46C05-07-002	mmu-miR-467g	5'	2	Pos.	60	86,7	-27,3	
AUAUAUGUGUGUAUAUAUAU	02-46C05-07-002	dme-miR-4979-3p	3'	1	Neg.	87	89,7	-35,5	miR-467g *
AUAGAUAUGUAUGUAUAUAUAU	02-46C05-07-003	dme-miR-4979-3p	5'	4	Neg.	92	73,9	-29,5	
AUAUACAUACAUUAUGUAUAU	02-46C05-07-003	mmu-miR-467e-3p	5'	4	Pos.	151	88,1	-71,8	
AUAUAUGUAUGUGUAUAUAUAU	02-46C05-07-003	dme-miR-4979-3p	3'	4	Neg.	71	94,4	-33,4	miR-467g *
AUGCUGCGGGGACUCUUCAAA	02-46C05-07-005	stu-miR7997a	5'	2	Neg.	111	44,1	-44,0	

AUAUACAUACACACGCGUGUAU	02-52G10-07-001	mmu-miR-467e-3p	5'	3	Pos.	78	75,6	-33,2	mmu-miR-669l-3p mmu-miR-467g
AUUUGUGUGUGUGUGUGUGU	02-52G10-07-001	mmu-miR-669n	3'	2	Pos.	74	55,4	-33,6	
UACACACACACACACACAC	02-52G10-07-001	mmu-miR-466h-3p	5'	2	Neg.	81	53,1	-42,8	
AUUCACACACACACACACAC	02-52G10-07-001	rno-miR-466b-3p	3'	3	Neg.	113	47,8	-63,6	
AUUCACACACACACACACAC	02-52G10-07-001	rno-miR-466b-3p	5'	3	Pos.	127	58,3	-54,9	
UACACACACACACACACAC	02-52G10-07-005	mmu-miR-466h-3p	3'	2	Neg.	129	49,6	-55,0	
AUAAUAAUAAUAAUAAUGA	02-52G10-07-005	gma-miR1533	3'	1	Pos.	75	90,7	-24,1	
AUAUACAUACAUUACAUACAU	03-09J04-03-001	mmu-miR-467e-3p	3'	4	Neg.	89	86,5	-30,1	mmu-miR-669l-3p
AUUUGUGUGUGUGUGUGUGU	03-09J04-03-002	mmu-miR-669n	3'	2	Pos.	83	60,2	-37,9	
AUAUAUGUACUGUAUAUAUAU	04-03C15-05-001	dme-miR-4979-3p	3'	2	Pos.	60	90,0	-21,8	
AUAUACAUACAUAGAUUACA	04-08A23-03-001	rno-miR-466b-2-3p	5'	3	Pos.	119	85,7	-47,2	mmu-miR-467d-3p
UAUGUGUGUGUAUAUAUAUAU	04-08A23-03-002	mmu-miR-466p-5p	3'	3	Neg.	69	79,7	-27,1	
GUAUAUAUGUACGUACAUAUG	04-08A23-03-002	dme-miR-9388-5p	3'	2	Neg.	60	75,0	-23,3	

UGAUGUGUGUGUAUAUGUAUUAU	04-08A23-03-002	mmu-miR-466b-5p	5'	2	Neg.	83	75,9	-32,7	
CAACCUCAUUUCUGAGGAUCUGG	04-12D24-04-001	gga-miR-7462-5p	5'	2	Pos.	113	55,8	-36,2	
UUCUGAGCUGAGGCCAG	04-12D24-04-001	hsa-miR-4303	5'	1	Pos.	77	40,3	-33,9	
UCAUUUUUGUGACUAUGCAACU	04-12D24-04-001	aca-miR-153-5p	5'	2	Pos.	60	58,3	-20,8	mdo-miR-153-1-5p spu-miR-153-5p ssa-miR-153b-5p
CAGCUGUUUCAACUACGGACAA	04-12D24-04-002	bmo-miR-2807b	3'	4	Neg.	66	54,5	-27,3	
UAGACACACACGCACACACAG	04-12N15-04-002	mmu-miR-466d-3p	5'	3	Pos.	84	48,8	-29,8	
UGUAUUUGCACCUGUACUUAUAC	04-12N15-04-009	bdi-miR530a	5'	3	Pos.	74	55,4	-26,1	osa-miR530-5p
GGCAAGUUGUCAAUUGACUAC	04-36H02-06-001	ath-miR169a-3p	5'	3	Pos.	83	54,2	-27,5	
UGUACGUACAUAUAUAUAUGU	04-36J02-06-002	dme-miR-9388-3p	3'	1	Neg.	93	80,6	-30,1	
UCAUCAUCAUCAUCAUCA	04-36J02-06-002	ath-miR414	5'	2	Neg.	211	59,7	-73,5	
AUAUAUGUAUAUGUAUAUAUAU	04-46B02-07-001	dme-miR-4979-3p	3'	4	Neg.	115	91,3	-63,0	
AUAUAUGUAUAUGUAUAUAUAU	04-46B02-07-001	dme-miR-4979-3p	5'	4	Pos.	88	92,0	-33,2	
AUAUAUGUAUAUGUAUAUAUAU	04-46B02-07-001	dme-miR-4979-3p	3'	4	Neg.	75	93,3	-35,3	

