



UNIVERSIDAD DE OVIEDO

**Departamento de Biología Funcional
Área de Genética**

Programa de Doctorado Biología Funcional y Molecular

TESIS DOCTORAL

**MEJORA GENÉTICA EN DORADA
(*Sparus aurata*): OPTIMIZACIÓN DE
MARCADORES GENÓMICOS**

Carmen García Fernández

Oviedo, 2014



UNIVERSIDAD DE OVIEDO

**Departamento de Biología Funcional
Área de Genética**

Programa de Doctorado Biología Funcional y Molecular

TESIS DOCTORAL

**MEJORA GENÉTICA EN DORADA
(*Sparus aurata*): OPTIMIZACIÓN DE
MARCADORES GENÓMICOS**

Carmen García Fernández

Oviedo, 2014

detectaron 10 reproductores con contribución preferencial hacia descendientes con crecimiento rápido (fenotipo de interés productivo) y 15 hacia lento. Asumiendo que bajo contribuciones preferenciales hacia fenotipos opuestos subyacen componentes genéticos diferentes, la selección de individuos que contribuyan a potenciar el fenotipo de interés conllevará una mejora genética. Asimismo, la comparación genética de individuos con un componente genético hacia fenotipos opuestos, resulta útil para la identificación de variantes genómicas asociadas al carácter.

La detección de estas variantes genómicas se ve favorecida por el conocimiento previo de los genes implicados en el control del carácter, razón por la cual la primera parte del *Capítulo III* se dirige a la búsqueda de genes candidatos para el carácter “calidad de huevo”. En esta ocasión, para 11 de los 16 genes analizados por qRT-PCR, se observó un patrón de expresión propio de genes maternos. Debido al papel crucial que los ARNm maternos juegan sobre la calidad de huevo, estos 11 genes constituyen indicadores potenciales para el carácter. La segunda parte de este capítulo describe la caracterización del gen transferrina (*Tf*), un gen multifunción propuesto en numerosas especies como candidato para resistencia a enfermedades y crecimiento. A partir de su secuenciación se obtuvo una estructura altamente conservada respecto a otras especies de peces, compuesta por una región no codificante de 5.495pb (16 intrones) y una región codificante de 2.076pb (17 exones) de la que se deduce una proteína de 691aa. La caracterización de *Tf* se completó con un análisis de expresión génica a través del cual se refleja una expresión constitutiva dependiente de tejido y ontogenia.

Buscando esa conexión variante-carácter, el *Capítulo IV* se dedica a la búsqueda de marcadores SNPs asociados a crecimiento. A partir del análisis de la secuencia de 11 genes candidatos para crecimiento en 30 individuos, se generó un panel final compuesto por 164 SNPs y 3 polimorfismos de longitud intrónica (ILPs). Para determinar la existencia de una asociación marcador-peso corporal, se genotiparon 2 ILPs y 41 SNPs en 53 reproductores y 389 descendientes procedentes de dos lotes de cultivo, y se infirió el diplotipo de cada individuo. Finalmente, el análisis estadístico estableció para los diplotipos GCDT/GCDT del gen PRL, DGG/DGG del gen MyoD-1 y ACA/CCA del gen MyoD-2, una asociación significativa con el peso corporal, por lo que el análisis de estos 10 marcadores podría resultar de interés en futuros programas de selección asistidos por marcadores moleculares (MAS) o genes (GAS).

RESUMEN (en Inglés)

Leaving behind the concept of complementary secondary activity to fishing, aquaculture has evolved into an indispensable productive sector to meet the current demand for food. However, despite its large expansion, the future stability of the sector requires improvements to achieve a more favorable cost/benefit balance. Thus, the management plan based on genetic criteria is a useful tool to enhance productivity. In this context, the present doctoral dissertation addresses different topics of interest for the genetic improvement of gilthead sea bream (*Sparus aurata*), one of the most important aquaculture species in the national and European market.

Given the importance of achieving a reproductive control that promotes a long-term stability of the fish culture, the *Chapter I* shows how, from the microsatellite parentage analysis carried out from spontaneous mass spawning, the reproductive dynamics in captivity of gilthead sea bream leads to a rate of inbreeding (ΔF) higher than expected, jeopardizing the viability of the culture. Nevertheless, the monitoring over time of this dynamics suggested that the unification of offspring from independent spawning event, as well as, the reproductive management based on a temporal division of the standard broodstock ($N \geq 40$) in *mini-broodstocks* ($N \approx 8$) should be considered strategies of interest to minimize the generated ΔF .

Underlined the importance of the inbreeding control on the future of fish farming, the remaining chapters are based on the increase of productivity through genetic improvement of productive traits.

Chapter II deals with the identification of breeders carriers of a favourable genetic component related with a certain phenotype, key information to accelerate the response to selection. Two broodstocks and their offspring collected from a single spawning event and divided into fast- and slow- growth groups were genotyped using microsatellite *loci*. Through parentage assignment testing, breeders with significantly different contribution to the fast- (10 breeders) and to the slow- (15 breeders) growth progeny groups were identified. Assuming that different genetic components are hidden after this differential parental contribution, the selection of fast-growth (phenotype of productive interest) progeny as future breeders will lead to genetically improved culture. Additionally, genotype comparison of breeders with preferential contributions to opposite phenotypes facilitate the identification of genomic variants linked to traits.

The detection of these variants is enhanced by the previous knowledge of the genes involved in the trait control. Based on this, first part of Chapter III is focused on the search of candidate genes for egg quality. On this occasion, 11 out of 16 genes analyzed by qRT-PCR showed a characteristic gene expression pattern pinpointing to the possible function as maternal genes. According to the crucial role of maternal mRNAs in egg quality, these 11 genes may function as

indicators for this trait. The second part of the chapter describes the characterization of transferrin gene (*Tf*), a multi-functional gene which has been proposed as candidate gene disease resistance and growth in diverse species. Gene structure obtained via PCR consisted of a coding region of 2076 nucleotides divided into 17 exons, and no-coding region that includes 16 introns and spans 5495 nucleotides. The deduced Tf protein is composed of 691 amino acids and consists of an initial peptide and two lobes. This gene and protein structure is similar to those previously described in other species. Gene characterization was completed with expression analysis in 7 different tissues from two developmental stages (pre-ongrowing and adult) which showed tissue- and stage-specific *Tf* pattern.

On the search for that variant-trait connection, *Chapter IV* is based on the identification of SNPs markers associated with body weight. A final panel of 164 SNPs and 3 Intron Length Polymorphisms (ILPs) were designed from the sequence analysis of 11 candidate growth genes in 30 unrelated individuals. The putative existence of marker-body weight associations was determined through genotyping of 2 ILPs and 41 SNPs in 53 breeders and 389 offspring from two broodstocks. Genotyping data were used to determine the diplotype configuration (maternal and paternal haplotype) of each individual. Finally, the diplotypes GCDD/GCDD, DGG/DGG and ACA/CCA in *PRL*, *MyoD-1* y *MyoD-2* genes, respectively, were significantly associated with body weight. These results suggest that these 10 markers should be considered as potential markers for marker (MAS) or gene assisted selection (GAS) in future breeding programs.

1. Introducción general	1
1.1. Importancia del sector acuícola: la dorada como especie de interés	3
1.2. La genética como herramienta para potenciar la eficiencia de los cultivos acuícolas	6
1.2.1. Creación y manejo de lotes reproductores	6
1.2.2. Mejora genética de caracteres de interés	13
2. Objetivos	23
3. Capítulo I: Optimización del manejo reproductivo	27
3.1. Introducción	29
3.2. Materiales y métodos	31
3.2.1. Muestras biológicas.....	31
3.2.2. Extracción de ADN	31
3.2.3. Creación de la PCR multiplex	32
3.2.4. Evaluación del estadio de desarrollo embrionario óptimo para amplificación por PCR.....	33
3.2.5. Análisis de variabilidad y diversidad genética.....	33
3.2.6. Asignación parental.....	34
3.2.7. Estimación del tamaño efectivo e incremento de endogamia.....	35
3.2.8. Análisis complementarios	36
3.3. Resultados.....	37
3.3.1. Características del periodo de puesta	37
3.3.2. Estadio de desarrollo embrionario óptimo para amplificación por PCR	37
3.3.3. Puesta a punto de la PCR múltiplex <i>OVIDOPLEX-R</i>	37
3.3.4. Caracterización de la diversidad genética.....	38
3.3.5. Contribución parental	39
3.3.6. Tamaño efectivo.....	41
3.4. Discusión.....	42
3.4.1. <i>OVIDOPLEX-R</i> : una herramienta eficaz de genotipado.....	42
3.4.2. Evaluación de la dinámica reproductiva a partir de descendientes en estadio de desarrollo temprano: ventajas y dificultades	43
3.4.3. Características de la dinámica reproductiva.....	44
3.4.4. Consecuencias genéticas de la dinámica reproductiva: propuesta de manejo.....	46
3.5. Tablas	51

3.6. Figuras.....	59
4. Capítulo II: Uso de loci microsatélites para la identificación de reproductores de Interés productivo	67
4.1. Introducción.....	69
4.2. Materiales y métodos.....	71
4.2.1. Muestras biológicas	71
4.2.2. Extracción de ADN y desarrollo de la PCR multiplex OVIDORPLEX.....	72
4.2.3. Estimación de parámetros genéticos	73
4.2.4. Evaluación del crecimiento y contribución parental.....	73
4.3. Resultados	75
4.3.1. Creación de grupos de descendencias con crecimiento diferencial	75
4.3.2. PCR multiplex OVIDORPLEX	75
4.3.3. Caracterización de la diversidad genética.....	75
4.3.4. Tamaño efectivo y participación parental	76
4.4. Discusión.....	78
4.4.1. Sistema de genotipado OVIDORPLEX.....	78
4.4.2. Dinámica reproductiva: implicaciones de una contribución parental diferencial.....	79
4.5. Tablas	85
4.6. Figuras	89
5. Capítulo III: Identificación y caracterización de genes candidatos para calidad de huevo y resistencia a enfermedades	93
5.1. Introducción	95
5.2. Materiales y métodos.....	99
5.2.1. Búsqueda de genes candidatos para calidad de huevo	99
5.2.1.1. Muestras biológicas.....	99
5.2.1.2. Extracción de ARN y síntesis de ADNc.....	99
5.2.1.3. Búsqueda y análisis de secuencias	100
5.2.1.4. Diseño de iniciadores	101
5.2.1.5. Cuantificación de la expresión génica	101
5.2.1.6. Análisis estadístico	102
5.2.2. Caracterización del gen codificante para transferrina (<i>Tf</i>) candidato para resistencia a enfermedades	103

5.2.2.1. Muestras biológicas.....	103
5.2.2.2. Extracción de ADN, ARN y síntesis de ADNc.....	103
5.2.2.3. Búsqueda y análisis de secuencias	104
5.2.2.4. Diseño de iniciadores, amplificación y secuenciación	104
5.2.2.5. Análisis del patrón de expresión génica.....	106
5.2.2.6. Análisis filogenético	106
5.3. Resultados.....	108
5.3.1. Identificación y análisis de expresión de genes candidatos para calidad de huevo.....	108
5.3.2. Caracterización del gen <i>Tf</i>	109
5.4. Discusión.....	111
5.4.1. Identificación de genes candidatos para calidad de huevo	111
5.4.2. Caracterización gen <i>Tf</i> : un primer paso hacia la búsqueda de polimorfismos funcionales	116
5.5. Tablas	121
5.6. Figuras	129
6. Capítulo IV: Búsqueda de SNPs funcionales en genes candidatos para crecimiento.....	139
6.1. Introducción.....	141
6.2. Materiales y métodos	146
6.2.1. Muestras biológicas.....	146
6.2.2. Extracción de ADN	146
6.2.3. Búsqueda de variación SNP	146
6.2.4. Determinación del genotipo.....	147
6.2.4.1. Genotipado masivo de SNPs	147
6.2.4.2. Genotipado para deleciones/inserciones de gran tamaño	149
6.2.5. Identificación de haplotipos y análisis de desequilibrio de ligamiento.....	149
6.2.6. Análisis estadísticos	151
6.3. Resultados.....	152
6.3.1. Construcción y caracterización del panel de SNPs.....	152
6.3.2. Identificación de ILPs	153
6.3.3. Diseño y validación de ensayos de genotipado masivo de SNPs e ILPs.....	153
6.3.4. Reconstrucción de haplotipos y análisis de desequilibrio de ligamiento	154
6.3.5. Asociación entre SNPs/ILPs y el peso corporal.....	154
6.4. Discusión.....	156

6.4.1. Panel de SNPs en genes candidatos para crecimiento.....	156
6.4.2. Validación SNPs e ILPS	159
6.4.3. Asociación entre SNPs/ILPs y el peso corporal.....	161
6.5. Tablas	167
6.6. Figuras	185
7. Conclusiones	193
7(Bis). Conclusions	197
8. Bibliografía	201
9. Anexo: artículos científicos publicados asociados a esta tesis doctoral	223

1. Introducción general

1.1. Importancia del sector acuícola: la dorada como especie de interés

Desde su origen en 1950 como una actividad socio-económica relevante, la acuicultura, entendiéndose como tal el cultivo de organismos acuáticos con técnicas encaminadas a hacer más eficiente su producción, ha experimentado a nivel mundial un crecimiento continuo (Figura 1), con una velocidad de progresión superior al resto de los sectores de producción de alimentos de origen animal.

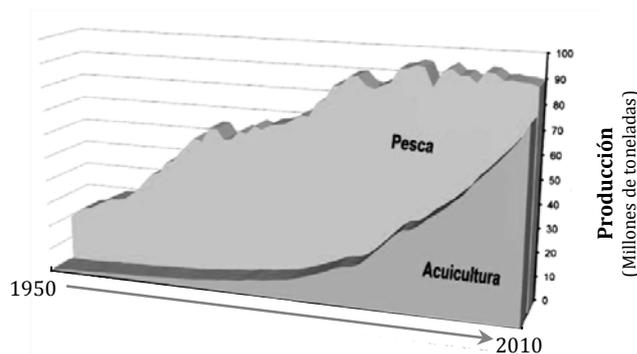


Figura 1. Evolución de la producción acuática (pesca y acuicultura) mundial durante el periodo comprendido entre los años 1950 y 2010 (gráfico modificado del informe APROMAR, 2013)

El motor que mueve esta rápida expansión se fundamenta en la limitación que presentan los recursos naturales como fuente de alimento, incapaces por si solos de hacer frente a la demanda mundial generada por una población humana en continuo crecimiento. Concretamente, los últimos datos estadísticos presentados por la FAO (Marzo de 2013) registraron para el año 2011 un incremento de la producción mundial acuícola animal del 6,2% respecto al año anterior, y prevén que dicha intensidad de crecimiento se mantenga estable durante el año 2012 (próximo informe de la FAO Marzo de 2014). No obstante, el dato que quizás mejor representa la importancia que ha adquirido el sector acuícola reside en que hoy en día más de la mitad de los productos acuáticos consumidos por la población mundial proceden de las granjas de cultivo (APROMAR, 2013). Por lo tanto, actualmente sería erróneo catalogar a la acuicultura como una actividad meramente complementaria a la pesca, ya que constituye un recurso alimenticio capaz de suministrar alimento de calidad en forma oportuna y estable.

A nivel mundial, el continente asiático se sitúa como la principal potencia acuícola animal, generando el 88% de la producción, seguido de forma muy distante de América y Europa, responsables respectivamente del 4,7% y 4,3% (FAO Fisheries and Aquaculture

Department, 2013). Tal y como se observa en la Tabla 1, dentro del conjunto de países productores que conforman el *Top-20* (responsables del 95% de la producción mundial acuícola animal) el 65% corresponde a países asiáticos, ocupando China un destacando puesto número uno; mientras que Europa solamente se encuentra representada por dos países, Noruega y España, cuyos valores de producción resultan entre 6 y 142 veces por debajo de la producción china, respectivamente (FAO Fisheries and Aquaculture Department, 2013).

Aunque la acuicultura animal es muy diversa e incluye tanto peces, moluscos, crustáceos, como otro tipo de invertebrados, su principal actividad se centra en el cultivo de peces, la cual actualmente representa dos tercios de la producción acuícola mundial animal y engloba 346 especies (FAO Fisheries and Aquaculture Department, 2013).

Si bien a lo largo de la última década el cultivo de peces ha doblado su producción mundial, dentro de la Unión Europea su desarrollo se ha visto seriamente retraído resultando necesario realizar una evaluación y mejora de la política pesquera común, creando un plan estratégico de actuación enfocado a los aspectos social, económico y ambiental, con la finalidad de desarrollar una acuicultura sostenible que obtenga un mayor aprovechamiento de la idoneidad de los recursos naturales (grandes extensiones de costa con ambientes propicios para el cultivo) y del apoyo científico-tecnológico que poseen sus regiones (APROMAR 2013; UICN, 2010).

Dentro del contexto acuícola marino europeo, la dorada (*Sparus aurata*, L.; perteneciente a la familia de los espáridos) es considerada la especie más importante, tanto por su volumen de producción como por su valor económico, seguida por la lubina (*Dicentrarchus labrax*) y el rodaballo (*Scophthalmus maximus*) (APROMAR, 2013). Como se puede observar en la Figura 2, el cultivo de dorada se extiende principalmente entorno a la región mediterránea, posicionándose Grecia, Turquía y España como países productores de referencia para la especie, generando en su conjunto más del 70% de la producción total. En España, las comunidades autónomas donde se realiza cultivo de dorada son por orden de importancia: Comunidad Valenciana, Región de Murcia, Islas Canarias, Cataluña y Andalucía, cuyas condiciones ambientales resultan especialmente óptimas para la especie. Preferentemente, la

Principales países productores	Producción (toneladas)
1 China	38.621.269
2 India	4.573.465
3 Vietnam	2.845.600
4 Indonesia	2.718.421
5 Bangladesh	1.523.759
6 Noruega	1.138.797
7 Tailandia	1.008.049
8 Egipto	986.820
9 Chile	954.845
10 Myanmar	816.820
11 Filipinas	767.287
12 Brasil	629.309
13 Japón	556.761
14 República de Corea	507.052
15 EEUU	396.841
16 Taiwan	314.363
17 Ecuador	308.900
18 Malasia	287.076
19 España	271.961
20 Irán	247.262
Producción Total (top-20)	59.474.657
Países Restantes	3.225.644
Producción Mundial	62.700.300

Tabla 1. Top-20 de los países con producción acuícola animal (tabla tomada del informe global FAO 2013).

creciente respecto a los últimos años, sus niveles (54,9 millones de unidades) resultan todavía insuficientes para suplir las necesidades exigidas por la producción anual de dorada, obligando al sector nacional a importar de otros países aproximadamente el 30% del volumen de alevines requerido (APROMAR 2013).

1.2. La genética como herramienta para potenciar la eficiencia de los cultivos acuícolas

Los actuales niveles de producción de dorada de crianza permiten un abastecimiento continuo de la especie en los mercados a un precio de venta económico [$\approx 5,0$ €/kg según los datos ofrecidos por el mercado de abastecimientos de Barcelona SA (www.mercarbana.es) para el periodo comprendido entre el 1/01/2013 y el 31/12/2013, en territorio nacional], permitiendo al consumidor acceder con facilidad a un producto alimenticio de calidad. Sin embargo, desde el punto de vista industrial la percepción de la situación es muy diferente, ya que los elevados niveles de producción desencadenan una gran competencia entre las empresas del sector que deriva en una importante caída de los precios del mercado, ocasionando a las empresas un margen de beneficios limitado (el coste medio de producción se estima entre 3,80 y 4,30 euros/kg; APROMAR 2012). Por consiguiente, para que el cultivo de dorada se convierta en una actividad económicamente sostenible, las empresas precisan de mejoras en la eficiencia de sus cultivos que les permitan realizar una correcta gestión de la biología de la especie y les conduzca a la obtención de una producción más rentable.

Ante este reto, la genética se presenta como una herramienta de gran potencial, pues se ha estimado que la utilización de lotes reproductores mejorados genéticamente para caracteres de interés productivo, pueden contribuir a una reducción del 50% de los costes de producción (Gjedrem y col., 2012).

1.2.1. Creación y manejo de lotes reproductores

El éxito de cualquier programa reproductivo dirigido hacia una mejora genética de los individuos se fundamenta en la selección reproductores “óptimos”, caracterizados por poseer niveles de variación genética suficientemente elevados para producir una respuesta estable en el tiempo a la selección aplicada (Falconer y Mackay, 2001; Borrell y col., 2007). De esta manera, la selección de los individuos maduros que fundarán el lote reproductor constituye un proceso indispensable para poder obtener con éxito individuos mejorados genéticamente (Moav y Wohlfart, 1976; Hulata y col., 1986).

Dentro de los diferentes tipos de marcadores moleculares disponibles a nivel de ADN, los *loci* microsatélites o repeticiones en tándem de fragmentos genómicos de pequeña longitud

(entre 1 y 6pb), gracias a su amplia distribución genómica (aunque se pueden localizar en cualquier punto del genoma, poseen una mayor predisposición a localizarse dentro de regiones no codificantes), abundancia, carácter multi-alélico, herencia mendeliana, codominancia, tasa de mutación elevada y alta conservación entre especies cercanas (Chistiakov y col., 2006), se han convertido en la herramienta utilizada por excelencia para determinar la diversidad genética existente entre individuos, y garantizar que el programa reproductivo planteado fija sus bases desde el punto de vista genético sobre una población *ideal* (Brown y col., 2005a; Borrell y col. 2007; Vogiatzi y col., 2011; Loukovitisy col., 2012). Reflejo de la gran utilidad de estos marcadores, el número de *loci* microsatélites identificados en dorada asciende actualmente a más de 1.000 secuencias (Vogiatzi y col., 2011; Calduch-Giner y col., 2013).

Por otro lado, cualquier programa reproductivo efectivo además de requerir partir de valores de diversidad genética elevados, precisa de un manejo adecuado del cultivo que maximice la conservación de la variación genética inicial a través de las generaciones. Para conseguir este segundo objetivo, se debe evitar en la medida de lo posible la endogamia (F) o cruzamiento entre individuos emparentados, ya que a nivel genético este fenómeno actúa aumentando la similitud entre individuos (aumenta la homocigosis) a la vez que disminuye la diversidad genética (disminuye la heterocigosis) inicial, restringiendo así la capacidad de respuesta de las nuevas generaciones a la selección aplicada, y por tanto, a su mejora. Además, la presencia de altos niveles de endogamia en un cultivo ($F \approx 0,25$) arrastra con frecuencia a la depresión endogámica, la cual disminuye la eficacia biológica de los individuos comprometiendo su fisiología (aparición de deformaciones, reducción en la tasa de crecimiento y supervivencia) y/o su capacidad reproductiva (reducción de la fecundidad, producción de huevos de menor tamaño y disminuciones en tasa de eclosión larvaria) (Dunham, 2004; Fjalestad, 2005).

Debido a la correlación inversa que se establece entre el incremento de endogamia (ΔF) y el número de reproductores que desde el punto de vista genético contribuyen al establecimiento de la nueva generación (*tamaño efectivo* o N_e) [$\Delta F = 1/2N_e$ (Falconer y Mackay, 2001)], la presencia de endogamia dentro de un programa de selección se plantea como un hecho inevitable, pues es el propio proceso selectivo quién, a través de sus cuellos de botella, propicia su aparición. Sin embargo, aunque la aparición de endogamia no se pueda evitar, sus efectos se pueden limitar mediante la adquisición de un conocimiento y control exhaustivo de la reproducción de la especie.

El manejo que indiscutiblemente posibilita el mayor control sobre la reproducción se basa en el establecimiento de cruzamientos dirigidos macho-hembra, donde el acuicultor diseña esquemas de cruces individuales específicos con la finalidad de generar familias

concretas, evitando así el cruzamiento no deseado entre individuos emparentados. Este sistema requiere que cada una de las líneas descendientes obtenidas se mantenga aislada del resto en un compartimento independiente hasta que sea posible proceder a un marcaje físico que no afecte al crecimiento y a la mortalidad de los individuos; por ejemplo, en dorada se establece que utilizando el sistema de marcaje PIT (*Passive Implant Transponder*) los alevines pueden ser marcados óptimamente una vez alcanzado un peso corporal de al menos tres gramos (Navarro y col., 2006).

Si a la capacidad de realizar cruzamientos específicos, se le añade la elevada fecundidad que presentan algunas especies como la dorada cuyas hembras son capaces de liberar entre 20.000-80.000 huevos al día durante un periodo de tiempo de aproximadamente cuatro meses (Colloca y Cerasi, 2005), las familias aislada tenderían a estar representadas por un elevado número de descendientes, lo cual permitiría estimar con precisión los parámetros genéticos (valor reproductivo de los individuos, heredabilidad y establecimiento de correlaciones genéticas entre caracteres) necesarios para direccionar los procesos de selección de forma mucho más efectiva.

Pese a la aparente idoneidad de este manejo reproductivo, y a pesar de que se ha logrado obtener cruzamientos individuales viables en dorada (Knibb y col., 1997), la baja eficiencia de su producción unida a los elevados costes que subyacen especialmente tras las exigencias de infraestructura derivadas del mantenimiento independiente de los individuos, hacen que su instauración bajo condiciones industriales resulte impracticable para esta especie acuícola. Por ello, las granjas de dorada se ven obligadas a utilizar un modelo reproductivo basado en la creación de lotes de cultivos cuyo tamaño censal suele oscilar entre los 40-60 individuos (guardando generalmente en una relación de sexos 2:1 a favor de los machos) donde se simula una reproducción “natural” que proporciona a las empresas niveles de producción adecuados, a un coste mucho más reducido. Este modelo, aunque permite una reducción en los costes de producción, limita al acuicultor la capacidad de control sobre el proceso reproductivo, ya que en ningún momento se dirigen los cruzamientos que se establecen entre los reproductores (dentro de cada tanque las hembras maduras desovan libremente y de forma conjunta mediante puestas masales intermitentes, al tiempo que los machos compiten entre sí por fecundar los huevos producidos), ignorando con ello el grado de contribución de cada reproductor a la puesta, así como las relaciones de parentesco existentes entre los descendientes. Por lo tanto, ante esta situación, resulta imposible estimar de forma directa el ΔF producido en cada generación.

No obstante, el vacío de información familiar ocasionado por este sistema de manejo, se puede suplir gracias a una combinación eficaz entre el marcaje físico y genético de los

individuos. Concretamente, se asume que la información derivada de un conjunto de 10-20 *loci* microsatélites polimórficos, a pesar de representar una porción mínima de la variación existente entre individuos, es suficiente para asignar con alta probabilidad y un éxito superior al 95%, una única pareja parental a cada descendiente (Wenne y col., 2007). En el caso concreto de dorada, se dispone de una amplia documentación que certifica el elevado poder resolutivo que posee este tipo de marcadores para reconstruir las relaciones de parentesco (Batargias y col., 1999; Launey y col. 2003; Brown y col., 2005b; Castro y col., 2007; Borrell y col., 2007; Loukovitisy col., 2012). No obstante, pese a la valiosa información proporcionada por los *loci* microsatélites, su aplicación a nivel industrial se ve frenada por el coste económico que ocasiona el genotipado independiente de cada uno de los individuos, resultando imprescindible generar una herramienta de genotipado de elevada eficacia y manejo sencillo, que proporcione la información necesaria en una forma fácilmente inteligible al menor coste posible. De acuerdo a este requerimiento, diversos estudios han centrado su objetivo en el desarrollo de sistemas de genotipado simultáneo que combinan un número variable de *loci* microsatélites (Launey y col., 2003; Brown y col., 2005b; Navarro y col., 2008; Porta y col., 2010; Lee-Montero y col. 2013). Entre las diferentes propuestas, se destaca por su gran aproximación al concepto de herramienta de genotipado óptima que demanda la industria, el sistema propuesto recientemente por Lee-Montero y col. (2013), basado en la combinación de 22 *loci* microsatélites (localizados exclusivamente en regiones no-codificantes) distribuidos equitativamente en dos reacciones de PCR múltiplex (SMSa1 y SMSa2), propuesto ante la *Sociedad Internacional de Genética Animal* (ISAG) para su aceptación como panel de marcadores microsatélites de uso internacional para dorada, y cuyo objetivo principal se destina a facilitar la comparación de las bases de datos generadas entre las diferentes empresas y centros de investigación.

Por otro lado, aunque su uso todavía no está muy extendido, los SNPs o polimorfismos de un solo nucleótido (Single Nucleotide Polymorphism) presentan un futuro prometedor dentro de la acuicultura moderna. Este tipo de polimorfismo hacen referencia a mutaciones puntuales del genoma que aparecen alternativamente ocupadas por al menos dos nucleótidos distintos con una frecuencia en la población igual o superior al 1% (Vignal y col., 2002). No obstante, aunque inicialmente cualquier posición del ADN sería susceptible de ser ocupada por cualquiera de los cuatro nucleótidos posibles, los SNPs muestran preferentemente un carácter bi-alélico. Además de las sustituciones nucleotídicas, a nivel práctico dentro del concepto de SNPs comúnmente se integran también pequeñas inserciones y deleciones. Los SNPs son marcadores moleculares de última generación que comparten con los *loci* microsatélites las características de herencia mendeliana, codominancia y amplia cobertura del genoma (Vignal

y col., 2002). Sin embargo, los SNPs presentan una densidad genómica mucho más elevada, atribuyéndoles por ejemplo hasta el 90% del polimorfismo existente en el genoma humano (Collins y col., 1998). Otra ventaja de los SNPs es su fácil combinación para generar sistemas con alta capacidad de genotipado simultáneo, existiendo plataformas de genotipado de capacidad media como ILPEX™ Gold MassArray® (Sequenom) que permite analizar simultáneamente hasta un máximo de 36 SNPs por individuo, o bien, plataformas de alta capacidad como la ofrecida por la tecnología Axiom™ Genotyping Taqman OpenArray (Applied Biosystem) que posee una capacidad de análisis entre 1.500 y 2,6 millones de SNPs (www.cegen.org). Además, a estos marcadores se les asocian otras dos ventajas importantes, la primera de ellas un bajo porcentaje de errores de genotipado, y la segunda, que sus ensayos son fácilmente transferibles entre laboratorios (Anderson y Garza, 2006). Al igual que todos los marcadores, los SNPs también tienen su desventaja, y reside principalmente en el limitado poder resolutivo que de forma independiente ofrece cada SNP. Para compensar este déficit informativo de SNP por *locus*, y poder así generar una herramienta de genotipado de SNPs con una eficacia resolutiva equiparable a la proporcionada por *loci* microsatélites, se propone trabajar siempre con un número de SNPs superior al número de *loci* microsatélites necesarios, pero ¿cuántos SNPs son realmente suficientes? (Smouse, 2010). Dando respuesta a esta cuestión, Santure y col. (2010) y Fernández y col. (2013) estimaron que en general para la adquisición de un poder de exclusión similar al obtenido con un *locus* microsatélite se precisa de media la combinación de 2-3 SNPs. Sin embargo a nivel práctico, a la hora de inferir las relaciones de parentesco las características de los individuos analizados pueden demandar el empleo de un número de SNPs superior a la media estimada por los anteriores autores (Hauser y col., 2011; Trong y col., 2013).

Gracias al desarrollo tecnológico, la actual disponibilidad de técnicas de genotipado masivo capaces de analizar un elevado número de SNPs a bajo coste ($\approx 0,02-0,30\text{€}/\text{SNP}$, depende de la tecnología utilizada y del número de SNPs analizados: a mayor número de SNPs analizados, menor coste por SNP), hace que la necesidad de generar paneles de genotipado amplios que incluyan un extenso número de variantes, se convierta en una desventaja de impacto mínimo, situando con ello a los SNPs, junto a los *loci* microsatélites, en una posición preferencial para promover el uso de marcadores genéticos en la industria acuícola.

En el caso de la dorada, a diferencia de los *loci* microsatélites, la información referente a los SNPs hasta el momento es bastante limitada, encontrándose disponibles solamente dos paneles SNPs, el primero de ellos compuesto por 76 variantes localizadas en 36 secuencias genómicas independientes no anotadas (Cenadelli y col., 2007), y el segundo por 59 variantes anotadas a partir de ESTs (secuencias expresadas, *Expresión Sequence Tag*) (Parati y col.,

2012), a los que se suman un pequeño número de SNPs descritos de forma independiente en diferentes genes de función conocida (Nadjar-Boger y Funkenstein, 2011). No obstante, atendiendo simplemente a la gran cantidad de recursos genómicos depositados actualmente en la base de datos del NCBI [1.583 secuencias nucleotídicas; 79.216 ESTs y 92.468 GSS (secuencias reconocidas del genoma, *Genome Survey Sequences*), con fecha 25.11.2013] se augura para dorada un rápido e inminente desarrollo de este tipo de marcadores.

Gracias a la información derivada de los marcadores genéticos, la reconstrucción de las relaciones familiares dentro del conjunto de descendientes procedentes generalmente de un único día/evento de puesta ha permitido establecer que bajo condiciones de cautividad los cultivos cuya reproducción transcurre a través de un sistema de puestas masales conjuntas, presentan una propensión a la reducción en número de reproductores que contribuyen a la siguiente generación respecto al tamaño censal del lote (Brown y col., 2005b; Fessehay y col., 2006; Piñera, 2009; Porta y col., 2010). De acuerdo a las estimas clásicas de tamaño efectivo para una población:

$$N_e = 4(N_h + N_m) / (N_h * N_m) \text{ (Falconer y Mackay, 2001)}$$

donde N_h y N_m representa respectivamente al número de hembras y machos que se reproducen y dejan descendientes, el cultivo de estas especies predispone a valores reducidos de N_e ($N_e < N$), quedando expuesto a la aparición de una tasa de incremento de endogamia superior a la estimada inicialmente por el acuicultor. Curiosamente, mediante esta estima clásica de N_e , también se deduce que, en contra del pensamiento general, la existencia de una relación de sexos macho:hembra asimétrica repercute negativamente sobre N_e , mientras que el mantenimiento de una proporción lo más cercana posible al equilibrio (1:1) tiende a maximizar su valor. En este sentido, cabe destacar que la dorada, al igual que otros espáridos, presenta un hermafroditismo proterándrico (Zohar y col., 1978) el cual predispone de forma natural al desequilibrio entre sexos favoreciendo que a lo largo de tiempo los lotes reproductores tiendan a mostrar un déficit de machos y una acumulación de hembras. Concretamente, este tipo de hermafroditismo se caracteriza porque todos los individuos inicialmente nacen bajo la condición sexual masculina pero alcanzado un momento determinado de su vida, se asume que debido a factores sociales y hormonales (Zohar y col., 1984), la inmensa mayoría sufre un proceso de reversión sexual basado en un desarrollo del tejido ovárico y una reabsorción/re-estructuración de las espermatogonias (Sadovy de Mitcheson y Liu, 2008) que les transforma definitivamente en hembras maduras. Normalmente, aunque es difícil de estimar con precisión, se acepta que en los machos la maduración sexual ocurre una vez alcanzados 20-30cm de longitud corporal

(aproximadamente a los 2 años de edad) y en las hembras a los 33-40cm (a los 2-3 años de edad) (Colloca y Cerasi, 2005). De esta manera, la capacidad de los individuos para actuar en un primer periodo reproductivo como machos activos, y posteriormente comportarse como hembras maduras, puede comprometer en mayor medida el tamaño efectivo de los lotes respecto a especies no hermafroditas. Por lo tanto, la condición biológica de hermafrodita proterándrico, supone una complicación añadida al manejo del cultivo, obligando a los acuicultores a renovar anualmente entre el 5 y 20% de sus individuos reproductores (Sola y col., 2006).

Por otro lado, además de la tendencia a la reducción en número de reproductores que toman parte en la puesta, en los estudios anteriormente citados, el establecimiento de pedigrís desveló que los reproductores participantes tienden a su vez a mostrar un éxito reproductivo diferente, contribuyendo a la nueva generación con un número de descendientes muy dispar. A diferencia de las estimas clásicas de N_e , formulaciones más complejas como las propuestas por (1) Chevassus (1989), o (2) Woolliams J. (comunicación personal descrita en Borrell y col., 2008), toman en consideración a este nuevo e importante factor, ofreciendo un cálculo de N_e , o bien directamente de ΔF , más preciso:

$$(1) N_e = 4(N-2) / ((K_m + V_m / K_m) + (K_h + V_h / K_h) - 2)$$

donde N corresponde al número de descendientes analizados, K representa el número de descendientes esperados bajo una contribución equitativa para cada macho (K_m) y hembra (K_h), y V marca la varianza de tamaño familiar dentro de los machos (V_m) y hembras (V_h).

$$(2) \Delta F = \sum C_i(m)^2 / 8 + 1 / (32m) + \sum C_i(h)^2 / 8 + 1 / (32h)$$

donde $C_i(m)$ y $C_i(h)$ representan respectivamente la contribución fraccional de descendientes de los machos y hembras, mientras m y h hacen referencia al número de machos y hembras participantes, respectivamente.

De este modo, los diversos factores que pueden condicionar seriamente el valor de N_e reflejan como los cultivos desarrollados a partir de sistema de puestas masales conjuntas, se exponen continuamente a incrementos de endogamia generacionales intensos. Por lo tanto, se lleve o no a cabo un programa de mejora genética, para salvaguardar a lo largo de las generaciones la efectividad y la calidad de la producción obtenida, el acuicultor debería en todo momento disponer de un control preciso del tamaño efectivo de sus lotes, y así, poder introducir medidas correctoras oportunas que confluyan en un riesgo de endogamia lo más restringido posible.

1.2.2. Mejora genética en caracteres de interés

Dentro de los posibles caracteres de mejora, a nivel industrial resultan especialmente interesantes todos aquellos caracteres cuya optimización conduce a una producción económicamente más rentable. Entre estos caracteres se destaca (López-Fanjul y Toro, 1990):

- *La tasa de crecimiento e índice de conversión de alimento en peso.* La mayoría de los programas de selección enmarcan a estos dos caracteres como su objetivo principal debido al gran impacto que originan sobre los costes de producción ya que, respectivamente, representan el tiempo que tardan los individuos en adquirir la talla comercial (\uparrow tiempo= \uparrow coste económico) y la cantidad de alimento necesaria para alcanzarla (\uparrow alimento= \uparrow coste económico).
- *La resistencia a enfermedades.* Los cultivos con frecuencia se ven expuestos a invasiones patogénicas que derivan en tasas de mortalidad elevadas provocando grandes pérdidas económicas. Por lo tanto, la adquisición de cultivos que posean una mayor resistencia al ataque de sus principales patógenos conferirá una seguridad especial a la producción.
- *La presencia/ausencia de deformidades.* Las deformidades causan anomalías que afectan principalmente a la forma del cuerpo, esqueleto, pigmentación o escamas, afecciones que a pesar de no resultar letales, provocan un efecto negativo sobre la producción correlacionándose con la calidad final del producto, ocasionando un impacto especialmente negativo en aquellas especies donde el producto comercializado se centra en el individuo entero y no en productos procesados (Afonso y Roo, 2007). Además de la repercusión sobre la calidad de la producción, la presencia de malformaciones puede también comprometer a la producción al repercutir negativamente sobre otros caracteres de interés como por ejemplo sobre el crecimiento, donde pueden provocar una ralentización significativa del mismo (Afonso y Roo, 2007).
- *Características asociadas al rendimiento y calidad de la carne.* Dentro de este punto se incluyen entre otras características: el peso del filete, la grasa visceral, cualidades nutritivas, la textura, sabor y color. A diferencia de los caracteres anteriores, en esta ocasión la mejora no se destina a la obtención de una reducción en los costes de producción, sino a la adquisición de un producto de calidad superior que al adquirir en el mercado un valor económico más elevado, reporta un margen de beneficios mayor.
- *La edad a la madurez sexual.* El interés que despierta este carácter surge a consecuencia del impacto que origina sobre otros caracteres productivos ya que una vez desarrollada la madurez sexual, el crecimiento tiende a ralentizarse, la mortalidad se incrementa y la

eficacia de conversión alimentaria y la calidad del filete tienden a disminuir, por lo que, desde el punto de vista productivo retrasar su aparición a la adquisición del tamaño comercial resultaría altamente ventajoso.

- *La calidad del huevo.* Debido a su relación con la tasa de fecundación (a mayor calidad, mayor porcentaje de huevos fecundados) y con la supervivencia durante las primeras fases del desarrollo, la mejora de este carácter pretende incrementar los niveles y la calidad de la producción de alevines.

Por lo tanto, la combinación de la mejora de todos estos caracteres se dirige en su conjunto a la obtención de un volumen de producción mayor, con una calidad superior, a un coste menor y un beneficio mayor.

Generalmente, el desarrollo de los programas de selección para estos caracteres es complejo, pues todos ellos representan caracteres cuantitativos cuya expresión se encuentra habitualmente bajo el control de numerosos genes (poligenes) cada uno de los cuales ejerce un efecto específico, y la suma del efecto de todos sus genes constituye la base genética responsable del fenotipo del individuo. A su vez, estos caracteres se encuentran sometidos a una importante influencia ambiental cuyo efecto es el desencadenante de que un mismo fenotipo esconda diferentes genotipos, conllevando con ello a un desajuste en las asociaciones genotipo-fenotipo.

A consecuencia del control genético múltiple, de la fuerte influencia ambiental y del efecto derivado de la interacción entre ambos factores, los caracteres cuantitativos suelen presentar una variación fenotípica elevada entre individuos, distribuyéndose sus valores fenotípicos de forma más o menos continua a través de un amplio rango de dispersión. Por ello, ante este tipo de caracteres el fenotipo de un individuo se define en base a un valor numérico, mientras que para un determinado grupo de individuos portadores de un mismo genotipo su fenotipo se designa de acuerdo al valor medio y a la desviación típica de sus fenotipos (Tave, 1996). De este modo, la selección aplicada en un programa de mejora dirigido a estos caracteres tiene como objetivo desplazar a lo largo de las generaciones el valor fenotípico medio de la población en la dirección deseada, maximizando así la expresión del fenotipo de interés al tiempo que se origina una población cada vez más homogénea para el carácter seleccionado. Dependiendo de la intensidad y precisión de la selección aplicada, el valor fenotípico medio de cada generación se distancia en mayor o menor medida de la media fenotípica establecida en las generaciones anteriores, y a este diferencial de medias fenotípicas comúnmente se le denomina ganancia genética generacional (ΔG) (Stem, 1964), y se puede calcular de acuerdo a la siguiente fórmula (Meuwissen, 2003):

$$\Delta G = \frac{\text{Intensidad de selección} * \text{Precisión de selección} * \text{Desviación genética estándar}}{\text{Intervalo generacional}}$$

Pese a la amplia variación fenotípica que poseen estos caracteres, los programas de mejora se valen exclusivamente de la proporción de la varianza debida al componente genético, pues solamente aquellos individuos cuya superioridad para el carácter resida en su constitución genética podrán transmitir dicha superioridad a la siguiente generación. Por ello, cualquier proceso de selección efectivo debe confluir en la identificación de individuos que presenten un genotipo superior al resto con independencia del ambiente que les rodee. De este modo, todos aquellos caracteres cuyo componente genético contribuya notablemente al fenotipo final, al poseer una elevada concordancia genotipo-fenotipo, mostrarán una capacidad de mejora superior. Para estimar el grado de contribución que posee el componente genético sobre el carácter se utiliza la heredabilidad (h^2), cuyo valor representa la proporción de la variación fenotípica total del carácter debida exclusivamente al efecto aditivo de los genes.

En dorada, Boulton y col. (2011) determinaron para diversos caracteres morfométricos valores de h^2 moderados-altos, al igual que Navarro y col. (2012) estimaron valores de h^2 moderados-bajos para características de crecimiento, rendimiento, composición de la carne y presencia/ausencia de deformidades, reforzando en diversas ocasiones los valores estimados para algunos de estos caracteres en estudios previos (referencias en Navarro y col., 2012). Respecto a la resistencia a enfermedades, Antonello y col. (2009) dedujeron que existe una cierta determinación genética en la resistencia a pasteurelosis o pseudotuberculosis, enfermedad producida por la bacteria gram negativa *Photobacterium damsela* subsp. *piscicida*, cuyo ataque afecta fuertemente al cultivo de dorada, llegando a provocar en los lotes de cultivo infectados una mortalidad del 90-100%. Por lo tanto, todos estos estudios reafirman que utilizando como recurso la varianza genética aditiva los caracteres de producción de esta especie pueden ser optimizados notoriamente a través de la instauración de programas de selección adecuados a las características de cada carácter.

De forma simple, los programas clásicos de selección fundamentan su criterio de selección sobre el valor de fenotípico de los individuos, asumiendo que tras un fenotipo de interés subyace una base genética (genotipo) favorable. Así, en los casos de selección masal o individual, a partir de la media fenotípica del lote de cultivo se fija un valor límite que marcará el umbral de selección por encima del cual se seleccionaran todos aquellos individuos que muestren un valor fenotípico superior, sin atender al grado de consanguinidad existente entre los individuos seleccionados (Tave, 1996). De esta manera, cuanto más estricto sea el umbral de selección impuesto, más intenso será el cuello de botella generado y en consecuencia, el

riesgo de endogamia será mayor. Por lo tanto la selección individual representa un sistema de mejora genética de fácil implantación y coste reducido, expuesto a rápidos incrementos de endogamia.

Dado que la selección individual se rige exclusivamente en base al fenotipo, su efectividad se ve fuertemente reducida en aquellos caracteres que poseen una determinación genética baja ($h^2 < 0,15$) (Tave, 1996). Ante esta situación, dentro de los métodos de clásicos, los esquemas selección familiar ofrecen una repuesta más eficaz a la selección, ya que en este tipo de esquemas el valor fenotípico de los individuos adquieren significado únicamente en base al valor fenotípico medio de su propia familia, lo cual minimiza la variación fenotípica debida al componente ambiental, y favorece la selección de individuos con un componente genético de interés (Tave, 1996). Los procesos de selección familiar se pueden dividir a su vez en interfamiliares (entre las familias) e intrafamiliares (dentro de las familias). En los esquemas interfamiliares el criterio de selección establecido se centra en la elección de aquellas familias que registran el valor fenotípico medio más elevado, sin la necesidad de seleccionar familias enteras, aislando a partir de cada familia seleccionada una muestra aleatoria de individuos. Por otro lado, en los esquemas intrafamiliares la selección se realiza dentro de cada familia de forma independiente, eligiendo como futuros reproductores a los mejores individuos de cada familia. En ambos sistemas, para asegurar que todas las familias se encuentran representadas por igual en el nuevo lote reproductor, la proporción de individuos seleccionados (número y relación sexual) debe ser idéntica entre ellas. Las selecciones inter- e intrafamiliar se puede combinar dentro de un mismo programa de mejora, y curiosamente su actuación combinada reporta una respuesta superior a la obtenida a partir de cada uno de estos sistemas de forma independiente (Gjedrem, 1983). Los sistemas familiares además de resultar especialmente ventajosos en caracteres de baja h^2 , también resultan de gran utilidad para caracteres cuya medición exige el sacrificio del animal como es el caso del rendimiento canal o la calidad de la carne, ya que de cada familia solamente se sacrifica una proporción individuos, de manera que si para algunas familias se determina una variación heredable de interés, su acervo genético se encontrará preservado en el conjunto de individuos no sacrificados de dicha familia (Tave, 1996). A pesar de que los sistemas de selección familiar suponen un esfuerzo económico mayor que los sistemas individuales, su efectividad en la mejora de un amplio rango de caracteres es la responsable de que actualmente se encuentren en activo más 100 programas en especies muy diversa, adquiriendo especial intensidad dentro el grupo de las tilapias (género *Oreochromis*) y los salmónidos (familia *Salmonidae*) (Gjedrem y col., 2012).

Actualmente, la información genómica disponible, junto a el elevado número de marcadores moleculares identificados, permiten optimizar los esquemas clásicos de selección

y desarrollar nuevos métodos. Así ya, Doyle y Herbiner (1994) proponen, para restringir el riesgo de endogamia, un sistema de selección denominado “*walk-back selection*”, donde la información de parentesco se combina con la selección individual. De acuerdo a este sistema, los individuos se clasifican en base a su valor fenotípico y se genotipan (vía loci microsatélites o SNPs) aquellos individuos que registran los valores más elevados. Seguidamente se selecciona el individuo situado en la primera posición de la clasificación y se compara con el individuo que ocupa la segunda posición de la clasificación fenotípica, el cual será seleccionado siempre y cuando no se corresponda con un hermano o medio hermano del individuo anteriormente seleccionado. Siguiendo este mismo criterio, el procedimiento se repite N veces hasta alcanzar el tamaño censal necesario. Si bien es cierto que el genotipado de los individuos conlleva un coste añadido al sistema, en esta ocasión su repercusión es limitada, ya que no se requiere el genotipado de todos los individuos sino solamente de aquellos que previamente fueron seleccionados por mostrar los valores fenotípicos más elevados.

Los nuevos métodos de selección parten de una misma premisa: *selección de marcadores moleculares cuyos alelos se asocian a una expresión fenotípica diferencial*. Mediante estos métodos los reproductores candidatos se seleccionan en base a su composición genotípica, ofreciendo, respecto a los métodos clásicos, una precisión de selección mayor que permite acelerar el proceso de selección especialmente en aquellos caracteres de baja h^2 . Además, estos sistemas eliminan también la necesidad de sacrificar a los individuos para determinar su fenotipo en caracteres como el rendimiento y calidad de la carne, y permiten anticipar el proceso de selección en caracteres cuya expresión tiene lugar de forma tardía en el desarrollo, como la madurez sexual. Dentro estos métodos se incluye la selección asistida por marcadores moleculares (MAS, *Marker Assisted Selection*), la selección asistida por genes (GAS, *Gene Assisted Selection*) y selección genómica (GS, *Genomic Selection*).

Concretamente en el sistema MAS el proceso de selección actúa a nivel de marcadores moleculares indirectos cuya asociación con el carácter de interés surge a consecuencia de su proximidad física en el cromosoma a alguna de las regiones genómicas implicadas en control genético del carácter denominadas QTLs (*Quantitative Trait Locus*). De este modo, la implementación de sistemas de selección MAS requiere previamente de un complejo proceso de identificación de regiones QTLs asociadas al carácter de interés (Broman y Speed, 1999), así como de la caracterización de marcadores moleculares que muestren un ligamiento o asociación física con dichas regiones, dependiendo la precisión de la selección aplicada del grado de conexión existente entre ambos, ya que cuanto más intenso sea el ligamiento entre marcador y QTL, menor será la probabilidad de que entren en juego fenómenos de recombinación que interfieran en la conexión genotipo-fenotipo previamente fijada. No

obstante, dado que dichos procesos de recombinación son inevitables, Sonesson y Meuwissen (2009) recomiendan re-estimar cada 2-3 generaciones la asociación entre las variantes alélicas del marcador y el fenotipo de interés. Aproximadamente, se asume que para garantizar un progreso selectivo elevado la distancia entre marcador y QTL no debería exceder de los 5 centimorgan (cM), no obstante una separación de 15-50cM sería suficiente para obtener una respuesta a la selección media-alta (López-Závala y col., 2007). En dorada, la identificación de QTLs se ha visto favorecida por el desarrollo de un mapa de ligamiento genético basado en 204 *loci* microsatélites (Franch y col., 2006) y un mapa físico por radiación híbrida (RH) compuesto por aproximadamente 1.000 marcadores que incluyen tanto *loci* microsatélites, secuencias génicas y fragmentos de ADN genómico conocidos o STS (*Sequence Tagged-Site*) (Senger y col., 2006; Sarropoulou y col., 2007). Actualmente para esta especie ya se han descrito QTLs asociados a caracteres morfométricos (Boulton y col., 2011; Loukovitis y col., 2013), crecimiento (Massault y col., 2011; Loukovitis y col., 2011; 2013), determinación sexual (Loukovitis y col., 2011) y resistencia a enfermedades (Massault y col., 2011).

No obstante, la precisión de selección del sistema MAS además de estar condicionada por la intensidad de ligamiento existente entre marcador y QTL, depende del número de QTLs analizados, pues un solo carácter puede estar controlado por 100-200 QTLs ejerciendo cada uno de ellos un efecto determinado sobre el fenotipo, y en algunos casos este efecto puede llegar a ser muy limitado. De media, se asume que para poder justificar el 50% de la variación genética del carácter se precisan entre 10 y 40 QTLs (Bao, 2010). En vista de estos hechos, y a consecuencia de que los sistemas MAS frecuentemente se valen de un pequeño número de marcadores moleculares que representan a un número limitado de QTLs, la ganancia genética obtenida procederá de la explotación de una determinada fracción de la variación genética total que controla el carácter, la cual será función del número de QTLs analizados, del efecto que dichos QTLs provoquen sobre el fenotipo, así como de la intensidad del vínculo de unión entre QTL y marcador.

A diferencia del sistema de selección MAS, el sistema GAS utiliza como base haplotipos o genotipos concretos de variantes localizadas en un gen/es con influencia sobre el carácter, evitando con ello que procesos de recombinación puedan ocasionar una fragmentación de la asociación genotipo-fenotipo y desvirtuar la selección aplicada. Por lo tanto, el sistema GAS permite ejercer una presión de selección superior al sistema MAS, ya que asienta sus bases sobre marcadores funcionales directos cuyos alelos poseen una implicación directa sobre el fenotipo; sin embargo, al igual que el anterior sistema su implementación depende del grado de conocimiento existente acerca del control genético del carácter en la especie, y de igual modo, se enfrenta también a la restricción de poder mejorar únicamente una proporción de la

variación genética que controla el carácter, la cual en este caso dependerá del número de genes incluidos en programa de selección, así como de su grado de implicación en el fenotipo.

Finalmente, el sistema GS ofrece dos ventajas importantes frente a los sistemas MAS y GAS: (1) no precisar información previa sobre las regiones cromosómicas (QTLs o genes) que controlan la expresión del carácter, y (2) como su nombre indica, su análisis se extiende por todo el genoma permitiendo capturar toda la variación genética asociada al carácter (Meuwiss y col., 2003). De forma simple, la selección genómica se define como una combinación de los sistemas MAS y GAS reproducidos a gran escala, y su concepto se proyecta en la utilización de un elevado número de marcadores moleculares (paneles de alta densidad) distribuidos de forma más o menos equitativa a lo largo de todo el genoma, asegurando con ello que todos los QTLs se encuentre bajo una situación de desequilibrio de ligamiento con al menos uno de marcadores analizados, o bien, se correspondan directamente con una variación localizada dentro de la propia secuencia de algún gen control (Meuwiss y col., 2003). El requerimiento de un panel de marcadores de alta densidad genómica convierte a los SNPs en una herramienta clave para el desarrollo de este tipo de programas de selección, y hace que una vez más sea necesario cuestionarse *¿cuántos marcadores son necesarios?* Por medio de simulaciones Meuwissen y col. (2009) establecen que el número de marcadores requeridos para llevar a cabo un sistema GS se correlaciona simultáneamente con el tamaño efectivo de la población y con la longitud del genoma de la especie (medido en Morgans), de manera que el número de marcadores necesarios asciende en la medida en que estos dos factores incrementan su valor. Por ejemplo, en ganado, concretamente en la raza Holstein (comúnmente conocida como la frisona) cuyo genoma alcanza los 30 Morgans de longitud, se estimó que para implementar un sistema GS adecuado, manejando un $N_e=100$, se precisa al menos 30.000 marcadores (Hayes y Goddard, 2010).

Por otro lado, bajo este último sistema de selección el valor reproductivo de cada individuo se determina a partir de la suma de los efectos de todos los marcadores analizados, y dada su elevada densidad, para poder determinar con precisión las relaciones genotipo-fenotipo que servirán de base para fijar los criterios de selección dentro del programa de mejora, se requiere de poblaciones de referencia muy numerosas (más 1.000 individuos que incluyen un elevado número de familias) (Hayes y Goddard, 2010). Como ejemplo del tamaño poblacional de referencia necesario, y recurriendo nuevamente a la raza de ganado frisona, VanRaden y col. (2009) estimaron con alta precisión el valor reproductivo de los individuos a partir del genotipado de 38.416 SNPs en 3.576 individuos.

Si bien los programas de selección genómica cuentan ya con una amplia difusión en el sector ganadero (Pena, 2012), dentro del contexto acuícola sus datos se limitan hasta el

momento a estudios de simulación (Nielsen y col., 2009; Sonesson y Meuwissen, 2009; Villanueva y col., 2010) cuyos resultados atribuyen a este novedoso sistema de selección un gran potencial para la mejora del cultivo, ejerciendo una selección que supera hasta en un 33% la precisión determinada mediante métodos de selección clásicos (Nielsen y col., 2009). A su vez, el sistema de selección GS al posibilitar el cálculo de las relaciones de parentesco a partir de la diversidad del genoma de los individuos, permite monitorizar con elevada precisión los niveles de endogamia generados, deduciéndose respecto a los métodos clásicos, reducciones en la tasa de endogamia del 81% (Sonesson y Meuwissen, 2009).

El desfase existente entre sector ganadero y acuícola surge principalmente a consecuencia de que en los organismos marinos las investigaciones en el campo de la genética han despertado interés de forma más tardía que en los animales terrestres. Sin embargo, durante las últimas décadas la acuicultura ha experimentado una gran revolución genómica permitiendo actualmente disponer de la información del genoma completo de especies modelo como el pez cebra (*Danio reiro*) y medaka (*Oryzias latipes*), así como de diversas especies entre las que destacan por su importancia acuícola la tilapia del Nilo (*Oreochromis Niloticus*) (ID GenBank: AERX00000000.1; pendiente de publicación) y el bacalao (*Gadus morhua*) (ID GenBank: CAEA00000000.1; Star y col., 2011), encontrándose además en una fase de desarrollo avanzada el proyecto de secuenciación del genoma completo de salmón Atlántico (*Salmo salar*) (ID GenBank: AGKD01000000; Davidson y col., 2010), carpa común (*Cyprinus carpio*) (ID GenBank: 202478), y del pez gato (*Ictalurus punctatus*) (Jiang y col., 2013). En dorada, Kuhl y col. (2011) en una primera aproximación han logrado secuenciar aproximadamente 75% de su genoma (60,6Mbp), generando un gran avance en la información disponible para la especie. Además de esta secuenciación parcial del genoma, otras investigaciones han optado por la caracterización del transcriptoma de dorada (García de la Serrana y col., 2012; Yúfera y col. 2012; Calduch-Giner y col., 2013), entendiendo como tal a la parte del genoma que se expresa en una célula concreta, en una etapa específica del desarrollo del individuo. Si bien el genoma es un elemento común a todas la células de los distintos tejidos celulares que conforman un organismo, el transcriptoma puede diferir notablemente de unos tejidos a otros e incluso dentro de un mismo tejido entre las diferentes fases del desarrollo, ya que depende de la combinación de ARNs mensajeros (ARNm) procedentes de los genes que se estén expresando en un tejido y en un momento concreto. Mientras García de la Serrana y col. (2012) y Yúfera y col. (2012) se centraron respectivamente en la secuenciación del transcriptoma específico del tejido muscular (musculo esquelético rápido) y del estado larvario, anotando un gran número de genes y factores de transcripción claves para el desarrollo muscular y larvario, diferenciación, morfología y crecimiento, Calduch-Giner y col.

(2013) optaron por combinar la información procedente del músculo esquelético, intestino, riñón y sangre para secuenciar y ensamblar el transcriptoma, lo que les permitió obtener un transcriptoma más completo (probablemente representa más del 75% de los genes activamente transcritos), con un mayor número de genes anotados dentro de los cuales se encuentran genes asociados al crecimiento, digestión de nutrientes y respuesta inmune. Por lo tanto en dorada, pese a la ausencia de una secuencia completa de su genoma, se dispone de una fuente importante de recursos genómicos con capacidad suficiente para propulsar el desarrollo de sistemas de selección basados en métodos que trabajan directamente sobre variaciones genéticas responsables o asociadas a variaciones en la expresión del carácter objeto de selección.

Curiosamente a pesar de la amplia documentación existentes sobre la capacidad de mejora genética que poseen las especies acuícolas, la introducción de cualquier tipo de programa de selección dentro del sector está siendo un proceso lento, y en la actualidad se asume que tan solo entorno al 10% de la producción acuícola proviene de cultivos mejorados genéticamente (Gjedrem y col., 2012), incluso a pesar de que su rentabilidad se ha puesto de manifiesto en especies tan relevantes para la acuicultura como el salmón Atlántico o la tilapia del Nilo, donde se ha determinado que los beneficios derivados de la implementación de sistemas de selección multiplican entre 8 y 60 veces la inversión económica inicial (Ponzoni y col., 2007; Gjedrem y col., 2012). Sin embargo, el constante crecimiento sobre el conocimiento de las características genéticas de las especies, fusionado con el rápido desarrollo tecnológico que pretende ofrecer sistemas de análisis cada vez más precisos y eficaces a coste reducido, predispone al sector acuícola a un inminente futuro donde las empresas a través de la organización de su producción en base a un criterio genético consigan optimizar sus cultivos y potenciar su rentabilidad económica.

2. Objetivos

En la actualidad, desde el punto de vista genético, la producción y el rendimiento del cultivo de dorada muestran una amplia capacidad de optimización a través de estrategias de mejora que comprendan tanto acciones pasivas, destinadas a la conservación de la variabilidad genética y al control de la endogamia, como acciones activas donde el conocimiento de las bases genéticas que regulan los caracteres productivos posibilitan el desarrollo de sistemas de selección que conllevan a la mejora de dichos caracteres. Por esta razón, la presente tesis doctoral centra su interés en el desarrollo y aplicación de marcadores genómicos encaminados a la mejora genética (pasiva y activa) de la especie, planteándose los siguientes objetivos:

1. Analizar la dinámica reproductiva a lo largo del periodo de puesta, para evaluar con precisión el incremento de endogamia derivado del manejo del cultivo, y proponer medidas correctoras de manejo que conduzcan a una maximización del tamaño efectivo de los lotes.
2. Identificar entre los reproductores de un lote de cultivo, aquellos individuos que, al contribuir de forma preferencial hacia descendientes con fenotipo *crecimiento-rápido*, muestren un valor reproductivo superior.
3. Identificar y caracterizar genes candidatos asociados a los caracteres calidad de huevo y resistencia a enfermedades.
4. Generar un panel de SNPs localizados en genes candidatos para crecimiento.
5. Evaluar la posible asociación entre marcadores SNPs identificados en genes candidatos para crecimiento y el peso corporal de los individuos.

3. Capítulo I

Optimización del manejo reproductivo

3.1. Introducción

A pesar del gran avance que ha supuesto la introducción de marcadores moleculares en la acuicultura, controlar el incremento de endogamia generacional resulta aún una tarea compleja para aquellas especies cuya reproducción en cultivo transcurre mediante un sistema de puestas masales, puesto que, bajo condiciones de cautividad, parece ser su propia dinámica reproductiva quién conduce a importantes incrementos de endogamia (ΔF) (Fessehaye y col., 2006; Borrell y col., 2008; Blonk y col., 2009; Piñera, 2009).

En el caso del cultivo de dorada, la dinámica reproductiva descrita por diversos autores (Brown y col., 2005b; Fessehaye y col., 2006; Herlin y col., 2008; Chavanne y col., 2014), se ha caracterizado por presentar una reducción significativa en el N_e del lote, a consecuencia de la ausencia de participación de algunos reproductores y a la existencia de una contribución parental diferencial, generándose una situación que predispone a corto plazo a un fuerte aumento de la endogamia no deseada. Sin embargo, los estudios desarrollados por estos autores poseen una importante limitación temporal, ya que en todos ellos la dinámica reproductiva se ha inferido exclusivamente a partir de un único evento/día de puesta, provocando que cuestiones de interés relevante como *¿permanece constante el número de reproductores activos a lo largo del periodo de puesta?*, así como, *¿existen fluctuaciones en las relaciones de dominancia establecidas entre los reproductores?* queden sin resolver. La incapacidad de dar respuesta a estas cuestiones, unida al extenso periodo de puesta que presenta esta especie (aproximadamente 3 meses), lleva a cuestionar si realmente el análisis de la progenie procedente de un único evento de puesta recolectado en un momento concreto del periodo reproductivo, es realmente representativo de la dinámica reproductiva desarrollada por un lote reproductor.

En base a este planteamiento, y dado el importante papel que juega la variación genética en la implementación de programas de selección efectivos dirigidos a mejorar la producción industrial de la especie (Gjedrem, 2000), desde el punto de vista genético, el desarrollo de estrategias de manejo que minimicen el incremento de endogamia y retarden la pérdida de variación genética producida en cultivo requiere de una profundización en el comportamiento reproductivo de la especie.

En este primer capítulo, por medio de la evaluación de descendencias recolectadas en diferentes momentos del periodo de puesta en dos lotes reproductores experimentales, se pretende establecer de forma precisa las consecuencias genéticas derivadas de la dinámica reproductiva adoptada por la especie bajo condiciones de cultivo. Para dar alcance a este

objetivo, se desarrollará un protocolo de análisis basado en el genotipado de un conjunto de seis *loci* microsatélites y adecuado a descendientes en fases embrionarias del desarrollo, que permita esclarecer de forma precoz la dinámica reproductiva del lote.

3.2. Materiales y Métodos

3.2.1. Muestras biológicas

En las instalaciones del IRTA (*Institut de Recerca i Tecnologia Agroalimentaries*) de Sant Carles de la Ràpita (Tarragona), se crearon dos lotes reproductores experimentales de dorada (lotes A y B) con un tamaño censal de ocho individuos (4♀ y 4♂).

Mediante la técnica de identificación por chip (*PIT, Passive Integrated Transponder*) se realizó el marcaje de los reproductores; proceso que además se aprovechó para la recolección de un pequeño fragmento de aleta dorsal de cada individuo, el cual se almacenó en etanol al 96% y se envió a la Universidad de Oviedo donde se preservó a 4°C hasta su utilización.

El mantenimiento de los lotes se realizó en tanques de cultivo circular de 4.000L de capacidad con recirculación, bajo condiciones controladas de temperatura (19±1°C), luz (fotoperiodo de 16h luz-8h oscuridad), salinidad (35±1ppt) y una dieta alimenticia basada en pienso comercial diseñado específicamente para individuos reproductores (*MAR-Vitalis Rebro. Skretting*).

Durante el periodo de puesta (24/12/2007 al 5/03/2008), mediante sistemas colectores dispuestos en los tanques, se procedió a la recolección de los huevos producidos y a su separación en las categorías de huevos flotantes (viables) y no-flotantes, desechándose esta última categoría. Una fracción de los huevos flotantes recolectados en cada puesta, se almacenó en tubos de 1,5mL con etanol al 70%, y se envió nuevamente a la Universidad de Oviedo, donde se preservó a 4°C hasta su utilización.

En total, se analizaron 670 descendientes, 410 procedentes de cinco días de puesta (S1A, S2A, S2A, S4A y S5A) distribuidos a lo largo del periodo reproductivo del lote A, y 260 correspondientes a tres puestas (S1B, S2B y S3B) del lote B (Tabla 3a, Figura 3a).

3.2.2. Extracción ADN

La extracción de ADN genómico, se realizó siguiendo el protocolo estándar de resina Chelex®100 (Biorad) (500µL al 10%) y proteinasa K (10µL a 20mg/mL) (Walsh y col., 1991). Para los individuos reproductores, se empleó un pequeño fragmento obtenido a partir de aleta dorsal, sin embargo, los individuos descendientes se analizaron en un estado de desarrollo muy temprano (huevos en estadio de desarrollo inferior a las 51:00h a 18,5°C, Figura 3b), requiriéndose para la extracción del ADN la utilización del huevo completo. En el caso de las

descendencias, debido a que la cantidad de ADN que se puede obtener a partir de este tipo de muestra biológica es más limitada, para obtener una concentración de ADN óptima para amplificación por PCR, el volumen de resina Chelex®100 empleado se disminuyó a 100µL. Las soluciones de ADN obtenidas se conservaron a 4°C hasta su utilización.

3.2.3. Creación de la PCR mutliplex

A partir de un conjunto de nueve *loci* microsatélites formado por *SaGT1*, *SaGT26*, *SaGT41b* (Batargias y col., 1999), *Dxd44*, *Dpt3* (De la Herrán y col., 2006), *Pb-OVI-A3*, *Pb-OVI-B2*, *Pb-OVI-D22* y *Pb-OVI-D106* (Piñera y col., 2006), se procedió al diseño de una PCR multiplex utilizando el programa Multiplex Manager 1.0 (Holleley y Geerts, 2009). Este programa diseña PCRs multiplex optimizadas utilizando como información basal el rango de distribución alélica de cada *locus*, el número de alelos por *locus*, la heterocigosidad, la temperatura de hibridación (T_a) de los iniciadores, junto a las características derivadas de sus secuencias.

El rango de distribución alélica de los *loci* microsatélites previamente seleccionados, se determinó mediante la amplificación individual de cada *loci* en cada uno de los reproductores. Las reacciones de amplificación se realizaron en un termociclador XP-Cycler (Bioer), empleando para un volumen final de 10µL las siguientes condiciones: 1µL ADN genómico, 0,5U Taq polimerasa (Ecogen), 0,2µM para cada iniciador, 200µM de cada dinucleótido (dNTP), MgCl₂ a 1,5mM y tampón de reacción al 1X. El programa de PCR utilizado se basó en: una fase de desnaturalización a 94°C durante 2 minutos, seguida por 40 ciclos de amplificación [1 minuto a 94°C, 30 segundos a 60°C (excepto para los loci *SaGT41b*, *Pb-OVI-B2* y *Pb-OVI-D22* que será a 55°C) y 30 segundos a 72°C], finalizando con una fase de extensión a 72°C durante 5 minutos. Todas las reacciones se comprobaron mediante electroforesis en un gel de agarosa al 2% con bromuro de etidio. Los productos de PCR obtenidos se caracterizaron mediante fluorescencia en un secuenciador automático ABI 3130XL Genetic Analyzer. Finalmente, el establecimiento de genotipos se realizó utilizando el programa GenMapper® 4.0 (Applied Biosystems).

A partir de los diversos diseños generados por el programa Multiplex Manager 1.0, se seleccionó el diseño de PCR multiplex compuesto por los *loci* microsatélites *SaGT1-VIC*, *SaGT26-NED*, *SaGT41b-NED*, *Dxd44-6-FAM*, *Dpt3-6-FAM*, *Pb-OVI-D106-VIC*. Esta PCR multiplex, se utilizó para llevar a cabo el genotipado de las descendencias, empleando para ello el kit comercial QIAGEN Multiplex PCR, siguiendo, para un volumen final de 15µL, las siguientes condiciones de reacción: 1X para el tampón de reacción (incluye tampón de reacción, MgCl₂ y HotStartTaq ADN polimerasa), 1X para la mezcla de iniciadores (0,5µL para el locus *Pb-OVI-*

D106 y 0,3 μ M para el resto de *loci*) y 1 μ L de ADN genómico. Las reacciones de PCR se llevaron a cabo en un termociclador XP-Cycler (Bioer) utilizando el siguiente programa de amplificación: desnaturalización a 95 $^{\circ}$ C durante 15 minutos, seguida de 30 ciclos de amplificación (30 segundos a 94 $^{\circ}$ C, 1 minuto y 30 segundos a 54 $^{\circ}$ C y 1 minuto a 72 $^{\circ}$ C) y una extensión final de 10 minutos a 60 $^{\circ}$ C. Todas las reacciones se comprobaron en un gel de agarosa al 2% con bromuro de etidio, y posteriormente se analizaron mediante fluorescencia de acuerdo al protocolo mencionado anteriormente.

Para excluir cualquier error de genotipado en los individuos parentales, cada lote reproductor se utilizó como control de sus respectiva descendencia, de modo que los genotipos parentales se fijaron tras un proceso de varias repeticiones (cinco repeticiones en el lote A y tres en el lote B), aceptando con ello que en el caso de existir errores de genotipado, estos quedarían limitados exclusivamente a los genotipos descendientes.

3.2.4. Evaluación del estadio de desarrollo embrionario óptimo para amplificación por PCR

Con ayuda de una lupa óptica se evaluó el estadio de desarrollo de los huevos fertilizados (Figura 3b) estableciéndose cuatro categorías: 2-4 células, 16 células, gástrula y embrión visible. A partir del análisis de los *loci* microsatélites *Dxd44* y *Dpt3* (incluidos en la PCR multiplex) en 100 huevos (25 por categoría), mediante el conteo del número de amplificaciones positivas obtenidas en cada categoría, se determinó el estadio de desarrollo óptimo para garantizar un éxito elevado en la amplificación por PCR. Las amplificaciones se realizaron en un termociclador XP-Cycler (Bioer), empleando para un volumen final de 10 μ L las siguientes condiciones de PCR: 1 μ L de ADN genómico, 0,5U de Taq polimerasa (Ecogen), 0,2 μ M para cada iniciador, 200 μ M de cada dinucleótido (dNTPs), 1,5mM de MgCl₂ y 1X de tampón de reacción. El programa de amplificación empleado se basó en una fase inicial de desnaturalización a 94 $^{\circ}$ C durante 2 minutos, seguida por 40 ciclos de amplificación (1 minuto a 90 $^{\circ}$ C, 30 segundos a 60 $^{\circ}$ C y 30 segundos a 72 $^{\circ}$ C), finalizando con una fase de extensión a 72 $^{\circ}$ C durante 5 minutos. La visualización de los productos de PCR, se realizó en un gel de agarosa al 2% con bromuro de etidio.

3.2.5. Análisis de variabilidad y diversidad genética

La variabilidad genética presente en cada lote, se estimó atendiendo a los parámetros: número de alelos por *locus* (N_a), frecuencias alélicas, heterocigosidad observada (H_o) y heterocigosidad esperada (H_E), calculados mediante el programa CERVUS v.3.0 (Marshall y col.,

1998; Kalinowski y col., 2007).

La diferenciación genética entre los lotes, así como entre reproductores y descendencias dentro de cada lote, se evaluó mediante el índice de fijación de Wright (1951) (F_{IT}) a través de sus dos componentes (F_{IS} y F_{ST}), obtenidos según el método descrito por Weir y Cockerham (1984):

$$(1-F_{IT}) = (1-F_{IS}) (1-F_{ST})$$

- F_{IT} : pérdida de heterocigosidad de un individuo tomado del conjunto de subpoblaciones a causa del apareamiento no aleatorio y a la deriva originada por un tamaño subpoblacional pequeño.
- F_{IS} : mide la pérdida de heterocigosidad de un individuo tomado del conjunto de subpoblaciones, a causa del apareamiento no aleatorio (utilizado para evaluar el ajuste al modelo de equilibrio de Hardy-Weinberg).
- F_{ST} : mide la pérdida de heterocigosidad originada por la subdivisión en poblaciones y la deriva.

Para ello, se empleó el programa estadístico Fstat v2.93 (Goudet, 1995; 2001) basado en métodos de distribución aleatorios. Los niveles de significación de F_{IS} se determinaron a través de la generación aleatoria de alelos (2.000 permutaciones) dentro de las muestras que componen cada lote. De igual modo, los valores de significación de F_{ST} se basaron en 2.000 permutaciones de genotipos multi-*locus* entre pares de muestras, aplicando la corrección conservativa de Bonferroni (Rice, 1989; Goudet, 2001).

3.2.6. Asignación parental

La asignación parental se realizó mediante el programa CERVUS v.3.0 (Marshall y col., 1998; Kalinowski y col., 2007), infiriéndose la pareja parental más probable para cada descendiente mediante análisis de máxima probabilidad, empleando como información basal los genotipos parentales y descendientes obtenidos. Además de diversos parámetros de diversidad genética, este programa permite, a partir de la información genotípica de los reproductores, calcular el contenido de información polimórfica (PIC) y el poder de no-exclusión (probabilidad de no excluir a un parental candidato no emparentado, seleccionado al azar, para uno o varios *loci* de un descendiente) total y de forma individualizada para cada *locus*.

A partir de las frecuencias alélicas observadas en el lote reproductor, mediante el programa CERVUS v.3.0, se realizaron simulaciones teóricas de los procesos de asignación de

parentesco enfrentando genotipos generados como parentales a genotipos generados como descendientes. La probabilidad de asignación de una pareja parental es función de la complementariedad alélica del trío descendiente-parental(1)-parental(2), siendo mayor cuanto mayor sea la probabilidad de que los parentales cedan el alelo correcto al descendiente, primando el genotipo homocigoto sobre el heterocigoto. El programa repite el proceso tantos ciclos como descendientes previamente se le hayan indicado, obteniéndose finalmente una distribución de probabilidades de acierto que delimita el valor de significación de futuras asignaciones con descendencias reales. Estos procesos de simulación permiten evaluar la capacidad de los marcadores seleccionados como herramientas de identificación parental en función del polimorfismo detectado en el conjunto de reproductores.

Además de las simulaciones teóricas incluidas en el CERVUS v.3.0, como medida adicional para evaluar la capacidad de asignar correctamente la pareja parental, se procedió a la generación de descendencias simuladas mediante el programa PROBMAXG (Danzmann, 1997). La creación de estas descendencias virtuales permite trabajar con un conjunto de descendientes cuyo pedigrí resulta ser un dato conocido. El enfrentamiento del genotipo de cada uno de los 8.000 individuos descendientes generados en la simulación (500 individuos por cada uno de los 16 cruzamientos posibles) con el conjunto de genotipos que constituyen el lote reproductor, permite contabilizar el porcentaje de parejas parentales sugeridas por CERVUS v.3.0 que se corresponden con la pareja parental real de cada descendiente.

3.2.7. Estimación del tamaño efectivo e incremento de endogamia

La determinación del tamaño efectivo reproductor (N_e) se llevó a cabo bajo dos supuestos diferentes. En una primera aproximación, el N_e se estimó atendiendo únicamente al número de hembras (N_h) y machos (N_m) que dejan descendientes viables, mediante la fórmula clásica:

$$N_{e(1)} = 4(N_h \times N_m) / (N_h + N_m) \text{ (Falconer y Mackay, 2001)}$$

En una segunda aproximación, se calculó el N_e bajo el supuesto de una contribución parental diferencial mediante la siguiente ecuación matemática (Chevassus, 1989):

$$N_{e(2)} = 4(N-2) / ((Kd + Vd/Kd) + (Ks + Vs/Ks) - 2)$$

donde N corresponden al número de descendientes analizados, K representa el número de descendientes esperados bajo una contribución equitativa por cada hembra (Kd) y macho (Ks), y V representa la varianza en del tamaño familiar en el grupo de las hembras (Vd) y machos (Vs).

En ambos casos, el incremento de endogamia generacional (ΔF) se estimó por medio de su relación inversa con el N_e :

$$\Delta F = 1/(2 N_e) \text{ (Falconer y Mackay, 2001)}$$

3.2.8. Análisis complementarios

La existencia de una contribución diferencial (a nivel global y dentro de cada puesta analizada) de los reproductores al establecimiento de la nueva generación, se evaluó mediante la prueba estadística Chi-cuadrado (χ^2), utilizando el programa SPSS v.15.0. Esta misma prueba estadística se empleó también para evaluar la hipótesis nula de independencia entre la fecha de puesta y la contribución parental.

3.3. Resultados

3.3.1. Características del periodo de puesta

Ambos lotes reproductores mostraron un periodo de puesta similar, contabilizándose un total de 68 días de puesta para el lote A y 65 días para el lote B (Figura 3a). El lote A, registró la mayor producción de huevos flotantes, presentando volúmenes de puesta diarios generalmente superiores al lote B durante todo el periodo reproductivo (Figura 3a; Tabla 3b). Respecto al lote B, cabe destacar una intensa reducción de la producción diaria de huevos durante los últimos 26 días de puesta (Figura 3a), donde se recolectó un volumen medio de tan solo 50,30 ($\pm 37,61$) huevos flotantes (Tabla 3b).

En referencia al patrón de puestas registrado, el periodo reproductivo de cada lote se dividió en tres fases: inicial (lote A: 24/12/07-11/01/08; lote B: 24/12/07-04/01/08), intermedia (lote A: 12/01/08-17/02/08; lote B: 05/01/08-01/02/08) y final (lote A: 18/02/08-05/03/08 lote B: 02/02/08-05/03/08)(Figura 3a).

3.3.2. Estadio de desarrollo embrionario óptimo para amplificación por PCR

El éxito de la amplificación de los *loci* microsatélites *Dpt3* y *Dxd44* presentó diferencias significativas entre los diferentes estadios de desarrollo embrionario analizados (Tabla 3c). El menor número de amplificaciones positivas se registró en los estadios de desarrollo más tempranos (embriones formados por 2-4 y 16 células), mientras el mayor éxito de amplificación se alcanzó a partir de embriones en un estadio igual o superior a las 18:00-20:00H desarrollo (perfil del embrión visible, Figura 3b) con un porcentaje de amplificaciones positivas superior al 90% (Tabla 3c).

En base a estos resultados, para el análisis de descendencias se seleccionaron solamente aquellos huevos que presentaban un estadio de desarrollo embrionario mínimo de 18:00-20:00H. Este criterio de selección, imposibilitó el análisis de las muestras recolectadas durante la fase inicial del periodo de puesta, pues la mayoría o la totalidad de los huevos recibidos mostraban un estadio inferior a las 18:00H de desarrollo requeridas para el análisis.

3.3.3. Puesta a punto de la PCR multiplex OVIDOPLEX-R

Las simulaciones realizadas modificando el número de *loci* microsatélites analizados, mostraron que la PCR multiplex posee una capacidad resolutive de asignación parental elevada e idéntica para los supuestos realizados utilizando entre 9 y 6 *loci* microsatélites (Tabla 3d).

Cuando las simulaciones se realizaron a partir del genotipado de 4 y 5 *loci* microsatélites, la capacidad de asignación parental aún conservaba un porcentaje superior al 90%, sin embargo el empleo de menos de 4 *loci* disminuía su éxito por debajo del 75% (Tabla 3d). En base a estos porcentajes, se estimó que el desarrollo de una PCR multiplex compuesta por solamente 6 *loci* microsatélites, sería suficiente para obtener (aun cuando existieran fallos de amplificación en uno o dos *loci* microsatélites) una capacidad de asignación elevada.

Inicialmente, se seleccionaron los 6 *loci* microsatélites que en ambos lotes reproductores presentaban los valores más elevados para los parámetros N_a , PIC y poder de exclusión: *SaGT1*, *SaGT26*, *SaGT41*, *Dxd44*, *Dpt3* y *Pb-OVI-D106* (Tabla 3e). Estos *loci* mostraron una elevada capacidad de resolución a nivel de individuo en el proceso de asignación parental, logrando la asignación correcta del 100% de los descendientes simulados con pedigrí conocido.

Finalmente el programa Multiplex Manager determinó que los 6 *loci* seleccionados se podían combinar en una única reacción de PCR, proporcionando 5 diseños de PCR multiplex posibles, de los cuales se seleccionó el diseño OVIDORPLEX-R (Figura 3c). Sin embargo, pese a que todos los *loci* microsatélites presentaron en los dos grupos parentales una señal de amplificación buena, durante el genotipado de las descendencias del lote A, se detectó para *locus SaGT1* un porcentaje de amplificación inferior al 40% (Tabla 3f). En base a las estimas realizadas a partir de las simulaciones previas para 5 *loci* microsatélites, se optó por la exclusión del *locus SaGT1* de OVIDORPLEX-R, y consecuentemente, de los análisis de variabilidad genética y asignación de paternidades.

En total, mediante OVIDORPLEX-R (5 *loci* microsatélites) se obtuvo, para aproximadamente el 92% de la descendencia analizada (598/646 descendientes), un perfil genético compuesto por al menos 4 *loci* microsatélites, registrándose para ambos lotes un éxito de genotipado similar (lote A: 93,5%; lote B: 90,9%) (Tabla 3f).

3.3.4. Caracterización de la diversidad genética

Los parámetros de diversidad genética deducidos en el estudio mostraron valores de variabilidad elevados, ligeramente superiores en el lote reproductor A ($N_a=9,6$; $H_E=0,915$; $H_0=0,950$) respecto al B ($N_a=8,4$; $H_E=0,892$; $H_0=0,900$) (Tabla 3g). Los análisis de diferenciación genética determinaron la existencia de diferencias genéticas entre lotes ($F_{ST(A-B)}=0,0852$, P-valor=0,001), encontrándose ambos conjuntos de reproductores en equilibrio de Hardy-Weinberg (lote A: $F_{IS}=-0,053$, P-valor=0,236; lote B: $F_{IS}=-0,040$, P-valor=0,365) (Tabla 3g).

