Mejora de la competitividad del sector de la dorada (*Sparus aurata*), a través de la selección genética

Álvaro Lorenzo Felipe

PhD Thesis 2021

PhD program in Sustainable Aquaculture and Marine Ecosystems (ACUISEMAR®)

Grupo de Investigación en Acuicultura (GIA®)

Institute of Aquaculture and Sustainable Marine Ecosystems (IU-ECOAQUA®)

Universidad de Las Palmas de Gran Canaria

Directores

Mª Jesús Zamorano Serrano

Hyun Suk Shin

Improvement of the competitiveness of the gilthead seabream sector (*Sparus aurata*), using genetic selection

Álvaro Lorenzo Felipe

PhD Thesis 2021

PhD program in Sustainable Aquaculture and Marine Ecosystems (ACUISEMAR®)

Grupo de Investigación en Acuicultura (GIA®)

Institute of Aquaculture and Sustainable Marine Ecosystems (IU-ECOAQUA®)

Universidad de Las Palmas de Gran Canaria

Supervisors

Mª Jesús Zamorano Serrano

Hyun Suk Shin

LIST OF CONTENTS

LIST OF ABBREVIATIONS

Alpha-linolenic acid	(ALA)	Least-squares discriminant analysis	(*PLS-DA*)
Best linear unbiased prediction	(BLUP)	Linoleic acid	*(LA)*
Biological process gene ontology	(GO-BP)	Liter	(L)
Celsius degrees	(ºC)	Lordosis, Scoliosis and Kyphosis	(LSK)
Complementary deoxyribonucleic acid	(cDNA)	Marker-assisted selection	*(MAS)*
Days post hatching	(dph)	Meters	(m)
Deoxyribonucleic acid	*(DNA)*	Microliter	(µL)
Diethylpyrocarbonate	*(DEPC)*	Micrometer	(µm)
Diferentially expressed transcript	(DET)	Milliliter	(mL)
Differentially expressed gene	*(DEG)*	Millimeter	(mm)
Docosahexaenoic acid	(DHA)	Million	(M)
EBV for the prevalence of deformity	(EBVdef)	Minute	(min)
Eicosapentaenoic acid	(EPA)	Over-Representation Analysis	*(ORA)*
Estimated breeding value	(EBV)	Pair-base	(pb)
European Union	(EU)	Parts per million	*(ppm)*
Euros	*(€)*	Percent	*(%)*
Expressed sequence tag	(EST)	Polymerase chain reaction	*(PCR)*
Expression quantitative trait locus	(eQTL)	Quatitative trait locus	(QTL)
False discovery rate	(FDR)	Relative centrifugal force	*(RCF)*
Fold change	(FC)	Ribonucleic acid	*(RNA)*
Gene Set Enrichment Analysis	(GSEA)	Ribonucleic acid sequencing	(RNAseq)
Genetically deformed	(gD)	Sequence read archive	(SRA)
Genetically normal	(gN)	Single nucleotide polymorphism	(SNP)
Genome-wide	(GW)	Super-scaffolds	(SS)

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

Durante el transcurso de los estudios recogidos en esta tesis han colaborado, en distinto grado, un gran número de personas sin las cuales habría sido imposible obtener las conclusiones a las que se ha llegado finalmente. Dado que sería imposible nombrarlos a todos, quiero agradecer de manera general a todo aquel que haya contribuido en la construcción de esta investigación.

Por otro lado, es cierto que existe un grupo de personas que merecen una mención especial por haber realizado aportaciones de mayor índole y que han definido la trayectoria de este trabajo de investigación, y es por ello que no puedo finalizar este periodo sin agradecérselo.

En primer lugar quiero agradecer a mis directoras y a mi tutor, por su gran calidad científica, su implicación en esta tesis y sobre todo por su calidad humana. Gracias María Jesús Zamorano, por contribuir con tu experiencia y conocimiento sobre marcadores moleculares y por haber servido como punto de anclaje a la realidad y estabilidad emocional. Gracias Hyun Suk Shin, por transmitirme tu gran capacidad de trabajo, por tu profesionalidad y por sacrificarte en ocasiones para que lograra concluir mis experimentos. Gracias Juan Manuel Afonso, por tu capacidad investigadora, por permitirme estar siempre en primera línea, incluso fuera del ámbito de mi tesis, y por ayudarme a abrir mi mente y desarrollar la capacidad de poner cada resultado en perspectiva y no tomarlo como un hecho aislado o concluyente. Ha sido un verdadero privilegio poder realizar esta investigación bajo su tutela.

En segundo lugar quiero agradecer a dos personas que durante este período han estado día tras día, en el ámbito laboral y personal, realizando tareas incluso fuera de su ámbito profesional, por el hecho de ayudarme. Gracias, Guacimara Alejandro y Sergio León Bernabeu por su apoyo y amistad. Esta tesis es en gran parte gracias a ustedes.

En tercer lugar quiero agradecer a todos los miembros del instituto ECOAQUA, y especialmente a los miembros del GIA con los que he podido compartir mucho tiempo de trabajo. Gracias a los Seniors por su pluriempleabilidad, por hacer de junta de unión y conseguir que, dentro de sus posibilidades, todo funcione. Gracias a los doctorandos y alumnos, con los que no sólo he compartido tiempo material, sino sufrimiento, vivencias, reflexiones y aprendizajes. Gracias a los técnicos, por su necesaria labor, por servir de guía y facilitar el trabajo, por su inmensa ayuda y comprensión. Dentro de este tan diverso grupo de personas he encontrado en estos años apoyo, compañerismo y amistades, y por qué no decirlo, una segunda familia.

En último lugar dentro del ámbito laboral, quiero dar las gracias a otros grupos de la universidad y entidades externas a ella que han contribuido a que opte a ser doctor. Gracias a Aquanaria por permitirme desarrollar parte de la tesis en sus instalaciones, a Biogemar y el CAT por acogerme en las estancias, al CSIC y la UMA por los servicios de secuenciación y análisis sin los que no habría tesis, y a otros grupos de la ULPGC, INIA e IFAPA por sus orientaciones y ayuda en experimentos concretos. Bajo el nombre de estas instituciones que he querido nombrar de modo general, espero llegue mi más sentido agradecimiento a personal humano concreto que me ha formado laboralmente y me ha marcado personalmente.

Para terminar este capítulo de la tesis, quiero agradecer cómo no podía faltar, a mis AMIGOS y FAMILIA de sangre y de corazón, por estar ahí, por pese a ser ajenos a lo acontecido, comprenderme, apoyarme y tratar de animarme, por darme fuerzas y ayudar a evadirme y recordar que el trabajo no lo es todo. Gracias en especial a mi MADRE, PADRE y HERMANOS y muy en especial a mi PAREJA. Laurita, a ti te ha tocado sufrir todos y cada uno de mis fracasos y celebrar todos y cada uno de mis éxitos, en muchas ocasiones incluso con mayor efusividad que yo, gracias por estar siempre ahí.