UAUGUGUGUGUAUAUGUAUAU	04-46B02-07-001	mmu-miR-466p-5p	3'	2	Neg.	116	82,8	-53,6	miR-467g *
AGGCAUGGUGGUGUAGUGGUUA	04-46B02-07-002	mdo-miR-885	3'	3	Neg.	191	69,1	-61,8	
GAGAGAGUGUGUGUGUGUGUGU	04-46B02-07-003	mmu-miR-3082-5p	5'	2	Neg.	76	59,2	-25,5	
AUAUAUGUAUGUGUAUAUAUAU	04-46B02-07-005	dme-miR-4979-3p	3'	4	Neg.	70	90,0	-34,2	
AUAUAUGUAUAUGUAUAUAUAU	06-10K23-06-002	dme-miR-4979-3p	5'	4	Pos.	131	93,1	-68,0	
AUAUAUGUAUAUGUAUAUAUAU	06-10K23-06-002	dme-miR-4979-3p	5'	4	Pos.	69	94,2	-32,1	
UGAGGUGCCUUUUUUACGUGUA	06-20D18-04-002	prd-miR-7966a-3p	5'	3	Neg.	132	50,0	-46,2	
AGUUUGUGUGUGUGUGUGUGUGU	06-20D18-04-003	mmu-miR-669f-5p	5'	4	Neg.	145	44,8	-56,3	mmu-miR-669l-5p
AGUUUGUGUGCAUGUGUGUGUGU	06-20D18-04-003	mmu-miR-669f-5p	5'	2	Neg.	99	54,5	-31,9	mmu-miR-669l-5p
UACACACACACACACACAC	07-19H19-04-001	mmu-miR-466h-3p	3'	2	Pos.	60	50,0	-26,0	
AUACAUUAUAUGUAUAUAUAU	09-32B08-04-002	dme-miR-4979-3p	3'	4	Neg.	108	92,6	-46,1	
AUAUACAUACAUUAUACAUUAU	09-32B08-04-002	mmu-miR-467e-3p	3'	3	Neg.	108	92,6	-46,1	mmu-miR-669l-3p miR-467g *
UACACACACACACACACAC	09-32B08-04-005	mmu-miR-466h-3p	3'	2	Pos.	60	51,7	-26,2	

AUAUAUAUGUAUGUAUAUAUAU	12-35D17-07-001	dme-miR-4979-3p	5'	4	Neg.	79	92,4	-38,4	
CGGGGAGGGAGGGAGGGAGGG	12-35D17-07-001	gga-miR-2127	3'	3	Pos.	100	37,0	-53,3	
CAGCAGCAGCAGCAGCAGCAGA	12-35D17-07-001	dme-miR-4968-3p	3'	1	Pos.	111	40,5	-45,9	
UGUGGGUGUGUGUGUGUGUGUG	12-35D17-07-004	bta-miR-669	3'	3	Neg.	108	49,1	-38,5	
AUAUAUGUGUGUGUAUAUAUAU	13-29D04-03-001	dme-miR-4979-3p	5'	3	Pos.	65	92,3	-29,3	mmu-miR-466a-5p mmu-miR-466p-5p
AUAUAUGUGUGUGUAUAUAUAU	13-29D04-03-001	dme-miR-4979-3p	5'	3	Pos.	70	92,9	-32,8	mmu-miR-466p-5p
UAUGUGUGUGUGUAUAUAUAUA	13-29D04-03-002	mmu-miR-466a-5p	3'	4	Neg.	102	85,3	-52,3	
UAUAUAUAUACACACAUAUAU	13-29D04-03-002	mmu-miR-467g	5'	2	Pos.	71	85,9	-26,9	miR-467g *
AUAUAUGUGUGUAUAUAUAU	13-29D04-03-002	dme-miR-4979-3p	3'	1	Neg.	71	85,9	-34,5	
CAUAUACAUAUAUACAUAUGUA	13-29D04-03-003	mdo-miR-7399-5p	5'	4	Neg.	118	92,4	-45,6	
CAGCCUGACAGGACCAG	13-29D04-03-004	hsa-miR-4293	5'	1	Neg.	95	44,2	-39,5	
AUUUGUGUGUGUGUGUGUGUGU	15-36E03-06-001	mmu-miR-669n	3'	2	Neg.	148	55,4	-51,2	
UAUGUGUGUGUGUGUGUAUAUA	15-36E03-06-005	mmu-miR-466a-5p	5'	4	Pos.	100	88,0	-46,5	

AUAUAUGUGUAUGUAUAUAUAU	15-36E03-06-005	dme-miR-4979-3p	3'	3	Neg.	84	94,0	-40,6	
UAUAUAUACAUACACAUUAUAU	15-36E03-06-005	miR-467g	5'	2	Neg.	87	88,5	-38,1	
AUAUACAUACACAUUAUAUAUAU	15-36E03-06-005	mmu-miR-467e-3p	5'	3	Pos.	84	90,5	-28,5	mmu-miR-669l-3p mmu-miR-467g
UUUGUUUUCAUUUUACAUUUUUUC	15-36E03-06-006	rgl-miR7797b	5'	3	Pos.	114	50,9	-45,9	
GUAUGUAUGUAUAUAUGUAUAUG	16-09N08-05-002	dme-miR-9388-5p	3'	4	Neg.	93	82,8	-42,3	
AUAUACAUACAUACAAUACA	16-09N08-05-002	rno-miR-466b-2-3p	5'	3	Pos.	165	78,2	-52,9	mmu-miR-669l-3p
UAUGUGUGUGUAUGUAUAUAU	16-09N08-05-002	mmu-miR-466a-5p	5'	1	Neg.	93	82,8	-42,3	
AUAUACAUACACACACAUUAUAU	16-09N08-05-002	mmu-miR-467e-3p	5'	1	Pos.	86	73,3	-36,8	mmu-miR-669l-3p mmu-miR-467g
UAUACAUACACACACAUUAUAU	16-09N08-05-002	miR-467g	5'	2	Neg.	86	73,3	-36,8	
AGCUCAGUUACUUUUUAACCA	16-52E07-07-001	hsa-miR-548v	3'	2	Pos.	63	58,7	-20,2	
AUGACUGACAUGCACU	17-31N01-06-001	bta-miR-2917	5'	1	Pos.	99	41,4	-50,2	
GAUGCUGCUGCUGCUGCU	17-31N01-06-001	ppy-miR-1322	5'	1	Pos.	114	38,6	-52,1	
GUAUUGAUUGGGUUUGUU	17-31N01-06-002	chi-miR-1814	5'	2	Neg.	122	52,5	-44,7	