En ambos lotes, las descendencias mostraron una ligera reducción en los parámetros de

diversidad génica N_a , H_E y H_0 (reducción lote A: $N_a=8,3\%$, $H_E=6,9\%$ y $H_0=6,4\%$; reducción lote B: $N_a=0\%$, $H_E=13\%$ y $H_0=6,5\%$) frente a su correspondiente lote reproductor (Tabla 3g), no obstante, las diferencias detectadas no resultaron significativas ($P\text{-valor}>0,005$).

Tanto la descendencia procedente del lote A ($F_{IS}=-0,044$, $P\text{-valor}=0,005$), como la procedente del lote B ($F_{IS}=-0,083$, $P\text{-valor}=0,005$), mostraron un desequilibrio de Hardy-Weinberg significativo debido a un exceso de heterocigotos (Tabla 3g).

Por otra parte, dentro del lote A, la comparación del conjunto de reproductores con su descendencia total (lote A: $F_{ST(\text{parental-descendencia})}=-0,0237$, $P\text{-valor}=1,00$), así como entre el lote reproductor y la descendencia procedente de diferentes fechas de puesta (Tabla 3h), no presentaron diferencias genéticas significativas. Sin embargo, la comparación de las descendencias recolectadas en diferentes fechas de puesta, sí mostró diferencias significativas (excepto entre las descendencias S1A y S2A) (Tabla 3h). En el lote B, aunque la comparación entre el conjunto reproductor y la descendencia global no reveló constituciones genéticas diferentes ($F_{ST(\text{parental-descendencia})}=-0,004$, $P\text{-valor}=0,70$), la comparación entre reproductores y las descendencias recolectadas en distintas fechas, así como los conjuntos descendientes entre sí, sí mostraron una diferenciación genética significativa (Tabla 3h).

3.3.5. Contribución parental

La asignación parental presentó valores superiores a los esperados bajo las simulaciones previamente desarrolladas (92-94%) (Tabla 3d), identificándose una pareja parental con máxima probabilidad para el 100% de la descendencia genotipada ($N_{\text{total}}=598$: $N_A=376$ y $N_B=222$).

A partir del análisis de parentesco y establecimiento de familias, se deduce, a nivel global, una dinámica reproductiva similar para ambos lotes reproductores, donde todos los individuos contribuyen al establecimiento de la siguiente generación con diferencias significativas en el grado de contribución parental (lote A: $\chi^2_{\text{hembras}}=38,06$, $p\text{-valor}<0,001$ y $\chi^2_{\text{machos}}=47,10$, $p\text{-valor}<0,001$; lote B: $\chi^2_{\text{hembras}}=107,33$, $p\text{-valor}<0,001$ y $\chi^2_{\text{machos}}=210,32$, $p\text{-valor}<0,001$) (Tabla 3i). Dentro del lote A, las hembras H033 y H217 presentaron una contribución notablemente superior a las hembras H131 y H155 (Tabla 3i). En el grupo de los machos, 3 individuos (P102, P108 y P109) contribuyeron con una intensidad similar mientras el cuarto individuo (P111) presentó una contribución aproximadamente 2,5 veces inferior al valor esperado bajo el supuesto de una contribución equilibrada. Las diferencias de contribución parental alcanzaron especial intensidad en el lote B, donde la hembra H279 y el macho P112 (responsables de aproximadamente el 48% y 65% de la descendencia analizada, respectivamente) mostraron

un claro comportamiento reproductivo dominante (Tabla 3i).

Aunque la descendencia analizada en el lote A reflejó todos cruzamientos parentales posibles, se observó que aproximadamente el 50% de la descendencia analizada se agrupa en tan solo seis de las 16 familias posibles (H033-P102, H033-P108, H033-P109, H217-P109, H217-P108 y H217-P102) (Figura 3d). En el caso del lote B las diferencias de contribución a nivel del pareja parental resultaron más acusadas, pues a pesar de que todos los reproductores contribuyeron al establecimiento de la siguiente generación, solamente se detectó el 81% de los cruzamientos parentales posibles, asignándose el 71% de los descendientes analizados a tres únicas parejas parentales (H036-P112, H279-P112 y H279-P105) (Figura 3e).

El análisis de contribución parental a partir de la agrupación de las diferentes puestas analizadas en función de la fase del periodo de puesta en el cual se recolectaron, mostró una tendencia a la reducción en número de familias establecidas (un 25% y 50% para el lote A y B, respectivamente). Dentro de las fases, las diferencias de contribución parental entre los individuos reproductores continúan presentes (lote A: (1) fase intermedia: $\chi^2_{(\text{hembras})}=17,17$, p-valor=0,001 y $\chi^2_{(\text{machos})}=42,30$, p-valor<0,001; (2) fase final: $\chi^2_{(\text{hembras})}=55,46$, p-valor<0,001 y $\chi^2_{(\text{machos})}=72,82$, p-valor<0,001; lote B: (1) fase intermedia: $\chi^2_{(\text{hembras})}=35,18$, p-valor<0,001 y $\chi^2_{(\text{machos})}=28,16$, p-valor<0,001; (2) fase final: $\chi^2_{(\text{hembras})}=71,28$, p-valor<0,001 y $\chi^2_{(\text{machos})}=119,77$, p-valor<0,001), registrándose una mayor variación en los tamaños familiares generados durante la fase final respecto a la fase intermedia (Figuras 3d y 3e). La valoración de la contribución parental por fases del periodo de puesta, permitió a su vez detectar la restricción de determinados cruzamientos parentales a fases concretas del periodo de puesta (Figuras 3d y 3e), y a pesar de que un mismo cruzamiento puede tener lugar en diferentes fases, su contribución puede resultar muy diferente (Figuras 3d y 3e).

Cuando se valoró la contribución parental (tanto a nivel de individuo como de familias) de forma independiente para cada fecha de puesta analizada, se observó una dinámica reproductiva *diaria*, donde no todos los reproductores participaron en los diferentes días de puesta analizados, y los comportamientos reproductivos dominantes adquirieron generalmente mayor intensidad, pese a que en el caso del lote A durante S2A y S5A, las diferencias de contribución parental dentro del grupo de las hembras, no resultaron significativas (S2A: $\chi^2=5,95$, p-valor=0,11 y S5A: $\chi^2=4,07$, p-valor=0,13). Los comportamientos dominantes resultaron de nuevo especialmente acusados en el lote B, donde una sola hembra dio lugar a más del 90% de la descendencia generada durante S2B y S3B (H279 y H036, respectivamente), y un solo macho (P112) contribuyó con más del 90% de la progenie de S3B (Tabla 3i).

La hipótesis nula de independencia entre la fecha de puesta y la contribución parental se rechazó para ambos sexos en los dos lotes de cultivo (lote A: (hembras) $\chi^2=93,58$, p-valor $<0,001$ y (machos) $\chi^2=104,00$, p-valor $<0,001$; lote B: (hembras) $\chi^2=238,30$, p-valor $<0,001$ y (machos) $\chi^2=43,19$, p-valor $<0,001$).

A pesar de que en diversos momentos del periodo reproductivo todos los reproductores participaron en el establecimiento de la siguiente generación, no se producen todos los cruzamientos posibles, variando el número de parejas reproductoras que se establecen en cada día de puesta. En el caso del lote A, el número de parejas establecidas varió entre el 56-81%, mientras que el lote B osciló entre el 31-75%. Tal y como se observa en las Figuras 3d y 3e, a nivel de pareja reproductora también existe una contribución desigual, identificándose cruzamientos parentales dominantes. Los individuos que conformaron las parejas dominantes fueron variando a lo largo del periodo reproductivo (Tabla 3i; Figuras 3d y 3e).

3.3.6. Tamaño efectivo

Atendiendo únicamente al número de hembras y machos reproductores participantes, el $Ne_{(1)}$ obtenido para cada lote, mostró un valor idéntico al tamaño censal ($N=8$) durante la fase intermedia y el global del periodo de puesta (Tabla 3j). No obstante, aunque atendiendo a la fase final del periodo de puesta el $Ne_{(1)}$ puede sufrir reducciones como ocurre en lote B (Tabla 3j), es al considerar cada día puesta de forma independiente, donde el $Ne_{(1)}$ experimenta las mayores diferencias respecto al tamaño censal (Tabla 3j), debido a una disminución en el número de reproductores activos, llegando a registrarse reducciones hasta del 60% en el lote B.

Cuando además del número de reproductores participantes, se tiene en cuenta las variaciones de contribución entre los parentales, las dominancias detectadas en los diferentes momentos del periodo reproductivo se traducen en una notable reducción del $Ne_{(2)}$ (Tabla 3j).

Estas reducciones del $Ne_{(2)}$ en referencia al tamaño censal del lote, conducen a importantes incrementos de endogamia, y debido a las extremas dominancias desarrolladas por los reproductores del lote B, se deducen valores en el incremento de endogamia generacional ($\Delta F_2 \approx 0,12$) que aproximadamente duplica el valor esperado bajo el supuesto de una contribución equitativa de los reproductores ($\Delta F_2/\Delta F_1=1,93$) (Tabla 3j).

3.4. Discusión

3.4.1. OVIDOPLEX-R: una herramienta eficaz de genotipado

El número de *loci* microsatélites necesario para desarrollar con éxito los procesos de caracterización genética y asignación parental, fluctúa en función del nivel del polimorfismo propio de cada *loci*, así como, de la variabilidad genética presente en el conjunto individuos objeto de análisis. De esta manera, para cada estudio se debe previamente optimizar el número mínimo de *loci* microsatélites necesarios para garantizar la obtención de la información requerida. En nuestro caso, los altos niveles de variabilidad presentes en los dos lotes reproductores (Tabla 3e), junto con la selección y combinación de un conjunto de cinco *loci* microsatélites altamente polimórficos en una única reacción de PCR (OVIDORPLEX-R), permitió deducir con éxito el genotipo de más del 90% de los individuos analizados, logrando asignar al 100% de individuos genotipados, una única pareja parental. Por lo tanto, OVIDORPLEX-R, supuso una herramienta de análisis altamente eficaz para el estudio debido a su alta capacidad de resolución en la identificación del origen de los alelos a nivel de individuo, fácil manejo, reducción en el tiempo de análisis requerido y a su rentabilidad (Tabla 3d).

La obtención de un porcentaje de asignación parental del 100% se debe principalmente a elevada variabilidad detectada en los lotes reproductores, y al conocimiento previo del sexo de los individuos, el cual simplifica el proceso de identificación parental (Wilson y Ferguson, 2002), ya que una vez asignado uno de los parentales, todas aquellas parejas constituidas por individuos del mismo sexo quedan excluidas, restringiéndose la búsqueda del segundo parental a un número de individuos candidatos mucho más reducido. No obstante, otros factores pueden entrar en juego y afectar a la capacidad de asignación parental, destacando por su frecuencia: la amplificación preferencia de alelos de pequeño tamaño (fenómeno conocido como efecto “*drop-out*”), el tipo y/o estado de conservación del material biológico empleado, y la aparición de alelos nulos (alelos que no pueden ser amplificados utilizando los iniciadores diseñados). A modo de ejemplo de cómo estos factores pueden influir sobre el éxito en la asignación parental, Piñera y col. (2006) establecieron a partir de un lote de cultivo de besugo (*Pagellus bogaraveo*), que el genotipado erróneo de individuos (parentales y descendientes) heterocigotos como homocigotos derivado de la presencia de alelos nulos, provoca una disminución de la capacidad de asignación parental correcta superior al 6%. Por este motivo, dado que la escasa y débil amplificación obtenida en nuestro estudio para el *locus SaGT1* podría predisponer fuertemente a este tipo de errores de genotipado, se optó por

eliminar dicho marcador de OVIDORPLEX-R, sin ello afectar a la capacidad de identificación de la pareja parental.

3.4.2. Evaluación de la dinámica reproductiva a partir de descendientes en estadio de desarrollo embrionario: ventajas y dificultades

Desde el punto de vista de las técnicas genéticas moleculares como la amplificación por PCR, el trabajo con muestras biológicas correspondientes a embriones formados por un número reducido de células, lleva implícito la dificultad técnica de trabajar con muestras donde la cantidad de ADN disponible puede ser bastante limitada, conllevando con ello a una reducción en el éxito de amplificación. De acuerdo al presente estudio, dicha dificultad técnica puede ser solventada mediante la selección de descendencias que muestren un tiempo de desarrollo embrionario igual o superior al establecido previamente como estadio de desarrollo mínimo necesario. Concretamente, en base a los porcentajes de amplificación positiva obtenidos para los diferentes tiempos de desarrollo evaluados, en dorada se fija como estadio de desarrollo mínimo necesario, el estadio mostrado por Kamaci y col. (2005) para las 18:00-20:00H de desarrollo (Figura 3b), a través del cual se logra un éxito de amplificación final similar al obtenido habitualmente en nuestro laboratorio a partir de tejido adulto (Tabla 3c y 3f). El porcentaje obtenido resulta significativamente superior al obtenido por Borrell y col. (2008) en dentón (*Dentex dentex*), donde solamente se consiguió amplificar el 65% de la progenie muestreada. Aunque Borrell y col. (2008) atribuyen el bajo éxito de amplificación obtenido a la posible presencia de huevos no fertilizados en las puestas, en base a los resultados derivados de nuestro estudio donde la utilización de embriones en estadio de gástrula (10:00-12:00H) originó una disminución aproximada del 30% en el porcentaje de amplificaciones positivas (Tabla 3c), el resultado obtenido por estos autores puede surgir también a consecuencia de la utilización de descendientes con un tiempo de desarrollo inferior al mínimo necesario. Por lo tanto, la selección de puestas bajo el criterio de un tiempo de desarrollo mínimo, serviría de garantía para lograr un alto éxito en el genotipado precoz de las progenies.

Con motivo de la estrategia de puesta asincrónica desarrollada por las hembras en dorada (Francescon y col, 1994), la eliminación de aquellos individuos con un tiempo de desarrollo inferior al establecido como umbral, podría ser interpretada como un sesgo, implicando la eliminación de familias procedentes de una o varias hembras cuya puesta pudo tener lugar pocas horas antes del proceso de recolección (<18:00-20:00H). No obstante, en nuestro estudio, el número de descendientes excluidos dentro de cada una de las puestas analizada, no superó el 5%. No obstante, sí se detectaron grandes diferencias en los tiempos de desarrollo de progenies recolectadas en diferentes días de puestas, y algunas muestras como

las correspondientes a la fase inicial del periodo de puesta, no pudieron ser evaluadas debido al elevado porcentaje de huevos no fertilizados y/o escaso desarrollo embrionario (estadios comprendidos entre de 2 células y mórula) que presentaban. Por lo tanto, *a priori*, en nuestro caso, la selección de descendientes aplicada condicionó la selección de las puestas empleadas para el estudio (descartándose todas aquellas puestas cuya progenie presentaba mayoritariamente un tiempo de desarrollo inferior a las 18:00-20:00H), pero dentro de las puestas analizadas no causó sesgo alguno sobre la dinámica reproductiva deducida. Por consiguiente, el protocolo establecido permite deducir con precisión de forma rápida, sencilla y a bajo coste, el comportamiento reproductivo adquirido por un lote de reproductores transcurrido un intervalo de tiempo muy pequeño desde el momento de la fertilización, posibilitando con ello la identificación y selección de puestas que bajo el criterio genético resulten óptimas para preservar los niveles de variación y controlar el incremento de endogamia entre generaciones. Sin embargo, se debe remarcar que debido a la escasa tasa de supervivencia larvaria en la especie (20-35%) (Basurco y col., 2011) y al impacto que este hecho puede provocar sobre la composición final de individuos que conforman el lote, la realización de un estudio de contribución parental durante las fases tempranas del desarrollo no excluye la necesidad de realizar, como paso previo a la selección de individuos que conformarán el nuevo lote reproductor, un análisis de paternidades durante la fase adulta.

3.4.3. Características de la dinámica reproductiva

El establecimiento de pedigríes y la agrupación de las progenies en familias, permitió reconstruir el comportamiento reproductivo adoptado por los reproductores que conformaban cada lote, describiendo una dinámica reproductiva diaria, de fase y global, a partir de progenies en estadio de desarrollo embrionario procedentes de puestas independientes que representan diferentes momentos del periodo de puesta.

Los dos lotes reproductores analizados, presentaron un patrón reproductivo similar, caracterizado por una participación de todos los individuos a nivel global del periodo reproductivo, con diferencias significativas (más acusadas en el lote B) en su grado de contribución. Dentro de cada fase del periodo de puesta, el número de reproductores activos presenta oscilaciones, registrándose una tendencia a la reducción en el número de individuos que toma parte en la puesta durante la fase final, similar a la descrita por Herlin y col. (2008) en bacalao (*Gadus morhua*). No obstante, las mayores variaciones se detectan al analizar de forma aislada los diferentes eventos de puesta, existiendo puestas diarias donde todos los individuos reproductores contribuyen a la progenie, mientras que en otras ocasiones, el

número de participantes se limita a tan solo el 62,5% (Tabla 3i). A diferencia de otros estudios donde se registra una fuerte reducción en el número de machos que contribuye a la progenie recolectada (Barbaro y col., 1997; Bekkevold y col., 2002; Brown y col., 2005b; Herlin y col. 2008), en esta ocasión, aunque el grupo de los machos tiende a mostrar generalmente varianzas de contribución superiores al grupo de las hembras, la ausencia de participación parental parece afectar a ambos sexos por igual.

Por otra parte, incluso en aquellas puestas donde se establece una participación reproductora máxima, los porcentajes de contribución pueden llegar a ser muy diferentes entre individuos, detectándose con frecuencia comportamientos dominantes para ambos sexos (Tabla 3i). Además, el grado de contribución que muestra un individuo reproductor sufre modificaciones a lo largo del periodo reproductivo, estableciéndose para ambos sexos, una asociación de dependencia significativa entre la fecha de puesta y el grado de contribución parental. Por tanto, la participación y grado de contribución que muestra un individuo reproductor se ven condicionados por el intervalo del periodo reproductivo seleccionado, pues individuos que en una determinada puesta/fase se muestran inactivos o con bajos niveles de contribución parental, pueden haber desarrollado comportamientos dominantes en etapas anteriores o desarrollarlos en eventos posteriores a la misma.

Cuando la evaluación de la dinámica reproductiva se realiza a nivel de pareja reproductora, las conductas dominantes vuelven a tomar presencia identificándose parejas con contribuciones dominantes durante uno o varios momentos del periodo de puestas. Esta dominancia aunque habitualmente surge a consecuencia del cruce entre la hembra y el macho dominante de la puesta, también se puede originar a partir de una estrategia de cruce preferencial donde individuos con contribuciones inferiores a los individuos dominantes, logran imponerse como pareja reproductora al establecer un único tipo de cruzamiento parental en la puesta, el cual produce un porcentaje de descendientes superior al resto de los cruzamientos múltiples [Tabla 3i (macho P111), Figura 3d (pareja reproductora P111-H155)] desarrollados por el resto de los reproductores.

No obstante, debido a la alternancia de los comportamientos dominantes entre las parejas reproductoras, la agrupación de puestas recolectadas en diferentes momentos del periodo de puesta, al implicar la combinación de dominancias ejercidas por diferentes parejas reproductivas, tiende a atenuar la intensidad de dicha dominancia, resultando un patrón de cruzamientos parentales con porcentajes de contribución más equilibrados (Figura 3d). No obstante, cuando las dominancias son muy intensas, como ocurre en el lote B, los efectos de la dominancia continúan estando presente en el global de periodo reproductivo, aunque, con menor magnitud (Figura 3e). De igual modo, a consecuencia de que el número y el tipo de

cruzamientos parentales que se establecen dentro de cada evento de puesta muestran también fluctuaciones, la combinación de diferentes puestas contribuye a su vez a incrementar el número de familias presentes en el lote descendiente (Figuras 3d y 3e).

Por lo tanto, a través del presente estudio se evidencia como, atendiendo al número de reproductores participantes, al número de cruzamientos parentales ocasionados y a su respectivo grado de contribución a la siguiente generación, cada día de puesta puede llegar a ser únicamente representativo de sí mismo; por lo que el seguimiento de los reproductores a lo largo de diferentes fechas de puesta permite deducir una dinámica reproductiva global, más precisa, donde el comportamiento reproductor de cada individuo no se evalúa a partir de un suceso puntual, sino a lo largo del tiempo.

3.4.4. Consecuencias genéticas de la dinámica reproductiva: propuesta de manejo

Asumiendo el global del periodo reproductivo, la contribución de todos los reproductores al establecimiento de la nueva generación, conduce a una relación óptima entre $Ne_{(1)}/N$ igual a uno (Tabla 3j). Sin embargo, las diferencias existentes en el grado de contribución parental de los reproductores, traen en consecuencia fuertes reducciones sobre el $Ne_{(2)}$ estimado respecto al tamaño censal de los lotes A y B, disminuyendo su valor un 12,8% y 48,8%, respectivamente. No obstante, a consecuencia de la participación discontinua de algunos individuos reproductores y/o debido a las intensas variaciones en la tasa de contribución parental, los valores estimados para $Ne_{(2)}$ se ven especialmente afectados dentro de cada evento de puesta, registrándose un rango de reducción del Ne más amplio, cuyas oscilaciones se sitúan entre el 12% (S2A) y el 78% (S3B) (Tabla 3j).

Esta intensa disminución del Ne , se podría fácilmente interpretar como consecuencia del establecimiento de lotes reproductores a partir de un número reducido de individuos ($N=8$). No obstante, los datos obtenidos en dorada y en otras especies acuícolas a partir de lotes cuyo tamaño censal ($N=35-131$) resulta notablemente superior al establecido en nuestro estudio (Brown y col., 2005b; Fessehaye y col., 2006; Piñera, 2009; Porta y col., 2010 y Chavanne y col., 2014), registran desde el punto de vista genético, reducciones en el tamaño censal que varían entre un 56% [$N \approx 16$, $Ne/N=0,43-0,45$ (Fessehaye y col., 2006)] y aproximadamente un 92% [$N=131$, $Ne/N=0,073$ (Porta y col., 2012)]. Por consiguiente, no existen motivos para concluir que el empleo de lotes reproductores de tamaño censal reducido pueda influir negativamente sobre la relación Ne/N , sino más bien todo lo contrario, pues la creación de lotes reproductores de pequeño tamaño parece favorecer la participación de los reproductores en la puesta, ya que a diferencia de los estudios anteriormente citados, en nuestro caso todos los reproductores

tomaron parte en la puesta, no solo a nivel global del periodo de puesta, sino también dentro de algunos de los eventos de puesta (principalmente correspondientes a la fase intermedia del periodo) analizados (Tabla 3i).

A la vista de estos resultados, se deduce que la creación de lotes reproductores a partir de descendientes procedentes de un único evento de puesta puede, a través del intenso grado de afectación del N_e , conducir a incrementos de endogamia elevados, capaces de repercutir sobre la eficacia biológica de los individuos, pudiendo condicionar a lo largo de las generaciones la continuidad del programa de reproducción establecido. Sin embargo, si bien es cierto que la creación de lotes reproductores a partir de un único evento de puesta supone siempre un riesgo elevado, se debe mencionar que dada la tendencia de N_e a mostrar un grado de conservación ligeramente superior en las puestas procedentes de la fase intermedia, la habitual práctica de manejo consistente en la selección de un único evento de puesta de dicha fase, como semilla para establecer el nuevo lote reproductor, supone, dentro del riesgo, la opción más acertada.

Ante esta situación, existe una necesidad evidente de mejorar el mecanismo de manejo reproductivo de la especie a través de instauración de medidas correctoras destinadas a optimizar el N_e de los lotes.

Como medida dirigida a minimizar la ausencia de participación parental, Brown y col. (2005b) a partir de la asociación deducida entre el peso de los individuos reproductores y la contribución parental, proponen incorporar el peso corporal de los individuos como factor a tener en cuenta en la creación de lotes reproductores. No obstante, estos autores destacan que posiblemente la asociación observada entre peso y contribución no refleje a una asociación directa entre estos dos caracteres, sino que posiblemente esté enmascarando una asociación "real" entre la edad de los individuos y la contribución parental; una hipótesis que toma fuerza con los resultados obtenidos por Jerez y col. (2012), quienes evaluando la influencia de la edad de la hembra sobre los parámetros de fecundación y fertilización, proponen la edad de las hembras reproductoras como factor de interés para minimizar las diferencias en el grado de contribución parental.

En nuestro caso, los resultados obtenidos muestran como la creación de lotes reproductores a partir del reagrupamiento de huevos flotantes recolectados en diferentes momentos de la fase intermedia del periodo de puestas, puede contribuir positivamente a optimizar el N_e de los lotes, al verse favorecido el incremento en el número de familias obtenidas y atenuada la intensidad de las fuertes dominancias ejercidas por algunas parejas reproductoras (Figura 3d y 3e). No obstante, para evitar que grandes diferencias de tamaño entre individuos (debidas a un tiempo de desarrollo desigual) puedan desembocar en

comportamientos de canibalismo, solamente se deberá unificar las descendencias cuando las diferencias de tamaño entre los individuos sean mínimas, por lo que se deben evitar que entre los eventos de puesta recolectados existan largos intervalos de tiempo.

De igual modo, en base a el elevado porcentaje de reproductores activos detectado dentro de los lotes A y B, especialmente en el transcurso de la fase intermedia del periodo de puesta donde, dentro de cada evento se registró como mínimo el 75% de los cruzamientos parentales posibles, se propone una nueva estrategia de manejo fundamentada en el reemplazamiento del sistema tradicional de reproducción, por un sistema de manejo donde los actuales lotes reproductores ($N=40-60$) se dividen (de forma previa al inicio del periodo reproductivo con antelación suficiente para garantizar la correcta aclimatación de los individuos) en grupos de pequeño tamaño censal (por ejemplo, $N=8$ individuos en una relación de sexos 1:1) denominados "*mini-lotes*". A través de este manejo, el futuro lote reproductor se constituirá a partir de la reagrupación en un único tanque de puestas procedentes de la fase intermedia de cada uno de los *mini-lotes*. La estrategia propuesta, además de favorecer el número de reproductores activos que contribuyen al establecimiento de la nueva generación, permite minimizar el efecto de las conductas dominantes adquiridas por algunos reproductores, ya que la dominancia ejercida por una pareja reproductora dentro de un *mini-lote* se irá diluyendo al combinarse con la dominancia ejercida por las parejas procedentes de otros *mini-lotes*.

Como requisito indispensable para llevar a cabo óptimamente este manejo, se precisa de la sincronización del periodo de puesta de los *mini-lotes*, la cual permitirá recolectar de forma simultánea en todos ellos, puestas procedentes de la fase intermedia con diferencias mínimas en el tiempo de desarrollo.

Además, de acuerdo a la ventaja que supone la recolección de diferentes eventos de puesta dentro de un mismo lote, la selección de al menos dos eventos de puesta diferentes por *mini-lote*, podrían contribuir a incrementar el número de familias y atenuar las diferencias de contribución parental.

No obstante, el manejo reproductivo propuesto condiciona el número máximo de familias posibles que se puede establecer por lote, pues a modo de ejemplo, la división de un conjunto de 16 individuos reproductores en dos lotes compuestos por 8 individuos (4 hembras y 4 machos), reduce los 64 cruzamientos parentales iniciales a un máximo de 16 cruzamiento por lotes, obteniéndose en total un máximo de 32 cruzamientos diferentes y por consiguiente, un máximo de 32 familias. No obstante, el número de familias podría ser fácilmente incrementado aumentando el número de *mini-lotes*, por lo tanto, inicialmente solamente sería necesario hacer una estima inicial del número de familias deseadas, y en base a ello, calcular el

número de *mini-lotes* requeridos. Además, a pesar de que lotes de tamaño censal estándar ofrecen la posibilidad teórica de obtener un mayor número de familias, los resultados obtenidos por Chavanne y col. (2014) muestran claramente como dentro de un único evento de puesta, el número de familias obtenido se ve fuertemente disminuido, resultando hasta tres veces inferior al valor máximo esperado.

Por lo tanto, la creación de un sistema de *mini-lotes* que maximice la participación parental y genere tamaños familiares más homogéneos, predispone de manera sencilla y a bajo coste a la optimización precoz del N_e , reteniendo los niveles de endogamia no deseada por debajo del umbral aceptado para preservar la eficacia biológica y evitar la depresión endogámica. A su vez, la maximización de N_e favorece la conservación de la variabilidad genética a lo largo de los ciclos de reproducción, lo cual resulta indispensable para instaurar programas de mejora genética de caracteres de interés productivo que contribuyan a una explotación óptima de la especie. Si bien es cierto que la implementación de protocolos de evaluación de la contribución parental a partir de la información proporcionada por un conjunto de *loci* microsatélites como el desarrollado en este estudio, ofrecería la garantía de seleccionar solamente aquellas puesta que en su conjunto contribuirían a conservar un N_e elevado, su desarrollo requiere una inversión económica mayor, por lo que el acuicultor deberá establecer un balance entre el beneficio que supone poder ejercer un control más preciso sobre la reproducción y el coste económico que conlleva dicho conocimiento.

3.5. Tablas

Tabla 3a. Muestras analizadas en el estudio.

Grupo		Individuos analizados	Fecha de recolección	Día de Puesta	
Lote A	Reproductores	8			
	Descendencia	S1A	72	21/01/2008	29
		S2A	66	31/01/2008	39
		S3A	88	12/02/2008	51
		S4A	88	21/02/2008	60
		S5A	88	03/03/2008	71
Total	410				
Grupo		Individuos analizados	Fecha de recolección	Día de Puesta	
Lote B	Reproductores	8			
	Descendencia	S1B	84	22/01/2008	30
		S2B	82	04/02/2008	43
		S3B	78	14/02/2008	53
	Total	260			

S1, S2, S3, S4 y S5: eventos de puesta analizados; *Día de puesta*: número de días transcurridos desde el comienzo del periodo de puesta.

Tabla 3b. Volúmenes de puesta recolectados en los eventos de puesta analizados.

	Nº máx. huevos/día	Nº medio huevos/día		Volumen puestas analizadas *				
		Flotantes (± SD)	No flotantes (± SD)	S1	S2	S3	S4	S5
Lote A	530	254,12 (±137,63)	37,84 (±41.52)	430 (80)	390 (70)	230 (60)	310 (40)	170 (40)
Lote B	480	143,04 (±121,81)	40,74 (±37,03)	330 (70)	250 (120)	90 (40)	-	-

SD: Desviación estándar; * Dentro de cada casilla se indica el número de huevos recolectado en cada fecha, mostrando el número de huevos flotantes y no flotantes (entre paréntesis).

Tabla 3c. Eficacia de amplificación por PCR en diferentes estados del desarrollo embrionario.

	2-4 Células	16-24 Células	Gástrula	Embrión visible
	1:15H-1:45H	2:30H-3:00H	10:00H-12:00H	18:00H-20:00H
				
Éxito de amplificación	0-5%	4-10%	63-70%	91-96%

Tiempos de desarrollo estimados a una temperatura de desarrollo de 18,5°C. Imágenes del desarrollo embrionario de dorada tomadas de Kamaci y col., (2005).

Tabla 3d. Porcentaje de asignación parental obtenido mediante el programa CERVUS v.3.0 (Marshall y col., 1998; Kalinowski y col., 2007), modificando el número de *loci* microsatélites empleados en el análisis.

	SIM-1P (%)	SIM-2P (%)	SIM-PP (%)
9 Microsatélites	96-97	96-98	98-99
8 Microsatélites	96-97	96-98	98-99
7 Microsatélites	96-97	96-97	98-99
6 Microsatélites	96-97	97-98	98-99
5 Microsatélites	91-93	92-93	92-94
4 Microsatélites	88-92	89-91	90-93
3 Microsatélites	67-75	68-75	74-75

SIM-1P: porcentaje de asignación maternal; *SIMP-2P*: porcentaje de asignación paternal; *SIM-PP*: porcentaje de asignación de pareja parental.

Tabla 3e. Diversidad genética deducida para los lotes reproductores a partir de 9 *loci* microsatélites.

<i>Locus</i>	Lote reproductor A										Lote reproductor B			
	N_a	H_0	H_E	PIC	NE-1P	NE-2P	NE-PP	N_a	H_0	H_E	PIC	NE-1P	NE-2P	NE-PP
<i>Dpt3</i>	10	1,000	0,930	0,860	0,556	0,379	0,196	8	1,000	0,890	0,820	0,496	0,326	0,152
<i>Dxd44</i>	12	1,000	0,960	0,890	0,404	0,251	0,096	7	0,880	0,82	0,740	0,614	0,435	0,242
<i>Pb-OVI-A3</i>	2	0,250	0,400	0,310	0,340	0,204	0,066	4	0,250	0,640	0,530	0,810	0,675	0,519
<i>Pb-OVI-B2</i>	7	0,880	0,690	0,630	0,930	0,848	0,761	5	0,380	0,610	0,540	0,815	0,640	0,447
<i>Pb-OVI-D22</i>	6	1,000	0,820	0,730	0,736	0,546	0,332	5	0,380	0,790	0,700	0,665	0,487	0,304
<i>Pb-OVI-D106</i>	7	0,880	0,860	0,780	0,630	0,452	0,269	7	0,880	0,890	0,820	0,500	0,329	0,155
<i>SaGT1</i>	11	1,000	0,950	0,880	0,363	0,221	0,076	9	0,880	0,910	0,840	0,457	0,294	0,125
<i>SaGT26</i>	8	0,880	0,880	0,800	0,520	0,347	0,165	9	0,750	0,890	0,820	0,483	0,314	0,137
<i>SaGT41b</i>	11	1,000	0,950	0,880	0,363	0,221	0,076	11	1,000	0,950	0,880	0,363	0,221	0,076
Total 9 <i>loci</i>	8,22	0,875	0,826	0,750	0,002	6,88E-05	8,00E-08	7,22	0,708	0,821	0,741	0,005	0,0002	5,30E-07
OVIDORPLEX-R	9,30	0,958	0,921	0,849	0,005	0,003	1,18E-06	8,50	0,896	0,892	0,817	0,012	0,0009	7,48E-06

N_a : número de alelos; H_0 : heterocigosidad observada; H_E : Heterocigosidad esperada; PIC: contenido de información polimórfica; NE-1P: probabilidad media de no-exclusión para un parental candidato; NE-2P: probabilidad media de no-exclusión para un parental candidato siendo conocido el genotipo del parental de sexo opuesto; NE-PP: probabilidad media de no-exclusión para una pareja parental candidata.

Tabla 3f. Evaluación del éxito de amplificación por PCR en el conjunto de descendencias.

		Descendencia A						Descendencia B					
		S1A	S2A	S3A	S4A	S5A	T	E (%)	S1B	S2B	S3B	T	E (%)
N_{an}		72	66	88	88	88	402		84	82	78	244	
N_g	<i>Dpt3</i>	67	61	82	82	80	372	92,5	74	75	73	222	91,0
	<i>Dxd44</i>	67	57	80	79	80	363	90,3	71	75	73	219	89,7
	<i>Pb-OVI-106</i>	69	60	81	83	80	373	92,8	72	75	73	220	90,1
	<i>SaGT1</i>	35	28	31	24	39	157	39,0	-	-	-	-	-
	<i>SaGT26</i>	69	61	82	84	79	375	93,3	73	75	71	219	89,7
	<i>SaGT41b</i>	67	61	82	83	80	373	92,8	71	75	73	219	89,7
	$N_{\geq 4 \text{ loci}}$	69	61	82	84	80	376	93,5	74	75	73	222	91,1

S1, S2, S3, S4 y S5: eventos de puesta; *T*: total individuos; *E*: éxito de amplificación (N_g / N_{an}); N_{an} : número de individuos analizados; N_g : número de individuos con genotipo determinado; $N_{\geq 4 \text{ loci}}$: número de individuos con un número mínimo de 4 *loci* genotipados.

Tabla 3g. Diversidad genética para los lotes reproductores y sus descendencias a partir de 5 *loci* microsatélites.

Grupo		N	N_A	H_E	H_O	F_{IS}	P-valor	
Lote A	Reproductores	8	9,6	0,915	0,950	-0,053 n,s	(0,2417)	
	Descendencia A	S1A	69	8,8	0,822	0,950	-0,158*	(0,0017)
		S2A	61	8,8	0,846	0,860	-0,016 n,s	(0,2600)
		S3A	82	8,8	0,819	0,857	-0,047*	(0,0100)
		S4A	84	8,6	0,842	0,887	-0,054*	(0,0017)
		S5A	80	8,2	0,847	0,892	-0,053*	(0,0033)
	Total	276	8,8	0,852	0,889	-0,044*	(0,0050)	
Grupo		N	N_A	H_E	H_O	F_{IS}	P-valor	
Lote B	Reproductores	8	8,4	0,892	0,900	-0,040 n,s	(0,3650)	
	Descendencia B	S1B	74	8,4	0,832	0,835	-0,004 n,s	(0,4050)
		S2B	75	5,6	0,675	0,835	-0,238*	(0,0025)
		S3B	73	5,6	0,682	0,850	-0,249*	(0,0025)
	Total	222	8,4	0,776	0,841	-0,083*	(0,0050)	

S1, S2, S3, S4 y S5: eventos de puesta; *N*: número de individuos; N_A : número medio de alelos; H_E : heterocigosidad esperada; H_O : heterocigosidad observada; F_{IS} : evaluación del equilibrio Hardy-Weinberg; * P-valor<0,05; n,s: no significativo).

Tabla 3h. Valores de F_{ST} obtenidos en el análisis de diferenciación genética por pares de muestras.

Lote A	S1A	S2A	S3A	S4A	S5A
Reproductores	-0,029 (0,856)	0,0203 (1,000)	0,0140 (0,986)	0,0113 (0,763)	0,0141 (0,596)
S1A		0,0059 (0,055)	0,0128 (0,003)*	0,0402 (0,003)*	0,0351 (0,003)*
S2A			0,007 (0,016)*	0,0257 (0,003)*	0,0248 (0,003)*
S3A				0,0369 (0,003)*	0,0033 (0,003)*
S4A					0,0174 (0,003)*

Lote B	S1B	S2B	S3B
Parentales	-0,0194 (1,000)	0,0580 (0,008)*	0,0576 (0,008)*
S1B		0,0433 (0,008)*	0,0939 (0,008)*
S2B			0,1324 (0,008)*

S1, S2, S3, S4 y S5: eventos de puesta Entre paréntesis se indica el correspondiente valor de significación; (*) P-valor<0,05.

Tabla 3i. Valores de contribución parental.

Lote A	Descendencia					
Parentales	S1A	S2A	S3A	S4A	S5A	T
N	69	61	82	84	80	376
H033	34	18	24	25	35	136
H131	14	18	25	14	0	71
H155	8	7	25	0	21	61
H217	13	18	8	45	24	108
P102	43	23	21	13	21	121
P108	12	18	33	27	20	110
P109	8	10	10	41	39	108
P111	6	10	18	3	0	37

Lote B	Descendencia				
Parentales	S1B	S2B	S3B	T	
N	74	75	73	222	
H036	1	7	69	77	
H123	17	0	1	18	
H154	19	0	0	19	
H279	37	68	3	108	
P101	3	0	0	3	
P104	14	17	4	35	
P105	23	13	2	38	
P112	34	45	67	146	

S1, S2, S3, S4 y S5: eventos de puesta; T: Total; N: número de individuos descendientes; ♀: hembras; ♂: machos

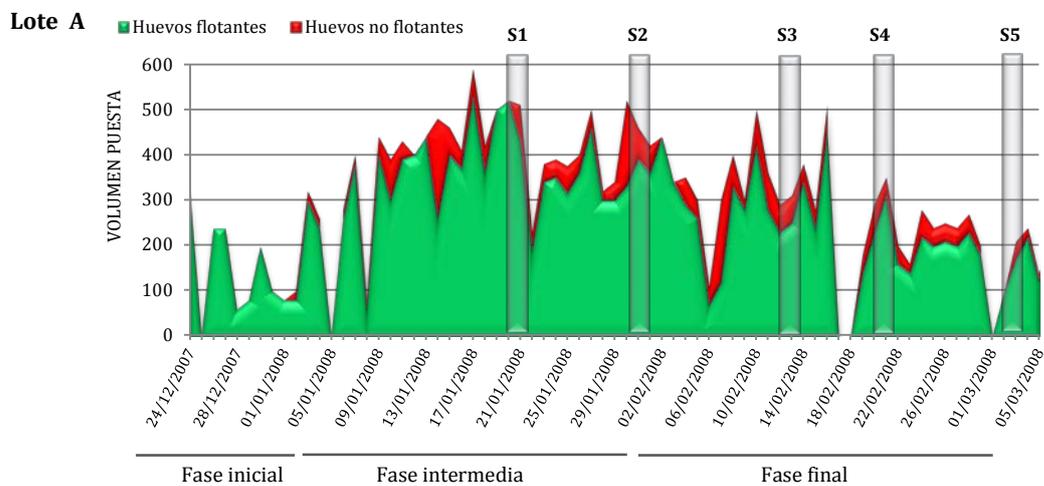
Tabla 3j. Valores de tamaño efectivo (N_e) estimado para cada fecha de puestas y para el global del periodo reproductivo.

	Lote A							Lote B							
	S1A	S2A	S3A	S4A	S5A	Total	FI	FF	S1B	S2B	S3B	Total	FI	FF	
$N_e^{(1)}$	8,00	8,00	8,00	6,85	6,00	8,00	8,00	8,00	8,00	4,80	6,00	8,00	8,00	8,00	6,00
$N_e^{(2)}$	4,65	7,11	6,83	4,49	4,15	6,97	6,78	5,25	5,13	2,81	2,07	4,09	5,13	3,15	
$N_e^{(2)}/N$	0,58	0,88	0,85	0,56	0,52	0,87	0,84	0,65	0,64	0,35	0,26	0,51	0,64	0,52	
ΔF_1	0,06	0,06	0,06	0,07	0,08	0,06	0,06	0,06	0,06	0,10	0,08	0,06	0,06	0,08	
ΔF_2	0,10	0,07	0,07	0,11	0,12	0,07	0,07	0,09	0,09	0,18	0,24	0,12	0,09	0,15	
$\Delta F_2/\Delta F_1$	1,71	1,12	1,17	1,57	1,37	1,14	1,15	1,50	1,55	1,80	2,89	1,95	1,55	1,90	

S1, S2, S3, S4 y S5: eventos de puesta; $N_e^{(1)}$: Tamaño efectivo calculado mediante la fórmula $N_e^{\wedge}=4(Nh Nm)/(Nh+ Nm)$ (Falconer, y Macka, 2001); $N_e^{(2)}$: tamaño efectivo calculado mediante la fórmula $N_e^{\wedge}=4(N-2)/((Kh+Vh/Kh)+(Km+Vm/Km)-2)$ (Chevassus B., 1989); N: número de individuos que conforman el lote reproductor; ΔF_1 : incremento de endogamia calculado utilizando $N_e^{(1)}$ ($\Delta F_1=1/(2 N_e^{(1)})$); ΔF_2 : incremento de endogamia calculado utilizando $N_e^{(2)}$ ($\Delta F_2=1/(2 * N_e^{(2)})$); (*) Unificación de la descendencia obtenida a partir de los lotes A y B, en un único conjunto de descendientes; FI: fase intermedia; FF: fase final.

3.6. Figuras

Figura 3a. Volúmenes de puesta recolectados a lo largo de periodo reproductivo en los lotes A y B. Las fechas de puesta analizadas se muestran remarcadas con barras verticales.



Lote B

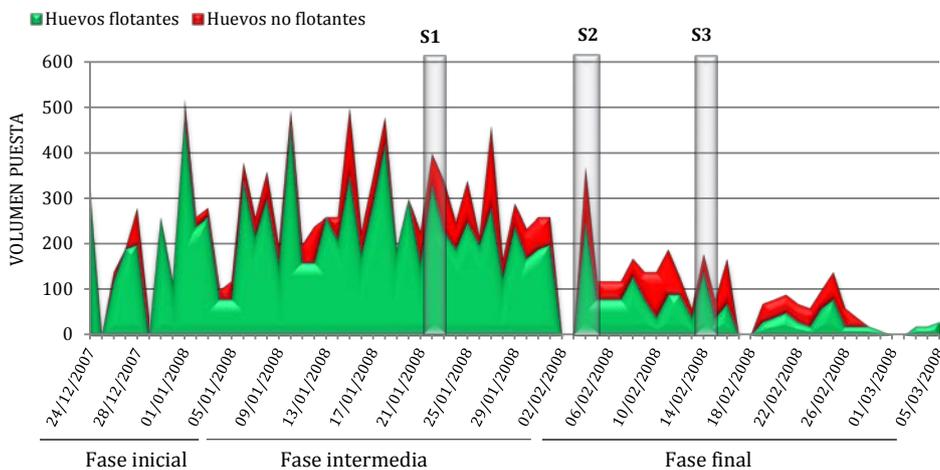
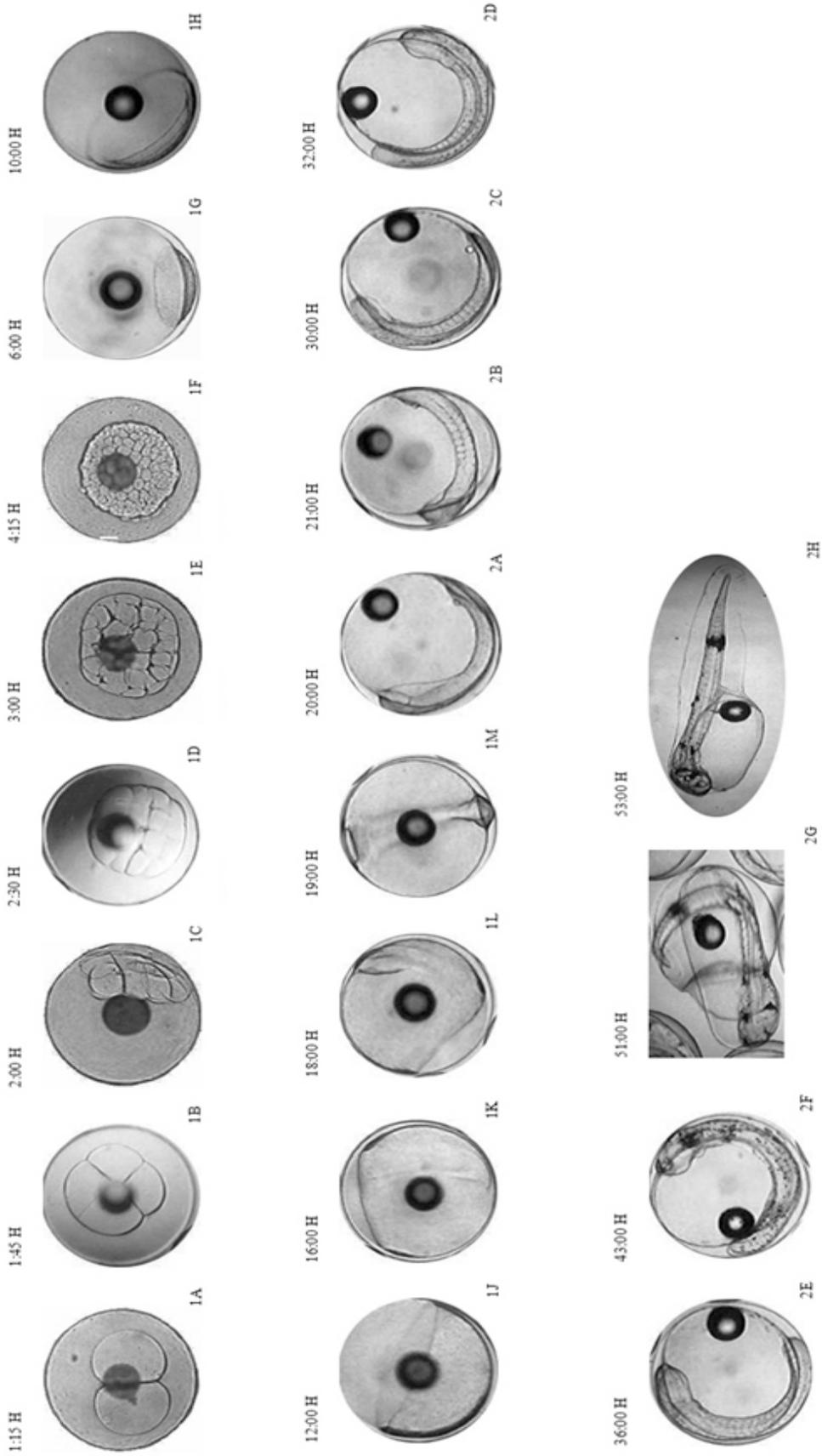


Figura 3b. Desarrollo embrionario de dorada (Kamaci y col., 2005).



Temperatura de desarrollo de 18,5°C.

Figura 3c. PCR multiplex *OVIDORPLEX-R*. **A:** diseño de PCR multiplex (*OVIDORPLEX-R*) propuesto por el programa Multiplex Manager (Holleley and Geerts, 2009). Para cada *loci* microsatélite se indica su rango alélico (mediante los números situados a ambos extremos de cada *loci*) y el tipo fluoróforo empleado en su marcaje. **B:** Electroforetrograma obtenido para *OVIDORPLEX-R* en dos individuos.

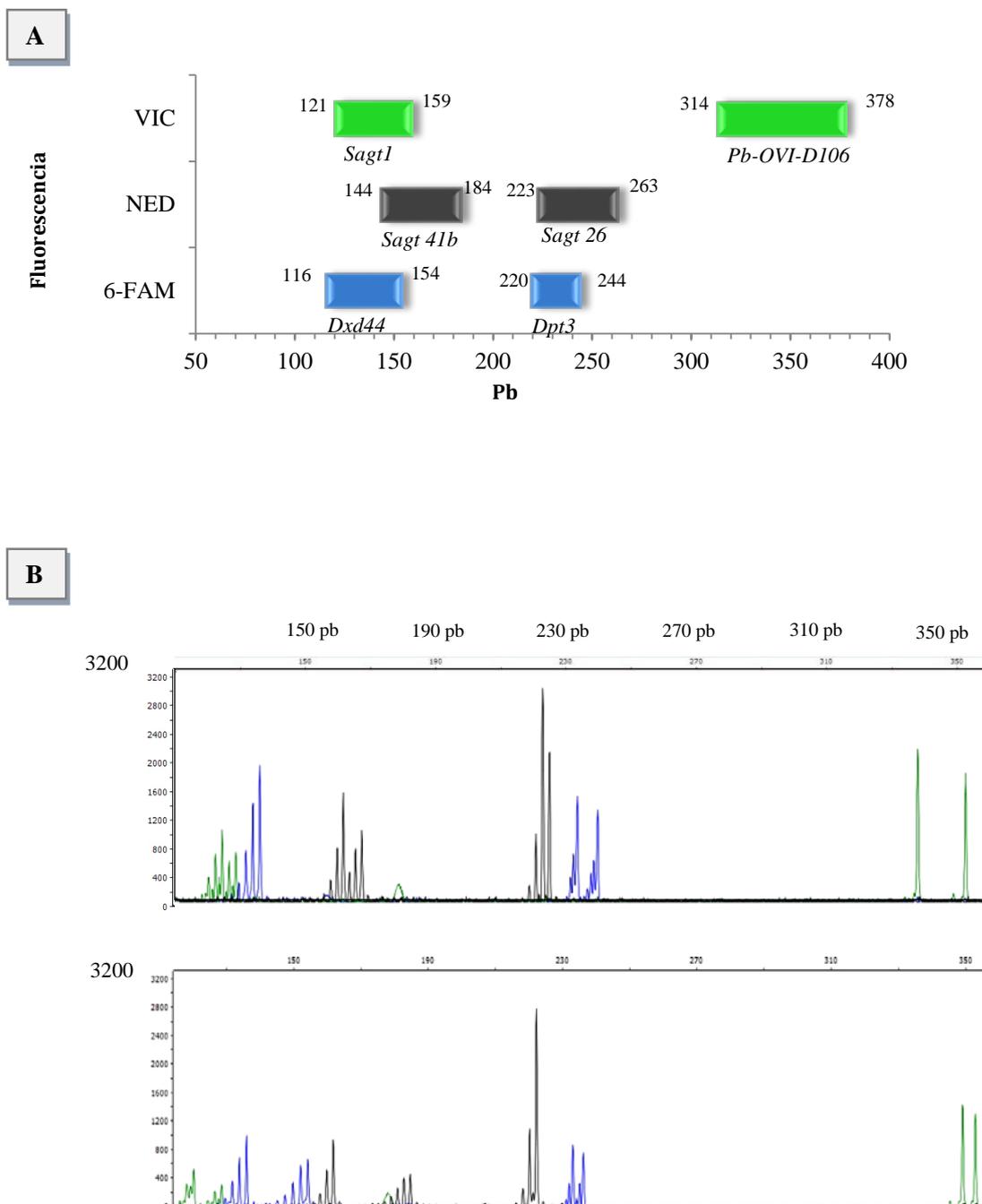
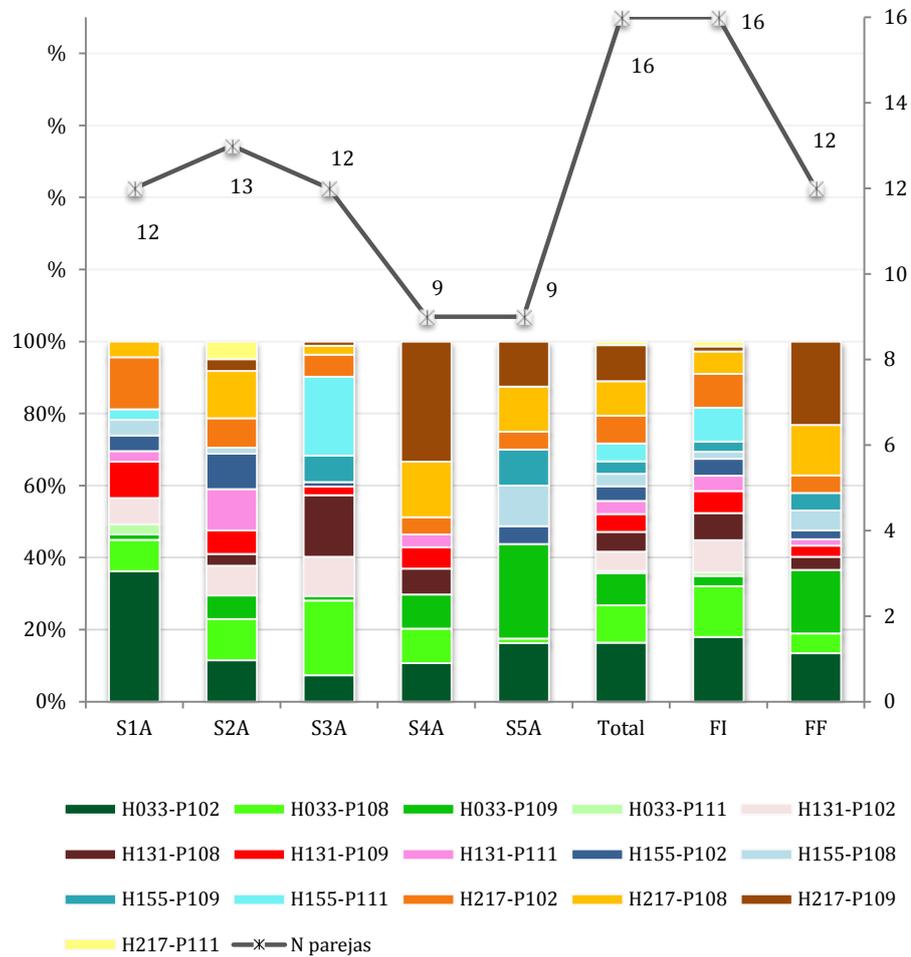
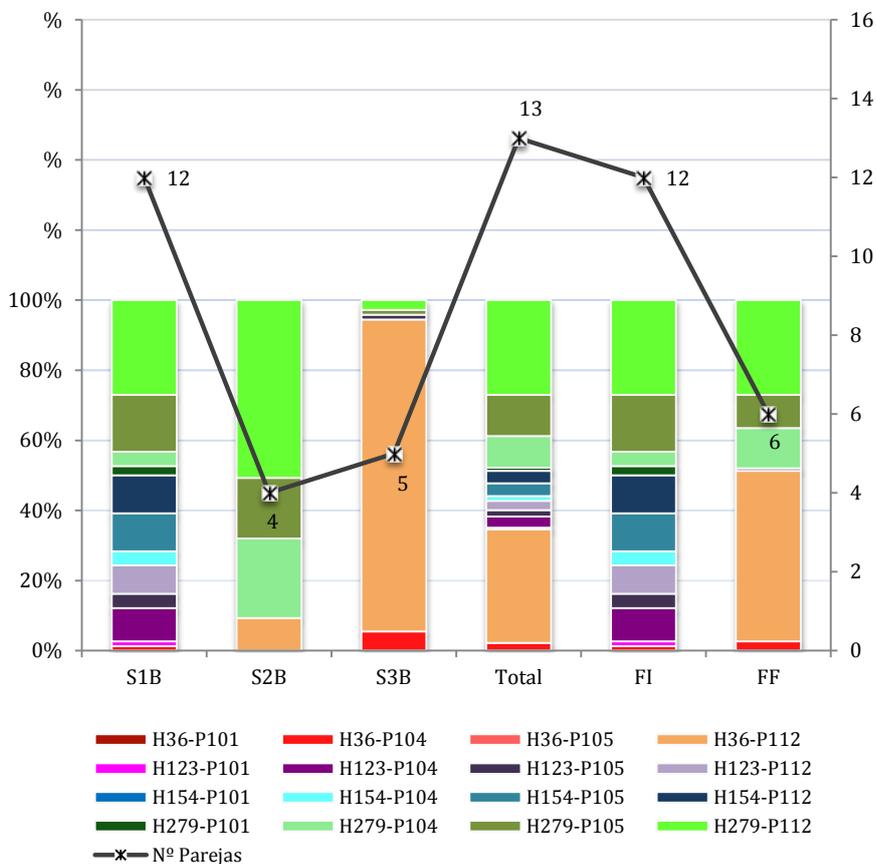


Figura 3d. Número de parejas reproductoras participantes y patrón de contribución parental obtenido en el lote A.



Eje abscisas: eventos de puesta analizados; Eje ordenadas (primario): porcentajes de contribución parental; Eje ordenadas (secundario): número de parejas reproductoras participantes; FI: fase intermedia; FF: fase final.

Figura 3e. Número de parejas reproductoras participantes y patrón de contribución parental obtenido en el lote B.



Eje abscisas: eventos de puesta analizados; Eje ordenadas (primario): porcentajes de contribución parental; Eje ordenadas (secundario): número de parejas reproductoras participantes; FI: fase intermedia; FF: fase final

4. Capítulo II

Uso de loci microsatélites para la identificación de reproductores de interés productivo

4.1. Introducción

En acuicultura, la optimización de procesos destinados a la obtención de un manejo correcto de la especie como son la creación/control de un ambiente de cultivo adecuado, el suministro de una dieta alimentaria que cumpla con las necesidades nutricionales, el control exhaustivo sobre los niveles de endogamia, etc., conduce de forma pasiva a una mejora en la productividad del cultivo. No obstante, los beneficios de esta producción se pueden acrecentar activamente mediante la mejora genética de aquellos caracteres cuya optimización conlleve a una reducción en los costes producción (por ejemplo, la tasa de crecimiento, la eficacia de conversión alimentaria o la resistencia a enfermedades), o bien, a la obtención de un producto final de mayor calidad (por ejemplo, la calidad de la carne o la textura del filete) que alcance en el mercado un valor económico superior.

La industria acuícola desde siempre ha mostrado un gran interés en el desarrollo de sistemas de mejora destinados a maximizar las tasas de crecimiento de los individuos. Basándose en el hecho de que tasas de crecimiento elevadas habitualmente se asocian a una conversión de alimento más eficaz (Thodesen y col., 1999), la industria, a través de la obtención de individuos con una alta capacidad de aprovechamiento de los recursos alimenticios, busca generar un abaratamiento en los costes de alimentación del cultivo, los cuales, en dorada, se estima que representan aproximadamente el 40% de los gastos de producción (Jover, 2007). Además, la selección hacia crecimiento rápido resulta de interés porque en algunas especies se ha detectado que de forma simultánea e indirecta, favorece la selección de individuos con comportamientos no agresivos (Doyle y Tabot, 1986), y predispone a un incremento en la resistencia a enfermedades y en la tasa de supervivencia (Fjalestad y col, 1993; Gjedrem y Thodesen, 2005).

En general, la tasa de crecimiento de las especies acuáticas muestra una buena capacidad de respuesta a la selección individual o masal, debido principalmente al elevado coeficiente de variación genética (20-35%) que presenta el carácter, así como a la alta fecundidad que poseen las especies, la cual permite aplicar una presión de selección intensa (Gjedrem, 1983). No obstante, la selección masal prevalente en acuicultura se fundamenta en un criterio de selección puramente fenotípico, que al no reparar en el grado de consanguinidad existente entre los individuos seleccionados, predispone a importantes incrementos de endogamia (Bentsen y Olesen, 2002). Por lo tanto, tal y como se mostró en el *Capítulo I*, el uso de marcadores moleculares para esclarecer las relaciones de consanguinidad existentes entre los individuos, constituye el complemento perfecto para el desarrollo de los esquema de

selección, posibilitando que los individuos reproductores sean seleccionados simultáneamente en base a su fenotipo y a su grado de parentesco. A su vez, la combinación de procesos de asignación parental, junto con una clasificación de la descendencia de acuerdo al fenotipo mostrado para un determinado carácter, permite evaluar la capacidad que posee cada reproductor para contribuir a la siguiente generación con descendientes portadores del fenotipo de interés.

Dentro de este contexto, se engloba el objetivo de este segundo capítulo donde, a partir de la descendencia obtenida para dos lotes reproductores y aplicando un sistema de selección exclusivamente fenotípico, se pretende establecer dos categorías de crecimiento con diferencias significativas (*crecimiento-rápido* y *crecimiento-lento*), para posteriormente, a través de un análisis de asignación de parental, determinar si entre los reproductores existe una contribución parental diferencial hacia cada categorías. Bajo este supuesto, la identificación de individuos reproductores que aporten preferentemente descendientes a la categoría de *crecimiento-rápido*, contribuirá a la creación de lotes reproductores capaces de mejorar la rentabilidad del cultivo.

4.2. Materiales y Métodos

4.2.1. Muestras biológicas

En Diciembre de 2006, a partir de un conjunto de 101 individuos adultos de dorada, cultivados en las instalaciones de la empresa Piscicultura Marina Mediterránea S.L. (Burriana, Castellón, España), se crearon dos lotes reproductores (lote C y D) compuestos cada uno de ellos por aproximadamente 40 individuos [lote C: 29 hembras (peso medio $1998,2 \pm 43,5\text{g}$) y 11 machos (peso medio $1877 \pm 83,4\text{g}$); lote D: 28 hembras (peso medio $1992,8 \pm 50,1\text{g}$) y 10 machos ($1950,0 \pm 76,3\text{g}$)] de acuerdo al procedimiento descrito por Borrell y col. (2007). El procedimiento de selección aplicado se basó en una aproximación de optimización combinatoria dirigida a la creación de los dos mejores grupos de individuos en base a un criterio de mínimo coeficiente de relación (coeficiente R) (Queller y Goodnight, 1989) y máxima variabilidad genética; empleando el programa Relatedness (Queller y Goodnight, 1989) para la determinación del coeficiente R. Teóricamente, individuos con una relación de parentesco de hermanos completos, deberían presentar valores medios de R cercanos a 0,5, puesto que de media comparten el 50% de su genoma, mientras que en el caso de los medios-hermanos y entre individuos no emparentados, el valor medio esperado para R sería del 0,25 y cero, respectivamente. Sin embargo, los valores del coeficiente R obtenidos varían entre -1 y +1, indicando el signo positivo una relación genética entre individuos mayor de la esperada, mientras que en contraposición la obtención de valores negativos indicaría una relación genética menor de la esperada (Queller y Goodnight, 1989). Para el conjunto global de individuos (101 individuos iniciales) los valores del coeficiente R oscilaron entre $-0,39 < R < +0,83$, y mediante el proceso de selección aplicado, estos valores quedaron restringidos a $-0,28 < R < +0,54$ en el lote C y $-0,25 < R < +0,43$ en el lote D.

Durante el la fase intermedia del periodo de puesta (Marzo 2007), con el objetivo de evitar que las diferencias en la tasa de crecimiento entre los individuos descendientes se debieran a tiempos de desarrollo diferentes, para cada lote se recolectó e incubó bajo condiciones intensivas, la puesta procedente de un único día (diferente día para cada lote). Transcurridos 86 (lote C) y 78 (lote D) días desde el momento de la eclosión, se procedió mediante el empleo de diferentes tamices de criba, a una primera división de los individuos descendientes en base al tamaño corporal, generándose las categorías *crecimiento-rápido* y *crecimiento-lento* (Tabla 4a).

A la edad de 165 días para el lote C y 157 para el lote D, los individuos se enviaron a la Universidad Politécnica de Valencia, donde se mantuvieron en tanques de 1.750L de capacidad

con recirculación, filtrado mecánico rotacional, biofiltrado por gravedad y aireación, bajo condiciones controladas de temperatura ($22,2 \pm 2,7^{\circ}\text{C}$), fotoperiodo natural, similares condiciones de luz en todos los tanques y una dieta alimentaria a base de pienso comercial (Dibaq S.A., Segovia, España).

Dentro de cada categoría, alcanzado un estado de crecimiento intermedio (transcurridos 311 y 303 días desde la eclosión para el lote C y D, respectivamente), se procedió a una segunda fase de cribado basada en peso corporal de los individuos. De acuerdo a este proceso, dentro de la categoría *crecimiento-rápido* de cada lote, se seccionaron los individuos de mayor peso (Tabla 4a), estableciéndose las categorías de *crecimiento-rápido* C12F y D12F en los lotes C y lote D, respectivamente. En la categoría de *crecimiento-lento* se aplicó el criterio de selección totalmente opuesto, constituyéndose las categorías C12S y D12S a partir de los individuos de menor peso corporal (Tabla 4a). Cada una de estas cuatro nuevas categorías, se transfirió a un tanque de cultivo de mayor capacidad (4.000L) donde, bajo condiciones intensivas, se mantuvieron hasta alcanzar la talla comercial.

A la edad de 528 días para lote C y 520 días para lote D, mediante el pesaje de aproximadamente 540 individuos incluidos en cada de las cuatro categorías anteriores, se establecieron cuatro nuevos grupos: C18F (N=150) (*crecimiento-rápido* lote C); C18S (N=150) (*crecimiento-lento* lote C); D18F (N=120) (*crecimiento-rápido* lote D) y D18S (N=120) (*crecimiento-lento* lote D). Para cada uno de estos individuos descendientes incluidos en estas categorías, así como para el total de reproductores empleados en el estudio, se recolectó en etanol al 96% una pequeña muestra de aleta caudal, para los correspondientes análisis genéticos.

4.2.2. Extracción de ADN y desarrollo de la PCR multiplex OVIDORPLEX

La extracción de ADN genómico, se realizó siguiendo el protocolo estándar de resina Chelex®100 (Biorad) (500 μL al 10%) y proteinasa K (Walsh y col., 1991) (10 μL a 20mg/mL) a partir de un pequeño fragmento de aleta caudal.

Utilizando nuevamente el programa Multiplex Manager 1.0 (Holleley y Geerts, 2009) (véase *Capítulo I*, apartado 3.2.3., pág. 32), empleando como referencia la información de marcadores microsatélites obtenida en un estudio previo para más de 800 individuos (Borrell y col., 2007), se desarrolló una PCR multiplex compuesta por nueve *loci* microsatélites: *SaGT1*, *SaGT26* y *SaGT41b* (Batargias y col., 1999); *Pb-OVI-D106* y *Pb-OVI-D102* (Piñera y col., 2006); *Dxd44* y *Dpt3* (De la Herrán y col., 2003); $\mu 184$ y $\mu 190$ (Power y col., 2003).

Las reacciones de PCR se realizaron en un termociclador XP-Cycler (Bioer), utilizando el

kit comercial QIAGEN Multiplex PCR, con concentraciones finales 1X para tampón de reacción (incluye tampón de reacción, $MgCl_2$ y HotStartTaq ADN polimerasa) y 1X para la mezcla de iniciadores (0,6 μ M para SaGT1 y Pb-OVI-D106; 0,5 μ M para SaGT41b; 0,4 μ M para Pb-OVI-D102, Dxd44 y Dpt3; 0,3 μ M para SaGT26, μ 184 y μ 190), empleando 1 μ L de ADN genómico para un volumen final de reacción de 15 μ L. El programa de amplificación utilizado consistió de una fase de desnaturalización inicial a 95°C (15 minutos), seguida de 30 ciclos de amplificación [94°C (30 segundos), 52°C (1 minuto y 30 segundos) y 72°C (1 minuto)], finalizando con una fase a 60°C (10 minutos). Todas las reacciones se comprobaron en un gel de agarosa al 2% con bromuro de etidio, y posteriormente se analizaron mediante fluorescencia en un secuenciador automático ABI 3130XL Genetic Analyzer, utilizando el programa GenMapper® 4.0 (Applied Biosystems) para el establecimiento de los correspondientes genotipos.

4.2.3. Estimación de parámetros genéticos

Los análisis de variabilidad y diversidad (parámetros N_a , H_o , H_e , frecuencias génicas, índices de fijación F_{IS} y F_{ST}), las estimas de tamaño efectivo ($Ne_{(1)}$ y $Ne_{(2)}$) y la asignación parental, se realizaron utilizando los procedimientos descritos en los apartados 3.2.5, 3.2.6 y 3.2.7 incluidos en el *Capítulo I* (pág. 33-35).

4.2.4. Evaluación del crecimiento y contribución parental

Las diferencias en el peso medio entre las categorías de crecimiento se evaluaron mediante la prueba estadística de normalidad de Kolmogorov-Smirnov (prueba K-S) incluida en el programa estadístico SPSS 15.0. La prueba K-S es una prueba no paramétrica que se utiliza para determinar la bondad de ajuste de dos distribuciones de probabilidad entre sí, a partir de dos muestras independientes, empleando el estadístico Z.

Para evaluar la contribución de los reproductores a cada una de las categorías de crecimiento (*rápido/lento*), una vez asignada a cada descendiente la pareja parental más probable, utilizando el programa estadístico SPSS v15.0, se contrastó para cada lote la hipótesis nula de: contribución equitativa a las categorías de *crecimiento-rápido* y *lento*. Para evaluar la existencia de la posible contribución diferencial de machos y hembras dentro de cada categoría de crecimiento, se utilizó un método de re-muestreo (Patefield, 1981; Corral, 2005; Carleos, 2010) donde de forma pseudo-aleatoria y manteniendo invariables los totales marginales, se generaron 100.000 tablas de contingencia, cuantificando estadísticamente la probabilidad de obtener para un reproductor dado las frecuencias observadas (o frecuencias más extremas) bajo la hipótesis nula de contribución equitativa. De este modo, valores de significación inferiores a 0,05, indican una baja probabilidad de las frecuencias observadas para el supuesto

planteado en la hipótesis nula. El análisis de re-muestreo se desarrolló en su totalidad mediante el programa estadístico R statistical environment (R Development Core Team, 2009).

4.3. Resultados

4.3.1 Creación de grupos de descendencias con crecimiento diferencial

Los diferentes pasos de cribado aplicados sobre el conjunto de descendientes obtenidos, conllevaron al establecimiento de las categorías *crecimiento-rápido* y *crecimiento-lento* dentro de cada lote.

A través de la comparación del peso medio de los descendientes a los 18 meses de edad (Tabla 4b), se observaron diferencias significativas entre las dos categorías de crecimiento que conforma cada lote (lote C: $Z_{C18F-C18S}=6,83$, P-valor $<0,001$; lote D: $Z_{D18F-D18S}=6,91$, P-valor $<0,001$), y entre cada una de la categorías de crecimiento del lote C y D ($Z_{C18F-D18F}=4,42$, P-valor $<0,001$; $Z_{C18S-D18S}=3,28$, P-valor $<0,01$).

4.3.2 PCR multiplex OVIDORPLEX

El sistema OVIDORPLEX desarrollado para este estudio (Figura 4a) resultó altamente eficaz, permitiendo genotipar más del 99% de los individuos analizados (78/78 reproductores y 520/540 descendientes).

Durante el proceso de lectura de los electroforetogramas obtenidos a partir de OVIDORPLEX, se detectaron tres pequeños artefactos de amplificación (VIC-164pb, NED-167pb y 6-FAM-199pb) (Figura 4a), pero su presencia no afectó, ni tampoco dificultó el proceso de genotipado.

4.3.3 Caracterización de la diversidad genética

Los dos lotes reproductores C y D, se caracterizaron por presentar una alta diversidad genética (Tabla 4c), mostrando valores elevados en el número medio de alelos por *locus* ($N_A \geq 14$) y niveles de heterocigosidad ($H_o \geq 0,850$), aunque su composición genética no mostró diferencias significativas ($F_{ST} = -0,005$; P-valor=0,996). Los valores de F_{IS} obtenidos indicaron que ambos lotes de reproductores se encuentran en equilibrio de Hardy-Weinberg, pues no se detectan desviaciones significativas hacia un exceso/déficit de heterocigotos (Tabla 4c).

En referencia a las descendencias, la comparación de las categorías *crecimiento-rápido* y *crecimiento-lento* dentro de cada lote, así como su comparación entre lotes, revelaron para los parámetros N_a y H_o , valores similares entre sí y respecto a su correspondiente generación parental (Tabla 4c). Sin embargo, la obtención de valores de F_{IS} negativos, indicó la existencia de un exceso de heterocigotos en la generación descendiente de ambos lotes, ocasionando un

desajuste de las frecuencias genotípicas respecto a las esperadas en el equilibrio de Hardy-Weinberg, con niveles de significación superiores en la categoría *crecimiento-rápido* (P-valor<0,001) respecto a la categoría *crecimiento-lento* (P-valor<0,01) (Tabla 4c).

De acuerdo a los valores de F_{ST} obtenidos al comparar las dos categorías de crecimiento dentro de los lotes ($F_{ST\ C18F-C18S}=0,023$, P-valor=0,003; $F_{ST\ D18F-D18S}=0,016$, P-valor=0,003) y entre los lotes ($F_{ST\ C18F-D18F}=0,043$, P-valor=0,003; $F_{ST\ C18S-D18S}=0,036$, P-valor=0,003), se determinó que el conjunto de individuos descendientes que componen cada una de las categorías de cada lote debería ser considerado genéticamente diferente (diferenciación dentro de lote), y de igual manera, el conjunto de individuos pertenecientes a cada una de las categorías de crecimiento de cada lote, también debería ser considerado genéticamente diferente del conjunto de individuos que componen su misma categoría en el otro lote (diferenciación entre lotes).

4.3.4 Tamaño efectivo y participación parental

El proceso de asignación parental mostró un éxito similar a las estimas realizadas previamente mediante simulaciones teóricas a partir de la información genotípica parental (99,9% para ambos lotes reproductores), logrando asignar la pareja parental más probable a cada uno de los 516 descendientes genotipados.

Atendiendo únicamente al número de reproductores que constituyen cada lote (lote C: 29 hembras y 11 machos; lote D: 28 hembras y 10 machos), el tamaño efectivo estimado se corresponde con $Ne_{(1)}\approx 32$ y $Ne_{(1)}\approx 29$, para el lote C y D respectivamente (Tabla 4c). Sin embargo, en ambos lotes se registró una ausencia de participación parental (Tabla 4c), provocando una sobrestimación en los valores de Ne deducidos. En lote C, la ausencia de participación parental alcanzó el 27,5% quedando así excluidos del estudio 1 macho y 10 hembras, reduciéndose un 18,7% el Ne estimado inicialmente (Tabla 4c). En el lote D, el porcentaje de individuos que no contribuyeron a la descendencia analizada ascendió hasta el 31,5%, viéndose afectadas 11 hembras y 1 macho, deduciéndose un valor de $Ne_{(1)}\approx 23$ (Tabla 4c).

Dentro de cada categoría de crecimiento el Ne sufre generalmente reducciones más intensas, ($Ne_{(1)C18F}\approx 18$ y $Ne_{(1)C18S}\approx 26$; lote D: $Ne_{(1)D18F}\approx 20$ y $Ne_{(1)D18S}\approx 22$), ya que algunos de los reproductores activos solamente aportan descendientes a una de las dos categorías (Tabla 4c). Además, la existencia de contribuciones significativamente diferentes entre los reproductores, repercutió negativamente sobre el valor de $Ne_{(1)}$ estimado anteriormente ($Ne_{(2)C18F}\approx 13$; $Ne_{(2)C18S}\approx 25$; $Ne_{(2)D18F}\approx 14$ y $Ne_{(2)D18S}\approx 18$) (Tabla 4c). En las dos estimas de Ne realizadas ($Ne_{(1)}$ y $Ne_{(2)}$), Ne mostró siempre un valor inferior dentro de la categoría de *crecimiento-rápido*,

deduciéndose un ΔF más elevado dentro de esta categoría (Tabla 4c).

En ambos lotes se detectó una contribución diferencial de los reproductores a las dos categorías de crecimiento [(Hembras lote C: $\chi^2=114,2$ (P-valor<0,001); Machos lote C: $\chi^2=69,2$ (P-valor<0,001); Hembras lote D: $\chi^2=84,3$ (P-valor<0,001); Machos lote D: $\chi^2=50,6$ (P-valor<0,001)]. En el caso del lote C, 8 de las 19 hembras y 6 de los 10 machos participantes (responsables de 76% de la descendencia) mostraron una contribución parental sesgada para *crecimiento-rápido/-lento* (Figura 4b). En el lote D la contribución diferencial afectó a 7 de las 17 hembras, junto con 4 de los 9 machos participantes (responsables del 74% de los descendientes) (Figura 4b). En total, se identificaron 30 reproductores con una contribución equitativa hacia las dos categorías de crecimiento, 10 reproductores con una contribución significativamente superior a la categoría *crecimiento-rápido* (lote C: 2 hembras y 2 machos; Lote D: 4 hembras y 2 machos) y 15 reproductores con contribución significativamente superior a la categoría *crecimiento-lento* (lote C: 7 hembras y 2 machos; lote D: 3 hembras y 2 machos).

4.4. Discusión

4.4.1. Sistema de genotipado OVIDORPLEX

La incorporación de tres nuevos *loci* microsatélites (*Pb-OVI-D102*, $\mu 190$ y $\mu 184$) a la multiplex OVIDORPLEX-R utilizada en el *Capítulo I* (véase pág. 37, apartado 3.3.3.), conllevó a la obtención de una nueva versión de PCR multiplex (OVIDORPLEX) optimizada para combinar nueve *loci* microsatélites en una única reacción de PCR. La introducción de nuevos marcadores a OVIDORPLEX-R, surge ante la necesidad de incrementar el poder de discriminación de la herramienta de genotipado a nivel de individuo, debido a que los lotes reproductores C (N=40) y D (N=38) presentan un tamaño censal que supera hasta en cinco veces el tamaño de los lotes reproductores A y B (N=8) empleados en el *Capítulo I*. De este modo, la creación de OVIDORPLEX supuso el desarrollo de una herramienta de análisis genético que responde satisfactoriamente a las exigencias del nuevo estudio, resolviendo con éxito el genotipo de más del 99% de los individuos analizados.

La caracterización de los lotes C y D mediante OVIDORPLEX, pone de manifiesto la eficacia del protocolo desarrollado por Borrell y col. (2007) para crear lotes reproductores que presenten una variabilidad genética elevada, la cual en este caso se refleja a través del elevado número medio de alelos por locus ($N_A > 14$) y de niveles de heterocigosidad observados ($H_o > 0,800$); variabilidad genética, que además de ser base fundamental para lograr una estimación correcta de los parámetros genéticos y reconstruir con precisión las relaciones genealógicas, constituye un factor crítico para desarrollar con éxito programas de selección, puesto que de ella depende la capacidad de respuesta a la selección aplicada, así como los límites de la respuesta a largo plazo (Falconer y Mackay, 2001).

Los valores de variabilidad determinados mediante OVIDORPLEX, resultaron superiores a los descritos por Navarro y col. (2008) en la evaluación de 148 individuos de dorada (66 reproductores y 82 descendientes) utilizando las PCR multiplex Rim A y Rim B, compuestas por 10 y 7 *loci* microsatélites, respectivamente. De igual modo, la variabilidad detectada con OVIDORPLEX, alcanzó valores superiores a los presentados por Porta y col. (2010) a partir del análisis de tres lotes reproductores con tamaños comprendidos entre los 30 y 130 individuos, mediante 10 *loci* microsatélites combinados en las PCRs multiplex SaGT6 y SaGT4, compuestas por 6 y 4 *loci* microsatélites, respectivamente. Sin embargo, a pesar de que la variabilidad mostrada por SaGT6 y SaGT4 (Porta y col., 2010) es elevada, la imposibilidad de combinar los 10 marcadores en única reacción de PCR, supone una desventaja frente a Rim A y OVIDORPLEX que al combinar un mayor número de *loci* polimórfico por

reacción, ofrecen una optimización mayor del sistema de genotipado, incrementando las ventajas a nivel de reducción de costes y tiempo necesario para el procesamiento de muestras. Concretamente, Navarro y col., 2008 estiman que una reacción multiplex compuesta por 10 *loci* microsatélites supone una reducción de más de seis veces los costes asociados al sistema tradicional de genotipado independiente de los *loci* a nivel de personal, consumibles y de análisis.

Nuevamente, a través de este estudio se valida la extracción de ADN genómico mediante el protocolo con resina Chelex, como un método de extracción eficaz, rápido y económico. No obstante, es posible que la introducción de protocolos que permitan aislar un ADN genómico de mayor calidad, podría ayudar a reducir o eliminar los artefactos de amplificación detectados en los electroforetogramas de OVIDORPLEX (Figura 4a). Sin embargo, se debe resaltar que en esta ocasión el balance entre la mejora ocasionada y el incremento en los costes de análisis no resultaría favorable, ya que los artefactos identificados no dificultan ni afectan a la capacidad de determinación del genotipo de los individuos, por lo que el incremento en los costes de análisis que conlleva la obtención de un ADN de mayor calidad tendría como única finalidad resolver un efecto óptico derivado de la técnica de análisis.

4.4.2. Dinámica reproductiva: implicaciones de una contribución parental diferencial

En concordancia con el *Capítulo I*, la reconstrucción de las relaciones de parentesco a partir de descendientes procedentes de un único evento de puesta, pese a proceder de la fase intermedia del periodo de puesta, registra una importante ausencia de participación parental en ambos lotes (27,5% y 31, 5% en el lote C y D, respectivamente) (Tabla 4c). No obstante, en esta ocasión la ausencia de participación parental detectada probablemente se encuentre sobrestimada, pues se debe tener en consideración que el conjunto de descendientes analizados procede de un proceso de clasificación fenotípica que derivó en el establecimiento de las categorías de *crecimiento-rápido* y *-lento*, por lo que, posiblemente algunos individuos reproductores que sí contribuyeron a la puesta pudieron quedar excluidos del estudio, al no ajustarse el fenotipo de su descendencia a las categorías de crecimiento establecidas. De esta manera, el manejo realizado puede ser también el desencadenante de la fuerte reducción observada en el número de hembras que toman parte en la puesta (Tabla 4c), la cual se contrapone a la tendencia general descrita para la especie donde el grupo de los machos suele ser frecuentemente quién sufre la mayores reducciones en el número de individuos activos debido al desarrollo de fuertes comportamientos dominantes (Barbaro y col., 1997; Bekkevold

y col., 2002; Brown y col., 2005b; Herlin y col. 2008).

Como era de esperar de acuerdo a las características generales de la dinámica reproductiva de la especie, los individuos reproductores que participan en la puesta presentan diferencias notables en el grado de contribución. Al igual que en el *Capítulo I*, las diferencias de contribución parecen adquirir mayor intensidad dentro del grupo de los machos (con excepción de la categoría D18S), estimándose varianzas entre dos y cinco veces superior a las estimadas en el grupo de las hembras, y su presencia posiblemente sea un reflejo de las conductas dominantes que frecuentemente suelen desarrollarse entre los machos, tal y como se mencionó anteriormente.