FUNDING

The research was partly funded by the PerformFISH project no. 727610, Consumer driven production: Integrating innovative approaches for competitive and sustainable performance across the Mediterranean aquaculture value chain (EU Horizon 2020 programme)

The research was partly funded by the PROGENSA®-III project, Mejora de la Competitividad del Sector de la Dorada a Través de la Selección Genética (JACUMAR program, European Maritime and Fisheries Fund [EMFF])

The research was partly funded by the MORFOGEN project no. RTA2017-00054-C03-02, Estudio de los factores genéticos y moleculares que regulan la morfología en dorada (Sparus aurata) [Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA) funded from MCIU/AEI/FEDER, UE]

The research was co-financed by the Agencia Canaria de Investigación, Innovación y Sociedad de la Información de la Consejería de Economía, Industria, Comercio y Conocimiento, and by the European Social Fund (ESF) Integrated Operational Program of Canary Islands 2014–2020, Axis 3 Priority Theme 74 (85 %)

Chapter 1

Introduction

1.1 Economic impact of the gilthead seabream culture

The gilthead seabream is one of the most important aquaculture species in the Mediterranean countries, with a trade of 240,786 t in 2019 and a demand that has grown by 27.3 % in the last 5 years (APROMAR, 2016, 2017, 2018, 2019, 2020). In recent years, more than 95 % of the gilthead seabream's supply comes from aquaculture and the fishery volumes of this species have remained constant with an amount of 6,000-9,000 t per year. This means that market demand must be satisfied from aquaculture production, which resulted, in economic terms, in an amount of 1,251.9 M€ in 2020 only in the Mediterranean (APROMAR, 2021).

Gilthead seabream is currently farmed in 20 countries, and if we focus on EU data, its production in 2018 ranked third in farmed fish, after salmon and rainbow trout. In terms of economic value, it occupies the 4th position with an amount of 434.1 M€, having been surpassed in the last five years by sea bass (APROMAR, 2016, 2017, 2018, 2019, 2020).

The total production of seabream fingerlings in Europe (including Turkey) has remained at a value of approximately 700 million units since 2016, 20-25 % more than the quantity produced in the previous five years (2011-2015). The average estimated sale price in Spain is 0.22 €/unit, at an equivalent weight of 2 g per unit (APROMAR 2020), which would mean a market of 154 M€, if the price in the rest of the countries is similar.

1.2 Species under study

The gilthead seabream, *Sparus aurata* (*Linnaeus*, 1758), is a marine teleost belonging to the *Actinopterygii* class, *Perciformes* order, *Sparidae* family:

Kingdom: *Animalia*
Phylum: *Chordata*
Subphylum: *Vertebrata*
Class: *Actinopterygii*
Infraclass: *Teleostei*
Superorder: *Acanthopterygii*
Order: *Perciformes*
Family: *Sparidae*
Genus: *Sparus*
Species: *aurata*

It has an oval, deep and compressed body. It has a curved head profile, small eyes and a low, slightly oblique mouth with thick lips. It exhibits 4-6 canine-type teeth in the first row of teeth, followed by 2-4 rows of increasingly blunt teeth or molars. It has strong jaws and the teeth of the outer rows are stronger than the inner ones. It has a silvery-gray color and a large black spot at the origin of the lateral line that extends over the upper margin of the operculum. In addition, it has a characteristic golden front stripe between the eyes, bordered by two dark areas; longitudinal dark lines on the sides of the body; a dark band on dorsal fin; and the fork and caudal fin tips with a black border (Figure 1).

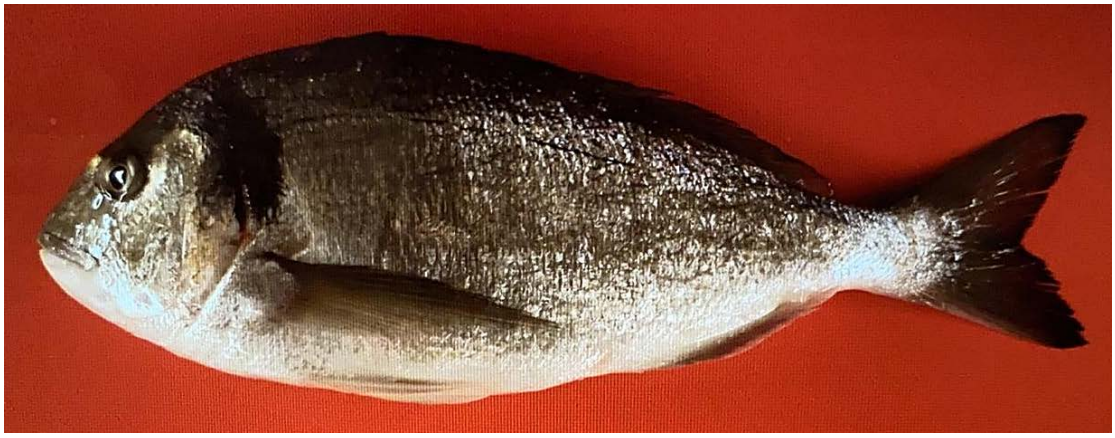


Figure 1. External appearance of a gilthead seabream adult (lateral view)

It is geographically distributed in subtropical coastal areas of the Mediterranean, Black Sea and eastern Atlantic, from the coasts of the United Kingdom and the North Sea to the coasts of Senegal (Froese & Pauly, 2019) (Figure 2). Due to its eurihalin and eurythermic habits, gilthead seabream is found both in marine environments and in coastal lagoons and estuaries, in addition to having been introduced in other regions by aquaculture. It inhabits in shallow waters, in the algae of rocky areas and sandy surfaces, generally up to 50 m deep, but can reach 150 m. Born offshore, the juveniles migrate to shallow coastal waters during the spring and return to deeper waters in late fall, when the adults reproduce (Moretti et al., 1999).

This species is a protandric hermaphrodite, reaching sexual maturity as a male at 2 years of age and changing to a female at 2–3 years, conditioned by social and hormonal factors. Females spawn in intermittent batches, being able to lay between 20,000 and 80,000 eggs each day for a period of up to 4 months. Their reproductive season begins when the days are short and the hours of light decrease, generally from October to December in the western Mediterranean (Arias, 1980) and between November and February in the eastern Mediterranean (Ben-Tuvia, 1979). The eggs are planktonic, spherical and transparent, approximately 1 mm in diameter and with one or more lipid droplets inside (Moretti et al., 1999).