UUUACCUAGUAGAGGAGGAGGAG	17-31N01-06-003	gga-miR-7466-3p	5'	2	Pos.	61	50,8	-21,9	
UAUACACACACACAUUAU	17-31N01-06-003	mmu-miR-467g	3'	0	Neg.	94	78,7	-32,6	miR-467g *
AUGUACAUACACACACACAU	17-31N01-06-003	mmu-miR-467e-3p	3'	3	Pos.	94	78,7	-30,4	mmu-miR-669l-3p
GGAGGAGGAGGAGGCGGGGG	17-63A07-02-002	hsa-miR-3960	5'	3	Neg.	118	39,0	-56,8	
UAUACAUACACAUUAUUAU	17-63A07-02-003	mmu-miR-467g	3'	2	Neg.	124	87,9	-49,9	miR-467g *
CAUAUACAUACACAUUAUUAU	17-63A07-02-003	mmu-miR-669b-3p	3'	3	Neg.	124	87,9	-49,9	mmu-miR-467e-3p mmu-miR-669l-3p
AUAUACAUACAUAAAUUAUUAU	17-63A07-05-002	mmu-miR-467e-3p	5'	4	Pos.	90	94,4	-37,5	mmu-miR-669l-3p
AUAUAUGUAUAUGUAUAUUAU	17-63A07-05-002	dme-miR-4979-3p	5'	4	Pos.	99	94,9	-46,9	
GGAGGAGGAGGAGGCGGGGG	17-63A07-05-002	hsa-miR-3960	5'	3	Neg.	118	39,0	-56,8	
AUACAUUAUAUGUAUAUUAU	17-63A07-05-002	dme-miR-4979-3p	3'	4	Neg.	68	91,2	-31,9	
AUAUAUAUGUAUGUAUAUUAU	17-63A07-05-002	dme-miR-4979-3p	5'	4	Pos.	65	92,3	-27,6	
AUACAUUAUAUGUAUAUUAU	17-63A07-05-002	dme-miR-4979-3p	3'	4	Neg.	63	92,1	-27,8	
AUACAUUAUAUGUAUAUUAU	17-63A07-05-002	dme-miR-4979-3p	3'	4	Pos.	60	91,7	-28,8	

ACGUGUGUGUGUAUGUGCAUGU	18-36M02-06-001	mmu-miR-466f	5'	1	Pos.	125	52,8	-41,0	
GUAUGUAUGUAUAUAUGUAUAUG	19-12K06-05-001	dme-miR-9388-5p	3'	4	Neg.	93	82,8	-42,3	
AUAUACAUAACAUAACAAUACA	19-12K06-05-001	rno-miR-466b-2-3p	5'	3	Pos.	165	78,2	-52,9	mmu-miR-669l-3p
UAUGUGUGUGUAUGUAUAUA	19-12K06-05-001	mmu-miR-466a-5p	5'	1	Neg.	93	82,8	-42,3	
AUAUACAUAACACACACAUUAU	19-12K06-05-001	mmu-miR-467e-3p	5'	1	Pos.	86	73,3	-36,8	mmu-miR-669l-3p mmu-miR-467g
UAUGUGUGUGUAUAUAUAUAUA	19-13F04-05-001	mmu-miR-466a-5p	3'	3	Neg.	65	80,0	-23,4	mmu-miR-466p-5p
AUAUAUGUAUAUGUAUAUAUAU	19-13F04-05-002	dme-miR-4979-3p	5'	4	Neg.	62	93,5	-27,7	
AUAAUAAUAAUAAUAAUGA	19-31C01-06-001	gma-miR1533	3'	1	Pos.	70	95,7	-25,8	
AUAAUAAUAAUAAUAAUGA	19-31C01-06-001	gma-miR1533	3'	1	Neg.	94	96,8	-30,8	
AUAUACAUAACACAUUAUACA	19-31C01-06-001	rno-miR-466b-2-3p	3'	2	Neg.	96	84,4	-41,7	mmu-miR-297a-3p mmu-miR-467g
UACGUGUGUGUGUGUGUGCGUG	19-31C01-06-003	mmu-miR-466f-5p	5'	3	Neg.	96	50,0	-47,0	
AUAAUAAUAAUAAUAAUGA	19-31C01-06-003	gma-miR1533	3'	1	Pos.	63	98,4	-21,4	
GACAGAUAGUGUGUGUGUGUGUGU	19-31C01-06-004	mmu-miR-3082-5p	5'	1	Neg.	92	54,3	-32,4	