Nuevamente a consecuencia de la dinámica reproductiva adquirida por lo lotes, el valor de N_e se ve fuertemente comprometido, resultando especialmente afectado dentro de la categoría de *crecimiento-rápido* de ambos lotes, donde su valor representa solamente el 20-22% de los individuos reproductores que contribuyen a esta categoría (Tabla 4c); hecho que desafortunadamente conduce a un incremento de la endogamia más intenso dentro de la categoría fenotípica de mayor interés productivo (Tabla 4c). No obstante, a pesar de que los incrementos de endogamia deducidos tanto para la categoría de *crecimiento-rápido* ($\Delta F=6\%$) como *-lento* ($\Delta F=3-4\%$) distan considerablemente del nivel de endogamia generacional aceptado como regla general en programas de reproducción ($\Delta F=0,5\%$) (Sonesson y col., 2005), se sitúan próximos al valor estimado previamente por Porta y col. (2010) a partir de un lote reproductor de gran tamaño censal ($N=131$; $\Delta F=5,2\%$). Sin embargo, otros autores como Brown y col. (2005b) ($N=48-58$) o Chavenne y col. (2012) ($N=57$), utilizando lotes reproductores cuyo un tamaño censal se asemeja más a los lotes C y D, deducen incrementos de endogamia más reducidos (2,7-3,27%), lo que permite concluir que si bien la creación de nuevos lotes a partir de un único día de puesta predispone a importantes incrementos de endogamia, los niveles de endogamia deducidos a partir de diferentes lotes reproductores pueden presentar una cierta diversidad. No obstante, a diferencia de los trabajos anteriormente citados, en nuestro estudio los descendientes analizados sufrieron previamente un filtrado fenotípico, de modo que el cuello de botella generado por la estrategia de manejo desarrollada podría ser, en parte, el responsable de la endogamia estimada para ambas categorías de crecimiento. Por otro lado, si el carácter *crecimiento-rápido/crecimiento-lento* posee un componente genético, bajo un mismo ambiente, solamente aquellos individuos reproductores que posean el componente genético adecuado podrán contribuir a una u otra categoría, y por consiguiente, el N_e esperado dentro de cada categoría de crecimiento debería ser inferior al N_e deducido para el global de la descendencia (categoría de *crecimiento-rápido* + categoría de *crecimiento-lento*), obteniendo como resultado final valores de ΔF más elevados

dentro de las categorías.

Desde el punto de vista industrial, con la finalidad de mejorar la producción del cultivo, la creación de nuevos lotes reproductores a partir de los individuos incluidos en la categoría *crecimiento-rápido*, podría resultar de gran interés, no obstante, su manejo requiere medidas destinadas a minimizar el incremento de endogamia. Entre las medidas correctoras, una práctica utilizada con frecuencia para contrarrestar el efecto negativo de la endogamia, se basa en la introducción de individuos no emparentados procedentes habitualmente del medio natural. Esta práctica, además estar en contraposición con el desarrollo de una acuicultura sostenible independiente de las poblaciones naturales, afecta también a la respuesta de selección acumulada, introduciendo ruido (alteración) en el fondo genético establecido con el transcurso de varias generación de selección. Asimismo, las ventaja de introducir individuos totalmente ajenos al sistema de selección desarrollado puede llegar a ser limitada, puesto que en el caso de que la presión de selección aplicada en cultivo resulte superior a la presión existente en el medio de procedencia de los individuos introducidos como nuevos reproductores, probablemente en el siguiente proceso de selección, sus descendientes no serían seleccionados, por lo que quedarían relegados del estudio sin ejercer efecto alguno sobre el sistema de selección (Gjedrem y Thodesen, 2005). Dado que ya se ha fijado el pedigrí de los individuos, otro esquema de selección a considerar sería la implantación de un sistema de cruzamiento rotacional (Kincaid, 1977; Tave, 1999) basado en la combinación de individuos pertenecientes a la categoría de crecimiento rápido del lote C y D siempre y cuando genéticamente no presente relación de parentesco alguna. Por consiguiente, el intercambio de individuos entre lotes ayudaría a restablecer la variabilidad genética y neutralizar la endogamia acumulada. Del mismo modo, el establecimiento de un sistema de selección basado en cohortes, permitiría prevenir la endogamia durante un número limitado de generaciones (se estima que podría ser útil para prevenir la endogamia durante 4-6 generaciones), pero una vez que la endogamia entre en juego en el sistema reproductivo, su incremento sería muy rápido y marcaría el fin del programa de selección (Tave, 1999). De acuerdo a los resultados presentados en el *Capítulo I*, otra posible medida para hacer frente al incremento de endogamia podría reparar en un manejo reproductivo basado en *mini-lotes*. Sin embargo, al tratarse de un sistema de selección para crecimiento, en esta ocasión, la reunificación de eventos de puesta recolectados a diferentes tiempos, quedaría descartada al dificultar seriamente el proceso de selección fenotípica, ya que las diferencias de crecimiento entre los descendientes estarían fuertemente condicionadas por la existencia de individuos con diferentes tiempos de desarrollo (días transcurridos desde la eclosión de la larva), y por lo tanto, individuos de tamaño superior podría representar simplemente un tiempo de desarrollo mayor.

Por otra parte, para lograr desarrollar con éxito un programa de selección, además de realizar un minucioso control sobre los niveles de endogamia acumulada, otro factor que requiere de un control exhaustivo son las variaciones ambientales, que en el caso de las características asociadas a crecimiento juegan un papel importante, llevando a confundir su efecto con diferencias de crecimiento heredables. Dupont-Nivet y col. (2008) analizando las interacciones genotipo-ambiente en lubina (*Dicentrarchus labrax* L.) mediante la distribución de 253 familias de hermanos completos bajo ambientes diferentes representados por granjas de cultivo localizadas en 4 países diferentes (Israel, Italia, Francia y Portugal), observaron que las diferencias en la tasa de crecimiento se acentuaban a medida que se intensificaban las diferencias de temperatura de su cultivo. En este mismo sentido, Navarro y col. (2009) determinaron que en dorada el tiempo de adquisición de tamaño comercial (350-500g) se encuentra condicionado por la temperatura de desarrollo del cultivo. No obstante, la temperatura no es el único factor ambiental con capacidad para influenciar sobre las características de crecimiento, y un ejemplo de ello se encuentra en el estudio desarrollado por Wohfarth y col. (1983), donde por medio de la evaluación de la tasa de crecimiento de descendencias procedente de dos razas del carpa y del híbrido originado a partir del cruce entre ambas, cultivadas bajo tres ambientes diferentes denominados pobre (alta densidad de cultivo y una dieta empobrecida), rico (baja densidad de cultivo y dieta enriquecida) e intermedio, determinaron que la raza dominante para las características de crecimiento variaba en función del ambiente de cultivo. A su vez, otro importante factor a considerar es el momento en el cual los individuos son expuestos a variaciones ambientales, puesto que las interacciones genotipo-ambiente parece que tienden a intensificarse cuando los individuos se exponen a condiciones ambientales diferentes desde fases tempranas del desarrollo (por ejemplo, una vez transcurrido el proceso de fertilización) (Saillant y col., 2006; Dupont-Nivet y col., 2008).

En nuestro estudio, el mantenimiento de las diferentes categorías de crecimiento en tanques de cultivo separados, puede estar contribuyendo a las diferencias observadas entre las dos categorías debido a la existencia de posibles diferencias en las condiciones ambientales entre tanques. Además de la posible existencia de este “efecto tanque” (Speare y col., 1995), el hecho de no poder evaluar todo el conjunto de descendientes producidos (debido al sistema de selección aplicado), dificulta la capacidad para estimar la heredabilidad (h^2) del carácter de interés *crecimiento-rápido*. No obstante, en base a la correlación positiva que parece existir entre la diferenciación genética cuantitativa (Q_{ST}) y la diferenciación genética a nivel de marcadores moleculares neutrales (F_{ST}) (Merila y Cronkrak, 2001), la información genética proporcionada por los marcadores incluidos en OVIDORPPLEX podría ser interpretada como

una medida del grado de diferenciación entre genes codificantes para caracteres cuantitativos (QTL) (Roff, 2003; Leinonen y col., 2008), y en consecuencia, la diferenciación genética identificada en nuestros lotes entre las categorías de crecimiento ($F_{ST}=0,016-0,023$; $P<0,01$) podría indicar una diferenciación a nivel de QTL. Por otra parte, la detección de 10 reproductores que contribuyen de forma preferencial a la categoría de crecimiento rápido y 15 reproductores con una contribución preferente hacia la categoría de crecimiento lento, refuerzan la hipótesis de que el fenotipo de crecimiento mostrado por los individuos pertenecientes a estas categorías se encuentre bajo el control de un componente genético.

En base a la moderada heredabilidad ($h^2=0,2-0,4$) estimada para el carácter crecimiento rápido en peces (Gjedrem y Thodesen, 2005), la implantación de un sistema de evaluación de la descendencia, podría ayudar a incrementar la intensidad de la selección aplicada en programas reproductivos basados en sistemas de puestas masales (Sonesson, 2003). Esta evaluación de descendencias permitiría generar una clasificación donde los individuos reproductores se ordenan en función del valor reproductivo mostrado para el carácter de interés, facilitando así en futuros ciclos de reproducción, la selección de aquellos individuos reproductores que posean un valor reproductivo superior al resto. Además, la identificación de reproductores que de forma diferencial proporcionan descendientes de fenotipo opuesto, resulta un recurso de gran utilidad para la búsqueda de marcadores moleculares asociados significativamente a dichos fenotipos, cuya información genotípica posibilite el reemplazamiento del sistema de selección fenotípico por un sistema genotípico que efectúe una selección más precisa sobre el carácter.

Por otro lado, a consecuencia del comportamiento multifuncional de los genes, la selección hacia crecimiento rápido habitualmente conlleva de forma involuntaria, una selección indirecta sobre otros caracteres. Esta selección no intencionada puede potenciar la mejora generada, por ejemplo complementando la obtención de individuos de rápido crecimiento con comportamientos menos agresivos, una capacidad de resistencia a enfermedades más elevada o simplemente una tasa de supervivencia superior (Fjalestad y col., 1993; Gjedrem y Thodesen, 2005). No obstante, la selección indirecta también pueden actuar en contra de los intereses del programa de selección establecido, derivando en la obtención de fenotipos no deseados como puede ser el caso de una maduración sexual precoz de los individuos (Longalon y col., 1999), desencadenante de una ralentización en la tasa de crecimiento y una disminución de la calidad de la canal (Afonso y Roo, 2007), causando con ello un incremento en el tiempo de producción necesario para que los individuos alcancen la talla comercial, a la vez que se compromete la calidad y valor añadido del producto finalmente comercializado.

Ante esta situación, cabe destacar que si bien la maximización del crecimiento de los individuos representa un objetivo primordial para incrementar la rentabilidad de la producción y convertir el cultivo de la especie en una actividad económica más rentable, esta mejora debe supeditar una conservación de la calidad del producto final, pues no hay que olvidar que en último término el éxito de toda producción depende de la demanda de mercado existente, la cual a su vez se encuentra influenciada por la calidad del producto ofertado, resultando necesario hoy en día satisfacer las necesidades de un consumidor cada vez más exigente, que demanda productos de calidad a un precio extremadamente competitivo. En este sentido, para evaluar si las diferencias de crecimiento detectadas entre las categorías de crecimiento rápido y lento establecidas en los lotes C y D, influyen significativamente de forma positiva/negativamente sobre la calidad del filete, Mazzeo y col. (2013) analizaron la composición de la canal de acuerdo al procedimiento propuesto por la asociación oficial de química analítica (AOAC, *Association of Official Analytical Chemist*) y realizaron una prueba sensorial donde un grupo de 40 personas no entrenadas evaluaron mediante cata individuos de ambos lotes, procedentes de ambas categorías de crecimiento. Como resultado del estudio, los autores determinaron que en el lote C no existen diferencias significativas en la composición del filete entre las categorías de crecimiento, sin embargo en el lote D, donde las diferencias de peso entre las categorías de crecimiento rápido y lento son más intensas, se detectó que los individuos pertenecientes a la categoría de crecimiento rápido poseen una riqueza en grasa superior y un porcentaje de humedad inferior a los individuos de crecimiento lento. Estas diferencias de composición entre las categorías de crecimiento de lote D, derivaron en una preferencia sensorial hacia los individuos de crecimiento rápido, al resultar su carne más apreciada por ser más sabrosa y jugosa. Por lo tanto, inicialmente se rechaza que el empleo de individuos pertenecientes a esta categoría pueda provocar un impacto negativo sobre la calidad de la carne, al mismo tiempo que se demuestra como el consumidor es capaz de identificar sensorialmente variaciones en la composición del filete, hecho que ratifica la necesidad de asegurar que la calidad del producto obtenido no se ve afectada por el proceso de selección desarrollado.

4.5. Tablas

Tabla 4a. Criterios de selección empleados en el estudio.

	Crecimiento rápido		Crecimiento lento	
	Lote C	Lote D	Lote C	Lote D
1ª Selección	> 4,5 mm	> 3,5 mm	< 3,5 mm	< 2,5 mm
2ª Selección	> 41 g	> 31 g	< 15 g	< 12 g

1ª selección: selección realizada durante la fase larvaria (\approx día 80 de desarrollo) en base al tamaño corporal; 2ª selección: selección realizada durante la fase juvenil (\approx 12 meses) en base al peso corporal.

Tabla 4b. Parámetros de crecimiento.

	Lote C			Lote D		
	P	C18F	C18S	P	D18F	D18S
N	40	139	139	38	118	120
Peso medio (g)	1965,0	295,4	195,7	1981,5	323,8	184,1
SD	249,4	52,3	29,5	256,6	52,7	32,4
SE	34,9	4,4	2,5	41,6	4,8	2,9

P: parentales; *C18F*: categoría *crecimiento-rápido* a los 18 meses de edad en el lote C; *C18S*: categoría *crecimiento-lento* a los 18 meses de edad en el lote C; *D18F*: categoría *crecimiento-rápido* a los 18 meses de edad en el lote D; *D18S*: categoría *crecimiento-lento* a los 18 meses de edad en el lote D; *N*: número de individuos; *SD*: desviación estándar; *SE*: error estándar.

Tabla 4c. Parámetros de variabilidad genética y tamaño efectivo (N_e) estimados en el estudio a partir de 9 loci microsátelites (OVIDORPLEX), a la edad de 18 meses.

	Lote C			Lote D		
	P	C18F	C18S	P	D18F	D18S
N_a	14,6	12,2	14,0	14,0	12,2	11,8
H_e	0,867	0,801	0,849	0,872	0,832	0,842
H_o	0,883	0,867	0,865	0,858	0,893	0,859
F_{IS}	-0,028	-0,085	-0,019	0,011	-0,073	-0,022
p-valor	-0,0843 n.s	-0,0009***	-0,0407*	-0,73 n.s	-0,0009***	-0,0315*
Sim	99,9%	-	-	99,9%	-	-
Reprod.	29♀x11♂	11♀x8♂	19♀x10♂	28♀x10♂	14♀x8♂	14♀x9♂
$Ne_{(1)}$	31,2	18,5	26,2	29,4	20,3	21,9
$\Delta F_{(1)} (\%)$	1,6	2,7	1,9	1,7	2,46	2,28
$Ne_{(2)}$	-	8,14	16,12	-	8,45	11,29
$\Delta F_{(2)} (\%)$	-	6,1	3,8	-	5,9	4,4

P: parentales; *C18F*: categoría *crecimiento-rápido* a los 18 meses de edad en el lote C; *C18S*: categoría *crecimiento-lento* a los 18 meses de edad en el lote C; *D18F*: categoría *crecimiento-rápido* a los 18 meses de edad en el lote D; *D18S*: categoría *crecimiento-lento* a los 18 meses de edad en el lote D; N_a : número de alelos por locus; H_e : heterocigosidad esperada; H_o : heterocigosidad observada; F_{IS} : desviación de las frecuencias esperadas bajo el equilibrio de Hardy-Weinberg dentro de las categorías y su nivel de significación (p-valor): (*) $P < 0,05$; (***) $P < 0,001$; (n.s) no significativo; *Sim*: porcentaje con el cual se asignó la verdadera pareja parental después de 10000 ciclos de simulación bajo un intervalo de confianza del 95%; *Reprod.*: número de machos y hembras que contribuyeron a la descendencia; $Ne_{(1)}$: tamaño efectivo estimado mediante la fórmula $Ne^{\wedge}4(N_h \times N_m) / (N_h + N_m)$ (Falconer y Mackay, 2001); $Ne_{(2)}$: tamaño efectivo estimado mediante fórmula $Ne=4(N-2)/((Kd+Vd/Kd) + (Ks+Vs/Ks)-2$ (Chevassus, 1989); $\Delta F_{(1)}$: incremento de endogamia generacional deducido a partir de $Ne_{(1)}$; $\Delta F_{(2)}$: incremento de endogamia generacional deducido a partir de $Ne_{(2)}$.

4.6. Figuras

Figura 4a. PCR multiplex *OVIDORPLEX*. *A*: diseño de PCR multiplex (*OVIDORPLEX*) propuesto por el programa Multiplex Manager (Holleley y Geerts, 2009). Para cada *loci* microsatélite se indica su rango alélico (mediante los números situados a ambos extremos de cada *loci*) y el tipo fluoróforo empleado en su marcaje. *B*: Electroforetrograma obtenido para *OVIDORPLEX* en dos individuos. Los tres artefactos detectados en el estudio se indican mediante flechas.

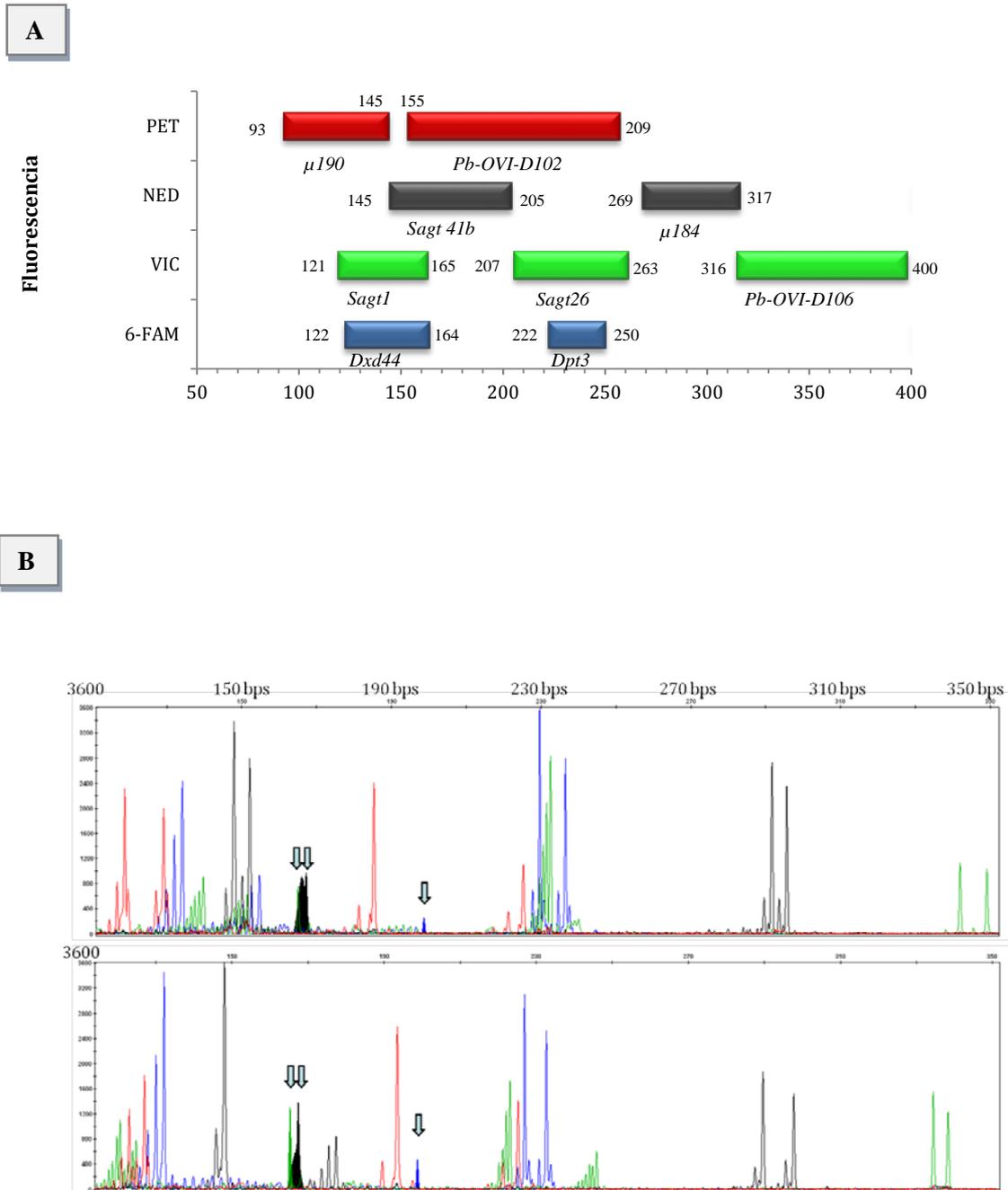
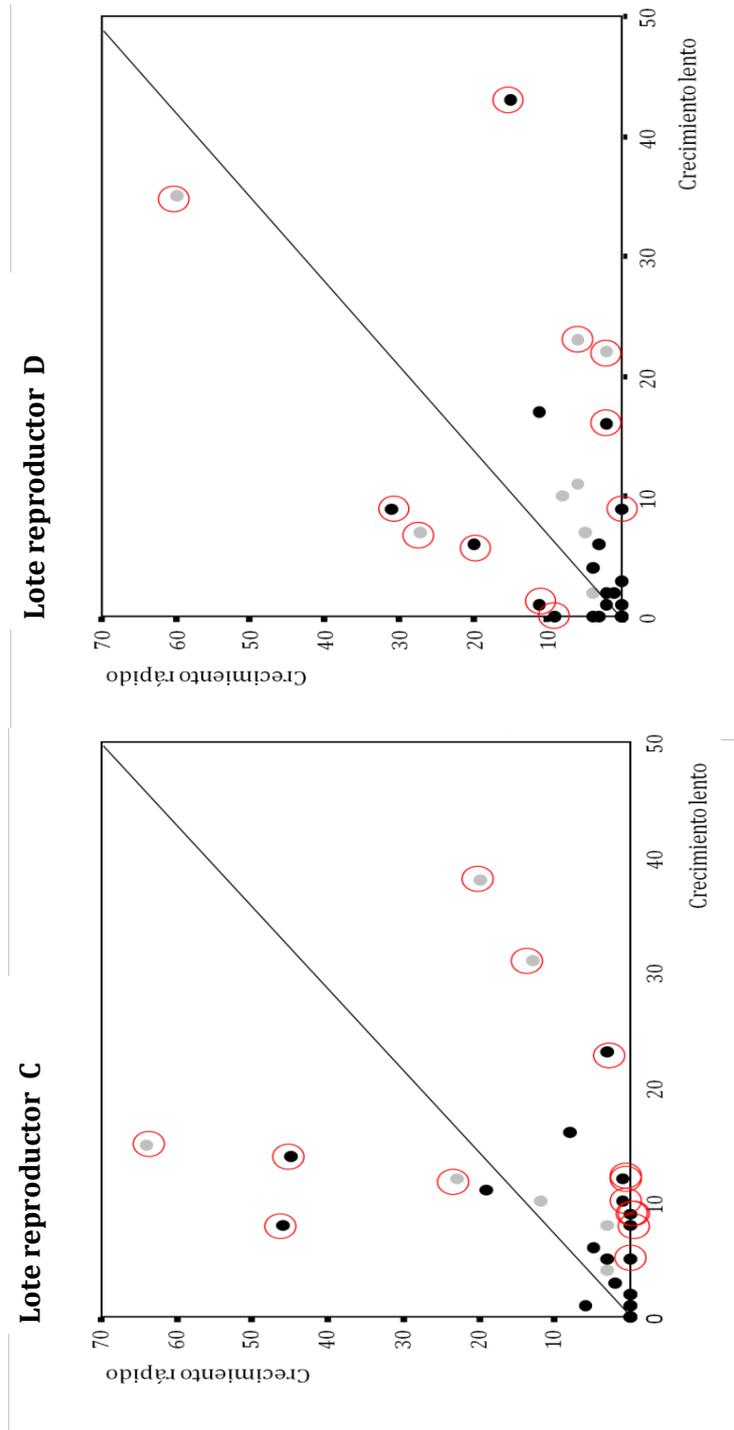


Figura 4b. Evaluación de la contribución parental a las diferentes categorías de crecimiento (*crecimiento-rápido y crecimiento-lento*). En el eje de coordenadas se representa la categoría de *crecimiento-lento*, y en el eje de abscisas se representa la categoría de *crecimiento-rápido*. Los individuos reproductores con contribución parental diferencial estadísticamente significativa ($p < 0,05$) se remarcan mediante círculos rojos.



5. Capítulo III

*Identificación y caracterización de genes
candidatos para calidad de huevo y
resistencia a enfermedades*

5.1. Introducción

A través de simulaciones teóricas se ha demostrado que la información derivada de marcadores genéticos puede mejorar notablemente la precisión del valor reproductivo estimado para cada individuo, incrementando consecuentemente, la respuesta genética a la selección aplicada (Meuwissen y col., 2001). Estos marcadores genéticos, frecuentemente se encuentran asociados a regiones del genoma (QTLs) con capacidad para ejercer un efecto significativo sobre el carácter de interés, destacando dentro de las especies acuícolas las regiones genómicas implicadas en el control del crecimiento (Reid y col., 2005; Sánchez-Molano y col., 2011), la resistencia a enfermedades (Ozaki y col., 2001; Moen y col., 2004; Cnaani y col., 2004), el estrés (Massault y col., 2011; Sauvage y col., 2012) o la determinación sexual (Loukovitis y col., 2011).

La asociación entre QTL y carácter, se debe entender como resultado del efecto ejercido por el gen o genes localizados en esa región cromosómica, cuya función se encuentra estrechamente correlacionada con la expresión del fenotipo de interés. Por consiguiente, el objetivo final de toda búsqueda de QTLs debería concluir con la identificación del gen o genes candidatos para el carácter localizados dentro del QTL. Sin embargo, los QTLs con frecuencia hacen referencia a regiones genómicas extensas (alcanzan tamaños comprendidos entre los 10-20 centiMorgans de longitud), con capacidad para albergar un gran número de genes (Kleeberger y Schwartz, 2005), resultando necesario incrementar la densidad de marcadores moleculares para obtener un poder de resolución más elevado, que permita realizar una localización mucho más precisa del QTL y limite el tamaño de la búsqueda del gene/s candidato/s a un tamaño de secuencia más manejable (entre los 0,5-5 centiMorgans de longitud) (Cañón, 2006).

Alternativamente a la identificación de regiones QTLs, habitualmente genes de función conocidas son directamente propuestos como genes candidatos para el carácter debido a la conexión existente entre el proceso fisiológico donde desarrollan su función y el carácter de interés. Bajo esta premisa, los estudios de expresión génica suponen una herramienta óptima para detectar de forma sencilla posibles genes candidatos (Lu y col., 2011; Quinn y col., 2011), puesto que si un gen determinado posee realmente influencia sobre el carácter, variaciones en su expresión se deberían correlacionar con variaciones a nivel fenotípico.

No obstante, el origen de las variaciones de expresión génica entre individuos reside en pequeñas modificaciones localizadas en la secuencia del gen candidato, de modo que una vez identificados los genes candidatos y validada su posible influencia sobre el carácter, la

búsqueda de estas posiciones polimórficas capaces de justificar las diferencias de expresión génica observadas entre individuos requiere de la caracterización de la secuencia genómica (completa o parcial) de dicho gen. Una vez identificadas las diferentes variantes alélicas, para poder evaluar estadísticamente la relación entre genotipo y variación fenotípica se precisa de una población de estudio adecuada, que posibilite la detección de una asociación genotipo-fenotipo significativa y de un sentido funcional al polimorfismo identificado, contribuyendo así, a un mayor conocimiento sobre el control genético del carácter. Sin embargo, se debe tener presente que en algunas ocasiones, la variación detectada puede no representar una asociación directa gen-carácter, pues el gen analizado no ejerce un efecto directo sobre el carácter, simplemente su proximidad en el cromosoma al gen con verdadera influencia, es quien le permite actuar como marcador indirecto (De-Santis y Jerry, 2007).

De acuerdo a estas directrices, el interés de este tercer capítulo se centra en la identificación y caracterización de genes candidatos para calidad de huevo y resistencia a enfermedades, ambos, caracteres de gran importancia para la industria acuícola.

La calidad de huevo, entendiendo como tal la capacidad que muestra un huevo para ser fertilizado y consecuentemente dar lugar a un embrión normal (Bobe y Labbe, 2010), es un carácter altamente variable en peces, de difícil control, influenciado por numerosos factores, pero que resulta de gran importancia para la acuicultura, pues el éxito del cultivo de una especie se encuentra en gran medida condicionado por la capacidad para obtener huevos de buena calidad (Brooks y col., 1997). Dentro de un cultivo, se clasifican como huevos de buena calidad a aquellos huevos que presentan tasas de mortalidad bajas durante los estadios comprendidos entre la fertilización y la primera alimentación larvaria (Bromage y col., 1992). Entre los diversos factores que condicionan a este carácter, se destaca el estado endocrino de la hembra durante la formación del oocito, la dieta de los individuos, las condiciones físico-químicas del medio externo donde tiene lugar la fertilización y posterior incubación de los huevos, así como el conjunto de biomoléculas transferidas al oocito durante el proceso de oogénesis, entre las cuales han despertado especial interés moléculas de ARNm de origen materno, denominadas ARNm maternos (Brooks y col., 1997). Durante las primeras fases del desarrollo embrionario, a pesar de que toda la maquinaria celular del cigoto permanece en un estado funcional, su genoma no se encuentra activo, y el control/regulación del desarrollo del embrión (desde el momento de la fecundación hasta el proceso de activación de la transcripción del genoma del cigoto) se atribuye a los ARNm maternos previamente depositados en el citoplasma del oocito (Traverso y col., 2012). Por lo tanto, estos ARNm maternos desarrollan un papel clave durante la embriogénesis, función que a su vez, les vincula estrechamente con el carácter calidad de huevo (Mommens y col., 2010).

Referente a la resistencia a enfermedades, el sistema inmune innato juega un papel de gran importancia en peces (organismos ectotérmicos) donde se ha demostrado que la exposición a bajas temperaturas puede desencadenar una supresión del sistema inmune adaptativo (Bly y Clem., 1992; Hayman y col., 1992; Le Morvan y col., 1998; Ellis, 2001), actuando el sistema inmune innato como sistema de defensa principal ofreciendo una respuesta no-específica, rápida y relativamente independiente de temperatura. Dentro de los factores solubles que participan en la inmunidad humoral innata, se incluye la transferrina (Tf) una proteína multifuncional que desarrolla un papel clave en el metabolismo del hierro. Concretamente Tf se encarga de transportar el hierro de forma segura (estado redox) a través del organismo desde el lugar de absorción hasta los centros de utilización o almacenamiento (Aisen, 1980; Gomme y McCann, 2005). Esta importante función transportadora, es a su vez el desencadenante de la asociación entre Tf y sistema inmune, pues la capacidad de esta proteína para regular los niveles de hierro libre en plasma le permite crear ambientes con una concentración de hierro reducida donde los microorganismos invasores ven fuertemente disminuida su capacidad patogénica (Suzumoto y col., 1977; Chen y col., 1999). Además de esta asociación, estudios desarrollados en la familia de los ciprínidos sugieren que en respuesta a un daño y/o infección los productos generados a partir de la escisión proteolítica de la Tf pueden actuar como alarma para el sistema inmune, provocando una activación de los macrófagos (Stafford y col., 2001; Matzinger, 2002; Stafford y Belosevic, 2003; Jurecka y col., 2009a). Por lo tanto, la estrecha relación existente entre Tf y el sistema inmune, convierte al gen *Tf* en un candidato de gran interés para marcar resistencia a enfermedades.

En base a lo expuesto, para este tercer capítulo se plantean dos objetivos bien diferenciados. El primero, se dirige a la búsqueda de genes candidatos asociados a calidad de huevo en dorada, a través de la identificación de genes que manifiesten un patrón de expresión maternal durante el desarrollo embrionario temprano. En total, se evaluaron 16 genes (*gapdh-1*, *gapdh-2*, *catepsina-D*, *catepsina-L*, *catepsina-S*, *catepsina-Z*, *tubulina-β*, *ferritina-M*, *ciclina-A2*, *ppar-α*, *ppar-β*, *ppar-γ*, *erk-1*, *jnk-1*, *p38-α* y *p38-δ*) seleccionados en base a su participación en la regulación de procesos celulares básicos, indispensables para el correcto desarrollo del embrión (Tabla 5a), y en algunos casos, debido a los indicios existentes en otras especies sobre su posible viculación con calidad de huevo (Howe y col., 1995; Braissant y Wahli, 1998; Aegerter y col., 2005; Krens y col., 2006; Transverso y col., 2012).

Como segundo objetivo, se propone la caracterización del gen de transferrina (*Tf*) de dorada, seleccionado principalmente por su posible papel como gen candidato para la mejora de la susceptibilidad a patógenos, y especialmente importante por su escasa documentación en la especie. La caracterización del gen *Tf* se realiza atendiendo principalmente a su estructura

genómica, y a su patrón expresión constitutiva en diferentes tejidos celulares pertenecientes a las fases de desarrollo juvenil y adulta.

5.2. Materiales y Métodos

5.2.1. Búsqueda de genes candidatos para calidad de huevo

5.2.1.1. Muestras biológicas

Para los 16 genes (Tabla 5a), se evaluó la expresión génica en un total de siete estadios de desarrollo embrionario temprano [estadio de 2-4 células, estadio de 16 células, estadio hacia mórula (mórula temprana), estadio de mórula, estadio avanzado de blástula, estadio correspondiente al comienzo de la gastrulación (gástrula temprana) y estadio medio de epibolia] (Figura 5a), procedentes de ocho lotes de reproductores mantenidos durante el periodo de puesta transcurrido entre Diciembre de 2009 y Marzo de 2010, en las instalaciones del Instituto de Biología Marina, Biotecnología y Acuicultura del *Hellenic Center for Marine Research (HCMR)* en Gournes Padiados (isla de Creta, Grecia). La conservación de las muestras se realizó por inmersión en nitrógeno líquido, seguida de una preservación a -80°C hasta su utilización.

5.2.1.2. Extracción de ARN y síntesis de ADNc

La extracción de ARN se realizó combinando el protocolo estándar de Trizol, con el sistema de tubos Lysing Matrix D (MP), que favorecen el lisado de los tejidos por rotura mecánica mediante unas pequeñas esferas de cerámica (1,4mm \varnothing) presentes en los tubos. El tejido se degradó incorporando a cada tubo Lysing Matrix, 500 μL de Trizol, 5 μL de β -mercaptoetanol y 50mg de muestra de huevos embrionados, efectuándose el proceso de lisado y homogenización de las muestras en un TissueLyser II (Qiagen), utilizando un programa basado en tres ciclos de 40s a 20Hz. La separación de fases se realizó por medio de una centrifugación a 12.000g durante cinco minutos a 4°C , tras la cual, la fase superior se transfirió a un nuevo tubo Eppendorf de 1,5mL, donde se combinó con 300 μL de cloroformo y se dejó reposar a temperatura ambiente durante cinco minutos. Seguidamente, se procedió a una segunda centrifugación a 12.000g durante cinco minutos a 4°C , transfiriendo nuevamente la fase superior a un tubo Eppendorf con 500 μL de etanol 100% frío y se incubó a -20°C durante toda la noche. La precipitación de ARN se realizó mediante un proceso de centrifugación a 12.000g durante 60 minutos a 4°C . Tras un proceso de lavado del ARN con etanol al 75%, el ARN precipitado se disolvió en 40 μL de agua libre de ARNasas. Las concentraciones de ARN se estimaron utilizando un espectrofotómetro NanoDrop ND-100 (NanoDrop Technologies Inc., Wilmington, DE, USA) por medio de la medición de absorbancia a 260nm, determinándose su

calidad en base a la relación de absorbancias 260/280nm y su visualización en un gel de agarosa al 1% con bromuro de etidio.

Para cada una de las muestras la síntesis del ADNc (ADN complementario) total se realizó por retro-transcripción o transcripción inversa (RT-PCR) siguiendo el protocolo estándar del kit Superscript II First Strand Synthesis (Invitrogen/VWR, Tromsø, Noruega), utilizando iniciadores de tipo *random hexamers* (iniciadores de hexanucleótidos al azar) y 1µg de ARN total, para un volumen final de 19µL. El programa de amplificación utilizado se basó en 10 minutos a 25°C, 50 minutos a 42°C y 15 minutos a 72°C. En todas la muestras se incluyeron controles negativos para detectar la posible contaminación de ADN genómico. El ADNc obtenido se almacenó a -20°C hasta su uso, en una dilución 1:50 con agua Merck desionizada (Whitehouse Station, NJ, USA).

5.2.1.3. Búsqueda y análisis de secuencias

A partir de la bibliografía se obtuvieron los iniciadores para siete (*gapdh-1*, *gapdh-2*, *ppar-α*, *ppar-β*, *ppar-γ*, *cathepsina-D*, *cathepsina-L*) de los 16 genes seleccionados en el estudio, así como para los genes *L13a* y *S18* utilizados como control (Tabla 5a). Para los nueve genes restantes (*cathepsina-S*, *cathepsina-Z*, *tubulina-β*, *ferritina-M*, *ciclina-A2*, *erk-1*, *jnk-1*, *p38-α* y *p38-δ*) se carecía de información en dorada, por lo que fue necesario realizar un primer paso de identificación de un fragmento de su secuencia codificante sobre el cual posteriormente diseñar iniciadores óptimos para PCR a tiempo real (qRT-PCR). Para cada uno de estos genes, el proceso de identificación se realizó a partir de la búsqueda de secuencias ESTs de dorada en la base de datos pública (no redundante) del NCBI y una base de datos privada específica de la especie (obtenida en estudios previos y propiedad del HCMR), utilizando como molde secuencias homólogas descritas en especies cercanas. Cada una de las posibles ESTs específicas identificadas en dorada, se utilizó como secuencia molde en una nueva búsqueda de secuencias nucleotídicas (BlastN, busca una secuencia de nucleótidos en una base de datos de secuencias nucleotídicas) y secuencias de proteínas homólogas (BlastX, busca una secuencia de nucleótidos traducida en la base de datos de proteínas), para confirmar que el fragmento de secuencia identificado presenta una similitud elevada con las secuencias descritas para estos genes en otras especies. Seguidamente, mediante el programa ORF finger (<http://www.ncbi.nlm.nih.gov>), se aisló para cada secuencia un fragmento con pauta de lectura abierta (*ORFs*, *Open Reading Fragments*), y se verificó nuevamente la correcta identificación de la secuencia por medio de un BlastP realizado contra la base de datos de proteínas (no redundante) del NCBI, observándose una vez más, la existencia de una similitud elevada con

las correspondientes secuencias de proteínas caracterizadas en otras especies.

5.2.1.4. Diseño de iniciadores

El diseño de iniciadores se llevó a cabo mediante el programa FastPCR (Kalendar y col., 2009), diseñándose una única pareja para cada uno de los nueve nuevos genes identificados en el estudio. Posteriormente, la idoneidad teórica de estos iniciadores para estudios de qRT-PCR, se evaluó por medio del programa de diseño de iniciadores y sondas específicos para qRT-PCR, Beacon designer 7.5 (<http://www.premier-biosoft.com>). El resto de los iniciadores empleados en el estudio se tomaron directamente de las referencias bibliográficas (Tabla 5b). La especificidad e idoneidad de las parejas de iniciadores utilizados, se comprobó mediante PCR convencional realizada a partir ADN genómico y ADNc procedente de los diferentes estadios de desarrollo evaluados en el estudio, así como de tejido de bazo y corazón, visualizándose el producto de PCR obtenido en un gel de agarosa al 2% con bromuro de etidio.

5.2.1.5. Cuantificación de la expresión génica

Las reacciones de qRT-PCR se realizaron en un termociclador DNA Engine Opticon 2 Continuous Fluorescence Detector (MJ Research, MA, USA). Las condiciones de reacción empleadas consistieron de 5 μ L de ADNc de una dilución 1:50, 12,5 μ L de SYBR Green PCR Master Mix (Applied Biosystems, Warrington, UK) y 7,5 μ L de agua Merck desionizada. Cada una de las 56 muestras incluidas en el estudio [ocho lotes (réplicas biológicas) y siete estadios de desarrollo diferentes por lote] se analizó por triplicado (réplicas técnicas). Como medida de control, en el análisis de cada gen, se incluyeron controles negativos carentes de ADNc. El programa de qRT-PCR empleado se basó en una fase inicial de desnaturalización a 95°C durante 10 minutos, seguida de 40 ciclos de amplificación consistentes en 15 segundos a 95°C y un minuto a 57°C, realizando después de cada ciclo una lectura de la señal de fluorescencia emitida. Nuevamente se comprobó la especificidad de los iniciadores y la ausencia de dímeros mediante la aplicación de una fase final con curvas de desnaturalización o disociación (gradiente de temperatura comprendido entre 30 y 99°C), basándose en la hipótesis de que diferentes amplicones tenderán a mostrar diferentes temperaturas de fusión (T_m). Además, como medida adicional para evitar que aquellas amplificaciones inespecíficas y formación de dímeros de iniciadores con iguales T_m pasen inadvertidos, todas las muestras se visualizaron posteriormente en un gel de agarosa al 2% con bromuro de etidio.

Los datos de fluorescencia brutos obtenidos se analizaron mediante el programa Miner (Zaho y Fernald, 2005), cuantificando directamente los resultados de qRT-PCR a partir de la

cinética de cada reacción individual de la PCR realizada. Mediante este programa se determinó el punto de inicio (Ct, ciclo umbral) y final de la fase de amplificación exponencial de la PCR, estimándose además la eficiencia de amplificación de cada pareja de iniciadores a través de un algoritmo de regresión no lineal, sin necesidad de recurrir a la curva de expresión estándar.

5.2.1.6. Análisis estadístico

Los niveles de expresión de los genes analizados se evaluaron y normalizaron frente a los genes de referencia *L13a* (gen codificante para la proteína ribosómica L13a) y *S18* (gen codificante para la proteína ribosómica S18) (Tablas 5a y b). Por medio del programa GeNorm se evaluó la estabilidad de expresión de los dos genes de referencia y se calculó el correspondiente factor de normalización (Vandesompele y col., 2002). Para cuantificar la estabilidad de expresión se empleó un algoritmo fundamentado en la teoría de que la expresión de dos genes de referencia debe guardar el mismo índice de relación en todas las muestras, independientemente de las condiciones ambientales y/o el tipo celular analizado; por lo tanto, desviaciones en el índice de relación de la expresión de los genes control, implicaría que al menos uno de los genes no se estaría expresando de forma constante a lo largo de las diferentes muestras evaluadas. De acuerdo con este planteamiento el programa GeNorm determina la variación de expresión de cada gen control, comparándolo con los valores de expresión del segundo gen control. Si la expresión de los genes de referencia es estable, se deduce un factor de normalización para cada muestra empleando la media geométrica de los valores relativos estimados para los genes control, el cual posteriormente se utiliza para normalizar los niveles de expresión de los genes objeto de estudio.

La cuantificación de la expresión relativa de cada uno de los genes analizados para cada uno de los estadios de desarrollo embrionario evaluados, se realizó por medio de la comparación de los valores Ct obtenidos utilizando el método ΔCt , normalizando los valores de expresión mediante su división entre el factor de normalización previamente estimado con el programa GeNorm. Los niveles de expresión relativa deducidos, se confirmaron mediante un segundo programa de cuantificación de la expresión génica [REST, Relative Expression Software Tool (Pfaffl y col., 2002)] que emplea un método de análisis similar al ΔCt corregido para las eficiencias de los diferentes genes.

La hipótesis nula de igualdad de medias entre los valores de expresión relativa obtenidos para cada gen en los diferentes estadios de desarrollo analizados se contrastó mediante un análisis múltiple de la varianza (ANOVA) de un factor, utilizando el programa estadístico SPSS v.15. El análisis estadístico se completó aplicando la prueba de comparaciones múltiples a posteriori o *post-hoc* de Turkey, que permite una vez rechaza la hipótesis nula general de

ANOVA de igualdad de medias, averiguar qué medias en concreto difieren entre sí.

5.2.2. Caracterización del gen codificante para transferrina (*Tf*) candidato para resistencia a enfermedades

5.2.2.1. Muestras biológicas

Para la caracterización del gen *Tf* se emplearon siete individuos: tres juveniles ($30\pm 0,5\text{g}$) y cuatro hembras adultas ($3,5\pm 0,25\text{kg}$), criados y mantenidos en las instalaciones del IRTA (Institut de Recerca i Tecnologia Agroalimentaries) en Sant Carles de la Ràpita (Tarragona, España).

Todos los individuos se sacrificaron de acuerdo al protocolo de eutanasia basado en 100mg L^{-1} de MS222 (Sandoz), diseccionando y recolectando muestras de siete diferentes tejidos: hígado, riñón, bazo, cerebro, estómago, corazón y ovario (solo en las hembras adultas), en RNAlater® (Ambion) para preservar la integridad del ARN, conservándose a -20°C hasta su utilización. Además, durante el proceso de disección, se recolectó en etanol al 96% un pequeño fragmento de aleta dorsal, que se almacenó a 4°C hasta su utilización para el proceso de extracción de ADN genómico.

5.2.2.2. Extracción de ADN, ARN y síntesis del ADNc

La extracción de ADN genómico, se realizó a partir de un pequeño fragmento de aleta dorsal procedente de individuos adultos, siguiendo el protocolo estándar del kit comercial DNeasy Blood & Tissue (Qiagen).

La extracción de ARN total se realizó a partir de muestras de tejidos preservadas en RNAlater® (Ambion) utilizando el protocolo estándar propuesto para el kit comercial High Pure RNA Tissue (Roche), un sistema rápido para aislar ARN total, donde el proceso de rotura y homogenización del tejido se realiza en presencia de un fuerte tampón desnaturizante que contiene clorhidrato de guanidina que inactiva las enzimas ARNasas de forma inmediata, facilitando la obtención de un ARN de elevada integridad. Tanto el aislamiento como la purificación del ARN total se realizaron a través de un sistema de diferentes pasos de lavado con columna de alta capacidad de unión al ARN, aislando el ARN en un único paso de elución. Con ayuda de un PicoDrop, se determinó la concentración del ARN extraído a partir de mediciones de absorbancia a 260nm , valorándose su calidad por medio de la relación de absorbancia $260/280\text{nm}$ y comprobaciones en gel de agarosa al 1% con bromuro de etidio.

Las síntesis de ADNc se llevó a cabo mediante RT-PCR utilizando el kit comercial

Transcriptor FirstStrand cDNA synthesis (ROCHE) para cada una de las muestras de acuerdo a las condiciones estándar descritas en su protocolo, utilizando 1µg de ARN total y empleando iniciadores de tipo anchored-oligo (dT)₁₈. El programa de RT-PCR consistió en 30 minutos a 55°C, seguidos de una fase de inactivación de 5 minutos a 85°C. La posible contaminación por ADN genómico se descartó mediante la introducción de controles negativos entre las reacciones de RT-PCR. Las muestras de ADNc total obtenidas, se almacenaron a -20°C hasta su utilización.

5.2.2.3. Búsqueda y análisis de secuencias

Al comienzo del estudio la única información disponible en la base de datos pública del NCBI referente al gen *Tf* de dorada, consistió en dos ESTs (*Expressed Sequence Tags*) de 291pb cada una (GenBank ID: CX734969 y CX734865). El alineamiento de estas dos ESTs, utilizando el programa BioEdit v.7.0.5.3 (Hall, 1999) determinó que ambas secuencias correspondían al mismo fragmento codificante del gen *Tf*, presentando entre sí una similitud del 98,97%. Empleando estas EST como molde, se identificaron mediante un BlastN dos secuencias de ARNm con una similitud superior al 90% con la secuencia molde, correspondientes a la secuencia codificante del gen *Tf* del besugo japonés (*Acanthopagrus shlegelii*) (GenBank ID: AY36505) y la dorada japonesa (*Pagrus major*) (GenBank ID: AY335444), ambas especies del grupo espáridos al que pertenece la dorada. Estas secuencias de ARNm contienen la secuencia exónica completa del gen *Tf*, junto con fragmentos parciales de las regiones 5' y 3' no transcritas (UTRs, *Untranslated regions*). Para determinar el fragmento de secuencia exacto que corresponde a cada uno de los exones que constituyen el gen *Tf*, mediante el programa BioEdit v.7.0.5.3 (Hall, 1999) se realizó un alineamiento múltiple entre estas dos secuencias y la secuencia de ADN genómico (incluye regiones codificantes y no-codificantes) del gen *Tf* perteneciente a tilapia del Nilo (*Oreochromis niloticus*) (GenBank ID: DQ272465) (Rengmark y Frode, 2007), seleccionada por ser la especie evolutivamente más cercada a la dorada donde la secuencia de ADN genómico completa del gen *Tf* había sido caracterizada. A partir del alineamiento de estas tres secuencias se estimó la distribución de los exones en las correspondientes secuencias de ARNm, definiéndose además todas aquellas regiones que presentaban un alto grado de conservación entre las especies.

5.2.2.4. Diseño de iniciadores, amplificación y secuenciación

El diseño de iniciadores se realizó mediante el programa FastPCR (Kalendar y col., 2009). Concretamente, para la amplificación y secuenciación del ADN genómico de gen *Tf* y su

correspondiente secuencia de ADNc, se diseñaron un total de 20 iniciadores organizados en 12 parejas (Tabla 5c).

Para el análisis del patrón de expresión del gen *Tf* se diseñaron los iniciadores específicos de dorada SaTf-FW/SaTf-RV (Tabla 5c). En esta ocasión, el iniciador SaTf-FW se diseñó a partir de la de secuencia de unión entre dos exones (E-XIV y E-XV) para que en caso de existir contaminación de ADN genómico su amplificación se viera dificultada, a la vez que su detección se viera facilitada mediante la visualización de los productos de PCR en gel de agarosa, observándose en estos casos una banda de amplificación con un tamaño superior al esperado, evitando de esta manera la obtención de una falsa señal de expresión.

Las condiciones de PCR empleadas para el conjunto de reacciones desarrollados para el aislamiento de la secuencia genómica del gen *Tf*, consistieron en: 0,25µL (5U/µL) del enzima TaKaRa Taq™HS (Takara Biotechnoloy (Dalian) CO., LTD), 5µL (10X) del tampón de reacción, 4µL de mezcla de dNTPs (concentración de cada dNTP de 2,5mM), 0,4µL de iniciador (10pmol/µL) y aproximadamente 100ng de ADN genómico, para un volumen final de 50µL. Todas las reacciones de PCR se realizaron en un termociclador XP-Cycler (BIOER). El programa de PCR empleado consistió en 35 ciclos de amplificación consistentes en 30s a 94°C, 30s a la temperatura de hibridación específica de cada pareja de iniciadores (Tabla 5c) y un paso final de extensión de 1min a 72°C. Posteriormente todos los productos de PCR se corrieron en gel de agarosa al 2%, se purificaron mediante el kit comercial Ilustra™ GFX™ PCR DNA and Gel Band Purification (GE HEALTHCARE), y se comprobaron nuevamente en un gel de agarosa al 2%. El producto de ADN purificado obtenido se envió a MACROGEN Inc. (<http://dna.macrogen.com/eng/>) para su secuenciación. Finalmente la secuencia genómica del gen *Tf* se obtuvo por medio del alineamiento y ensamblaje de los diferentes productos de PCR obtenidos.

La caracterización del ADNc específico del gen *Tf* se realizó a partir de ADNc obtenido de hígado de individuo adulto, bajo las condiciones de PCR descritas anteriormente, utilizando una única pareja de iniciadores (*SaTf-F1* y *SaTf-R17*, Tabla 5c) para su amplificación y 4 parejas (*SaTf-F1/ SaTf-R5.2*, *SaTf-F4/ SaTf-R8*, *SaTf-F5/ SaTf-R13* y *SaTf-F11/R17*, Tabla 5c) para su secuenciación (Macrogen Inc.), generando productos de PCR solapantes, de aproximadamente 1000pb de longitud. Una vez obtenida la secuencia nucleotídica del ADNc, su alineamiento con la secuencia de ADN genómico previamente aislada, permitió determina la estructura génica (número, longitud y posición de las regiones intrónicas y exónicas) del gen *Tf*. La estructura de la proteína codificada, se dedujo mediante el programa SMART (Simple Modular Architecture Research Tool) (<http://www.smart.embl-heidelberg.de/>) proporcionado por el EMBL (European Molecular Biology Laboratory) como herramienta para

la identificación de dominios en proteínas.

5.2.2.5. Análisis del patrón de expresión génica

A partir las muestras de ADNc, se determinó la presencia de transcritos del gen *Tf* para los diferentes tejidos analizados visualizando los productos de PCR obtenidos por medio de PCR convencional en un gel de agarosa al 1,5% con bromuro de etidio, utilizando para ello la pareja de iniciadores específicos SaTf-FW/SaTf-RV (Tabla 5c) junto con las condiciones y el programa de PCR descritos en el apartado anterior. Como gen de referencia, se seleccionó el gen codificante para β -actina (Tiago y col., 2008) (Tabla 5c).

5.2.2.6. Análisis filogenético

La secuencia codificante completa del gen *Tf* (ADNc) de dorada se comparó con las secuencias codificantes descritas para este gen en 14 especies de peces, un anfibio y tres especies de mamíferos, todas ellas disponibles en la base de datos del NCBI. Inicialmente, mediante el programa BioEdit v.7.0.5.3 (Hall, 1999), se dedujo la correspondiente secuencia de aminoácidos para cada una de la secuencias codificantes analizadas. A través del programa Muscle v.1.5 (Edgar, 2004), utilizando los parámetros de configuración predeterminada, se realizó un alineamiento múltiple de las secuencias de aminoácidos obtenidas para las diferente especies. A partir de dicho alineamiento y dado que a lo largo de la secuencia del gen existen tanto zonas conservadas como no conservadas, se utilizó un modelo de búsqueda heurística implementado en el programa Trimal (opción automated1) (Capella-Gutiérrez y col., 2009), para proceder a la identificación de aquellas regiones más informativas, descartándose las que presentan tasas de variación elevadas que podrían desvirtuar el resultado del análisis filogenético. El alineamiento final se empleo como base para los sucesivos análisis mediante cálculos de máxima probabilidad, incluidos en el programa ProtTest (Abascal y col., 2005), se determinó el modelo de evolución proteico más adecuado para la variación presente en las secuencias de aminoácidos. De igual modo, se determinaron los valores medios para diferentes parámetros evolutivos: posiciones invariables (+I), tasa de variación entre posiciones (+G) y frecuencias aminoacídicas (+F). Utilizando el programa PhyML (Guindon y Gascuel, 2003) se construyó el árbol filogenético a partir de un modelo evolutivo en base a la matriz WAG, ajustando la longitud de las ramas mediante procedimientos de máxima probabilidad, y evaluando el soporte de la topología resultante en un entorno estadístico sólido por medio un test *bootstrap* basado en 10.000 replicas. Finalmente, el árbol filogenético deducido se visualizó y editó a través del visor gráfico Figtree v1.3.1 (<http://tree.bio.ed.ac.uk>

/software/figtree/).

5.3. Resultados

5.3.1 Identificación y análisis de expresión de genes candidatos para calidad de huevo

La búsqueda realizada en las base de datos para ESTs del NCBI y del HCMR permitió identificar, por primera vez en dorada, parte de la secuencia nucleotídica perteneciente a los genes *catepsina-S*, *catepsina-Z*, *tubulina- β* , *ferritina-M*, *erk-1*, *ciclina-A2*, *jnk-1*, *p38- α* y *p38- δ* (Tabla 5d).

Los genes de referencia *L13a* y *18S*, empleados en la normalización de los datos de expresión, mostraron una amplificación específica, detectándose en el gel de agarosa una única banda de intensidad elevada para el producto de PCR obtenido por medio de PCR convencional a partir del ADN genómico y ADNc procedente de diferentes estadios de desarrollo temprano, así como de los tejidos de bazo y corazón, utilizados como control. Posteriormente, las medidas de fluorescencia obtenidas por qRT-PCR para cada gen control en un total 168 reacciones (8 lotes x 7 estadios de desarrollo embrionario x 3 réplicas técnicas por muestra), determinaron que la expresión de ambos genes se mantenía estable a lo largo de los diferentes estadios de desarrollo embrionario, por lo que, sus valores de expresión se utilizaron para generar un factor de normalización dirigido a corregir dentro de los genes de interés las variaciones de expresión causadas por posibles imprecisiones cometidas en la cantidad de ADNc depositado dentro de cada reacción.

Por medio de la visualización en gel de agarosa del fragmento de ADN amplificado por PCR convencional a partir de muestras de ADNc procedente de estadios de desarrollo embrionario temprano, bazo y corazón, se observó para los genes *gapdh-2*, *ppar- γ* , y *catepsina-S*, en los estadios de desarrollo embrionario temprano no se detectó amplificación alguna, mientras que en las muestras procedentes de bazo y corazón, se detecta una banda de amplificación específica e intensa (Figura 5b). En el caso del gen *ppar- α* , tampoco se detecta expresión durante el desarrollo temprano, y la señal registrada en bazo y corazón es muy débil (Figuras 5b).

En los 12 genes restantes (*gapdh-1*, *catepsina-D*, *catepsina-L*, *catepsina-Z*, *tubulina- β* , *ferritina-M*, *ciclina-A2*, *ppar- β* , *erk-1*, *jnk-1*, *p38- α* y *p38- δ*) se observó para cada uno de los estadios de desarrollo evaluado, una amplificación clara y específica, así como también una aparente ausencia de dímeros de iniciadores (datos no mostrados), lo cual permitió aceptar a sus respectivos iniciadores como válidos para cuantificar la expresión génica mediante qRT-PCR. Finalmente, la cuantificación de la expresión reveló para estos 12 genes, con excepción

del gen *p38-δ*, un patrón de expresión común caracterizado por una expresión inicial casi constante, con pequeñas variaciones de expresión durante la fase de desarrollo comprendida entre los estadios de 2-4 células y blástula, seguida de una drástica disminución de los niveles de expresión una vez sobrepasado el estadio de desarrollo de blástula avanzada (transcurridas aproximadamente unas seis horas y media desde el momento de la fecundación) (Figuras 5c). La excepción, el gen *p38-δ*, presentó una expresión reducida a nivel general, siendo especialmente limitada durante los estadios de 2-4 células, mórula, gástrula temprana y media-epibolia (Figuras 5c).

5.3.2 Caracterización del gen *Tf*

A partir del alineamiento y ensamblaje de los diferentes fragmentos de PCR amplificados (Tabla 5c), se obtuvo una secuencia de ADN genómico (GenBank ID: JF309047) y ADNc (GenBank ID: JF309047) con una longitud de 7.855pb y 2.320pb, respectivamente (Figuras 5d). Por medio del alineamiento de estas dos secuencias, se estableció para el gen *Tf* una estructura génica compuesta por 17 exones separados por 16 intrones (Tabla 5e, Figura 5d), y se determinó que la secuencia de ADN genómico aislada incluye, además de las regiones exónicas e intrónicas, fragmentos parciales de las regiones 5' (18pb) y 3' (274pb) UTR. Todas las uniones exón-intrón concordaron con la regla de corte y empalme GT-AG.

La comparación de la estructura deducida para el gen *Tf* de dorada con las estructuras previamente descritas para pez gato (*Ictalurus punctatus*), tilapia del Nilo (*Oreochromis niloticus*), salmón plateado (*Oncorhynchus kisutch*) y medaka (*Oryzias latipes*), reveló una idéntica organización del gen y una elevada conservación en la longitud de los exones, detectándose el mismo tamaño para 6 (EI, EII, EVI, EX, EXI y EXIV) de los 17 exones (Tabla 5e). El mayor grado de similitud se observó entre dorada y tilapia del Nilo, con una longitud de secuencia idéntica en 15 exones. En contraposición, las regiones intrónicas presentaron rangos de variación muy amplios entre todas las especies comparadas (Tabla 5e).

A partir de la secuencia codificante del gen *Tf* en dorada, se dedujo una proteína compuesta por 691 aminoácidos (Tabla 5e, Figura 5e), distribuidos en un péptido señal (localizado entre los aminoácidos cuatro y 14) y dos lóbulos: el lóbulo-N (localizado entre los aminoácidos 24 y 338) y el lóbulo-C (localizado entre los aminoácidos 340 y 681). Dentro de la estructura de proteína, se identificó una conservación máxima en los residuos de unión al hierro (Asp-73, Tyr-103, Tyr 200, His-256, Asp-394, Tyr-429, Tyr-524 and Arg-592), y a aniones (Thr-128, Lys-132, Thr-454 and Arg-458) descritos por Lambert y col. (2005) (Figura 5e).

El análisis filogenético desarrollado, permitió situar la proteína Tf de dorada dentro de un contexto evolutivo. Las distancias genéticas, estimadas en función del número de sustituciones de aminoácidos, mostraron valores moderados-bajos, y la robustez de la topología deducida se apoyó con fuertes valores de bootstrap (Figuras 5f). Dentro del grupo de los peces, tal y como era de esperar, la dorada se incluyó dentro del subgrupo de los espáridos (pargo de cabeza negra y pargo japonés) con altos valores de bootstrap, mientras tilapia del Nilo y medaka constituyeron un nuevo subgrupo, los salmónidos y los ciprínidos formaron sus respectivos grupos, y el bacalao, pez gato y el falso halibut de Japón, no se integraron dentro de ningún grupo. Por su parte, las especies de vertebrados terrestres formaron un grupo principal compuesto por la única especie de anfibios empleada en el análisis y el subgrupo de los mamíferos.

El análisis de expresión mostró que el gen *Tf* posee una expresión constitutiva especialmente intensa en el tejido hepático, con niveles inferiores en los tejidos correspondientes al riñón, bazo, estómago, cerebro y ovario (analizado solamente en hembras adultas) y ausente en el corazón (Figura 5g). Los patrones de expresión obtenidos en individuos juveniles y adultos presentaron diferencias notables en bazo y estómago (Figura 5g); en el caso del bazo las muestras de tejido de individuos adultos mostraron una expresión más elevada que las recolectadas a partir de individuos juveniles, mientras en el caso del estómago, la expresión más intensa se observó a partir de tejido juvenil (Figura 5g).

5.4 .Discusión

5.4.1 Identificación de genes candidatos para calidad de huevo

La hipótesis de la posible participación de moléculas de ARNm maternal sobre el control y la regulación de la calidad del huevo (Brooks y col., 1997), hace que desde el punto de vista productivo despierte especial interés la identificación de genes que poseen una expresión maternal. Sin embargo, entre la gran cantidad de transcritos maternos que se transcriben y trasladan al oocito durante la ovogénesis, solamente algunos de ellos resultaran vitales para el desarrollo del embrión, y será en ellos donde, la aparición de patrones de expresión anómalos conlleve a problemas en el desarrollo embrionario, provocando una disminución en la calidad de huevo (Brooks y col., 1997). Por lo tanto, la búsqueda de genes maternos con posible influencia sobre calidad de huevo, se debe enfocar principalmente hacia genes cuya función biológica se corresponde con el control/regulación de procesos celulares básicos indispensables para garantizar el correcto desarrollo del organismo. Por ello, en nuestro caso la búsqueda de posibles genes candidatos se dirigió específicamente a genes cuya función se implica dentro de las rutas de obtención de energía, el ciclo celular, la estructura celular, la degradación del vitelo, rutas de señalización destinadas a la regulación del crecimiento, proliferación y diferenciación celular, así como supervivencia, apoptosis e inmunidad (Tabla 5a). Este criterio de selección se considera el responsable de que 11 (*gapdh-1*, *cathepsina-D*, *cathepsina-L*, *cathepsina-Z*, *tubulina-β*, *ferritina-M*, *ciclina-A2*, *ppar-β*, *erk-1*, *jnk-1* y *p38-α*) de los 16 genes evaluados presenten el perfil de expresión propio de genes maternos. Este perfil de expresión génica se caracteriza porque los niveles de los transcritos se mantienen (la intensidad de la expresión varía en función del gen evaluado) entre los estadios de huevo no fertilizado (fase no analizada en el estudio) y blástula, y a partir de este último estadio, coincidiendo con el punto en cual se ha estimado que transcurre la fase de transición maternal-zigótica (Kane y Kimel, 1993), se origina un proceso de degradación de los ARNm maternos (Mathavan y col., 2005; Giraldez y col., 2006), visualizado mediante la técnica de qRT-PCR como una reducción drástica de la señal de expresión (Figura 5c). Aunque desde el punto de vista fisiológico la destrucción de una gran cantidad de ARNm maternos previamente almacenados en el oocito, supone un derroche energético hasta el momento inexplicable, existen evidencias que postulan el agotamiento de las reservas de estas moléculas, como un proceso necesario e indispensable para la correcta activación del genoma del cigoto (Macaulay y col., 2011).

La PCR a tiempo real es una técnica ampliamente utilizada hoy en día para cuantificar

expresión génica. La precisión de los resultados obtenidos mediante esta técnica se encuentra condicionada por la calidad del ARN empleado, la especificidad de los iniciadores, y por la utilización de genes control apropiados que permita cuantificar de forma correcta la expresión relativa del gen objeto de estudio (Mitter y col., 2009). En nuestro caso el sistema de extracción de ARN basado en la combinación del protocolo estándar de Trizol junto con los tubos Lysing Matrix D, permitió obtener un ARN de alta calidad con mínima degradación (dato no mostrado). Así mismo, la realización de curvas de disociación como paso final en las reacciones de PCR a tiempo real, junto con la comprobación de los productos de PCR obtenidos en un gel de agarosa, ofrecieron un control exhaustivo que nos permite garantizar que la señal de expresión registrada no se encuentra afectada por la presencia de dímeros de iniciadores o amplificaciones inespecíficas que pudieran ocasionar una “falsa” o secundaria señal de expresión. Por otro lado, hasta hace poco tiempo la selección de los genes que actúan como control interno, se dirigía hacia aquellos genes cuya transcripción se creía que permanecía constante con independencia de las condiciones experimentales y el tipo de tejido/muestra analizado. Sin embargo, hoy en día, este concepto de genes control, también denominados genes “housekeeping” se ha catalogado de irreal, por lo que se recomienda para cada estudio verificar previamente la validez de los genes seleccionados como control, sugiriéndose además la utilización de un mínimo dos genes control por estudio (Mitter y col., 2009). Siguiendo estas recomendaciones, en nuestro caso los genes *18S* y *L13a* seleccionados como control, cumplieron las exigencias demandas presentando una expresión prácticamente invariable a lo largo de los diferentes estadios de desarrollo analizados, concordando con los resultados obtenidos por Mitter y col. (2009) en lubina [especie que guarda una estrecha cercanía evolutiva con la dorada (Figura 5f)].

Un ejemplo de la desaparición del concepto de gen “housekeeping”, lo encontramos en el gen *gapdh*, cuyas isoformas *gapdh-1* y *gapdh-2* se han incluido en este estudio como posibles candidatos a genes maternos. Durante mucho tiempo, el gen *gapdh* se utilizó como gen control estándar para estudios de expresión, ya que al codificar para el enzima gliceraldehído-3-fosfato deshidrogenasa, participante en una ruta metabólica esencial para la formación de energía en la célula como es la glicólisis, se asumió erróneamente que su expresión era ubicua y constante en los diferentes tejidos del organismo, y a lo largo de las diferentes fases del desarrollo (Sarropoulou y col., 2011). Sin embargo, el descubrimiento de su comportamiento multifuncional, participando no solo en la glicólisis sino también en procesos como la apoptosis y la neurodegeneración (Chuang y col., 2005; Sirover, 2005), junto con las variaciones de expresión registradas durante los estadios iniciales del desarrollo observadas en nuestro

estudio, y anteriormente por Sarropoulou y col. (2011), han puesto seriamente en duda su validez como gen control.

Concordando con nuestros resultados, Sarropoulou y col. (2011) detectaron transcritos del gen *gapdh-1* en huevos no fertilizados y en estadios iniciales del desarrollo embrionario de dorada y lubina, proponiendo por primera vez la actuación de este gen como gen maternal. Aunque se necesitan más datos para establecer una precisa correlación entre este gen y el carácter calidad de huevo, la participación del enzima GAPDH-1 en procesos de apoptosis, unido a la relación que parece existir entre esta y la calidad de huevo (Carnevali y col., 2003), nos permite especular sobre el posible papel del gen *gapdh-1* como biomarcador de la calidad de huevo (Sarropoulou y col., 2011). En referencia a la isoforma *gapdh-2*, la ausencia de expresión detectada durante las primeras fases del desarrollo embrionario de dorada, reafirma el patrón de expresión propuesto nuevamente por Sarropoulou y col. (2011), quienes describen en dorada una expresión más tardía para la isoforma *gapdh-2* que para *gapdh-1*. Por lo tanto, la concordancia entre el patrón de expresión obtenido para estas dos isoformas y el patrón previamente descrito en dorada por Sarropoulou y col. (2011), ha supuesto que estos dos genes actúen en nuestro estudio como un buen control interno.

El segundo grupo de genes incluidos en el estudio lo conforman los genes codificantes para catepsinas. Las catepsinas, son enzimas proteolíticas lisosomales encargados de catalizar la hidrólisis de proteínas a péptidos y, concretamente, se encuentran implicados en la regulación de la presentación de antígenos, así como en procesos de degradación de materiales endocitados en los lisosomas (Honey y Rudensky, 2003; Hsing y Rudensky, 2005). De los cuatro genes analizados (*catepsina-D*, *-L*, *-S*, y *-Z*), el gen *catepsina-S* (Figuras 5b), fue el único que no mostró expresión alguna durante la fase inicial del desarrollo embrionario, lo cual induce a pensar que la necesidad de esta proteína durante desarrollo temprano es muy limitada. En mamíferos para esta proteína se ha descrito un importante papel durante la fase inicial de reconocimiento de la respuesta inmune (Turnsek y col., 1975), sin embargo, pese a que en peces teleósteos la regulación de sus transcritos en respuesta a una infección permite especular sobre su posible implicación en el sistema inmune (Yeh y Klesius, 2008), su función fisiopatológica aún resulta desconocida (Sun y col., 2011).

Dentro de superfamilia de las catepsinas, las catepsinas-B, -D y -L desarrollan un papel importante durante la oogénesis, participando en el proceso de degradación del vitelo (Carnevali y col., 1999a y b), proporcionando los nutrientes esenciales para el correcto desarrollo del embrión. Por consiguiente, el funcionamiento de estas enzimas se debería considerar un proceso clave para alcanzar el éxito reproductivo (Carnevali y col., 2001). Carnevali y col., (1999b y 2001), a partir de estudios dirigidos hacia la evaluación del carácter

calidad de huevo en dorada y en lubina, determinaron que la actividad del enzima catepsina-D es más elevada en huevos flotantes (buena calidad) que en huevos no flotantes (mala calidad). De igual modo, Aegerter y col. (2005) detectaron en trucha arco iris (*Oncorhynchus mykiss*) que los niveles de expresión de *catepsina-Z* resultan superiores en los huevos clasificados con mala calidad frente a huevos de buena calidad. Sin embargo, a diferencia de *catepsina-D*, el sentido biológico de la asociación entre *catepsina-Z* y calidad de huevo, aún no se ha determinado.

Por su parte, el grupo compuesto por los genes *ciclina-A2*, *tubulina β* y *ferritina M*, también presentó una posible expresión maternal en dorada. Las ciclinas, son proteínas implicadas en el progreso del ciclo celular que actúan uniéndose a quinasas, que a su vez, mediante fosforilación, activan a otros enzimas generándose una señal en la célula que conlleva la síntesis de las proteínas necesarias para regular el ciclo celular. En el pez cebra (*Danio reiro*), a través de estudios de expresión realizados mediante microarrays, se determinó que la mayoría de los genes pertenecientes a la familia de las ciclinas comenzaban su expresión de forma muy temprana en el desarrollo, manteniendo el nivel de sus transcritos hasta el estadio de blástula, momento a partir del cual su expresión se veía fuertemente disminuida (Mathavan y col., 2005). Además del patrón de expresión observado para el gen *ciclina-A2* en nuestro estudio, un factor adicional que contribuye a reafirmar la clasificación del este gen como un gen de expresión maternal, es la detección de sus transcritos en el huevos no fertilizados de trucha arco iris (Aegerter y col., 2005). Dada la importante función de la familia de las ciclinas durante el ciclo celular, una expresión anómala podría provocar alteraciones en el ciclo celular durante el desarrollo del embrión, y por consiguiente, se correlacionaría con una reducción de la calidad del huevo (Aegerter y col., 2005).

De igual manera que el éxito en el proceso de división celular resulta indispensable para el correcto desarrollo del embrión, la correcta estructuración del citoesqueleto de la célula debería ser otro factor clave. En base a esta hipótesis, variaciones en la expresión de los genes codificantes para los componentes del citoesqueleto, también se podrían vincular con variaciones en calidad de huevo. Este tipo de correlación ha sido establecida en trucha arco iris por Aegerter y col. (2005), quienes han detectado que los niveles de expresión en el gen *tubulina- β* se muestran más reducidos en huevos de baja calidad respecto a huevos de calidad buena-intermedia.

En referencia a la proteína ferritina cabe destacar su función en el metabolismo del hierro, almacenando el hierro transportado por la transferrina en el interior de la célula. Hasta la fecha, para esta proteína se han descrito tres tipos de subunidades: H, L y M. La subunidad

M (subunidad analizada en el estudio) se ha identificado en peces y anfibios (Dickey y col., 1987; Andersen y col., 1995; 1998; Mignogna y col., 2002; Giorgi y col., 2008), y aunque su función específica aún no se ha esclarecido, recientemente se le ha vinculado una posible actividad antimicrobiana (Wang y col., 2011). La elevada expresión registrada para el gen *ferritina-M* en nuestro estudio, sugiere que durante el desarrollo temprano del embrión este gen juega una importante función, reforzándose la hipótesis de su actuación como gen maternal, y posible biomarcador de calidad de huevo.

En cuarto lugar, se sitúan los genes *erk-1* (MAPK3), *jnk-1* (MAPK8), *p38- α* (MAPK14a) y *p38- δ* (MAPK13) que representan a las tres sub-familias que constituyen la familia de las proteínas quinasas activadas por mitógenos o MAPKs [ERK (quinasas reguladas por señales extracelulares), JNK (c-Jun-quinasa terminal, se une y fosforila a la proteína c-Jun, que al combinarse con la proteína c-Fos forma el factor de transcripción AP1) y p38]. Estas proteínas además de desarrollar un papel importante en la regulación de procesos celulares como el crecimiento, la proliferación, la diferenciación, la supervivencia y la inmunidad innata (Bogoyevitch y Court, 2004; Roux y Blenis, 2004; Johnson y col., 2005), intervienen como reguladores del balance entre la supervivencia y la apoptosis (Marshall, 1995; Ballif y Blenis, 2001). Las MAPKs, se activan a través de una gran variedad de señales, en general, la sub-familia ERK es principalmente activada por factores de crecimiento, p38 por factores de estrés, y JNK por factores de estrés, diferenciación y crecimiento (Krens y col., 2006). Por medio de una cascada de fosforilación de diversos sustratos la señal recibida por la célula se integra y se transforma en una respuesta biológica apropiada. De los cuatro genes codificantes para MAPKs analizados en el estudio, los genes *erk-1*, *jnk-1* y *p38- α* , presentaron al igual que los anteriores, el patrón de expresión propuesto para genes con posible función maternal, mientras *p38- δ* presentó un perfil de expresión diferente. Así, resultados obtenidos para *erk-1*, *jnk-1* y *p38- α* , concuerdan con los resultados propuestos por Krens y col. (2006), en el pez cebra, donde un amplio conjunto de MAPKs (entre los que se encuentran los genes *erk-1*, *jnk-1*, y *p38- α*) mostraron una clara expresión maternal durante el desarrollo embrionario. Por el contrario, el patrón de expresión no maternal detectado para el gen *p38- δ* se asemeja al patrón de expresión descrito en el pez cebra para *p38 β* (Krens y col., 2006) el cual mostró una expresión de intensidad reducida, limitada al espacio de tiempo comprendido entre los estadios de 16 células y blástula. Hasta la fecha, p38- δ es una de las isoformas MAPKs peor caracterizadas, existiendo muy poca información sobre ella.

Finalmente, el último grupo de genes analizados lo componen las isoformas α , β y γ de los receptores activados por factores de proliferación de peroxisomas o PPARs, pertenecientes a la superfamilia de receptores nucleares de hormonas. Leaver y col. (2005) analizaron la

expresión de los genes *ppar- α* y *ppar- γ* en dorada, y describieron para la isoforma *ppar- α* un patrón similar al observado en mamíferos, caracterizado por una expresión elevada en tejidos con alta capacidad para degradar ácidos grasos por β -oxidación, mientras que para la isoforma *ppar- γ* , se establece una expresión mucho más amplia en dorada que en mamíferos, identificándose sus transcritos en un mayor número de tejidos. Sin embargo, la información referente a la expresión génica de *ppars* durante el desarrollo de peces es muy limitada o inexistente. De acuerdo a nuestros resultados, a partir de la ausencia de transcritos de las isoformas *ppar- α* y *ppar- γ* se deduce para estas proteínas una función nula o restringida durante las primeras fases del desarrollo embrionario, concordando con el patrón de expresión tardío (13,5 días de desarrollo embrionario) publicado por Braissant y Wahli (1998) en rata. Por consiguiente, si el patrón de expresión de dorada guarda una estrecha similitud con el patrón propuesto en rata, la expresión de estas dos isoformas debería de ser evaluada en estadios de desarrollo embrionario más avanzados. A diferencia de las anteriores isoformas, el gen *ppar- β* sí parece desempeñar una función importante durante el desarrollo embrionario temprano, al mostrar una vez más el característico perfil de expresión génica maternal (Figura 5c), coincidiendo con el patrón descrito en rata para esta isoforma (Brassiant y Wahli, 1998).

En base a los resultados obtenidos, aunque se precisa de estudios funcionales que verifique la asociación significativa entre gen y calidad de huevo más allá de un patrón de expresión maternal en las primeras fases del desarrollo, se propone a los genes *gapdh-1*, *catepsina-D*, *catepsina-L*, *catepsina-Z*, *tubulina- β* , *ferritina-M*, *ciclina-A2*, *ppar- β* , *erk-1*, *jnk-1* y *p38- α* como posibles genes candidatos asociados a calidad de huevo en dorada, existiendo fuertes evidencias de ello especialmente para los genes *catepsina-D* (Carnevali y col., 1999 y 2001), *catepsina-Z* y *tubulina- β* (Aegerter y col., 2005).

El futuro desarrollo de marcadores moleculares asociados a la calidad de huevos contribuirá a asentar las bases genéticas de este carácter especialmente relevante para dorada, donde la gran variabilidad existente en la calidad de huevo, y la obtención de un porcentaje de eclosión que suele oscilar entre el 10 y 15% del total de huevos producidos (Carrillo y col., 1989), pueden llegar a suponer un importante cuello de botella en la gestión de la reproducción y comprometer la producción de su cultivo.

5.4.2 Caracterización de gen *Tf*: un primer paso hacia la búsqueda de polimorfismos funcionales.

Identificados los genes candidatos para un determinado carácter, el desarrollo de estudios funcionales dirigidos a evaluar cómo la variación alélica existente en la secuencia de

estos genes puede conducir a la obtención de un fenotipo diferente, precisa de la caracterización detallada del gen o genes en cuestión, que proporcione no solamente el recurso necesario para llevar a cabo la búsqueda de estas posiciones polimórficas (secuencia de ADN genómico o ADNc), sino también, la información biológica necesaria para dar un sentido funcional a la variación identificada (por ejemplo, el tipo de región génica donde se localiza la variación, y su posible implicación en la alteración de la expresión o estructura de la proteína codificada). Dentro de este contexto, se sitúa la caracterización del gen *Tf* de dorada desarrollada en esta parte del estudio, basada en el aislamiento de su secuencia de genómica utilizando la información de secuencias homólogas descritas en especies cercanas, en la obtención de su secuencia codificante (ADNc) y en la descripción del patrón de expresión constitutiva observado para siete tejidos diferentes dentro de las fases juvenil y adulta.

La estructura deducida para el gen *Tf* de dorada se compone de una región codificante de 2.076pb organizada en 17 exones, separados por 16 intrones cuya longitud de secuencia se extiende 5.495pb, resultando una longitud de secuencia intrónica que supera en aproximadamente en 2,6 veces la longitud de la secuencia codificante (Tabla 5f). La comparación del gen *Tf* de dorada con las estructuras previamente descritas para otras especies de peces, mostró una elevada conservación del tamaño exónico, existiendo una conservación especialmente elevada entre dorada y tilapia del Nilo (Tabla 5e). Esta conservación en la longitud de la secuencia codificante, conllevó a la deducción de proteínas con pequeñas variaciones en la longitud de su secuencia aminoacídica, presentando el pez gato y tilapia del Nilo las proteína de mayor y menor tamaño, con 679 y 694 aminoácidos, respectivamente. La proteína *Tf* deducida para dorada se compone de 691 aminoácidos, que se organizan en una estructura bi-lobular (lóbulo N- y C-terminal) altamente conservada.

La reconstrucción filogenética realizada a partir de la secuencia de aminoácidos deducida para *Tf* en diferentes especies, reunió dentro de un mismo grupo (con cortas distancias genéticas y altos valores de bootstrap) a la dorada junto al resto de especies de espáridos comparadas. A este mismo grupo, a pesar de no pertenecer a la familia de los espáridos, se unió con alta confianza la lubina, perteneciente a la familia Moronidae (Orrel y Carpenter, 2004) (Figuras 3g). Esta reconstrucción evolutiva, permitió confirmar que el grado de conservación medio-alto de la proteína *Tf* detectado dentro de las especies de peces, se extiende también al grupo de los vertebrados. Esta elevada conservación probablemente deriva de su importante función en el metabolismo del hierro y en la respuesta del sistema inmune innato, dos procesos de vital importancia para el organismo que podrían estar causando fuertes restricciones evolutivas sobre la proteína.

Tal como se esperaba, la expresión del gen *Tf* se detectó, tanto en individuos juveniles

como adultos, en un amplio espectro de tejidos, observándose para el tejido hepático una intensidad especialmente elevada. La comparación del perfil de expresión obtenido entre individuos adultos y juveniles, reveló para el tejido estomacal y esplénico una expresión de *Tf* dependiente de ontogenia. Aunque estos resultados deben ser tomados con precaución, resultando necesario establecer con mayor precisión mediante PCR a tiempo real, las diferencias de expresión observadas entre las dos fases de desarrollo, la escasa presencia de transcritos de *Tf* en el tejido esplénico juvenil podría ser consecuencia del escaso desarrollo y de la falta de funcionalidad de los órganos linfoides durante este estado de desarrollo prematuro. Esta premisa concuerda con la sensibilidad especialmente elevada que presentan los individuos juveniles (peso corporal <50g) a infecciones como la pasteurelosis causada por la bacteria *Photobacterium damselea* subespecie *piscicida* (*Phdp*) (Antonello y col., 2009) que afecta principalmente a las granjas de cultivo de dorada y lubina.