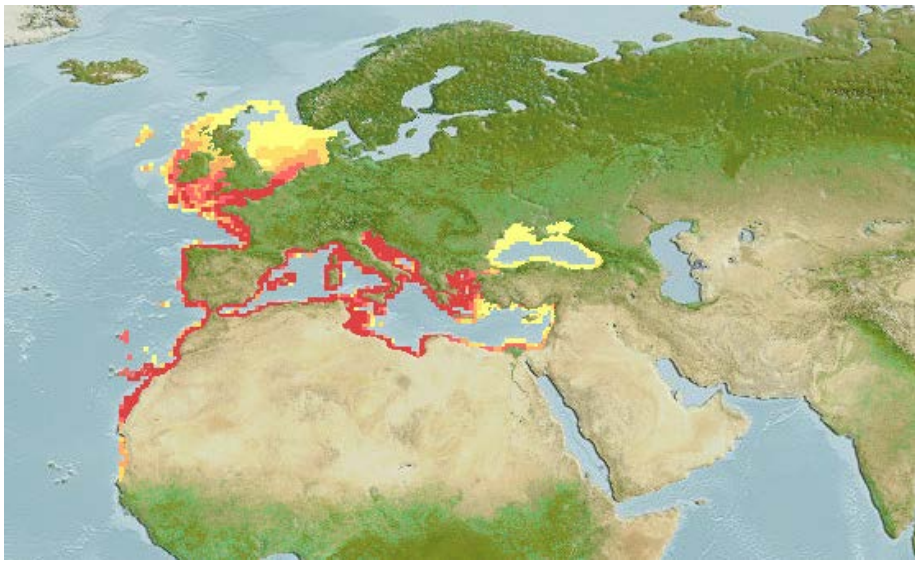


Figure 2. Geographical distribution of Sparus aurata in 2019. The range of colours shows the adequacy of the habitat, which can be interpreted as the relative probability of occurrence (Image taken from www.aquamaps.org).

1.3 Industrial impact of morphological deformities in fish aquaculture

At present, it is common to find individuals with skeletal deformities during the industrial culture process of the main aquaculture species with commercial importance. These deformities have great economic consequences for aquaculture companies by affecting their market value and increasing production costs. Previous studies have estimated that the economic loss due to skeletal deformities in Europe is at least 50 M€ per year (Haga et al., 2011). In addition, although the incidence of skeletal deformities is species dependent, incidences of between 50-100 % of production have been reported in species such as Japanese snapper (*Pagrus major*), seabream (*Sparus aurata*), sea bass (*Dicentrarchus labrax*), Atlantic cod (*Gadus morhua*), grouper (*Epinephelus*), Japanese sole (*Paralichthys olivaceus*), amberjack (*Seriola*), Senegalese sole (*Solea senegalensis*), Atlantic salmon (*Salmo salar*) and rainbow trout (*Oncorhynchus mykiss*) (Haga et al ., 2011).

Economic losses produced for deformities occur at two levels, hatchery and ongrowing. The appearance of deformed animals in hatcheries reduces the survival and growth rate of deformed larvae (Fernández et al., 2008). This forces them to introduce additional costs in healthy animals by having to oversize their production and/or carry out manual sorting procedures to eliminate deformed animals, since ongrowing companies do not accept batches with a prevalence of deformed animals greater than 5 % (Afonso & Roo, 2007), when some deformed fish could remain undetected by external observation. At the level of juvenile to growth in ongrowing companies, deformed fish must be discarded or sold below the market prices, which implies a supplementary production cost in addition to that produced by their lower production yields (such as swimming capacity, conversion rate, growth rate, higher mortality rate and susceptibility to stress, pathogens and bacteria) (Andrades et al., 1996; Boglione et al., 2001; Fernández et al., 2008; Karahan et al., 2013). On the other hand, the presence of deformities generates a negative impact on animal welfare and on the consumer's perception of aquaculture products (Komen et al., 2002). Therefore, the presence of deformities is the second character in economic importance for the industrial production of gilthead seabream, due to the reduction in the marketing value of the product throughout the value chain, aggravated by its sale as whole fish (Bardon et al., 2009;

Georgakopoulou et al., 2010; Boglione & Costa, 2011; Boglione et al., 2013a, 2013b). For this reasons, morphological quality traits have acquired the same importance as growth traits in European genetic breeding programs, in terms of frequency (Janssen et al., 2017).

1.4 Types of morphological deformities and their importance in aquaculture

Morphological abnormalities are deviations of the fish morphology from a standard quality. The prevalence of these anomalies in industrial batches is associated with the intensification of the cultivation of fish species with commercial interest and affect their integral quality (Divanach et al., 1996; Afonso & Roo, 2007). The main anomalies that have industrial importance are those that affect the shape of the body of the fish, since the shape is the first visual criterion for evaluating the quality of a species, especially in those that are marketed as whole fish (Divanach et al., 1996; Afonso & Roo, 2007). In general, in fish farms, animals that are recognized as having an abnormal body shape have vertebral deformities (Gjedrem & Baranski, 2009). Spinal cord deformities, with neurocranial and appendicular skeleton deformities, are classified within the group of skeletal deformities and they are the main cause of variation in fish shape (Afonso & Roo, 2007). In addition to skeletal deformities, morphological anomalies can be classified as scale malformations and pigmentation problems, but these anomalies generally have less economic impact, particularly in the studied species.

In gilthead seabream, the most economically significant skeletal deformities include severe abnormalities of the opercular complex, neurocranium, vertebral column and appendicular skeleton abnormalities, given that fish are sold mainly as whole fish and customers rarely accept fish showing malformations (Koumoundouros et al., 1997b; Boglione et al., 2001; Lee-Montero et al., 2015; Boursiaki et al., 2019).

1.5 Etiology of morphological deformities in fish

The origin and factors that cause most of the deformities are not precisely known and in most species, the deformities appear during embryological, larval or early adult development (Boglione et al. 2009; Loizides et al. 2014; Azevedo, et al. 2017; Fragkoulis et al. 2017; Thuong et al. 2017). In general, the increased prevalence of deformities is associated with unfavorable abiotic conditions, such as excessive handling, the presence of xenobiotics or nutritional

imbalances, and biotic conditions such as parasites. However, although the deformities have a strong environmental component, there are several studies that suggest an interesting genetic component to be exploited in breeding programs.

1.5.1 Non-genetic factors

There are numerous studies that associate unfavorable management or culture conditions with a higher incidence of deformed animals in fish species. Thus, high densities of culture in gilthead seabream have been related to the appearance of a higher proportion of caudal anomalies and vertebral deformities in general at an industrial level (Koumoundouros et al., 1997b; Prestinicola et al., 2013). Chaptain and Ounais-Guschemann (1990) reported the occurrence of 90 % prevalence of lordosis in gilthead seabream larvae caused by non-inflation of the swim bladder due to oil layers on the water surface or high hydrodynamic conditions (Chatain & Ounais-Guschemann, 1990; Chatain, 1994). Although lordosis can be associated with swim bladder inflation problems, it can also develop with fully functional swim bladders (Andrades, 1996). The appearance of deformities due to high water currents has also been described in other species close to gilthead seabream such as sea bass and red seabream (Chatain, 1994; Divanach et al., 1997; Kihara et al., 2002). Temperature is another factor that in gilthead seabream can also increase the prevalence of deformities, as well as the mortality rate during early larval stages (Polo et al., 1991; Georgakopoulou et al., 2010). This also occurs in other sparids such as the common pandora (Sfakianakis et al., 2004) and species such as sea bass (Koumoundouros et al., 2001; Abdel et al., 2004; Sfakianakis et al., 2006; Georgakopoulou et al., 2007).