CGUCUUUGUGUGUAUGUUACAG	19-31C01-06-008	dsi-miR-2580	5'	3	Pos.	121	51,2	-43,8	
AUUUGUGUGUGUGUGUGUGU	20-04M14-02-002	mmu-miR-669n	5'	0	Pos.	131	55,0	-45,7	
AUUUGUGUGUGUGUGUGUGU	21-30H22-04-001	mmu-miR-669n	3'	2	Pos.	116	74,1	-37,2	
UACGCACACACACACACAC	21-30H22-04-001	mmu-miR-466h-3p	3'	1	Pos.	97	53,6	-33,3	
AUACACACACACACACACAC	21-30H22-04-001	rno-miR-466b-3p	3'	2	Neg.	71	64,8	-52,7	mmu-miR-466h-3p rno-miR-466b-3p
AUAUAUAUGUAUACACAC	21-30H22-04-001	bta-miR-568	3'	2	Pos.	66	66,7	-45,9	
AUACACACACACACACACAC	21-30H22-04-001	rno-miR-466b-3p	3'	2	Pos.	64	67,2	-48,2	mmu-miR-466h-3p
AUAUAUAUGUAUGUAUAUAU	21-30H22-04-002	dme-miR-4979-3p	5'	4	Neg.	60	93,3	-26,5	
AUAUAUAUGUAUGUAUAUAU	21-30H22-04-002	dme-miR-4979-3p	5'	4	Neg.	85	91,8	-43,6	
AUAUACAUAUAUAUAUACA	21-30H22-04-002	rno-miR-466b-2-3p	3'	3	Pos.	84	90,5	-26,9	
AUAUAUAUGUAUGUAUAUAU	21-30H22-04-002	dme-miR-4979-3p	5'	4	Neg.	65	92,3	-31,2	
AUAUACAUAUAUAUAUACA	21-30H22-04-002	rno-miR-466b-2-3p	3'	3	Pos.	75	92,0	-24,7	
UACACACACACACACACAC	21-30H22-04-004	mmu-miR-466h-3p	3'	2	Pos.	126	47,6	-76,6	

CGGCUCGGCUCAGCUCGGCUC	21-30H22-04-004	gga-miR-1648-5p	5'	2	Neg.	78	48,7	-34,3	
AUAUAUGUAUAUGUAUAUAUAU	21-30H22-04-004	dme-miR-4979-3p	3'	4	Pos.	63	93,7	-21,3	
AUAGAUUAAGAUUAAGAUUAU	21-30H22-04-004	hsa-miR-620	5'	3	Neg.	98	90,8	-39,8	
GGCGGCGGCGGCUGCGGCGG	21-63A03-02-001	hsa-miR-3960	3'	3	Pos.	133	27,1	-69,6	
AUGCAGAGCAUCCAUCGUGC	21-63A03-02-002	cin-miR-3598-3p	3'	3	Pos.	169	49,1	-60,4	
UAUGUGUGUGUAUAUGUAUAU	21-63A03-02-009	mmu-miR-466p-5p	3'	2	Neg.	70	82,9	-25,9	miR-467g *
UAUGUGUGUGUAUAUACAUA	21-63A03-02-009	mmu-miR-466a-5p	3'	1	Pos.	62	80,6	-20,1	
ACAUGUAUGUACAUGUAUAUGU	21-63A03-02-076	mmu-miR-669d-5p	5'	3	Pos.	69	78,3	-32,7	
AUAUAUAUACGUGUAUAUAUAU	21-63A03-05-001	dme-miR-4979-3p	3'	4	Neg.	110	91,8	-46,0	
UAUGUGUGUGUAUAUACAUA	21-63A03-05-005	mmu-miR-466a-5p	3'	1	Neg.	62	80,6	-20,1	miR-467g *
UAUGUGUGUGUAUAUGUAUAU	21-63A03-05-005	mmu-miR-466p-5p	3'	2	Pos.	70	82,9	-25,9	
ACAUGUAUGUACAUGUAUAUGU	21-63A03-05-012	mmu-miR-669d-5p	5'	3	Pos.	69	78,3	-32,7	
UAUGUGUGUGUAUAUGUAUAUA	21-63A03-05-013	mmu-miR-466a-5p	5'	2	Neg.	64	81,3	-24,2	mmu-miR-466p-5p miR-467g *

UAUACAUAUACACACACAUAUAU	21-63A03-05-013	mmu-miR-467g	5'	0	Pos.	76	82,9	-35,8	
UAUACACACACACACAUAUAU	21-63A03-05-013	mmu-miR-467g	5'	1	Neg.	64	81,3	-24,2	miR-467g *
GGCGGGCGGGCGGUCGGCGGG	21-63A03-05-017	hsa-miR-3960	3'	3	Pos.	133	27,1	-69,6	
UAGCUCAGUGGUAGAG	99-08O07-04-001	miR-5112	5'	4	Pos.	83	49,4	-29,0	
UGUGGGUGUGUGUGUGUGUGUG	99-08O07-04-001	bta-miR-669	5'	3	Pos.	61	54,1	-26,0	
AUAUAUGUAUAUGUAUAUAUAU	99-08O07-04-003	dme-miR-4979-3p	3'	4	Pos.	66	93,9	-30,2	
AUAUAUAUGUAUAUGUAUAUAUAU	99-08O07-04-003	dme-miR-4979-3p	3'	4	Pos.	80	92,5	-38,7	
AUAUAUGUAUAUGUAUAUAUAU	99-08O07-04-003	dme-miR-4979-3p	3'	4	Pos.	136	91,9	-70,8	
CAGUAAUGAUCUCUGCUGGGC	99-16E16-03-008	hsa-miR-8079	5'	3	Pos.	85	56,5	-28,6	
AUAGAUUAAGAUUAAGAUUAU	99-16E16-05-002	hsa-miR-620	3'	3	Pos.	114	90,4	-39,6	
AUAUGACUAUUUAUUUGUGCU	99-19J21-03-026	bfl-miR-4875a-5p	3'	3	Pos.	71	56,3	-26,8	
UCGGAGGUGUUUGUGUACCUGU	99-19J21-03-029	bta-miR-6522	3'	2	Pos.	64	51,6	-25,8	
AAACAGGCUCUGUCAAGCA	99-19J21-03-050	oan-miR-218-3p	5'	1	Neg.	85	57,6	-27,6	
AUAUACAUAUAUAUAUAUAU	99-21O23-05-001	mmu-miR-467e-3p	3'	4	Neg.	92	96,7	-33,3	mmu-miR-669l-3p