Al igual que en bacalao (Devonan-Wright y col., 1996), lubina (Neves y col., 2009), carpa (Sahoo y col., 2009) y pez gato (Liu y col., 2009), la dorada mostró una expresión local de *Tf* en cerebro. El significado biológico de esta expresión cerebral se puede explicar aludiendo a la posible dificultad que esta proteína podría encontrar para atravesar la barrera hematoencefálica cuya función reside en facilitar el paso de nutrientes y de oxígeno al sistema nervioso y a su vez, dificultar el paso de sustancias de gran tamaño molecular y/o tóxicas. Sin embargo, esta hipótesis pierde veracidad ante la posible existencia de una expresión de *Tf* cerebral especie-específica, al no detectarse transcritos de *Tf* en el tejido cerebral en especies como el salmón común (*Salmo salar*) (Kvingedal y col., 1993). Neves y col. (2009), sugieren que la síntesis local de *Tf* en el cerebro podría estar asociada a un mecanismo de defensa de inmunidad innata del hospedador dentro del propio tejido cerebral, pues al profundizar sobre la expresión cerebral de *Tf* en lubina, observaron que en respuesta a una infección bacteriana la proteína *Tf* aumentan rápidamente su concentración, comportándose como una proteína de fase aguda positiva.

Ante una infección, a nivel hepático las especies parecen responder con estrategias de defensa diferentes en cuanto a la regulación del gen *Tf*, generando un aumento o bien, una disminución de su expresión; no obstante ambas estrategias tienen como finalidad bloquear la utilización del hierro del hospedador al huésped. Por ejemplo, en lubina, en respuesta a un ataque bacteriano la expresión hepática de *Tf* se ve fuertemente disminuida comportándose como una proteína de fase aguda negativa (Neves y col., 2009), mientras que en pez gato, su actuación es propia de una proteína de fase aguda positiva, respondiendo con un incremento en la producción de sus transcritos (Liu y col., 2009). Es posible que estas diferentes

estrategias tengan su origen en la diferente presión evolutiva ejercida que los patógenos principales de cada una de estas especies. De esta manera, aquellas especies cuyos patógenos principales fundamenten su sistema de captura de hierro en sideróforos capaces de adherirse al complejo hierro-Tf del organismo hospedador, podría haber desarrollado a lo largo de la evolución un sistema de defensa inmune basado en una disminución de la expresión *Tf*, capaz de originar ambientes con niveles reducidos de Tf donde estos patógenos verán mermada su capacidad de supervivencia y desarrollo. En contraposición, aquellas especies cuyos invasores principales sean capaces de tomar el hierro libre directamente del plasma del hospedador, un incremento en los niveles de Tf sería la respuesta más eficaz, ya que reducirá la cantidad de hierro libre en plasma disponible para estos patógenos. En este sentido, al ser la dorada una especie especialmente sensible a los mismos patógenos que afectan al cultivo de la lubina (por ejemplo, *Photobacterium damselae*), ante una invasión patogénica sería de esperar una reacción de Tf propia de una proteína de fase aguda negativa; no obstante para esclarecer como la regulación del gen *Tf* pueden actuar como alarma de defensa para el sistema inmune innato de dorada, se requiere de estudios específicos que determinen el patrón de expresión del gen *Tf* en respuesta a un ataque patogénico.

En resumen, a partir de la caracterización del gen *Tf* de dorada se deduce respecto a otras especies, una preservación de la estructura génica y una conservación elevada de la proteína codificada, describiéndose a su vez un patrón de expresión constitutiva dependiente de tejido y ontogenia. Además, la información referente a las secuencias de ADN genómico (GenBank ID: JF309047) y ADNc (GenBank ID: JF309047) depositadas en la base de datos del NCBI, abren la puerta al desarrollo de estudios dirigidos a la búsqueda de polimorfismos funcionales (variaciones nucleotídicas con capacidad para influenciar sobre la expresión del carácter) que posibiliten al igual que en otras especies de peces (Suzumoto y col., 1977; Winter y col., 1980; Jurecka y col., 2009a, b y c), establecer correlaciones significativas entre la presencia de diferentes genotipos del gen *Tf* con una mayor/menor susceptibilidad a patógenos, u otros posibles caracteres asociados al gen.

5.5. Tablas

Tabla 5a. Clasificación de los genes incluidos en el estudio de determinación de la expresión génica durante el desarrollo embrionario temprano.

	Gen	Función
1	<i>gapdh-1</i>	Procesos catabólicos celulares básicos
2	<i>gapdh-2</i>	
3	<i>catepsina-D</i>	Enzimas proteolíticas localizadas en los lisosomas con una importante función en la digestión celular
4	<i>catepsina-L</i>	
5	<i>catepsina-S</i>	
6	<i>catepsina-Z</i>	
7	<i>tubulina-β</i>	Componente esencial del citoesqueleto
8	<i>ferritina-M</i>	Metabolismo del hierro
9	<i>ciclina-A2</i>	Enzima implicada en la regulación del ciclo celular
10	<i>ppar-α</i>	Receptores activadores de proliferación de peroxisomas: son factores de transcripción que desarrollan una importante función en la homeostasis lipídica
11	<i>ppar-β</i>	
12	<i>ppar-γ</i>	
13	<i>erk-1</i>	Proteín quinasas activadas por mitógenos: enzimas que participan en la transducción de señales intracelulares, que regulan múltiples procesos celulares entre ellos cambios en los patrones de expresión génica, el metabolismo, la proliferación celular y la muerte celular programada.
14	<i>jnk-1</i>	
15	<i>p38-α</i>	
16	<i>p38-δ</i>	
17	<i>L13a (control)</i>	Proteínas ribosomales
18	<i>S18 (control)</i>	

Tabla 5b. Pares de iniciadores utilizados en el estudio de determinación de la expresión génica durante el desarrollo embrionario temprano.

Gen	Inicador 5'-3'	Iniciador 3'-5'	Ta	Amplión (pb)	Bibliografía
<i>gadh1</i>	CCAGCCAGAACATCATCC	GCAGCCTTGACGACCTTC	57	172	Sarropoulou y col., 2011
<i>gadh2</i>	CATGAAGCCAGCAGAGATCC	GCTGGCCGGTCAATATTC	57	190	
<i>ppara</i>	GCAGCCTGTGAGTCTTGTGAGTGA	CTCCATCAGGTCTCCACACAGC	57	121	
<i>pparβ</i>	CGTGTTCGGGATTCGGGACT	CACCCCTGCTGCTGCTCTGTA	62	186	Leaver y col. 2005
<i>ppary</i>	CGGAGAGAGAAGCAAGAACAAGAA	GAGGAGGAGGAGATGGAGGTGTA	57	213	
<i>catepsina D</i>	CTGGCATCACGTTTGCAGTG	CATGGCTGTGTCAAACACG	57	101	Carnevali y col.,1999b
<i>catepsina L</i>	TGGAGCCCAACTTCCTGGA	AGCCCCAGCAAGAACCACA	62	100	Carnevali y col.,2008
<i>catepsina S</i>	AGGAATGAATGTGGCGTAAG	CGAGGTGGAGGAAGTGAG	57	171	
<i>catepsina Z</i>	GGAGACCCAATCGTCCCCAA	GGCACTGAACGTGACCTGA	57	180	
<i>tubulina β</i>	TCCGTGAAGAGTACCCCGA	GCAGGTAGTGACACCCGCTCA	57	254	
<i>ferritina M</i>	AAGCCTCCAAGGGCCACAT	TGGAGTCCCAAGTCCGTGAG	57	142	Presente estudio
<i>erk1</i>	GCTCTATGGCAAGGCTGAC	TGCCTGGAAAACGAGCTGTT	57	238	
<i>ciclina A2</i>	ATGGCAGCAGTCCCAACC	GCAGCCTGACATCACC AAC	57	177	
<i>jnk1</i>	TCTCCAGCACCCCTTATATCAAC	TGTCCTCTCTTCCCAGTCC	57	157	
<i>P38α</i>	GGCTCACTCCTACTTCTC	TAATCGTTTCCACTCTTCG	57	112	
<i>P38δ</i>	CGAAGTGGGAGGTCATC	CGGTTTACAGCCAAGTTTCC	57	133	
<i>L13a (control)</i>	TCTGGAGGACTGTCAAGGGCATGC	AGACGCACAAATCTTAAGAGCAG	57	148	Mitter y col., 2009
<i>S18 (control)</i>	AGGGTGTGGCAGACCGTTAC	TTTCTGCCTGTTGAGGAACC	57	154	

Ta: Temperatura de hibridación específica de cada pareja de iniciadores en °C; amplicón: tamaño en pb del producto de PCR obtenido.

Tabla 5c. Parejas de iniciadores utilizados para la caracterización del gen *Tf*.

Nombre	Secuencia (5' ->3')	Nombre	Secuencia (5' ->3')	Ta (°C)	Amplificación PCR (pb)
<i>SaTf</i> -F1	TGCTGACCTGGAGACCTCGA	<i>SaTf</i> -R3	AGTCTGTAGTTGTTTCAGTCCAG	62	1357
<i>SaTf</i> -F2	AGAAAGCCTACTGTTTGGCCT	<i>SaTf</i> -R3	AGTCTGTAGTTGTTTCAGTCCAG	62	859
<i>SaTf</i> -F3	CTGGTGAAGCAGACGACCATCACTC	<i>SaTf</i> -R5.2	TGGAAGGCTCCACCCTAGTCA	62	1184
<i>SaTf</i> -F4	CAGGTGCTGAAGACAGCCCGGT	<i>SaTf</i> -R5.2	TGGAAGGCTCCACCCTAGTCA	62	823
<i>SaTf</i> -F5	CTGTGAGCAACTTCTTCAGTGCC	<i>SaTf</i> -R8	GAACATCAGGTTTTTTGGAGGT	64	1698
<i>SaTf</i> -F6	TGACTCCTCAGCCAAGCAAGCT	<i>SaTf</i> -R8	GAACATCAGGTTTTTTGGAGGT	64	1147
<i>SaTf</i> -F7	ACCTGGCCAAAGTACCAGTCAATG	<i>SaTf</i> -R9	AAGGCACCCGTCAACTGTGGCAGC	64	620
<i>SaTf</i> -F9	TGGAGCATCAACACAGCATGACTG	<i>SaTf</i> -R13	GCTCAATACTACTGGCTTTTGCA	62	1626
<i>SaTf</i> -F11	GCAATGCGGCCAACCAATGGA	<i>SaTf</i> -R13	GCTCAATACTACTGGCTTTTGCA	64	1246
<i>SaTf</i> -F12	AGTTCTTCAGTGAGGGCTGTGC	<i>SaTf</i> -R14	ACCACTAGCCCATGCTGGACCA	64	1008
<i>SaTf</i> -F14	ACCAGGAACATTACCCACATG	<i>SaTf</i> -R15	CCACAGTTGTACAAAACCTGAG	64	883
<i>SaTf</i> -F15	GATCCCTCGTTCAAACCTGTCCA	<i>SaTf</i> -R17	AAGTCAGGCATGTGTTTGGCAATAC	64	639
<i>SaTf</i> -FW	CAGTAACGGCAATGGTCCAGC	<i>SaTf</i> -RV	CCAGCTGAAACATCCTACAGAC	60	314
β -actina FW (control)	CTTCCTCGGTATGGAGTCCTGCGG	β -actina RV (control)	TCCTGCTTGCTGATCCACATCTGCT	60	367

Ta: Temperatura de hibridación específica de cada pareja de iniciadores en °C; amplicón: tamaño en pb del producto de PCR obtenido.

Tabla 5d. Nuevos genes identificados en dorada.

Gen	Resultados BLASTX	Numero de acceso base de datos ESTs
<i>catepsina-S</i>	AC082388.1 cathepsin S [<i>Lutjanus argentimaculatus</i>]	AM955494; HS988673; HS988821
<i>catepsina-Z</i>	AA064476.1 cathepsin Z precursor [<i>Fundulus heteroclitus</i>]	FM145771; FM160311; AM961674
<i>tubulina-β</i>	XP_003455078.1 PREDICTED: tubulin beta-1 chain-like [<i>Oreochromis niloticus</i>]	HS985002; HS987551; SATP0002B02_F
<i>ferritina-M</i>	ACY75476.1 ferritin M subunit [<i>Larimichthys crocea</i>]	HS987130, HS983799, SATP0001D20_F; SATP0001H01_F; SATP0001H16_F
<i>erk-1</i>	NP_001167267.1 Mitogen-activated protein kinase 3 [<i>Salmo salar</i>]	FM146973; FM147422; FG267570
<i>ciclina-A2</i>	AAF82778.1 cyclin A2 [<i>Carassius auratus</i>]	AM951618; AM976083
<i>jnk-1</i>	AET86807.1 jnk-1 [<i>Carassius auratus red var.</i>]	HS984271; HS984888
<i>p38-α</i>	NP_001117170.1 mitogen activated protein kinase p38a [<i>Salmo salar</i>]	AM966839, AM966839
<i>p38-δ</i>	XP_003449202.1 mitogen-activated protein kinase 13-like [<i>Oreochromis niloticus</i>]	AM979198

En negrita se indican los números de acceso de las ESTs identificadas en la base de datos de ESTs del HCMR (*Hellenic Center For Marine Research*)

Tabla 5e. Comparación de la estructura del gen codificante para Tf de diferentes especies de peces.

Exones	<i>S. aurata</i>	<i>I. punctatus</i>	<i>O. niloticus</i>	<i>O. kisutch</i>	<i>O. latipes</i>	Intrones	<i>S. aurata</i>	<i>I. punctatus</i>	<i>O. niloticus</i>	<i>O. kisutch</i>	<i>O. latipes</i>
E-I	40	40	40	40	40	IN-I	83	905	33	111	66
E-II	149	137	152	152	149	IN-II	989	1203	715	-	364
E-III	103	103	103	103	103	IN-III	96	508	128	265	85
E-IV	192	189	192	192	192	IN-IV	663	348	784	-	287
E-V	130	127	130	130	130	IN-V	157	100	96	189	103
E-VI	56	56	56	56	56	IN-VI	1019	80	363	169	693
E-VII	164	161	173	164	164	IN-VII	118	590	106	701	150
E-VIII	154	151	154	151	154	IN-VIII	107	983	85	108	86
E-IX	164	167	164	164	164	IN-IX	103	96	91	307	84
E-X	97	97	97	97	97	IN-X	92	1550	98	109	97
E-XI	21	21	21	21	21	IN-XI	471	343	604	146	550
E-XII	165	165	165	162	169	IN-XII	467	136	218	-	404
E-XIII	151	142	151	154	147	IN-XIII	666	1152	332	174	838
E-XIV	68	68	68	68	68	IN-XIV	99	223	85	188	102
E-XV	182	179	182	182	181	IN-XV	208	155	878	-	189
E-XVI	193	193	193	193	194	IN-XVI	149	524	123	102	658
E-XVII	47	44	44	35	44	-	-	-	-	-	-
Total	2076	2040	2085	2064	2073	Total	5487	8896	4739	2568	4756
Proteína	691	679	694	687	690						

5.6. Figuras

Figura 5a. Estados de desarrollo embrionario temprano empleados en la determinación de expresión de los genes *gapdh-1*, *gapdh-2*, *ppar- α* , *ppar- β* , *ppar- γ* , *catepsina-S*, *catepsina-D*, *catepsina-L*, *catepsina-Z*, *tubulina- β* , *ferritina-M*, *ciclina-A2*, *erk-1*, *jnk-1*, *p38- α* y *p38- δ* .

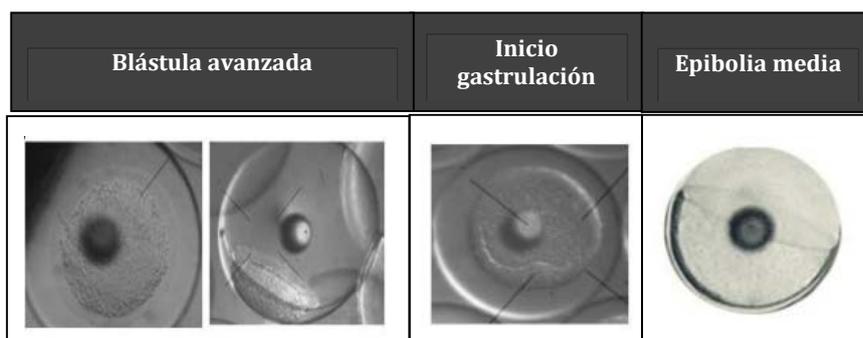
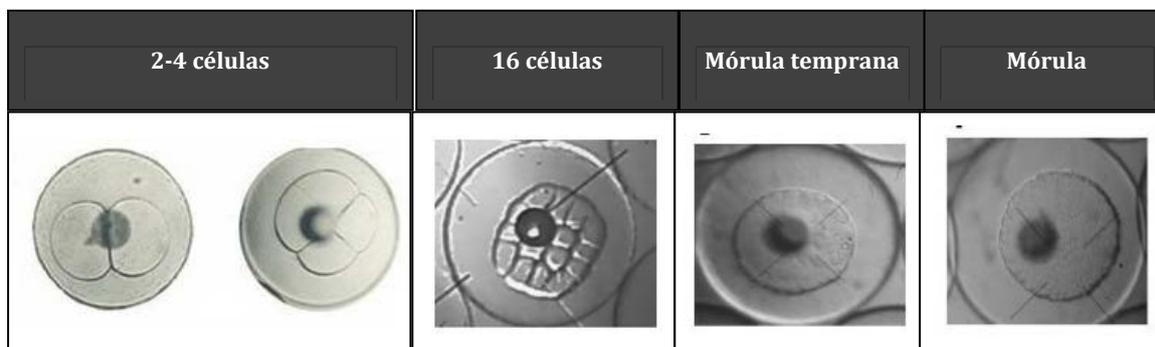
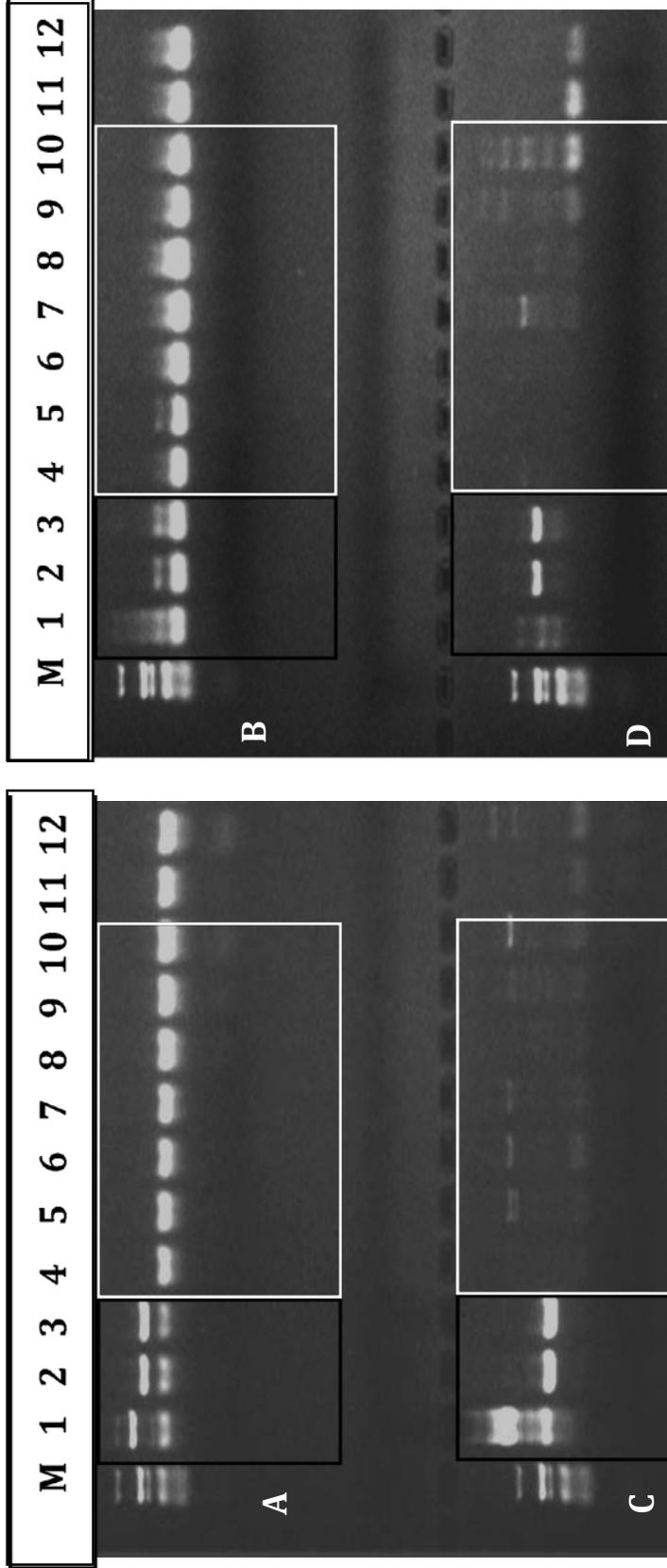
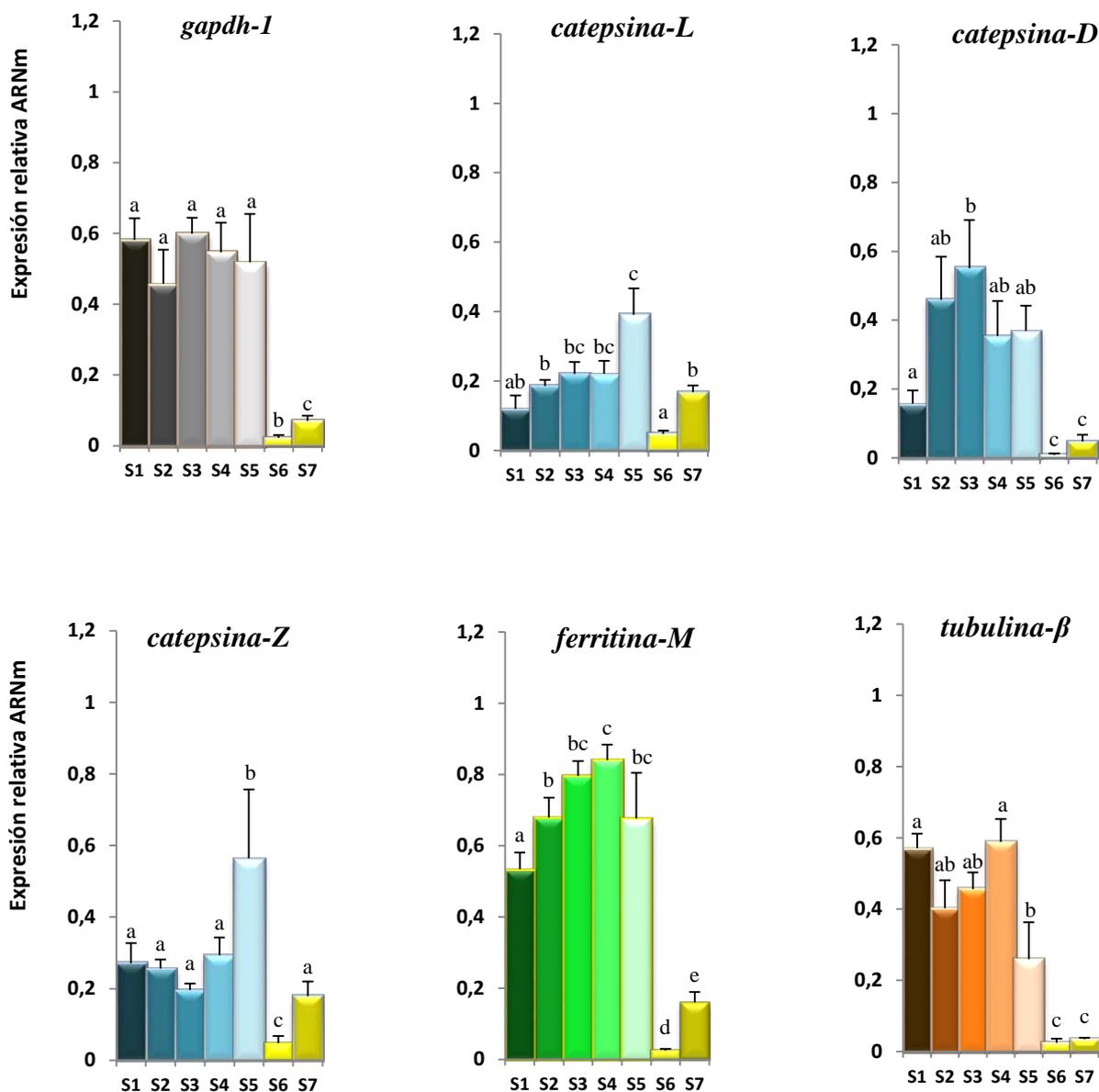


Figura 5h Visualización en gel de agarosa al 2% de la RT-PCR obtenida para los genes *nanhdh-2* (A) *nnar-α* (B) *catensina-S* (C) y *nnar-γ* (D)



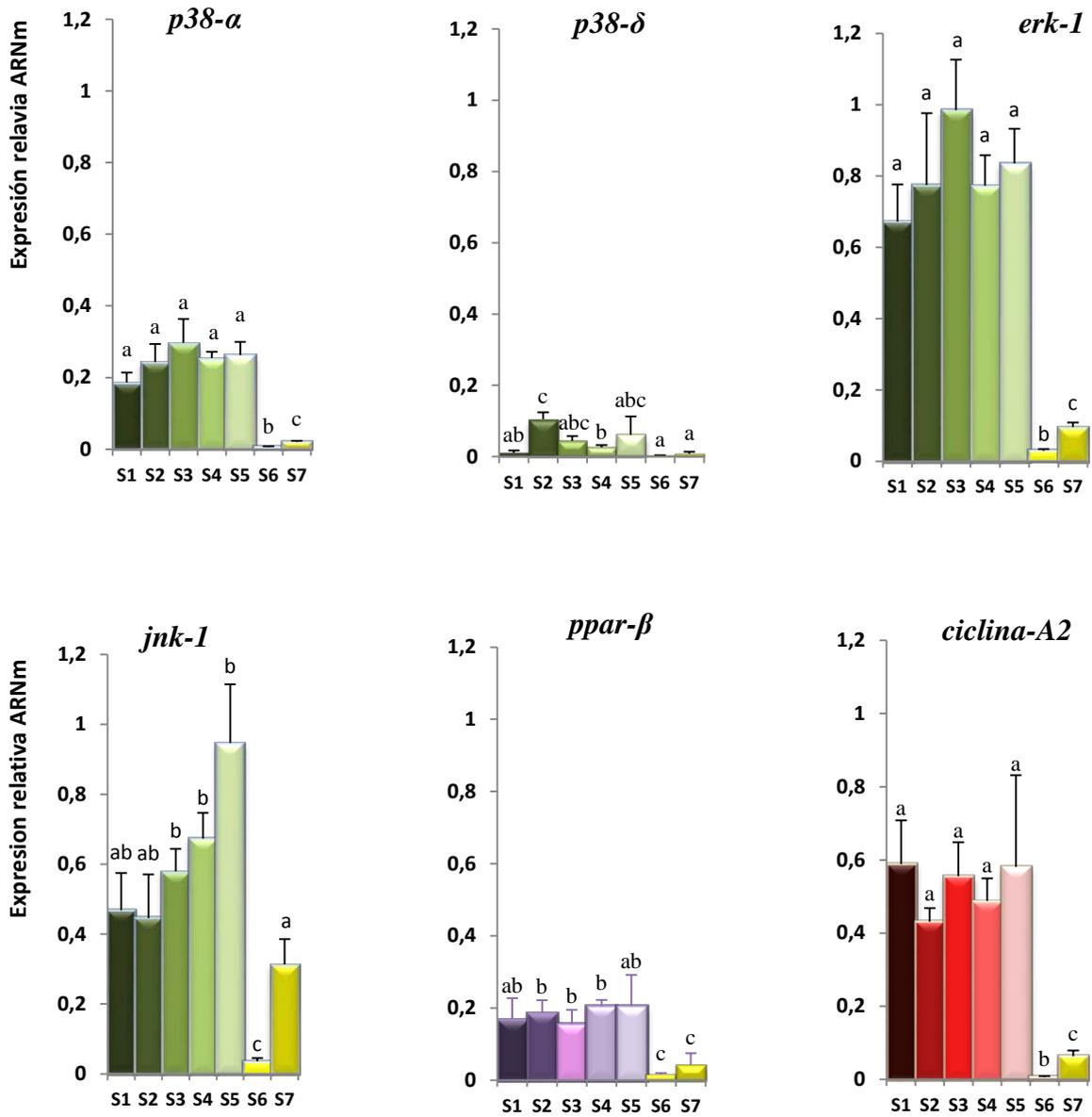
[1] ADN genómico; [2] ADNc de corazón; [3] ADNc de bazo; [4] ADNc estadio de 2-4 célula; [5] ADNc estadio de 16 células; [6] ADNc de estadio mórula temprana; [7] ADNc estadio de mórula; [8] ADNc estadio de blástula avanzada; [9] ADNc estadio correspondiente al inicio de gastrulación; [10] ADNc de epibolia media; [11; 12] control negativo (solo iniciadores).

Figura 5c. Representación gráfica de los niveles de expresión de los genes *gapdh-1*, *catepsina-L*, *catepsina-D*, *catepsina-Z*, *ferritina-M*, *tubulina-β*, *p38-α*, *p38β*, *erk-1*, *jnk-1*, *ppar-β* y *ciclina-A2*.



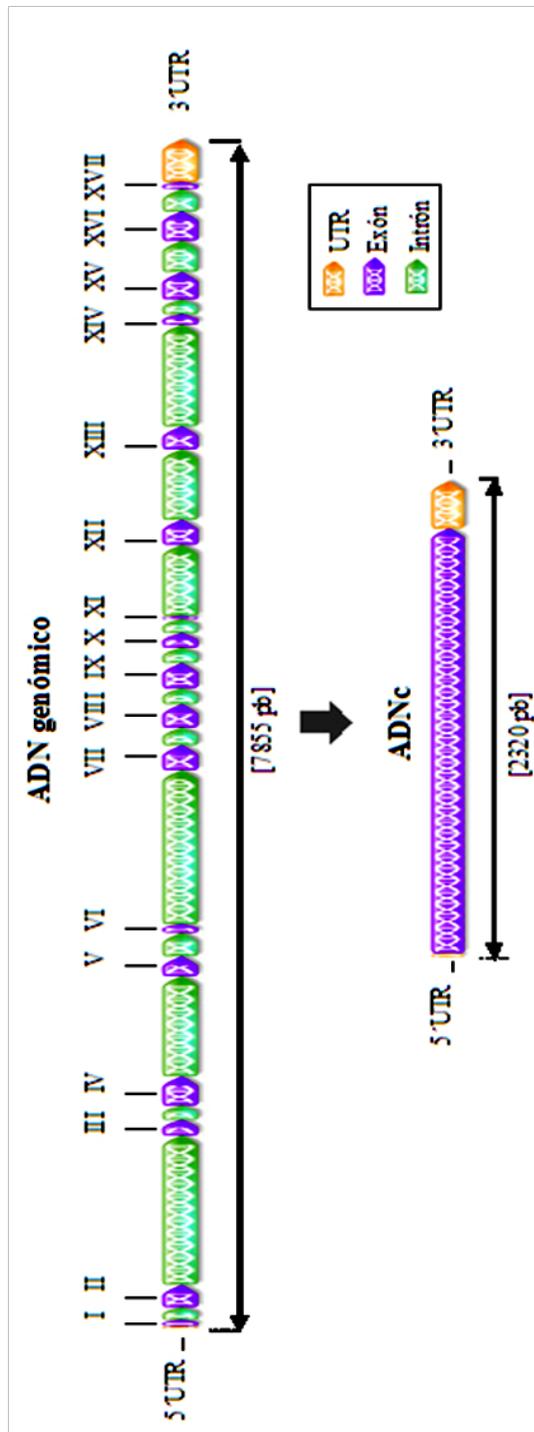
S1: estadio de 2-4 células; S2: estadio de 16 células; S3: estadio mórula temprana; S4: estadio mórula; S5: estadio blástula avanzada; S6: estadio correspondiente al inicio de gastrulación; S7: estadio epibolia media; a, b, c, d y e: representa estadios con diferencias significativas en los niveles de expresión.

Figura 5c. Continuación.



S1: estadio de 2-4 células; S2: estadio de 16 células; S3: estadio mórula temprana; S4: estadio mórula; S5: estadio blástula avanzada; S6: estadio correspondiente al inicio de gastrulación; S7: estadio epibolia media; a, b, c, d y e: representa estadios con diferencias significativas en los niveles de expresión.

Figura 5d. Esquema gráfico de la estructura deducida para la secuencia de ADN genómico del gen *Tf* aislada en el estudio.



UTR: región no transcrita; ADNc: ADN complementario. En números romanos se indica la numeración de los exones.

Figura 5f. Árbol filogenético deducido a partir de la secuencia de amino ácidos de la proteína Tf caracterizada para 15 taxones. Las distancias evolutivas se midieron a través de una matriz WAG (+I=0,08; +G=1,43), y se representan en función del número de cambios o sustituciones de amino ácido por posición. Los números adyacentes a las ramas representan los valores de bootstrap obtenidos para 1000 réplicas.

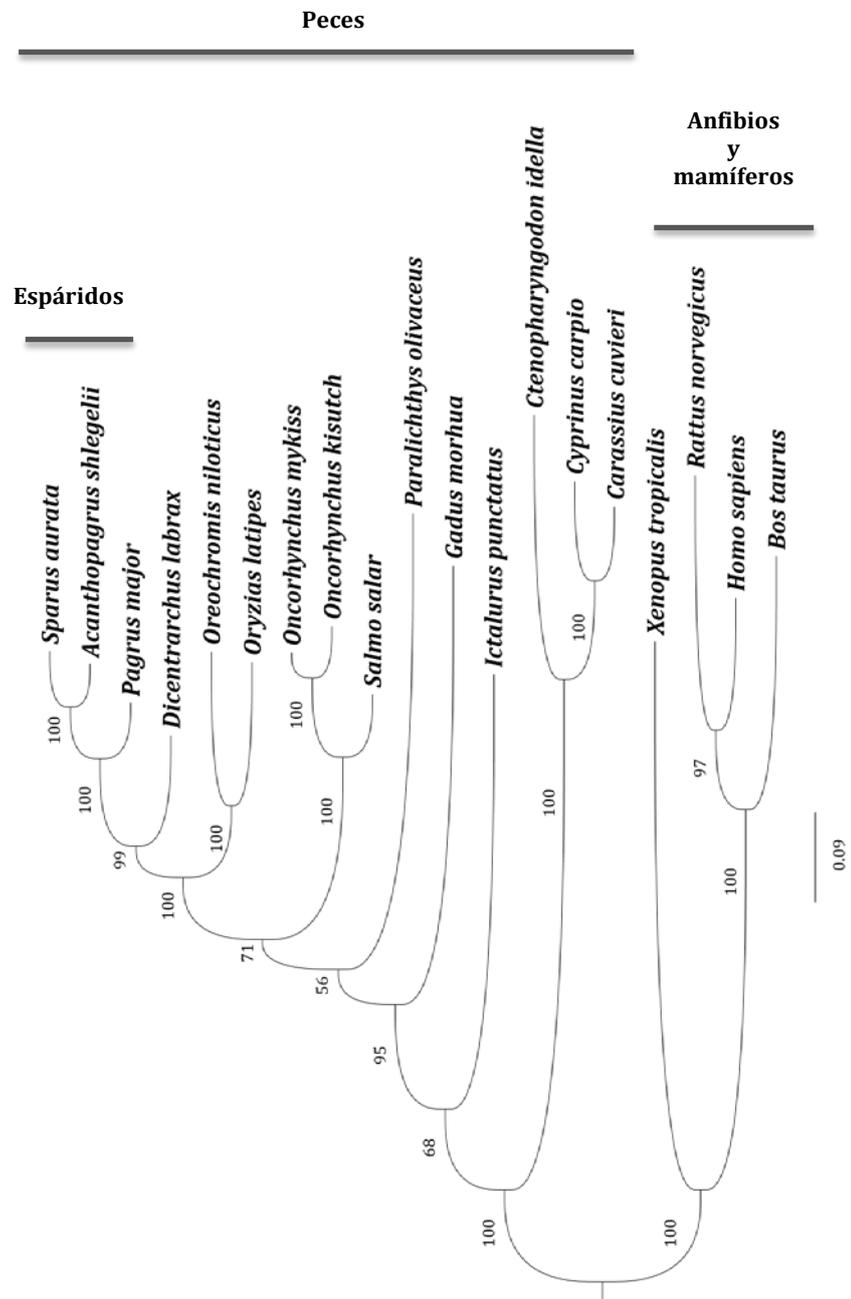
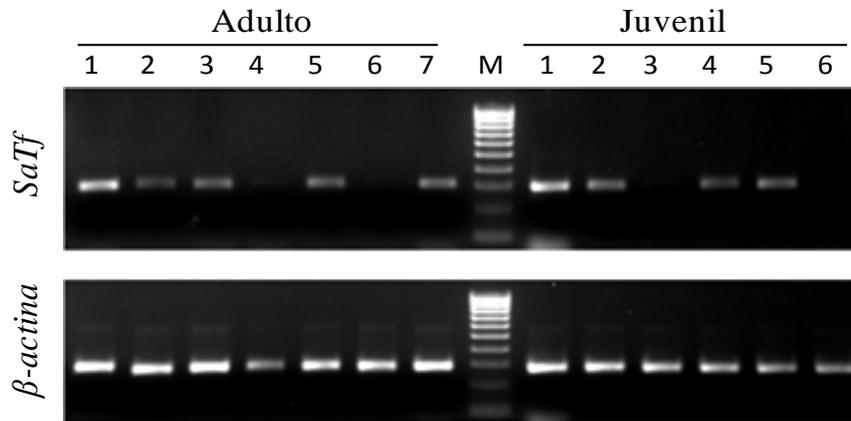


Figura 5g. Resultado RT-PCR obtenido para el gen *Tf* en diferentes tejidos recolectados a partir de individuos adultos y juveniles, visualizado en gel de agarosa al 1,5%, empleando el gen β -actina (Tiago y col., 2008) como gen control.



[1] Hígado; [2] riñón; [3] Bazo; [4] estómago; [5] cerebro; [6] corazón; [7] Ovario (solamente en adultos); [M] Marcador de peso molecular (patrón de 10 bandas separadas regularmente cada 100pb, desde las 100pb hasta las 1000pb.)

6. Capítulo IV

*Búsqueda de SNPs funcionales en genes
candidatos para crecimiento*

6.1. Introducción

A consecuencia del gran avance tecnológico experimentado durante las últimas décadas en el campo de la genómica y al gran interés que existe por establecer dentro de las especies económicamente relevantes para la industria acuícola, conexiones genotipo-fenotipo que permitan llevar a cabo una identificación prematura de individuos con fenotipo deseado, multitud de estudios de investigación centran su objetivo en la identificación y caracterización de genes implicados en funciones fisiológicas asociadas a caracteres de interés productivo, generando con ello un recurso de gran utilidad para la búsqueda de polimorfismos funcionales. Tal y como se ha mencionado en capítulos anteriores, entre los caracteres de interés productivo el crecimiento ocupa una posición preferencial; por ello, dada la importancia económica de la dorada en acuicultura, no es de extrañar que en la base de datos pública del NCBI se encuentren disponibles para esta especie numerosas secuencias de genes vinculados a la regulación del crecimiento somático, disponiéndose incluso de un transcriptoma específico de músculo esquelético que contiene información referente a factores de transcripción, moléculas de señalización y proteínas estructurales implicadas en el proceso de formación y crecimiento del tejido muscular (García de la Serrana y col., 2012).

Los peces teleósteos experimentan un patrón de crecimiento peculiar, con un incremento en la longitud y masa corporal continuado a lo largo de su vida (Johnston y col., 2011), presentando diferencias notables en el proceso de formación del tejido muscular somático respecto a mamíferos. Sin embargo, los genes implicados en la regulación del crecimiento se encuentran altamente conservados a lo largo de la escala evolutiva que comprende desde peces teleósteos hasta mamíferos (Johnston y col., 2003).

Aunque el crecimiento, como buen carácter cuantitativo, se define como un carácter poligénico gobernado por el efecto acumulativo de numerosos genes, se sabe que el efecto de estos genes no es proporcional, y algunos de ellos poseen la capacidad de ejercer una influencia superior sobre el control del carácter, lo cual les convierte en el blanco perfecto para comenzar la búsqueda de polimorfismos funcionales relacionados con el crecimiento. De-Santis y Jerry (2007) clasifican a estos genes candidatos en dos bloques bien diferenciados, correspondientes con el eje hormonal somatotrópico y los factores de crecimiento transformante, los cuales forman dos redes génicas bien diferenciadas conectadas entre sí, encargadas respectivamente del control endocrino del crecimiento y del control genético del desarrollo muscular. Los genes pertenecientes al eje hormonal somatotrópico codifican para hormonas que, en base a una señal externa provocada por algún factor ambiental y recogida por la glándula pineal,

conforman una cascada de señalización vía hipotálamo-glándula pituitaria capaz de transformar dicha señal en una respuesta de estímulo o inhibición de crecimiento dirigida a determinados tejidos diana, entre los que destaca el músculo. Ante la señal enviada por el eje hormonal somatotrópico, el tejido muscular reacciona a través de los factores de crecimiento transformantes estimulando o bloqueando la diferenciación de células musculares embrionarias (mioblastos) hacia fibras musculares (De-Saint y Jerry, 2007). Inicialmente, el crecimiento muscular se fundamenta en un proceso de hiperplasia mediante el cual se incrementa el número de fibras musculares (Rowlerson y col., 1995), y una vez alcanzado el número máximo de fibras el proceso se bloquea, y solamente se activará de nuevo en caso de que el tejido sufra algún daño (Rowlerson y col., 1997). A partir de este momento, el crecimiento queda exclusivamente confinado a un proceso de desarrollo hipertrófico que conduce a un incremento en la masa muscular por medio del engrosamiento y alargamiento de las fibras musculares (Rowlerson y col., 1995).

El tejido muscular ocupa la mayor parte del cuerpo de los peces, llegando a representar aproximadamente el 65% de su masa corporal (López-Albors y col., 1998), y actúa como reservorio de energía para el organismo, función que se hace claramente visible durante el periodo de reproductivo donde el elevado gasto energético que conlleva el desarrollo gonadal, se traduce en una notable disminución de la masa muscular de los individuos (Johnston y col., 2011).

En base a la importancia del crecimiento en la mejora de la producción, y al elevado número de genes de secuencia conocida propuestos como genes candidatos para crecimiento, este capítulo centra su primer objetivo en la creación de un panel de SNPs de dorada, localizados específicamente en genes candidatos para crecimiento. Siguiendo las recomendaciones de De-Santis y Jerry (2007), se seleccionó un total de 11 genes previamente descritos en dorada, entre los cuales se incluyen siete genes asociados al control genético del eje hormonal de crecimiento somático, mientras que los cuatro genes restantes corresponden a factores de crecimiento que intervienen en el control genético a nivel muscular. Dentro de los genes vinculados al eje hormonal somatotrópico, se seleccionó el gen *GH* que codifica para la hormona de crecimiento (GH), la cual desarrolla un papel básico en la estimulación del crecimiento somático en vertebrados, y concretamente, dentro del grupo de los peces, su función se correlaciona con el crecimiento lineal, la conversión del alimento y el apetito (Almuy y col., 2000), además de participar en otras funciones fisiológicas como la osmoregulación (Laiz-Carrion y col., 2009). Esta hormona, sintetizada en la glándula pituitaria desarrolla su acción modulando la expresión de numerosos genes por medio de una cascada de procesos de fosforilación desencadenados tras su unión con los receptores específicos GHR-

I y GHR-II (receptor de la hormona de crecimiento tipo I y II, respectivamente) situados en la superficie celular (Argetsinger y col., 1996). El gen *GHR-I* responsable del receptor GHR-I, localiza su expresión más intensa en el hígado, órgano que a su vez actúa como principal productor del factor de crecimiento insulínico tipo I (IGF-I), encargado de mediar el efecto de GH en numerosos procesos de crecimiento (Pérez-Sánchez y Le Bail, 1999). Por su parte, el gen *GHR-II* presenta una expresión mucho más limitada en el tejido hepático, sin embargo ante condiciones de privación de alimento presenta en el músculo niveles de expresión variables, que permiten especular sobre su posible participación en el mantenimiento de la supervivencia del organismo al controlar la regulación del crecimiento muscular bajo condiciones no óptimas (Saera-Vila y col., 2005). En base a lo expuesto anteriormente, no es de extrañar que tanto los genes *GHR-I* y *GHR-II*, como *IGF-I*, se presenten como dianas de interés para este estudio.

Por otro lado, también se seleccionaron los genes *PRL* y *SL*, codificantes para las hormonas prolactina y somatolactina, respectivamente. Estas dos hormonas pertenecen a la misma familia de GH, ya que aunque sus rutas de expresión y secreción difieren entre sí, todas ellas se sintetizan en la glándula pituitaria, comparten una estructura génica similar y se encuentran bajo el control del mismo factor de transcripción (Pit-1) (Astola y col., 2005). A esta familia de hormonas se les considera posibles transformadores de las señales ambientales en una señal molecular legible para el organismo (Vargas-Chacoff y col., 2009). Dentro de los vertebrados a la hormona PRL se le atribuyen más de 300 funciones biológicas, organizadas en las categorías de reproducción, crecimiento, desarrollo, regulación del balance iónico, inmunoregulación y comportamiento (Bole-Feysot y col., 1998). Sin embargo para la hormona SL, en peces teleósteos todavía no se conoce un papel funcional bien definido, pues aunque su actividad se ha vinculado a diversas funciones, estas varían de unas especies a otras. Entre estas funciones, dada la relación existente entre SL y GH, algunos autores como Sánchez-Ramos y col. (2012) clasifican al gen responsable de esta hormona como un posible gen candidato para crecimiento.

El último gen analizado en relación al eje hormonal somatotrópico es el gen *Tf*, caracterizado en el *Capítulo III*. Aunque en peces la relación entre transferrina y crecimiento ha pasado bastante desapercibida, y los únicos datos disponibles se remontan a citas antiguas (Gary, 1977; Stajkov y Yablanski, 1992), en mamíferos diversos estudios propone a este gen como marcador potencial para estudios destinados a la mejora de las características de crecimiento (Nyström y col., 1997; Kmiec M., 1999a y b). No hay que olvidar que pese al gran interés que despierta la vinculación de Tf con el sistema inmune innato en peces, su función principal se centra en transportar el hierro de forma segura a través de organismo, actuando en consecuencia como transportador de un elemento indispensable para el metabolismo, el

cual además posee cierta influencia sobre el crecimiento, pues se ha demostrado que un suministro complementario de hierro en la dieta de los peces puede condicionar su crecimiento (Noga, 2000). Además, de acuerdo a los resultados obtenidos por Weinzimer y col. (2001) en humanos, Tf podría estar conectada directamente con el eje hormonal GH/IGF-I, al comportarse como una proteína de unión a IGFs (proteínas de unión a IGFs) las cuales además de encargarse del transporte de IGFs a través del plasma e incrementar su vida media, poseen la capacidad de inhibir o potenciar el crecimiento, dificultando o favoreciendo respectivamente, la unión de estas hormonas a su receptor (Conover y Powell, 1991; Conover, 1992). La unión TF-IGFBPs muestra importantes consecuencias sobre la proliferación y la apoptosis celular, y dado que dicha unión se encuentra influenciada por los niveles de hierro libre (unión hierro dependiente), esta interacción podría representar un sistema de control del crecimiento celular condicionado, una vez más, por el estatus nutricional del organismo (Weinzimer y col., 2001).

Dentro de los factores de crecimiento producidos en músculo, se fijaron para este estudio como diana los genes *MSTN-1*, *Myog*, *MyoD-1* y *MyoD-2*, codificantes para miostatina (isoforma tipo 1), miogenina, y los factores de diferenciación miogénica tipo 1 y 2, respectivamente. La *MSTN-1*, también conocida como factor de diferenciación de crecimiento tipo 8 (GFD-8), pertenece a la superfamilia de factores de crecimiento transformante beta (TGF- β), y actúa como un regulador negativo del crecimiento del músculo esquelético. Mutaciones localizadas en la región codificante de *MSTN-1*, conducen a incrementos importantes en el nivel de masa muscular dando lugar al fenotipo conocido como doble-músculo descrito en mamíferos (McPherron y col., 1997; Clop y col., 2006; Schuelke y col., 2004; Mosher y col., 2007) y pez cebra (*Danio reiro*) (Lee y col., 2009). No obstante, a diferencia de los mamíferos donde la expresión del gen *MSTN-1* se encuentra confinada exclusivamente en el tejido muscular, los peces registran para esta isoforma de *MSTN* un patrón de expresión ubicuo (Maccatrozzo y col., 2001), indicando que dentro de este grupo su acción fisiológica podría ir más allá de una actuación muscular.

Los tres genes restantes (*Myog*, *MyoD-1* y *MyoD-2*) forman parte de la familia de factores de regulación miogénica (MFRs), encargados de la determinación y formación de las células del tejido muscular durante el desarrollo embrionario y durante el mantenimiento y reparación del músculo esquelético maduro (Gilbert, 2000). Estos factores se caracterizan por poseer un dominio estructural hélice-bucle-hélice (HLH, helix-loop-helix), que por un lado posibilita la unión a motivos específicos del ADN permitiendo a estos factores regular la expresión de otros genes (Murre y col., 1989; Chaudhary y Skinner, 1999), y por otro lado, les permiten formar heterodímeros con otras proteínas HLH e inactivar así su actividad como

factores de transcripción (Cabrera y col., 1994). A diferencia de otros factores incluidos dentro de la familia MFRs, los factores MyoD y Myog presentan una expresión diferencial en el tiempo, y no desarrollan funciones redundantes. Concretamente, los genes *MyoD-1* y *MyoD-2*, se expresan de forma muy temprana en el desarrollo y son un pieza clave para la iniciación del programa de desarrollo del músculo esquelético en el embrión, mientras que el gen *Myog* registra una expresión más tardía y su función se asocia a procesos de diferenciación y maduración de las fibras esqueléticas (Tan y Du, 2002).

En una segunda fase dentro de este mismo estudio, utilizando como material biológico los lotes reproductores y descendencias caracterizados en el *Capítulo II*, se plantea como objetivo final, evaluar si alguna de las variantes polimórficas incluidas en el panel de SNPs desarrollado en el objetivo anterior, se correlaciona con una variación en el peso corporal de los individuos, permitiendo clasificar a dicho polimorfismo como un marcador potencial para estudios de crecimiento en la especie.

6.2. Materiales y Métodos

6.2.1. Muestras biológicas

Para la búsqueda de variaciones SNP en genes candidatos para crecimiento, se seleccionó un conjunto de 30 individuos con mínimo grado de parentesco (pedigrí conocido), procedentes de las descendencias de los lotes de cultivo C (N=10) y D (N=10) descritas en el *Capítulo II* (apartado 4.2.1), y de un tercer lote denominado lote E (N=10).

Los análisis de asociación entre el polimorfismo detectado y el peso corporal de los individuos, se realizaron sobre una población de estudio generada a partir de los lotes C y D, compuesta en total por 53 parentales (Lote C: 18♀ y 9♂; Lote D: 16♀ y 10♂) y 389 descendientes derivados de las categorías de *crecimiento rápido* (Lote C: 107; Lote D: 92) y *lento* (Lote C: 103; Lote D: 87), que representan un total de 82 familias.

6.2.2. Extracción de ADN

La extracción de ADN genómico se llevó a cabo de acuerdo al protocolo descrito en el apartado 5.2.2.2 del *Capítulo III* (pág. 103). Con ayuda de un PicoDrop se estimó la concentración de ADN de cada muestra, fijándose para todas ellas una concentración final de 10ng/μL.

6.2.3. Búsqueda de variación SNP

Exceptuando para el gen *Tf* caracterizado previamente en el *Capítulo III*, las secuencias nucleotídicas de los genes seleccionados para el estudio se obtuvieron a partir de la base de datos pública del NCBI (Tabla 6a), y se utilizaron como molde para el diseño de iniciadores específicos utilizando el programa FastPCR (Kalendar y col., 2009). Las parejas de iniciadores diseñados se muestran en la Tabla 6b. Para los genes *Myog*, *MyoD-1*, *MyoD-2*, *GH*, *PRL*, *SL* y *MSTN-1*, se diseñaron iniciadores a lo largo de toda la secuencia nucleotídica disponible (desde la región promotora-5'UTR hasta la región 3'UTR), realizando una cobertura máxima de la región codificante, y evitando todas aquellas regiones ricas en repeticiones que pudieran dificultar el proceso de secuenciación. Para los genes *GHR-I* y *GHR-II*, debido a la gran longitud de su secuencia, al igual que para *IGF-I* del cual todavía no se dispone de una secuencia nucleotídica completa, el diseño de iniciadores se limitó principalmente a la región promotora y 5'UTR. Por último, para el análisis del gen *Tf* se emplearon los iniciadores diseñados previamente en el *Capítulo III* (Tabla 5c, pág. 125).

Utilizando el ADN genómico aislado para cada individuo y de acuerdo al protocolo descrito en el apartado 5.2.2.4 (pág. 104), se amplificaron, purificaron y secuenciaron los productos de PCR obtenidos para las diferentes regiones génicas. Mediante al programa BioEdit v.7.0.5.3 (Hall, 1999), a través de la visualización de los electroferogramas, se revisó la calidad (intensidad de señal y presencia de ruido/señales secundarias) de las secuencias de ADN obtenidas, y por medio del alineamiento múltiple de las secuencias obtenidas para una misma región génica en diferentes individuos, se detectó el polimorfismo SNPs (sustitución e inserción/delección de bases). Cada una de las variaciones nucleotídicas detectadas se describió en base a su posición en la secuencia génica y la frecuencia de sus alelos (MAF, *Minor Allele Frequency*). Las variaciones localizadas en regiones codificantes (exones), se clasificaron como sinónimas y no-sinónimas, determinándose para estas últimas el cambio de aminoácido que originan en la correspondiente secuencia proteica.

A partir del conjunto de las posibles variaciones SNP identificadas, mediante el programa HaploSNPer (Tang y col., 2008) (utilizando los parámetros establecidos por defecto) se seleccionaron todos aquellos SNP clasificados como SNPs “*reales o más probables*”. HaploSNPer selecciona los SNPs atendiendo principalmente a dos factores: (1) la redundancia del polimorfismo, y (2) la calidad obtenida para la secuencia donde se localiza el SNP. Frecuentemente, la calidad de las secuencias obtenidas se ve disminuida hacia los extremos 5' y 3', siendo posible que gran parte de la variación detectada en los extremos de las secuencias amplificadas puedan ser debidas a simples errores de secuenciación, y por lo tanto, la probabilidad de encontrar un SNP “*real*” en estas regiones será menor que en el resto de la secuencia donde los niveles de calidad resultan superiores.

6.2.4. Determinación del genotipo

6.2.4.1 Genotipado masivo de SNPs

El proceso de genotipado masivo se desarrolló en el nodo del Centro Nacional de Genotipado localizado en el hospital clínico de Santiago de Compostela (CEGEN-ISCIH USC), empleando la tecnología ILPEX™ Gold MassArray® de Sequenom. La principal característica de esta tecnología, es la utilización de la espectrometría de masas MALDI-TOF (*Matrix Assited Laser Desorption/Ionization-Time of Flight-Mass Spectrometry*), que permite realizar una discriminación alélica de variantes puntuales del genoma de tipo SNP, así como, inserciones y deleciones (con longitud máxima inferior a 40pb) en función de su masa molecular.

A partir del panel final de SNPs, el sistema diseña diversas combinaciones, seleccionándose para el estudio el diseño que implicaba el genotipado del mayor número de

SNPs utilizando el menor número de ensayos. Brevemente, el proceso de genotipado realizado se basó en las fases siguientes:

- Amplificación de las muestras-. Las muestras se someten a una PCR multiplex dirigida a la amplificación de las regiones que albergan las variaciones de interés. Para el diseño de los correspondientes iniciadores se utilizó el programa MassArray Assay designer v4.0., obteniéndose una lista de iniciadores agrupados de acuerdo a sus características. Las concentraciones finales de los reactivos de PCR empleados para un volumen final de 5 μ L de reacción consistieron en: 1,25X de tampón de PCR, 1,625mM de la solución MgCl₂, 500 μ M de la mezcla de dNTPs, 0,5U del enzima Hostar Taq®, 100nM de la mezcla de iniciadores y 20ng de ADN. El programa de PCR se basó en una primera fase de desnaturalización de 15 minutos a 94°C, seguida de una segunda fase de amplificación compuesta por 45 ciclos de 20 segundos a 94°C, 30 segundos a 56°C y un minuto a 72°C, finalizando el programa con un fase de extensión de 3 minutos a 72°C.
- Purificación-. Mediante el enzima SAP (*Shrimp Alkaline Phosphatase*), que hidroliza el hidroxilo (-OH) del grupo fosfato en el extremo 5' de los dNTPs, se desactivan los dNTPs no incorporados durante la reacción de PCR anterior, evitando así, que la ADN polimerasa pueda utilizarlos en fases posteriores. La reacción de purificación se llevó a cabo empleando una concentración final 10X para el tampón de reacción y 1,7U/ μ L del enzima SAP, para un volumen total de 2 μ L. La inactivación de los nucleótidos se obtuvo mediante una primera incubación a 37°C durante 40 minutos, seguida de una segunda incubación a 85°C durante 5 minutos, finalizando posteriormente la reacción a 4°C.
- Discriminación alélica-. La identificación de variantes alélicas se realizó por medio de una reacción de extensión de una única base (*Single Base Extension*, SBE). Para ello se emplearon los cuatro tipos de nucleótidos terminadores modificados, que incrementan la diferencia de masa entre los nucleótidos, evitando así un solapamiento de sus picos en el espectro de masas (para que no exista solapamiento entre picos, las masas deben diferenciarse como mínimo en 15Da) y permite cubrir el espectro de masas de forma amplia, de manera que todos los SNPs del ensayo puedan ser representados en un solo espectro (Figura 6a). El diseño de iniciadores de extensión se realizó mediante el programa MassArray designer v4.0, sin embargo en esta ocasión, para cada variante se diseñó solamente un único iniciador. La reacción de extensión se desarrolló para un volumen final de 2 μ L utilizando las siguientes concentraciones finales de los reactivos ILPEX: tampón de reacción ILPEX plus 10X, mezcla de terminadores ILPEX 0,5X, mezcla

de iniciadores 0,625 μ M (iniciadores de bajo peso molecular) y 1,25 μ M (iniciadores de alto peso molecular), y enzima ILPEX a 0,5X. El programa de extensión consistió en una fase inicial de 30 segundos a 94°C, seguida de una fase de amplificación compuesta por 40 ciclos de 5 segundos a 94°C, 5 segundos a 52°C y 5 segundos a 80°C, finalizando con una fase de extensión de 3 minutos a 72°C.

- ***Lectura de datos***-. Los productos obtenidos en la reacción anterior, se purificaron mediante 6mg de una resina de limpieza y 16 μ L de agua Mili-Q, eliminando la presencia de sales externas que puedan interferir en el posterior análisis de espectrometría de masas. Para el análisis mediante MALDI-TOF (modelo MA4 de Sequenom) los productos de la reacción se transfirieron a un SpectroChip utilizando un robot RS100. Finalmente, los datos obtenidos se analizaron mediante el programa MassArray TYPER Analyzer v4, revisando la calidad de los espectros y los correspondientes controles (concordancia entre muestras duplicadas, controles negativos, porcentaje de genotipado por muestra y SNP, entre otros).

6.2.4.2. Genotipado para deleciones/inserciones de gran tamaño

Entre las variaciones polimórficas identificadas, se aislaron deleciones/ inserciones (*ILPs, Intronic Length Polymorphisms*) cuyo tamaño superaba las 40pb admitidas por la tecnología ILPEX™ Gold MassArray® de Sequenom. Su amplificación se realizó utilizando las parejas de iniciadores PRL_F3/PRL_R3 y PRL_F4/PRL_R5 (Tabla 6b). El genotipo se determinó mediante la visualización de los productos de PCR obtenidos en un gel de agarosa al 2% con bromuro de etidio, empleando como controles positivos individuos previamente genotipados como: (1) homocigoto para la variante larga, (2) homocigoto para la variante corta y (3) heterocigoto.

6.2.5. Identificación de haplotipos y análisis de desequilibrio de ligamiento

A partir del genotipo obtenido para cada una de las variantes analizadas (SNPs e inserciones/deleciones), para los genes *MyoD-1*, *MyoD-2*, *PRL*, *MSTN-1* y *Tf*, donde se analiza más de una posición polimórfica por gen, mediante el programa Phase v.2.0.2 (Stephens y Donnelly, 2003), aplicando un método de análisis basado en estadística bayesiana, se llevó a cabo una reconstrucción de haplotipos. Para cada individuo se estableció la pareja de haplotipos (diplotipo) más probable de cada gen, y su correcta asignación se validó a nivel de familia por medio de la comparación del trío padre-madre-descendiente, identificando para cada descendiente el haplotipo procedente de cada parental.

Para cada variante nucleotídica se determinó mediante el programa Haploview v.4.2 (Barrett y col., 2005) la frecuencia del alelo menos común (MAF, *Minor Allele Frequency*), y se evaluó su ajuste al HWE aplicando una prueba χ^2 con un grado de libertad bajo un intervalo de confianza del 95%. Utilizando este mismo programa, se evaluó el desequilibrio de ligamiento entre SNPs localizados en un mismo gen, a partir de las estimas de los parámetros D' (coeficiente de desequilibrio de ligamiento normalizado) (Lewontin, 1964) y r^2 (coeficiente de correlación al cuadrado) (Pritchard y Przeworski, 2001). El parámetro D' mide la asociación entre dos SNPs e informa cómo de aleatoria es su distribución, es decir, estima su herencia conjunta. Considerando un haplotipo compuesto por dos alelos (A y B) de dos SNPs diferentes, la estimación de D' se obtiene por medio de la siguiente fórmula:

$$D'=D/D_{max} \text{ (Lewontin, 1964)}$$

$$D= p(AB)-p(A) \times p(B) \text{ (Lewontin y Kojima, 1960)}$$

donde $p(AB)$ representa la frecuencia observada del haplotipo, $p(A) \times p(B)$ representa la frecuencia esperada para el haplotipo en ausencia de desequilibrio de ligamiento, y D_{max} es el valor máximo que puede alcanzar D dadas las frecuencias alélicas. Los valores de D' oscilan entre 0 y 1, reflejando $D'=0$ una independencia completa entre los dos SNPs, y $D'=1$ el ligamiento completo.

Por su parte, el parámetro r^2 mide el valor de redundancia estadística que supone analizar dos SNPs, y se estima mediante la siguiente fórmula:

$$r^2= D^2/p_A \times q_A \times p_B \times q_B \text{ (Pritchard y Przeworski, 2001)}$$

donde D representa el coeficiente de desequilibrio de ligamiento, y p q las frecuencias alélicas de los SNPs A y B. Los valores de r^2 oscilan también entre 0 y 1, indicando un valor igual a $r^2=1$ que un SNP actúa como perfecto marcador del segundo SNP, resultando redundante el genotipado de ambos marcadores. A diferencia del parámetro D' , donde $D'=1$ se obtiene siempre que el valor del alelo de menor frecuencia de un locus se encuentre asociado a uno de los dos alelos del otro locus, el parámetro r^2 es más limitante, pues solamente tomará el valor de $r^2=1$ siempre y cuando los dos alelos de un SNP se encuentren en completo desequilibrio de ligamiento con los alelos del otro SNP.

Para estimar si los diferentes genes analizados se encuentran en desequilibrio de ligamiento, se empleó el Test-G o Prueba de la Razón de Verosimilitud utilizando el algoritmo de las cadenas de Márkov de Raymond y Rousset (1995) (condiciones: dememorization: 10.000; batches: 200; iterations: 5.000), incluido el programa GENEPOP v.4 (Rousset, 2008).

6.2.6. Análisis estadístico

Para estimar la asociación entre variación genética (SNPs e ILPs) y diferencias fenotípicas en el peso corporal de los individuos, se realizó un análisis de la varianza (ANOVA) de un factor, contrastando la hipótesis de igualdad de medias entre los genotipos/diplotipos comparados, bajo un nivel de significación de 0,05. Todos aquellos contrastes donde se rechazó la hipótesis nula de igualdad de medias, se sometieron a un segundo análisis estadístico basado en el test de Duncan, consistente en un análisis de comparaciones múltiples *post hoc* o *a posteriori*, donde, de forma similar a la prueba de Turkey empleada en el *Capítulo III* (apartado 5.2.1.6, pág. 106), por medio de una comparación de medias dos a dos, se trata de identificar entre cuales de los grupos comparados reside realmente la diferencia de medias detectada previamente en el ANOVA. La principal diferencia del test de Duncan respecto a otras pruebas *post hoc* de similar índole, radica en que esta prueba utiliza un valor umbral cambiante, el cual depende del número de medias comprendido entre las medias a comparar. Para esta prueba se considera que dos grupos difieren estadísticamente de forma significativa cuando se alcance un p-valor igual o inferior a 0,05. Ambas pruebas estadísticas se desarrollaron utilizando el programa estadístico SPSS v.15.0.

6.3. Resultados

6.3.1. Construcción y caracterización del panel de SNPs

A partir del diseño de 103 iniciadores combinados en 57 parejas (Tabla 6b; Tabla 5c, *Capítulo III*, pág. 125), se obtuvo una cobertura de secuencia total de aproximadamente 30.500pb por individuo (N=30), distribuidas en un total de 11 genes (Tabla 6a). La mayor cobertura génica se obtuvo para los genes *Myog*, *PRL*, *MSTN-1* y *Tf*, en los cuales se analizó más del 75% de la secuencia nucleotídica disponible (Tabla 6a).

La comparación de las secuencias obtenidas en el análisis de la población utilizada (N=30), permitió identificar 357 “*posibles*” variaciones SNP. Sin embargo, la comparación de las secuencias obtenidas con las respectivas secuencia molde disponibles en la base de datos del NCBI, localizaron otras 41 *posibles* posiciones SNP, las cuales resultaron de carácter monomórfico para los individuos analizados, pero podrían representar posiciones polimórficas en otras poblaciones. Por lo tanto, la búsqueda de variación SNP en genes candidatos para crecimiento concluyó con la anotación de 398 *posibles* SNPs.

Posteriormente, la evaluación de estos *posibles* SNPs mediante el programa HaploSNPer (Tang y col., 2008), concluyó en la creación de un panel final compuesto por un total de 164 SNPs bialélicos aceptados como SNPs “*reales*” o altamente probables, distribuidos en 10 genes candidatos para crecimiento (*GHR-I*, *GHR-II*, *IGF-I*, *MyoD-1*, *MyoD-2*, *GH*, *PRL*, *SL*, *MSTN-1* y *Tf*), ya que el filtrado realizado eliminó todos los SNPs *posibles* detectados para el gen *Myog* (Tabla 6c). Dentro del panel final, las transiciones (cambio purina-purina o pirimidina-pirimidina) representan un 52%, y las transversiones (cambio purina-pirimidina o pirimidina-purina) un 48%. Por otro lado, el 87% de estos SNPs se localizaron dentro de las regiones promotoras-5' UTR e intrónicas, mientras que las regiones exónicas y 3'UTRs albergaron solamente el 13% de la variantes (Tabla 6e).

De los nueve SNP codificantes, denominados así por estar localizados en regiones exónicas, solamente cuatro (*SaMyoD-1_C679G*, *SaMyoD-2_G1066A*, *SaMyoD-2_A2031T* y *SaGH_C2210G*) dieron lugar a una sustitución no sinónima, generando cambios en la secuencia de aminoácidos de la correspondiente proteína (Tabla 6d). Los SNP codificantes *SaMyoD-1_C679G* y *SaMyoD-2_G1066*, ocasionaron los re-emplazamientos menos conservativos, implicando un intercambio de aminoácidos con diferencias en las características principales de su cadena lateral (Tabla 6d). Por otra parte, la secuencia molde tomada de la base de datos del NCBI para el gen *MyoD-1* (GenBank ID: AF478569.1), reveló que el SNP codificante *SaMyoD-1_C679G* podría ser un SNP tri-alélico, ya que la en la secuencia molde la posición se encuentra

representada por un nucleótido de timina, mientras que en nuestra población de estudio dicha posición se encuentra ocupada por un nucleótido de citosina o guanina (Tabla 6d).

Atendiendo a la longitud de secuencia analizada, se deduce para dorada una frecuencia global de 1 SNP por cada 185pb (Tabla 6e). No obstante, esta frecuencia se modifica dentro de cada región génica, detectándose las densidades de SNPs más elevadas en regiones no codificantes (promotora/5'UTR, 3'UTR e intrones) donde los valores resultan hasta 4,8 veces superiores a la densidad detectada en exones (Tabla 6e).

El análisis individual de cada gen, mostró que las mayores diferencias nucleotídicas entre individuos se localizaron en los genes *MSTN-1*, *PRL* y *MyoD-1*, los cuales, de forma conjunta, aportaron más del 70% de los SNPs que conforman el panel final (Tabla 6e); sin embargo, la distribución de los SNPs difiere entre estos tres genes. En el caso del gen *MSTN-1*, la variación tiende a concentrarse dentro de las región promotora/5'UTR (Tabla 6e), mientras que *PRL* registra un polimorfismo más intenso en la región intrónica, y *MyoD-1* presenta para ambas regiones una tasa de variación similar (Tabla 6e). En contraposición, los genes *GHR-I* y *SL*, mostraron secuencias altamente conservadas, aportando cada gen un único SNP al panel final (Tabla 6e).

6.3.2. Identificación de ILPs

Además de los SNPs detectados, los análisis de secuencia realizados, permitieron identificar la existencia de tres variaciones ILPs consistentes en inserciones/delecciones de gran tamaño que producen polimorfismos en la longitud de los intrones. De los tres ILPs detectados, dos se localizaron en el gen *PRL* (Figura 6b), afectando el primero de ellos (*SaPRL_ILP1524-1621*) a una región de 98 nucleótidos del primer intrón, mientras el segundo ILP (*SaPRL_ILP2947-2999*) implicó una variación de 53pb en el tercer intrón. El tercer ILP se localizó en el primer intrón del gen *MyoD-2* (*SaMyoD-2_ILP1954-1976*) y causó una variación de tamaño de 23pb (Figura 6b).

6.3.3. Diseño y validación de ensayos de genotipado masivo de SNPs e ILPs

Mediante la tecnología ILPEX™ Gold MassArray® de Sequenom, a partir del panel final de 164 SNPs y el ILP *SaMyoD-2_ILP1954-1976* cuyo tamaño se encuentra dentro del rango óptimo de análisis de esta tecnología, se generó un diseño de genotipado masivo optimizado para 75 SNPs distribuidos en tres ensayos (SaOVI-1, SaOVI-2 y SaOVI-3)(Tabla 6f).

De los tres ensayos diseñados, se validaron los SaOVI-1 y SaOVI-2 en una población de estudio compuesta por 442 individuos organizados en 82 familias, obteniéndose un porcentaje

de genotipado superior al 98%. De los 54 SNPs analizados (27 por ensayo), tres SNPs (*SaGHR-II_C751T*, *SaMSTN-1_A760G* y *SaPRL_C994G*) presentaron fallos en su amplificación por PCR, cuatro (*SaMSTN-1_A93T*, *SaGH_C2210G*, *SaGHR-II_C275T* y *SaPRL_INS1002C*) resultaron monomórficos, y seis (*SaGH_C4705T*, *SaPRL_C565T*, *SaPRL_C2050G*, *SaMSTN-1_A180T*, *SaMyoD-1_A2450T* y *SaMyoD-1_C2524T*) presentaron un patrón de genotipado anómalo que dificulta el proceso de asignación de genotipos debido a la presencia de grupos de genotipos indefinidos situados de forma intermedia entre los genotipos correspondientes al homocigoto1, heterocigoto y homocigoto2 (Figura 6c), validándose finalmente un total de 41 SNPs.

De forma independiente, se determinó el genotipo de los ILPs *SaPRL_ILP1524-1621* y *SaPRL_ILP2993-3039* para el mismo conjunto de individuos.

Las frecuencias alélicas obtenidas para cada uno de los 43 marcadores analizados (SNPs+ILPs) en los lotes de reproductores C y D, se muestra en la Tabla 6g.

6.3.4. Reconstrucción de haplotipos y análisis de desequilibrio de ligamiento

La relación entre el número de posiciones polimórficas analizadas, y los haplotipos y diplotipos detectados por gen para cada lote se muestran en la Tabla 6h. Ambos lotes mostraron una diversidad haplotípica similar para los genes *MyoD-1*, *MyoD-2*, *Tf*, *PRL* y *MSTN-1*, con haplotipos específicos de lote en *MyoD-1*, *PRL* y *MSTN-1* (Tabla 6i). Todos los haplotipos asignados en la generación descendiente se correspondieron con haplotipos previamente identificados en la generación parental. Por medio de la comparación de la información haplotípica obtenida para el trío descendiente-parental1-parental2, se estableció para cada descendiente el haplotipo de origen materno y paterno que conforman su diplotipo, y se determinó que el porcentaje de errores de la asignación parental realizada a partir de la información de los *loci* microsatélites incluidos en *OVIDOPLEX* (Capítulo II, apartado 4.2.2) se situaba dentro del 5% aceptado por el programa CERVUS v.3.0 (Marshall y col., 1998; Kalinowski y col., 2007) como tasa de error en el proceso de asignación.

No se detectaron desequilibrios de ligamiento entre los genes analizados, pero sí se establecieron fuertes relaciones de ligamiento entre SNPs situados dentro de un mismo gen, identificándose bloques haplotípicos caracterizados por altos valores de D' y r^2 , en los genes *MyoD-1*, *Tf* y *PRL* (Figura 6d).

6.3.5. Asociación entre SNPs/ILPs y el peso corporal

De forma individual, ninguno de los marcadores analizados (SNPs e ILPs) mostró una

correlación significativa con las variaciones en el peso de los individuos para toda las categorías de crecimiento (C18S, C18F, D18S y D18S), como máximo se detectaron diferencias significativas para una única categoría en un único lote, tal y como se observa para el SNP *SaGH_C411T* (Tabla 6j). Sin embargo, en los genes donde se dispone de más de un polimorfismo por gen, cuando la relación entre gen y carácter se evaluó utilizando la información derivada de los diplotipos (como mínimo cada diplotipo comparado registró descendientes procedentes de dos familias sin parentesco y varias familias de medios hermanos) fijados para cada uno de ellos, se detectaron asociaciones significativas en *PRL*, *MyoD-1* y *MyoD-2* (Tablas 6k y 6l). En *PRL* el diplotipo GCDT/GCDT formado a partir de dos SNPs (*SaPRL_G489C* y *SaPRL_C707T*) con un fuerte desequilibrio de ligamiento (Figura 6d) situados en la región promotora, un ILP de gran tamaño en el primer intrón (*SaPRL_ILP1524*) y un SNP localizado en el tercer intrón (*SaPRL_C2958T*), respectivamente, presentó para las cuatro categorías de crecimiento (C18F, C18S, D18F y D18S) un peso corporal medio significativamente superior respecto a otros diplotipos (Tabla 6k), provocando incrementos en el peso medio que varían entre 10% (C18S) y el 26% (D18S). En el caso de *MyoD-1*, el diplotipo DGG/DGG (D hace referencia a una delección) derivado del agrupamiento de tres SNPs localizados en el segundo intrón (*SaMyoD-1_D2100A*, *SaMyoD-1_A2143G* y *SaMyoD-1_A2404G*), difiere significativamente para las categorías C18F, C18S y D18S de otros diplotipos, y mostró el peso medio más bajo de todos diplotipos comparados, con reducciones máximas que oscilan entre el 21% (D18S) y 24% (C18F) (Tabla 6l). En *MyoD-2*, el diplotipo ACA/CCA generado a partir de la combinación de tres SNPs distribuidos en la región promotora (*SaMyoD-2_A785C*), primer intrón (*SaMyoD-2_C1982T*) y segundo exón (*SaMyoD-2_A2031T*), mostró también diferencias significativas con otros diplotipos dentro de las categorías C18F, C18S y D18S (Tabla 6l), presentado en todas ellas el peso corporal medio más bajo. En esta ocasión, las reducciones máximas de peso variaron entre 14% (C18S) y el 18% (D18S).

En el caso de los genes *MSTN-1* y en *Tf*, las numerosas comparaciones realizadas para cada gen entre múltiples combinaciones de diferentes SNPs, no reflejaron correlación significativa alguna con el peso corporal (datos no mostrados).

6.4. Discusión

6.4.1. Panel de SNPs en genes candidatos para crecimiento

La búsqueda de variación SNP concluyó con la anotación de 398 SNPs localizados en 11 genes candidatos para crecimiento. Sin embargo, el proceso de filtrado realizado posteriormente excluyó aproximadamente al 58% de la variación anotada inicialmente, reduciendo el tamaño del panel final a 164 SNPs *reales* pertenecientes a 10 genes (Tabla 6c).

A pesar de la importante reducción generada sobre el número de posiciones polimórficas que pasan a formar parte del panel final, el criterio selectivo de SNPs aplicado se considera un paso indispensable para la obtención de paneles que posean una capacidad resolutive máxima en la caracterización de la población de estudio (Hubert y col., 2010), ya que favorece la incorporación de posiciones polimórficas reales con una frecuencia de aparición elevada, minimiza la incorporación de polimorfismos *reales* que debido su baja frecuencia pueden resultar poco informativos, y a su vez, reduce la probabilidad de incorporar falsos positivos originados a partir de errores derivados de la propia técnica de secuenciación.

La densidad de SNPs deducida para dorada en este estudio en base a los 164 SNPs que conforman el panel final (1SNP/185pb), es equivalente a la densidad descrita por Hemmer-Hansen y col. (2011) en un estudio similar desarrollado a partir de poblaciones naturales en bacalao (*Gadus morhua*) (1SNP/111pb). No obstante, la deducción de estas densidades se encuentra fuertemente condicionada por el tipo de genes o región del genoma analizada, así como por el conjunto de individuos empleados para la detección de SNPs. Referente a los genes analizados, nosotros detectamos niveles de variación nucleotídica intragénica diferentes, presentando los genes *MyoD-1* y *MSTN-1* las densidades de SNPs más elevadas, alcanzando frecuencias superiores a un SNP cada 100pb, mientras que en contraposición, el gen *SL* registró una secuencia altamente conservada (Tabla 6d); por lo tanto, parece evidente que dependiendo de los genes analizados la frecuencia global estimada para la especie puede sufrir modificaciones. El segundo factor a considerar nos permite especular sobre una posible infravaloración de la frecuencia de SNPs estimada, pues si bien la utilización de individuos procedentes de lotes de cultivo diferentes constituidos a partir de individuos con un origen heterogéneo y un grado de parentesco mínimo, tiende a favorecer la diversidad nucleotídica, el limitado número de individuos que componen dicha población muestral (N=30), puede provocar que por azar, la constitución genotípica de los individuos seleccionados no represente a la población global, y SNPs reales con una frecuencia de aparición moderada pueden ser excluidos del estudio al clasificarse incorrectamente como SNPs poco probables.

La mayoría de los SNPs detectados se localizaron en regiones génicas no codificantes (Tabla 6e), concordando con los resultados publicados en otras especies acuícolas como bacalao (*Gadus morhua*) (Moen y col., 2008; Hubert y col., 2009, 2010; Hemmer-Hansen y col., 2011), rodaballo (*Scophthalmus maximus*) (Vera y col., 2011) y salmón (*Salmo Salar*) (Hayes y col., 2007). No obstante, el porcentaje de SNPs codificantes obtenido en dorada (6,7%), se aleja de forma considerable de los porcentajes registrados en las especies anteriormente citadas, cuyos valores oscilan entre el 17% y el 26% (Hubert y col., 2009; 2010). Estas diferencias se pueden deber a que en todos estos estudios, a excepción de Hemmer-Hansen y col., (2011), la búsqueda de SNPs se realiza a partir de secuencias EST procedentes de un elevado número de genes, realizando una cobertura de regiones codificantes mucho más extensa que en nuestro caso; y además, algunas de estas regiones codificantes podrían encontrarse bajo una presión selectiva menor conduciendo hacia una diversidad nucleotídica superior. Por otra parte, de los 11 SNPs codificantes identificados, solamente cuatro se clasificaron como variaciones no-sinónimas al implicar un cambio de aminoácido en la proteína correspondiente (Tabla 6d). Esta proporción entre variaciones sinónimas y no-sinónimas, resulta idéntica a la establecida en bacalao por Hubert y col. (2010), quienes a su vez, proponen esta aparición preferencial de SNPs sinónimos como el resultado de una posible presión evolutiva que actuaría en contra de las variaciones no-sinónimas, a consecuencia de la capacidad que este último tipo de sustituciones posee para desencadenar mutaciones deletéreas. Sin embargo, el potencial que poseen las variaciones no-sinónimas para provocar cambios drásticos en las proteínas, es la razón por la que a pesar de su baja frecuencia en el genoma, se les considera puntos dianas claves para estudios de asociación genotipo-fenotipo, pues todas aquellas variaciones que afecten a la estructura/funcionalidad de la proteína sin resultar deletéreas para el organismo, podrían justificar directamente las variaciones fenotípicas observadas entre los individuos. En nuestro caso, de las cuatro variantes no-sinónimas detectadas (Tabla 6d), los SNPs *SaMyoD-1_C679G* y *SaMyoD-2_G1066*, ocasionaron los re-emplazamientos menos conservativos, implicando un intercambio de aminoácidos con diferencias en las características principales de su cadena lateral, no obstante, su presencia no parece repercutir sobre la estructura final y/o función de la proteína implicada, y por lo tanto, no es de esperar que se comporten como responsables directos de variaciones fenotípicas.

Por su parte, los SNPs no-codificantes aislados se localizaron principalmente en la región promotora-5'UTR (48%) y en las regiones intrónicas (39%). El bajo porcentaje obtenido en la región 3'UTR (6%) tal vez se deba a la baja cobertura de secuencia (<2000pb) realizada para esta región (Tabla 6a). Aunque los SNPs no-codificantes no repercutan en la estructura de la proteína, despiertan interés por su capacidad para provocar alteraciones en la regulación de la

expresión génica. Un claro ejemplo de ello son los SNPs localizados dentro de la región promotora cuya presencia puede producir modificaciones en las secuencias de unión a factores y complejos de transcripción y ocasionar variaciones en los niveles de expresión del gen en cuestión (De-Santis y Jerry, 2007). Otros SNPs, como algunos de los emplazados en las regiones 5' y 3'UTR, pueden afectar a la estabilidad del ARNm y ejercer un efecto a nivel post-transcripcional, repercutiendo sobre la síntesis de proteína (Garnacho y col., 2005). De igual modo, la variación identificada en los intrones, también puede repercutir sobre la expresión génica, pues en contra del pensamiento clásico que hace referencia a estas regiones de genoma como ADN "basura", se ha visto que en vertebrados, numerosos intrones codifican pequeñas moléculas de ARN (18-24 nucleótidos de longitud) no codificante, altamente conservadas entre las especies, denominadas microARNs (miARNs). Estas moléculas de ARN se unen a proteínas específicas, formando complejos capaces de reprimir o degradar a ARNs mensajeros diana por medio de su anclaje a secuencias específicas ubicadas dentro de la región 3'UTR de dichos ARNm (Okamura y col., 2004) participando así de forma activa en la regulación de multitud de procesos fisiológicos (Bartel, 2004; Lin y col., 2006; Xu y col., 2013). En la actualidad, se han identificado numerosos miARNs con expresión elevada en el tejido muscular (Xia y col., 2011; Yan X. y col., 2012; Xu y col., 2013), y Huang y col. (2012) en tilapia (*Oerochromis niloticus*) asocian 16 miARNs a la característica de crecimiento corporal, poniendo de manifiesto que estas moléculas de ARN pueden jugar un papel importante en la regulación del crecimiento de los peces. Dado que los miARNs localizan sus dianas de unión dentro de la región 3'UTR del ARNm, la presencia de polimorfismos SNPs en estas secuencias específicas, podrían ser el origen de la inestabilidad anteriormente mencionada que algunos SNPs situados en la región 3'UTR provocan sobre el ARNm. Hariharan y col. (2009), basándose en el hecho de que la estructura secundaria de la región 3'UTR del ARNm es quien parece gobernar la interacción entre miARN y su secuencia diana, proponen que además de los SNPs localizados en las dianas de reconocimiento, todas aquellas variaciones que se sitúen próximas a ellas se deberían considerar también de interés, pues su presencia podría alterar la estructura secundaria de la región 3'UTR, permitiéndoles actuar como puntos llave de apertura o cierre para acceso de los miARN a las secuencias de reconocimiento.