Regarding xenobiotics, an increased incidence of deformities has been reported in fish exposed to pesticides for a long time (Chun et al., 1981; Whittle et al. 1992). Similarly, there are reports of a higher incidence of deformities in fish from water contaminated with agricultural, industrial and urban effluents (Slooff, 1982; Browder et al., 1993), as well as with selenium residues (Lemly, 1993).

The prevalence of deformities in fish is also associated with nutritional imbalances during larval development (Boglione et al., 2013b; Cahu et al., 2003b). Thus, an excess of vitamin A is the cause of a higher proportion of deformities in gilthead seabream (Fernández et al., 2008) and other fish species (Haga et al., 2011). Dominguez et al., (2022) suggested that an appropriate K3 levels in gilthead seabream fingerlings reduce the prevalence of skeletal disorders. Cahu et al. (2003a) reported that a diet with inadequate phospholipid values results in a higher prevalence of mouth and spinal deformities in sea bass. Similarly, using high levels of docosahexaenoic acid in the diet was able to reduce the incidence of skeletal deformities in red porgy by 50 % (Izquierdo et al., 2010).

As biotic effects that can trigger the appearance of deformities in fish, Lom et al. (1991) have described that parasitization by the myxosporea family can cause skeletal anomalies due to ectopic agglomeration along the connective tissues of the central canal that surround the spinal nerve. Similarly, jaw anomalies have been associated with mouth parasites in halibut (Morrison & MacDonald, 1995).

1.5.2 Genetic factors

Skeletal deformities have also been associated with a genetic basis. Thus, it has been reported that inbreeding depression, which is a consequence of the level of consanguinity, affects the presence of malformations. In rainbow trout, Austal and Kittelsen (1971) found a prevalence of deformed fingerlings 38 times higher in families of full siblings. On the other hand, Kincaid (1983) detected a prevalence of deformities of 37.6 % compared to families of unrelated individuals.

In gilthead seabream, Andrades et al. (1996) proposed a genetic origin to explain the prevalence of lordosis in a cultured population. Afonso et al. (2000) reported for the first time a statistical association of an idiopathic disease with a specific family from a genetic breeding program. This disease is a triple skeletal malformation (Lordosis, Scoliosis and Kyphosis (LSK)), repeated from head to tail. Astorga et al. (2004) proposed the segregation of a major gene by considering all deformities as a single trait (presence-absence of any deformity), after obtaining very high heritability estimates using an animal model and Bayesian analysis (0.78 and 0.85,

respectively). Afterwards, Lee-Montero et al. (2015) estimated medium-low heritabilities (0.07-0.26) for the presence-absence of any deformity from fingerlings to adults, but with high genetic correlations between them (0.83-0.99), using an animal model. At the same time, Negrín-Báez et al. (2015b) reported a significantly higher number of offspring with operculum deformity, lordosis, and spinal fusion from directed crosses of parents with the same deformities with heritability estimates between 0.34-0.46 for these three deformities. García-Celdrán et al. (2015) confirmed this high heritability estimates for spinal deformities (0.38-0.56) and operculum (0.43-0.46) through Bayesian analysis at 163 and 690 days post hatching (dph), respectively. Using an animal model at 163 dph, heritability estimates of 0.53 and 0.37 for lordosis and lack of operculum, respectively (García-Celdrán et al., 2016). In this last study, a medium heritability estimates was also determined for the non-inflation of the swim bladder, which also showed a positive genetic correlation with lordosis (0.48) (García-Celdrán et al., 2016). Also in larvae of 39 dph, was estimated a medium heritability for fusion of maxillae with premaxillae (0.22) (Fragkoulis et al., 2018) and for abnormalities in the caudal fin (0.15-0.27) (Fragkoulis et al., 2020). More recently in adults, Fragkoulis et al., (2021) used gilthead seabream juveniles from 155 dph to study body shape variations until 589 dph, determining a substantial heritability ranged between 0.17-0.51 and a genetic correlation between ages ranged between 0.22-0.76 (Fragkoulis et al., 2021).

1.6 **Evolution of breeding programs in gilthead seabream**

The gilthead seabream culture started in the 1970s with the fattening of larvae captured from the environment (Janssen et al., 2017), accelerating the development of this industry in 1990 thanks to the improvement of reproduction techniques (Divanach & Kentouri, 2000). The first genetic improvement trials in the species arrived in the mid-1990s (Knibb et al., 1998, 1997) and the first industrial-level genetic breeding program began in the early 2000s (Thorland et al., 2007; Chavanne et al., 2016; Janssen et al., 2017). The main objective of the breeding programs is to increase the profitability and sustainability of the productive companies, while maintaining the genetic variability of the culture stocks (Martinez, 2007). Traditionally, the main trait of industrial interest in breeding programs has been growth, but over time other traits have gained industrial

importance and have been added, such as deformity or resistance to pathogens (Chavanne et al. 2016; Janssen et al. 2017; Vandeputte et al., 2019; Boudry et al., 2021).

Currently, there are companies that carry out breeding programs for gilthead seabream in France, Greece, Italy, Spain, Croatia, Israel and Turkey (Boudry et al., 2021). Most of them use mass selection as a selection method, although there are companies that use family selection methods, including best linear unbiased prediction methodology (BLUP) through molecular pedigree or separate family growth (Boudry et al., 2021). The use of molecular markers, such as microsatellites and single nucleotide polymorphisms (SNPs), improves the efficiency of selection methods, thanks to the linkage disequilibrium association of these markers, located in specific regions of the genome, with the phenotype of quantitative traits of interest, thus determining a quantitative trait locus (QTL) (Pérez-Enciso y Toro, 2007; Yue, 2014). Traditional QTL detection approaches allow finding those QTLs with a medium-high effect on the character of interest, but not those with a low effect (Yue et al., 2014). This allows marker-assisted selection (MAS) to be carried out, improving selection accuracy and genetic gain, especially in characters that are difficult to measure or have low heritability, such as deformity (Meuwissen, 2003; Martinez, 2007; Yue et al., 2014; Boudry et al., 2021). At present, the improvement of massive sequencing techniques and their lower cost has allowed the availability of a large number of molecular markers, as well as the genome and transcriptome of many of the culture species. This has permitted the development of techniques such as RNAseq or genomic selection.