AUAUAUAUGUAUGUAUAUAUAU	99-21023-05-001	dme-miR-4979-3p	5'	4	Pos.	91	96,7	-44,7	
AUAUAUGUGUGUAUAUAUAU	99-21023-05-001	dme-miR-4979-3p	5'	1	Pos.	69	88,4	-32,1	
AGUUGUGUGUGUGUGUGUGUGUGU	99-30J04-03-003	mmu-miR-669f-5p	3'	4	Pos.	64	53,1	-22,7	mmu-miR-669l-5p
GGGAGGCGUCCAGGAGGC	99-30J04-03-004	rno-miR-3584-5p	3'	1	Pos.	66	36,4	-37,1	
AUACAUAUAUAUGUAUAUAUAU	99-30J04-03-004	dme-miR-4979-3p	5'	4	Pos.	122	92,6	-58,8	
AUAUAUGUAUAUGUAUAUAUAU	99-30J04-03-004	dme-miR-4979-3p	5'	4	Pos.	67	95,5	-31,0	
UAUGUGUGUGUAUAUAUAUAUA	99-30J04-03-005	mmu-miR-466a-5p	5'	3	Neg.	85	89,4	-38,3	mmu-miR-466p-5p
UUUUUJAGGGUUCGAGCAAC	99-48K07-07-002	sly-miR5303	5'	1	Pos.	61	63,9	-21,5	
UCAGGGAGUCAACGGAGCGC	99-48K07-07-004	hsa-miR-4270	5'	3	Pos.	61	49,2	-21,7	
AUAUAUAUGUAUGUAUAUAUAU	na-00000-01-005	dme-miR-4979-3p	3'	4	Neg.	60	96,7	-22,7	
GUGGGUUGGGGGGGGGGG	na-00000-01-005	oan-miR-1346	5'	1	Neg.	129	36,4	-61,1	
UAUGUGUGUGUAUAUAUAUAUA	na-00000-01-013	mmu-miR-466a-5p	5'	3	Pos.	70	91,4	-32,8	mmu-miR-466p-5p
UGUGUGUGCGUGUGUGUACAUG	na-00000-01-016	mmu-miR-466d-5p	5'	3	Neg.	67	53,7	-34,1	
AUAAGUGUGUGUGUGUGUGUGU	na-00000-01-016	mmu-miR-467h	3'	4	Pos.	113	50,4	-42,7	

UAUGUGUGUGUAUGUGUGCAU	na-00000-01-017	mmu-miR-466p-5p	5'	3	Pos.	129	46,5	-48,0	
UAUGUGUGUGUAUAUAUAUAUA	na-00000-01-060	mmu-miR-466a-5p	5'	3	Pos.	60	83,3	-29,2	mmu-miR-466p-5p
UAUGUGUGUGUAUAGUAUAUA	na-00000-01-090	mmu-miR-466a-5p	5'	2	Pos.	153	92,2	-65,8	
GGGAAAGAGGUGGGUGUGGU	na-00000-01-314	lja-miR7535	5'	1	Pos.	114	61,4	-36,7	
AUUUGUGUGUGUGUGUGUGUGU	na-00000-01-880	mmu-miR-669n	3'	2	Neg.	67	53,7	-26,2	
UGUGCAUGUGUGCAUGUGUGUG	na-11O20-03-001	bta-miR-669	5'	3	Neg.	119	55,5	-44,4	
UGGUUUUAUAUGUGUGUGG	na-11O20-03-001	cin-miR-4194-3p	5'	1	Pos.	179	53,6	-57,5	
UGUGCAUGUGUGCAUGUGUGUG	na-11O20-04-001	bta-miR-669	5'	3	Neg.	119	55,5	-44,4	
UGGUUUUAUAUGUGUGUGG	na-11O20-04-001	cin-miR-4194-3p	5'	1	Pos.	179	53,6	-57,5	
UGUGCAUGUGUGCAUGUGUGUG	na-11O20-06-001	bta-miR-669	5'	3	Neg.	119	55,5	-44,4	
UGGUUUUAUAUGUGUGUGG	na-11O20-06-001	cin-miR-4194-3p	5'	1	Pos.	179	53,6	-57,5	
UAUACAUACACACACAUUAU	na-12K16-02-001	mmu-miR-467g	5'	0	Neg.	86	73,3	-36,8	miR-467g *
AUAUACAUACACACACAUUAU	na-12K16-02-001	mmu-miR-467e-3p	5'	1	Neg.	86	73,3	-36,8	mmu-miR-669l-3p
UAUGUGUGUGUAUGUAUAUA	na-12K16-02-001	mmu-miR-466a-5p	5'	1	Pos.	93	82,8	-42,3	