A su vez, los SNPs localizados en intrones, pueden afectar a las zonas de corte y empalme intrón-exón, posibilitando que a partir de un mismo precursor de ARNm se originen diferentes ARNm maduros. Así, Tao y Boulding (2003) en trucha alpina (*Salvelinus alpinus* L.), identifican en la zona de empalme del cuarto intrón con los exones cuatro y cinco del precursor de ARNm común para gen PACAP (Pituitaria Adenilato Ciclasa activada por Polipéptidos) y GHRH (Hormona liberadora de la hormona del crecimiento), un SNP (A/G) fuertemente asociado a

crecimiento, que provoca un *splicing* alternativo que origina dos ARNs mensajeros con diferente longitud. Los autores proponen que la relación entre SNP y crecimiento podría ser debida a una implicación funcional del alelo G, cuya presencia podría favorecer la expresión del ARNm *PACAP* provocando que los individuos portadores de este alelo posean una capacidad de crecimiento superior.

Por otro lado, aunque dentro de cada gen la frecuencia de SNPs más alta tiende a residir en los intrones (Tabla 6d), en el caso de *MSTN-1*, es la región promotora-5'UTR quien presenta los mayores niveles de variación (1SNP/34pb) mostrando una región promotora-5'UTR altamente polimórfica que, de acuerdo a los resultados de Nadjar-Boger y Funkenstein (2011) para el gen *MSTN-2*, parece ser una característica común para las dos isoformas *MSTN* expresadas en dorada (*MSTN-1* y *MSTN-2*).

6.4.2. Validación de SNPs e ILPs

El panel final de SNPs se compone de variantes genéticas aceptadas por herramientas bioinformáticas como posiciones polimórficas altamente probables, de manera que la aceptación de estas variaciones como SNPs se basa en cierto modo, en un criterio exclusivamente teórico. Para poder aceptar con precisión que estos SNPs representan marcadores moleculares de utilidad en la especie, se precisa un proceso de validación adicional, como puede ser la determinación de su transmisión hereditaria. En nuestro caso, mediante la tecnología ILPEX™ Gold MassArray® de Sequenom, se sometieron a análisis los ensayos SaOVI-1 y SaOVI-2 compuestos cada uno de ellos por 27 SNPs (Tabla 6f), sin embargo dentro de SpOVI-1 se eliminaron del estudio las variantes *SaGHR-IL_C751T*, *SaMSTN-1_A760G* y *SaPRL_C994G* al presentar sus iniciadores fallos de incompatibilidad no detectados durante el diseño experimental; quedando finalmente los ensayos SaOVI-1 y SaOVI-2 compuestos cada uno de ellos por 24 y 27 variantes, respectivamente. Aunque ambos ensayos se sitúan próximos al número medio de variantes incluidas por ensayo que se registra para esta técnica (N=24), se encuentran ligeramente por debajo de las 36 variantes/ensayo fijadas como su capacidad máxima de genotipado simultáneo (comunicación persona Ana Torres, CEGEN). La limitación para generar ensayos de genotipado con un grado de optimización más elevado, se fundamenta en que el número de variantes que se incorporan en cada ensayo se encuentra fuertemente condicionado por el número de variaciones que constituyen el panel inicial a partir del cual se realiza el diseño, así como también, de las características propias de la secuencia donde se localizan las variaciones. Por lo tanto, cuanto mayor sea el tamaño del panel de SNPs obtenido, mayor será la probabilidad de generar ensayos de genotipado que combinen un mayor número

de variantes, con la consiguiente reducción en el coste económico de los ensayos/análisis de genotipado múltiple de SNPs.

En nuestro caso, por medio de la comparación de trío descendiente-parental1-parental2 realizada para los 389 descendientes analizados, se validó el 76% de los SNPs analizados (Tabla 6f). Las variantes *SaMSTN-1_A93T*, *SaGH_C2210G*, *SaGHR-IL_C275T* y *SaPRL_INS1002C* se clasificaron como SNPs no validados al resultar monomórficas para el lote C y D, sin embargo, una revisión de los datos derivados del análisis de la población muestral empleada en la búsqueda de SNPs, determinó que estas variaciones se identificaban exclusivamente en individuos procedentes del lote E, cuyas descendencias no se incluyeron en el proceso de validación; de esta manera, estos cuatro SNPs podrían representar polimorfismos específicos de lote. Este resultado refleja nuevamente la importancia de diseñar paneles de SNPs que no traten de albergar el mayor número de variaciones existentes en la especie, sino que engloben solamente aquellas posiciones polimórficas que resultan de interés para los lotes o poblaciones objeto de estudio, aunque ello derive en la creación de paneles que poseen una capacidad resolutoria menor en estudios poblacionales más amplios (Hubert y col., 2010).

A su vez, los SNPs *SaGH_C4705T*, *SaPRL_C565T*, *SaPRL_C2050G*, *SaMSTN-1_A180T*, *SaMyoD-1_A2450T* y *SaMyoD-1_C2524T* aunque en este estudio se clasifican como SNPs no validados, representan posiciones polimórficas reales que se transmiten de generación en generación, pero la dificultad que plantean a la hora de asignar el genotipo a determinados individuos (Figura 6c) hace que se precise de más análisis para poder considerarlos marcadores de utilidad en la especie. Pese a que para justificar el origen de los patrones de genotipado anómalo originados por estas variantes se requiere de una caracterización más precisa de las regiones nucleotídicas que albergan a dichos SNPs, basándose en el hecho de que al menos uno de los parentales de los descendientes con genotipo anómalo también se ve afectado, se presentan dos posibles hipótesis que argumentan la aparición de estos genotipos indeterminados: (1) el propio SNP analizado representa en sí mismo un cambio nucleotídico tri-alélico no identificado en el estudio inicial realizado, o bien, (2) existe una interacción de señal entre el SNP analizado y polimorfismos localizados en sus proximidades. Ambos supuestos implicarían pequeños cambios de secuencia que ocasionarían alteraciones en el peso molecular de los alelos, y provocando que los individuos portadores de estos alelos se situasen dentro del espectro de masas en posiciones intermedias a las esperadas, tal y como se observa en nuestros resultados (Figura 6c).

En base a estos resultados, la eliminación de los SNPs no validados deja los ensayos SaOVI-1 y SaOVI-2 abiertos a un nuevo rediseño dirigido a incrementar su optimización mediante la incorporación de nuevas variantes presentes en el panel, y ofrecer un análisis de

genotipado simultáneo que implique un mayor número de variantes informativas por ensayo. Dentro de variantes del panel de SNPs, que no se incluyeron inicialmente en SaOVI-1 y SaOVI-2, se considera de especial interés la adicción de SNPs localizados en los genes *GH*, *GHR-II* y *IGF-I*, para los cuales, en este estudio, solamente dispone de un único SNP validado.

Además de los 41 SNPs, también se validaron los dos ILPs descritos para el gen *PRL* (Figura 6b). Frecuentemente, los ILPs se suele identificar a través de la técnica EPIC-PCR (*Exon-Primed Intron-Crossing*), basada en el diseño de iniciadores ubicados en regiones exónicas localizadas a ambos lados del intrón. Después de los SNPs, las inserciones/delecciones se consideran el segundo tipo de polimorfismo más frecuente en el genoma, y en el caso de los ILPs, su análisis suponen una importante ventaja a nivel de reducción de costes frente a los marcadores SNPs, ya que al implicar variaciones de gran tamaño en la longitud de los fragmentos amplificados, se convierten en un marcador fácilmente detectable mediante la simple combinación de las técnicas de amplificación por PCR y electroforesis en gel. Por otro lado, atendiendo a la tendencia general hacia la conservación exónica entre especies cercanas, el análisis basado a partir de iniciadores diseñados en estas regiones codificantes favorece que a partir de una única pareja de iniciadores se puedan estudiar especies separadas filogenéticamente. Hasta el momento los ILPs se han empleado principalmente como marcadores de gran utilidad en estudios de identificación de especies y en estudios poblacionales (Touriya y col., 2003), pero en esta ocasión, su ubicación en un gen candidato para crecimiento, permite especular sobre su posible actuación como marcadores de crecimiento. En este sentido, recientemente Hu y col. (2013) han identificado en la región 3' UTR del gen *IGF-I* de carpa común (*Cyprinus carpio* L.) una inserción/delección de 79pb, cuyos genotipos parecen marcar diferencias significativas en el crecimiento de los individuos, demostrando con ellos que las inserciones/delecciones de gran tamaño pueden actuar perfectamente como marcadores funcionales de interés, a la vez que refuerza el principio de que todo polimorfismo localizado en las regiones flanqueantes del gen (5' y 3' UTR) puede conducir a variaciones fenotípicas detectables.

6.4.3. Asociación entre SNPs/ILPs y el peso corporal

La disponibilidad de una población estructurada en cuatro grupos bien definidos como son C18F, C18S, D18F y D18S, donde no solo existe una clara diferenciación fenotípica en cuanto al peso medio corporal, sino también una diferenciación genética avalada por los valores de *Fst* estimados a partir de la información derivada de *loci* microsatélites (Capítulo II, apartado 4.3.3), ofrece la posibilidad de realizar una primera evaluación sobre la posible

asociación entre polimorfismos SNPs/ILPs y el carácter bajo cuatro contextos genéticos diferentes. Si un marcador o marcadores presentan un efecto notable sobre el peso corporal, dicho efecto se debería detectar dentro de cada categoría con independencia del valor de peso corporal medio obtenido para cada una de ellas como consecuencia del proceso selectivo aplicado. Además, dado que cada una de las categorías se desarrolló en un tanque de cultivo independiente, la evaluación de la relación polimorfismo-peso corporal a través de estas categorías, permite valorar si las asociaciones observadas pueden ser el simple reflejo del efecto ejercido por la presencia de variaciones ambientales entre tanques (“efecto tanque”) (Speare y col., 1995) con cierta influencia sobre el carácter (por ejemplo, aquellos casos donde solamente se detecta la asociación polimorfismo-peso corporal en una única categoría), o si bien, representan una asociación altamente probable entre componente genético y carácter (la asociación polimorfismo-peso corporal se detecta en todas o casi todas las categorías).

A la hora de evaluar asociaciones entre polimorfismo SNP y carácter, diversos autores (Day y col., 2001; Lei y col., 2005; He y col., 2006, 2008; Li, H. y col., 2012; Li, X. y col., 2012) proponen que las correlaciones se establezcan a partir de los haplotipos o diplotipos originados a partir de SNPs localizados en un mismo gen, pues la información derivada de su interacción ofrece una mejor aproximación que la inferida a partir de un único SNP, la cual puede llegar a ser muy limitada y conducir a deducciones erróneas. Quizá este sea el motivo por el cual el análisis individual realizado para cada uno de los 43 marcadores (SNPs e ILPs) en referencia al peso corporal de los individuos, no reveló ninguna asociación de interés, mientras que el análisis de diplotipos derivados de la combinación de diferentes marcadores localizados en un mismo gen, sí reportó asociaciones significativas entre gen y carácter.

En nuestro caso, el análisis de asociación basado en diplotipos se vió fuertemente condicionado por la elevada cantidad de haplotipos detectada (Tablas 6h y 6i), pues incluso cuando los haplotipos se establecen a partir de tan solo 3 ó 4 posiciones polimórficas (SNPs e ILPs), el número de combinaciones diplotídicas obtenido continúa siendo elevado, por lo que en múltiples ocasiones dentro de los distintos grupos descendientes (C18F, C18S, D18F y D18S) no se observan todas las combinaciones de haplotipos posibles y/o algunas de ellas se encuentran representadas por un número muy reducido de individuos el cual limita el poder de los análisis estadísticos aplicados. No obstante, los análisis de asociación realizados permitieron establecer para los genes *PRL*, *MyoD-1* y *MyoD-2*, correlaciones significativas entre determinados diplotipos y el peso corporal de los individuos.

Concretamente, en el caso del gen *PRL*, el diplotipo GCDT/GCDT originado a partir de tres SNPs no-codificantes (*SaPRL_G489C*, *SaPRL_C707T* y *SaPRL_C2958T*) y un ILP de gran tamaño (aproximadamente 100pb) (*SaPRL_ILP1524*) mostró para las cuatro categorías de

crecimiento una significativa superioridad hacia la ganancia de peso respecto a otros diplotipos (Tabla 6k), conduciendo de media a un incremento del 18% sobre el peso medio de cada categoría. Anteriormente, son pocos estudios que han logrado establecer en peces asociaciones entre marcadores genéticos de *PRL* y características de crecimiento. Sin embargo, nuestros resultados se asemejan a los obtenidos por He X. y col. (2012) en el pez barramundi (*Lates calcarifer*, perteneciente al orden de los peciformes) donde diplotipos generados a partir del polimorfismo derivado de la interacción entre cuatro SNPs no codificantes y un SNP codificante (sinónimo) del gen *PRL*, se correlacionaron con variaciones en diversas características de crecimiento, entre las cuales se incluye el peso corporal de los individuos.

Otras aproximaciones entre el polimorfismo del gen *PRL* y el crecimiento se establecen a través de los trabajos desarrollados inicialmente por Steelman y Kocher (2002) y de forma más reciente por Blel y col. (2010), donde se correlaciona la variación en la longitud de los alelos de *loci* microsatélites (localizados respectivamente en la región promotora y el primer intrón) de *PRL* con una respuesta diferencial de crecimiento, bajo ambientes con diferente salinidad, estableciéndose una conexión entre *PRL*, crecimiento y salinidad. Blel y col. (2010) postulan que la relación observada entre *PRL* y crecimiento no se debe interpretar como un marcaje directo del polimorfismo de *PRL* sobre el crecimiento, sino como un efecto secundario derivado de la actuación del polimorfismo de *PRL* como marcador genético de la capacidad de aclimatación de los individuos a ambientes con diferente salinidad. Este razonamiento se basa en el importante papel que desarrolla la hormona PRL en el proceso de osmoregulación, y en el elevado coste energético que supone para el organismo dicho proceso. Bajo este planteamiento, aquellos individuos que posean una capacidad de adaptación al medio superior, desarrollan un control osmótico más eficaz que se asocia con un gasto energético menor, permitiendo al organismo derivar una mayor cantidad de energía hacia el crecimiento; mientras que por el contrario, un menor grado de adaptación implica un gasto energético mayor comprometiendo la capacidad de crecimiento de los individuos. Por consiguiente, aquel genotipo del gen *PRL* que marque una mejor adaptación al medio, lleva implícito el marcaje de individuos con un crecimiento superior.

Sin embargo, dado que en nuestro caso los valores de salinidad empleados se basaron siempre en condiciones de cultivo estándar, la relación establecida entre *PRL* y crecimiento parece más probable que sea el reflejo de la actuación de su polimorfismo como marcador genético indirecto del efecto desencadenado por variaciones localizadas en otro gen o genes próximos, que la mera existencia de un haplotipo que represente una mejor adaptación al medio de cultivo, o una consecuencia de un posible efecto biológico generado por el propio polimorfismo de *PRL* sobre el crecimiento, pues si bien las variantes *SaPRL_G489C* y

SaPRL_C707T se encuentran localizadas dentro de la región promotora del gen *PRL*, su posición inicialmente no compromete a ninguna de las secuencias reguladoras previamente descritas por Astola y col. (2003). Aunque a partir de nuestros datos no se ha deducido ningún desequilibrio de ligamiento entre el gen *PRL* y el resto de genes, los resultados obtenidos por Sarropoulou y col. (2007) en dorada, sitúan al gen *PRL* y a otros genes candidatos para crecimiento dentro del mismo grupo de ligamiento, hecho que podría justificar la posible capacidad del polimorfismo de *PRL* para actuar como marcador indirecto de la variación genética con verdadera influencia sobre el peso corporal presente en otros genes localizados dentro de su mismo grupo de ligamiento.

Lo cierto es que, pese a que dentro del grupo de los peces se considera al gen *PRL* como un gen candidato para crecimiento (He X. y col., 2012; Sanchez-Ramos y col., 2012), existe muy poca información sobre la actuación directa de la hormona PRL sobre el eje somatotrópico, y aunque se ha demostrado una cierta capacidad de la hormona PRL para la inducción de la producción de IGF-I, todavía se desconoce cómo la hormona PRL se conecta con la cascada GH/IGF-I para activar dicha producción de IGF-I en el hígado y generar así un efecto sobre el crecimiento (Power, 2005).

De igual modo, el polimorfismo presente en los genes *MyoD-1* y *MyoD-2* marcó diferencias en peso corporal de los individuos. Concretamente, el diplotipo DGG/DGG derivado de la combinación de tres SNPs (*SaMyoD-1_D2100A*, *SaMyoD-1_A2143G* y *SaMyoD-1_A2404G*) localizados en el segundo intrón de *MyoD-1*, reflejó un efecto negativo hacia el incremento en el peso corporal, ocasionando de media una reducción del 22,5% (Tabla 6I). De forma similar, el diplotipo ACA/CCA generado a partir de la asociación de tres SNP localizados respectivamente en la región promotora (*SaMyoD-2_A785C*), primer intrón (*SaMyoD-2_C1982T*) y segundo exón (*SaMyoD-2_A2031T*) de *MyoD-2*, mostró una menor capacidad de crecimiento respecto a otros diplotipos, provocando de media una reducción del 16% (Tabla 6I). En ambas isoformas, la asociación entre diplotipo y peso corporal se determinó para las categorías C18F, C18S y D18S, mientras que en D18F, aunque se aprecia esa misma tendencia, las diferencias no resultaron significativas (Tabla 6I). No obstante, la reiteración del efecto de estos diplotipos sobre el peso corporal en las categorías de crecimiento-lento C18S y D18S, donde se registran menores diferencias de peso corporal entre los individuos y por consiguiente, resultaría más complejo establecer grupos diferenciales, supone un buen refuerzo para considerar de interés estas relaciones entre polimorfismo y carácter.

A diferencia del gen *PRL*, las isoformas *MyoD-1* y *MyoD-2* se vinculan directamente con el crecimiento, pues codifican factores miogénicos que juegan un papel clave en la diferenciación y desarrollo del tejido muscular, participando durante las fases tempranas del

desarrollo en la determinación de las líneas celulares del músculo esquelético a partir de las células del mesodermo (Biga, 2009). En la dorada, Tan y Du (2002), describen para estas isoformas un patrón de expresión diferencial en los dos tipos de tejido muscular que conforman el músculo: músculo lento o rojo (se emplea en movimientos de velocidad lenta o media), y músculo rápido o blanco (se emplea en movimientos rápidos que se mantiene un intervalo de tiempo muy corto debido al gran gasto energético que conlleva su actividad). Al comienzo del desarrollo, los dos genes *MyoD* se expresan de forma simultánea en las células precursoras de ambos tipos de tejido muscular, pero durante el desarrollo *MyoD-2* va reduciendo progresivamente su expresión en el músculo lento, y en adulto su expresión queda limitada al músculo rápido; mientras que *MyoD-1* se expresa en ambos tejidos tanto a nivel embrionario como en el adulto (Tan y Du, 2002). Dado que este patrón de expresión diferenciado presenta variaciones entre las especies, Tan y Du (2002) proponen que las diferencias de expresión entre las dos isoformas, puedan estar reflejando modelos de crecimiento diferentes que permitan a cada especie desarrollar la estructura muscular que mejor se adecue a su hábitat.

A pesar de la evidente participación de los genes *MyoD* en el control genético del desarrollo muscular, y a las correlaciones significativas detectadas entre SNPs no-codificantes y características de crecimiento en otros grupos de vertebrados como mamíferos (Verner y col., 2007; Bhuiyan y col., 2009; Chu y col., 2012) y ánades (Wu y col., 2012), los resultados obtenidos para dorada en este trabajo suponen la primera aproximación entre la posible asociación de las variantes alélicas de los genes *MyoD-1* y *MyoD-2* con el crecimiento en peces; aunque previamente algunos autores como De-Santis y col. (2007), Hemmer y col. (2011) o Yu y col. (2009), ya habrían indicado la posible potencialidad de su polimorfismo.

Pese a la fuerte vinculación de los genes *MyoD-1* y *MyoD-2* con el desarrollo muscular y por consiguiente con el crecimiento, las correlaciones deducidas entre diplotipo y peso corporal, se plantean, al igual que en el caso de *PRL*, como una asociación indirecta entre marcador y carácter, pues si bien en *MyoD-2* el SNP *SaMyoD-2_A785C* se localiza dentro de la región promotora, inicialmente no afecta a ninguna región reguladora (Tan y Du, 2002), y a pesar de que *SaMyoD-2_A2031T* supone el re-emplazamiento de aminoácidos serina (A)-cisteína (T), dicha sustitución es altamente conservativa ya que se intercambian dos aminoácidos con similares características (Tabla 6d), por lo que solamente se vería afectada la estructura primaria de la proteína, sin verse afectado el dominio HLH (Hélice-Bucle-Hélice) altamente conservado (Tan y Du, 2002). Por otro lado, los SNPs *SaMyoD-1_D2100A*, *SaMyoD-1_A2143G*, *SaMyoD-1_A2404G* y *SaMyoD-2_C1982T*, localizados todos ellos en regiones intrónicas, no afectan a ninguna región de corte y empalme exón-exón, por lo que en principio

tampoco implicarían modificaciones a nivel de regulación, pues si bien se sabe que *MyoD* se encuentra regulado por miARNs, sus dianas se localizan a nivel de la región 3'UTR (Yan B. y col., 2013), y hasta el momento no se ha descrito que sus intrones codifiquen para miARNs.

En base a los resultados derivados del presente estudio de asociación, se evidencia el poder que presenta un diseño experimental basado en la identificación de posiciones polimórficas (principalmente variantes de tipo SNP) localizadas en genes candidatos, para vincular la presencia de determinadas variantes alélicas con una mayor predisposición a la obtención del fenotipo deseado. No obstante, estos resultados se deben entender como una aproximación hacia el establecimiento de marcadores genéticos asociados a características de crecimiento en dorada, sirviendo de base para el desarrollo de estudios de asociación dirigidos a corroborar la relación establecida entre los genes *PRL*, *MyoD-1*, *MyoD-2* y el crecimiento, evaluando de forma específica el impacto individual de sus diplotipos sobre las variaciones en el crecimiento en los individuos, y valorando el efecto aditivo causado por las interacciones entre los diplotipos de dichos genes. Aunque la aceptación final de estos polimorfismos como marcador de crecimiento requiere de nuevos estudios que corroboren los resultados obtenidos, las asociaciones entre el peso corporal y el polimorfismo de los genes *PRL*, *MyoD-1* y *MyoD-2* deducidas en este estudio, ofrecen expectativas prometedoras para el uso de estos marcadores como herramientas genéticas a emplear en futuros programas de selección destinados a la mejora de las características de crecimiento en dorada.

6.5. Tablas

Tabla 6a. Información de las secuencia nucleotídica de los 11 genes incluidos en el estudio.

Gen	Nº acceso GenBank	Bibliografía	Longitud Secuencia Disponible						Longitud Secuencia Analizada						Cob.
			P + 5'	Ex	In	3'	T	P + 5'	Ex	In	3'	T			
<i>GHR_I</i>	AH014067.3	Saera-Vila y col. (2005; 2007)	1589	2326	7755	-	11670	1512	115	-	-	1627	13,9		
<i>GHR_II</i>	AH014068.3		1262	1752	6535	86	9635	1112	-	-	-	1112	11,5		
<i>IGF-I</i>	FJ827498.1 FJ827499	Sánchez-Ramos y col. (2012)	922	192	1324	-	2438	878	48	677	-	1603	65,7		
<i>Myog</i>	EF462192.1	Codina y col. (2008)	3120	753	761	1007	5641	2751	753	761	208	4473	79,2		
<i>MyoD-1</i>	AF478568.1	Tan y Du (2002)	573	894	1339	348	3154	512	699	757	0	1968	62,4		
<i>MyoD2</i>	AF478569.1		873	828	902	810	3413	354	828	684	97	1963	57,5		
<i>GH</i>	AY138985 AF195646	Almuly y col. (2000; 2005)	1784	614	2834	772	6009	980	605	1094	92	2771	46,1		
<i>PRL</i>	AJ509807.1 FJ827502	Astola y col. (2003); Sánchez Ramos y col. (2012)	1021	641	2259	669	4590	978	641	2259	493	4371	95,2		
<i>SL</i>	AJ272342 AJ508042 L49205	Astola y co. (2005)	1386	666	5360	0	7412	1296	322	1145	0	2763	37,2		
<i>MSTN-1</i>	EU881511 AF258447	Maccatrozzo y col. (2001); Funkenstein y col. (2009)	1369	1158	1070	918	4515	1323	1158	733	831	4045	89,5		
<i>Tf</i>	JF309046	García-Fernández y col. (2011)	18	2076	5487	274	7855	18	2076	5487	274	7855	100		

La longitud de las secuencias se muestra en pb. P+5': promotor y/o región 5' UTR; Ex: exones; In: intrones; T: total; Cob: porcentaje de cobertura total de secuencia realizada.

Tabla 6b. Iniciadores diseñados en el estudio.

GEN	ID [F]	Secuencia (5' → 3')	ID [R]	Secuencia (5' → 3')	Ta (°C)
<i>MyoD-1</i>	MyoD-1-F1	CCTGATCGTTACACCTCAAAG	MyoD-1-R1	CCTTCTGCCTGGTGTGGAG	52
	MyoD-1-F1	CCTGATCGTTACACCTCAAAG	MyoD-1-R1.2	TTCAGCCCCACACACACACAGAA	60
	MyoD-1-F1.2	CTGAGTAGTGTGATACTGGTCCT	MyoD-1-R1	CCTTCTGCCTGGTGTGGAG	60
	MyoD-1-F3	AAACCACCGTCAGCAGTTTG	MyoD-1-R3	GCTCCACGATGCTGGACACACAG	60
<i>MyoD-2</i>	MyoD2-F1	ACAAAACCTTCTGCACCTGTA	MyoD2-R1	TCCTCTGGGATAGGACGCTG	64
	MyoD2-F2	GACACTTTGGGCACACTGTA	MyoD2-R2	CAGCAGTTGAA CCAATTCAGC	62
	MyoD2-F2	GACACTTTGGGCACACTGTA	MyoD2-R3	GATCCGCTCTACGATGCTGG	64
	MyoD2-F4	GCTCCTTCAGCAACCAGACAG	MyoD2-R4	CAGTGAGGGAATGTCAGCTC	62
<i>Myog</i>	Myog-F1	AGTGACTCTGACACACTGCA	Myog-R2	AGAGGCATGGCCCTTCGTT	62
	Myog-F3	ATGTCA GCCCTGTGTGAGA	Myog-R3	CTAAGTTCGGGATGCCAAG	62
	Myog-F4	GAACCTTCAGCCACAACA	Myog-R4	ATGGCACTCCGCCAAGATC	55
	Myog-F5	TCAAGGGCAGGACCGAA	Myog-R5	GTATTCCAGCACACAGACA	62
	Myog-F6	CAGCAACTTCTGTGTGCA	Myog-R6	CGGCCTCATTTGGGAACAGGA	60
	<i>MSTN-1</i>	MSTN-1-F1	GGTCCTGGTTCATATGAGC	MSTN-1-R1	CGTTACTGTTTTGAGTGGCAGGTG
MSTN-1-F2		GATGGCACCATCTTCTCAGGCT	MSTN-1-R2	AAGTCCACACTGACAGGCATCAT	65
MSTN-1-F3		CCAACTAGACAAGACCTGGCCTGTA	MSTN-1-R3	ATCGTCTCCGTGATGGCGTG	62
MSTN-1-F4		TCAGGTCCAGCACAA CCA	MSTN-1-R4	CGAAAACCTCACGATCCGCTG	60
MSTN-1-F5		AATGGCTTTTACGCAGG	MSTN-1-R5	TCTCCAGCAGTGAGACGCA	60
MSTN-1-F6		ATCACACGGCTCAGATGC	MSTN-1-R7	TCCGGTATGTTGTTGCGGT	60

Tabla 6b. Continuación

GEN	ID [F]	Secuencia (5' → 3')	ID [R]	Secuencia (5' → 3')	Ta (°C)
<i>GHR-I</i>	GHR-I-F1	TGCTAAACACAAAATGCCAATAGCAG	GHR-I-R1	ACAGGTGGGAATCTCGTCCGAC	62
	GHR-I-F2	ACCAGGACAACTGCCAAACTGCA	GHR-I-R2	GCTAGCTGCCACGGCTCTGTG	66
	GHR-I-F4	ACAAGTCGCCACAGAGACAC	GHR-I-R4	AGGTTTCGTGCCCGCTTTC	60
	GHR-I-F1	ATGCAGGCAATGCTCATTACGC	GHR-I-R3	ACACCCCTGCTTTACATGCAGCTGGG	65
<i>IGF-I</i>	IGF-I-F1	GTCATTTGCAGACCAACTAGAG	IGF-I-R2	AGGGCTAGACATCCCCAC	60
	IGF-I-F3	GCAAATGCTGCCCCAGCT	IGF-I-R3	GGGTTCACCCCTGTGCCA	60
	GH-F1	CCAGAGAAACACTACAGACACC	GH-R1	CACAGATGTCACTGATCTCAGGTCA	65
<i>GH</i>	GH-F2	ATCAGCGTGTAAACCAGTGG	GH-R2	CTCTGCATGTTGACGGGTGAACTGA	64
	GH-F4	ACCAGAACTGAACCAGACAT?	GH-R4	TGGCTGCAGTCGGACACTGT	65
	GH-F3	GGCTTAGGTGTTGAAGCTGC	GH-R3	ACATCCTCCACCCTGAAACAG	60
	GH-F7	CTGCAGAAACAGCAGCCGCC	GH-R7	CGTAGGTTGCTCTCAGCGA	60
	SL-F1	CAAGGTTTGACACACTTTCAGG	SL-R1	TTCTGCCTCTGCCACATCAA	54
	SL-F2	ATATGACCCCTGTAGCCACAA	SL-R2	CACCTTTGCCGAGGTTGCGT	55
	SL-F2	ATATGACCCCTGTAGCCACAA	SL-R3	GCATGCTGGATGACTCGGTC	55
<i>SL</i>	SL-F3	CCAACAACATTTTGGGGGT	SL-R3	GCATGCTGGATGACTCGGTC	60
	SL-F5	CGCAACCTGGCAAGGT	SL-R5	AGGCAGACTTCCACAGC	52
	SL-F8	CCAACTGCCAGGCTGAA	SL-R8	TGCCTCTATGCACACTGACACA	60
	SL-F9	CCGCTGAATGATGGGCTGT	SL-R9	GGTCTTGTGAGCAGCATGT	54
	SL-F10	AGACTTTGGGATAATGTGCCAG	SL-R9	GGTCTTGTGAGCAGCATGT	54

Tabla 6b. Continuación

GEN	ID [F]	Secuencia (5' → 3')	ID [R]	Secuencia (5' → 3')	Ta (°C)
PRL	PRL_F1	CAGTAGAGGTTGTTCCTGT	PRL_R1	CACACCAATGAGCTGAGAGA	52
	PRL_F2	GTGCCAACAACCACTTGGGA	PRL_R2	GGTTTCTCTGTGAGCCATCT	58
	PRL_F3	CAGCACACACTGACAAAAGCTGCTGCA	PRL_R3	AAGGCAACCAGCAAAGCCAAAGAGCCA	62
	PRL_F4	ATCACGTCACCCTGTAGGCTGGACGA	PRL_R5	GCCGATGTCATTTGGACCCTCTGTAGGGCAG	52
	PRL_F6	GCTCTGACATGCTGCACT	PRL_R6	AGGACATCTGCGGTGGAA	60
	PRL_F7	ACCCATCCAAAAGCAGCA	PRL_R7	TCAAAGCCCTGCAACCCAG	60
	PRL_F1.2*	AGCCATTGAGGTTTACCGTGACA			

ID: nombre del iniciador; Ta: temperatura de hibridación

Tabla 6c. Listado de los 164 SNPs que componen el panel final.

SNP	MAF	He	SNP	MAF	He
<i>SaGH_C411T</i>	T=0,262	0,387	<i>SaMyoD-1_D2100A</i>	A=0,474	0,499
<i>SaGH_G2215C</i>	C=0,111	0,198	<i>SaMyoD-1_A2138C</i>	C=0,318	0,434
<i>SaGH_T2371A</i>	A=0,417	0,486	<i>SaMyoD-1_G2143A</i>	A=0,455	0,496
<i>SaGH_C4450T</i>	T=0,140	0,241	<i>SaMyoD-1_T2356A</i>	A=0,340	0,449
<i>SaGH_C4699T</i>	T=0,192	0,311	<i>SaMyoD-1_C2385T</i>	T=0,180	0,295
<i>SaGH_A4719T</i>	T=0,115	0,204	<i>SaMyoD-1_C2387T</i>	T=0,313	0,430
<i>SaGH_C4877T</i>	T=0,333	0,444	<i>SaMyoD-1_G2404A</i>	A=0,460	0,497
<i>SaGHR-I_G936A</i>	A=0,375	0,469	<i>SaMyoD-1_G2432A</i>	A=0,440	0,493
<i>SaGHR-II_A60T</i>	T=0,286	0,408	<i>SaMyoD-1_A2450T</i>	T=0,460	0,497
<i>SaGHR-II_T102A</i>	A=0,214	0,337	<i>SaMyoD-1_G2474A</i>	A=0,440	0,493
<i>SaGHR-II_T116C</i>	C=0,262	0,387	<i>SaMyoD-1_T2524C</i>	C=0,440	0,493
<i>SaGHR-II_A216C</i>	C=0,262	0,387	<i>SaMyoD-1_A2579G</i>	G=0,460	0,497
<i>SaGHR-II_G441C</i>	C=0,214	0,337	<i>SaIGF1_T195C</i>	C=0,104	0,187
<i>SaGHR-II_C449T</i>	T=0,262	0,387	<i>SaIGF1_T201C</i>	C=0,104	0,187
<i>SaGHR-II_C621T</i>	T=0,262	0,387	<i>SaIGF1_A219T</i>	T=0,316	0,432
<i>SaGHR-II_T753C</i>	C=0,262	0,387	<i>SaIGF1_A234T</i>	T=0,132	0,229
<i>SaGHR-II_C989T</i>	T=0,333	0,444	<i>SaIGF1_G250A</i>	A=0,342	0,450
<i>SaGHR-II_AC277T</i>	T=0,262	0,387	<i>SaIGF1_G251T</i>	T=0,079	0,145
<i>SaSL_A988T</i>	T=0,472	0,498	<i>SaIGF1_T264A</i>	A=0,132	0,229
<i>SaMyoD-2_A785C</i>	C=0,429	0,490	<i>SaIGF1_G342T</i>	T=0,105	0,188
<i>SaMyoD-2_G1066A</i>	A=0,385	0,473	<i>SaIGF1_G344C</i>	C=0,417	0,486
<i>SaMyoD-2_A1692C</i>	C=0,433	0,491	<i>SaIGF1_T1276C</i>	C=0,167	0,278
<i>SaMyoD-2_C1982T</i>	T=0,452	0,495	<i>SaTf_G739A</i>	A=0,214	0,337
<i>SaMyoD-2_A2031T</i>	T=0,114	0,201	<i>SaTf_G967T</i>	T=0,375	0,469
<i>SaMyoD-1_T104A</i>	A=0,235	0,360	<i>SaTf_C992T</i>	T=0,500	0,500
<i>SaMyoD-1_T139A</i>	A=0,265	0,389	<i>SaTf_G2125A</i>	A=0,400	0,480
<i>SaMyoD-1_A153G</i>	G=0,382	0,472	<i>SaTf_A2436G</i>	G=0,395	0,478
<i>SaMyoD-1_A166T</i>	T=0,382	0,472	<i>SaTf_A3223G</i>	G=0,231	0,355
<i>SaMyoD-1_A205T</i>	T=0,265	0,389	<i>SaTf_C3321T</i>	T=0,250	0,375
<i>SaMyoD-1_G409T</i>	T=0,125	0,219	<i>SaTf_A6100C</i>	C=0,200	0,320
<i>SaMyoD-1_A427G</i>	G=0,467	0,498	<i>SaTf_A6433T</i>	T=0,233	0,358
<i>SaMyoD-1_T528C</i>	C=0,300	0,420	<i>SaTf_C6782T</i>	T=0,225	0,349
<i>SaMyoD-1_C679G</i>	G=0,148	0,252	<i>SaTf_A7431G</i>	G=0,500	0,500
<i>SaMyoD-1_C771T</i>	T=0,167	0,278	<i>SaTf_C5681T</i>	T=0,132	0,229
<i>SaMyoD-1_G873A</i>	A=0,278	0,401	<i>SaTf_C6702T</i>	T=0,200	0,320
<i>SaMyoD-1_T1873A</i>	A=0,263	0,388	<i>SaTf_T6732C</i>	C=0,200	0,320
<i>SaMyoD-1_G1928A</i>	A=0,421	0,488	<i>SaPRL_G284A</i>	A=0,250	0,375
<i>SaMyoD-1_T2061C</i>	C=0,368	0,465	<i>SaPRL_A460G</i>	G=0,500	0,500

Tabla 6c. Continuación.

SNP	MAF	He	SNP	MAF	He
<i>SaPRL_G489C</i>	C=0,056	0,105	<i>SaPRL_A4190C</i>	C=0,448	0,495
<i>SaPRL_A538T</i>	T=0,444	0,494	<i>SaMSTN-1_C90A</i>	A=0,346	0,453
<i>SaPRL_T565C</i>	C=0,444	0,494	<i>SaMSTN-1_A93T</i>	T=0,154	0,260
<i>SaPRL_A629T</i>	T=0,500	0,500	<i>SaMSTN-1_G132A</i>	A=0,385	0,473
<i>SaPRL_C707T</i>	T=0,313	0,430	<i>SaMSTN-1_T134C</i>	C=0,231	0,355
<i>SaPRL_C994G</i>	G=0,417	0,486	<i>SaMSTN-1_C142T</i>	T=0,231	0,355
<i>SaPRL_CS1002D</i>	D=0,145	0,248	<i>SaMSTN-1_C144T</i>	T=0,308	0,426
<i>SaPRL_A1168T</i>	T=0,500	0,500	<i>SaMSTN-1_C150A</i>	A=0,385	0,473
<i>SaPRL_T1191A</i>	A=0,094	0,170	<i>SaMSTN-1_G172C</i>	C=0,231	0,355
<i>SaPRL_G1196C</i>	C=0,361	0,461	<i>SaMSTN-1_A180T</i>	T=0,462	0,497
<i>SaPRL_T1199C</i>	C=0,353	0,457	<i>SaMSTN-1_A299C</i>	C=0,231	0,355
<i>SaPRL_T1209C</i>	C=0,088	0,161	<i>SaMSTN-1_A337C</i>	C=0,308	0,426
<i>SaPRL_A1212C</i>	C=0,389	0,475	<i>SaMSTN-1_T342C</i>	C=0,346	0,453
<i>SaPRL_T1341C</i>	C=0,333	0,444	<i>SaMSTN-1_G350T</i>	T=0,308	0,426
<i>SaPRL_A1370T</i>	T=0,375	0,469	<i>SaMSTN-1_G352A</i>	A=0,308	0,426
<i>SaPRL_A1408G</i>	G=0,125	0,219	<i>SaMSTN-1_A354G</i>	G=0,231	0,355
<i>SaPRL_A1430G</i>	G=0,150	0,255	<i>SaMSTN-1_A364C</i>	C=0,346	0,453
<i>SaPRL_A1624G</i>	G=0,241	0,366	<i>SaMSTN-1_T426A</i>	A=0,385	0,473
<i>SaPRL_G1919C</i>	C=0,313	0,430	<i>SaMSTN-1_C437T</i>	T=0,423	0,488
<i>SaPRL_C2050G</i>	G=0,417	0,486	<i>SaMSTN-1_C576T</i>	T=0,143	0,245
<i>SaPRL_C2204G/A*</i>	G=0,155	0,392	<i>SaMSTN-1_A580C</i>	C=0,071	0,133
	A=0,086		<i>SaMSTN-1_G657A</i>	A=0,167	0,278
<i>SaPRL_G2370A</i>	A=0,300	0,420	<i>SaMSTN-1_A659T</i>	T=0,125	0,219
<i>SaPRL_T2422C</i>	C=0,156	0,264	<i>SaMSTN-1_T687A</i>	A=0,469	0,498
<i>SaPRL_C2720G</i>	G=0,176	0,291	<i>SaMSTN-1_A688C</i>	C=0,500	0,500
<i>SaPRL_C2729T</i>	T=0,156	0,264	<i>SaMSTN-1_T694C</i>	C=0,250	0,375
<i>SaPRL_T2959G</i>	G=0,409	0,483	<i>SaMSTN-1_G695T</i>	T=0,125	0,219
<i>SaPRL_G3108T</i>	T=0,125	0,219	<i>SaMSTN-1_G714C</i>	C=0,344	0,451
<i>SaPRL_C3125A</i>	A=0,455	0,496	<i>SaMSTN-1_C736A</i>	A=0,469	0,498
<i>SaPRL_C3144T</i>	T=0,152	0,258	<i>SaMSTN-1_G760G</i>	G=0,471	0,498
<i>SaPRL_C3311T</i>	T=0,167	0,278	<i>SaMSTN-1_T772A</i>	A=0,175	0,289
<i>SaPRL_C3322T</i>	T=0,204	0,324	<i>SaMSTN-1_C778T</i>	T=0,206	0,327
<i>SaPRL_A3357C</i>	C=0,296	0,417	<i>SaMSTN-1_G783C</i>	C=0,206	0,327
<i>SaPRL_T3378G</i>	G=0,241	0,366	<i>SaMSTN-1_A804T</i>	T=0,382	0,472
<i>SaPRL_A3931T</i>	T=0,480	0,499	<i>SaMSTN-1_G827C</i>	C=0,353	0,457
<i>SaPRL_G3932T</i>	T=0,480	0,499	<i>SaMSTN-1_G830A</i>	A=0,333	0,444
<i>SaPRL_A3933G</i>	G=0,480	0,499	<i>SaMSTN-1_G884A</i>	A=0,184	0,301
<i>SaPRL_G4020C</i>	C=0,300	0,420	<i>SaMSTN-1_G886T</i>	T=0,184	0,301

Tabla 6c. Continuación.

SNP	MAF	He	SNP	MAF	He
<i>SaMSTN-1_A896G</i>	G=0,206	0,327	<i>SaMSTN-1_A3181G</i>	G=0,200	0,320
<i>SaMSTN-1_C961T</i>	T=0,184	0,301	<i>SaMSTN-1_A3891G</i>	G=0,107	0,191
<i>SaMSTN-1_A2251G</i>	G=0,350	0,455	<i>SaMSTN-1_C4191T</i>	T=0,361	0,461
<i>SaMSTN-1_T2488C</i>	C=0,381	0,472	<i>SaMSTN-1_T4318C</i>	C=0,278	0,401
<i>SaMSTN-1_C2545T</i>	T=0,400	0,480	<i>SaMSTN-1_C4337T</i>	T=0,361	0,461
<i>SaMSTN-1_T3138G</i>	G=0,475	0,499	<i>SaMSTN-1_C4372T</i>	T=0,083	0,153

SNP: nombre del SNP (*Sa*=*Sparus aurata* + nombre del gen donde se localiza el polimorfismo +alelo más frecuente+ posición que ocupa el polimorfismo en la secuencia (pb)+alelo menos frecuente). MAF: Alelo menos frecuente; He: Heterocigosidad esperada. D: deleción; I: inserción.

*SNP trialélico.

Tabla 6d. SNPs codificantes identificados.

SNP	Cambio nucleótido	Posición aa	Cambio aa	
<i>SaMyoD-1_C679G</i>	T (molde)	36	W (Triptófano)	Aromático-Apolar
	C		R (Arginina)	Básico-Apolar
	G		G (Glicina)	Alifático-Polar
<i>SaMyoD-1_C771T</i>	C	66	S (Serina)	Neutro-Polar
	T			
<i>SaMyoD-1_G873A</i>	G	100	K (Lisina)	Básico (carga positiva)
	A			
<i>SaMyoD-1_T2061C</i>	C	219	F (Fenilalanina)	Aromático-Apolar
	T			
<i>SaMyoD-2_G1066A</i>	G	65	E (Glutamato)	Ácido (carga negativa)
	A		K (Lisina)	Básico (carga positiva)
<i>SaMyoD-2_A2031T</i>	A	196	S (Serina)	Neutro-Polar
	T		C (Cisteína)	
<i>SaGH_G2215C</i>	G	10	V (Valina)	Neutro -Apolar
	C		L (Leucina)	
<i>SaPRL_G2370A</i>	G	85	L (Leucina)	Neutro -Apolar
	A			
<i>SaMSTN-1_A2251G</i>	G	182	L (Leucina)	Neutro -Apolar
	A			

En gris se remarcar los SNPs que implican un cambio de aminoácido con diferencias en las características principales de la cadena lateral.

Tabla 6e. Distribución de la variación SNP incluida en el panel final.

Total SNPs "reales"										
Gen	Nº Global	Frec. Global	Nº Pro/5'UTR	Frec. Pro/5'UTR	Nº Exones	Frec. exones	Nº Intrones	Frec. intrones	Nº 3' UTR	Frec. 3'UTR
<i>GHR_I</i>	1	1SNP/1627pb	1	1SNP/1512pb	-	-	-	-	-	-
<i>GHR_II</i>	10	1SNP/111pb	10	1SNP/111pb	-	-	-	-	-	-
<i>IGF_I</i>	10	1SNP/160pb	9	1SNP/97pb	-	-	1	1SNP/677pb	-	-
<i>MyoD-1</i>	26	1SNP/394pb	8	1SNP/44pb	4	1SNP/207pb	14	1SNP/49pb	-	-
<i>MyoD-2</i>	5	1SNP/75pb	1	1SNP/512pb	2	1SNP/349pb	2	1SNP/378pb	-	-
<i>GH</i>	7	1SNP/396pb	1	1SNP/980pb	1	1SNP/605pb	5	1SNP/219pb	-	-
<i>PRL</i>	41	1SNP/106pb	9	1SNP/140pb	1	1SNP/320pb	26	1SNP/87pb	5	1SNP/99pb
<i>SL</i>	1	1SNP/2763pb	1	1SNP/1296pb	-	-	-	-	-	-
<i>MSTN-1</i>	49	1SNP/82pb	39	1SNP/34pb	1	1SNP/1158pb	4	1SNP/183pb	5	1SNP/166pb
<i>Tf</i>	14	1SNP/561pb	-	-	2	1SNP/1038pb	12	1SNP/203pb	-	-
Total	164	1SNP/185pb	79	1SNP/131pb	11	1SNP/553pb	64	1SNP/201pb	10	1SNP/114pb

Nº: número de SNPs; Frec.: frecuencia; Pro/5'UTR: SNPs localizados en la región promotora y/o 5' UTR.

Tabla 6f. Ensayos de genotipado múltiple diseñados utilizando el programa MassArray designer v.4.0.

<i>SaOVI_1</i>	<i>SaOVI_2</i>	<i>SaOVI_3</i>
<i>SaMSTN-1_A132G</i> ●	<i>SaMSTN-1_A180T</i> ●	<i>SaGH_A4696T</i>
<i>SaMSTN-1_C364A</i> ●	<i>SaMSTN-1_A299C</i> ●	<i>SaGH_C4883T</i>
<i>SaMSTN-1_C576T</i> ●	<i>SaMSTN-1_C437T</i> ●	<i>SaGHR-II_C114T</i>
<i>SaMSTN-1_A3891G</i> ●	<i>SaMSTN-1_A687T</i> ●	<i>SaIGF-I_C1276T</i>
<i>SaMSTN-1_C4191T</i> ●	<i>SaMSTN-1_G827C</i> ●	<i>SaMSTN-1_A337C</i>
<i>SaMSTN-1_T4318C</i> ●	<i>SaMSTN-1_C961T</i> ●	<i>SaMSTN-1_A659T</i>
<i>SaMSTN-1_A760G</i> ●	<i>SaMSTN-1_T3138G</i> ●	<i>SaMSTN-1_C2545T</i>
<i>SaMSTN-1_A93T</i> ●	<i>SaMSTN-1_C4337T</i> ●	<i>SaMSTN-1_C783G</i>
<i>SaMYOD-1_C679G</i> ●	<i>SaPRL_G460A</i> ●	<i>SaMSTN-1_G3138T</i>
<i>SaMYOD-1_D2100A</i> ●	<i>SaPRL_A4190C</i> ●	<i>SaMyoD-1_A1873T</i>
<i>SaMYOD-1_A2404G</i> ●	<i>SaPRL_INS1002C</i> ●	<i>SaMyoD-1_A2138C</i>
<i>SaMYOD-2_A785C</i> ●	<i>SaPRL_C2050G</i> ●	<i>SaMyoD-1_A2356T</i>
<i>SaMYOD-2_G1060A</i> ●	<i>SaMYOD-1_A104T</i> ●	<i>SaMyoD-1_A2579G</i>
<i>SaMYOD-2_C1692A</i> ●	<i>SaMYOD-1_C528T</i> ●	<i>SaMyoD-2_ILP1954</i>
<i>SaMYOD-2_A2031T</i> ●	<i>SaMYOD-1_A1928G</i> ●	<i>SaPRL_A3356C</i>
<i>SaPRL_G284A</i> ●	<i>SaMYOD-1_A2143G</i> ●	<i>SaPRL_A629T</i>
<i>SaPRL_G489C</i> ●	<i>SaMYOD-1_C2450T</i> ●	<i>SaPRL_ATA3930TGG</i>
<i>SaPRL_C707T</i> ●	<i>SaMYOD-1_C2524T</i> ●	<i>SaPRL_C1339T</i>
<i>SaPRL_C2958T</i> ●	<i>SaMYOD-2_C1982T</i> ●	<i>SaPRL_C1919G</i>
<i>SaPRL_G4020C</i> ●	<i>SaTf_A2125G</i> ●	<i>SaTf_C3321T</i>
<i>SaPRL_C565T</i> ●	<i>SaTf_A3223G</i> ●	<i>SaTf_C992T</i>
<i>SaPRL_C994G</i> ●	<i>SaTf_A6433T</i> ●	
<i>SaIGF-I_G342T</i> ●	<i>SaTf_A7431G</i> ●	
<i>SaGH_C411T</i> ●	<i>SaGHR-I_A936G</i> ●	
<i>SaGH_C4705T</i> ●	<i>SaGHR-II_C275T</i> ●	
<i>SaSL_A988T</i> ●	<i>SaGHR-II_C987T</i> ●	
<i>SaGHR-II_C751T</i> ●	<i>SaGH_C2210G</i> ●	

- SNPs validados
 ● SNPs con más de tres clusters de genotipos
- SNPs monomórficos en el lote C y D
 ● SNPs con PCR fallida
- } SNPs no validados

Tabla 6g. Caracterización de los marcadores analizados en los reproductores de los lotes C y D.

MARCADOR	Parentales Lote C		Parentales Lote D	
	MAF	He	MAF	He
<i>SaMSTN-1_A132G</i>	G=0,463	0,630	G=0,500	0,462
<i>SaMSTN-1_A299C</i>	C=0,296	0,593	C=0,327	0,500
<i>SaMSTN-1_C364A</i>	A=0,500	0,704	A=0,500	0,462
<i>SaMSTN-1_C437T</i>	T=0,352	0,630	T=0,365	0,500
<i>SaMSTN-1_C576T</i>	C=0,481	0,667	T=0,442	0,500
<i>SaMSTN-1_A687T</i>	T=0,241	0,407	T=0,288	0,346
<i>SaMSTN-1_G827C</i>	C=0,463	0,704	C=0,423	0,385
<i>SaMSTN-1_C961T</i>	T=0,315	0,630	T=0,231	0,462
<i>SaMSTN-1_T3138G</i>	G=0,463	0,630	T=0,481	0,577
<i>SaMSTN-1_A3891G</i>	G=0,185	0,148	G=0,250	0,423
<i>SaMSTN-1_C4191T</i>	T=0,167	0,333	T=0,192	0,308
<i>SaMSTN-1_T4318C</i>	C=0,389	0,481	T=0,462	0,538
<i>SaMSTN-1_C4337T</i>	T=0,407	0,444	T=0,192	0,385
<i>SaPRL_G284A</i>	A=0,444	0,519	A=0,346	0,308
<i>SaPRL_G460A</i>	A=0,426	0,630	A=0,462	0,538
<i>SaPRL_G489C</i>	C=0,259	0,444	C=0,212	0,192
<i>SaPRL_C707T</i>	T=0,241	0,407	T=0,192	0,231
<i>SaPRL_ILP1524-1621</i>	I=0,037	0,074	I=0,096	0,192
<i>SaPRL_C2958T</i>	C=0,222	0,222	C=0,212	0,346
<i>SaPRL_ILP2993</i>	D=0,370	0,370	D=0,404	0,500
<i>SaPRL_G4020C</i>	C=0,185	0,370	C=0,135	0,192
<i>SaPRL_A4190C</i>	C=0,407	0,667	C=0,308	0,538
<i>SaMyoD-1_A104T</i>	A=0,111	0,222	A=0,058	0,115
<i>SaMyoD-1_C528T</i>	C=0,111	0,222	C=0,192	0,385
<i>SaMyoD-1_C679G</i>	G=0,019	0,037	G=0,038	0,077
<i>SaMyoD-1_A1928G</i>	G=0,389	0,556	G=0,385	0,615
<i>SaMyoD-1_D2100A</i>	D=0,352	0,704	D=0,308	0,385
<i>SaMyoD-1_A2143G</i>	G=0,315	0,630	G=0,250	0,423
<i>SaMyoD-1_A2404G</i>	A=0,204	0,407	A=0,115	0,231

Tabla 6g. Continuación.

MARCADOR	Parentales Lote C		Parentales Lote D	
	MAF	He	MAF	He
<i>SaMyoD-2_A785C</i>	C=0,389	0,630	C=0,423	0,615
<i>SaMyoD-2_G1060A</i>	A=0,019	0,037	A=0,058	0,115
<i>SaMyoD-2_C1692A</i>	A=0,074	0,148	A=0,096	0,192
<i>SaMyoD-2_C1982T</i>	T=0,056	0,111	T=0,154	0,308
<i>SaMyoD-2_A2031T</i>	T=0,111	0,222	T=0,115	0,231
<i>SaTf_A2125G</i>	A=0,481	0,519	A=0,385	0,462
<i>SaTf_A3223G</i>	G=0,370	0,444	G=0,212	0,269
<i>SaTf_A6433T</i>	T=0,111	0,222	T=0,173	0,346
<i>SaTf_A7431G</i>	A=0,352	0,481	A=0,423	0,538
<i>SaGH_C411T</i>	T=0,220	0,440	T=0,333	0,583
<i>SaGHR-I_A936G</i>	A=0,460	0,440	G=0,480	0,640
<i>SaGHR-II_C987T</i>	T=0,220	0,280	T=0,260	0,440
<i>SaIGF-I_G342T</i>	T=0,060	0,120	T=0,040	0,080
<i>SaSL_A988T</i>	A=0,420	0,600	A=0,500	0,520

MAF: Alelo menos frecuente; He: Heterocigosidad observada; D: delección; I: inserción.

Tabla 6h. Relación de haplotipos y diplotipos identificados en los lotes C y D.

Gen	SNPs	Haplotipos Total	Lote C			Lote D		
			Haplotipos		Diplotipos	Haplotipos		Diplotipos
			Repro	Descen	Descen	Repro	Descen	Descen
<i>Tf</i>	4	5	5	5	14	5	5	12
<i>MyoD-1</i>	6	12	11	10	30	10	10	34
<i>MyoD-2</i>	5	6	6	6	12	6	6	16
<i>PRL</i>	9	35	26	19	67	23	18	62
<i>MSTN-1</i>	13	40	28	24	71	27	22	64

Repro: reproductores; Descen: descendientes.

Tabla 6i. Haplotipos identificados en los lotes C y D.

Gen	Haplotipo	Repro_C	Repro_D	Gen	Haplotipo	Repro_C	Repro_D
PRL	GAGCDTIGC	0,019	-	<i>Tf</i>	AGAG	0,315	0,192
	GAGCDCIGA	0,019	-		AGAA	0,056	0,019
	AGGCDCIGA	0,019	-		AATA	0,111	0,173
	AGCTDTIGC	0,019	-		GAAG	0,333	0,385
	AAGCDTDCA	0,019	-		GAAA	0,185	0,231
	AGGCDTIGC	0,037	-	<i>MyoD1</i>	ATCGAAA	0,037	-
	GAGCDTIGA	0,148	0,096		ATCAAAG	0,074	0,058
	GAGCDTDGA	0,148	0,135		TCCAAAG	0,019	0,096
	GAGCDCIGC	0,037	0,077		TCCADGG	0,074	0,038
	GAGCDCICA	0,019	0,038		TCGAAAA	0,019	0,038
	GGGCDTIGC	0,074	0,038		TTCAAAA	0,148	0,077
	GGGCDTDGA	0,037	0,096		TTCAAAG	0,222	0,250
	GGGCITIGC	0,019	0,058		TTCADAG	0,037	0,058
	GGCTDTIGA	0,037	0,019		TTCGAAG	0,130	0,173
	AGGCDTIGA	0,019	0,019		TTCGDGG	0,222	0,192
	AGGCDTICA	0,037	0,019	TCCGDGG	-	0,019	
	AGCTDTDCC	0,056	0,058	TTCADGG	0,019	-	
	AGCTDCIGA	0,037	0,019	<i>MyoD2</i>	AGCCA	0,611	0,577
	AGCTDCIGC	0,056	0,038		CAACA	0,019	0,058
	GGGCDTIGA	-	0,077		CGACA	0,056	0,038
	AGCCDTIGA	-	0,019		CGCCA	0,148	0,058
	AGCTDTDGA	-	0,019		CGCCT	0,111	0,115
	AGCTICIGA	-	0,019	CGCTA	0,056	0,154	
	AAGCDDAGA	-	0,058	<i>MSTN1</i>	AACTTACCGACTT	0,037	-
	AAGCDTIGC*	0,019	-		AACTCTCTGGCTC	0,019	-
	AGGCDTDCC*	0,037	-		AAACCTGTTACTT	0,019	-
	AGGCDCDGA*	0,019	-		AAATTAGCTACTT	0,019	-
	AGCCDTIGC*	0,019	-		ACACCAGTTACCC	0,019	-
	AGCTDCDGA*	0,019	-		GACCCAGTTACTT	0,019	-
	AGCTITDCC*	0,019	-		GAATTTGCTACCC	0,019	-
	AGGCDTDGA*	0,019	0,019		GCACCTGCGACCC	0,019	-
	GGGCITIGA*	-	0,019		GCATCAGTTACTT	0,019	-
	AGCTDTDGC*	-	0,019		AACCTAGCTACTT	0,148	0,058
	AAGCDTICC*	-	0,019		AACCTACCTACTT	0,056	0,038
	AAGCDCIGA*	-	0,019		AACCTACCGATTC	0,019	0,019

Tabla 6i. Continuación.

Gen	Haplotipo	Reprod_C	Reprod_D
<i>MSTN-1</i>	AACCCAGCTGCCC	0,019	0,038
	AACTTACCGACCC	0,093	0,096
	AACTTACCGGCC	0,037	0,019
	AACTCTGTTACCC	0,019	0,058
	AAACCTCTACTT	0,037	0,019
	GACTCACTTGCCC	0,019	0,019
	GAACCTCCTACCC	0,037	0,038
	GAATCAGTGACTT	0,037	0,019
	GCACTACCGATTC	0,093	0,135
	GCACCAGTTATTC	0,056	0,038
	GCACCTGCTGCCC	0,019	0,038
	GCACCTGCGGCC	0,056	0,058
	GCACCTGCTACTT	-	0,038
	AACCTAGCGACCC	-	0,038
	AACCCTGTTACCC	-	0,019
	AACTTACCGGCTC	-	0,038
	GAACCTGCTGCTC	-	0,019
	GAATCAGCTACCC	-	0,019
	GACCCAGTTACCC	-	0,019
	GACCCACCTACTT	0,019	-
	GAATCAGTGACCC	0,019	-
	GCACCAGTGGCCC	0,019	-
	GAATCAGTGACTC	0,019	0,019
	GCACCAGCTACCC	-	0,019
	AACTCAGCGACCC	-	0,038
	AAATCAGTGACTT	-	0,019
	GAATCAGCGGCC	-	0,019

Repr_C y *D*: frecuencia del haplotipo en los reproductores del lote C y D, respectivamente; *D*: delección; *I*: inserción.

Tabla 6j. Relación de los SNPs localizados en los genes *GH*, *GHR-I*, *GHR-II*, *IGF-I* y *SL* con el peso corporal en las categorías C18F, C18S, D18F y D18S.

GEN	SNPs	Diplotipo	Peso medio (g) ± Desviación estándar							
			C18F	C18S	D18F	D18S				
			Frec.	Frec.	Frec.	Frec.				
<i>GH</i>	<i>SaGH_C411T</i>	TT	0,08	259,25 ± 51,95 ^a	0,09	209,22 ± 45,11	0,15	306,31 ± 61,19	0,18	171,87 ± 22,96
		CT	0,41	279,95 ± 54,35 ^{ab}	0,46	193,39 ± 31,25	0,67	325,49 ± 53,28	0,60	185,64 ± 33,50
		CC	0,51	309,53 ± 54,65 ^b	0,45	200,97 ± 45,11	0,18	325,25 ± 39,34	0,22	189,55 ± 34,48
<i>GHR-I</i>	<i>SaGHR_LA936G</i>	AA	0,10	280,18 ± 53,76	0,07	187,57 ± 13,90	0,23	308,00 ± 54,49	0,10	197,33 ± 32,22
		AG	0,58	286,92 ± 60,38	0,47	195,83 ± 35,10	0,50	325,11 ± 51,14	0,57	185,04 ± 29,65
		GG	0,32	309,70 ± 47,43	0,46	202,47 ± 35,22	0,27	325,66 ± 58,32	0,32	177,50 ± 34,46
<i>GHR-II</i>	<i>SaGHRIL_C987T</i>	CC	0,66	283,48 ± 54,24	0,56	200,03 ± 31,30	0,65	317,02 ± 51,03	0,73	182,34 ± 31,14
		CT	0,33	316,38 ± 53,84	0,44	196,15 ± 33,23	0,33	329,2 ± 59,70	0,19	178,00 ± 30,15
<i>IGF-I</i>	<i>SaIGF-I_G342T</i>	GG	0,09	296,16 ± 54,82	0,84	200,61 ± 27,52	0,92	322,49 ± 54,48	0,96	184,19 ± 32,59
		GT	0,91	264,67 ± 69,40	0,16	185,19 ± 49,49	0,08	303,71 ± 48,79	0,04	179,33 ± 6,11
<i>SL</i>	<i>SaSL_G342T</i>	AA	0,13	262,71 ± 53,29	0,40	192,85 ± 29,75	0,12	327,63 ± 38,82	0,30	182,92 ± 28,58
		AT	0,61	297,67 ± 55,81	0,46	198,47 ± 32,12	0,53	317,96 ± 54,03	0,45	185,33 ± 33,97
		TT	0,25	299,53 ± 53,83	0,14	212,36 ± 33,58	0,35	323,29 ± 59,59	0,25	182,45 ± 28,07

Frec.: frecuencia del genotipo; C18F: categoría *crecimiento-rápido* a los 18 meses de edad en el lote C; C18D: categoría *crecimiento-lento* a los 18 meses de edad en el lote C; D18F: categoría *crecimiento-rápido* a los 18 meses de edad en el lote D; D18D: categoría *crecimiento-lento* a los 18 meses de edad en el lote D. Los pesos medios que difieren estadísticamente de forma significativa se indican con diferente superíndice.

Tabla 6k. Asociación entre haplotipos de *PRL* y el peso corporal en las categorías C18F, C18S, D18F y D18S.

GEN	SNPs	Diplotipo	Peso medio (g) ± Desviación estándar							
			Frec.	C18F	Frec.	C18S	Frec.	D18F	Frec.	D18S
PRL	<i>SaPRL_G489C</i> <i>SaPRL_C707T</i> <i>SaPRL_ILP1524</i> <i>SaPRL_C2958T</i>	GCDT/GCDT	0,45	310,14 ± 50,69 ^b	0,18	209,22 ± 25,39 ^b	0,28	328,56 ± 46,34 ^b	0,17	185,06 ± 30,58 ^b
		GCDT/CTDT	0,40	279,90 ± 54,87 ^a	0,26	195,33 ± 30,02 ^{ab}	0,13	281,77 ± 61,33 ^a	0,12	168,72 ± 33,80 ^{ab}
		GCDT/GCDC	-	-	0,14	197,85 ± 37,32 ^{ab}	0,12	336,66 ± 42,95 ^b	0,11	188,4 ± 35,36 ^b
		GCDT/CTDC	-	-	0,15	200,66 ± 28,58 ^{ab}	-	-	0,10	177,77 ± 27,79 ^{ab}
		GCDT/GCDT	-	-	-	-	-	-	0,09	203,00 ± 25,83 ^b
		GCDT/CTIC	-	-	-	-	0,14	321,23 ± 42,02 ^{ab}	0,07	173,33 ± 21,34 ^{ab}
		GCDC/CTDT	-	-	0,08	175,00 ± 19,26 ^a	-	-	-	-
		CTDT/CTDC	-	-	0,06	181,833 ± 19,26 ^a	0,14	305,53 ± 54,46 ^{ab}	0,06	190,40 ± 38,14 ^b
		CTDC/CTIC	-	-	-	-	-	-	0,08	196,28 ± 30,12 ^b
		GCIT/GCDT	-	-	-	-	-	-	0,07	175,33 ± 29,68 ^{ab}
		GCDC/GCDT	-	-	-	-	-	-	0,04	146,5 ± 25,99 ^a
		CTDT/GCDT	-	-	-	-	-	-	0,03	186,66 ± 12,22 ^b

Frec.: frecuencia del genotipo; *C18F*: categoría *crecimiento-rápido* a los 18 meses de edad en el lote C; *C18D*: categoría *crecimiento-lento* a los 18 meses de edad en el lote C; *D18F*: categoría *crecimiento-rápido* a los 18 meses de edad en el lote D; *D18D*: categoría *crecimiento-lento* a los 18 meses de edad en el lote D. Los pesos medios que difieren estadísticamente de forma significativa se indican con diferente superíndice. *D*: delección; *f*: inserción.

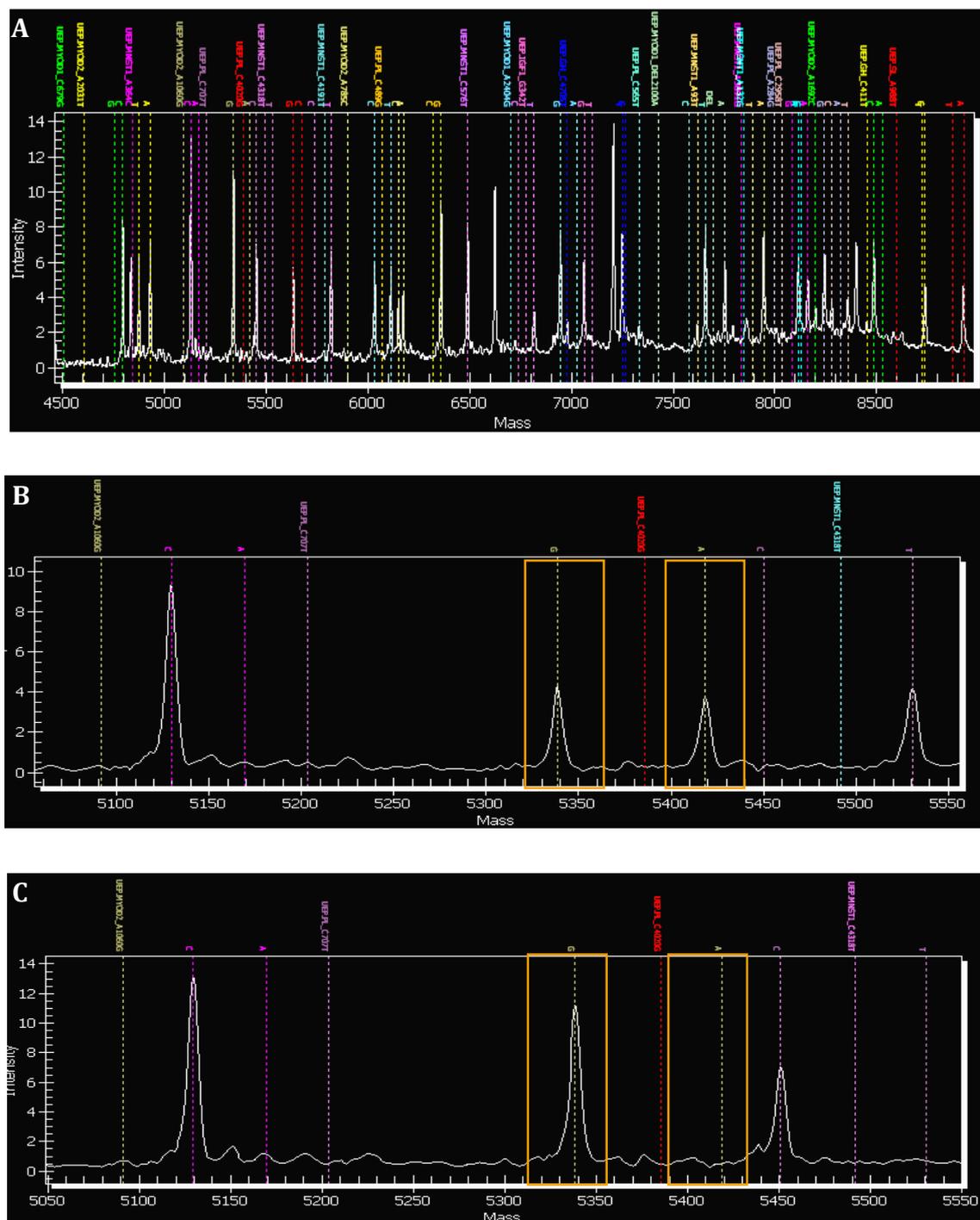
Tabla 6I. Asociación entre haplotipos de *MyoD-1* y *MyoD2* y el peso corporal en las categorías C18F, C18S, D18F y D18S.

Gen	SNPs	Diplotipo	Peso medio (g) ± Desviación estándar								
			Frec.	C18F	Frec.	C18S	Frec.	D18F	Frec.	D18S	
<i>MyoD-1</i>	SaMyoD-1_D2100A SaMyoD-1_A2143G SaMyoD-1_A2404G	AAG/AAG	0,31	310,93 ± 42,86 ^{ab}	0,27	197,82 ± 27,00 ^b	0,20	327,33 ± 37,39	0,24	193,33 ± 30,95 ^b	
		AAG/DGG	0,40	281,78 ± 53,30 ^a	0,43	197,20 ± 28,33 ^b	0,17	338,80 ± 39,40	0,13	188,90 ± 31,44 ^{ab}	
		DGG/DGG	0,12	277,00 ± 54,67 ^a	0,09	175,55 ± 39,40 ^a	0,06	302,80 ± 39,40	0,09	163,50 ± 20,41 ^a	
		AAG/AAA	0,07	344,57 ± 60,68 ^b	0,09	213,77 ± 17,24 ^b	0,17	343,20 ± 17,24	0,05	-	
		AAG/DAG	-	-	-	-	-	-	-	0,20	183,17 ± 26,20 ^{ab}
		DGG/DAG	-	-	-	-	0,08	303,42 ± 36,29	0,11	188,90 ± 36,29 ^{ab}	
<i>MyoD-2</i>	SaMyoD2_A785C SaMyoD2_C1982T SaMyoD2_A2031T	ACA/ACA	0,54	299,67 ± 43,63 ^{ab}	0,35	203,63 ± 29,57 ^b	0,24	333,81 ± 52,48	0,29	194,00 ± 34,75 ^b	
		ACA/CCA	0,14	272,4 ± 53,09 ^a	0,27	178,28 ± 22,83 ^a	0,24	322,09 ± 42,14	0,19	164,11 ± 23,00 ^a	
		ACA/CCT	0,08	301,75 ± 46,99 ^{ab}	0,10	198,80 ± 25,01 ^{ab}	0,24	326,11 ± 51,17	0,18	184,5 ± 32,77 ^{ab}	
		ACA/CTA	0,12	316,16 ± 51,58 ^b	0,11	204,09 ± 30,50 ^b	0,12	344,54 ± 46,21	0,06	186,8 ± 26,59 ^{ab}	

Frec.: frecuencia del genotipo; C18F: categoría *crecimiento-rápido* a los 18 meses de edad en el lote C; C18D: categoría *crecimiento-lento* a los 18 meses de edad en el lote C; D18F: categoría *crecimiento-rápido* a los 18 meses de edad en el lote D; D18D: categoría *crecimiento-lento* a los 18 meses de edad en el lote D. Los pesos medios que difieren estadísticamente de forma significativa se indican con diferente superíndice.
D: delección; /: inserción.

6.6. Figuras

Figura 6a. Espectro de masas obtenido para el ensayo SaOVI_1 mediante la tecnología ILPEX™ Gold MassArray® de Sequenom.



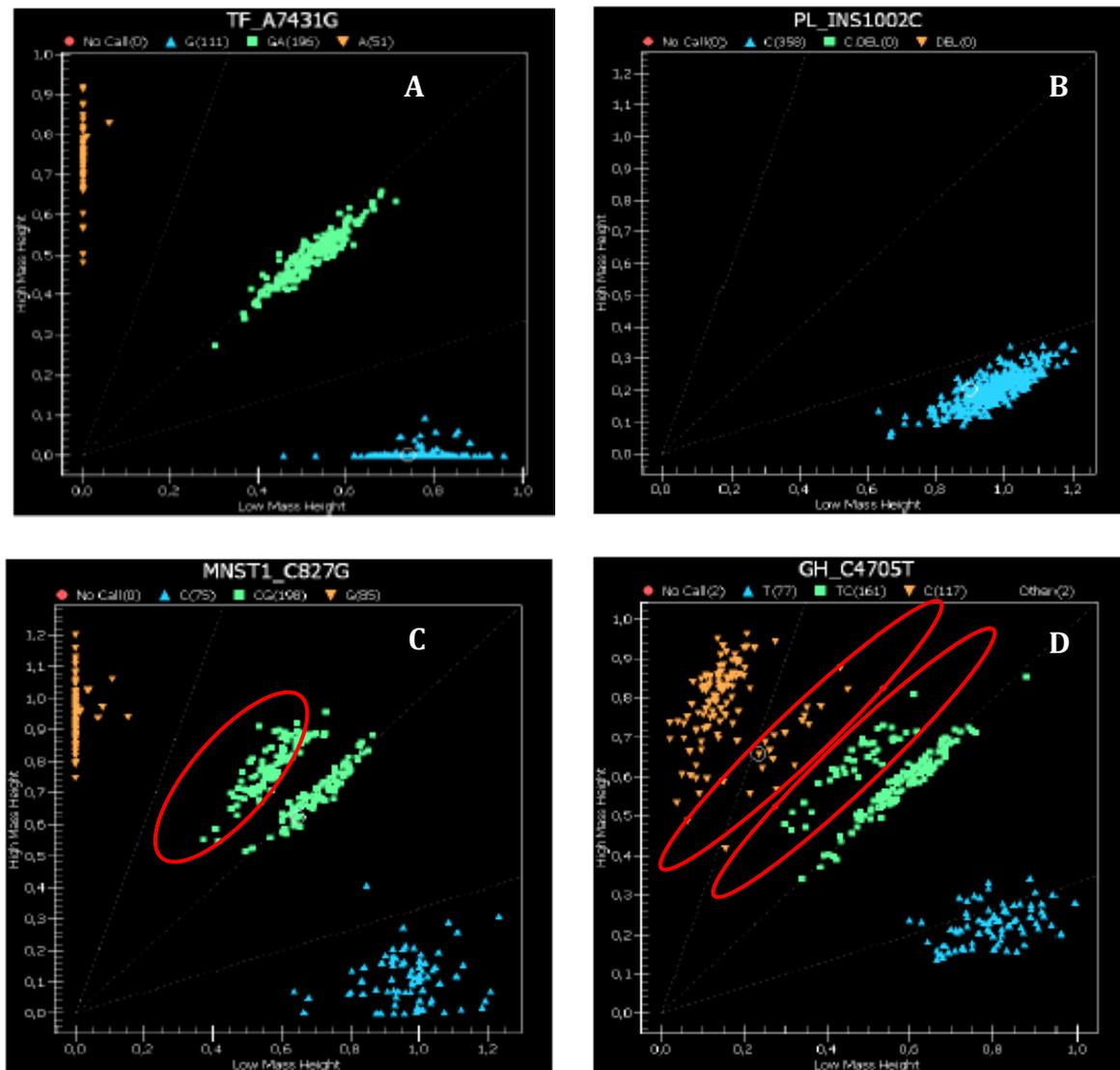
A: representación del espectro de masas obtenido para el conjunto de 24 SNPs amplificados en el ensayo 1. Cada color representa un SNP; B: individuo heterocigoto AG para el SNP *SaMyoD-2_A1060G* remarcado en naranja; C: individuo homocigoto GG para el SNP *SaMyoD-2_A1060G* remarcado en naranja.

Figura 6b. ILPs identificadas en los genes *PRL* y *MyoD-2*.



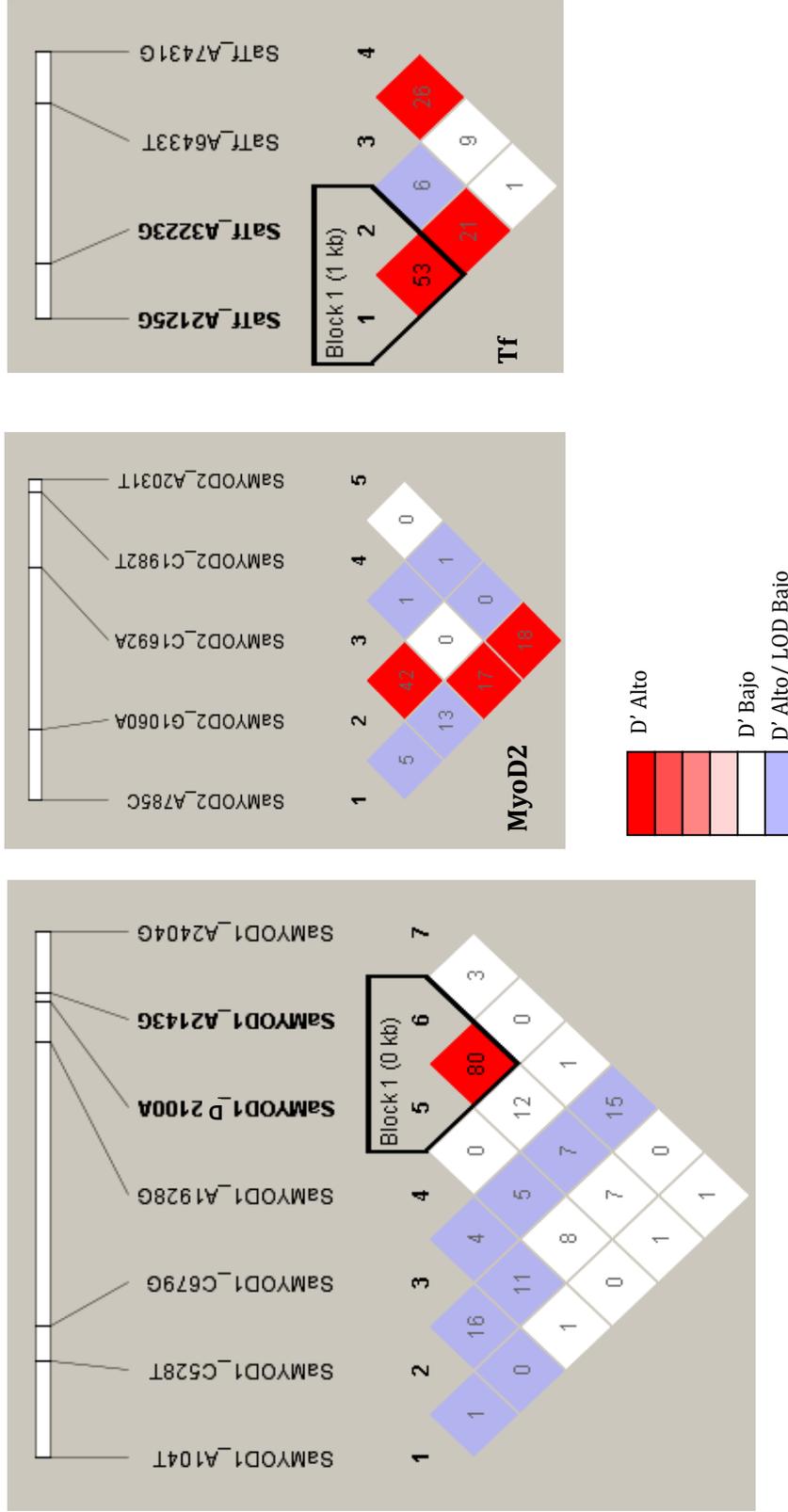
D: delección; I: inserción.

Figura 6c. Modelos de clusters obtenidos para los ensayos SaOVI-1 y SaOVI-2.