RNAseq is a technique that allows the total RNA of a sample to be sequenced, mapped, annotated and quantified at once (Saroglia & Liu, 2012; Qian et al., 2014). In fish aquaculture, it has been used to assemble new transcriptomes, annotate genomes, detect SNPs and perform expression analysis (Qian et al., 2014). RNAseq is a cheaper technology than whole genome sequencing (WGS) and although it is limited to the coding regions of the genome, it can be used to reliably detect genomic variants such as SNPs (Piskol et al., 2013). There are markers that are associated with genomic loci that regulate the expression levels of certain genes that could be of interest, these regions are known as eQTLs (Saroglia y Liu, 2012). The analysis of eQTLs of complex traits allows to study the molecular processes that underlie the pathobiology (Ogura et al., 2015).

Genomic selection (GS) was proposed by Meuwissen et al. in 2001 and dictates that with a sufficiently high number of molecular markers spread throughout the genome (genome-wide (GW)), most quantitative trait loci would be in strong linkage disequilibrium with at least one marker, which would allow adding the effects of each of them to predict the behavior of this trait (Meuwissen et al., 2001). In this way, it can be summarized that the BLUP selection calculates the genetic potential of the breeders (estimated breeding value (EBV)) through the information of their relatives, the MAS selection uses molecular markers associated with QTLs with a medium-high effect on the trait of interest and the selection GS predicts the genetic potential through a model made from the sum of effects of the GW markers (Figure 3).

Although MAS and GS selection have been shown to be potentially effective in controlled experiments, in an industrial aquaculture context they are not widespread selection methods, probably due to the lack of studies on the balance of extra costs and extra benefits that implementing these techniques (Yue et al., 2014; Boudry et al., 2021).

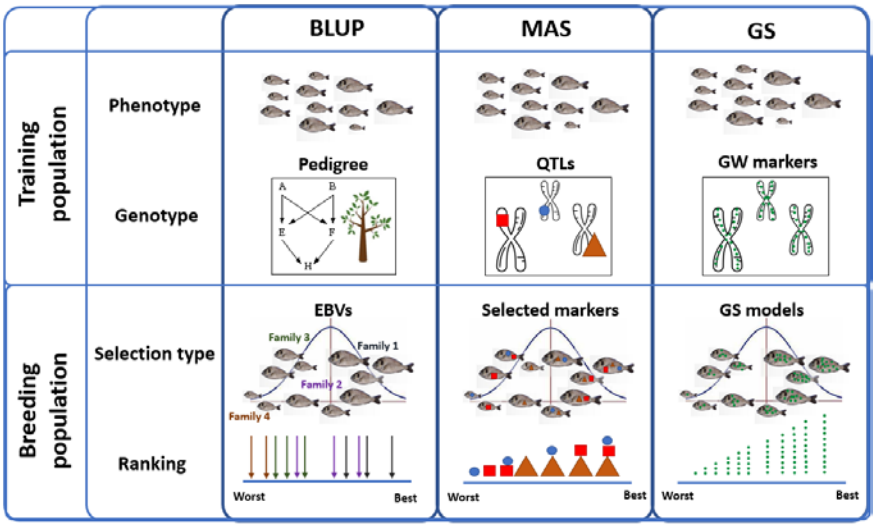


Figure 3. Scheme of the selection systems (BLUP, MAS and GS). Scheme is based on the phenotypic and genotypic information in the training population and the methodology used to estimate the genetic value and perform the ranking in the breeding population

1.7 **Molecular markers and genomic resources available in gilthead seabream**

The industrial importance of gilthead seabream has made it a species on which numerous studies are focused, which has allowed numerous genomic resources and molecular markers for this species to exist over time (Manchado et al., 2016). Some of these genomic resources were

carried out within the framework of the BRIDGEMAP project (2000–04), financed with European Community funds. In 2005, Sarropoulou et al. constructed the first cDNA microarray containing 10,176 clones from a cDNA library from a mixture of embryonic and larval stages.

In 2006, Senger et al. constructed a preliminary radiation hybrid map. The same year, it was reported by Franch et al. (2006) the first linkage map, with 26 linkage groups, made from 204 microsatellites. On the other hand, numerous studies focused their efforts on obtaining expressed sequence tags (ESTs) from different tissues and moments of development in order to study not only markers in the genome, but also discover which genes are expressed (Manchado et al., 2016). Already in the 2010s, different authors reported additional transcriptomes for samples of skeletal muscle, skeletal tissues, intestine, blood and kidney from juvenile seabream, which resulted in a global transcriptome of 113,927 contigs, of which approximately 50 % were annotated (García de la Serrana et al., 2012; Calduch-Giner et al., 2013; Vieira et al., 2013; Manchado et al., 2016). At the end of this same decade, genomes of this species began to be available, complete enough to be assembled into chromosomes. Dray et al. (2016) reported the mitochondrial genome and later in 2018, within the framework of the AquaTrace (2012-2016) and FISHBOOST (2014-2019) projects, Pauletto et al. achieved the complete genome. In 2019, Sánchez et al., within the framework of the AQUAEXEL2020 project (2015-2020), also reported another complete assembly of the gilthead seabream genome.

All this genomic information makes it very close to obtaining a stable transcriptome and genome that can be used to perform robust RNAseq analyzes and a standardized annotation that allows comparison with other species and carry out a correct functional analysis. Although RNAseq does not require a priori knowledge of the species genome, it can clearly benefit from the abundant information contained in genomes by reducing their complexity in a practical way (Manchado et al., 2016).

It should be noted that the succession of molecular markers reported in this species and the advancement of molecular and genomic techniques led to the beginning of QTL studies for different characters in seabream, including morphological traits and skeletal deformities. Boulton et al. (2011) found a significant QTL that explains 13-23 % of the phenotypic variance of their

population for morphometric traits. Between 2015 and 2016 Negrín-Báez et al. carried out different experiments in which they reported different QTLs associated with skeletal deformities. In 2015, they found three significant QTL that explained 1.6-11.4 % of the variance, one for fusion of vertebrae (QTLFV3), one for lordosis (QTLOR1) and another for mandibular deformity (QTLJW1), the latter located very close to 2 of the microsatellites used (CId-26-H, CId-03-F) (Negrín-Báez et al., 2015a). In 2016, they performed a QTL analysis for LSK deformity using linear regression and linear mixed model, detecting four significant QTLs (QTLSK3, QTLSK6, QTLSK12, QTLSK14) at the genome level, which showed a large effect in both methods (>35 %), jointly explaining 65-94 % of the variance and also finding 2 microsatellite markers (DId-03-T and Bt-14-F) strongly linked to this deformity (Negrín-Báez et al., 2016b). That same year, using linear regression methodology, they detected 2 significant QTLs (QTLOP1 and QTLOP2) for operculum deformity, with an effect of around 27 % (Negrín-Báez et al., 2016a).