AUAUACAUAACAUAACAAUACAUA	na-12K16-02-001	mmu-miR-669l-3p	5'	4	Neg.	165	78,2	-52,9	rno-miR-466b-2-3p
GUAUGUAUGUAUAUAUGUAUAUG	na-12K16-02-001	dme-miR-9388-5p	3'	4	Pos.	93	82,8	-42,3	
UGGGUCAGACGGGGCGUCUCAG	na-31B01-07-001	bbe-miR-4857-3p	3'	4	Neg.	87	25,3	-50,5	
GCUGGCGGGAGCCCCGGGGAG	na-31B01-07-001	mmu-miR-2137	5'	2	Pos.	139	33,8	-68,2	
GACAUGAGUGUGUGUGUGUGUGU	na-31O01-06-001	mmu-miR-3082-5p	3'	1	Pos.	133	52,6	-50,9	
AUAUAUGUAUGUGUAUAUAUAU	na-31O01-06-002	dme-miR-4979-3p	3'	4	Pos.	84	89,3	-39,3	miR-467g *
AUAUAUGUGUAUGUAUAUAUAU	na-31O01-06-002	dme-miR-4979-3p	3'	3	Neg.	67	86,6	-31,2	miR-467g *
UAUGUGUGUGUAUAUAUAUAUA	na-31O01-06-002	mmu-miR-466a-5p	5'	3	Pos.	61	88,5	-27,3	mmu-miR-466p-5p
AGAGGCAGUAGAUACUGUUUU	na-38N10-08-005	gga-miR-1617	3'	2	Pos.	79	60,8	-25,6	
UAGCUCAGUGGUAGAG	na-38N10-08-007	miR-5112	3'	4	Pos.	82	45,1	-31,7	
UAUGUGUGUGUAUAUAUAACAUA	na-46P22-08-011	mmu-miR-466p-5p	3'	2	Neg.	60	80,0	-19,3	
UGAUUUUGUUGUUGUUGGAUU	na-46P22-08-013	bta-miR-2293	5'	1	Pos.	60	60,0	-20,6	
AGUUGUGUGUGUGUGUGUGUGU	na-46P22-08-027	mmu-miR-669f-5p	5'	4	Neg.	87	47,1	-39,8	mmu-miR-669l-5p
UGUGGCCUGUGUGUGUGUCC	na-50K03-08-001	efu-miR-9318a	5'	3	Pos.	72	50,0	-27,2	

CUAUACAGUCUAUUACCUUCCU	na-50K03-08-014	gga-let-7j-3p	3'	1	Pos.	85	57,6	-40,9	oha-let-7a-3-3p
UGAGGUAGUAGAUUGAAUAGUUGUG	na-50K03-08-014	efu-let-7f	5'	1	Pos.	71	59,2	-27,2	aca-let-7a-5p aca-let-7c-5p aca-let-7e-5p aca-let-7f-5p aca-let-7g ame-let-7 ccr-let-7g cfa-let-7j gga-let-7g-5p ggo-let-7c ggo-let-7f hsa-miR-1827 lva-let-7-5p mmu-miR-1961 mse-let-7a ola-let-7e ola-let-7g pma-let-7b-5p pma-let-7c pma-let-7d ssa-let-7f-5p ssa-let-7j-5p xtr-let-7b xtr-let-7e

UAACUAUGCAAUCUACUGUCUUUC	na-50K03-08-014	aca-let-7e-3p	3'	1	Pos.	72	58,3	-27,2	aca-let-7a-3p aca-let-7f-1-3p bta-let-7a-3p chi-let-7a-3p hsa-let-7f-2-3p ssa-let-7a-3-3p
UGAGGUAGUAGGUUGUAUAGUUU	na-50K03-08-014	prd-let-7-5p	5'	0	Pos.	85	57,6	-40,9	aca-let-7a-5p aca-let-7c-5p aca-let-7e-5p aca-let-7f-5p aae-let-7 aca-let-7g ame-let-7 bta-let-7b bta-let-7e ccr-let-7g cfa-let-7e cin-let-7b-5p csa-let-7c gga-let-7g-5p ggo-let-7c ggo-let-7f lva-let-7-5p mse-let-7a ola-let-7e

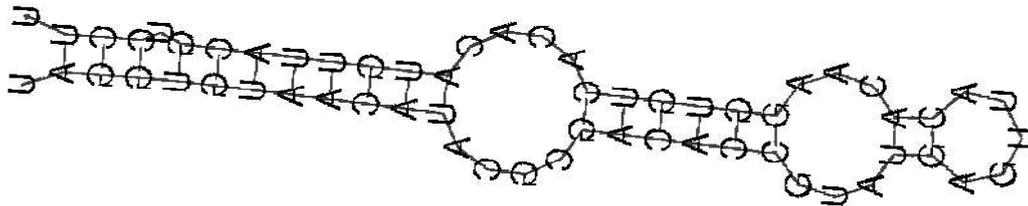
									ola-let-7g pma-let-7b-5p pma-let-7c pma-let-7d ssa-let-7f-5p ssa-let-7j-5p xtr-let-7b xtr-let-7e
CGUCGCGUCGUCGUUCGCUCUG	na-53D20-08-001	sbi-miR5386	5'	3	Neg.	97	26,8	-50,6	
AUAUAUGUGUGUAUAUAUAU	na-53D20-08-001	dme-miR-4979-3p	5'	1	Pos.	112	87,5	-46,7	
GGGGGCGUGUGGGUGGGGCGGGGGGG	na-53D20-08-005	hsa-miR-6089	3'	4	Pos.	114	31,6	-58,5	
AUAUAUGUAUAUGUAUAUAUAU	na-53D20-08-005	dme-miR-4979-3p	3'	4	Neg.	86	95,3	-40,1	
AUAUAUGUAUAUGUAUAUAUAU	na-71N11-08-002	dme-miR-4979-3p	3'	4	Pos.	62	88,7	-23,0	
UACACACACACACACACAC	na-71N11-08-013	mmu-miR-466h-3p	5'	2	Neg.	136	51,5	-80,6	
UAUGUGUGUGUGUAUAUAUAUA	na-71N11-08-033	mmu-miR-466a-5p	5'	4	Pos.	68	76,5	-23,1	
AUACAUAUAUAUGUAUAUAUAU	na-71N11-08-034	dme-miR-4979-3p	3'	4	Pos.	67	74,6	-22,9	
AUACAUAUAUAUGUAUAUAUAU	na-71N11-08-034	dme-miR-4979-3p	5'	4	Pos.	60	86,7	-27,1	

GAUGAUGCUGCUGCUGCUG	na-71N11-08-045	hsa-miR-1322	5'	1	Pos.	101	40,6	-43,4	ppy-miR-1322
AUAUAUGUGUGUGUAUAUAU	na-71N11-08-048	dme-miR-4979-3p	3'	3	Pos.	105	85,7	-37,2	

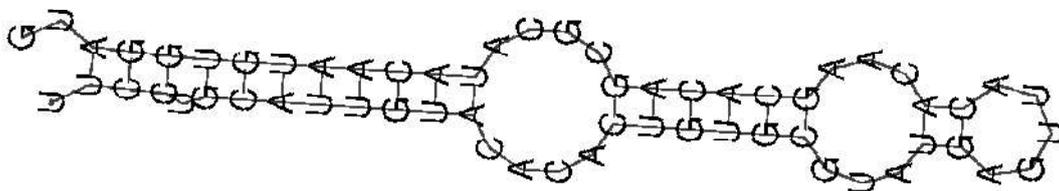
* Secuencias con similitud a miRNAs maduros descritos por Campos y otros autores (2014) en *Solea senegalensis*.