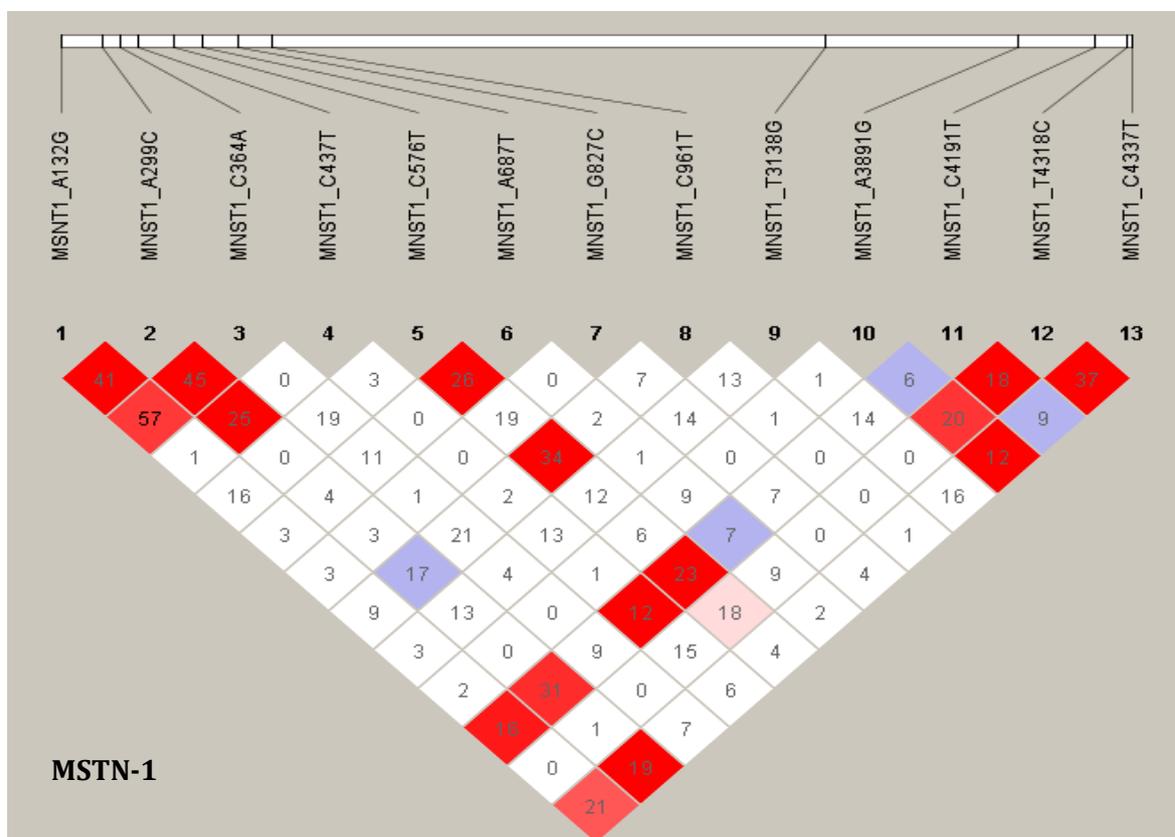
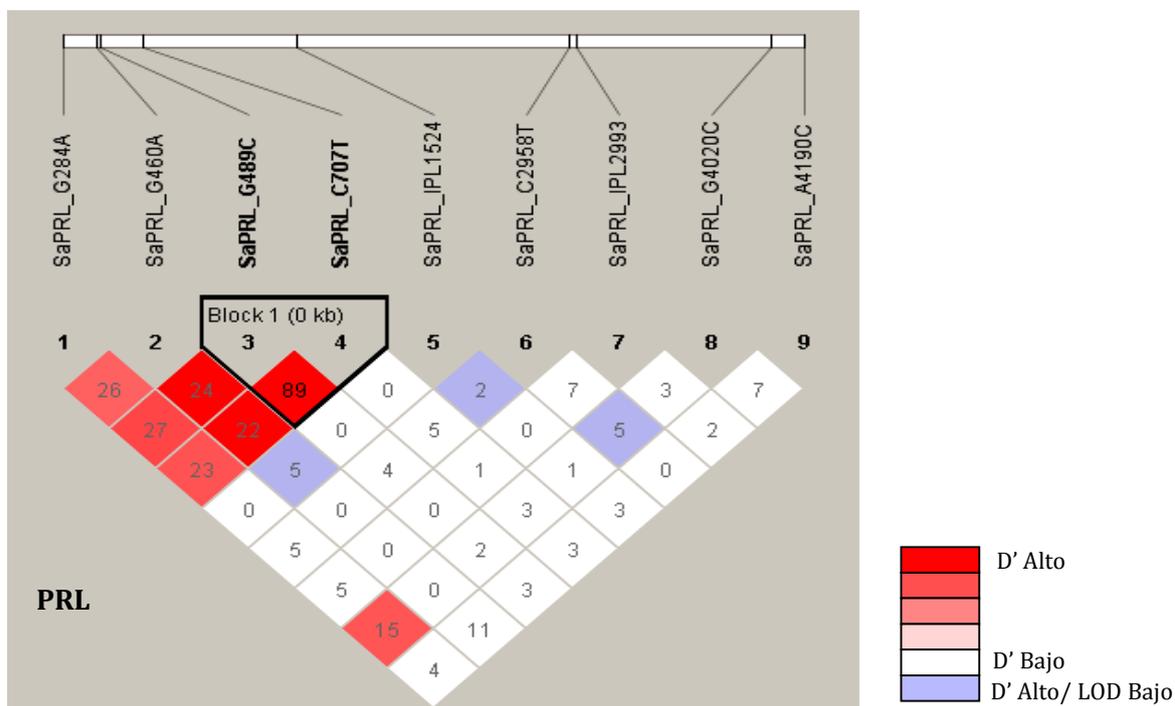
A: cluster estándar donde se observan los dos grupos de individuos homocigotos (en naranja y azul) separados por el grupo de individuos heterocigotos (en verde); B: cluster correspondiente a un SNP monomórfico; C: cluster anómalo donde se identifican dos grupos de heterocigotos (remarcados en rojo); D: cluster anómalo donde se identifican dos grupos de genotipo indefinido (remarcados en rojo) entre el grupo de homocigoto CC (naranja) y el grupo de heterocigoto CT (verde).

Figura 6d. Análisis de desequilibrio de ligamiento entre SNPs e ILPs localizados en un mismo gen.



El desequilibrio de ligamiento se muestra de acuerdo al valor del coeficiente de desequilibrio de ligamiento normalizado D' (representado en color) y al correlación R^2 (indicado en número). Los bloques haplotípicos remarcados en negro representan marcadores que se encuentran en un fuerte ligamiento.

Figura 6d. Continuación



7. Conclusiones

En base a los resultados obtenidos a partir del presente trabajo de investigación, se concluye que:

1. El protocolo de análisis desarrollado para llevar a cabo asignaciones de parentesco a partir de descendientes en estadio de desarrollo embrionario, unido a la creación de una PCR mutliplex con capacidad para combinar óptimamente un número variable de *loci* microsatélites (OVIDOPLEX-R y OVIDOPLEX), permite, tal y como demanda la empresa, conocer de forma precoz, rápida y sencilla, la dinámica reproductiva del lote de cultivo a un coste económico reducido.
2. A lo largo del periodo reproductivo, se evidencia la existencia de patrones de contribución parental muy diversos, marcados por fluctuaciones en el número de reproductores activos y variaciones en los porcentajes de contribución de los reproductores, que conducen a un incremento de endogamia superior al esperado.
3. De acuerdo a su capacidad para potenciar la participación parental y atenuar las diferencias en los porcentajes de contribución entre reproductores, se recomienda la reunificación de varias puestas independientes y el manejo reproductivo basado en *mini-lotes*, como medidas correctoras de interés para maximizar el tamaño efectivo del lote.
4. El acoplamiento de un análisis de parentesco a un sistema de selección fenotípica de descendencias, resulta de gran utilidad para identificar reproductores que contribuyen significativamente a producir descendientes con fenotipo de interés. Asumiendo que tras la existencia de una contribución parental diferencial hacia fenotipos opuestos subyacen componentes genéticos diferentes, la selección de aquellos reproductores que contribuyan en mayor medida a potenciar el fenotipo de interés, conducirá a una mejora genética del cultivo.
5. En base al patrón de expresión maternal obtenido, se propone a los genes *gapdh-1*, *catepsina-L*, *catepsina-D*, *catepsina-Z*, *ferritina-M*, *tubulina-β*, *p38-α*, *p38-δ*, *erk-1*, *jnk-1*, *ppar-β* y *ciclina-A2*, como candidatos para calidad de huevo en la especie.
6. El gen *Tf* de dorada presenta una estructura génica altamente conservada respecto a otras especies de peces, acompañada de un patrón de expresión específico de tejido y dependiente de ontogenia.
7. El análisis genómico de 11 genes candidatos para crecimiento permitió identificar un total de 398 posibles SNPs y 3 ILPs. A través de herramientas bioinformáticas, se generó un panel de 164 SNPs altamente probables, deduciéndose para la especie una densidad

genómica media de 1 SNP cada 185pb, con fuertes variaciones intra and intergénicas.

8. El análisis de la transmisión hereditaria de los dos ILPs localizados en el gen *PRL* (*SaPRL_ILP1524* y *SaILP_2993*) y 41 SNPs procedentes de los genes *GH*, *GHR-I*, *GHR-II*, *IGF-I*, *SL*, *PRL*, *MSTN*, *Tf*, *MyoD-1* y *MyoD-2*, no detectó desequilibrio de ligamiento alguno entre estos genes. No obstante, dentro de los genes *PRL*, *Tf* y *MyoD-1* se confirma la presencia de bloques haplotípicos originados a consecuencia de la segregación conjunta de SNPs próximos.
9. Los diplotipos GCDT/GCDT del gen *PRL*, DGG/DGG del gen *MyoD-1* y ACA/CCA del gen *MyoD-2*, debido a su asociación significativa con el peso corporal de los individuos, representan marcadores potenciales para crecimiento en la especie.

7 (bis). Conclusions

Based on the obtained results is concluded that:

1. The analysis protocol developed to carry out parentage assignment from embryonic development stage, together with the designed PCR multiplex assay (OVIPLEX-R and OVIPLEX) which allows to analyse simultaneously a variable number of microsatellite *loci*, provides an early, fast, easy and low cost method to assess the reproduction dynamics of the broodstock in the hatchery as demand by the industry.
2. Along reproductive period, parental contributions pattern marked by fluctuations in the number of active breeders and variable parental contributions are evidenced, giving rise to higher than expected inbreeding.
3. According to the ability to enhance the potential contribution and attenuate the different of parent contribution between breeders, the unification of offspring from independent spawning events and the *mini-broodstock* management constitute corrective measurements to maximize effective broodstock size.
4. The connection between pedigree analysis and phenotype selection process represents a useful tool for detecting breeders that preferentially contribute to offspring with a desirable phenotype. Assuming that after a preferential parental contribution to opposite phenotypes different genetic components are hidden, the selection of breeders that contribute primarily to the phenotype of interest will lead to genetic improvement for the trait.
5. According to the observed maternal expression pattern, the *gapdh-1*, *cathepsin-L*, *cathepsin-D*, *cathepsin-Z*, *ferritin-M*, *tubulin - β* , *p38- α* , *p38- δ* , *erk-1*, *jnk-1*, *ppar- β* and *cyclin-2* are proposed as candidate genes for egg quality in gilthead sea bream.
6. Gilthead sea bream *Tf* shows a highly preserved gene structure with a typical arrangement of seventeen exons/sixteen introns, which is accompanied by tissue- and stage-specific expression.
7. Genomic sequence analysis of 11 candidate growth genes detected 398 putative SNPs and 3 ILPs. Bioinformatic tools accepted 164 of them as highly probable SNPs, from which is established a mean density of 1SNP per 185bp with strong intra- and intergenic variations.

8. No linkage disequilibrium between analysed candidate growth genes is evidenced. However, inside the *PRL*, *Tf* and *MyoD-1* gene, are detected SNP haplotype blocks come from linked SNPs.
9. The GCDT/GCDT, DGG/DGG and DGG/DGG diplotype from *PRL* *MyoD-1* and *MyoD-2* gene, respectively, are significantly associated with body weight thus being potential molecular genetic markers for future breeding selection.

8. Bibliografía

- Abascal F., Zardoya R. y Posada D. 2005. ProtTest: selection of bestfit models of protein evolution. *Bioinformatics* 21 (9): 2104-2105.
- Aegerter S., Jalabert B. y Bobe J. 2005. Large scale real-time PCR analysis of mRNA abundance in rainbow trout eggs in relationship with egg quality and post-ovulatory ageing. *Molecular Reproduction and Development* 72: 377-385.
- Afonso J.M. y Roo F.J. 2007. Anomalías morfológicas en peces cultivados: heredabilidad y selección. En: *Genética y Genómica en acuicultura*. Tomo I. Págs. 215-239. (Coordinadores Figueras Huerta A. y Martínez Portela P.). Publicaciones Científicas y Tecnológicas de la Fundación Observatorio Español de Acuicultura. Mundi-Prensa. Madrid, España.
- Aisen P. 1980. Iron transport and storage protein. *Annual Review of Biochemistry* 49: 357-393.
- Almuni R., Cavari B., Ferstman H., Kolodny O. y Funkenstein B. 2000. Genomic structure and sequence of the gilthead seabream (*Sparus aurata*) growth hormone-encoding gene: Identification of minisatellite polymorphism in intron I. *Genome* 43: 836-845.
- Almuni R., Poleg-Danin Y., Gorshkov S., Gorshkova G., Rapoport B., Soller M., Kashi Y. y Funkenstein B. 2005. Characterization of 5' flanking region of the growth hormone gene of the marine teleost, gilthead sea bream (*Sparus aurata*): analysis of polymorphism microsatellite in the proximal promoter. *Fisheries Science* 71 (3): 479-490.
- Andersen O., Dehli A., Standal H., Giskegjerde T.A., Karstensen R. y Rorvik K.A. 1995. Two ferritin subunits of Atlantic salmon (*Salmo salar*): Cloning of the liver cDNAs and antibody preparation. *Molecular Marine Biology and Biotechnology* 4: 164-170.
- Andersen O., Pantopoulos K., Kao H.T., Muckenthaler M., Youson J.H. y Pieribone V. 1998. Regulation of iron metabolism in the sanguivore lamprey *Lampetra fluviatilis*—molecular cloning of two ferritin subunits and two iron-regulatory proteins (IRP) reveals evolutionary conservation of the iron-regulatory element (IRE)/ IRP regulatory system. *European Journal of Biochemistry* 254: 223-229.
- Anderson E.C. y Garza J.C. 2006. The Power of single nucleotide polymorphisms for large-scale parentage inference. *Genetics* 172 (4): 2567-2582.
- Antonello J., Massault C., Franch R., Haley C., Pellizzari C., Bovo G., Patarnello T., Koning D. y Bargelloni L. 2009. Estimates of heritability and genetic correlation for body length and resistance to fish pasteurellosis in the gilthead sea bream. *Aquaculture* 298 (1-2): 29-35.
- APROMAR 2012. La acuicultura marina en España. Disponible en www.apromar.es.
- APROMAR 2013. La acuicultura en España. Disponible en www.apromar.es y www.esacua.com.
- Argentsinger L.S. y Carter-Su C. 1996. Mechanism of signaling by growth hormone receptor. *Physiological Reviews* 76: 1089-1107.
- Astola A., Caldach-Giner J.A., Ortiz M., Pérez-Sánchez J. y Valdivia M.M. 2005. Genomic structure and functional analysis of promoter region of somatolactin gene of sea bream (*Sparus aurata*). *Marine Biotechnology* 6: 411-418.
- Astola A., Ortiz M., Caldach-Giner J.A., Pérez-Sánchez J. y Valdivia M.M. 2003. Isolation of *Sparus aurata* prolactin gene and activity of the cis-acting regulatory elements. *General and Comparative Endocrinology* 134: 57-61.
- Ballif B.A. y Blenis J. 2001. Molecular mechanisms mediating mammalian mitogen-activated protein kinase (MAPK) kinase (MEK)-MAPK cell survival signals. *Cell and Growth Differentiation* 12: 397-408.

- Bao Z. 2010. Comparison of index selection, BLUP, MAS and whole genome selection. En: Next generation sequencing and whole genome selection in aquaculture. Págs. 185-218. (Ed. Liu Z.). Wiley-Blackwell, Oxford, UK.
- Barbaro A., Francescon A., Bozzato G., Merlin A., Belvedere P. y Colombo L. 1997. Induction of spawning in gilthead seabream *Sparus aurata* L. by a long-acting GnRH agonist and its effect on egg quality and daily timing of spawning. *Aquaculture* 154: 349-359.
- Barrett J.C., Fry B., Maller J. y Daly M.J. 2005. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21 (2): 263-265.
- Bartel D.P. 2004. MicroRNAs: genomics, biogenesis, mechanism and function. *Cell* 116 (2): 281-297.
- Basurco B., Lovatelli A. y García B. 2011. Current status of Sparidae aquaculture. En: Sparidae: Biology and aquaculture of gilthead sea bream and other species. Págs. 4-14. (Eds. Pavlidis M.A. y Mylonas C.C.). Wiley-Blackwell, Oxford, UK.
- Batargias C., Dermitzakis E., Magoulas A. y Zouros E. 1999. Characterization of six polymorphic microsatellite markers in gilthead seabream *Sparus aurata*. *Molecular Ecology* 8 (5): 897-898.
- Bekkevold D., Hansen M.M. y Loeschcke V. 2002. Male reproductive competition in spawning aggregations of cod (*Gadus morhua* L.). *Molecular Ecology* 11: 91-102.
- Bentsen H.B. y Olesen I. 2002. Designing aquaculture mass selection programs to avoid high inbreeding rates. *Aquaculture* 204: 349-359.
- Bhuiyan M.S.A, Kim N.K., Cho Y.M., Yoon D., Kim K.S., Jeon J.T. y Lee J.H. 2009. Identification of SNPs in *MyoD* gene family and their associations with carcass traits in cattle. *Livestock Science* 126: 292-297.
- Biga P.R. 2009. Muscle regulation. En: Molecular research in aquaculture. Págs. 279-314. (Ed. Ken Overturf). Wiley-Blackwell, Iowa, USA.
- Blel H., Panfili J., Guinand B, Berrebi P., Said K. y Durand J. 2010. Selection footprint at the first intron of the *Prl* gene in natural populations of the flathead mullet (*Mugil cephalus*, L. 1758). *Journal of Experimental Marine Biology and Ecology* 387: 60-67.
- Blonk R.J.W., Komen J., Kamstra A., Crooijmas R.P.M.A. y van Arendonk J.A.M. 2009. Levels of inbreeding in group mating captive broodstock populations of common sole (*Solea solea*) inferred from parental relatedness and contribution. *Aquaculture* 289: 26-31.
- Bly J.E. y Clem L.W. 1992. Temperature and teleost immune functions. *Fish Shellfish and Immunology* 2: 159-171.
- Bobé J. y Labbe C. 2010. Egg and sperm quality in fish. *General and Comparative Endocrinology* 165 (3): 535-548.
- Bole-Feysot C., Goffin V., Edery M., Binart N. y Kelly P.A. 1998. Prolactin (PRL) and its receptor: actions, signal transduction pathways and phenotypes observed in PRL receptor knockout mice. *Endocrine Reviews* 19: 225-268.
- Bogyevitch M.A. y Court N.W. 2004. Counting on mitogen-activated protein kinases—ERKs 3, 4, 5, 6, 7 and 8. *Cell Signal* 16: 1345-1354.
- Borrell Y.J, Blanco G., Vázquez E., Piñera J.A., Gimenez G., Estévez A. y Sánchez J.A. 2008. Assessing the spawning season in common dentex (*Dentex dentex*) using microsatellites. *Aquaculture Research* 39: 1258-1267.

- Borrell Y.J., Carleos C.E., Asturiano J.F., Bernardo D., Vázquez E., Corral N., Sánchez J.A. y Blanco G. 2007. Use of microsatellites and a combinatorial optimization approach in the acquisition of gilthead seabream (*Sparus aurata* L.) broodstocks for hatcheries. *Aquaculture* 269 (1-4): 200-210.
- Botstein D., White R., Skolnik M. y Davis R.W. 1980. Construction of a genetic linkage map using restriction fragment length polymorphisms. *American Journal of Human Genetic* 32: 314-331.
- Boulton K., Massault C., Houston R.D., Jan de Koning D., Haley C.S., Bovenhuis H., Batargias C, Canario A.V.M., Kotoulas G. y Tsigenopoulos C.S. 2011. QTL affecting morphometric traits and stress response in the gilthead seabream (*Sparus aurata*). *Aquaculture* 319 (1-2): 58-66.
- Braissant O. y Wahli W. 1998. Differential expression of peroxisome proliferator-activated receptor- α , - β and - γ during rat embryonic development. *Endocrinology* 139: 2748-2754.
- Bromage N., Jones J., Randall C., Thrush M., Davies B., Springate J., Duston J. y Barker G. 1992. Broodstock management, fecundity, egg quality and the timing of egg production in the rainbow trout (*Oncorhynchus mykiss*). *Aquaculture* 100: 141-166.
- Broman K.W. y Speed T.P. 1999. A review of methods for identifying QTLs in experimental crosses. En: *Statistics in Molecular Biology and Genetics. Lecture Notes-Monograph Series, Volumen 33: 114-142.* (Ed. Seillier-Moiseiwitsch F.). Institute of Mathematical Statistics, Hayward, USA.
- Brooks S., Tyler C.R. y Sumpter J.P. 1997. Egg quality in fish: what makes a good egg? *Reviews in Fish Biology and Fisheries* 7: 387-416.
- Brown R. C., Tsalavouta M., Terzoglou V., Magoulas A. y McAndrew B. J. 2005a. Additional microsatellites for *Sparus aurata* and cross-species amplification within the Sparidae family. *Molecular Ecology Notes* 5: 605-607.
- Brown R.C., Woolliams J.A. y McAndrew B.J. 2005b. Factors influencing effective populations size in commercial populations of gilthead seabream *Sparus aurata*. *Aquaculture* 247: 219-225.
- Cabrera C.V., Alonso M.C. y Huikeshoven H. 1994. Regulation of scute function by extramacrochaete in vitro and in vivo. *Development* 120 (12): 3595-603.
- Calduch-Giner J. A, Bermejo-Nogales A., Benedito-Palos L., Estensoro I., Ballester-Lozano G., Sitjà-Bobadilla A. y Pérez-Sánchez J. 2013. Deep sequencing for de novo construction of a marine fish (*Sparus aurata*) transcriptome database with a large coverage of protein-coding transcripts. *BioMed Central Genomics* 14: 178.
- Cañón J. 2006. Utilización de información molecular en programas de mejoramiento animal. *Revista Corpoica-Ciencia y Tecnología Agropecuaria* 7 (1): 5-15.
- Capella-Gutiérrez S., Silla-Martínez J.M. y Gabaldón T. 2009. TrimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25 (15): 696-704.
- Carleos C., 2010. Pij-valor: cómo expresar la rareza de la frecuencia de una cierta celda en una tabla de contingencia. *Teleskopo* 2: 36-46.
- Carnevali O., Carletta R., Cambi A., Vita A. y Bromage N. 1999a. Yolk formation and degradation during oocyte maturation in sea bream *Sparus aurata*. Involvement of two lysosomal proteinases. *Biology of Reproduction* 60: 140-146.
- Carnevali O., Centonze F., Brooks S., Marota I. y Sumpter J.P. 1999b. Molecular cloning and expression of ovarian cathepsin D in sea bream *Sparus aurata*. *Biology of Reproduction* 61: 785-791.
- Carnevali O., Cionna C., Tosti L., Cerda J. y Gioacchini G. 2008. Changes in cathepsin gene expression and

- relative enzymatic activity during gilthead sea bream oogenesis. *Molecular Reproduction and Development* 75: 97–104.
- Carnevali O., Mosconi G., Cambi A., Ridolfi S., Zanuy S. y Polzonetti-Magni A. 2001. Changes of lysosomal enzyme activities in sea bass (*Dicentrarchus labrax*) eggs and developing embryos. *Aquaculture* 202: 249–256.
- Carnevali O., Polzonetti V., Cardinali M., Pugnali A., Natalini P., Zmora N., Mosconi G. y Polzonetti-Magni A.M. 2003. Apoptosis in sea bream *Sparus aurata*. *Molecular Reproduction and Development* 66: 291–296.
- Carrillo, M., Bromage N.R., Zanuy S., Serrano R. y Prat F. 1989. The effects of modifications in photoperiod on spawning time, ovarian development and egg quality in the sea bass (*Dicentrarchus labrax* L.). *Aquaculture* 81: 351–365.
- Castelló Orvay F. 1993. Cría intensiva de la dorada (*Sparus aurata*) y la lubina (*Dicentrarchus labrax*). En: *Acuicultura marina: Fundamentos biológicos y tecnológicos de la producción*. Págs. 389–402. (Ed. Universidad de Barcelona). *Ciències Experimentals I Matemàtiques* 4. Barcelona, España.
- Castro J., Pino A., Hermida M., Bouza C., Chavarrías D., Merino P., Sánchez L. y Martínez P. 2007. A microsatellite marker tool for parentage assessment in gilthead seabream (*Sparus aurata*). *Aquaculture* 272 (S1): 210–216.
- Cenadelli S., Maran V., Bongioni G., Fusetti L., Parma P. y Aleandri R. 2007. Identification of nuclear SNPs in gilthead seabream. *Journal of Fish Biology*, 70: 399–405.
- Chaudhary J. y Skinner M.K. 1999. Basic helix-loop-helix proteins can act at the E-box within the serum response element of the c-fos promoter to influence hormone-induced promoter activation in Sertoli cells. *Molecular Endocrinology* 13 (5): 774–786.
- Chavanne H., Parati K., Cambuli C., Capoferri R., Aguilera Jiménez C.A. y Galli A. 2014. Microsatellite markers to depict the reproductive and genetic patterns of farmed gilthead seabream (*Sparus aurata*): illustration by a case study on mass spawning. *Aquaculture Research* 45 (4): 577–590.
- Chen D., McMichael J.C., VanDerMeid K.R., Masi A.W., Bortell E., Caplan J.D., Chakravarti D.N. y Barniak V.L. 1999. Evaluation of a 74-kDa transferrin-binding protein from *Moraxella (Branhamella) catarrhalis* as a vaccine candidate. *Vaccine* 18: 109–118.
- Chevassus B. 1989. Constitution of aquacultural stocks: genetic aspects. *AQUACOP IFREMER actes de Colloque* 9. Págs. 569–592. Tahití, France.
- Chistiakov D.A., Hellemans B. y Volckaert F.A.M. 2006. Microsatellites and their genomic distribution, evolution, function and applications: A review with special reference to fish genetics. *Aquaculture* 255: 1–29.
- Chu M., Yang P., Liang C., Guo X., Pei J., Ding X., Bao P. y Zhu X. 2012. Association between polymorphism of *MyoD1* gene and growth traits in Yak. *China Animal Husbandry and Veterinary Medicine* 39 (2): 111–113.
- Chuang D.M., Hough C. y Senatorov V.V. 2005. Glyceraldehyde-3-phosphate dehydrogenase, apoptosis, and neurodegenerative diseases. *Annual Review Pharmacology and Toxicology* 45: 269–290.
- Clop A., Marcq F., Takeda H., Pirottin D., Tordo X., Bibé B., Bouix J., Caiment F., Elsen J.M., Eychenne F., Larzul C., Laville E., Meish F., Milenkovic D., Tobin J., Charlier C. y Georges M. 2006. A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep. *Nature Genetics* 38 (7): 813–818.

- Cnaani A., Zilberman N., Tinman S., Hulata G. y Ron M. 2004. Genome-scan analysis for quantitative trait loci in an F2 tilapia hybrid. *Molecular Genetics and Genomics* 272 (2): 162-172.
- Codina M., Bian Y.H., Gutierrez J. y Du S.J. 2008. Cloning and characterization of myogenin from seabream (*Sparus aurata*) and analysis of promoter muscle specificity. *Comparative Biochemistry and Physiology- Part D Genomics and Proteomics* 3 (1): 128-139.
- Collins F.S., Brooks L.D. y Chakravarti A. 1998. A DNA polymorphism discovery resource for research on human genetic variation. *Genome Research* 8: 1229-1231.
- Colloca F. y Cerasi S. 2005-2009. Programa de información de especies acuáticas *Sparus aurata*. Departamento de Pesca y Acuicultura de la FAO (Disponible en: http://www.fao.org/fishery/culturedspecies/Sparus_aurata/en). Roma.
- Conover C.A. 1992. Potentiation of insulin-like growth factor (IGF) action by IGF-binding protein-3: studies of underlying mechanism. *Endocrinology* 130: 3191-3199.
- Conover C.A. y Powell D.R. 1991. Insulin-like growth factor (IGF)-binding protein-3 blocks IGF-I-induced receptor down-regulation and cell desensitization in cultured bovine fibroblasts. *Endocrinology* 129: 710-716.
- Corral N., 2005. Detección de Alelos Atípicos en Estudios de Homogeneidad. X Conferencia Española de Biometría, celebrada del 25 al 27 de Mayo en Oviedo, España.
- Danzmann R.G. 1997. PROBMAX: a computer program for assigning unknown parentage in pedigree analysis from known genotypic pools of parents and progeny. *Journal of Heredity* 88 (4): 333.
- Davidson W.S., Koop B.F., Jones S.J., Iturra P., Vidal R., Maass A., Jonassen I., Lien S. y Omholt S.W. 2010. Sequencing the genome of the Atlantic salmon (*Salmo salar*). *Genome Biology* 11 (9): 403.
- Day M.J., Rioux D. y Schaff S.F. 2001. High-resolution haplotype structure in the human genome. *Nature Genetics* 29: 229-232.
- De la Herrán R., Magoulas A., Garrido-Ramos M.A., Ruiz-Rejón C., Ruiz-Rejón M. y Zouros E. 2003. Desarrollo de microsatélites en tres especies de espáridos con interés en la acuicultura. IX Congreso Nacional de Acuicultura, celebrado del 12 al 16 de Mayo. Págs 113-136. Consejería de Agricultura y Pesca. Junta de Andalucía, Sevilla, España.
- De-Santis C. y Jerry D.R. 2007. Candidate growth genes in finfish – where should we be looking? *Aquaculture* 272: 22-38.
- Devonan-Wright E.M., Ramsey N.B., McCormick C.J., Lazier C.B. y Wright J.M. 1996. Nucleotide sequence of transferrin cDNA and tissue-specific expression of the transferrin gene in Atlantic cod (*Gadus morhua*). *Comparative Biochemistry and Physiology* 113B (2): 269-273.
- Dickey L.F., Sreedharan S., Theil E.C., Didsbury J.R., Wang Y.H. y Kaufman R.E. 1987. Differences in the regulation of messenger RNA for housekeeping and specialized-cell ferritin. A comparison of three distinct ferritin complementary DNAs, the corresponding subunits, and identification of the first processed in amphibia. *The Journal of Biological Chemistry* 262: 7901-7907.
- Doyle R.W. y Herbinger C.M. 1994. The use of DNA fingerprinting for high-intensity within-family selection in fish breeding. En: Proceedings of the 5th World Congress on Genetics Applied to Livestock Production, celebrado el 7 al 12 de Agosto. Págs. 364-371. Department of Animal and Poultry Science, University of Guelph, Guelph, Canada.
- Doyle R.W. y Talbot A.J. 1986. Artificial selection on growth and correlated selection on competitive behaviour in fish. *Canadian Journal of Fisheries and Aquatic Sciences* 43: 1059-1064.

- Dunham R.A. 2004. Strategies for genetics, conservation gene banking and maintaining genetic quality. En: *Aquaculture and Fisheries Biotechnology: Genetic Approaches*. Págs. 248-253. (Ed. Dunham R.A.) CABI Publishing, Massachusetts, USA.
- Dupont-Nievet M., Vandeputte M., Vergnet A., Merdy O., Haffray P., Chavanne H, y Chatain B. 2008. Heritabilities and GxE interaction for growth in the European sea bass (*Dicentrarchus labrax*, L.) using a marker-based pedigree. *Aquaculture* 275 (1-4): 81-87.
- Edgar R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32 (5): 1792-1797.
- Ellis A.E. 2001. Innate host defense mechanisms of fish against viruses and bacteria. *Developmental and Comparative Immunology* 25: 827-839.
- Falconer D.S. y Mackay T.F.C. 2001. *Introducción a la genética cuantitativa*. Cuarta edición. Editorial Acribia, Zaragoza, España. ISBN 9788420009490.
- FAO Fisheries and Aquaculture Department. 2013. *Global Aquaculture Production Statistics for the year 2011*. Disponible en <ftp://ftp.fao.org/FI/news/GlobalAquacultureProductionStatistics2011.pdf>.
- Fernández M.E., Goszczynski D. E., Lirón J.P., Villegas-Castagnasso E.E., Carino M.H., Ripoli M.V., Rogberg-Muñoz A., Posik D.M., Peral-García P., y Giovambattista G. 2013. Comparison of the effectiveness of microsatellites and SNP panels for genetic identification, traceability and assessment of parentage in an inbred Angus herd. *Genetics and Molecular Biology*, 36 (2): 185-191.
- Fesseehaye Y., El-bialy Z., Rezk M.A., Crooijmans R., Bovenhuis H. y Komen H. 2006. Mating systems and male reproductive success in Nile tilapia (*Oreochromis niloticus*) in breeding hapas: A microsatellite analysis. *Aquaculture* 256: 148-158.
- Fjalestad K.T., Gjedrem T. y Gjerder B. 1993. Genetic improvement of disease resistance in fish: an overview. *Aquaculture* 111 (1-4): 65-74.
- Fjalestad K.T. 2005. Breeding strategies. En: *Selection and breeding programs in Aquaculture*. Págs. 145-158. (Ed. Trygve Gjedrem). Springer. Netherlands.
- Francescon A., Barbaro A., Colombo L., Bozzato G., Chiereghin S. y Belverdere S. 1994. Induction of multiple spawning in the gilthead seabream *Sparus aurata* L. by LH-RH analogue treatments their influence on egg quality. *Rivista Italiana di Acquacoltura* 29: 109-120.
- Franch R., Louro B., Tsalavouta M., Chatziplis D., Tsigenopoulos C.S., Sarropoulou E., Antonello J., Magoulas A., Mylonas C.C., Babucci M., Patarnello T., Power D.M., Kotoulas G. y Bargelloni L. 2006. A genetic linkage map of the hermaphrodite teleost fish *Sparus aurata* L. *Genetics* 174 (2): 851-861.
- Funkenstein B., Balas V., Rebhan Y. y Pliatner. 2009. Characterization and functional analysis of the 5' flanking region of *Sparus aurata* Myostatin-1 gene. *Comparative Biochemistry and Physiology- Part A: Molecular and Integrative Physiology* 153 (1): 55-62.
- Gallardo J.A. y Neira R. 2005. Environmental dependence of inbreeding depression in cultured coho salmon (*Oncorhynchus kisutch*): Aggressiveness dominance and intraspecific competition. *Heredity* 95 (6): 449-56.
- García de la Serrana D., Estévez A., Andree K. y Johnston I.A. 2012. Fast skeletal muscle transcriptome of the gilthead sea bream (*Sparus aurata*) determined by next generation sequencing. *BioMed Central Genomics* 13: 181.
- García García J., Rouco Yanez A. y García García B. 2002. Directrices generales de diseño de explotación de engorde de especies acuícolas en jaulas en mar. *Archivos de Zootecnia* 51: 469-472.

- García-Fernández C., Sánchez J.A. y Blanco G. 2011. Characterization of the gilthead seabream (*Sparus aurata* L.) transferrin gene: genomic structure, constitutive expression and SNP variation. *Fish and Shellfish Immunology* 1 (4): 548-356.
- Garnacho J., Garnacho M.C., Ortiz C. y Aldabó T. 2005. Polimorfismos genéticos en la sepsis. *Medicina Intensiva* 29 (3): 185-191.
- Gary L. Reinitz. 1977. Tests for Association of Transferrin and Lactate Dehydrogenase Phenotypes with Weight Gain in Rainbow Trout (*Salmo gairdneri*). *Journal of the Fisheries Research Board of Canada* 34 (12): 2333-2337.
- Gilbert S.F. 2000. Myogenesis: The Development of Muscle. En: *Developmental Biology*. 6th edition. (Ed. Wigg C.) Sunderland (MA): Sinauer Associates. Disponible en: <http://www.ncbi.nlm.nih.gov/books/NBK10006/>
- Giorgi A., Mignogna G., Bellapadrona G., Gattoni M., Chiaraluce R., Consalvi V., Chiancone E. y Stefanini S. 2008. The unusual co- assembly of H- and M-chains in the ferritin molecule from the Antarctic teleosts *Trematomus bernacchii* and *Trematomus newnesi*. *Archives of Biochemistry and Biophysics* 478: 69-74.
- Giraldez A.J., Mishima Y., Rihel J., Grocock R.J., Van Dongen S., Inoue K., Enright A.J. y Schier A.F. 2006. Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. *Science* 312: 75-79.
- Gjedrem T. 1983. Genetic variation in quantitative traits and selective breeding in fish and shellfish. *Aquaculture* 33 (1-4): 51-72.
- Gjedrem T. 2000. Genetic improvement of cold-water fish species. *Aquaculture Research* 31: 25-33.
- Gjedrem T. y Olesen I. 2005. Basic statistical parameters. En: *Selection and breeding programs in aquaculture*. Págs. 45-72. (Ed: Gjedrem, T.). Springer. Dordrecht. The Netherlands.
- Gjedrem T. y Thodesen J. 2005. Selection. En: *Selection and breeding programs in aquaculture*. Págs. 89-111. (Ed: Gjedrem, T.). Springer. Dordrecht. The Netherlands.
- Gjedrem T., Robinson N. y Rye M. 2012. The importance of selective breeding in aquaculture to meet future demands for animal protein: A review. *Aquaculture* 350-353: 117-129.
- Gjerde B., Gunnes K. y Gjedrem T. 1983. Effect of inbreeding on survival and growth in rainbow trout. *Aquaculture* 34: 327-332.
- Gomme P.T. y McCann K.B. 2005. Transferrin: structure, function and potential therapeutic actions. *Drug Discovery Today* 4: 267-273.
- Gorshkov S., Gordin H, Gorshkova G. y Knibb W. R. 1997. Reproductive constraints for family selection of the gilthead seabream *Sparus aurata*. *Israeli Journal of Aquaculture* 49: 57-66.
- Goudet J. 1995. FSTAT (version 1.2): A computer program to calculate F-statistics. *Journal of Heredity* 86 (6): 485-486.
- Goudet J. 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Disponible en <http://www2.unil.ch/popgen/softwares/fstat.htm>.
- Guindon S. y Gascuel O. 2003. A simple, fast and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology* 52 (5): 696-704.
- Hall T.A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41: 95-98.

- Hariharan M., Scaria V. y Brahmachari S. 2009. dbSMR: a novel resource of genome-wide SNPs affecting microRNAs mediate regulation. *BioMed Central Bioinformatics* 10: 108.
- Hauser L., Baird M., Hilborn R., Seeb L.W. y Seeb J.E. 2011. An empirical comparison of SNPs and microsatellites for parentage and kinship assignment in a wild sockeye salmon (*Oncorhynchus nerka*) population. *Molecular Ecology Resource* 11 (1): 150-161.
- Hayes B. y Goddard M. 2010. Genome-wide association and genomic selection in animal breeding. *Genome* 53: 576-883.
- Hayes B., Laerdahl J.K., Lien S., Moen T., Berg, P., Hindar K., Davidson W.S., Koop B.F., Adzhubei A. y Hoyheim B. 2007. An extensive resource of single nucleotide polymorphism markers associated with Atlantic salmon (*Salmo salar*) expressed sequences. *Aquaculture* 265: 82-90.
- Hayman J.R., Bly J.E., Levine R.P. y Lobb C.J. 1992. Complement deficiencies in channel catfish (*Ictalurus punctatus*) associated with temperature and seasonal mortality. *Fish Shellfish and Immunology* 2: 183-192.
- He F., Sun D.X., Yu Y., Wang Y.C. y Zhang Y. 2006. Association between SNPs within prolactin gene and milk performance traits in holstein dairy cattle. *Asian-Australasian Journal of Animal Sciences* 19 (10): 1384-1389.
- He F., Wen H.S., Dong S.L., Shi B., Chen C.F., Wang L.S., Yao J., Mu X.J. y Zhou Y.G. 2008. Identification of single nucleotide polymorphism cytochrome P450-c19a and its relation to reproductive traits in Japanese flounder (*Paralichthys olivaceus*). *Aquaculture* 279: 177-181.
- He X.P., Xia J.H., Wang C.M., Pang H.Y. y Yue G.H. 2012. Significant associations of polymorphisms in the prolactin gene with growth traits in Asian seabass (*Lates calcarifer*). *Animal Genetics* 43: 233-236.
- Hemmer-Hansen J., Nielsen E.E.G, Meldrup D. y Mittelholzer C. 2011. Identification of single nucleotide polymorphisms in candidate genes for growth and reproduction in a nonmodel organism, the Atlantic cod, *Gadus morhua*. *Molecular Ecology Resources* 11 (1): 71-80.
- Herlin M., Delghandi M., Wesmajervi M., Taggart J.B., McAndrew B.J. y Penman D.J. 2008. Analysis of the parental contribution to a group of fry from a single day of spawning from a commercial Atlantic cod (*Gadus morhua*) breeding tank. *Aquaculture* 274: 218-224.
- Hsing L. C. y Rudensky A. Y. 2005. The lysosomal cysteine proteases in MHC class II antigen presentation. *Immunology Reviews*. 207: 229-241.
- Holleley C.E. y Geerts P.G. 2009. Multiplex Manager 1.0: a crossplatform computer program that plans and optimizes multiplex PCR. *BioTechniques* 46 (7): 511-517.
- Honey K. y Rudensky A.Y. 2003. Lysosomal cysteine proteases regulate antigen presentation. *Nature Reviews Immunology* 3 (6):472-82.
- Howe J.A., Howell M., Hunt T. y Newport J.W. 1995. Identification of a developmental timer regulating the stability of embryonic cyclin A and a new somatic A-type cyclin at gastrulation. *Genes and Development* 9: 1164-1176.
- Hu X., Li C. y Shi L. 2013. A novel 79-pb insertion/deletion polymorphism in 3'-flanking region of *IGF-I* gene is associated with growth-related traits in common carp (*Cyprinus carpio* L.). *Aquaculture Research* 44 (10): 1632-1638.
- Huang C.W., Li Y.H., Hu S.Y., Chi J.R., Lin G.H., Lin C.C., Gong H.Y., Chen R.H., Chang S.J., Liu F.G., y Wu J.L. 2012. Differential expression patterns of growth-related microRNA in the skeletal muscle of Nile tilapia (*Oreochromis niloticus*). *Journal of Animal Science* 90 (12): 4266-4279.

- Hubert S., Higgins B., Curtis B.A., Borza T., Bowman S. 2010. Development of a SNP resource and a genetic linkage map for Atlantic cod (*Gadus morhua*). *BioMed Central Genomics* 11: 191.
- Hubert S., Tarrant B., Higgins B., Curtis B.A. y Bowman S. 2009. Development of single nucleotide polymorphism markers for Atlantic cod (*Gadus morhua*) using expressed sequences. *Aquaculture* 296 (1-2): 7-14.
- Hulata G., Wohlfart G.W. y Halevy A. 1986. Mass selection for growth rate in the Nile tilapia (*Oreochromis niloticus*). *Aquaculture* 57: 177-184.
- Jerez S., Rodríguez C., Cejas J.R., Martín V., Bolaños A. y Lorenzo A. 2012. Influence of age of female gilthead seabream (*Sparus aurata* L.) broodstock on spawning quality throughout the reproductive season. *Aquaculture* 350-353: 54-62.
- Jiang Y., Gao X., Liu S., Zhang Y., Liu H., Sun F., Bao L., Waldbieser G. y Liu Z. 2013. Whole genome comparative analysis of channel catfish (*Ictalurus punctatus*) with four model fish species. *BioMed Central Genomics* 14: 780.
- Johnson G.L., Dohlgan H.G. y Graves L.M. 2005. MAPK kinase kinases (MKKKs) as a target class for small-molecule inhibition to modulate signaling networks and gene expression. *Current Opinion in Chemical Biology* 9: 325-331.
- Johnston, I.A. y Ashton T.J. 2013. Fish selection for improved traits based on the detection of sequence variations in genes. Disponible en: <http://www.google.com/patents/WO2013054107A1?CI=en>.
- Johnston I.A., Bower N.I. y Macqueen D.J. 2011. Review: Growth and the regulation of myotomal muscle mass in teleost fish. *The Journal of Experimental Biology* 214: 1617-1628.
- Johnston I.A., Hall T.E. y Fernández A. 2003. Genes regulating the growth of myotomal muscle in teleost fish. En: *Aquatic Genomic: Steps Toward a Great Future*. Págs. 153-166. (Eds Shimizu N., Aoki T., Hirono I. y Takashima F.) Springer-Verlang. Tokio. Japón.
- Jover M. 2007. Alternativas de futuro para la producción de la dorada (*Sparus aurata*). XI Congreso Nacional de Acuicultura: "Cultivando el futuro", celebrado del 24 al 28 de Septiembre en Vigo, España.
- Jurecka P., Irnazarow I., Stafford J.L., Ruszczuk A., Taverne N., Belosevic M., Savelkoul H.F.J. y Wiegertjes G.F. 2009a. The induction of nitric oxide response of carp macrophages by transferrin is influence by the allelic diversity of the molecule. *Fish Shellfish and Immunology* 26: 632-638.
- Jurecka P., Wiegertjes G.F., Rakus K., Pilarczyk A. y Irnazarow I. 2009b. Genetic resistance of carp (*Cyprinus carpio* L.) to *Trypanoplasma borrelli*: influence of transferrin polymorphisms. *Veterinary Immunology and Immunopathology* 127 (1-2): 19-25.
- Jurecka P., Irnazarow I., Westphal A.H., Forlenza M., Arts J.A.J., Savelkoul H.F.J. y Wiegertjes G.F. 2009c. Allelic discrimination, three-dimensional analysis and gene expression of multiple transferrin alleles of common carp (*Cyprinus carpio* L.). *Fish Shellfish and Immunology* 26 (4): 573-581.
- Kalendar R., Lee D. y Schulman A.H. 2009. FastPCR software for PCR primer and probe design and repeat search. *Genes, Genomes and Genomics* 3 (1): 1-14.
- Kalinowski S.T., Taper M.L. y Marshall T.C. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology* (16): 1099-1106.
- Kamaci H.O., Saka S. y Firat K. 2005. The cleavage and embryonic phase of gilthead seabream (*Sparus aurata* Linnaeus 1758) eggs. *E.U. Journal of Fisheries and Aquatic Science* (22): 205-209.

- Kane D.A. y Kimme C.B. 1993. The zebrafish midblastula transition. *Development* 119: 447-456.
- Kincaid H.L. 1977. Rotational line crossing: an approach to the reduction of inbreeding accumulation in trout brood stocks. *Progressive Fish-Culturist* 39: 179-181.
- Kleeberger S.R. y Schwartz D.A. 2005. From quantitative trait locus to gen: a work in progress. *American Journal of Respiratory and Critical Care Medicine* 171: 804-805.
- Kmiec M. 1999a. Transferrin Polymorphism versus growth rate in lambs, polish long-wool sheep. I. Frequency of genes and genotypes of transferrin in flock of Polish Long-wool Sheep. *Archiv Tierzucht*. 42 (4): 393-402.
- Kmiec M. 1999b. Transferrin Polymorphism versus growth rate in lambs, polish long-wool sheep. II. Analysis of relation between transferrin polymorphism of lamb blood serum versus growth rate of lambs up to age of 5 months. *Archiv Tierzucht*. 42 (5): 469-479.
- Knibb W., Gorshkova G. y Gorshkov S. 1997. Growth in strains of gilthead seabream (*Sparus aurata* L.). *The Israeli Journal of Aquaculture-Bamidgeh*. 49: 43-56.
- Krens S.F., He S., Spaink H.P. y Snaar-Jagalska B.E. 2006. Characterization and expression patterns of the MAPK family in zebrafish. *Gene Expression Patterns* 6: 1019-1026.
- Kuhl H., Sarropoulou E., Tine M., Kotoulas G., Magoulas A. y Reinhardt A. 2011. A comparative BAC Map for the gilthead sea bream (*Sparus aurata* L.). *Journal of Biomedicine and Biotechnology* 2011: 329025.
- Kvingedal A.M., Rorvik K.A. y Alestrom P. 1993. Cloning and characterization of Atlantic salmon (*Salmo salar*) serum transferrin cDNA. *Molecular Marine Biology and Biotechnology* 2: 233-238.
- Laiz-Carrión R., Fuentes J., Redruello B., Guzmán J.M., Martín del Río M.P., Power D. y Mancera J.M. 2009. Expression of pituitary prolactin, growth hormone and somatolactin is modified in response to different stressors (salinity, crowding and food-deprivation) in gilthead sea bream *Sparus aurata*. *General and Comparative Endocrinology* 162 (3): 293-300.
- Lambert L.A., Perri H., Halbrooks P.J. y Mason A.B. 2005. Evolution of the transferrin family: conservation of residues associated with iron and anion binding. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* 142 (2): 129-141.
- Launey S., Krieg F., Haffray P., Bruant J.S., Vannier A. y Guyomard R. 2003. Twelve new microsatellite markers for gilted seabream (*Sparus aurata* L.): characterization polymorphism and linkage. *Molecular Ecology* 3: 457-459.
- Le Morvan C., Troutaud D. y Deschaux P. 1998. Differential effects of temperature on specific and nonspecific immune defences in fish. *The Journal of Experimental Biology* 201: 165-168.
- Leaver M.J., Boukouvala E., Antonopoulou E., Diez A., Favre-Krey L., Ezaz M.T., Bautista J.M., Tocher D.R. y Krey G. 2005. Three peroxisome proliferator-activated receptor isotypes from each of two species of marine fish. *Endocrinology* 146: 3150-3162.
- Lee C.Y., Hu S.Y., Gong H.Y., Chen M.H.C., Lu J.K. y Wu J.L. 2009. Suppression of myostatin with vector-based RNA interference causes a double-muscle effect in transgenic zebrafish. *Biochemical and Biophysical Research Communications* 387 (4): 766-771.
- Lee-Montero I., Navarro A. Borrell Y., García-Celdrán M., Martín M., Negrín-Báez D., Blanco G., Armero E., Berbel C., Zamorano M.J., Sánchez J.A., Estévez A., Ramis G., Manchado M. y Afonso J.M. 2013. Development of the first standardised panel of two new microsatellite multiplex PCRs for gilthead seabream (*Sparus aurata* L.). *Animal Genetics* 44 (5): 533-546.

- Lei M.M., Nie Q.H., Peng, Zhang D.X. y Zhang D.Q. 2005. Single nucleotide polymorphisms of the chicken insulin-like factor binding protein-2 gene associated with chicken growth and carcass traits. *Poultry Science* 84 (8): 1191-1198.
- Leinonen T., O'Hara R.B., Cano J.M. y Merila J. 2008. Comparative studies of quantitative trait and neutral marker divergence: a meta-analysis. *Journal of Evolutionary Biology* 21: 1-17.
- Lewontin R.C. y Kojima K. 1960. The evolutionary dynamics of complex polymorphisms. *Evolution* 14 (4): 458-472.
- Lewontin R. C. 1964. The interaction of selection and linkage. I. General considerations; heterotic models. *Genetics* 49 (1): 49-67.
- Li H., Fan J., Liu S., Yang Q., Mu G. y He C. 2012. Characterization of a myostatin gene (*MSTN1*) from spotted halibut (*Verasper variegatus*) and association between its promoter polymorphism and individual growth performance. *Comparative Biochemistry and Physiology, part B* 161: 315-322.
- Li X., Bai J., Hu Y., Ye X., Li S. y Yu L. 2012. Genotypes, haplotypes and diplotypes of *IGF-II* SNPs and their association with growth traits in largemouth bass (*Micropterus salmoides*). *Molecular Biology and Reproduction* 39: 4359-4365.
- Lin S.L., Miller J.D. y Ying S.Y. 2006. Intronic microRNA (miRNA). *Journal Biomedical and Biotechnology* 268 (18): 1-13.
- Liu H., Takano T., Abernathy J., Wang S., Sha Z., Jiang Y., Terhune J., Kucuktas H., Peatman E. y Liu Z. 2009. Structure and expression of transferrin gene of channel catfish, *Ictalurus punctatus*. *Fish and Shellfish Immunology* 28 (1): 1-8.
- Longalong F.M., Eknath A.E. y Bentsen H.B. 1999. Response to bidirectional selection for frequency of early maturing females for Nile tilapia (*Oreochromis niloticus*). *Aquaculture* 178: 13-25.
- López-Albors O., Gil F., Ramírez-Zarzosa G., Vázquez J.M., Latorre R., García-Alcázar A., Arencibia A. y Moreno F. 1998. Muscle development in gilthead sea bream (*Sparus aurata* L.) and sea bass (*Dicentrarchus labrax*, L.): further histochemical and ultrastructural aspects. *Anatomia, Histologia, Embryologia* 27(4): 223-229.
- López-Fanjul, C. y Toro, M. A. 1990. Fundamentos de la mejora genética en acuicultura. En: Mejora genética de peces y moluscos. Págs. 157-181. (Coordinadores Figueras Huerta A. y Martínez Portela P.). Mundi-prensa. Madrid, España.
- López-Závala R., Cano-Camacho H., Chassin-Noria O. y Závala-Páramo M.G. 2007. Selección asistida por marcadores genéticos moleculares en especies de animales de interés pecuario. *Ciencia Nicolaita* 46: 43-56.
- Loukovitis D., Batargias C., Sarropoulou E., Apostolidis A.P., Kotoulas G., Magoulas A., Tsigenopoulos C.S. y Chatziplis D. 2013. Quantitative trait loci affecting morphology traits in gilthead seabream (*Sparus aurata* L.). *Animal Genetics* 44 (4): 480-483.
- Loukovitis D., Sarropoulou E., Tsigenopoulos C.S., Batargias C., Magoulas A., Apostolidis A.P., Chatziplis D. y Kotoulas G. 2011. Quantitative trait loci involved in sex determination and body growth in the gilthead sea bream (*Sparus aurata* L.) through targeted genome scan. *PloS One* 6 (1): e16599.
- Loukovitis D., Sarropoulou E., Vogiatzi E., Tsigenopoulos C. S., Kotoulas G. Magoulas A. y Chatziplis D. 2012. Genetic variation in farmed populations of the gilthead sea bream *Sparus aurata* in Greece using microsatellite DNA markers. *Aquaculture Research* 43: 239-246.
- Lu J., Hou S., Huang W., Yu J. y Wang W. 2011. Polymorphisms in the myostatin gene and their association

- with growth and carcass traits in duck. *African Journal of Biotechnology* 10: 11309-11312.
- Macaulay A., Scantland S. y Robert C. 2011 RNA Processing During Early Embryogenesis: Managing Storage, Utilisation and Destruction. En: RNA Processing. (Ed. Grabowski P). Intech University of Pittsburgh, USA. ISBN: 978-953-307-557-0, DOI: 10.5772/20375. Disponible en: <http://www.intechopen.com/books/rna-processing/rna-processing-during-early-embryogenesis-managing-storage-utilisation-and-destruction>.
- Maccatrozzo L., Bargelloni L., Radaelli G., Mascarello F. y Patarnello T. 2001. Characterization of the myostatin gene in the gilthead seabream, *Sparus aurata*: sequence, genomic structure, and expression pattern. *Marine Biotechnology* 3: 224-230.
- Maccio I. By I. Mazzeo, Borrell Y.J., Gallego V., García-Fernández C., Sánchez J.A., Blanco G., Pérez L. y Asturiano J.F. 2013. Larval width as indicator of growth rate and effect of larval classification on final body composition and flesh quality in cultured gilthead seabream (*Sparus aurata*, L.). *Journal of applied Ichthyology* 1-7.
- Marshall C.J. 1995. Specificity of receptor tyrosine kinase signaling: transient versus sustained extracellular signal-regulated kinase activation. *Cell* 80: 179-185.
- Marshall T.C., Slate J., Kruuk L.E.B. y Pemberton J.M. 1998. Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology* 7: 639-655.
- Mathavan S., Lee S.G., Mak A., Miller L.D., Murthy K.R., Govindarajan K.R., Tong Y., Wu Y.L., Lam S.H., Yang H., Ruan Y., Korzh V., Gong Z., Liu E.T. y Lufkin T. 2005. Transcriptome analysis of zebrafish embryogenesis using microarrays. *PLoS Genetics* 1: 260-276.
- Matzinger P. 2002. The danger model: a renewed sense of self. *Science* 296: 301-305.
- Massault C., Franch R., Haley C., De Koning D.J., Bovenhuis H., Pellizzari C., Patarnello T. y Bargelloni L. 2011. Quantitative trait loci for resistance to fish pasteurellosis in gilthead sea bream (*Sparus aurata*). *Animal Genetics* 42 (2): 191-203.
- McPherron A., Lawer A.M. y Lee S.J. 1997. Regulation of skeletal muscle mass in mice by a new TGF- β superfamily member. *Nature* 387: 83-90.
- Merila J. y Crnokrak P. 2001. Comparison of genetic differentiation at marker loci and quantitative traits. *Journal of Evolutionary Biology* 14: 892-903.
- Meuwissen T.H. 2003. Genomic selection: the future of marker assisted selection and animal breeding. Session II: MAS in animals, Congreso *Marker assisted selection: a fast track to increase genetic gain in plant and animal breeding?* FAO, celebrado del 17 y 18 de Octubre, en Turin, Italia.
- Meuwissen T.H. 2009. Accuracy of breeding values of 'unrelated' individuals predicted by dense SNP genotyping. *Genetics Selection Evolution* 41 (1): 35-43.
- Meuwissen T.H., Hayes B.J. y Goddard M.E. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157 (4): 1819-1829.
- Mignogna G., Chiaraluca R., Consalvi V., Cavallo S., Stefanini S. y Chiancone E. 2002. Ferritin from the spleen of the Antarctic teleost *Trematomus bernacchii* is an M-type homopolymer. *European Journal of Biochemistry* 269: 1600-1606.
- Mitter K., Kotoulas G., Magoulas A., Mulero V., Sepulcre P., Figueras A., Novoa B. y Sarropoulou E. 2009. Evaluation of candidate reference genes for QPCR during ontogenesis and of immune-relevant tissues of European seabass (*Dicentrarchus labrax*). *Comparative Biochemistry and Physiology - Part B: Biochemistry and Molecular Biology* 153: 340-347.

- Moav R. y Wohlfarth G.W. 1976. Two way selection for growth rate in the common carp (*Cyprinus Carpio* L.). *Genetics* 82: 83-101.
- Moen T., Hayes B., Nielsen F., Delghandi M., Fjalestad K.T., Fevolden S.E., Berg P.R. y Lien S. 2008. Identification and characterisation of novel SNP markers in Atlantic cod: evidence for directional selection. *BioMed Central Genetics* 9: 18.
- Moen T., Hoyheim B., Munck H. y Gomez-Raya L. 2004. A linkage map of Atlantic salmon (*Salmo salar*) reveals an uncommonly large difference in recombination rate between the sexes. *Animal Genetics* 35 (2): 81-92.
- Mommens M., Fernandes J.M., Bizuayehu T.T., Bolla S.L., Johnston I.A. y Babiak I. 2010. Maternal gene expression in Atlantic halibut (*Hippoglossus hippoglossus* L.) and its relation to egg quality. *BioMed Central Research Notes* 3:138
- Mosher D.S., Quignon P., Bustamante C.D., Sutter N.B., Mellersh C.S., Parker H.G. y Ostrander E.A. 2007. Mutation in the myostatin gene increases muscle mass and enhances racing performance in heterozygote dogs. *PLoS Genetics* 3 (5): e79, doi:10.1371/journal.pgen.003 0079.
- Murre C., McCaw P.S., Vaessin H., Caudy M., Jan L.Y., Jan Y.N., Cabrera C.V., Buskin J.N., Hauschka S.D., Lassar A.B., Weintraub H. y Baltimore D. 1989. Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell* 58 (3): 537-544.
- Nadjar-Boger E. y Funkenstein B. 2011. Myostatin-2 gene structure and polymorphism of the promoter and first intron in the marine fish *Sparus aurata*: evidence for DNA duplications and/or translocations. *BioMed Central Genetics* 12: 22.
- Navarro A., Badilla R., Zamorano M.J., Pasamontes V., Hildebrandt S., Sánchez J.J. y Afonso J.M. 2008. Development of two new microsatellite multiplex PCRs for three sparid species: Gilthead seabream (*Sparus aurata* L.), red porgy (*Pagrus pagrus* L.) and redbanded seabream (*P. auriga*, Valenciennes, 1843) and their application to paternity studies. *Aquaculture* 285: 30-37.
- Navarro A., Manchado M., Estévez A., Ramis G., Lee I., Ponce M., Sánchez J.A., Armero E., Negrín-Báez D., Puertas M.A., Borrell Y., García M., Sánchez J.J., Crespo A.M., Blanco G., Mariadolores E., Zamorano M.J., Martín N., Aguilera C., Roo J., Toro M.A. y Afonso J.M. 2012. Estimación de parámetros genéticos en dorada (*Sparus aurata* L.) para caracteres de interés industrial (PROGENSA ®). XVI Reunión Nacional de Mejora Genética Animal., celebrado del 31 de Mayo al 2 de Junio en Menorca, España.
- Navarro A., Oliva V., Zamorano M.J., Gines R., Izquierdo M.S., Astorga N. y Afonso J.M. 2006. Evaluation of PIT system as a method to tag fingerling of gilthead seabream (*Sparus aurata* L.): effects on growth, mortality and tag loss. *Aquaculture* 257: 309-315.
- Navarro A., Zamorano M.J., Hildebrandt S., Ginés R., Aguilera C. y Afonso J.M. 2009. Estimates of heritabilities and genetic correlations for body composition traits and G×E interactions, in gilthead seabream (*Sparus aurata* L.). *Aquaculture* 295: 183-187
- Neves J.V., Wilson J.M. y Rodrigues P.N.S. 2009. Transferrin and ferritin response to bacterial infection: the role of the liver and brain in fish. *Developmental and Comparative Immunology* 33: 848-857.
- Nielsen H.M., Sonesson A.K., Yazdi H. y Meuwissen T.H. 2009. Comparison of accuracy of genome-wide and BLUP breeding value estimates in sib based aquaculture breeding schemes. *Aquaculture* 289: 259-264.
- Noga E.J. 2000. Fish Disease: Diagnosis and treatment. Pág. 367. Iowa State University, Ames, Iowa.
- Norris A.T., Bradley D. y Cunningham E. 2000. Parentage and relatedness determination in farmed Atlantic salmon (*Salmon salar*) using microsatellite markers. *Aquaculture* 182: 73-83.

- Nyström P.E., Juneja R.K., Johansson K., Andersson-Eklund L. y Andersson, K. 1997. Association of the transferrin locus on chromosome 13 with early body weights in pigs. *Journal of Animal Breeding and Genetics* 114: 363–368.
- Okamura K., Ishizuka A., Siomi H. y Siomi M.C. 2004. Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. *Genes and Development* 18: 1655–1666.
- Orrell T.M. y Carpenter K.E. 2004. A phylogeny of fishes of the family Sparidae (Perciformes: Percoidei) inferred from mitochondrial sequence data. *Molecular Phylogenetics and Evolution* 32: 425-434.
- Ozaki A., Sakamoto T. Khoo S., Nakamura K., Coimbra M.R., Akutsu T. y Okamoto N. 2001. Quantitative trait loci (QTLs) associated with resistance/susceptibility to infectious pancreatic necrosis virus (IPNV) in rainbow trout (*Oncorhynchus mykiss*). *Molecular Genetics and Genomics* 265 (1): 23-31.
- Parati K., Cenadelli S., Bongioni G., Tsigenopoulos C., Galli A. y Chavanne H. 2012. Parentage allocation by different molecular markers for genetics selection in farmed sea bream. REPROSEL. XI Natinal Congress of S.I.R.A. (Italian Society of Animal Reproduction), celebrado del 19 al 22 de Junio en Ustica, Italia.
- Patefield W.M. 1981. Algorithm AS159. An efficient method of generating $r \times c$ tables with given row and column totals. *Journal of Applied Statistics* 30: 91–97.
- Pena J. 2012. La selección genómica en el ganado Holsteing. Portal informativo veterinario PV albéitar. Disponible en <http://albeitar.portalveterinaria.com>.
- Pérez-Sánchez J. y Le Bail P.Y. 1999. Growth hormone axis as marker of nutritional status and growth performance in fish. *Aquaculture* 177: 117-128.
- Pfaffl M.W., Horgan G.W. y Dempfle L. 2002. Relative expression software tool (REST) for groupwise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Research* 30 (9): e36. doi: 10.1093/nar/30.9.e36.
- Piñera J.A. 2009. Técnicas moleculares en la mejora genética de especies de interés en la acuicultura española. Tesis doctoral. Departamento de Biología Funcional, Universidad de Oviedo, Asturias, España.
- Piñera J.A., Bernardo D., Blanco G., Vázquez E. y Sánchez J.A. 2006. Isolation and characterization of polymorphic microsatellite markers in *Pagellus bogaraveo* and cross-species amplification in *Sparus aurata* and *Dicentrarchus labrax*. *Molecular Ecology Notes* (6): 33-35.
- Ponzoni R.W., Nguyen N.H. y Khaw H.L. 2007. Investment appraisal of genetic improvement programs in Nile tilapia (*Oreochromis niloticus*). *Aquaculture* 269: 187-199.
- Porta J., Porta J. M., Béjar J. y Álvarez M.C. 2010. Development of a microsatellite multiplex genotyping tool for the fish gilthead seabream (*Sparus aurata*): applicability in population genetics and pedigree analysis. *Aquaculture Research* 41: 1514–1522.
- Power D. 2005. Developmental ontogeny of prolactin and its receptor in fish. *General and Comparative Endocrinology* 142: 25-33.
- Power D., Almeida S. y Louro B.E.P. 2003. Seabream pituitary c-DNA EST-library. Genbank references: DN048405 y DN048399.
- Pritchard J.K. y Przeworski M. 2001. Linkage disequilibrium in humans: models and data. *American Journal of Human Genetics* 69(1): 1-14.
- Queller D.C. y Goodnight K.F. 1989. Estimating relatedness using genetic markers. *Evolution* 43: 258–275.

- Quinn N.L., McGowan C.R., Cooper G.A., Koop B.F. y Davidson W.S. 2011. Identification of genes associated with heat tolerance in arctic charr exposed to acute thermal stress. *Physiological genomics* 43 (11): 685-96.
- R Development Core Team. 2009. R: A language and environment for statistical computing. ISBN 3-900051-07-0. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org>.
- Raymond M. y Rousset F. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Heredity* 86: 248-249.
- Reid D.P., Szanto A., Glebe B., Danzmann R.G. y Ferguson M.M. 2005. QTL for body weight and condition factor in Atlantic salmon (*Salmo salar*): comparative analysis with rainbow trout (*Oncorhynchus mykiss*) and arctic charr (*Salvelinus alpinus*). *Heredity* 94 (2):166-172.
- Rengmark A.H. y Lingaas F. 2007. Genomic structure of the Nile tilapia (*Oreochromis niloticus*) transferrin gene and haplotype associated with saltwater tolerance. *Aquaculture* 272:146-155.
- Rice W.R. 1989. Analysing tables of statistical tests. *Evolution* 43: 223-225.
- Roff D. 2003. Evolutionary quantitative genetics. Are we in danger of throwing out the baby with the bathwater? *Annales Zoologici Fennici* 40: 315-320.
- Rousset F. 2008. Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Molecular Ecology Resources* 8: 103-106.
- Roux P.P. y Blenis J. 2004. ERK and p38 MAPK-activated protein kinases: a family of protein kinases with diverse biological functions. *Microbiology and Molecular Biology Reviews* 68: 320-344.
- Rowlerson A., Mascarello F., Radaelli G. y Veggetti A. 1995. Differentiation and growth of muscle in the fish *Sparus aurata* (L.): II. Hyperplastic and hypertrophic growth of lateral muscle from hatching to adult. *Journal of Muscle Research and Cell Motility* 16 (3): 223-236.
- Rowlerson A., Radaelli G., Mascarello F. y Veggetti A. 1997. Re-generation of skeletal muscle in two teleost fish: *Sparus aurata* and *Brachydanio rerio*. *Cell and Tissue Research* 289: 311-322.
- Sadovy de Mitcheson Y. y Liu M. 2008. Functional hermaphroditism in teleosts. *Fish and Fisheries* 9: 1-43.
- Saera-Vila A., Calduch-Giner J.A. y Pérez-Sánchez J. 2005. Duplication of growth hormone receptor (GHR) in fish genome: gene organization and transcriptional regulation of GHR type I and II in gilthead sea bream (*Sparus aurata*). *General and Comparative Endocrinology* 142: 193-203.
- Saera-Vila A., Calduch-Giner J.A. y Pérez-Sánchez J. 2007. Co-expression of IGFs and GH receptors (GHRs) in gilthead sea bream (*Sparus aurata* L.): sequence analysis of the GHR-flanking region. *The Journal of Endocrinology* 199 (2): 361-372.
- Sahoo P.K., Mohanty B.R., Kumari J., Barat A. y Sarangi N. 2009. Cloning, nucleotide sequence and phylogenetic analyses, and tissue-specific expression of the transferrin gene in *Cirrhinus mrigala* infected with *Aeromonas hydrophila*. *Comparative Immunology, Microbiology and Infectious Diseases* 32 (6): 527-237.
- Saillant E., Dupont-Nivet M., Haffray P. y Chatain B. 2006. Estimates of heritability and genotype-environment interactions for body weight in sea bass (*Dicentrarchus labrax* L.) raised under communal rearing conditions. *Aquaculture* 254: 139-147.
- Sánchez-Molano E., Cerna A., Toro M.A., Bouza C., Hermida M., Pardo B.G., Cabaleiro S., Fernández J. y Martínez P. 2011. Detection of growth-related QTL in turbot (*Scophthalmus maximus*). *BioMed Central Genomics* 12: 473.

- Sánchez-Ramos I., Cross I., Mácha J., Martínez-Rodríguez G., Krylov V. y Rebordinos L. 2012. Assessment of tools for marker-assisted selection in a marine commercial species: significant association between *Mstn-1* gene polymorphism and growth traits. *The Scientific World Journal*, doi:10.1100/2012/369802.
- Santure A.W., Stapley J., Ball A.D., Birkhead T.R., Burke T. y Slate J. 2010. On the use of large markers panels to estimate inbreeding and relatedness: empirical and simulation studies of a pedigreed zebra finch population typed at 771 SNPs. *Molecular Ecology* 19: 1439-1451.
- Sarropoulou E., Franch R., Louro B., Power D.M., Bargelloni L., Magoulas A., Senger F., Tsalavouta M., Patarnello T., Galibert F., Kotoulas G. y Geisler R. 2007. A gene-based radiation hybrid map of the gilthead sea bream *Sparus aurata* refines and exploits conserved synteny with *Tetraodon nigroviridis*. *BioMedCentral Genomics* 8: 44.
- Sarropoulou E., Nousdili D., Kotoulas G. y Magoulas A. 2011. Functional divergences of GAPDH isoforms during early development in two perciform fish species. *Marine Biotechnology* 13: 1115-1124.
- Sauvage C., Vagner M., Derôme N., Audet C. y Bernatchez L. 2012. Coding gene SNP mapping reveals QTL linked to growth and stress response in brook charr (*Salvelinus fontinalis*). *G3 (Bethesda, Md)* 2 (6): 707-720.
- Senger F., Priat C., Hitte C., Sarropoulou E., Franch R., Geisler R., Bargelloni L., Power D. y Galibert F. 2006. The first radiation hybrid map of a perch-like fish: the gilthead seabream (*Sparus aurata* L). *Genome* 87 (6): 793-800.
- Schuelke M., Wagner K.R., Stolz L.E., Hübner C., Riebel T., Kömen W., Braun T., Tobin J.F., Lee S.J. 2004. Myostatin mutation associated with gross muscle hypertrophy in a child. *The New England Journal of Medicine* 350 (26): 2682-2688.
- Sirover M.A. 2005. New nuclear functions of the glycolytic protein, glyceraldehyde-3-phosphate dehydrogenase, in mammalian cells. *Journal of Cellular Biochemistry* 95 (1): 45-52.
- Smouse P.E. 2010. How many SNPs are enough? *Molecular Ecology* 19 (7):1265-6.
- Sola L., Moretti A., Crosetti D., Karaiskou N., Magoulas A., Rossi A.R., Rye M., Triantafyllidis A. y Tsigenopoulos C.S. 2006. Gilthead seabream - *Sparus aurata*. En: "Genetic effects of domestication, culture and breeding of fish and shellfish, and their impacts on wild populations". (Eds Crosetti D., Lapègue S., Olesen I. y Svaasand T). GENIMPACT project: Evaluation of genetic impact of aquaculture activities on native populations. A European network. WP1 workshop "Genetics of domestication, breeding and enhancement of performance of fish and shellfish", celebrado del 12 al 17 Junio en Viterbo, Italia, <http://genimpact.imr.no/>.
- Sonesson A.K. 2003. Marker-assisted selection in fish- case studies. En: *Markers Assisted Selection*. Eds. Guimarães, E., Scherf, B., Sonnino, A., Dargie, J. Current Status and Future Perspectives in Crops, Livestock, Forestry and Fish. FAO, Rome. Disponible en www.fao.org/biotech/docs/Sonesson.pdf.
- Sonesson A.K. y Meuwissen, T.H.E. 2009. Testing strategies for genomic selection in aquaculture breeding programs. *Genetics Selection and Evolution*, 41: 37.
- Sonesson A.K., Woolliams J.A., Meuwissen T.H.E. 2005. Kinship, relationship and inbreeding. En: *Selection and Breeding Programs in Aquaculture*. Págs. 73-87. Ed. Gjedrem, T. Springer, Dordrecht, The Netherlands.
- Speare D.J., MacNair N. y Hammell K.L. 1995. Demonstration of tank effect on growth indices of juvenile rainbow trout (*Oncorhynchus mykiss*) during an ad libitum feeding trial. *American Journal of Veterinary Research* 56 (10): 1372-1379.

- Stafford J.L. y Belosevic M. 2003. Transferrin and the immune response of fish: identification of a novel mechanism of macrophage activation. *Developmental and Comparative Immunology* 27: 539-554.
- Stafford J.L., Neuman N.F. y Belosevic M. 2001. Products of proteolytic cleavage of transferrin induce nitric oxide response of goldfish macrophages. *Developmental and Comparative Immunology* 25: 101-115.
- Stajkov J. y Yablanski T. 1992. Relationships between transferrin genetic system and growth of fish of the species *Cyprinus carpio* L. at various breeding environments. *Genetika i Selekcija* 25 (2): 194-199.
- Star B., Nederbragt A.J., Jentoft S., Grimholt U., Malmstrøm M., Gregers T.F., Rounge T.B., Paulsen J., Solbakken M.H., Sharma A., Wetten O.F., Lanzén A., Winer R., Knight J., Vogel J.H., Aken B., Andersen O., Lagesen K., Tooming-Klunderud A., Edvardsen R.B., Tina K.G., Espelund M., Nepal C., Previti C., Karlsen B.O., Moum T., Skage M., Berg P.R., Gjøen T., Kuhl H., Thorsen J., Malde K., Reinhardt R., Du L., Johansen S.D., Searle S., Lien S., Nilsen F., Jonassen I., Omholt S.W., Stenseth N.C. y Jakobsen K.S. 2011. The genome sequence of Atlantic cod reveals a unique immune system. *Nature* 477 (7363): 207-210.
- Streelman J. M. y Kocher T.D. 2002. Microsatellite variation associated with prolactin expression and growth of salt-challenged tilapia. *Physiological Genomics* 9: 1-4.
- Stem K. 1964. Genética de poblaciones como base de selección. En: Reunión FAO/IUFRO sobre genética forestal. *Unasylva Revista de Silvicultura y Productos Forestales* 73-74.
- Stephens M. y Donnelly P. 2003. A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *American Journal of Human Genetics* 73: 1162-1169.
- Sun Y.N., Xu T.J., Wang J.X., Cheng Y.Z. y Wang R.X. 2011. Sequence and expression analysis of cathepsin S gene in the miiuy croaker *Miichthys miiuy*. *Fish Physiology and Biochemistry* 37: 761-765.
- Suzumoto B.K., Schreck C.B. y McIntyre J.D. 1977. Relative resistances of three genotypes of Coho Salmon (*Oncorhynchus kisutch*) and their hematological responses to bacterial kidney disease. *Journal of the Fisheries Research Board of Canada* 34: 1-7.
- Tan X. y Du S. J. 2002. Differential expression of two MyoD genes in fast and slow muscles of gilthead seabream (*Sparus aurata*). *Development Genes and Evolution* 212: 207-217.
- Tang J., Leunissen J.A.M., Voorrips R.E., van der Linden C.G. y Vosman B. 2008. HaploSNPer: a web-based allele and SNP detection tool. *BioMed Central Genetics* 9: 23.
- Tao W. J. y Boulding E.G. 2003. Associations between single nucleotide polymorphisms in candidate genes and growth rate in Arctic charr (*Salvelinus alpinus* L.). *Heredity* 91: 60-69.
- Tave D. 1996. Conceptos de genética. En: Programa de cría selectiva para piscifactorias de tamaño medio. Documento Técnico de Pesca 352. FAO. Roma.
- Tave D. 1999. Inbreeding and broodstock management. Documento Técnico de Pesca 392. Pág. 122. FAO. Roma.
- Thodesen J., Grisdale-Helland B., Helland S.J. y Gjerder B. 1999. Feed intake, growth and feed utilization of offspring from wild and selected Atlantic salmon (*Salmo salar*). *Aquaculture* 180 (3-4): 237-246.
- Tiago D.M., Laizé V. y Cancela M.L. 2008. Alternatively spliced transcripts of *Sparus aurata* insulin-like growth factor 1 are differentially expressed in adults tissues and during early development. *General and Comparative Endocrinology* 157: 107-115.
- Touriya A., Rami M., Cattaneo-Berrebi G., Ibañez C., Augros S., Boissin E., Dakkak A. y Berrebi, P. 2003.

- Primers for EPIC amplification of intron sequences for fish and other vertebrate population genetic studies. *BioTechniques* 35: 676-68.
- Traverso J.M., Fostier A. y Bobe J. 2012. Egg transcriptome, the maternal legacy to the embryo. En: *Aquaculture biotechnology*. Págs. 175-192. (Eds. Fletcher G.L y Rise M.L). Wiley-Blackwell, Oxford, UK.
- Trong T.Q., van Bers N., Crooijmans R., Dibbits B y Komen. 2013. A comparison of microsatellites and SNPs in parental assignment in the GIFT strain of Nile tilapia (*Oreochromis niloticus*): The power of exclusion. *Aquaculture* 388-391: 14-23.
- Turnsek T., Kregar I. y Lebez D. 1975. Acid sulphhydryl protease from calf lymph nodes. *Biochimica et Biophysica Acta* 403 (2): 514-520.
- UICN. 2010. Guía para el Desarrollo Sostenible de la Acuicultura Mediterránea 3. Acuicultura: Prácticas Responsables y Certificación. Gland, Suiza y Málaga, España: UICN. vi+78 pág.
- Vandesompele J., De Preter K., Pattyn F., Poppe B., Van Roy N., De Paepe A. y Speleman F. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology* 3 (7): RESEARCH0034.
- VanRaden P.M., Van Tassell C.P., Wiggans G.R., Sonstegard T.S., Schnabel R.D., Taylor J.F. y Schenkel F.S. 2009. Reliability of genomic predictions for North American Holstein bulls. *Journal of Dairy Science* 92 (1): 16-24.
- Vargas-Chacoff L., Astola A., Arjona F.J., Martín del Río M.P., García-Cózar F., Mancera J.M. y Martínez-Rodríguez G. 2009. Gene and protein expression for prolactin, growth hormone and somatolactin in *Sparus aurata*: Seasonal variations. *Comparative Biochemistry and Physiology Part B* 153: 130-135.
- Vera M., Álvarez-Dios J.A., Millán A., Pardo B.G., Bouza C., Hermida M., Fernández C., de la Herrán R., Molina-Luzón M.J. y Martínez P. 2011. Validation of single nucleotide polymorphism (SNP) markers from an immune Expressed SequenceTag (EST) turbot, *Scophthalmus maximus*, database. *Aquaculture* 313: 31-41.
- Verner J., Humpolicek P. y Knoll A. 2007. Impact of MYOD family genes on pork traits in Large White and Landrace pigs. *Journal of Animal Breeding and Genetics* 124: 81-85.
- Vignal A., Milan D., SanCristobal M. y Eggen A. 2002. A review on SNP and other types of molecular markers and their use in animal genetics. *Genetics, Selection and Evolution* 34 (3): 275-305.
- Villanueva B., Fernández J., García-Cortés L.A., Varona L., Daetwyler H.D. y Toro M.A. 2010. Accuracy of genome-wide evaluation for disease resistance in aquaculture breeding programmes. Proc 9th World Congress on Genetics Applied to Livestock Production, celebrado del 6 al 2 de Agosto en Liepzig, Alemania.
- Vogiatzi E., Lagnel J., Pakaki V., Louro B., Canario A.V.M., Reinhardt R., Kotoulas G., Magoulas A. y Tsigenopoulos C.S. 2011. In silico mining and characterization of simple sequence repeats from gilthead sea bream (*Sparus aurata*) expressed sequence tags (EST-SSRs), PCR amplification polymorphism evaluation and multiplexing and cross-species assays. *Marine Genomics* 4 (2): 83-91.
- Walsh P.S., Metzger D.A. y Higuchi R. 1991. ChelexR 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *BioTechniques* 10: 506-13.
- Wang W., Zhang M. y Sun L. 2011. Ferritin M of *Cynoglossus semilaevis*: an iron-binding protein and a broad-spectrum antimicrobial that depends on the integrity of the ferroxidase center and nucleation center for biological activity. *Fish and Shellfish Immunology* 31: 269-274.
- Weinzimer S.A., Gibson T.B., Collett-Solberg P.F., Khare A., Liu B. y Cohen P. 2001. Transferrin is an

- insulin-like growth factor-binding protein-3 binding protein. *The Journal of Clinical Endocrinology and Metabolism* 86 (4): 1806-1813.
- Weir B.S. y Cockerman C.C. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38 (6): 1358-1370.
- Wenne R., Boudry P., Hemmer-Hansen J., Lubieniecki K.P., Was A. y Kause A. 2007. What role for genomics in fisheries management and aquaculture? *Aquatic Living Resources* 20: 241-255.
- Wilson A.J. y Ferguson M.M. 2002. Molecular pedigree analysis in natural population of fishes: approaches applications and practical considerations. *Canadian Journal of Fisheries and Aquatic Sciences* 59: 1696-1707.
- Winter G.W., Scherck C.B. y McIntyre J.D. 1980. Resistance of different stocks and transferrin genotypes of coho salmon, *Oncorhynchus kisutch*, and steelhead trout, *Salmo gairdneri*, to bacterial kidney disease and vibriosis. *Fishery Bulletin* 77: 795-802
- Wohlfarth G., Moav R. y Hulata G. 1983. A genotype-environment interaction for growth rate in the common carp, growing in intensively manured ponds. *Aquaculture* 33:187-195.
- Wright S. 1951. The genetical structure of populations. *Annals of Eugenics* 15: 323-354.
- Wu Y.P., Pi J.S., Pan A.L., Pu Y.J., Du J.P., Shen J., Liang Z.H. y Zhang J.R. 2012. An SNP in the MyoD1 gene intron 2 associated with growth and carcass traits in the three duck populations. *Biochemical Genetics* 50 (11-12): 898-907.
- Xia J.H., He X.P., Bai Z.Y. y Yue G.H. 2011. Identification and characterization of 63 MicroRNAs in the Asian sea bass *Lates calcarifer*. *PLoS ONE* 6 (3): e17537, doi:10.1371/journal.pone.0017537.
- 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): e54174, doi:10.1371/journal.pone.0054174.
- Yan X., Ding L., Li Y., Zhang X., Liang Y., Sun X. y Teng C. 2012. Identification and profiling of MicroRNAs from skeletal muscle of the Common Carp. *PLoS ONE* 7 (1): e30925, doi:10.1371/journal.pone.0030925.
- Yan B., Guo J., Zhu C., Zhao L. y Zhao J. 2013. miRNA-203b: a novel regulator of MyoD expression in tilapia skeletal muscle. *The Journal of Experimental Biology* 216: 447-451.
- Yeh H.Y. y Klesius P. 2008. Molecular cloning, sequencing and characterization of channel catfish (*Ictalurus punctatus*, R. 1818) Cathepsin S gene. *Veterinary Immunology and Immunopathology* 126: 382-387.
- Yu Li., Bai J., Ye X., Li S. y Li X. 2009. Molecular cloning and identification of single nucleotide polymorphisms of *MyoD* in largemouth bass (*Micropterus salmoides*). *Journal of Fisheries of China* 33: 1-8.
- Yúfera M., Halm S., Beltran S., Fusté B., Planas J.V. y Martínez-Rodríguez G. 2012. Transcriptomic characterization of the larval stage in gilthead seabream (*Sparus aurata*) by 454 pyrosequencing. *Marine Biotechnology* 14 (4): 423-435.
- Zhao S. y Fernald R.D. 2005. Comprehensive algorithm for quantitative real-time polymerase chain reaction. *Journal of Computational Biology* 12: 1047-1064.
- Zohar Y., Abraham M. y Gordin H. 1978. The gonadal cycle of the captivity-reared hermaphroditic teleost *Sparus aurata* (L.) during the first two years of life. *Annales De Biologie Animale Biochimie Biophysique* 18 (4): 877-882.

Zohar Y., R. Billard y C. Weil 1984. La reproduction de la daurade (*Sparus aurata*) et du bar (*Dicentrarchus labrax*): connaissance du cycle sexuel et contrôle de la gamétogenèse et de la ponte. En: L'Aquaculture du bar et des Sparidés. Págs. 3-24. (Eds.: Barnabé G. y Billard R). INRA Publ. Paris.