The study in primary cultures of mesenchymal stem cells derived from gilthead seabream bone cells has made it possible to characterize the gene networks that regulate the osteogenesis process under different temperature conditions (Riera-Heredia et al., 2018) and treated with fatty acids (EPA, DHA, LA and ALA) (Riera-Heredia et al., 2019a). On the other hand, by comparing the tissues of deformed fish and normal fish, it was determined that the expression pattern of genes related to the maturation and mineralization of the bone extracellular matrix (*col1a1*, *op*, *ocn*, *mgp* and *tnap*) as well as the involved in bone resorption (*ctsk* and *mmp9*), seem to be key to distinguish between normal fish and fish with spinal deformities (lordosis and LSK), without finding a differential pattern in jaw and operculum deformities (Riera-Heredia et al., 2019b).

1.8 Objectives

The course of generations of aquaculture fish production has achieved the domestication of various farmed species by studying their biology and improving the methodologies for their culture and reproduction. Previously, the strategy followed in industrial production systems was culture intensification and the selection through growth traits to increase production and satisfy market demand. But over time, the increase in the competitiveness of the sector and the maturity of the market has stabilized or even decreased the price of these products, forcing the industrial

sector to reduce production costs or achieve added value to its products to maintain profitability. The seabream sector is a good example of this event. In this species, the improvement of the quality of the final product and the reduction of production costs have gained industrial interest, so reducing the rate of deformed animals cultivated is of vital importance.

Against this background, the objective of this study is to study the genetic structure underlying in the prevalence of skeletal deformities in aquaculture gilthead seabreams and to describe genetic markers that can be used to characterize the genetic potential of the breeder stocks in order to reduce the incidence of deformities with genetic origin in production.

To achieve this general objective it is necessary to establish these series of specific objectives:

-Analyze the EBV for the prevalence of deformity (EBVdef) in a large number of animals coming from industrial lots that guarantees a sufficient number of recombinants to elucidate the genetic mechanism through the comparison between animals with different EBVdef.

-Determine the existence of a differential expression pattern in target tissues from breeders of families with distant EBVdef (with high and positive EBVdef or genetically deformed (gD) and with high and negative EBVdef or genetically normal (gN)).

-Establish batches of phenotypically normal breeders with distant EBVdef (gD and gN) in order to analyze their spawning quality and to carry out the larval culture of their descendants to determine the existence of differences in the deformity prevalence and the moment of development in which they originate.

-To determine the existence of a differential expression pattern and sequence differences between the genes of both groups (gD and gN) at target stage of larvae development.

Chapter 2

Materials and methods

2.1 General design

The general design of the trials conducted in this thesis included the genetic evaluation for the presence of skeletal deformities of a commercial batch of gilthead seabream with commercial size of the 3rd generation of the National breeding programme PROGENSA®-III (Lee-Montero et al., 2015; León-Bernabeu et al., 2021), by characterizing the presence of deformities in the phenotype visually and connecting information between animals through genealogical reconstruction using molecular tools. After the evaluation, the selection of breeders was made to carry out 3 different experiments, all the selected breeders were phenotypically normal, but with divergent levels of EBVdef. Experiment 1 was based on the selection of 4 full sibs families with divergent EBVdef (2 families of 3 individuals with high and positive EBVdef and 2 families of 3 individuals with high and negative EBVdef). Then, differential expression pattern of total RNA coming from three target tissues (vertebra, white muscle and brain) was analysed two determine differences between gD and gN animals. Experiment 2 and experiment 3 were based on the selection of two different stocks per duplicate of breeders according to their divergent EBVdef. In experiment 2, spawning quality was monitored daily for a complete spawning season to determine if the deformity genetic background affect the quantity and quality of the spawning. Finally in experiment 3, the offspring of these breeder batches were produced per triplicate and separately following common commercial protocols up to 73 dph, taking samples daily between 0-35 dph and every other day between 35-73 dph for osteology and gene expression analysis. Figure 4 shows an overview of the experimental design carried out in this work.

The study was conducted according to the European Union Directive (2010/63/EU) on the protection of animals for scientific purposes and it was approved by the Bioethical Committee of the University of Las Palmas de Gran Canaria.

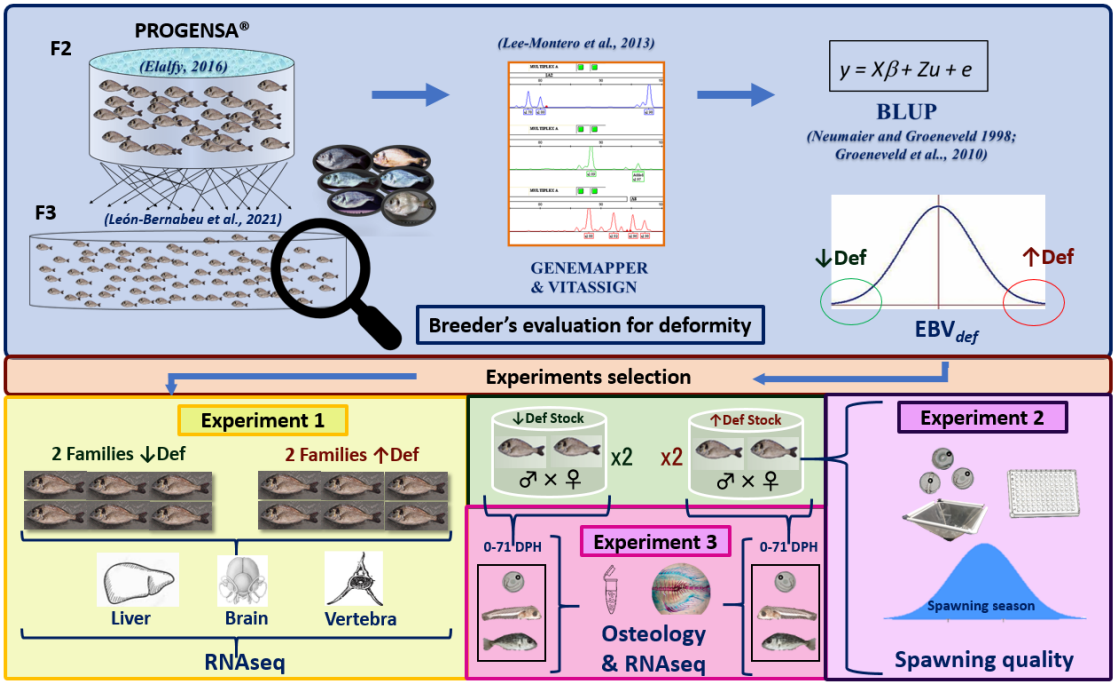


Figure 4.Overview of the experimental design followed in this work. In order the steps are the deformity breeder's evaluation and the animal selection for the 3 experiments carried out.