Abreviaturas: LP – longitud del precursor (nt), Energía Libre – Energía libre mínima de la estructura secundaria, A+U – Porcentaje de adenina y uracilo en la longitud de la secuencia del precursor del miRNA maduro.

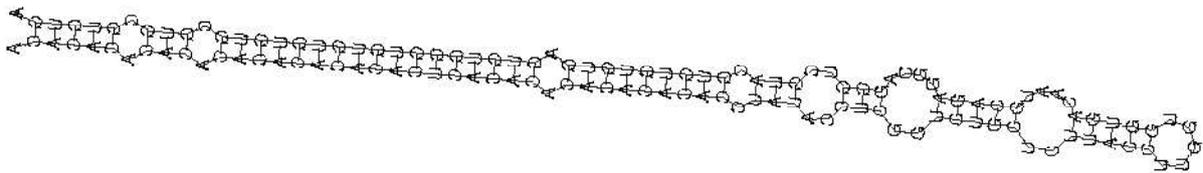
Anexo 3. Estructuras secundarias generadas con RNAfold de las secuencias identificadas con similitud a miRNAs previamente descritos en peces.



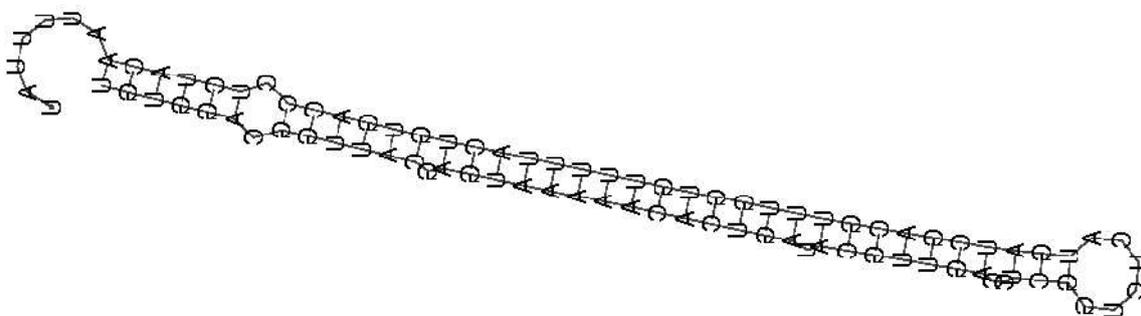
Secuencia 01-48P07-07-005, con similitud a ccr-miR-460-5p.



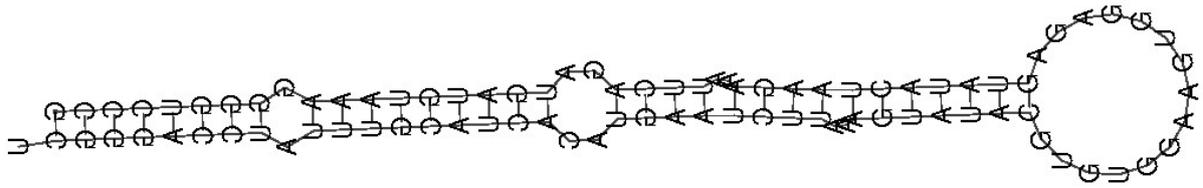
Secuencia 01-48P07-07-005, con similitud a ccr-miR-460-3p.



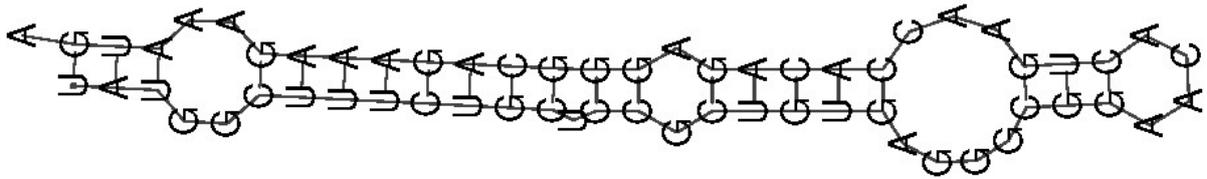
Secuencia 02-04D15-06-001, con similitud a ipu-miR-7562.



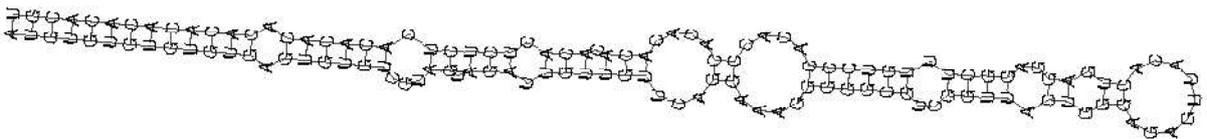
Secuencia 04-12D24-04-001, con similitud a ccr-miR-153b y dre-miR-153b-5p.