9. Anexo

*Artículos científicos publicados asociados
a esta tesis doctoral*



Assessment of parental contributions to fast- and slow-growing progenies in the sea bream *Sparus aurata* L. using a new multiplex PCR

Yaisel J. Borrell^a, Victor Gallego^b, Carmen García-Fernández^a, Ilaria Mazzeo^b, Luz Pérez^b, Juan F. Asturiano^b, Carlos E. Carleos^c, Emilia Vázquez^a, Jose A. Sánchez^a, Gloria Blanco^{a,*}

^a Laboratorio de Genética Acuícola, Departamento de Biología Funcional, Universidad de Oviedo, IUBA, 33071 Oviedo, Spain

^b Grupo de Acuicultura y Biodiversidad, Instituto de Ciencia y Tecnología Animal, Universidad Politécnica de Valencia, Camino de Vera s/n. 46022, Valencia, Spain

^c Grupo de Estadística Genómica, Departamento de Estadística, Universidad de Oviedo, 33007, Oviedo, Spain

ARTICLE INFO

Article history:

Received 12 July 2010

Received in revised form 10 January 2011

Accepted 20 January 2011

Available online 2 February 2011

Keywords:

Sparus aurata

Microsatellite markers

Multiplex PCR

OVIDORPLEX

Parentage

Inbreeding

ABSTRACT

Molecular tools to assist breeding programs in the gilthead sea bream (*Sparus aurata* L.) are scarce. A new multiplex PCR technique (OVIDORPLEX), which amplifies nine known microsatellite markers, was developed in this work. This multiplex system showed a high mean heterozygosity (>0.800) and a high mean number of alleles per marker (>14) when tested in two sea bream broodstocks (A: 40 breeders and B: 38 breeders). We tested this multiplex PCR for inferring parentage in a Spanish hatchery that graded the animals by size as part of their management procedure. The progeny of the broodstock were divided into fast- and slow-growth groups. Parentage studies revealed that this management procedure entailed a global reduction of the breeders' representation in progeny and that breeders' contributions were significantly unequal. Due to this, effective sample sizes fell to $N_e \approx 13$ –14 for fast- and $N_e \approx 18$ –24 for slow-growth progeny groups. These results imply a 3 to 4% rate of inbreeding per generation in the fast-growth group, which is more important to hatchery managers than the slow group. Not all the progeny were evaluated in this experiment (due to the discarding steps), and thus it is difficult to know if the phenotypic performance showed by the fast-growing progeny will be heritable. However, there were genetic differences between the differentiated growth progeny groups (fast vs. slow, F_{ST} values = 0.016 to 0.023; $P < 0.01$). We also identified breeders with significantly different contributions to the fast- (10 breeders) or to the slow- (15 breeders) growth progeny groups. Our results demonstrated that this new multiplex PCR could be useful for quantitative programs (breeding programs, detection of QTL, inbreeding control or reconstruction of fish genealogies) to improve the aquaculture of the gilthead sea bream (*S. aurata*).

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

In the last decade, geneticists have helped aquaculture managers by implementing selection and breeding programs (while preventing inbreeding) to obtain higher productivity and sustainability in fish hatcheries. The use of molecular markers has significantly helped this goal. In particular, the use of microsatellites has allowed the inference of effective breeding numbers (N_e) through parentage assignments in several species, including Atlantic salmon (*Salmo salar*) (Norris et al., 2000), turbot (*Scophthalmus maximus*) (Borrell et al., 2004), Atlantic cod (*Gadus morhua*) (Herlin et al., 2007, 2008), common dentex (*Dentex dentex*) (Borrell et al., 2008), and European sea bass (*Dicentrarchus labrax*) (Bardon et al., 2009). This approach allows pedigree reconstruction (indispensable for quantitative programs) and detection and avoidance of inbreeding in hatcheries. Fish are highly fecund, and with

only a few breeders it is possible to obtain high numbers of eggs and sperm. However, high levels of inbreeding can cause a significant decrease in growth rates or other production phenotypes (Kincaid, 1976, 1983; Gjerde et al., 1983; Sbordoni et al., 1986; Su et al., 1996; Evans et al., 2004).

The gilthead sea bream (*Sparus aurata* L.) is a member of the family Sparidae and one of the most important farmed fish in Europe (especially in the Mediterranean area). Its European production doubled between 1999 [46,353 tons (t)] and 2006 (86,665 t). The major producers of this fish are Greece (56%), Spain (24%) and Italy (10%) (Moretti et al., 1999; FEAP, 2008). Although traditionally considered a delicacy in the Mediterranean area, gilthead sea bream is now widely available in most parts of Europe. This fish is a protandrous hermaphroditic, mass-spawning species in which individuals are males during the first two years of life and then gradually become females. Members of the species breed once a year during a six- to eight-week period (Zohar et al., 1978, 1995). Some animals either delay or never attain sex reversal, possibly due to social, environmental, and/or genetic factors (Zohar et al., 1995). In its natural

* Corresponding author. Tel.: +34 985 103889; fax: +34 985 103534.
E-mail address: gloriablanc@uniovi.es (G. Blanco).

environment, the gilthead sea bream spawns between October and January, but spawning occurs later in the eastern Mediterranean, where mean water temperatures are higher. Spawning in the wild occurs in large groups or schools. A minimum of five to seven fish appears to be necessary to reduce stress and induce natural spawning in artificial environments (Brown, 2003 and references therein).

There is a lack of efficient breeding programs for the gilthead sea bream. Although estimates of genetic variance, heritability, and phenotypic and genetic correlations between traits are now becoming available for this species (e.g., Navarro et al., 2009), breeding programs require funding and infrastructure, both of which are limited. In the most recent and complete study yet, Navarro et al. (2009) found low to medium heritability estimates in gilthead sea bream at harvest (509 days) for body weight, fork length, condition factor, gutted body weight, filet weight, dressing percentage, and filet percentage. Other authors have found significant heritabilities for these and other traits (Knibb et al., 1997; Knibb, 2000; Batargias, 1998; Afonso et al., 2000; Thorland et al., 2007). These data on heritability have stimulated the gilthead sea bream industry to improve production through the exploitation of additive genetic variation.

Useful molecular tools (e.g., efficient multiplex PCRs and typing of SNPs and QTL) to assist breeding programs in gilthead sea bream aquaculture are still scarce. Franch et al. (2006) reported the first linkage map for the gilthead sea bream using 204 previously known and novel microsatellites, and Cenadelli et al. (2007) identified 76 SNPs in this species. The use of microsatellites in multiplex PCR has great value allowing geneticists to reduce the cost per reaction and optimize analyses of genealogy reconstruction and genetic variability within a population (Navarro et al., 2008). Navarro et al. (2008) developed the first two microsatellite multiplex systems to infer kinships in the gilthead sea bream, and the first QTL for resistance to fish pasteurellosis recently were reported by Massault et al. (2010).

In light of these circumstances, we designed a multiplex PCR system including nine previously known microsatellites (from gilthead sea bream and related species) that show high levels of genetic variation (Borrell et al., 2007). Our goal was a quick, inexpensive and efficient genetic analysis that could be useful for quantitative programs in the culture of this species (breeding programs, detection of QTL, inbreeding control or reconstruction of fish genealogies). We tested the multiplex PCR system in a Spanish Mediterranean hatchery (Piscicultura Marina Mediterránea, S.L. Burriana, Castellón, Spain). This hatchery possessed two gilthead sea bream broodstocks (group A: 40 breeders and group B: 38 breeders) that were established in 2006 using a combinatorial optimization approach for guaranteeing the maximum amount of genetic variation, which results in low inbreeding values in future matings (Borrell et al., 2007). Untagged progeny from a single day of spawning ($n = 520$ fish) from these two sea bream broodstocks were raised to 18 months of age with two successive steps of phenotypic selection for size (to obtain fast- and slow-growth progeny groups). Fast-growth, which is intrinsically related to better food conversion, is an important trait for the sea bream industry; feeding represents approximately 40% of the production cost for gilthead sea bream (Jover, 2007). Fast-growth is also correlated with disease resistance or survival in several fish species (Gjedrem and Olesen, 2005) and shows fairly high heritability ($h^2 = 0.2$ to 0.4) and considerably high percentages of genetic gain per generation (Gjedrem and Thodesen, 2005). The size-based grading system used by the hatchery is common when rearing fish in communal conditions to avoid negative social interactions such as competition for access to food. However, this kind of management could potentially introduce bias in estimates of genetic parameters because tank effects might be confounded with genetic potential (Blonk et al., 2010).

As a result of the management procedure discussed above, the hatchery had progeny from both broodstocks showing fast- and slow-growth rates during culture. This system of fish classification is based

only on phenotypic values and includes several progeny discarding steps, which makes it difficult to estimate additive genetic components. However, genetic analysis of these samples using the new multiplex PCR technique OVIDORPLEX and a parentage study could help to answer relevant questions: Are there quantitative or qualitative genetic differences between the differentiated growth rate progeny? Are all the breeders contributing equally to the progeny? Are there breeders showing a significant differential parental contribution to the differentiated progeny growth groups examined here?

2. Materials and methods

2.1. Fish

From a total of 101 *Sparus aurata* adult individuals (80 females and 21 males) that were available in December of 2006 in a hatchery in Spain (Piscicultura Marina Mediterránea, S.L. Burriana, Castellón, Spain), two broodstocks (Broodstock A: 29 females, average weight $1998.2 \pm$ (standard error) 43.5 g and 11 males, average weight 1877.2 ± 83.4 g; Broodstock B: 28 females, average weight 1992.8 ± 50.1 g, and 10 males, average weight 1950.0 ± 76.3 g) were grouped using a combinatorial optimization approach that followed the procedures described in Borrell et al. (2007). Briefly, the best possible aggregations of approximately 40 breeders per broodstock were selected, taking into account the relatedness coefficients (Queller and Goodnight, 1989) among the breeders. Values of the Queller and Goodnight's relatedness coefficient range from -1 to $+1$. Theoretically, full-sib individuals share, on average, 50% of their genome constitution and hence the relatedness coefficient (r) among them should be near 0.5, while for pairs of half-sibs, it is expected that $r \approx 0.25$ and among pairs of unrelated individuals $r \approx 0$. A positive value of relatedness would therefore indicate greater than expected relatedness, and a negative value would suggest that two individuals are more divergent than expected (Queller and Goodnight, 1989). Relatedness coefficients ranged from $-0.39 < r < +0.83$ in the complete stock and were restricted to $-0.28 < r < +0.54$ in Broodstock A and $-0.25 < r < +0.43$ in Broodstock B.

The breeders within each tank were allowed to spawn freely in March of 2007. Fertilized eggs coming from a single day were collected and incubated under identical conditions in two tanks until hatching. This process was carried out on different days for the two different broodstocks, but the harvest was always in the middle of the spawning period because it has been suggested that majority of the breeders are contributing to descendants at that time (C. García-Fernández, Univ. of Oviedo, unpublished). All larvae were reared in isolated large tanks until June of 2007, 86 and 78 days post-hatching for batches A and B, respectively. At this point, untagged individuals were separated in fast- and slow-growth groups according to their body width distributions using different sieves. In batch A, fish with a body width greater than 4.5 mm (fast) or smaller than 3.5 mm (slow), were separated, whereas the batch B fish with a body width higher than 3.5 mm or smaller than 2.5 mm formed fast- and slow-growth groups, respectively.

The fish were moved to Universidad Politécnica de Valencia (Valencia, Spain) at the age of 165 and 157 days post-hatching, respectively, for batches A and B. They were released into 1750-l fiberglass tanks, where they were reared under intensive conditions. The tanks were in a recirculating marine water system (30-m³ capacity) with a rotary mechanical filter and a gravity bio-filter. All tanks were aerated, and the water temperature was maintained at 22.2 ± 2.7 °C by a heat pump that was installed in the system. The photoperiod was natural, and all tanks had similar light conditions. Fish were fed using commercial fish feed (Dibaq S.A., Segovia, Spain) by hand twice a day until apparent satiation. At an intermediate stage of growth (311 and 303 days post-hatching for batches A and B, respectively), 600 juveniles

were separated by weight into four groups: A12-Fast (A12F>41 g), A12-Slow (A12S<15 g), B12-Fast (B12F>31 g) and B12-Slow (B12S<12 g). The fish were placed in four 4000-l tanks and reared under intensive conditions to commercial size.

Around 540 fish ($n=150$ for A18F, $n=150$ for A18S, $n=120$ for B18F, and $n=120$ for B18S) were weighed at 528 and 520 days post-hatching for batches A and B, respectively. A small piece of the caudal fin was preserved in 96% ethanol for genetic analysis. Mean weight differences among groups were evaluated using the Kolmogorov–Smirnov test (a non-parametric test using two independent samples and the Z statistic), available in the SPSS 15.0 statistic software.

2.2. Microsatellite amplification and scoring

Genomic DNA was purified from a small piece of the caudal fin using Chelex[®] 100 (Walsh et al., 1991). We developed a novel multiplex PCR (OVIDORPLEX) using the Multiplex Manager 1.0 software (Holleley and Geerts, 2009) and the gilthead sea bream microsatellite data previously collected from more than 800 gilthead sea bream individuals (Borrell et al., 2007) (Fig. 1). OVIDORPLEX included nine microsatellite loci and fluorescent dyes: *VIC-SaGT1* (0.6 μ M), *VIC-SaGT26* (0.3 μ M), and *NED-SaGT41b* (0.5 μ M) (Batargias et al., 1999); *VIC-Pb-OVI-D106* (0.6 μ M), and *PET-Pb-OVI-D102* (0.4 μ M) (Piñera et al., 2006); *FAM-Dxd44* (0.4 μ M), and *FAM-Dpt3* (0.4 μ M) (De la Herrán et al., 2005) and *NED- μ 184* (0.3 μ M), and *PET- μ 190* (0.3 μ M) microsatellites (Power et al., 2003). The QIAGEN multiplex PCR kit protocol with an annealing temperature of 52 °C was used to obtain 15 μ l of the OVIDORPLEX amplifications. Parental and offspring genotypes were scored after the analysis of the amplification products on the ABI 3130XL Genetic Analyzer using Genemapper 4.0.

2.3. Genetic diversity analysis

The number of alleles at each microsatellite locus (N_A), the percentage of polymorphic loci ($P_{0.95}$), the proportion of individual samples that were heterozygous (direct count heterozygosity, H_o) and the unbiased estimate of heterozygosity (H_e) for each group were assessed using Cervus 3.0 (Marshall et al., 1998). The Fstat statistical package (2.93 version) (Goudet, 1995, 2001) was used to estimate the total variation in gene frequencies (F_{IT}), partitioned into components of variation occurring within (F_{IS}) and among (F_{ST}) samples for each locus following Weir and Cockerham (1984). Significance levels of F_{IS} were assessed through randomization of alleles (1000 times) within samples (Fstat) for each broodstock. Pairwise F_{ST} values between samples and P -values were calculated using Fstat (for significance levels of F_{ST} , multi-locus genotypes were randomized between pairs of samples (1000 permutations), and then the significance after a strict Bonferroni correction was calculated (Rice, 1989; Goudet, 2001)).

2.4. Parentage assignments

We used the Cervus 3.0 software (Marshall et al., 1998; Kalinowski et al., 2007) to assign parentage. This program calculates both the *a priori* polymorphic information content (PIC) for every locus from each broodstock and the total exclusionary power (E). In addition, the program simulates parental assignments. The total exclusionary power is defined as the probability of excluding an arbitrary, unrelated parent candidate. When multiple parent candidates are not excluded, the exclusionary approach is inadequate (Cervus 3.0, Marshall et al., 1998). The parentage assignment simulations were carried out, taking into account the number of breeders per broodstock. Ten thousand cycles of simulated assignments were carried out using 95% confidence intervals. Finally, after genotyping, all the offspring were assigned to the most likely candidate parent pair with sexes known. In the assignment procedures, we allowed for typing errors (0.05), as this dramatically

reduces the impact of two other possible causes of mismatches in parent–offspring relationships: mutations and null alleles (Marshall et al., 1998).

2.5. Estimating effective breeding numbers (N_e)

The rate of inbreeding (ΔF) and the effective population size (N_e) are related as $\Delta F=1/2 N_e$ (Falconer, 1989). We first used the classical formula for estimating N_e : $N_e=4(Nm \times Nf)/(Nm+Nf)$ where Nm =the number of male breeders and Nf =the number of female breeders (Falconer, 1989). We also used an approach previously assessed by Brown et al. (2005) and reassessed by J. Woolliams (Roslin Institute, pers. comm.). This approach takes into account the proportion of descendants left by “presumed” unrelated parents to the next generation following Hill (1979) and Woolliams and Bijma (2000); Woolliams proposed $\Delta F=\sum c_{i(m)}^2/8+1/(32m)+\sum c_{i(f)}^2/8+1/(32f)$; with $c_{i(m)}$ being the fractional contribution of males ($c_{i(m)}$) and females ($c_{i(f)}$) to offspring and m and f the number of males and females, respectively. This formula is Woolliams’s derivation from the Hill (1979) formulation in terms of offspring contributions, and has no selection component. It simplifies to Wright’s formula for Poisson family sizes and, even when it is needed to be treated with caution, it is a useful approximation to the problem of estimating N_e when unequal breeders contributions to offspring occur (J. Woolliams, Roslin Institute, pers. comm.; Borrell et al., 2008).

2.6. Differential parental contribution to progenies

Contingency tables were used to (a) assess whether breeders contributed equally to progenies in a global sense and also inside the different growth groups in each broodstock, and (b) evaluate different contributions of male and female breeders to either the fast- or slow-growth groups. For the last test, a re-sampling method (Patefield, 1981; Corral, 2005; Carleos, 2010) was used with the contingency table frequencies. Under the null hypothesis of equal contribution and homogeneity, 100,000 contingency tables were pseudo-randomly generated, while keeping the marginal totals constant (Patefield, 1981). For each breeder, the proportion of tables where the frequency in the fast-growth group is more extreme than or so extreme as in the observed contingency table was recorded. Slow-growth group frequencies were handled similarly. This procedure quantifies how likely it is to obtain the observed frequencies (or more extreme frequencies) for the corresponding breeder under the null hypothesis. Thus, low values (less than 0.05) indicate unlikely frequencies under the null hypothesis. The entire re-sampling procedure was carried out in the R statistical environment (R Development Core Team, 2009).

3. Results

3.1. The OVIDORPLEX and the genetic characterization of broodstocks and differentiated growth progeny groups

In this work, progeny from two *Sparus aurata* broodstocks were separated twice by size in a hatchery. Samples were taken at 18 months of development, a point at which we have observed significant weight differentiation between fast- and slow-growth groups in both broodstocks (A – A18F: 295.4 ± 4.4 g, A18S: 195.7 ± 2.5 g, $Z \text{ weight}_{A18F-A18S}=6.83$, $P<0.001$; B – B18F: 323.9 ± 4.8 g, B18S: 184.1 ± 2.9 g, $Z \text{ weight}_{B18F-B18S}=6.91$, $P<0.001$) (Table 1). We found significant weight differences between the corresponding size classes of A and B progeny groups ($Z \text{ weight}_{A18F-B18F}=4.42$, $P<0.001$; $Z \text{ weight}_{A18S-B18S}=3.28$, $P<0.01$).

Almost all of the individuals (516/520 offspring and the 78 breeders) were successfully genotyped at nine microsatellite loci using OVIDORPLEX (Fig. 1). The system was reliable even with the quick and inexpensive Chelex DNA extraction method that results in

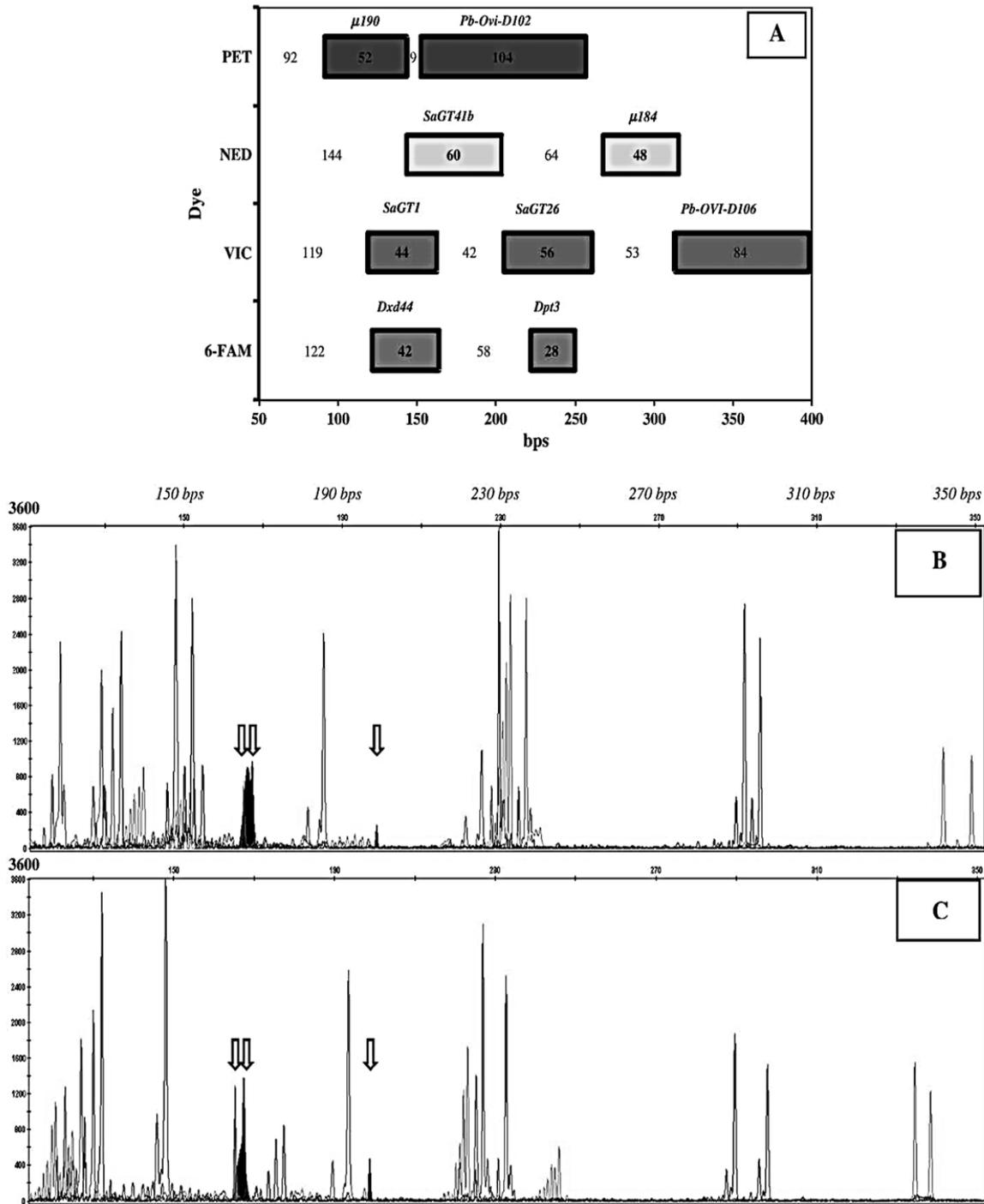


Fig. 1. The OVIDORPLEX system for genotyping *Sparus aurata* L. individuals with nine microsatellites. A: Diagram showing the design of the new multiplex PCR (OVIDORPLEX) using the Multiplex Manager software (Holleley and Geerts, 2009). The base-pair space between markers is indicated by numbers. B and C: Electrophoretograms of two individuals (breeders) using the ABI3130XL Genetic Analyzer and the Genemapper 4.0 software (three artifact amplifications found in this work are shown (indicated by arrows and filled): VIC-164bps, NED-167bps and 6-FAM-199bps).

low quality DNA. We detected three small artifact peaks (channels of VIC-165bps, NED-167bps and 6-FAM-199bps), but they did not complicate the genotyping process (Fig. 1).

The genetic characteristics of all the groups analyzed are presented in Table 1. Breeders showed high levels of genetic variation ($N_A > 14$ and $H_o > 0.800$). We also found agreement with the Hardy–Weinberg

expectations for populations under equilibrium (non-significant F_{IS} values). Both breeder groups were found to be genetically similar ($F_{ST} = -0.005$, $P = 0.996$). In terms of quantitative and/or qualitative genetic differences between the fast- and slow-growth groups in the two analyzed broodstocks, we did not observe heterozygosity level differences between them (Table 1). At 18 months of development, all

Table 1

Growth, genetic variability parameters and effective breeding number (N_e) estimations in *S. aurata* L. after genetic analysis using nine microsatellites included in the multiplex PCR OVIDORPLEX.

Broodstock	A			B		
	P	Fast	Slow	P	Fast	Slow
Collection day (approx. months)	–	528	528	–	520	520
n	40	139	139	38	118	120
Mean weight Δ (g)	1965.0	295.4	195.7	1981.5	323.8	184.1
St. dev.	249.4	52.3	29.5	256.6	52.7	32.4
St. error	39.4	4.4	2.5	41.6	4.8	2.9
N_a	14.6	12.2	14.0	14.0	12.2	11.8
H_e	0.867	0.801	0.849	0.872	0.832	0.842
H_o	0.883	0.867	0.865	0.858	0.893	0.859
F_{IS} (p values)	–0.028 n.s (0.0843)	–0.085*** (0.0009)	–0.019* (0.0407)	+0.011 n.s (0.7300)	–0.073*** (0.0009)	–0.022* (0.0315)
Sim	99.9%	–	–	99.9%	–	–
Breeders \square	29♀×11♂	11♀×8♂	19♀×10♂	28♀×10♂	14♀×8♂	14♀×9♂
$N_e(1)$	31.2	18.5	26.2	29.4	20.3	21.9
$\Delta F(1)$ (%)	1.60	2.70	1.90	1.70	2.46	2.28
$N_e(2)$	–	12.7	24.6	–	14.2	17.6
$\Delta F(2)$ (%)	–	3.91	2.02	–	3.50	2.83

P: parental breeders. n: number of individuals. Mean weight Δ : at December 2006 for breeders and at day of collection for progenies. St. dev.: standard deviation. St. error: standard error. N_a : number of alleles per locus; H_e : expected heterozygosity; H_o : observed heterozygosity. F_{IS} : degree of departure from expected Hardy–Weinberg proportions within groups. Sim: percentages of times that the true parental pair was found after ten thousand cycles of simulated assignments using 95% confidence intervals. Breeders \square : all the possible breeders among parents, breeders that left progenies after parental assignment assessments using microsatellites in progenies. $N_e(1)$: effective population sizes using $4(Nm \times Nf)/(Nm + Nf)$ (Falconer, 1989). $\Delta F(1)$: rate of inbreeding using $\Delta F = 1/2N_e$ (Falconer, 1989). $N_e(2)$: effective population sizes using $\Delta F(2) = \sum c_{im}^2/8 + 1/(32m) - \sum c_{if}^2/8 + 1/(32f)$ (Brown et al., 2005; Woolliams J., pers. comm.) and then $\Delta F = 1/2N_e$ (Falconer, 1989).

* $P < 0.05$.

*** $P < 0.001$. n.s = not significant.

samples showed significant H–W disequilibrium due to an excess of heterozygotes. This excess was more significant in the fast- ($P < 0.001$) than the slow-growth progeny ($P < 0.05$) (Table 1). Fast- and slow-growth progeny were significantly differentiated within broodstocks ($F_{ST\ A18F-A18S} = 0.023$, $P = 0.003$; $F_{ST\ B18F-B18S} = 0.016$, $P = 0.003$). Corresponding fast- or slow-growth groups were genetically distinct between broodstocks ($F_{ST\ A18F-B18F} = 0.043$, $P = 0.003$; $F_{ST\ A18S-B18S} = 0.036$, $P = 0.003$).

3.2. Parentage assignments and effective breeding number (N_e) estimations

The parentage assignment procedures carried out here revealed that the true parental pair was expected, by simulation, to be found 99.9% of the time using the multiplex PCR including nine microsatellite loci (Table 1). Almost all the progeny (516 of 520 collected after 18 months of development, 99.2%) were assigned to a parental pair with certainty. In the A broodstock, 11 breeders did not produce progeny (27.5%: 10 ♀ and 1 ♂), while in the B broodstock, 12 breeders failed to contribute to offspring (31.5%: 11 ♀ and 1 ♂).

The effective breeding numbers, taking into account the numbers of all males and females in the tank, would be $N_e = 31$ and $N_e = 29$ in the A and B broodstocks, respectively. However, not all the breeders contributed to offspring, and this implied reductions in the N_e estimates (A18F: $N_e \approx 18$, A18S: $N_e \approx 26$; B18F: $N_e \approx 20$, and B18S: $N_e \approx 22$) (Table 1). We observed a globally significantly unequal contribution of breeders to progeny in both broodstocks (A: 19 females, 14.6 descendants expected for each female assuming equal contributions, $\chi^2 = 356.3$, $P < 0.001$; 10 males, 27.8 descendants expected for each male if there are equal contributions, $\chi^2 = 221.6$, $P < 0.001$) (B: 17 females, 14.0 descendants are expected for each female assuming equal contributions, $\chi^2 = 281.1$, $P < 0.001$; 9 males, 26.4 descendants expected by each male if contributions are equal, $\chi^2 = 230.9$, $P < 0.001$). A similar result was found within growth groups. Unequal reproductive success affected the N_e estimates. The “global” effective breeding numbers calculated here using Woolliams’ approach revealed significantly lower N_e values than expected with equal contributions, primarily in the fast-growth groups (A18F: $N_e \approx 13$, A18S: $N_e \approx 24$; B18F: $N_e \approx 14$, and B18S: $N_e \approx 18$) (Table 1).

3.3. Differential parental contributions to fast- and slow-growth progeny groups

The breeders did not contribute equally to fast- and slow-growth progeny in this experiment (A: females $\chi^2 = 114.2$, $P < 0.001$; males $\chi^2 = 69.2$, $P < 0.001$) (B: females $\chi^2 = 84.3$, $P < 0.001$; males $\chi^2 = 50.6$, $P < 0.001$). This observation means that we found differential parental contributions to the growth performance groups under study in both broodstocks (Fig. 2).

The individual breeder’s χ^2 values were analyzed using a resampling method to identify the breeders responsible for differential contribution to the fast- or slow-growth progeny groups. At 18 months of age, eight out of 19 females and six out of 10 males left descendants at a differential rate in fast- and slow-growth groups in the A broodstock (14/29 = 48%, 14 breeders left 76% of the descendants), while seven out of 17 females and four out of nine males were identified in the B broodstock (11/26 = 42%, 11 breeders were responsible for 74% of the descendants) (Fig. 2). To summarize, we identified 10 breeders (six females and four males) that contributed significantly more to the fast-growth progeny and 15 that contributed significantly more to the slow-growth progeny; 30 breeders left descendants in equal proportions to the two growth groups under study, while 23 breeders had null contributions in this experiment.

4. Discussion

4.1. The multiplex PCR system

The multiplex PCR developed in this work (OVIDORPLEX) was highly effective using our samples, and it was less costly in terms of both time and money than scoring each marker with independent PCRs and gel electrophoreses. This multiplex system showed high mean heterozygosity (>0.800) and high mean number of alleles per marker (>14) when tested in two gilthead sea bream broodstocks (A: 40 breeders and B: 38 breeders). These variability values are higher than those reported for the Rim-A and Rim-B multiplex PCR by Navarro et al. (2008) and allow highly accurate determination of parentage, which is important for correctly reconstructing fish

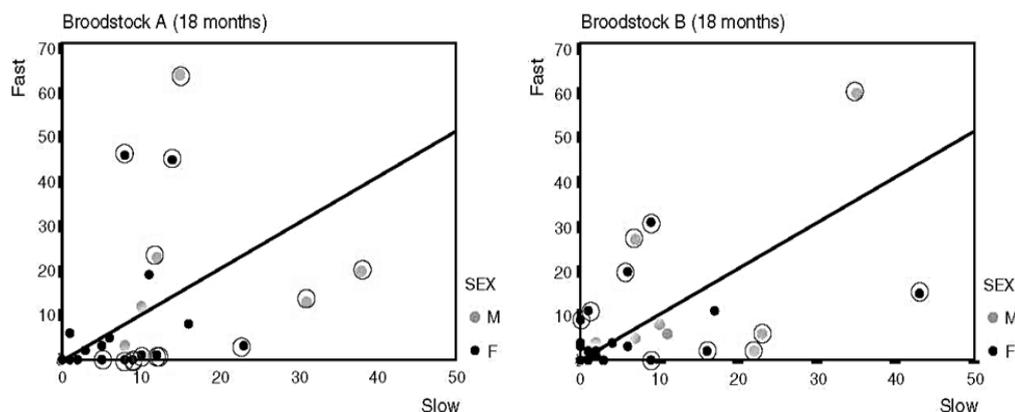


Fig. 2. The *Sparus aurata* L. breeder contributions to fast- (Y axis) and to slow-growth performances (X axis) in progenies (in numbers of offspring) at 18 months of development (lines represent identical contributions to both growth groups). In circles, those breeders showing extreme observed χ^2 values compared to the statistic obtained from data sets after runs of 100,000 iterations ($P < 0.05$). (M: males, F: females).

genealogies and for estimating genetic parameters in breeding programs.

It is possible to adjust OVIDORPLEX for use with high-quality DNA (e.g., that produced by the Qiagen DNA extraction kit). This adjustment entails the use of lower amounts of PCR primers and/or raising the annealing temperature (OVIDORPLEX works well in the 52 to 55 °C range), which should produce even better results by doubling the intensities of the microsatellite signals and diminishing the signals from artifact peaks. Amplification failures of any of the loci used in OVIDORPLEX would not significantly affect parentage studies. The use of OVIDORPLEX-R (with R standing for reduced) using only five of the nine microsatellites included in OVIDORPLEX (*Dxd44*, *Dpt3*, *SaGT41b*, *SaGT2*, and *Pb-OVI-D106*) in parentage assignment simulations resulted in approximately 95% of parent pairs being correctly identified in our two broodstocks. However, simulated data are not real data. Null alleles or degraded tissues that result in a low PCR amplification or failure to amplify some alleles can affect genotyping and parentage assignments (Borrell et al., 2004).

4.2. The breeding aspects

The hatchery Piscicultura Marina Mediterránea, S.L. (Burriana, Castellón, Spain) uses a size-based grading system for rearing fish. As a result of this system, the hatchery has progeny from two broodstocks potentially showing fast- and slow-growth rates during culture. We used the OVIDORPLEX system for performing parentage studies, which revealed that the size-based hatchery management procedure produces a significant reduction in the breeder's participation in the maintenance of the A and B broodstocks, and significantly unequal breeders' contributions to progeny. This outcome seriously affected the effective breeding number estimates. Effective sample sizes fell to $N_e \approx 13$ –14 for fast- (representing only 32 to 36% of all the breeders involved) and $N_e \approx 24$ (60%)/ $N_e \approx 18$ (47%) for slow-growth progeny groups in the A and B broodstocks, respectively. Our results imply a 3 to 4% rate of inbreeding in the fast-growth groups, which are more important to hatchery managers than the slow-growth groups. These inbreeding levels are similar to those previously found by Brown et al. (2005) in this species without a selection strategy, although they used a higher number of breeders per broodstock ($n = 48$ to 58; Brown et al., 2005). Probably, the initial broodstock composition (maximized for genetic variation in our case) had positively influenced the accumulation of inbreeding to levels comparable to those from Brown et al.'s study (2005), even when two steps of phenotypic selection were performed here. As a general rule, there should be only a 0.5% of rate of inbreeding per generation (ΔF) in breeding schemes (Sonesson et al., 2005). Tave (1999) has

suggested that managers from small, medium and large hatcheries wanting minimum risk should avoid 5% of inbreeding accumulation over generations, although this would require more management effort. More than a 12% accumulation of inbreeding in rainbow trout allows inbreeding depression of valuable fitness traits (Gjerde et al., 1983; Fjalestad, 2005). A rate of 5% of accumulated inbreeding was almost reached in this study in only one generation. If the progeny showing fast-growth are to be used as breeders in upcoming culture cycles, some strategies should be considered to diminish inbreeding. To introduce wild breeders into the culture cycles is a common procedure, although this approach affects not only the rates of inbreeding but probably also the genetic gain, which will likely be reduced. One plausible remedy could be the rotational line crossing (Kincaid, 1977; Tave, 1999) using genetically unrelated fast-growth progeny from the two broodstocks (the pedigree is now available). Cohorts can be quite useful for preventing inbreeding over four to six generations and minimizing inbreeding for several generations thereafter (Tave, 1999).

In addition to the challenge of avoiding inbreeding accumulation for upcoming culture cycles using this type of hatchery management strategy, the prevention of environmental sources of variation in growth from becoming confounded with heritable differences during selection is a serious concern. The most problematic issue is that not all the progeny were evaluated in this experiment (due to the discarding steps), and it is thus difficult to know whether the phenotypic performance showed by the fast-growing progeny would be heritable. In any case, the genetic characterization of the samples and the parentage study carried out here could yield insights into it as suggested below.

The genetic assessment of the broodstocks and their progeny (using nine randomly chosen microsatellite markers) revealed H–W equilibrium for the two broodstocks under study. However, significant excesses in heterozygosity were found in 18-month-old progeny; this is a characteristic usually seen in recently bottlenecked populations (Luikart and Cornuet, 1997) (it also may be attributable to heterozygotes advantage or selection acting against inbred (i.e., overly homozygous) individuals). More relevantly, fast- and slow-growth groups were genetically different in both broodstocks (F_{ST} values = 0.016 to 0.023; $P < 0.01$). It has been suggested that a correlation between Q_{ST} values (a measure of differentiation in quantitative loci) and F_{ST} values is generally expected (Merila and Cnokrak, 2001), which implies that molecular markers may be used as an index of the degree of differentiation at quantitative loci (Roff, 2003; Leinonen et al., 2008). Thus, genetic differences found between the fast- and slow-growth groups in these two broodstocks' progeny using neutral molecular markers also suggest differentiation at QTL. Moreover, we identified

breeders with significantly different contributions to fast- (10 breeders) and to slow- (15 breeders) growth groups. The rest of the breeders (67%) either did not contribute to progeny or contributed in an equal manner to both growth groups. These findings could lead us to expect that a genetic component is guiding the growth performance observed here (fast-growth often shows high heritability in fish; $h^2 = 0.2$ to 0.4 ; Gjedrem and Thodesen, 2005). However, our results have an implicit sampling effect (as a consequence of using only one-day spawning eggs, the posterior discarding steps, or even the sampling for genetic analyses). A more complete design should take into account that the spawning dynamics in *Sparus aurata* seem to be sequential (unpublished results). Thus, a mix of eggs from several days and several points during the spawning season would be the best choice for maintaining as much of the genetic variation of the parental stock as possible in the progeny. Preserving samples from all the phenotypic classes will help to correctly evaluate the breeders' values for fast-growth, also allowing an estimation of additive genetic components. Sonesson (2003) suggested that it is possible to raise the selection intensity in multi-spawners such as the gilthead sea bream using progeny testing. This increased selection intensity could help save space and money during culture by preserving only the valuable breeders for upcoming culture cycles.

In summary, this work demonstrated that a new multiplex PCR technique, OVIDORPLEX, is a quick and inexpensive method of genetic analysis that allows parentage studies that provide information about breeders' contributions. This information can be used to assess the population's N_B , which provides a good estimate of inbreeding, a process that should be controlled in hatcheries performing any kind of management strategies. Parentage studies using microsatellite loci also facilitate the identification of breeders that contribute significantly to the phenotypes of interest and allow an efficient genetic characterization of the samples. The findings of this study should be relevant for the establishment of successful breeding programs in aquaculture of the gilthead sea bream.

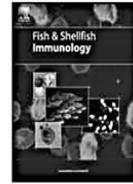
Acknowledgements

This work was carried out in collaboration with the fish farm Granja Marina Safor, S.L. (Gandía, Valencia, Spain) and the hatchery Piscicultura Marina Mediterránea, S.L. (Burriana, Castellón, Spain). It was financed by JACUMAR (the PROGensa project) and the Spanish Ministry of Science and Innovation (MICINN; National Program of Resources and Food and Agriculture Technologies, AGL2006-13411-C03-00, SELECTSPARUS, and AGL2007-64040-C03-00, SELECTBREAM, including European Regional Development Funds). V. Gallego was supported by a FPI scholarship financed by MICINN; C. García-Fernández was supported by a FPU scholarship financed by MICINN; and I. Mazzeo was supported by a FPI scholarship financed by Generalitat Valenciana. We are indebted to three anonymous referees and the journal editor for valuable comments.

References

- Afonso, J.M., Montero, D., Robaina, L., Astorga, N., Izquierdo, M.S., Ginés, R., 2000. Association of a lordosis-scoliosis-kyphosis deformity in gilthead seabream (*Sparus aurata*) with family structure. *Fish Physiol. Biochem.* 22, 159–163.
- Bardon, A., Vandeputte, M., Dupont-Nive, M., Chavanne, H., Haffray, P., Vergnet, A., Chatain, B., 2009. What is the heritable component of spinal deformities in the European sea bass (*Dicentrarchus labrax*)? *Aquaculture* 294 (3–4), 194–201.
- Batargias, C., 1998. Genetics of gilthead seabream (*Sparus aurata*). Study of microsatellites and their use for the estimation of genetic parameters of growth and other quantitative characters. Ph.D. Thesis, University of Crete, Greece.
- Batargias, C., Dermizakis, E., Magoulas, A., Zouros, E., 1999. Characterization of six polymorphic microsatellite markers in gilthead seabream, *Sparus aurata* (Linnaeus 1758). *Mol. Ecol. Notes* 8, 897–899.
- Blonk, R.J.W., Komen, H., Kamstra, A., Van Arendonk, J.A.M., 2010. Effects of grading on heritability estimates under commercial conditions: a case study with common sole, *Solea solea*. *Aquaculture* 300, 43–49.
- Borrell, Y.J., Álvarez, J., Vázquez, E., Fernández, C., Martínez, C., Sánchez, J.A., Blanco, G., 2004. Applying microsatellites to the management of turbot stocks (*Scophthalmus maximus* L.) in hatcheries. *Aquaculture* 241, 133–150.
- Borrell, Y.J., Carleos, C.E., Asturiano, J.F., Bernardo, D., Vázquez, E., Corral, N., Sánchez, J.A., Blanco, G., 2007. Use of microsatellites and a combinatorial optimization approach in the acquisition of gilthead seabream (*Sparus aurata* L.) broodstocks for hatcheries. *Aquaculture* 269, 200–210.
- Borrell, Y.J., Blanco, G., Vázquez, E., Piñera, J.A., Giménez, G., Estévez, A., Sánchez, J.A., 2008. Assessing the spawning season in common dentex (*Dentex dentex*) using microsatellites. *Aquacult. Res.* 39 (12), 1258–1267.
- Brown, R.C., 2003. Genetic management and selective breeding in farmed populations of gilthead seabream (*Sparus aurata*). Ph.D. thesis. University of Stirling, Stirling, UK.
- Brown, R.C., Woolliams, J.A., McAndrew, B.J., 2005. Factors influencing effective population size in commercial populations of gilthead seabream, *Sparus aurata*. *Aquaculture* 247, 219–225.
- Carleos, C., 2010. Pij-valor: cómo expresar la rareza de la frecuencia de una cierta celda en una tabla de contingencia. *Teleskopo* 2, 36–46.
- Cenadelli, S., Maran, V., Bongioni, G., Fusetti, L., Parma, P., Aleandri, R., 2007. Identification of nuclear SNPs in gilthead seabream. *J. Fish Biol.* 70, 399–405.
- Corral, N., 2005. Detección de Alelos Atípicos en Estudios de Homogeneidad. Libro de resúmenes, p. 110. X Conferencia Española de Biometría, Oviedo, Spain.
- De la Herrán, R., Magoulas, A., Garrido-Ramos, M.A., Ruiz-Rejón, C., Ruiz-Rejón, M., Zouros, E., 2005. Desarrollo de Microsatélites en tres Especies de Espáridos con Interés en Acuicultura. La acuicultura como actividad económica de las zonas costeras. IX Congreso Nacional de Acuicultura, pp. 113–136. Eds. Consejería de Agricultura y Pesca. Junta de Andalucía, Sevilla, Spain.
- Evans, F., Matson, S., Brake, J., Langdon, C., 2004. The effect of inbreeding on performance traits of adult Pacific oysters (*Crassostrea gigas*). *Aquaculture* 230, 89–98.
- Falconer, D.S., 1989. Introduction to Quantitative Genetics, 3rd ed. Longman, New York.
- FEAP (The Federation of European Aquaculture Producers), 2008. National Aquaculture Production. [Online]. available http://www.feap.info/production/countries/default_en.asp. [Accessed Jan. 10, 2010].
- Fjalestad, K.T., 2005. Breeding strategies. In: Gjedrem, T. (Ed.), Selection and Breeding Programs in Aquaculture. Springer, Dordrecht, The Netherlands.
- Franch, R., Louro, B., Tsalavouta, M., Chatzifilis, D., Tsigenopoulos, C.S., Sarropoulou, E., Antonello, J., Magoulas, A., Mylonas, C.C., Babbucci, M., Patarnello, T., Power, D.M., Kotoulas, G., Bargelloni, L., 2006. A genetic linkage map of the hermaphrodite teleost fish *Sparus aurata* L. *Genetics* 174 (2), 851–861.
- Gjedrem, T., Olesen, I., 2005. Basic statistical parameters. In: Gjedrem, T. (Ed.), Selection and Breeding Programs in Aquaculture. Springer, Dordrecht, The Netherlands.
- Gjedrem, T., Thodesen, J., 2005. Selection. In: Gjedrem, T. (Ed.), Selection and Breeding Programs in Aquaculture. Springer, Dordrecht, The Netherlands.
- Gjerde, B., Gunnes, K., Gjedrem, T., 1983. Effect of inbreeding on survival and growth in rainbow trout. *Aquaculture* 34, 327–332.
- Goudet, J., 1995. FSTAT (vers. 1.2): a computer program to calculate *F*-statistics. *J. Hered.* 86, 485–486.
- Goudet, J., 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available from <http://www.unil.ch/izea/software/fstat.html>
- Herlin, M., Taggart, J.B., McAndrew, B.J., Penman, D.J., 2007. Parentage allocation in a complex situation: a large commercial Atlantic cod (*Gadus morhua*) mass spawning tank. *Aquaculture* 272 (S1), S195–S203.
- Herlin, M., Delghandi, M., Wesmajervi, M., Taggart, J.B., McAndrew, B.J., Penman, D.J., 2008. Analysis of the parental contribution to a group of fry from a single day of spawning from a commercial Atlantic cod (*Gadus morhua*) breeding tank. *Aquaculture* 274, 218–224.
- Hill, W.G., 1979. A note on effective population size with overlapping generations. *Genetics* 92, 317–322.
- Holleley, C.E., Geerts, P.G., 2009. Multiplex Manager 1.0: a crossplatform computer program that plans and optimizes multiplex PCR. *Biotechniques* 46 (7), 511–517.
- Jover, M., 2007. Alternativas de futuro para la producción de la dorada (*Sparus aurata*). In: Cervoño-Eiroa, A., Guerra-Díaz, A., Pérez-Acosta, C. (Eds.), Cultivando el futuro: Proceedings of the XI Congreso Nacional de Acuicultura, Vigo, Spain, 24–28 September 2007, pp. 1337–1342.
- Kalinowski, S.T., Taper, M.L., Marshall, T.C., 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol. Ecol.* 16, 1099–1106.
- Kincaid, H.L., 1976. Effects of inbreeding on rainbow trout populations. *Trans. Am. Fish. Soc.* 105, 273–280.
- Kincaid, H.L., 1977. Rotational line crossing: an approach to the reduction of inbreeding accumulation in trout brood stocks. *Prog. Fish Cult.* 39, 179–181.
- Kincaid, H.L., 1983. Inbreeding in fish populations used for aquaculture. *Aquaculture* 33, 215–227.
- Knibb, W., 2000. Genetic improvement of marine fish — which method for the industry? *Aquacult. Int.* 31, 11–23.
- Knibb, W., Gorskova, G., Gorskov, S., 1997. Selection for growth in the gilthead seabream (*Sparus aurata*). *Isr. J. Aquacult. Bamidgeh* 49, 57–66.
- Leinonen, T., O'Hara, R.B., Cano, J.M., Merila, J., 2008. Comparative studies of quantitative trait and neutral marker divergence: a meta-analysis. *J. Evol. Biol.* 21, 1–17.
- Lukart, G., Cornuet, J.M., 1997. Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. *Conserv. Biol.* 12 (1), 228–237.
- Marshall, T.C., Slate, J., Kruuk, L.E.B., Pemberton, J.M., 1998. Statistical confidence for likelihood-based paternity inference in natural populations. *Mol. Ecol.* 7, 639–655.

- Massault, C., Franch, R., Haley, C., De Koning, D.J., Bovenhuis, H., Pellizzari, C., Patarnello, T., Bargelloni, L., 2010. Quantitative trait loci for resistance to fish pasteurellosis in gilthead sea bream (*Sparus aurata*). *Anim. Genet.* doi:10.1111/j.1365-2052.2010.02110.x.
- Merila, J., Crnokrak, P., 2001. Comparison of genetic differentiation at marker loci and quantitative traits. *J. Evol. Biol.* 14, 892–903.
- Moretti, A., Pedini Fernandez-Criado, M., Cittolin, G., Guidastri, R., 1999. Manual on Hatchery Production of Seabass and Gilthead Seabream. Volume 1. FAO, Rome.
- Navarro, A., Badilla, R., Zamorano, M.J., Pasamontes, V., Hildebrandt, S., Sánchez, J.J., Afonso, J.M., 2008. Development of two new microsatellite multiplex PCRs for three sparid species: Gilthead seabream (*Sparus auratus* L.), red porgy (*Pagrus pagrus* L.) and redbanded seabream (*P. auriga*, Valenciennes, 1843) and their application to paternity studies. *Aquaculture* 285, 30–37.
- Navarro, A., Zamorano, M.J., Hildebrandt, S., Ginés, R., Aguilera, C., Afonso, J.M., 2009. Estimates of heritabilities and genetic correlations for growth and carcass traits in gilthead seabream (*Sparus auratus* L.), under industrial conditions. *Aquaculture* 289, 225–230.
- Norris, A.T., Bradley, D.G., Cunningham, E.P., 2000. Parentage and relatedness determination in farmed Atlantic salmon (*Salmo salar*) using microsatellite markers. *Aquaculture* 182, 73–83.
- Patefield, W.M., 1981. Algorithm AS159. An efficient method of generating $r \times c$ tables with given row and column totals. *Appl. Statist.* 30, 91–97.
- Piñera, J.A., Bernardo, D., Blanco, G., Vázquez, E., Sánchez, J.A., 2006. Isolation and characterization of polymorphic microsatellite markers in *Pagellus bogaraveo*, and cross-species amplification in *Sparus aurata* and *Dicentrarchus labrax*. *Mol. Ecol. Notes* 6 (1), 33–35.
- Power, D., Almeida, S., Louro, B.E.P., 2003. Seabream pituitary c-DNA EST-library. Genbank references: DN048405, DN048399.
- Queller, D.C., Goodnight, K.F., 1989. Estimating relatedness using genetic markers. *Evolution* 43, 258–275.
- R Development Core Team, 2009. R: A Language and Environment for Statistical Computing. ISBN 3-900051-07-0. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org>.
- Rice, W.R., 1989. Analysing tables of statistical tests. *Evolution* 43, 223–225.
- Roff, D., 2003. Evolutionary quantitative genetics. Are we in danger of throwing out the baby with the bathwater? *Ann. Zool. Fennici* 40, 315–320.
- Sbordoni, V., De Mattheis, M., Cobolli Sbordoni, M., La Rosa, G., Mattoccia, M., 1986. Bottleneck effects and the depression of genetic variability in hatchery stocks of *Penaeus japonicus* (Crustacea, Decapoda). *Aquaculture* 57, 239–251.
- Sonesson, A.K., 2003. Marker-assisted selection in fish – case studies. In: Guimarães, E., Scherf, B., Sonnino, A., Dargie, J. (Eds.), *Markers Assisted Selection. Current Status and Future Perspectives in Crops, Livestock, Forestry and Fish*. FAO, Rome.
- Sonesson, A.K., Woolliams, J.A., Meuwissen, T.H.E., 2005. Kinship, relationship and inbreeding. In: Gjedrem, T. (Ed.), *Selection and Breeding Programs in Aquaculture*. Springer, Dordrecht, The Netherlands.
- Su, G.S., Liljedahl, L.E., Gall, G.A.E., 1996. Effects of inbreeding on growth and reproductive traits in rainbow trout (*Oncorhynchus mykiss*). *Aquaculture* 142, 139–148.
- Tave, D., 1999. Inbreeding and brood stock management. Fisheries Technical Paper. No. 392. FAO, Rome. 122p.
- Thorland, I., Papaioannou, N., Kottaras, L., Refstie, T., Papisolomontos, S., Rye, M., 2007. Family-based selection for production traits in gilthead sea bream (*Sparus aurata*) and European sea bass (*Dicentrarchus labrax*) in Greece. *Aquaculture* 272 (1), S314.
- Walsh, P.S., Metzger, D.A., Higuchi, R., 1991. Chelex[®] 100 as a medium for simple extraction of DNA for PCR-based typing for forensic material. *Biotechniques* 10, 506–510.
- Weir, B.S., Cockerham, C.C., 1984. Estimating *F*-statistics for the analysis of population structure. *Evolution* 38, 1358–1370.
- Woolliams, J.A., Bijma, P., 2000. Predicting rates of inbreeding in populations undergoing selection. *Genetics* 154, 1851–1864.
- Zohar, Y., Abraham, M., Gordin, H., 1978. The gonadal cycle of the captivity-reared hermaphroditic teleost *Sparus aurata* (L.) during the first two years of life. *Ann. Biol. Anim. Biochem. Biophys.* 18, 877–882.
- Zohar, Y., Harel, M., Hassain, S., Tandler, A., 1995. Gilthead seabream (*Sparus aurata*). In: Bromage, N.R., Roberts, R.J. (Eds.), *Broodstock Management and Egg and Larval Quality*. Blackwell Science, Oxford.



Characterization of the gilthead seabream (*Sparus aurata* L.) transferrin gene: Genomic structure, constitutive expression and SNP variation

Carmen García-Fernández, José A. Sánchez, Gloria Blanco*

Laboratorio de Genética Acuícola, Departamento de Biología Funcional, Universidad de Oviedo, IUBA, 33071 Oviedo, Spain

ARTICLE INFO

Article history:

Received 7 March 2011
Received in revised form
1 July 2011
Accepted 2 July 2011
Available online 20 July 2011

Keywords:

Transferrin
SNP
Iron metabolism
Disease resistance
Gilthead seabream (*Sparus aurata* L.)

ABSTRACT

Transferrin (Tf) is a multi-function protein with a central role in iron metabolism, and it is this function that is associated with a role in the innate immune system response. The clear link between Tf and immune defense mechanism lead to propose Tf as a candidate gene for disease resistance. In this study, genomic and cDNA sequences of Tf gene in gilthead seabream (*Sparus aurata* L.) (*SalTf* gene), were identified and characterized. *SalTf* gene structure consists of a coding region of 2076 nucleotides divided into 17 exons and a no-coding region that includes 16 introns and spans 5495 nucleotides. The deduced Tf protein for gilthead seabream is composed of 691 amino acids and consists of an initial peptide and two lobes (N- and C-lobes). This gene structure is similar to that of previously described Tf genes in other fish species. RT-PCR analyses carried out in different tissues and two developmental stages showed tissue- and stage-specific Tf expression in gilthead seabream. Finally, by sequencing the transferrin genomic sequences of 20 unrelated seabreams, 31 SNPs were identified. These data allowed the estimation of the frequency of nucleotide substitution in the *SalTf* gene as 1SNP per 253 bp. SNPs were detected in different regions of the genomic sequence but they were mainly localized in non-coding regions, specifically, SNP frequency in non-coding regions was fifteen-fold higher than within coding regions.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

The innate immune system plays an important role in the first phase of the body's response against pathogenic invasion. Its response is non-specific, rapid and relatively temperature-independent. Non-specific immunity is especially important in fish (ectothermic organisms), in which it has been shown that exposure to low temperatures can cause a suppression of adaptive immune defenses [1–4]. Humoral innate immunity is mediated by soluble factors that inhibit microorganism growth [5]. Of these, serum transferrin (Tf) has generated particular interest.

Tf is a multi-function protein with a primary role in transporting iron in a safe, redox-inactive state from absorption to utilization or storage sites around the body [6,7]. In addition to being essential for organismal development, iron levels must be tightly regulated. An excess of iron in the blood can result in oxidative damage by promoting free radical formation [8,9]. Therefore, Tf plays an important function in maintaining proper iron concentrations in the blood. Besides its iron affinity, Tf is able to transport other heavy

metals like cadmium or zinc. The increase of serum Tf levels in Nile tilapia (*Oreochromis niloticus*) as a result of exposure periods to these metals suggest the use of Tf as biomarker parameter of heavy-metal toxicity in fish [10].

With respect to protein structure, Tf is defined as a glycoprotein with two functional domains, the N- and C-lobes, separated by a short, inter-domain bridge, the length of which seems to be species-dependent [11,12]. Each lobe contains a high affinity reversible Fe^{3+} binding site. This structure has been proposed to result from a duplication and fusion of an ancestral gene, increasing the number of iron binding sites and therefore protein functionality [13,14]. Tf is mainly synthesized in the liver, but mRNA transcripts have also been isolated from different tissues [15], including the brain, where Tf expression seems to be species-dependent [16].

The association of Tf with the immune system derives from its ability to restrict serum free-iron levels, creating low-iron environments where the infection capacity of pathogenic microorganisms is limited [17,18]. Specific functional regions of the Tf protein seem to be undergone positive natural selection in salmonids and goldfish (*Carassius auratus*), showing a possible relationship between Tf and resistance to pathogens in fish [19,20]. In fact, in several fish species like salmonids and common carp (*Cyprinus carpio carpio* L.), associations between Tf polymorphism and

* Corresponding author. Tel.: +34 985 103889; fax: +34 985 103534.
E-mail address: gloriablanco@uniovi.es (G. Blanco).

susceptibility to pathogens have already been established [17,21–23]. On the other hand, studies in goldfish [24,25] and common carp [26], have suggested a new immune function for Tf in fish as a primary activator of macrophages. In this context, in response to cellular damage and/or infection, proteolytic cleavage products of Tf serve as “alarms” for the immune system, thereby enhancing the production of nitric oxide by macrophages [27].

Gilthead seabream (*Sparus aurata*) is a marine species that is extensively farmed in Mediterranean regions, and it is considered one of the most important species for the aquaculture industry. A reflection of the importance of this species is the start-up of a genomic sequencing program (www.bridgemap.tuc.gr), and the large amounts of genetic tools available, among which three genetic linkage maps (with more than 930 molecular markers [28–31]), two microarrays [32,33] and a database with more than 59000 ESTs available in GenBank are included. Due to the significant number of genetic and molecular tools, it has been possible to begin to establish associations between phenotypes for traits of interest and molecular markers. This process will ultimately allow for genic/genomic selection (GAS) in this teleost fish. But despite its importance to aquaculture and the large amount of genetic information available on gilthead seabream, the fact of its genomic sequencing program has not been completed yet, make that genes as Tf gene with relevant metabolic and immunologic implications, remain uncharacterized.

The present study is included in a large-scale project, whose global goal is the search for SNPs makers in candidate genes for traits of interest in gilthead seabream such as resistance to diseases. Among the previously candidate genes proposed, Tf gene was selected because of its strong relationship with the innate immune system. Due to lack of genetic information about this gene in gilthead seabream (*SaTf* gene), as a first step, in this study, we describe *SaTf* gene according to its genomic structure, constitutive expression pattern (analyzed in two different developmental stages), and the frequency of nucleotide substitution in its genomic sequence.

2. Material and methods

2.1. Fish rearing

All experimental fish (adults and pre-ongrowing) used in the present experiment were provided by the Institut de Recerca i Tecnologia Agroalimentaries (IRTA, San Carlos de la Ràpita (Spain)).

Adult fish were kept in 12000 L tanks under natural conditions of temperature and photoperiod, whereas pre-ongrowing fish were kept under controlled conditions using water recirculation systems (19 ± 1 °C temperature, 35 ± 1 ppt salinity and photoperiod 16hL:8hD). Both fish were fed commercial pre-ongrowing and adult diets (Opti-bream and Vitalis repro, Skretting, Spain).

2.2. Sequences, alignments and primer design

Only two expressed sequence tags (ESTs) encoding for *SaTf*, both 291 bp, were available from the GenBank database (GenBank ID: CX734969 and CX734865) at the beginning of this study. A BLAST search was used to identify Tf nucleotide sequences that had the highest similarity with available gilthead seabream ESTs (saESTs). BLAST results showed that the Tf mRNA sequences (complete coding sequence (CDS)) of black porgy (*Acanthopagrus Shlegelii*) (GenBank ID: AY36505) and red seabream (*Pagrus major*) (GenBank ID: AY335444), presented high nucleotide similarity with SaESTs of (94% and 92%, respectively). Both of these species are sparids, similar to gilthead seabream. Black porgy and red seabream Tf mRNA sequences were determined to consist of exons and untranslated regions (UTRs) through multiple alignments with the genomic DNA sequence of Nile tilapia Tf gene [34] (GenBank ID: DQ272465) using Sequence Alignment Editor BioEdit (V. 7.0.5.3) [35]. The genomic DNA sequence of Nile tilapia Tf was selected because it is one of the most closely related fish species to gilthead seabream, in which the Tf gene has been characterized to date. From alignment of these three sequences, regions of high conservation value were defined and used for subsequent primer design. Using FastPCR software [36] and the black porgy Tf mRNA sequence as a template sequence, a set of 20 primers were designed and organized into 12 pairs (Table 1) to use it in PCR *SaTf* gene amplification.

2.3. DNA extraction and PCR amplification of *SaTf* genomic DNA

Genomic DNA was purified from eye tissue of four adult gilthead seabream using ZymoBead™ Genomic DNA kit (ZYMO RESEARCH) according to the manufacturer's protocol. DNA samples were stored at -20 °C until used for PCR.

PCR reactions were performed under the following conditions: a final volume of 50 μ l PCR mixture contained 0.25 μ l (5 U/ μ l) of TaKaRa Taq™MS (TAKARA BIOTECHNOLOGY, (DALIAN) CO., LTD),

Table 1

Primer pairs used in the present study, including primer name, sequence, annealing temperature (Ta) and amplicon size (bp).

PCR primer pairs				Ta (°C)	PCR length
Forward primer		Reverse primer			
Name	Sequence (5'→3')	Name	Sequence (5'→3')		
<i>Sequencing of Tfβa gene</i>					
<i>SaTf</i> -F1	TGCTGACCTGGAGACCTCGA	<i>SaTf</i> -R3	AGGTCGTAGTTGTCAGTCCAG	62	1357
<i>SaTf</i> -F2	AGAAGCCTACTGTTGGCCT	<i>SaTf</i> -R3	AGGTCGTAGTTGTCAGTCCAG	62	859
<i>SaTf</i> -F3	CTGGTGAAGCAGACGACCATCACTC	<i>SaTf</i> -R5.2	TGGAAGGCTCCACCGTAGTCA	62	1184
<i>SaTf</i> -F4	CAGGTGCTGAAGACAGCCCGGT	<i>SaTf</i> -R5.2	TGGAAGGCTCCACCGTAGTCA	62	823
<i>SaTf</i> -F5	CTGTGAGCAACTTCTCAGTGCC	<i>SaTf</i> -R8	GAACATCAGGTTTTGGAGGT	64	1698
<i>SaTf</i> -F6	TGACTCCTCAGCCAAGCAAGCT	<i>SaTf</i> -R8	GAACATCAGGTTTTGGAGGT	64	1147
<i>SaTf</i> -F7	ACCTGGCCAAAGTACCAGCTCATG	<i>SaTf</i> -R9	AAGGCACCCGTCACCTGGCAGC	64	620
<i>SaTf</i> -F9	TGGAGCATCAACAGCATGACTG	<i>SaTf</i> -R13	GCTCAATACTACTGGCTTTGCA	62	1626
<i>SaTf</i> -F11	GCAATGCGGCACCAATGGA	<i>SaTf</i> -R13	GCTCAATACTACTGGCTTTGCA	64	1246
<i>SaTf</i> -F12	AGTTCTTCAGTGAGGCTGTGC	<i>SaTf</i> -R14	ACCCACTAGCCCATGCTGGACCA	64	1008
<i>SaTf</i> -F14	ACCAGGAACATTACCCACATG	<i>SaTf</i> -R15	CCACAGTTGTACAACCTGAG	64	883
<i>SaTf</i> -F15	GATCCCTCGTTCAAACGTTCCTCA	<i>SaTf</i> -R17	AAGTCAGGCATGTGTTTGAATAC	64	639
<i>RT-PCR</i>					
<i>SaTf</i> -FW	CAGTAACGGCAATGGTCCAGC	<i>SaTf</i> -RW	CCAGCTGAACATCCTACAGAC	60	314
β -actin FW	CTTCCTCGGTATGGAGTCTGCGG	β -actin RW	TCTTGCTGCTGATCCACATCTGCT	60	367

5 μ l (10 \times) PCR buffer, 4 μ l dNTP mixture (2.5 mM each), 0.4 μ l of each primer (10 pmol/ μ l) and approximately 100 ng of genomic DNA. The following PCR amplification program was used in an XP-Cycler (BIOER) thermocycler: 35 cycles at 94 °C for 30 s, annealing temperature (optimized for each primer pair, Table 1) for 30 s and 72 °C for 1 min. PCR products were run on a 2% agarose gel, purified with the Illustra™ GFX™ PCR DNA and Gel Band Purification Kits (GE HEALTHCARE), and analyzed on a new 2% agarose gel. Sequencing of all PCR products was performed in both directions by MACROGEN Inc. The complete *SaTf* genomic sequence was constructed by assembling the partially overlapping TF sequences obtained from each PCR reaction.

2.4. RNA extraction and *SaTf* cDNA synthesis

Fish were euthanized with 100 mg l⁻¹ MS222 (Sandoz) and dissected. Tissue samples were collected and preserved in RNAlater® (Ambion) at -20 °C until use. Total RNA from seven tissues (liver, kidney, spleen, brain, stomach, heart and ovary (female adults only)) of 3 pre-ongrowing (30 \pm 0.5 g body weight) and 3 adult individuals (3.5 \pm 0.25 kg body weight), were isolated using the High Pure RNA Tissue Kit (ROCHE) following the manufacturer's protocol. The concentration and purity of total RNA samples was quantified using a Picodrop by measuring absorbance at 260 nm and the ratio of 260 nm/280 nm, respectively. For each sample, 1 μ g of total RNA was reverse transcribed using Anchored-oligo(dT)₁₈ Primers. cDNA synthesis was performed using a Transcriptor First Strand cDNA Synthesis Kit (ROCHE) following the manufacturer's protocol, and cDNAs were stored at -20 °C for later use.

Preparations of cDNA from the liver were used to sequence the coding region of *SaTf* gene. Full-length cDNA sequences were obtained in one amplicon, using *SaTf*-F1/*SaTf*-R17 primers (Table 1). The resulting cDNA PCR product was sequenced using the following pairs of primers: (*SaTf*-F1/*SaTf*-R5.2, *SaTf*-F4/*SaTf*-R8, *saT*-F5/*SaTf*-R13 and *saT*-F11/*SaTf*-R17) (Table 1), creating partially overlapping cDNA sequences with sizes of approximately 1000 bp. PCR conditions were identical to those used for the characterization of *SaTf* genomic DNA sequences. Once the complete cDNA sequence was obtained, the *SaTf* protein structure was determined using the SMART program (http://www.smart.embl_heidelberg.de/).

2.5. RT-PCR analysis

RT-PCR analyses were carried out on total RNA extracted from different tissues of adult and pre-ongrowing gilthead seabream. Specific primers (*SaTf*-FW/*SaTf*-RW) (Table 1) were designed to amplify a 303 bp cDNA fragment of *SaTf*. To reduce the possibility of genomic DNA amplification, the *SaTf*-FW primer sequence was designed to span the border between exons XIV and XV. β -actin was used as housekeeping gene to test the quality of the extracted RNA and the efficiency of the RT-PCR reaction [37]. PCR products were resolved on 1.5% agarose gels with ethidium bromide.

2.6. SNP analysis

To assay SNPs in *SaTf*, genomic DNA from eye tissue of twenty seabream, previously identified as unrelated through microsatellite analyses [38], was isolated using ZymoBead™ Genomic DNA kit (ZYMO RESEARCH) according to the manufacturer's protocol. The complete *SaTf* genomic sequence characterized in this study was amplified and sequenced (as described above) for each of the selected individuals. Sequence alignment was performed with Sequence Alignment Editor BioEdit (V.7.0.5.3) [35], allowing for the

identification of putative SNPs. SNP frequency, distribution and possible involvement in amino acid substitution were analyzed.

2.7. Phylogenetic analyses

For phylogenetic analyses, gilthead seabream cDNA was compared to TF cDNA sequences from fourteen fish, one amphibian and three mammalian species. All sequences were retrieved from NCBI GenBank. Inference of coding regions and deduction of amino acid sequences were performed using Sequence Alignment Editor BioEdit (V.7.0.5.3) [35]. Amino acid sequences were initially aligned using the Muscle web program (v.1.5) [39] using the multiple alignment default settings. The alignment was then edited using a heuristic search for best representative positions (*automated1* option), implemented by TrimAL software [40]. The subsequent phylogenetic analysis was performed on this alignment.

The most appropriate model of protein evolution based on amino acid alignment was determined through maximum-likelihood methodologies using Prottest software [41]. Also determined were the average values of different evolutionary parameters, including invariable sites (+I), rate heterogeneity among sites (+G) and amino acid frequencies (+F). The resulting phylogenetic tree was constructed using an evolutionary model based on a WAG matrix (+I = 0.08; +G = 1.43). Branch length optimization was achieved using maximum-likelihood methodologies in the PhyML package [42]. Robustness of the topology was evaluated by bootstrap tests with 1000 replicates. The resulting phylogenetic tree was visualized and edited using Figtree software v.1.3.1 (<http://tree.bio.ed.ac.uk/software/figtree/>).

3. Results

3.1. Characterization of the *SaTf* gene

The specific PCR products from each one of the twelve unique primer pairs (Table 1), ranged in size from 620 bp (*SaTf*-F7/*SaTf*-R9) to 1698 bp (*SaTf*-F5/*SaTf*-R8) (Table 1). The alignment and assembly of these genomic fragments showed a *SaTf* genomic sequence of 7855 nucleotide in length [GenBank ID: JF309046]. By means of RT-PCR, by combining 8 out of primers (previously used in the sequencing of the genomic sequence) in four new primer pairs, a cDNA sequence of 2320 nucleotide in length was obtained [GenBank ID: JF309047]. From the alignment of the genomic and cDNA sequence the structure of *SaTf* gene was determined. The *SaTf* genomic DNA sequence described in this study consists of 17 exons and 16 introns, and the sequence is flanked by 5' (18 bp) and 3' (274 bp) partial UTR domains (Fig. 1, Table 2). All intron-exon boundary sequences conform to the GT-AG rule.

Comparison of the TF gene structures of gilthead seabream, catfish (*Ictalurus punctatus*), Nile tilapia, coho salmon (*Oncorhynchus kisutch*) and Japanese medaka (*Oryzias latipes*) showed high conservation with respect to exon size (Table 2). These species presented complete size conservation in six exons (EI, EII, EVI, EX, EXI and EXIV). A strong similarity in exon size was observed between gilthead seabream and Nile tilapia, which presented identical sizes for fifteen of the seventeen exons. However, this conservation did not extend to intron sizes, which showed a wide range of variation in all species compared (Table 2).

The *SaTf* protein deduced from the cDNA sequence consists of 691 amino acids (Fig. 2) and is composed of a signal peptide (4–14 aa) and two lobes, the N-lobe (24–338 aa) and C-lobe (340–681 aa). The Fe-binding (Asp-73, Tyr-103, Tyr 200, His-256, Asp-394, Tyr-429, Tyr-524 and Arg-592) and anion-binding (Thr-128, Lys-132, Thr-454 and Arg-458) residues described by Lambert et al. [43]. are conserved in gilthead seabream (Fig. 2).

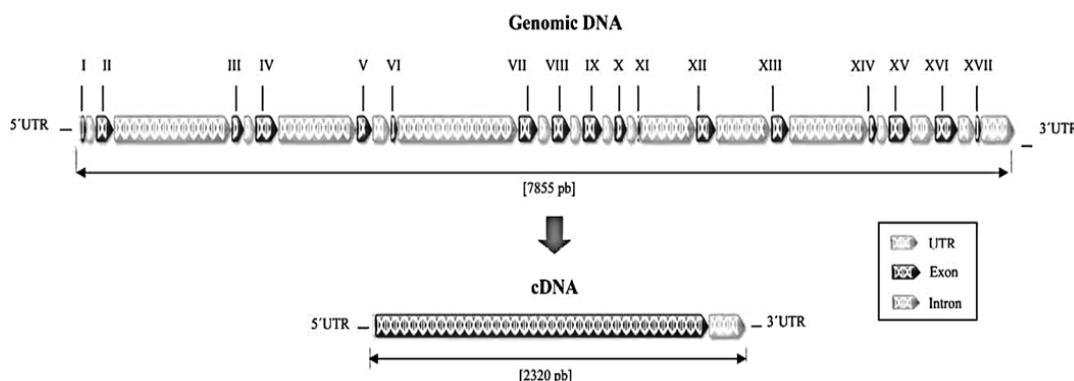


Fig. 1. Schematic graphic of *SaTf* genomic DNA and cDNA structures obtained in this study. Graphic was made by means of Gene Designer by DNA 2.0 Inc (<http://www.DNA20.com>) [57]. Each region is showed in proportional size.

3.2. Expression of the *SaTf* gene

Tf expression patterns in different tissues were analyzed by RT-PCR in pre-ongrowing and adult fish. These analyses showed high Tf expression in the liver in both developmental stages, whereas expression was lower in the kidney, spleen, stomach, brain and ovary (analyzed in adult females only) (Fig. 3). Heart tissue was the only tissue where transcripts of *SaTf* mRNA were not observed in either of the two developmental stages analyzed (Fig. 3). Differences between pre-ongrowing and adult *SaTf* expression patterns were detected in the spleen and stomach (Fig. 3). In the spleen, *SaTf* was expressed more abundantly in the adults than in pre-ongrowing individuals, whereas expression in the stomach Tf was higher in pre-ongrowings.

3.3. SNP variation

The sequencing of *SaTf* genomic DNA in twenty fish resulted in the detection of 31 polymorphic sites distributed along the 7855 bp constituting the gene (Table 3). Therefore, the estimation of SNP frequency in the *SaTf* genomic locus was 3.9 SNP per 1000 bp, which reduces to an approximate SNP frequency of 1 SNP per

253 bp. However, most of nucleotide variations were located in non-coding regions (27 SNPs within intronic regions and two SNPs in the 3'UTR domain), whereas only 2 SNP were identified in coding regions (Table 2). Therefore, the SNP density detected within intronic regions (1 SNP/203 bp) was higher than in exonic regions (1 SNP per 1038 bp). The two detected coding single nucleotide polymorphisms (cSNPs) resulted in synonymous substitutions and no changes in the amino acid sequence were detected (Table 3). The C/G nucleotide substitution never was detected. In addition, all SNPs were bi-allelic, each SNP site varied between only two nucleotides. The least frequently observed allele was always observed with a frequency greater than 1%, which is the fixed-threshold for consideration as a SNP.

3.4. Phylogenetic analyses

Phylogenetic analyses were performed to establish an evolutionary context for the *SaTf* gene. Genetic distances (measured as substitutions per site) showed moderate-low values, and the topology was well-supported by strong bootstrap values (Fig. 4). As expected, gilthead seabream was included within a sub-cluster of the sparids (black porgy, red seabream) with high bootstrap values.

Table 2

Comparison of transferrin gene structures obtained from gilthead seabream (*Sparus aurata*, GeneBank ID JF309046 and JF309047), catfish (*Ictalurus punctatus*, GeneBank ID FJ176741 and FJ176740), Nile tilapia (*Oreochromis niloticus*, GeneBank ID DQ272465), Coho salmon (*Oncorhynchus kisutch*, GeneBank ID AH008270.1 and D89084) and Japanese medaka (*Oryzias latipes*, GeneBank ID D64033). Exon and intron lengths are given in base pairs. Exons with fully conserved size are shaded in grey. Lengths of inferred amino acid sequences are indicated in the last row.

	Gilthead seabream	Catfish	Nile tilapia	Coho salmon	Japanese medaka		Gilthead seabream	Catfish	Nile tilapia	Coho salmon	Japanese medaka
E-I	40	40	40	40	40	IN-I	83	905	33	111	66
E-II	149	137	152	152	149	IN-II	989	1203	715	–	364
E-III	103	103	103	103	103	IN-III	96	508	128	265	85
E-IV	192	189	192	192	192	IN-IV	663	348	784	–	287
E-V	130	127	130	130	130	IN-V	157	100	96	189	103
E-VI	56	56	56	56	56	IN-VI	1019	80	363	169	693
E-VII	164	161	173	164	164	IN-VII	118	590	106	701	150
E-VIII	154	151	154	151	154	IN-VIII	107	983	85	108	86
E-IX	164	167	164	164	164	IN-IX	103	96	91	307	84
E-X	97	97	97	97	97	IN-X	92	1550	98	109	97
E-XI	21	21	21	21	21	IN-XI	471	343	604	146	550
E-XII	165	165	165	162	169	IN-XII	467	136	218	–	404
E-XIII	151	142	151	154	147	IN-XIII	666	1152	332	174	838
E-XIV	68	68	68	68	68	IN-XIV	99	223	85	188	102
E-XV	182	179	182	182	181	IN-XV	208	155	878	–	189
E-XVI	193	193	193	193	194	IN-XVI	149	524	123	102	658
E-XVII	47	44	44	35	44	–	–	–	–	–	–
Total	2076	2040	2085	2064	2073	Total	5487	8896	4739	2568	4756
Protein	691	679	694	687	690						

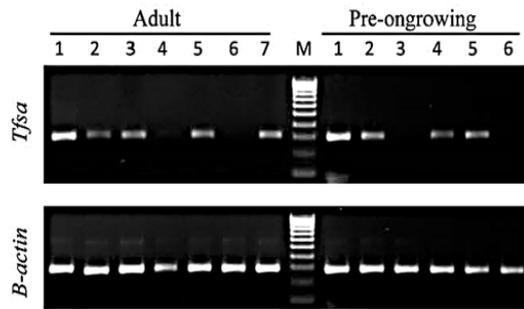


Fig. 3. RT-PCR assays performed on different tissues and developmental stages. Lanes correspond to liver (1), kidney (2), spleen (3), stomach (4), brain (5), heart (6), ovary (7), molecular weight ladder (M) (pattern of 10 regularly spaced bands ranging from 100 to 1000 bp).

Nile tilapia and medaka formed another sub-cluster, whereas salmonids and cyprinids were clustered in their own respective groups. Atlantic cod, channel catfish and Japanese flounder were not included in any group. The terrestrial vertebrates formed a major cluster composed of amphibian western clawed frog and mammalian sub-cluster.

4. Discussion

In the present study, genomic and cDNA sequences of *SaTf* were identified and characterized. The predicted gene structure of *SaTf* consists of a coding region of 2076 nucleotides divided into 17 exons and a non-coding region that includes 16 introns and spans

Table 3

Putative SNPs identified in the *SaTf* gene, including position, SNP ID (gene name and nucleotide position), type of nucleotide change [Homozygous 1 (Hom 1), Homozygous 2 (Hom 2) and Heterozygous (Het)], allelic frequency (Freq) and amino-acid change (amino-acid position indicated in brackets).

	Position	SNP ID	Hom 1	Freq	Hom 2	Freq	Het	Freq	Amino-acid change	
In-II		<i>SaTf</i> 0434	C	0.842	–	–	C/T	0.158		
		<i>SaTf</i> 0739	G	0.714	A	0.143	G/A	0.143		
		<i>SaTf</i> 0992	C	0.286	T	0.428	C/T	0.286		
In-IV		<i>SaTf</i> 0967	G	0.714	T	0.143	T/G	0.143		
		<i>SaTf</i> 2056	G	0.850	–	–	A/G	0.150		
		<i>SaTf</i> 2071	T	0.550	–	–	T/G	0.450		
		<i>SaTf</i> 2125	A	0.200	G	0.300	A/G	0.500		
Ex-V		<i>SaTf</i> 2153	C	0.800	–	–	A/C	0.200		
		<i>SaTf</i> 2436	A	0.367	G	0.158	A/G	0.475	Pro/Pro (196)	
In-VI		<i>SaTf</i> 3169	G	0.769	–	–	T/G	0.231		
		<i>SaTf</i> 3223	A	0.615	G	0.154	A/G	0.231		
		<i>SaTf</i> 3321	C	0.600	T	0.100	C/T	0.300		
In-VII		<i>SaTf</i> 3539	A	0.850	–	–	A/G	0.150		
		<i>SaTf</i> 3943	T	0.900	–	–	A/T	0.100		
In-XI		<i>SaTf</i> 4755	C	0.800	–	–	C/T	0.200		
		<i>SaTf</i> 4984	G	0.950	–	–	T/G	0.050		
In-XII		<i>SaTf</i> 5382	A	0.900	–	–	A/T	0.100		
		<i>SaTf</i> 5681	C	0.789	T	0.053	C/T	0.158		
		<i>SaTf</i> 5783	A	0.737	–	–	A/G	0.263		
In-XIII		<i>SaTf</i> 6039	G	0.895	–	–	T/G	0.105		
		<i>SaTf</i> 6100	A	0.579	C	0.368	A/C	0.053		
		<i>SaTf</i> 6234	A	0.818	–	–	A/T	0.182		
		<i>SaTf</i> 6238	T	0.727	–	–	A/T	0.273		
		<i>SaTf</i> 6239	T	0.692	–	–	T/G	0.308		
Ex-XIV		<i>SaTf</i> 6426	A	0.722	T	0.111	A/T	0.167		
		<i>SaTf</i> 6702	C	0.700	T	0.500	C/T	0.250	Asp/Asp (551)	
In-XIV		<i>SaTf</i> 6732	T	0.632	C	0.530	C/T	31.5		
		<i>SaTf</i> 6782	C	0.600	T	0.350	C/T	0.500		
In-XVI		<i>SaTf</i> 7424	A	0.316	G	0.210	A/G	0.474		
	3'UTR		<i>SaTf</i> 7586	G	0.733	–	–	A/G	0.267	
			<i>SaTf</i> 7767	C	0.889	–	–	A/C	0.111	

5495 nucleotides (Table 2). Comparison the *SaTf* gene structure with previously reported gene structures of other fish species shows a high conservation of exon size (Table 2), being particularly high between gilthead seabream and Nile tilapia. Conservation of coding sequence length between the compared species was also observed in the small variation in the amino acid sequence length of the deduced protein, with sizes ranging from 679 aa in catfish to 694 aa in Nile tilapia. The deduced Tf protein for gilthead seabream is composed of 691 aa and consists of an initial peptide and two lobes (N- and C-lobes). There was complete conservation of iron and anion-binding residues, which are established features of Tf in fish (Fig. 2).

In the phylogenetic reconstruction based on the deduced Tf amino acid sequences of different vertebrate species, gilthead seabream Tf clusters with other sparid species, showing low distances and high bootstrap values. European sea bass (*Dicentrarchus labrax*), which does not belong to Sparidae family, but comes from the Moronidae taxa [44], also clustered with the sparids with high confidence (Fig. 4). Therefore, Tf has apparently undergone few modifications within the Sparidae family throughout evolution. Phylogenetic analyses confirm that Tf is a medium-high conserved protein in vertebrate taxa, probably as a result of its important role in iron metabolism and the innate immune response, two processes that impose evolutionary constraints on the amino acid sequence.