2.2 Broodstock maintenance

The breeders selected for this three experiments came entirely from the 3rd generation of the National breeding program PROGENSA®-III, specifically from the backup populations of the Canary Islands (León-Bernabeu et al., 2021). The breeders come from a base population with an estimated medium-low heritability for any type of deformity (0.07-0.26) and a high genetic correlation between ages (0.83-0.99) (Lee-Montero et al., 2015).

All breeders were identified by an intraperitoneally passive integrated transponder (EID Iberica SA-TROVAN, Madrid, Spain) tagged at fingerling stage, when fin samples were also collected to stablish parental assignment (Navarro et al., 2006; Lee-Montero et al., 2013).

After genetic evaluation, selected breeders were moved to the FCPCT-ULPGC facilities, where were maintained in a 40 m^3 circular tank with 600 % water exchange rate, nature photoperiod and fed a commercial diet (Skretting, Spain). For experiment 1, selected breeders were separated and acclimated in a different tank for 1 month before sampling. The sampling date was chosen outside the spawning season so that the animals did not have developed gonads to minimize interference based on gender and maturity stage.

For experiment 2, breeders' gender was determined by a gentle abdominal massage to identify males that were spermiating and by cannulation to identify mature females, previously to the spawning season. Then, duplicate batches of gD and gN breeders were established into different experimental 1,000 L tanks and their spawning quality checked every day for the whole spawning season. For experiment 3, the same broodstocks from experiment 2 were used for rearing larvae. During the experiment, broodstock were fed twice daily, cleaned once a week and kept under natural photoperiod and thermoperiod.

2.3 Breeders genetic evaluation

To carry out the evaluation process using BLUP, it is necessary to measure the phenotypic traits to be quantified (yield matrix) and obtain genetic information (genetic matrix), in this case by reconstructing the pedigree using molecular markers and exclusion methodology.

For this purpose, fish deformities were visually assessed at commercial size in accordance with AquaExcel-ATOL (ATOL: 0000087) (AQUAEXCEL, 2013). Caudal fin samples were digested with proteinase K (Qiagen), then DNA was extracted by using DNeasy 96 protocol® in BioSprint96® robots (Qiagen) following the manufacturer's protocol. DNA quantity and quality was verified by measurement in NANODROP-8000 spectrophotometer (ThermoFisher) and by 1 % agarose gel electrophoresis. Robot TECAN FreedomEvo® was used to normalize the DNA samples to 30 ng/µl before genotyping by BioRad PCR with the 11 *loci* microsatellite supermultiplex SMsa1 (Lee-Montero et al., 2013) and followed by sequencing analysis using the MACROGEN fragments service. Electropherograms were read using *GeneMapper* software (v5.0) (Applied Biosystems) and parental assignments were carried out with *Vitassign* software (Vandeputte et al., 2006).

EBVdef were estimated for both data matrix (phenotypic and genotypic information) with BLUP methodology by using the restricted maximum likelihood method with *VCE 6.0* software (Neumaier & Groeneveld, 1998; Groeneveld et al., 2010) and a complementary program to help for managing input data and automatize processes

(*VCE_Executer, v3.0*) (León-Bernabeu et al., 2021). In all experiments EBVdef were used to select the breeders and phenotypically deformed fish were discarded to simulate the real situation at the hatchery companies.

2.4 Spawning quality assessment

Maturation and spawning occurred spontaneously under natural conditions in December 2019. During the spawning season, eggs were collected daily (20 h after release and fertilization) from each tank overflow in semi-submerged 500-µm mesh net egg collectors placed in a nearby incubation tank. The *oocyte yield* (ATOL:0001723), *fertilization rate* (ATOL:0001775), *viability rate* (ATOL:0001531), *hatching rate* (ATOL:0001531), *larvae survival rate* at yolk sac resorption (ATOL:0001531) were measure daily in accordance with AquaExcel-ATOL (Kjørsvik et al., 1990; Fernández-Palacios et al., 1995, 1997; AQUAEXCEL, 2013).

Briefly, eggs from each tank were collected in 5 L plastic beakers. Then, 3 randomized samples of 10 mL were taken using strong aeration to guarantee a correct homogenization. Egg samples were placed in Bogorov counting chamber, observed and counted under a binocular stereo microscope to estimate volumetrically the *oocyte yield*, *fertilization rate* and *viability rate*. The good egg viability was considered as transparent eggs, perfectly spherical with clear, symmetrical early cleavages.

In addition, to calculate the hatching rate and survival rate, two 96-well microplates (VWR, West Chester, USA) were used to hold 192 floating eggs (1 per well) with 200 µl filtered seawater and incubated at 19 ºC. After 24 and 96 h, the number of hatched eggs and the number of alive larvae was counted, respectively. The total number of hatched eggs and alive larvae in the spawn was determined based on the hatching rate and larval survival rate of these samples.

2.5 Larval rearing

Larvae of experiment 3 were reared per triplicate under controlled conditions using the protocols established in EcoAqua Institute facilities, after checking a good spawning quality (enough oocyte yield, a good fertilization and viability rates and a period with high hatching and larvae survival rates) as detailed below.

Following the methodology of the section 2.4, eggs were collected in 5 L plastic beakers and a good spawning quality was verified. Then, eggs floating phase was separated and counted, and dead eggs were discarded to avoid bacterial blooms associated with an excess of organic matter. Viable eggs were transferred to 500 L larval rearing tanks with an initial density of 100 eggs/L, calculated volumetrically after counts.

The larval rearing tanks had the water inlet from the bottom and the outlet at the top of the tank wall connected to a filter, with a mesh size adaptable to larvae size. They also have an aeration system in the center and an air blower skimmer. This system facilitates the maintenance of eggs and larvae in suspension in the early life stages and keeps the surface free of lipids, a prerequisite for correct inflation of the swim bladder (Chatain & Ounais-Guschemann, 1990). Tanks were matured with phytoplankton strain (*Nannochloropsis sp.*) 24 h before the eggs transfer to guarantee good water quality and prevent the growth of unwanted microorganisms. Temperatures and dissolved oxygen concentrations were measured daily, ranged between 19-22 °C and 6–12 ppm, respectively.