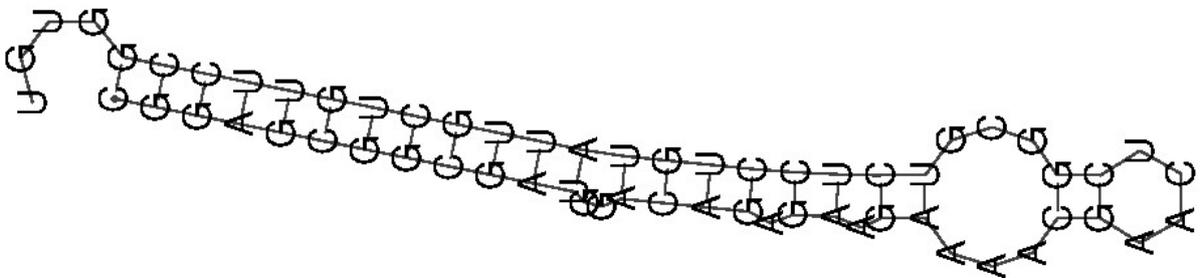
Secuencias 06-10K23-03-004 y 06-10K23-06-002, con similitud a hhi-miR-728.



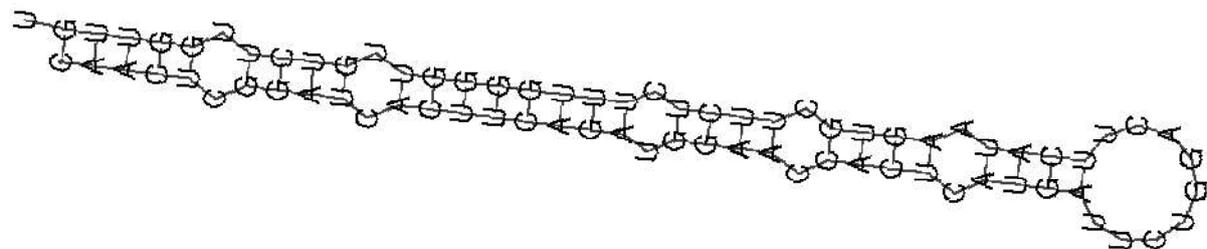
Secuencia 16-09N08-05-001, con similitud a ccr-miR-135c.



Secuencia 19-31C01-06-008, con similitud a ipu-miR-7562.



Secuencia na-00000-01-685, con similitud a dre-miR-457b-3p.



Secuencia na-09B02-02-015, con similitud a ssa-miR-430a-5p y hhi-miR-430b.

Anexo 4. Longitud en nucleótidos de cromosomas y BACs secuenciados.

Cromosoma	Longitud	Cromosoma	Longitud
1	1398343	13	198050
2	772879	14	0
3	139953	15	106719
4	887310	16	317315
5	0	17	253483
6	312908	18	77992
7	180593	19	275289
8	143889	20	275860
9	196621	21	367032
10	203659	99	879305
11	120960	na	3410464
12	135028		

BAC	Longitud	BAC	Longitud
00000	1827524	30P17	84700
01C02	44432	31A01	75031
02K18	63981	31A02	68858
03C15	78966	31B01	10108
03N10	54814	31C01	166654
04D15	73048	31F01	50345
04E10	70615	31N01	69460
04F12	65621	31O01	49077
04M14	275860	32B08	196621
05K05	575396	35D17	112174
06P22	273797	36D03	63491
07H22	109792	36E03	41098
08A23	100687	36H02	65994
08O07	175561	36H03	63473
09B02	24620	36I03	33278
09J04	139953	36J02	56196
09N08	50351	36K01	29618
10K23	92591	36M02	77992
10L10	65053	38F24	55286
11O20	89778	38N10	202923
12D22	45979	39F02	77120
12D24	183408	46B02	175688
12K06	19946	46C05	190621
12K16	17928	46P22	86895
12N15	162898	48K07	180050
13B02	64647	48P07	246624

13F02	22854	50K03	226224
13F04	88689	52C17	185957
13G01	14876	52E07	182264
15I19	203659	52G10	172517
16E16	43095	53D20	91167
19H19	70801	55B12	44536
19J21	54085	56H24	156535
20D18	220317	60P19	174674
21O23	53932	60P24	55628
28F09	24474	63A03	177015
29D04	198050	63A07	184023
30H22	190017	71N11	169060
30J04	372582		

Anexo 5. Relación de BACs con genes diana

BAC	Gen de interés	BAC	Gen de interés
10L10	calr	20D18	aqp3
16E16	dmrt2	07H22	fsh
05K05	h3	19H19	cyp19a1a
48P07	aqp1	08O07	sox6
48K07	dmrt1	31A01	anónimo
56H24	aqp10	31A02	anónimo
12D22	h4	32B08	sox3
36D03	anónimo	15I19	anónimo
01C02	anónimo	04E10	anónimo
52C17	sox17	31F01	anónimo
13G01	anónimo	30J04	tlr8
04D15	hmgb2 y vtc	35D17	sox9
06P22	nos3	13F02	anónimo
19J21	sox9a	29D04	tlr3
52G10	irf5	08O07	sox6
46C05	anónimo	21O23	dmrt4
36I03	anónimo	19J21	sox9a
21O23	dmrt4	04F12	anónimo
36K01	anónimo	16E16	dmrt2
09J04	anónimo	36E03	anónimo
46B02	sox21	30P17	anónimo
12N15	vasa	52E07	gapdh
30J04	tlr8	09N08	sema7a
08A23	mstn	31N01	anónimo
36J02	anónimo	63A07	anónimo
03C15	anónimo	36M02	anónimo

36H03	anónimo	12K06	mx
12D24	casq 2a	13F04	anónimo
36H02	anónimo	31C01	trim16
48K07	dmrt1	04M14	trhr3
16E16	dmrt2	63A03	anónimo
10K23	anónimo	30H22	amh