The overall SNP density identified in the *SaTf* gene from genomic DNA was found to be 1 SNP per 253 bp, similar to the described frequency for the Atlantic salmon (*Salmo salar*) Tf (1 SNP/217 bp) [45]. SNP frequency in non-coding sequences was fifteen-fold higher (93.5%) than within coding sequences (6.5%). Alternatively, determination of the position of the two cSNPs revealed that both affected the third position of the codon, resulting in synonymous changes. The fact that the nucleotide variations detected in the coding regions result in silent substitutions and consequently do not result in amino acid changes in the protein sequence could be a reflection of the important functions carried out by Tf, including iron metabolism, antimicrobial activity, growth and cellular differentiation. Drastic changes in the amino acid sequence of this multi-function protein could give rise to a non-functional protein, causing a negative effect on the development of organism. The low nucleotide variation detected in the coding regions and the absence of amino acid replacement in the protein, might be consequence of limited number of individuals analyzed who although were previously identified as unrelated individuals, all of them came from offspring of a unique brookstock. A greater number of cSNPs should be found from sequence alignment of individuals from populations subjected to different environmental conditions, showing a local adaptation to different environments. In fact, recently, by comparison of cDNA sequences from Northeast and Northwest Atlantic cod populations, Andersen et al. [46] describe 22 SNPs in Tf1, causing the replacement of 16 amino acids, of which 8 affect to surface residues. The authors suggest that this multiple surface residue changes, as a result of an evolutionary competition for transferrin-bound iron between the host and invading pathogens.

Although most SNPs identified in *SaTf* gene were found in non-coding regions (often incorrectly referred to as junk DNA), beside its potential usefulness as linkage or parentage assessment markers, it should be considered that these nucleotide substitutions could induce changes in the regulation of gene expression. In vertebrates, it has been proposed that numerous introns encode miRNAs [47,48], sequences of approximately 18–25 nucleotides in length associated with chromatin silencing mechanisms. The ability of miRNAs to inhibit protein translation or to directly degrade their target mRNA allow miRNAs to take part in the

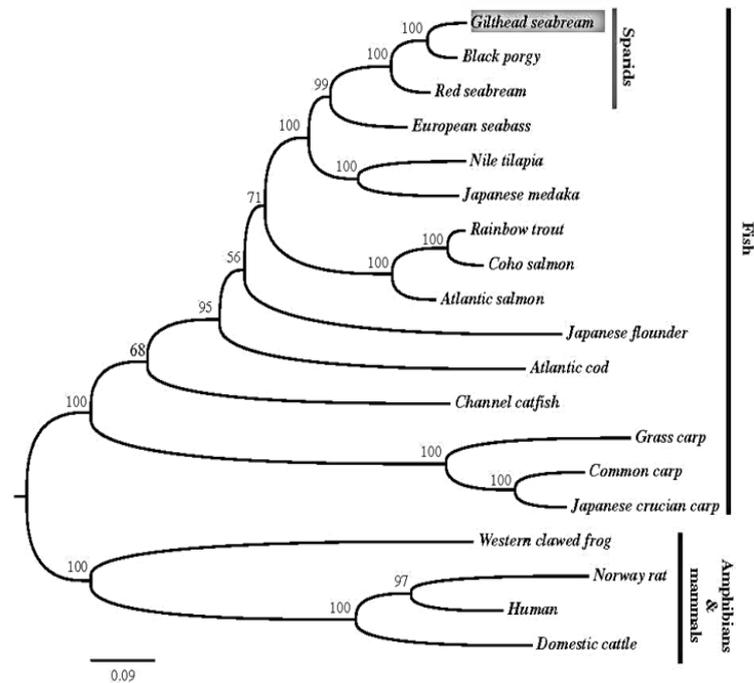


Fig. 4. Phylogenetic tree of the deduced amino acid sequences of transferrin proteins from 15 taxa. Tree topology was reconstructed using Maximum-Likelihood methodologies. Evolutionary distances were computed using the WAG matrix (+I = 0.08; +G = 1.43) and are in units of the number of amino acid substitutions per site. Numbers adjacent to the branches are percentage values derived from 1000 bootstrap replicates.

regulation of gene expression, for example, as a defense mechanism against viral infections [49,50]. On the other hand, SNPs detected in the introns of the *SaTf* gene do not appear to have a homogenous distribution, as evidenced by the fact that their distribution does not correlate with the number of base pairs analyzed. Long introns (II and VI introns) showed a SNP frequency of approximately 1 SNPs per 250 bp, whereas higher SNP frequencies were found in smaller introns [e.g., intron XIV (1 SNP per 50 bp)] (Table 3). This fact suggests that these regions are not subjected to the same degree of conservation and, therefore, that each region has a different likelihood of sequence modifications.

The importance of SNP polymorphism in Tf is shown up through its association with biological functions. In fish, this link has been established in Nile tilapia where two haplotype blocks comprised 34 SNPs (including SNPs located in coding and no-coding regions) were associated with a different saltwater tolerance [34]. Besides, the usefulness of SNPs as genetic tool to study the interactions between genotype and environment, make them the main genetic maker to immune response gene in fish. An example of this is found in common carp where a study to establish the genetic bases of resistance to cyprinid herpesvirus-3 (CyHV-3) and to develop SNP markers linkage of this disease, has already been started-up [51].

As expected, expression analyses showed that Tf expression is highest in the liver of both pre-ongrowing and adult individuals. Comparison between pre-ongrowing and adult Tf expression patterns revealed stage-dependent expression in some tissues. Specifically, Tf expression in the stomach decreased during development. In contrast, an increase in Tf expression was observed in the spleen between pre-ongrowing and adult stages. These differences in splenic Tf expression could be a consequence of a lower maturity of the lymphoid organs during this developmental stage. Moreover, these organs may not be fully functional in pre-ongrowing fish, which may be one of the reasons why larval and

juvenile fish are highly sensitive to infections (e.g., fish pasteurellosis) whereas individuals with greater weight, larger than 50 g, show higher resistance to infections [52].

Gilthead seabream, Atlantic cod (*Gadus morhua*), migral carp (*Cirrhinus mrigala*), sea bass and catfish show significant Tf expression in the brain [53–56]. The difficulty of Tf in crossing of the blood-brain barrier to supply iron to the brain could explain the need to synthesize Tf in this tissue. However, the fact that Tf expression in the brain is not detected in other fish species, including Atlantic salmon, calls this hypothesis into question. Neves et al. [55] reported that Tf in the sea bass brain acts as a positive acute phase protein, increasing its expression in response to bacterial infection. These authors suggest that locally produced brain Tf may be related to host defense mechanisms. Thus, the characterization of the Tf promoter region may provide a better understanding regarding the regulation of Tf expression and may help to elucidate the observed differential expression of Tf in the brains of different fish species.

Experimental bacterial infections carried out in sea bass and catfish indicate a differential regulation of Tf expression in response to infection. Sea bass seem to decrease Tf expression in the liver in response to bacterial infection [55], whereas catfish appear to up-regulate its expression in the liver, suggesting its role as a positive, acute phase protein [56]. This contradictory finding could be a consequence of the evolutionary pressure exerted by the primary strategy of the major pathogens of these fish species, which is iron uptake from the host. A down-regulation of Tf expression should be a limiting factor for pathogens whose major strategy for obtaining iron from host is through siderophores. However, up-regulating Tf expression in response to pathogenic microorganism invasion would enable the host to directly capture free iron, which would presumably be a great advantage for reducing the available iron to these microorganism, thereby restricting their growth. According

to this hypothesis, being gilthead seabream a species sensitive to the same major pathogens that affect sea bass farming (e.g., *Photobacterium damsela*), it would therefore be expected that gilthead seabream down-regulate Tf expression in response to an immune challenge. Future studies in gilthead seabream aimed at determining the mechanisms that regulate *SaTf* expression will provide insights into how changes in Tf expression can act as an arm of the host innate immune defense.

In summary, the *SaTf* sequences reported in the present study suggest conservation of the gene and protein structure with those previously reported for other fish species. Moreover, *SaTf* expression analyses show tissue-specific and ontogeny-dependent expression. Finally, the polymorphism analyses described here contribute to the collection of available SNPs in gilthead seabream and provide the preliminary information for subsequent studies that aim to identify possible SNP haplotypes linked to disease resistance.

Acknowledgments

This work was financed by the Spanish Ministry of Science and Innovation (MICINN; National Program of Resources and Food and Agriculture Technologies, AGL2006-13411-C03-00, SELECTSPARUS, and AGL2007-64040-C03-00, SELECTBREAM, including European Regional Development Funds). C. García-Fernández was supported by a FPU scholarship financed by MICINN. The authors wish to thank Alicia Estévez (Institut de Recerca i Tecnològica Agroalimentaria, IRTA) for providing the experimental fish used in this study.

References

- Ellis AE. Innate host defense mechanisms of fish against viruses and bacteria. *Dev Comp Immunol* 2001;25:827–39.
- Le Morvan C, Troutaud D, Deschaux P. Differential effects of temperature on specific and nonspecific immune defences in fish. *J Exp Biol* 1998;201:165–8.
- Hayman JR, Bly JE, Levine RP, Lobb CJ. Complement deficiencies in channel catfish (*Ictalurus punctatus*) associated with temperature and seasonal mortality. *Fish Shellfish Immunol* 1992;2:183–92.
- Bly JE, Clem LW. Temperature and teleost immune functions. *Fish Shellfish Immunol* 1992;2:159–71.
- Aoki T, Takano T, Santos MD, Kondo H, Hirono I. Molecular innate immunity in teleost fish: review and future perspectives. *Fisheries for Global Welfare and Environment, 5th World Fisheries Congress*; 2008. pp. 263–76.
- Gomme PT, McCann KB. Transferrin: structure, function and potential therapeutic actions. *DDT* 2005;4:267–73.
- Aisen P. Iron transport and storage protein. *Annu Rev Biochem* 1980;49:357–93.
- Crichton RR, Wilmet S, Legssyer R, Ward RJ. Molecular and cellular mechanisms of iron homeostasis and toxicity in mammalian cells. *J Inorg Biochem* 2002;91:9–18.
- Van Campenhout A, van Campenhout CM, Lagrou AR, Manuel-y-Keenoy B. Transferrin modifications and lipid peroxidation: implications in diabetes mellitus. *Free Radic Res* 2003;37:1069–77.
- Özgür F, Ferit K. Individual and combined effects of heavy metals on serum biochemistry of Nile tilapia *Oreochromis niloticus*. *Arch Environ Contam Toxicol* 2010;58:151–7.
- Retzer MD, Kabani A, Button LL, Yu RH, Schryvers AB. Production and characterization of chimeric transferrin for the determination of the binding domains for bacterial transferrin receptors. *J Biol Chem* 1996;271(2):1166–73.
- Baldwin GS. Comparison of transferrin sequences from different species. *Comp Biochem Physiol B* 1993;106:203–18.
- Park I, Schaeffer E, Sidoli A, Baralle FE, Cohen GN, Zakin MM. Organization of the human transferrin gene: direct evidence that it originated by gene duplication. *Proc Natl Acad Sci U S A* 1985;82:3149–53.
- Lambert LA, Perri H, Meehan TJ. Evolution of duplications in the transferrin family of proteins. *Comp Biochem Physiol* 2005;140:11–25.
- Zakin MM. Regulation of transferrin gene expression. *FASEB J* 1992;6:3253–8.
- Tu GF, Achen MG, Aldred AR, Southwell BR, Schreiber G. The distribution of cerebral expression of the transferrin gene is species-specific. *J Biol Chem* 1991;266:6201–8.
- Suzumoto BK, Schreck CB, McIntyre JD. Relative resistances of three genotypes of Coho Salmon (*Oncorhynchus kisutch*) and their hematological responses to bacterial kidney disease. *J Fish Res Board Can* 1977;34:1–7.
- Chen D, McMichael JC, VanDerMeid KR, Masi AW, Bortell E, Caplan JD, et al. Evaluation of a 74-kDa transferrin-binding protein from *Moraxella (Branchiella) catarrhalis* as a vaccine candidate. *Vaccine* 1999;18:109–18.
- Ford MJ. Molecular evolution of transferrin: evidence for positive selection in salmonids. *Mol Bio Evol* 2001;18:639–47.
- Yang L, Gui JF. Positive selection on multiple antique allelic lineages of transferrin in the polyploid *Carassius auratus*. *Mol Biol Evol* 2004;21:1264–77.
- Winter GW, Scherck CB, McIntyre JD. Resistance of different stocks and transferrin genotypes of coho salmon, *Oncorhynchus kisutch*, and steelhead trout, *Salmo gairdneri*, to bacterial kidney disease and vibriosis. *Fish Bull* 1980;77:795–802.
- Jurecka P, Wiegertjes GF, Rakus KL, Pilarczyk A, Irnazarow I. Genetic resistance of carp (*Cyprinus carpio* L.) to *Trypanoplasma borreli*: influence of transferrin polymorphisms. *Vet Immunol Immunopathol* 2009;127(1–2):19–25.
- Jurecka P, Irnazarow I, Westphal AH, Forlenza M, Arts JA, Savelkoul HFJ, et al. Allelic discrimination, three-dimensional analysis and gene expression of multiple transferrin alleles of common carp (*Cyprinus carpio* L.). *Fish Shellfish Immunol* 2009;26(4):573–81.
- Stafford JL, Neuman NF, Belosevic M. Products of proteolytic cleavage of transferrin induce nitric oxide response of goldfish macrophages. *Dev Comp Immunol* 2001;25:101–15.
- Stafford JL, Belosevic M. Transferrin and the immune response of fish: identification of a novel mechanism of macrophage activation. *Dev Com Immunol* 2003;27:539–54.
- Jurecka P, Irnazarow I, Stafford JL, Ruzsyczky A, Taverne N, Belosevic M, et al. The induction of nitric oxide response of carp macrophages by transferrin is influenced by the allelic diversity of the molecule. *Fish Shellfish Immunol* 2009;26:632–8.
- Matzinger P. The danger model: a renewed sense of self. *Science* 2002;296:301–5.
- Franch R, Louro B, Tsalavouta M, Chatzizelis D, Tsigonopoulos CS, Sarropoulou E, et al. A genetic linkage map of the hermaphrodite teleost fish *Sparus aurata* L. *Genetics* 2006;174:851–61.
- Senger F, Priat C, Hitte C, Sarropoulou E, Franch R, Geisler R, et al. The first radiation hybrid map of a perchlike fish: the gilthead seabream (*Sparus aurata* L.). *Genomics* 2006;87:793–800.
- Sarropoulou E, Franch R, Louro B, Power DM, Bargelloni L, Magoulas A, et al. A gene-based radiation hybrid map of the gilthead sea bream *Sparus aurata* refines and exploits conserved synteny with *Tetraodon nigroviridis*. *BMC Genomics* 2007;8:44.
- Sarropoulou E, Nousdili D, Magoulas A, Kotoulas G. Linking the genomes of nonmodel teleosts through comparative genomics. *Marine Biotechnol* 2008;10:227–33.
- Sarropoulou E, Kotoulas G, Power DM, Geisler R. Gene expression profiling of gilthead sea bream during early development and detection of stress-related genes by the application of cDNA microarray technology. *Physiol Genomics* 2005;23:182–91.
- Ferraresso S, Vitulo N, Mininni AN, Romualdi C, Cardazzo B, Negrissolo E, et al. Development and validation of a gene expression oligo microarray for the gilthead sea bream (*Sparus aurata*). *BMC Genomics* 2008;9:580.
- Rengmark AH, Lingaas Frode. Genomic structure of the Nile tilapia (*Oreochromis niloticus*) transferrin gene and haplotype associated with saltwater tolerance. *Aquaculture* 2007;272:146–55.
- Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser* 1999;41:95–8.
- Kalendar R, Lee D, Schulman AH. FastPCR software for PCR primer and Probe Design and Repeat search. *Genes, Genomes and Genomics* 2009;3(1):1–14.
- Tiago DM, Laizé V, Cancela ML. Alternatively spliced transcripts of *Sparus aurata* insulin-like growth factor 1 are differentially expressed in adults tissues and during early development. *Gen Comp Endocrinol* 2008;157:107–15.
- Borrell YJ, Gallego V, García-Fernández C, Mazzeo I, Pérez L, Asturiano JF, et al. Assessment of parental contributions to fast- and slow-growing progenies in the sea bream *Sparus aurata* L. using a new multiplex PCR. *Aquaculture* 2011;314:58–65.
- Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucl Acids Res* 2004;32(5):1792–7.
- Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 2009;25(15):1972–3.
- Abascal F, Zardoya R, Posada D. ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* 2005;21(9):2104–5.
- Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 2003;52(5):696–704.
- Lambert LA, Perri H, Halbrooks PJ, Mason AB. Evolution of the transferrin family: conservation of residues associated with iron and anion binding. *Comp Biochem Physiol Part B: Biochem Mol Biol* 2005;142(2):129–41.
- Orrell TM, Carpenter KE. A phylogeny of fishes of the family Sparidae (Perciformes: Percoidae) inferred from mitochondrial sequence data. *Mol Phyl Evol* 2004;32:425–34.
- Ryynänen HJ, Primmer CR. Single nucleotide polymorphism (SNP) discovery in duplicated genomes: intron-primed exon-crossing (IPEC) as a strategy for

- avoiding amplification of duplicated loci in Atlantic salmon (*Salmo salar*) and other salmonid fishes. *BMC Genomics* 2006;7:192.
- [46] Andersen O, De Rosa MC, Pirolli D, Tooming-Klunderud A, Petersen PE, Andre C. Polymorphism, selection and tandem duplication of transferrin genes in Atlantic cod (*Gadus morhua*) – Conserved synteny between fish monolobal and tetrapod bilobal transferrin loci. *BMC Genet*; 2011. doi:10.1186/1471-2156-12-51.
- [47] Lin SL, Miller JD, Ying SY. Intronic MicroRNA (miRNA). Review. *J Biomed Biotech* 2006;26818:1–13.
- [48] Carthew RW. Gene silencing by double-stranded RNA. *Curr Opin Cell Biol* 2001;13(2):241–7.
- [49] Ying SY, Chang CP, Lin SL. Intron-mediated RNA interference, intronic microRNAs, and applications. *Methods Mol Biol* 2010;629:205–37.
- [50] Brennecke J, Hipfner DR, Strak A, Russell RB, Cohen SM. Bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene hid in *Drosophila*. *Cell* 2003;113(1):25–36.
- [51] Kongchum P, Palti Y, Hallerman EM, Hulata G, David L. SNP discovery and development of genetic markers for mapping innate response genes in common carp (*Cyprinus carpio*). *Fish Shellfish Immunol* 2010;29:356–61.
- [52] Noya M, Magariños B, Lamas J. Interactions between peritoneal exudates cells (PECs) of gilthead seabream (*Sparus aurata*) and *Pasteurella piscida*. A morphological study. *Aquaculture* 1995;131:11–21.
- [53] Devonan-Wright EM, Ramsey NB, McCormick CJ, Lazier CB, Wright JM. Nucleotide sequence of transferrin cDNA and tissue-specific expression of the transferrin gene in Atlantic cod (*Gadus morhua*). *Comp Biochem Physiol* 1996;113B(2):269–73.
- [54] Sahoo PK, Mohanty BR, Kumari J, Barat A, Sarangi N. Cloning, nucleotide sequence and phylogenetic analyses, and tissue-specific expression of the transferrin gene in *Cirrhinus mrigala* infected with *Aeromonas hydrophila*. *Comp Immunol Microbiol Infect Dis* 2009;32(6):527–37.
- [55] Neves JV, Wilson JM, Rodrigues PNS. Transferrin and ferritin response to bacterial infection: the role of the liver and brain in fish. *Dev Comp Immunol* 2009;33:848–57.
- [56] Liu H, Takano T, Abernathy J, Wang S, Sha Z, Jiang Y, et al. Structure and expression of transferrin gene of channel catfish, *Ictalurus punctatus*. *Fish Shellfish Immunol* 2009;28(1):1–8.
- [57] Alan Villalobos A, Ness JE, Gustafsson C, Minshull J, Govindarajan S. Gene designer: a synthetic biology tool for constructing artificial DNA segments. *BMC Bioinformatics* 2006;7:285.

Expression of Developmental-Stage-Specific Genes in the Gilthead Sea Bream *Sparus aurata* L.

Carmen García Fernández · Chrysoula Roufidou ·
Efthimia Antonopoulou · Elena Sarropoulou

Received: 4 July 2012 / Accepted: 11 September 2012
© Springer Science+Business Media New York 2012

Abstract The mechanism of early fish development as well as the control of egg quality is of great importance for the ability of the oocyte to develop after fertilization. Embryonic development is initially regulated by maternally provided mRNAs and later by the zygotic genome. Maternal mRNAs have an important role in initiating processes crucial to patterning the developing fish embryo. Furthermore, it has been shown that maternal RNA plays an important role in egg quality. The identification and characterization of candidate maternal genes in non-model fish species with important aquaculture interest like the gilthead sea bream *Sparus aurata* L. is of importance for future studies related to egg quality. The broodstock of the gilthead sea bream produces large quantities of eggs with a high and non-controllable quality variation. In the present study, we have studied the gene expression of 16 genes (*gapdh 1* and *2*, *cathepsin D*, *L*, *S* and *Z*, *erk1*, *jnk1*, *p38 alpha* and *p38 delta*, *ppar alpha*, *beta*

and *gamma*, *tubulin beta*, *ferritin M*, *cyclinA2*) of different functional categories in seven developmental stages. The 16 genes were chosen based on their putative involvement in egg quality and regulation of early development. In total, 11 showed a characteristic gene expression pattern pinpointing to the possible function as maternal genes and thus may function as molecular biomarker for egg quality.

Keywords Isoforms · Development · Gene expression · Maternal genes · *Sparus aurata* L. · Perciformes

Introduction

The development of fish larvae is determined as in other animals during embryogenesis. At a molecular level, several studies have been performed using state-of-the-art technologies in order to assess the differential expression of the egg and the early embryo stage (Kimmel et al. 1995; Lo et al. 2003; Mathavan et al. 2005; Sarropoulou et al. 2005a; Yufera et al. 2012). It also has been shown that specific genes show certain patterns of RNA degradation during early development and that maternal RNA is present (Wen et al. 2005; Bai et al. 2007; Ferg et al. 2007). Furthermore, it has been demonstrated that the sudden decrease after germinal breakdown is characteristic for maternal genes (Mommens et al. 2010). Studies in fish addressing this hypothesis are however limited. The role of maternal mRNA in early development has been studied mainly in mammals and in model fish species. In the model fish *Danio rerio*, it has been shown that the transition of maternal mRNA to zygotic mRNA occurs at the mid-blastula stage and is known as the maternal-embryo transition (MET) or mid-blastula transition (MBT) (Kane and Kimmel 1993). Several genes in non-model fish species revealed that the characteristic gene expression pattern according to the MET

Electronic supplementary material The online version of this article (doi:10.1007/s10126-012-9486-8) contains supplementary material, which is available to authorized users.

C. G. Fernández · C. Roufidou · E. Sarropoulou (✉)
Institute of Marine Biology, Biotechnology and Aquaculture,
Hellenic Centre for Marine Research,
P.O. Box 2214, 71003 Heraklion, Crete, Greece
e-mail: sarris@hcmr.gr

C. G. Fernández
Laboratory of Aquaculture Genetic,
Department of Functional Biology, University of Oviedo,
Julián Clavería s/n,
33006 Oviedo, Spain

C. Roufidou · E. Antonopoulou
Laboratory of Animal Physiology, Department of Zoology,
School of Biology, Aristotle University of Thessaloniki,
GR 54124 Thessaloniki, Greece

Published online: 02 October 2012

 Springer

is followed. Among those, genes involved in several different molecular functions were described, such as temperature acclimation (*Wap65-1*; Sarropoulou et al. 2010), immune system (*cathepsin Z*; Aegerter et al. 2005), and also basic cellular functions (*gapdh-1*; Sarropoulou et al. 2011).

The gilthead sea bream (*Sparus aurata* L.) is one of the main aquaculture species in the Mediterranean. The genomic toolbox of the gilthead sea bream has been significantly enriched in the recent years, facilitating the identification and characterization of genes. A broad range of gene expression studies using up-to-date technologies have been performed (Darias et al. 2008; Sarropoulou et al. 2005b, 2009). In addition, several studies have been performed in order to better understand the role of genes implicated in several physiological mechanisms such as growth rate and immune response (Davey et al. 2011) along with disease resistance (Casani et al. 2009), regulation of spermatogenesis (Chaves-Pozo et al. 2005, 2007), and assessment of flesh quality (Menoyo et al. 2004; Picone et al. 2012). The most decisive period however in the production of marine fish has been determined to be during the larval stage (Pavlidis and Mylonas 2011), which is set up during embryogenesis.

The aim of the present work was to identify and study the expression of genes which may have maternal origin in the gilthead sea bream, a species of commercial interest. Thus, we describe the expression patterns during the development of 16 genes (*gapdh 1* and 2, *cathepsin D, L, S* and *Z, erk1, jnk1, p38 alpha* and *p38 delta, ppar alpha, beta* and *gamma, tubulin beta, ferritin M, cyclinA2*) from seven different functional categories (Table 1). Genes were chosen based

on their putative involvement in the regulation of early development and egg quality (Howe et al. 1995; Braissant and Wahli 1998; Ibabe et al. 2005; Aegerter et al. 2005; Krens et al. 2006; Traverso et al. 2012). Nine sequences of the selected genes in sea bream are described here for the first time.

Materials and Methods

Tissue Sample Collection

Embryonic and larval samples of the gilthead sea bream were obtained from the aquaculture facility of the Institute of Marine Biology, Biotechnology, and Aquaculture of the Hellenic Centre for Marine Research in Crete, Greece. In total, eight different batches (biological replicates) of seven stages early in development were collected: (1) two- to four-cell stage, (2) 16-cell stage, (3) towards morula, (4) morula, (5) high blastula, (6) start of gastrula, and (7) one half epibolie (Supp. 1 of the “Electronic Supplementary Material”). All collected samples were frozen immediately in liquid nitrogen and stored at -80°C prior to RNA extraction.

RNA Extraction and cDNA Synthesis

RNA was extracted using the Trizol protocol in combination with the Lysing Matrix tubes (D). In brief, 500 μl Trizol and 5 μl mercaptoethanol were mixed, and 50 mg of eggs was added. Homogenization was performed with TissueLyser II (Qiagen) in 3 cycles of 40 s and 20 Hz. After centrifugation at 12,000g at 4°C for 5 min, the upper phase was transferred to a new 1.5-ml Eppendorf tube. A total of 300 μl of chloroform was added and the mixture was incubated at room temperature for 5 min. After another round of centrifugation at 12,000g for 5 min at 4°C , the upper phase was again transferred into a new 1.5-ml Eppendorf tube. Precipitation of total RNA was achieved by adding 500 μl cold 100 % ethanol and followed by incubation at -20°C overnight. RNA concentrations were determined using NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies Inc., Wilmington, DE, USA), and the quality was assessed by electrophoresis on 1 % ethidium bromide agarose gel as well as by the $A_{260/280}$ ratio. For cDNA synthesis, 1 μg of total RNA from each sample was reverse-transcribed in a final volume of 19 μl for 10 min at 25°C , 50 min at 42°C , and 15 min at 70°C using Superscript II First Strand Synthesis transcriptase (Invitrogen/VWR, Tromsø, Norway) primed with random hexamers. Reverse transcriptase negative controls were included for all samples to test for genomic DNA contamination. The obtained cDNA was diluted 1:50 in deionized Merck water (Whitehouse Station, NJ, USA) and stored as 50- μl aliquots at -20°C prior to further use for quantitative real-time PCR (qRT-PCR).

Table 1 Gene classification according to the gene ontology molecular function categories

Gene	Function
GAPDH1	Basic cell catabolic processes
GAPDH2	
Cathepsin D	Lysosomal enzymes
Cathepsin L	
Cathepsin S	
Cathepsin Z	
Tubulin β	Microtubule element
Ferritin M	Iron storage
Cyclin A2	Regulator of CDK kinases
PPAR α	Ligand-activated transcription factors
PPAR β	
PPAR γ	
ERK1	Mitogen-activated protein kinases (signaling pathways)
JNK1	
p38 α	
p38 δ	
L13a (reference gene)	Ribosomal protein
S18 (reference gene)	Ribosomal protein

Database Mining and Sequence Analysis

To identify homologous sequences of the gilthead sea bream for each gene, sequences available for other fish species were used to search the NCBI non-redundant (nr) sequence database, the EST database as well as EST sequences from previous unpublished EST libraries. Each sequence was submitted to BlastN and BlastX search, and predicted ORF (ORF finder, <http://www.ncbi.nlm.nih.gov>) was verified by BlastP against NCBI non-redundant protein sequence database.

Primer Design and Evaluation

Primers (Table 1; Supp. 1 of the “Electronic Supplementary Material”) were designed by using PRIMER 3 and FastPCR software (Kalendar et al. 2009). All primer pairs were tested by Beacon designer 7.5 (<http://www.premierbiosoft.com/>) and NET primer (<http://PremierBiosoft.com/netprimer/>). Primers for reference genes are described in Mitter et al. (2009). All primers were tested with genomic DNA and cDNA on agarose gel prior to real-time PCR for the presence of primer dimers and for verification of product sizes. Primer pairs of genes with no expression (absence of PCR product on agarose gel, no signal when performing real-time PCR) in early stages were not included in the final qPCR analysis.

Quantitative Real-Time PCR

DNA Engine Opticon 2 Continuous Fluorescence Detector (MJ Research, MA, USA) was used for qRT-PCR. Each reaction filled up to an end volume of 25 μ l with deionized Merck water comprised 5 μ l of 50 \times diluted cDNA and 12.5 μ l SYBR Green PCR Master Mix (Applied Biosystems, Warrington, UK). All samples (seven stages and eight biological replicates) were run in triplicates (technical replicates) including one reverse-transcription negative control for each gene. After an initial denaturation step at 95 $^{\circ}$ C for 10 min, 40 amplification cycles of 15 s at 95 $^{\circ}$ C and 1 min at 57 $^{\circ}$ C followed. After each cycle, a plate read for fluorescent signal assessment was performed. Finally, a dissociation curve with a gradient of 30 to 99 $^{\circ}$ C was performed after thermo-cycling to evaluate the specificity of the amplification and to verify the absence of primer dimers.

Data Acquisition

C_t values as well as the PCR efficiency was calculated using the Miner software (Zhao and Fernald 2005) based on the exported raw fluorescence data. The Miner software is a method for quantifying qRT-PCR results using calculations based on the kinetics of individual PCR reactions without the need of standard curves. The raw fluorescence data is fit as a function of PCR cycles to identify the exponential

amplification phase of the reaction using a four-parameter logistic model. This algorithm is an objective and noise-resistant method to quantify qRT-PCR results.

Statistical Analyses

The expression levels for all transcripts under study were assessed and normalized against the two reference genes, *ribosomal protein L13a (L13a)* and *ribosomal protein S18 (S18)*. Using GeNorm, the suitability of the reference genes was tested and the normalization factors were calculated as the geometric average of the two reference genes (Vandesompele et al. 2002). Besides the present investigations, previous studies in developmental stages of fish like in the European sea bass (Mitter et al. 2009) and in the Atlantic halibut have shown that both reference genes have a stable expression throughout development. Fernandes et al. (2008) and McCurley and Callard (2008) also showed this in zebrafish as well as their suitability. After converting the C_t values into quantities using the delta C_t formula for transforming C_t values to relative quantities, gene expression levels were normalized by dividing the raw data by the appropriate normalization factors. Expression levels and significance were confirmed by REST (Pfaffl et al. 2002). Finally, one-way ANOVA was performed as well as Tukey’s post hoc test in order to determine the statistically significant differences between developmental stages.

Results and Discussion

Gene Identification

In total, nine new genes, namely, *cathepsin S*, *cyclin A2*, *jnk1*, *p38alpha*, *p38delta*, *cathepsin Z*, *tubulin beta*, *ferritin M*, and *erk1*, were identified for the gilthead sea bream by database mining of the NCBI EST database as well as of EST sequences originating from unpublished EST libraries (Table 2; Supp. 3 of the “Electronic Supplementary Material”). The remaining seven genes were previously published (Carnevali et al. 1999, 2008; Leaver et al. 2005; Sarropoulou et al. 2011). All 16 genes were classified to seven different categories according to the gene ontology molecular function categories: basic cell catabolic processes, lysosomal enzymes, microtubule element, iron storage, regulator of CDK kinases, ligand-activated transcription factors, and mitogen-activated protein kinases (Table 1).

Gene Expression Analysis

Using classical RT-PCR as well as qPCR, four out of the 16 genes, namely, *gapdh-2*, *cathepsin S*, *ppar alpha*, and *ppar gamma*, did not show any expression in the developmental stages (Fig. 1).

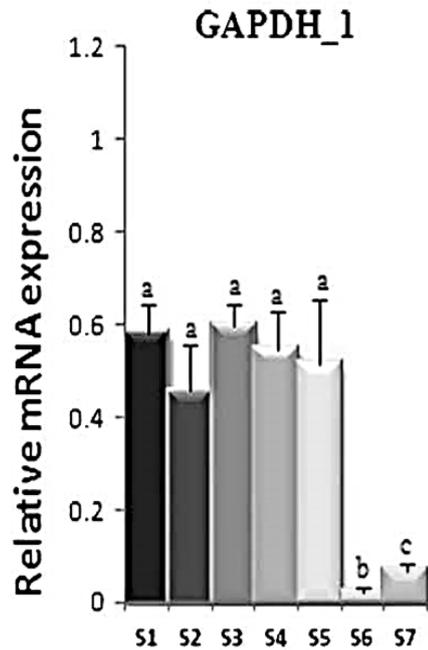


Fig. 2 Gene expression of GAPDH_1 during the gilthead sea bream development. S1 two- to four-cell stage, S2 16-cell stage, S3 towards morula, S4 morula, S5 high blastula, S6 start of gastrula, S7 one half epibolie

discussed earlier, the other three *cathepsins L, D, and Z* followed the expression pattern described earlier (Fig. 3). The same pattern showed also three out of the four mitogen-activated protein kinases (MAPK) under study (Fig. 4). MAPKs are important factors not only for the regulation of cellular activities like cell growth, proliferation, differentiation, survival, innate immunity, and development (Bogoyevitch and Court 2004; Roux and Blenis 2004; Johnson et al. 2005) but also as balance regulators between survival and apoptosis (Marshall 1995; Ballif and Blenis 2001). The MAPK family contains three subfamilies represented by *erk1* (extracellular signal-regulated kinases), *jnk1* (c-Jun amino-terminal kinases) and *p38 alpha* and *p38 delta* (p38 subfamily). Whereas *p38*

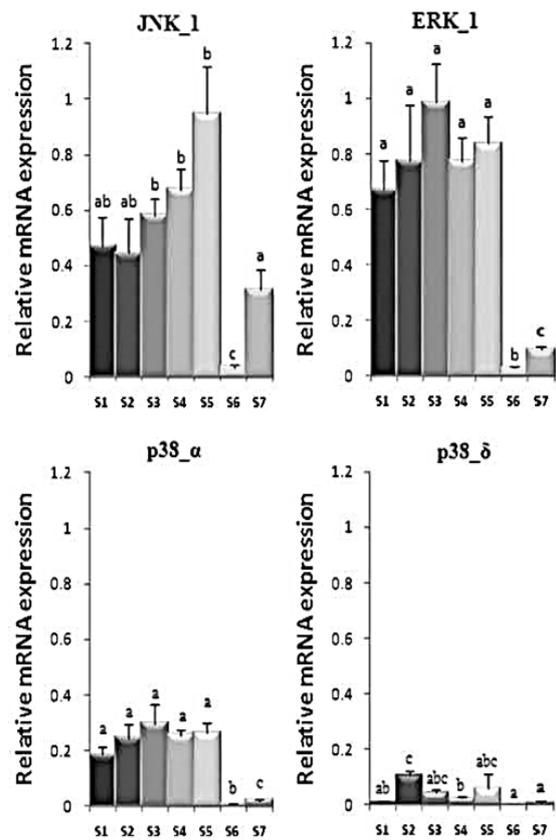
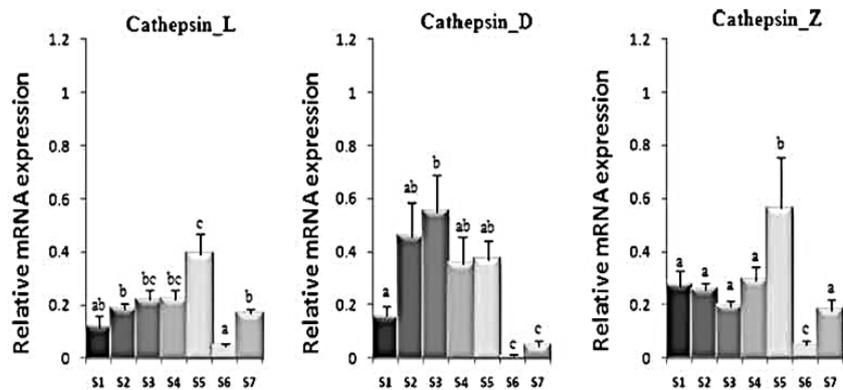


Fig. 4 Gene expression of *jnk1*, *erk1*, *p38alpha*, and *p38 delta* during the gilthead sea bream development. S1 two- to four-cell stage, S2 16-cell stage, S3 towards morula, S4 morula, S5 high blastula, S6 start of gastrula, S7 one half epibolie

alpha is the best characterized isoform among the p38 family members, only little information is available for *p38 delta*. In zebrafish, it has been shown that *p38 alpha* is required for symmetric and synchronous cleavage. *P38 delta* was cloned and characterized as the fourth member of the p38 MAPK, and expression was shown in different developmental stages in the

Fig. 3 Gene expression of *cathepsin L, D, and Z* during the gilthead sea bream development. S1 two- to four-cell stage, S2 16-cell stage, S3 towards morula, S4 morula, S5 high blastula, S6 start of gastrula, S7 one half epibolie



mouse embryo (Hu et al. 1999). Robust gene expression was found in the developing mouse epidermis, pinpointing the physical role in epidermal tissue.

Out of the three *ppar* studied, *ppar beta* revealed expression in the early stages of development and a typical sudden decrease after S5 (Fig. 5). Only little is known about the expression of *ppars* in fish development. Braissant and Wahli (1998) showed that *ppar beta* is expressed very early in mouse embryogenesis with a peak of expression in the developing nervous system, with a subsequent decrease. These results, in combination with the present ones, indicate the need of further functional research of *ppar beta* during development. The present study is, to our knowledge, the first study of *ppar* expression during early development in fish. The next two genes studied, namely, *tubulin beta* and ferritin, also had the typical pattern (Fig. 5). *Tubulin beta* is an essential component of cytoskeleton microtubules, and in rainbow trout the lower expression of *tubulin beta* pointed to low-quality eggs (Aegerter et al. 2005). The major function of ferritin is iron storage. In vertebrates, three types of

ferritin subunits have been described: H, L, and M. The M subunit type has been identified in lower vertebrates including teleosts (Andersen et al. 1995, 1998; Mignogna et al. 2002; Giorgi et al. 2008). The function of M ferritin-like has up until today not been extensively studied. One interesting finding recently in the smooth tongue sole showed a possible antimicrobial activity against a wide range of fish bacterial pathogens (Wang et al. 2011). In the present study, we show the potential function of M ferritin during development and following the pattern characteristic for maternal genes. The last group under study is the cyclin genes, represented by *cyclin A2*. In zebrafish, it has been shown that most of the cyclin genes start their expression maternally (Mathavan et al. 2005). Studies of maternal gene expression in the Atlantic halibut also found *cyclin A2* to be highly abundant (Mommens et al. 2010). In the present study, *cyclin A2* was used and the results were in concordance with the microarray data of Mathavan et al. 2005 (Fig. 5). Interestingly, studies in the Atlantic cod have shown that maternal genes are closely linked to egg quality (Mommens et al. 2010). However, in fish, the link of maternal mRNA contribution to the regulation of egg quality is limited and the possible contribution of maternal genes to the control of egg quality has only been hypothesized (Brooks et al. 1997). Up until today only little information is available for farmed fish concerning the identification of maternal genes and their regulation as being linked to egg quality. Fish populations are dependent on the production of good egg quality. Thus, one of the bottlenecks in commercial fish farming is the high variability of egg quality as it is closely linked to survival rates and to significant problems during the early production stages. Egg quality control is of significance especially for fish species like the gilthead sea bream where the broodstocks produce large quantities of eggs with a great and non-controllable quality variation (Kjorsvik et al. 1990). Studies examining egg quality in the gilthead sea bream as well as in the European sea bass showed that the amount of *cathepsin D* activity is much higher in sinking eggs than in good-quality eggs (Camevali et al. 1999, 2001). In the present study, *cathepsin D* follows the characteristic expression pattern for maternal genes, supporting the hypothesis that maternal genes may be closely linked to egg quality. Hence, in conjunction with the previous published results, *cathepsin D* may be one of the candidate biomarkers for egg quality. Similar studies for *tubulin beta* and *cathepsin Z* in rainbow trout showed that *tubulin beta* mRNA levels are lower and *cathepsin Z* mRNA levels are higher in low-quality eggs (Aegerter et al. 2005). However, besides several general parameters such as maternal age and condition factor, the timing of the spawning cycle, environmental factors, physicochemical conditions as well as genetic factors known to affect egg quality (Kjorsvik et al. 1990; Brooks et al. 1997), only little information is available concerning egg quality in sea bream.

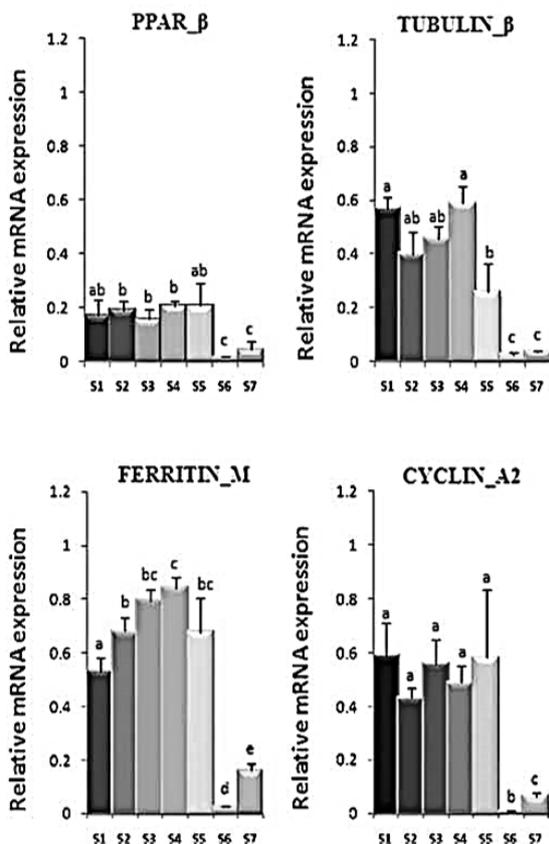


Fig. 5 Gene expression of *ppar-beta*, *tubulin-beta*, *ferritin M*, and *cyclin A2* during the gilthead sea bream development. S1 two- to four-cell stage, S2 16-cell stage, S3 towards morula, S4 morula, S5 high blastula, S6 start of gastrula, S7 one half epibolie

Conclusion

Early studies of embryonic development have shown that maternal factors are regulating several developmental processes (Tata 1986). It has also been shown that they are involved in germline establishment, pattern formation, gastrulation, neurogenesis as well as somatogenesis (for a review, see Traverso et al. (2012) and Pelegri (2003)). Consequently, the molecular assessment of gene expression in the early developmental stages of sea bream may be of significance to determine early offspring quality, which in turn is crucial for aquaculture practices. The identification and characterization of possible maternal genes in non-model fish species with important aquaculture interest may also be of importance for future studies related to egg quality.

In the present study, we have studied the gene expression of 16 candidate genes of different functional categories in seven developmental stages. Among them, 11 showed a characteristic gene expression pattern pinpointing to the possible function as maternal genes and thus may function as egg quality indicators, which however remain to be investigated in subsequent studies.

Acknowledgments Authors would like to thank the Department of Aquaculture and specifically Dr. Constantinos Mylonas and Dr. Pascal Divanach for providing embryonic and larvae stages. We also would like to thank FPU (MICINN, Spanish Ministry of Science and Innovation, AP2007-03178).

References

- Aegerter S, Jalabert B, Bobe J (2005) Large scale real-time PCR analysis of mRNA abundance in rainbow trout eggs in relationship with egg quality and post-ovulatory ageing. *Mol Reprod Dev* 72:377–385
- Andersen O, Dehli A, Standal H, Giskegerde TA, Karstensen R, Rorvik KA (1995) Two ferritin subunits of Atlantic salmon (*Salmo salar*): cloning of the liver cDNAs and antibody preparation. *Mol Mar Biol Biotechnol* 4:164–170
- Andersen O, Pantopoulos K, Kao HT, Muckenthaler M, Youson JH, Pieribone V (1998) Regulation of iron metabolism in the sanguivore lamprey *Lampetra fluviatilis*—molecular cloning of two ferritin subunits and two iron-regulatory proteins (IRP) reveals evolutionary conservation of the iron-regulatory element (IRE)/IRP regulatory system. *Eur J Biochem* 254:223–229
- Bai J, Solberg C, Fernandes JM, Johnston IA (2007) Profiling of maternal and developmental-stage specific mRNA transcripts in Atlantic halibut *Hippoglossus hippoglossus*. *Gene* 386:202–210
- Ballif BA, Blenis J (2001) Molecular mechanisms mediating mammalian mitogen-activated protein kinase (MAPK) kinase (MEK)-MAPK cell survival signals. *Cell Growth Differ* 12:397–408
- Bogoyevitch MA, Court NW (2004) Counting on mitogen-activated protein kinases—ERKs 3, 4, 5, 6, 7 and 8. *Cell Signal* 16:1345–1354
- Braissant O, Wahli W (1998) Differential expression of peroxisome proliferator-activated receptor- α , - β , and - γ during rat embryonic development. *Endocrinology* 139:2748–2754
- Brooks S, Tyler CR, Sumpter JP (1997) Egg quality in fish: what makes a good egg? *Rev Fish Biol Fish* 7:387–416
- Camevali O, Centonze F, Brooks S, Marota I, Sumpter JP (1999) Molecular cloning and expression of ovarian cathepsin D in seabream, *Sparus aurata*. *Biol Reprod* 61:785–791
- Camevali O, Mosconi G, Cambi A, Ridolfi S, Zanuy S, Polzonetti-Magni A (2001) Changes of lysosomal enzyme activities in sea bass *Dicentrarchus labrax* eggs and developing embryos. *Aquaculture* 202:249–256
- Camevali O, Cionna C, Tosti L, Cerda J, Gioacchini G (2008) Changes in cathepsin gene expression and relative enzymatic activity during gilthead sea bream oogenesis. *Mol Reprod Dev* 75:97–104
- Casani D, Randelli E, Costantini S, Facchiano AM, Zou J, Martin S, Secombes CJ, Scapigliati G, Buonocore F (2009) Molecular characterisation and structural analysis of an interferon homologue in sea bass (*Dicentrarchus labrax* L.). *Mol Immunol* 46:943–952
- Chaves-Pozo E, Mulero V, Meseguer J, Garcia Ayala A (2005) An overview of cell renewal in the testis throughout the reproductive cycle of a seasonal breeding teleost, the gilthead seabream (*Sparus aurata* L.). *Biol Reprod* 72:593–601
- Chaves-Pozo E, Liarte S, Vargas-Chacoff L, Garcia-Lopez A, Mulero V, Meseguer J, Mancera JM, Garcia-Ayala A (2007) 17 β -estradiol triggers postspawning in spermatogenically active gilthead seabream (*Sparus aurata* L.) males. *Biol Reprod* 76:142–148
- Darias MJ, Zambonino-Infante JL, Hugot K, Cahu CL, Mazurais D (2008) Gene expression patterns during the larval development of European sea bass (*Dicentrarchus labrax*) by microarray analysis. *Mar Biotechnol* (NY) 10:416–428
- Davey GC, Caldach-Giner JA, Houeix B, Talbot A, Sitja-Bobadilla A, Prunet P, Perez-Sanchez J, Cairns MT (2011) Molecular profiling of the gilthead sea bream (*Sparus aurata* L.) response to chronic exposure to the myxosporean parasite *Enteromyxum leei*. *Mol Immunol* 48:2102–2112
- Ferg M, Sanges R, Gehrig J, Kiss J, Bauer M, Lovas A, Szabo M, Yang L, Straehle U, Pankratz MJ, Olasz F, Stupka E, Muller F (2007) The TATA-binding protein regulates maternal mRNA degradation and differential zygotic transcription in zebrafish. *EMBO J* 26:3945–3956
- Fernandes JM, Mommens M, Hagen O, Babiak I, Solberg C (2008) Selection of suitable reference genes for real-time PCR studies of Atlantic halibut development. *Comp Biochem Physiol B Biochem Mol Biol* 150: 23–32
- Giorgi A, Mignogna G, Bellapadrona G, Gattoni M, Chiaraluce R, Consalvi V, Chiancone E, Stefanini S (2008) The unusual co-assembly of H- and M-chains in the ferritin molecule from the Antarctic teleosts *Trematomus bernacchii* and *Trematomus newnesi*. *Arch Biochem Biophys* 478:69–74
- Howe JA, Howell M, Hunt T, Newport JW (1995) Identification of a developmental timer regulating the stability of embryonic cyclin A and a new somatic A-type cyclin at gastrulation. *Genes Dev* 9:1164–1176
- Hu MC, Wang YP, Mikhail A, Qiu WR, Tan TH (1999) Murine p38-delta mitogen-activated protein kinase, a developmentally regulated protein kinase that is activated by stress and proinflammatory cytokines. *J Biol Chem* 274:7095–7102
- Ibabe A, Bilbao E, Cajaraville MP (2005) Expression of peroxisome proliferator-activated receptors in zebrafish (*Danio rerio*) depending on gender and developmental stage. *Histochem Cell Biol* 123:75–87
- Johnson GL, Dohlman HG, Graves LM (2005) MAPK kinase kinases (MKKKs) as a target class for small-molecule inhibition to modulate signaling networks and gene expression. *Curr Opin Chem Biol* 9:325–331
- Kalendar R, Lee D, Schulman AH (2009) FastPCR software for PCR primer and probe design and repeat search. *Genes Genomes Genomics* 3(1):1–14

- Kane DA, Kimmel CB (1993) The zebrafish midblastula transition. *Development* 119:447–456
- Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF (1995) Stages of embryonic development of the zebrafish. *Dev Dyn* 203:253–310
- Kjorsvik E, Mangor-Jensen A, Holmefjord I (1990) Egg quality in fishes. *Adv Mar Biol* 26:71–113
- Krens SF, He S, Spaink HP, Snaar-Jagalska BE (2006) Characterization and expression patterns of the MAPK family in zebrafish. *Gene Expr Patterns* 6:1019–1026
- Leaver MJ, Boukouvala E, Antonopoulou E, Diez A, Favre-Krey L, Ezaz MT, Bautista JM, Tocher DR, Krey G (2005) Three peroxisome proliferator-activated receptor isotypes from each of two species of marine fish. *Endocrinology* 146:3150–3162
- Lo J, Lee S, Xu M, Liu F, Ruan H, Eun A, He Y, Ma W, Wang W, Wen Z, Peng J (2003) 15000 unique zebrafish EST clusters and their future use in microarray for profiling gene expression patterns during embryogenesis. *Genome Res* 13:455–466
- Marshall CJ (1995) Specificity of receptor tyrosine kinase signaling: transient versus sustained extracellular signal-regulated kinase activation. *Cell* 80:179–185
- Mathavan S, Lee SG, Mak A, Miller LD, Murthy KR, Govindarajan KR, Tong Y, Wu YL, Lam SH, Yang H, Ruan Y, Korzh V, Gong Z, Liu ET, Lufkin T (2005) Transcriptome analysis of zebrafish embryogenesis using microarrays. *PLoS Genet* 1:260–276
- McCurley AT, Callard GV (2008) Characterization of housekeeping genes in zebrafish: male-female differences and effects of tissue type, developmental stage and chemical treatment. *BMC Mol Biol* 9:102
- Menoyo D, Izquierdo MS, Robaina L, Gines R, Lopez-Bote CJ, Bautista JM (2004) Adaptation of lipid metabolism, tissue composition and flesh quality in gilthead sea bream (*Sparus aurata*) to the replacement of dietary fish oil by linseed and soyabean oils. *Br J Nutr* 92:41–52
- Mignogna G, Chiaraluca R, Consalvi V, Cavallo S, Stefanini S, Chiancone E (2002) Ferritin from the spleen of the Antarctic teleost *Trematomus bernacchii* is an M-type homopolymer. *Eur J Biochem* 269:1600–1606
- Mitter K, Kotoulas G, Magoulas A, Mulero V, Sepulcre P, Figueras A, Novoa B, Sarropoulou E (2009) Evaluation of candidate reference genes for QPCR during ontogenesis and of immune-relevant tissues of European seabass (*Dicentrarchus labrax*). *Comp Biochem Physiol B Biochem Mol Biol* 153:340–347
- Mommens M, Fernandes JM, Bizuayehu TT, Bolla SL, Johnston IA, Babiak I (2010) Maternal gene expression in Atlantic halibut (*Hippoglossus hippoglossus* L.) and its relation to egg quality. *BMC Res Notes* 3:138
- Pavlidis M, Mylonas C (2011) Sparidae. Biology and aquaculture of gilthead seabream and other species. Wiley-Blackwell, Oxford
- Pelegri F (2003) Maternal factors in zebrafish development. *Dev Dyn* 228:535–554
- Pfaffl MW, Horgan GW, Dempfle L (2002) Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Res* 30:e36
- Picone G, Engelsen SB, Savorani F, Testi S, Badiani A, Capozzi F (2012) Metabolomics as a powerful tool for molecular quality assessment of the fish *Sparus aurata*. *Nutrients* 3:212–227
- Roux PP, Blenis J (2004) ERK and p38 MAPK-activated protein kinases: a family of protein kinases with diverse biological functions. *Microbiol Mol Biol Rev* 68:320–344
- Sarropoulou E, Kotoulas G, Power DM, Geisler R (2005a) Gene expression profiling of gilthead sea bream during early development and detection of stress-related genes by the application of cDNA microarray technology. *Physiol Genomics* 23:182–191
- Sarropoulou E, Power DM, Magoulas A, Geisler R, Kotoulas G (2005b) Comparative analysis and characterization of expressed sequence tags in gilthead sea bream (*Sparus aurata*) liver and embryos. *Aquaculture* 243:69–81
- Sarropoulou E, Sepulcre P, Poisa-Beiro L, Mulero V, Meseguer J, Figueras A, Novoa B, Terzoglou V, Reinhardt R, Magoulas A, Kotoulas G (2009) Profiling of infection specific mRNA transcripts of the European seabass *Dicentrarchus labrax*. *BMC Genomics* 10:157
- Sarropoulou E, Fernandes JM, Mitter K, Magoulas A, Mulero V, Sepulcre MP, Figueras A, Novoa B, Kotoulas G (2010) Evolution of a multifunctional gene: The warm temperature acclimation protein Wap65 in the European seabass *Dicentrarchus labrax*. *Mol Phylogenet Evol* 55:640–649
- Sarropoulou E, Nousdili D, Kotoulas G, Magoulas A (2011) Functional divergences of GAPDH isoforms during early development in two perciform fish species. *Mar Biotechnol* (NY) 13:1115–1124
- Sun YN, Xu TJ, Wang JX, Cheng YZ, Wang RX (2011) Sequence and expression analysis of cathepsin S gene in the miuiy croaker *Miichthys miuiy*. *Fish Physiol Biochem* 37:761–765
- Tata JR (1986) Coordinated assembly of the developing egg. *Bioessays* 4:197–201
- Traverso JM, Fostier A, Bobe J (2012) Egg transcriptome, the maternal legacy to the embryo. In: Garth MRA, Fletcher (eds) *Aquaculture biotechnology*. Blackwell, Chichester
- Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 3:RESEARCH0034
- Wang W, Zhang M, Sun L (2011) Ferritin M of *Cynoglossus semi-laevis*: an iron-binding protein and a broad-spectrum antimicrobial that depends on the integrity of the ferroxidase center and nucleation center for biological activity. *Fish Shellfish Immunol* 31:269–274
- Wen C, Zhang Z, Ma W, Xu M, Wen Z, Peng J (2005) Genome-wide identification of female-enriched genes in zebrafish. *Dev Dyn* 232:171–179
- Yufera M, Halm S, Beltran S, Fuste B, Planas JV, Martinez-Rodriguez G (2012) Transcriptomic characterization of the larval stage in gilthead seabream (*Sparus aurata*) by 454 pyrosequencing. *Mar Biotechnol* (NY)
- Zhao S, Fernald RD (2005) Comprehensive algorithm for quantitative real-time polymerase chain reaction. *J Comput Biol* 12:1047–1064



Larval width as indicator of growth rate and effect of larval classification on final body composition and flesh quality in cultured gilthead seabream (*Sparus aurata*, L.)

By I. Mazzeo¹, Y. J. Borrell², V. Gallego¹, C. García Fernández², J. A. Sánchez², G. Blanco², L. Pérez¹ and J. F. Asturiano¹

¹Grupo de Acuicultura y Biodiversidad, Instituto de Ciencia y Tecnología Animal, Universitat Politècnica de València, Valencia, Spain; ²Departamento de Biología Funcional, Laboratorio de Genética Acuicola, IUBA, Universidad de Oviedo, Oviedo, Asturias, Spain

Summary

The objective of this study was evaluating the effects on the growth rate and flesh quality of separating larvae according to their larval width in cultured gilthead seabream (*Sparus aurata*, L.). Progeny from two broodstocks (A and B) were divided into two groups according to larval width: heads (being the larger specimens) and tails. After 18 months both groups were analysed to evaluate growth, carcass traits and proximal composition. The head specimens reached a larger size and showed a greater level of well-being and degree of nourishment. Total body and fillet composition in general were found to be as expected for this species. Some differences were found in total body composition between head and tail specimens in both batches, and in the fillet composition in batch B (higher fat and lower moisture content in the head specimens). Sensory analyses were carried out using untrained panelists, who were unable to distinguish between the head and tail samples in batch A, whereas differences were noticeable in batch B. Samples from the head group were judged to be tastier and juicier, a consequence of their higher fat content. Hence, fish separation according to larval width is an effective tool to separate progeny into slow and fast growing groups, whereas the total body and fillet analyses and a sensory test ensure that the selection does not generate negative effects on the product quality.

Introduction

The gilthead seabream (*Sparus aurata* L.) represents one of the most important cultured species in Mediterranean Sea. Global production reached 151,346 tonnes in 2011, with Greece, Turkey, and Spain being the main producers (APROMAR, 2012). To improve production, companies usually focus their efforts on nutrition, stock management, disease prevention and facility location (Navarro et al., 2009a). However, little attention is given to genetic aspects, due to the high cost and no immediate results. Nevertheless, genetic selection and rearing of breeders with high genetic variability are important tools to improve progeny quality as well as avoid inbreeding, a bigger problem in fish farming than in livestock (Borrell et al., 2007).

Selection programmes focus on growth rates, food conversion efficiency, fecundity, meat quality and resistance to stress and diseases, whereby enhancing these characteristics can lead to shorter rearing times and lower mortality rates (Gjedrem, 1983; Afonso et al., 1998; Thorland et al., 2007; Dupont-Nivet et al., 2008; Antonello et al., 2009; Navarro et al., 2009a).

In selection programmes, composition analyses are required to ensure that characteristics of the final product are not altered through the selection process. Body composition and flesh quality, especially skeletal muscle composition and fat deposition, can influence consumer choice because of their effect on the fish appearance, smell, and taste (Grigorakis, 2007).

Favourable and unfavourable traits can be positively correlated, as in the case of Condition Factor (C_F) and visceral fat – desirable and undesirable, respectively – in *S. aurata* (Grigorakis and Alexis, 2005; Navarro et al., 2009b) or as in high harvest weight (favourable) and flesh fat content (unfavourable) in Atlantic salmon (*Salmo salar* L.) (Quinton et al., 2005).

In *S. aurata*, high heritability of visceral fat has been found (Navarro et al., 2009b), but consumers do not appreciate this undesirable characteristic (Grigorakis, 2007); therefore, its indirect selection should be monitored and avoided.

However, a positive correlation between size and fat deposition in *S. aurata* (Grigorakis and Alexis, 2005) as well as weight and length (Navarro et al., 2009b) has been found whereby an increase in both characteristics is desirable because they increase the size, appearance, taste and juiciness of the fish.

Efficient breeding programmes for gilthead seabream are lacking (Thorland et al., 2007). However, using samples from the progenies of the broodstocks used in the present study, it has been demonstrated that microsatellite analysis is a good molecular tool for assisting breeding programmes, as it allows us to identify the breeders primarily contributing to the fast-growing progenies (Borrell et al., 2011).

Complementing these previous results, the main aim of this experiment was to test other non-genetic tools (such

as the larval classification, and the evaluation of flesh composition) for complementary use as a part of selection programmes. Hence, differences were evaluated in growth rate, body parameters, proximate composition and flesh composition in progenies with different larval width obtained from two different broodstocks (A and B).

Materials and methods

Rearing and fish samplings

In December 2006, two broodstocks (Broodstock A: 29 females, average weight 1998.2 ± 43.5 g; 11 males, average weight 1877.2 ± 83.4 g; Broodstock B: 28 females, average weight 1992.8 ± 50.1 g; 10 males, average weight 1950.0 ± 76.3 g) from a total of 101 *S. aurata* adults (80 females and 21 males) from the hatchery Piscicultura Marina Mediterránea, S.L. Burriana, Castellón, Spain, were grouped using a combined optimization approach following procedures described in Borrell et al. (2007), ensuring the same genetic variability in both broodstocks.

In March 2007, the breeders were allowed to spawn freely. Fertilized eggs from one single day were collected from each broodstock and incubated under identical conditions in two tanks until hatching. Larvae were reared in large tanks and after 86 and 78 days post-hatch the offspring (batch A and B, respectively) were classified according to body width using different sieves, thus a slow and a fast growing group, namely tails and heads, respectively, were obtained from each batch. As measuring each of the fish was not viable, the selection size was determined according to the mean width after testing by using different sized sieves. In batch A and batch B, those fish with a body width greater than 4.5 mm or less than 3.5 mm were selected as heads and tails, respectively. The difference between batch A and B was the different ages of the larvae. Weights were obtained in the UPV facility with a precision balance and using absorbent tissue paper to minimise the effect of humidity (Table 1).

Ultimately, four fish groups were obtained: A-Heads (AH), A-Tails (AT), B-Heads (BH), and B-Tails (BT).

On days 165 and 157 post-hatch (batches A and B, respectively) a sample of approximately 1,000 fish per group (mean fish weight of head groups: 6.5 g; mean fish weight of tail groups: 2.0 g) was moved to the facilities at the Universitat Politècnica de València in Valencia, Spain. The fish were distributed in 1,750 L fiberglass tanks and reared under intensive conditions. The tanks were kept in a recirculating marine water system (30 m³ capacity) with a

rotary mechanical filter and a gravity biofilter. All tanks were equipped with aeration and with water temperature maintained at $22 \pm 3^\circ\text{C}$ by a heat pump installed in the system, and oxygen at 6.25 ± 0.44 ml L⁻¹. Photoperiod was natural. Fish were fed by hand using commercial fish feed (Dibaq S.A., Segovia, Spain) twice a day to apparent satiety. Feed size was chosen according to fish size. Rearing conditions were as much as possible similar to those present in fish farms.

At an intermediate growth stage (311 and 303 days post-hatch, batch A and B, respectively) 250 juveniles from each group (corresponding to 25% of the total) were chosen by weight, with the final aim of selecting the largest fish for the head groups and the smallest fish for the tail groups. Weight ranges were established according to the estimated mean weight in each group. In batch A the heads with a > 41 g weight and tails with a < 15 g weight were selected; in batch B the heads with a > 31 g weight and tails with a < 12 g weight were selected.

Finally, the head and tail groups were distributed into four 4,000-L tanks where they were reared under intensive conditions. At an intermediate stage of growth (411 days post-hatch batch A and 403 days post-hatch batch B), 100 fish from each tank were weighed to calculate biomass and growth tendency.

At the end of the growth cycle (528 and 520 days post-hatch, batch A and B, respectively) all fish were weighed to establish the growth rate.

Fifteen *S. aurata* from each group were randomly measured to calculate body parameters and stored at -20°C until carcass and fillet proximate analysis (moisture, fat, protein and ash content). In addition, ten *S. aurata* from each group were gutted and filleted.

For the sensory test, the fish were filleted and the dorsal filets stored at -20°C until used.

Growth analysis

Growth was calculated by using weights recorded at 311, 411 and 528 days post-hatch (batch A) and 303, 403 and 520 days post-hatch (batch B).

Between 311 and 528 days post-hatch (batch A) and 303 and 520 days post-hatch (batch B) the parameters calculated to evaluate the differences between tail and head groups were:

$$\text{Daily growth index (DGI, \% / d)} \\ = 100 \times (\text{final weight}^{1/3} - \text{initial weight}^{1/3}) / d$$

$$\text{Feed Intake (FI, \% / d)} = 100 \times \text{feed consumption (g)} / \\ \text{average biomass (g)} \times d$$

$$\text{Feed Conversion Ratio (FCR)} = \text{feed offered (g)} / \\ \text{weight gain (g)}$$

Table 1
Weights and longitude at the moment of first larval selection

Batch	Size (mm)	Heads	Size (mm)	Tails
A	>4.5	0.74 ± 0.20 g	<3.5	0.18 ± 0.24 g
		3.91 ± 0.30 cm		2.52 ± 0.24 cm
B	>3.5	0.31 ± 0.10 g	<2.5	0.06 ± 0.02 g
		3.91 ± 0.30 cm		1.89 ± 0.19 cm

Thermal Growth Coefficient (TGC × 1000)

$$= (\text{Final weight}^{1/3} - \text{Initial weight}^{1/3}) / (\sum \text{°C effective} \times d)$$

Body analysis

Fifteen *S. aurata* from each group were each measured for total length plus weights of total body, liver, gonads, mesenteric fat, guts, head, skin, fins and bones.

For each fish, the hepatosomatic index [I_H (%) = 100 × liver weight (g)/total weight (g)], gonadosomatic index [I_G (%) = 100 × gonad weight (g)/total weight (g)], condition factor [C_F = 100 × total weight (g)/total length³ (cm³)], viscerosomatic index [I_V (%) = 100 × viscera weight (g)/total weight (g)], mesenteric fat index [I_M (%) = 100 × mesenteric fat weight (g)/total weight (g)], fillet yield [F_Y (%) = 100 × (total body weight head viscera fins bone skin - gonad) (g)/total weight (g)] and head percentage [H_P (%) = 100 × head weight (g)/total weight (g)] were calculated.

Proximate analysis

Ten specimens from each group were analysed for carcass composition following the AOAC procedure (1990). Moisture levels were determined by drying samples (2.5–3 g) in porcelain cups at 104°C for 24 h. Ash levels were determined by incinerating the dried samples at 550°C for 5 h. Protein content was determined in 0.5 g of lyophilized and minced fish using the Kjeldhal method. Fat content was determined in 0.5 g of lyophilized and minced fish using Soxhlet extraction.

In five specimens per group, the analyses previously described were carried out only on the fish fillets with the aim to evaluate the composition of the edible parts. All analyses were performed in triplicate. All of these traits are expressed as a percentage of fresh weight.

Sensory analysis

A triangle test following ISO 1420 (1983) was performed to evaluate the sensory differences between head and tail specimens. Dorsal fillets were thawed at room temperature, vacuum-packed in plastic bags and cooked for 10 min using a water-bath at 60°C. Each fillet was then cut into nine pieces, each weighing approximately 3–4 g. The resulting equally sized pieces were coded, wrapped in aluminum foil and kept in 40°C thermo regulated boxes until tested. To ensure that possible differences were not due to the effects of the fillet portions, samples from a similar part of the fillet were compared in each test. At each session, three triangle tests were presented to each panelist. A total of 120 triangles for 40 untrained panelists were prepared for each group who were then asked to identify the odd piece. 'No difference' was not acceptable as a valid response according to Meilgaard et al. (1999). The minimum number of correct answers (x) for a statistically significant result ($P < 0.05$) was calculated according to $x = 0.4174 \times 1.64 \times (\sqrt{n}) + (2n + 3)/6$ where n = number of responses (International Organization for Standardization, 1983).

Statistical analysis

After establishing data normality, an ANOVA (Newman-Keuls test) was carried out to compare results. If normality failed (GSI and weights for growth evolution) with transformed data, a Kruskal-Wallis test was carried out. All values are expressed as mean ± standard error (SEM).

Growth curves were compared by regression analysis and regression coefficients (slopes) were compared using analysis of variance. Differences were considered significant at $P < 0.05$. All statistical procedures were run using Statgraphics Plus® 5.1.

Results

Growth parameters at the end of the feeding period are shown in Table 2. During the feeding period, the head weight of the specimens was significantly higher. The tail specimens presented higher FI values, whereas no differences were recorded in DGI, TGC or FCR. Comparison of the slopes (b) of the two regression lines confirmed a higher growth rate in the head group specimens.

Body parameters at the end of the growth period are shown in Table 3. In batch A, higher body weights, total length, C_F and I_G were recorded in the head specimens than in the tails, demonstrating a significant difference in size. However, all specimens were sexually immature. No differences were found in the remaining parameters (I_H , I_V , I_M , F_Y and H_P). In batch B the head specimens showed higher values in all measured parameters, except the head percentage.

Regarding body composition, analyses were carried out on the whole body in both batch A and batch B (Table 4). No differences were found in either ash or protein content between head and tail groups; however, the head specimens presented a higher fat content and a lower moisture percentage than the tail specimens.

Concerning fillet composition (Table 5), no differences were recorded in batch A, where both head and tail groups demonstrated the same moisture, ash, protein and fat content. Nevertheless, in batch B, coinciding with the results observed in the whole body analyses, no differences were found in ash or protein content, whereas a higher fat content and a lower moisture percentage were recorded in head specimens.

Sensory test results are given in Table 6, whereby 120 triangle tests were performed. In both batches four replies of 'no difference' were rejected, thus only 116 tests were considered.

In the 116 triangle tests, the number of correct answers for a $P < 0.05$ was 47. For batch A the number of correct answers was 42, demonstrating that the panelists were unable to distinguish between head and tail specimens. However, for batch B the number of correct answers was 48, showing that the panelists could distinguish between heads and tails. In terms of differences noted, the panelists described the head samples as tastier and juicier than the tail samples in both batches.

Discussion

At the end of the growth period (528 days post-hatch for batch A and 520 days for batch B), higher length, weight

Table 5
Fillet composition at end of growth period. Values reported as mean \pm SEM

	AH	AT	BH	BT
Moisture (%)	73.50 \pm 0.41	74.56 \pm 0.41	71.82 \pm 0.39a	73.35 \pm 0.39b
Ash (%)	1.37 \pm 0.03	1.43 \pm 0.03	1.43 \pm 0.02	1.43 \pm 0.02
Protein (%)	20.20 \pm 0.28	20.56 \pm 0.28	20.08 \pm 0.27	20.70 \pm 0.27
Fat (%)	4.58 \pm 0.20	4.23 \pm 0.22	6.52 \pm 0.38b	4.61 \pm 0.38a

AH, A Heads; AT, A Tails; BH, B Heads; BT, B Tails. $n_{AH, AT, BH, BT} = 5$. Different letters mean significant differences between head and tail specimens of a same batch.

Table 6
Sensory test results

Set	AH vs AT	BH vs BT
Number of triangles	116	116
Number of correct responses	42	48
Significance	>0.05	<0.05

AH, A Heads; AT, A Tails; BH, B Heads; BT, B Tails.

heads in the present experiment (TGC = 1.49). Moreover, a lower FCR of 1.40 was recorded in the head specimens in the present study than the FCR of 1.93 obtained by Martínez-Llorens et al. (2008).

However, the tail group specimens selected during the experiment showed a lower TGC (TGC = 1.67) than that obtained by Sitjà-Bobadilla et al. (2005) (TGC = 2.04). Moreover, tail specimens from the present experiment showed a higher FCR (FCR = 1.35) than that obtained by Sitjà-Bobadilla et al. (2005) and Benedito-Palos et al. (2008) (1.18 and 0.99, respectively).

Therefore, size classification allows the possibility of keeping animals with faster growth and discarding those with slow growth.

Differences in C_F indicate that the head specimens have a higher level of well being and degree of nourishment than the tail specimens. C_F can vary according to sexual maturation, age, and sex, but in this case it is due only to weight and length differences as the specimens were the same age and had not reached sexual maturation. In fact, I_G for mature *S. aurata* is higher than 1% for females and 0.70 1% for males (Zohar et al., 1984), whereas the maximum I_G recorded at the end of the experiment was 0.2%. Sexual maturity was not reached, despite the age of the fish, probably because of the water temperature, which was maintained at 22°C, while *S. aurata* spawns during winter with an optimal temperature range of 15–17°C (Moretti et al., 1999). In any case, C_F results are consistent with those in the literature (Navarro et al., 2009a,b; Sánchez-Lozano et al., 2009).

In batch A, body parameter differences were limited in length, weight, C_F and I_G . However, in batch B, the head specimens presented differences in all considered parameters

except the head percentage, probably as consequence of the greater difference between head and tail specimens in this batch.

Results reported here for I_M and I_H are similar to those shown by Grigorakis et al. (2002) for specimens reared at the same temperature. In the same way, I_V obtained in this experiment is consistent with the dressed percentage reported by Navarro et al. (2009a), as these parameters are inversely proportional.

The differences in the lipid content between head and tail specimens, recorded in the whole body analyses in both batches A and B, are due to their difference in weight. In fact, it has been demonstrated that muscle, perivisceral and peritoneal fat increase with size (Grigorakis and Alexis, 2005), as animals first use energy for their muscles and then to deposit fat (Gjedrem, 1997). In accordance with Navarro et al. (2009b) our experiments also demonstrated that the percentage of fat is inversely proportional to the percentage of moisture (Tables 4 and 5). When analysing only the fillets, no differences were found between the head and tail groups in batch A, whereas in batch B fillets from the head specimens presented a higher fat content and a lower moisture percentage than the tail specimens.

The differences between batch A and B with regard to body parameters and body and fillet composition are due to the greater weight differences between the head and tail specimens in batch B as opposed to batch A.

Protein and ash content maintained constant values in both batches A and B, both in whole body and fillet composition. In general, ash and protein are considered stable components in the fish body and do not seem to depend on weight (Navarro et al., 2009b). In salmonids, the mechanism for protein homeostasis seems to be more effective than that for lipid homeostasis, and body protein percentage shows low phenotypic and genetic variations (Kause et al., 2009). *Salmo salar* carcass protein values stabilize when fish reach 100 g in weight (Shearer et al., 1994). As highlighted by Grigorakis et al. (2002), no studies have been carried out to determine the relationship between body weight and protein percentage in Mediterranean fish.

In cultured *S. aurata* the values for fillet yield and moisture, ash, fat and protein in fillet are 31.3 48%, 68 76%, 1.2 1.6%, 2.5 11% and 18 23%, respectively (Grigorakis, 2007). Hence, the values in the present experiment (Tables 3 and 5) are those expected for this species.

During the sensory test, the panelists were unable to distinguish between the head and tail fish from batch A. When sampling fish from batch B, the panelists distinguished head from tail samples. In general, head samples were better appreciated because they were considered tastier and juicier. Taste and juiciness depend on the percentage of fat, and the opinion of the panelists is consistent with proximate analyses. The ability to distinguish head and tail samples in batch B and not in batch A is consistent with the differences in fillet composition. These results show that consumers can detect the differences in fish having different levels of fat content, confirming the importance of carrying out composition and quality analyses during selection programmes, as selection itself should not jeopardize the final product quality.

Due to the limitations of the experimental design, in particular related to the lack of true replicates, results obtained here have to be considered as preliminary. However conditions under which experiments were carried out are very similar to those of fish farms, especially the limited number of holding tanks. Moreover, the study aimed to identify a valid, effective and immediate tool that could be easily used in fish farms to optimize the productive cycle and obtaining fish with rapid growth.

Also, the aim was to study not only the growth performance of an experimental group but also the individual growth performance of some subsamples. Analysing fish individually allowed establishing the parental contribution of the breeders to the progeny, as shown in the parallel article by Borrell et al. (2011). The evidence that some breeders contribute more to fast growing progeny reinforces the idea that progeny is an important element for breeder selection.

Acknowledgements

Funded by the Coordinated Project SELECTBREAM (Spanish Ministry of Science and Innovation; AGL2007-64060-C03-00, including European Regional Development Funds); I. Mazzeo was supported by a FPI scholarship financed by Generalitat Valenciana; V. Gallego was supported by a FPI scholarship financed by the Spanish Ministry of Science and Innovation; C. García-Fernández was supported by a FPU scholarship financed by the Spanish Ministry of Science and Innovation.

References

- Afonso, J. M.; Montero, D.; Robaina, L.; Fernández, H.; Izquierdo, M.; Ginés, R., 1998: Selection programmes for stress tolerance in fish. *CIHEAM – Options Méditerranéennes* **34**, 235–245.
- Antonello, J.; Massault, C.; Franch, R.; Haley, C.; Pellizzari, C.; Bovo, G.; Patarnello, T.; de Koning, D. J.; Bargelloni, L., 2009: Estimates of heritability and genetic correlation for body length and resistance to fish pasteurellosis in the gilthead sea bream (*Sparus aurata* L.). *Aquaculture* **298**, 29–35.
- AOAC, 1990: Official Methods of Analysis, vols I and II. 15th edn. Association of Official Analytical Chemists, Arlington, 1298 pp.
- APROMAR, 2012: La acuicultura marina en España. <http://www.apromar.es/Informes/informe%202012/APROMAR%20Informe%20Anual%202012.pdf>. (In Spanish).
- Benedito-Palos, L.; Navarro, J. C.; Sitjà-Bobadilla, A.; Bell, J. C.; Kaushik, S.; Pérez-Sánchez, J., 2008: High levels of vegetable oils in plant protein-rich diets fed to gilthead sea bream (*Sparus aurata* L.): growth performance, muscle fatty acid profiles and histological alterations of target tissues. *Br. J. Nutr.* **100**, 992–1003.
- Borrell, Y. J.; Carleos, C. E.; Asturiano, J. F.; Bernardo, D.; Vázquez, E.; Corral, N.; Sánchez, J. A.; Blanco, G., 2007: Use of microsatellites and a combinatorial optimization approach in the acquisition of gilthead seabream (*Sparus aurata*, L.) broodstocks for hatcheries. *Aquaculture* **269**, 200–210.
- Borrell, Y. J.; Gallego, V.; García-Fernández, C.; Mazzeo, I.; Pérez, L.; Asturiano, J. F.; Carleos, C. E.; Vázquez, E.; Sánchez, J. A.; Blanco, G., 2011: Assessment of parental contributions to fast and slow growth progenies in the sea bream *Sparus aurata* L. using a new multiplex PCR. *Aquaculture* **314**, 58–65.
- Dupont-Nivet, M.; Vandeputte, M.; Vergnet, A.; Merdy, O.; Haffray, P.; Chavanne, H.; Chatain, B., 2008: Heritabilities and GxE interactions for growth in the European sea bass (*Dicentrarchus labrax* L.) using a marker-based pedigree. *Aquaculture* **275**, 81–87.
- Gjedrem, T., 1983: Genetic variation in quantitative traits and selective breeding in fish and shellfish. *Aquaculture* **33**, 51–72.
- Gjedrem, T., 1997: Flesh quality improvement in fish through breeding. *Aquacult. Int.* **5**, 197–206.
- Grigorakis, K., 2007: Compositional and organoleptic quality of farmed and wild gilthead sea bream (*Sparus aurata*, L.) and sea bass (*Dicentrarchus labrax*) and factors affecting it: a review. *Aquaculture* **272**, 55–75.
- Grigorakis, K.; Alexis, M. N., 2005: Effects of fasting on the meat quality and fat deposition of commercial-size farmed gilthead sea bream (*Sparus aurata* L.) fed different dietary regimes. *Aquacult. Nutr.* **11**, 341–344.
- Grigorakis, K.; Alexis, M. N.; Taylor, K. D. A.; Hole, M., 2002: Comparison of wild and cultured gilthead sea bream (*Sparus aurata*); composition, appearance and seasonal variations. *Int. J. Food Sci. Technol.* **37**, 477–484.
- International Organization for Standardisation, ISO 4120, 1983: Sensory Analysis. Triangle Test. International Organisation for Standardisation, Geneva.
- Kause, A.; Quinton, C. D.; Ruohonen, K.; Koskela, J., 2009: Genetic potential for the regulation of variability in lipid and protein content of European whitefish (*Coregonus lavaretus*). *Br. J. Nutr.* **101**, 1444–1451.
- Martínez-Llorens, S.; Tomás Vidal, A.; Moñino, A. V.; Gómez Ader, J.; Pla Torres, M.; Jover Cerdá, M., 2008: Blood and haemoglobin meal as protein sources in diets for gilthead sea bream (*Sparus aurata*): effects on growth, nutritive efficiency and fillet sensory differences. *Aquacult. Res.* **39**, 1028–1037.
- Meilgaard, M.; Civille, G. V.; Carr, B. T., 1999: Sensory evaluation techniques, 3rd edn. CRC Press, Boca Raton, 387 pp.
- Moretti, A.; Pedini Fernandez-Criado, M.; Cittolin, G.; Guidastri, R., 1999: Manual on Hatchery Production of Seabass and Gilthead Seabream. *FAO I*, 194.
- Navarro, A.; Zamorano, M. J.; Hildebrandt, S.; Ginés, R.; Aguilera, C.; Afonso, J. M., 2009a: Estimates of heritabilities and genetic correlations for growth and carcass traits in gilthead seabream (*Sparus aurata* L.), under industrial conditions. *Aquaculture* **289**, 225–230.
- Navarro, A.; Zamorano, M. J.; Hildebrandt, S.; Ginés, R.; Aguilera, C.; Afonso, J. M., 2009b: Estimates of heritabilities and genetic correlations for body composition traits and G x E interactions, in gilthead seabream (*Sparus aurata* L.). *Aquaculture* **295**, 183–187.
- Quinton, C. D.; McMillan, I.; Glebe, B. D., 2005: Development of an Atlantic salmon (*Salmo salar*) genetic improvement program: genetic parameters of harvest body weight and carcass quality traits estimated with animal models. *Aquaculture* **247**, 211–217.
- Sánchez-Lozano, N. B.; Martínez-Llorens, S.; Tomás-Vidal, A.; Jover Cerdá, M., 2009: Effect of high-level fish meal replacement by pea and rice concentrate protein on growth, nutrient utilization and fillet quality in gilthead seabream (*Sparus aurata*, L.). *Aquaculture* **298**, 83–89.
- Shearer, K. D.; Åsgård, T.; Andorsdóttir, G.; Aas, G. H., 1994: Whole body elemental and proximate composition of Atlantic salmon (*Salmo salar*) during the life cycle. *J. Fish Biol.* **44**, 785–797.
- Sitjà-Bobadilla, A.; Peña-Llopis, T. S.; Gómez-Requeni, P.; Médale, F.; Kaushik, S.; Pérez-Sánchez, J., 2005: Effect of fish meal replacement by plant protein sources on non-specific defence, mechanisms and oxidative stress in gilthead sea bream (*Sparus aurata*). *Aquaculture* **249**, 387–400.
- Thorland, I.; Papaioannou, N.; Kottaras, L.; Refstie, T.; Papisolomontos, S.; Rye, M., 2007: Family based selection for production traits in gilthead seabream (*Sparus aurata*) and European sea bass (*Dicentrarchus labrax*) in Greece. *Aquaculture* **272S**(1), 314.
- Zohar, Y.; Billard, R.; Weil, C., 1984: La reproduction de la daurade (*Sparus aurata*) et du bar (*Dicentrarchus labrax*) connaissance du cycle sexuel et contrôle de la gamétogenèse et de la ponte.

In: L'Aquaculture Du Bar et Des Sparidés. G. Barnabé and R. Billard (eds), Paris, INRA, pp. 3–24.

Author's address: Juan F. Asturiano, Grupo de Acuicultura y Biodiversidad, Instituto de Ciencia y Tecnología Animal,

Universitat Politècnica de València, Camino de Vera s/n. 46022, Valencia, Spain.
E-mail: jfastu@dca.upv.es



UNIVERSIDAD DE OVIEDO

**Departamento de Biología Funcional
Área de Genética**