The water renewal was adjusted from 10 to 40 % of the volume per h until the establishment of microdiet feeding (~46 dph). The homogenization of the larval distribution into the tanks, and the reoxygenation, were achieved by an external air lift recirculation pump creating a gentle vertical convection current from the tank bottom. Until hatching, high aeration is maintained to keep a good eggs dispersion, then is reduced enough to keep the larvae suspended in passive movement until 3 dph (mouth opening). At this time, aeration was adjusted at a low level to avoid microbubbles formation and larvae gas hypersaturation derived problems. The ‘green water’ technique was applied during the period 3–28 dph by using the

same phytoplankton strain (250 ± 100 × 10^3 cells/mL). This technique improves the vision of the larvae to hunt rotifers, keeps them enriched more time and improve water quality. Regarding feeding, larvae were fed twice a day with rotifers (*Brachionus plicatilis*) enriched with ORI-GREEN (Skretting). Rotifer density was progressively increased from 5 to 10 rotifers per mL. Artemia nauplii (EG, INVE, Belgium) were offered twice a day to larvae from 17 to 22 dph, in increasing density from 0.5 to 2 nauplii per mL, and 2 days enriched-metanauplii from 20 to 40 dph(1 to 5 metanauplii per mL). From 36 dph to the end of the experiment (71 dph), larvae were progressively weaned onto dry feed, with Gemma Micro (size range: 75 to 500 μm; Skretting, Spain) with manual feeding each hour from 08:00 to 16:00 and automatic feeding 3 times from 17:00 until 19:00, maintaining one feeding with enriched-metanauplii at 12:00. During this period passive aeration induced by water flow was enough and not grading or scaling was performed along the whole experiment.

2.6 Sampling procedures

2.6.1 Breeders

Selected breeders for experiment 1 were fasted and anesthetized with clove oil (10 ppm clove oil:methanol, 1:1 volume in sea water). Then, each fish was slaughtered with an excess of clove oil, certifying the death by exsanguination with a cut in the gills, to collect white muscle, brain and vertebra samples for the RNAseq analysis. Surfaces and sampling material were cleaned with propano AF and rinsed with water treated with diethylpyrocarbonate 0.01 % (DEPC) before sampling and between samples. Samples of all tissues were taken from the same region for all animals. Until RNA isolation, Vertebra samples were quickly frozen with liquid nitrogen and kept at −80 ◦C and white muscles and brain samples were conserved in RNAlater (Sigma-Aldrich) in a ratio 1:5 (sample:solvent) and trimming the tissue sample to be less than 0.5 cm in at least one dimension, according to the manufacturer's recommendations. Then, were kept overnight at 4 °C and stored at − 80 °C until RNA isolation.

2.6.2 Larvae

Larvae samples were taken in experiment 3. The sampling was carried out daily from the egg phase until day 35 dph and alternate days from 35 to 71 dph for osteology and RNAseq analysis. For osteological analysis in each sample point (day and tank) a minimum of 10 larvae were taken, and for RNAseq analysis a minimum of 10 larvae and 100 mg of sample.

Sampling is a fundamental step to guarantee a representative sample of our population. During the "green water" phase, the larvae were collected randomized by taking water samples at different points in the tank, since the larvae are homogeneously distributed. After this period, larvae were collected with the help of a net, quickly, during the feeding time, also taken the samples in different points of the tank randomized and taken all the size represented in the tank to obtain a representative sample. If necessary, the larvae were grouped using a mesh to reduce the volume of water.

Slaughtering was carried out immediately after taking the samples, quickly and following the current legislation to reduce, as much as possible, animal suffering and possible alterations derived from stress. Samples were sacrificed by thermal shock by immersing them in seawater with ice at 0 ºC and keeping there until 5 min after the last larvae movement with the help of a plastic strainer.

Finally, larvae were washed with DEPC-treated water and the excess water was dried on filter paper to avoid altering the solvents. In case of osteology, samples were conserved with 4 % formalin buffered with phosphate at 7.2 pH in a 1:10 ratio (sample:solvent) and stored in darkness at room temperature until staining and osteological analysis. In case of RNAseq, samples were conserved in RNAlater (Sigma-Aldrich) in a ratio 1:5 (sample:solvent) and guaranteeing a size less than 0.5 cm in at least one dimension, according to the manufacturer's recommendations. Then, samples were kept overnight at 4 °C, stored at − 80 °C and sent with dry ice to the Institute of Aquaculture Torre de la Sal (IATS-CSIC) for RNAseq analysis.

2.7 Osteology analysis

For the examination of skeletal deformities, a mean of 22 formalin fixed specimens per condition (high and low EBVdef) and per sampling day, from day 35 to day 73, were stained for bone and cartilage using a double-staining protocol with some modifications (Potthoff, 1984; Taylor & Van Dyke, 1985). In brief, specimens were washed in distilled water to eliminate formalin, followed by a bleaching step in peroxide (3 %) and KOH (0.5 %). The bleaching process removes the pigments and contributes to degrading the muscle fibers, allowing the dyes to penetrate to the bone. Then, specimens were dehydrated through a graded series of ethanol (25–95 %) and stained with Alcian blue 8GX (0·02 % in 70 % alcohol and 30 % glacial acetic acid). This dye has a high affinity for cartilage, but it is a step that must be carried out with great care not to overstain, since the acidic pH can demineralize the bone. Therefore, the next step was the neutralization of the samples using a 1 % KOH aqueous solution. In this step the pH is checked regularly and the solution is changed in case it becomes acidic, keeping the samples inside until the elements of the skeleton are clearly visible. Later, series of trypsin digestion (35 % saturated sodium borate, 65 % distilled water and trypsin powder) and staining with Alizarin red S (stock solution: 1 % Alizarin red in 1 % KOH). Trypsin digests the muscle fibers and is required by large larvae to facilitate dye penetration into the bone and Alizarin red stains the ossified bone matrix red (calcium). Finally, samples were cleared by using trypsin and KOH (1 %) and preserved in glycerin with thymol. Staining time was variable and depended on the larvae size.

In accordance with the degree of mineralization, samples from 39 dph to 73 dph were used for the examination of severe skeletal deformities. In total a number of 808 specimens were evaluated, 453 of low EBVdef and 355 of high EBVdef. The analysis of deformities was based on the normal osteological phenotype of the species (Divanach et al., 1996; Koumoundouros et al., 1997a; Faustino & Power, 1998; Fragkoulis et al., 2018) and on the existing descriptions of skeletal deformities in marine fish larvae (Divanach et al., 1996; Boglione et al., 2001; Boglione & Costa, 2011; Boglione et al., 2013a, 2013b).

Larvae samples were analysed under a light microscope (CX41, Olympus, Tokyo, Japan), each larvae was placed with the head to the left and the tail to the right, as the dorsal part upwards and the ventral part downwards. Each section was individually analyzed (Skull, cranial and pre-haemal vertebrae, haemal and caudal, the pectoral, dorsal, anal and caudal fins).

2.8 RNA sequencing analysis

2.8.1 RNA isolation

For experiment 1, Samples (60–70 mg) were extracted using the RNeasy Mini Kit (Qiagen). The homogenization process was carried out with Tissue Lyzer-II (Qiagen, Hilden, Germany) in 1 mL TRI Reagent (Sigma-Aldrich) in case of brain and white muscle samples. A specific protocol was carried out on vertebra samples, they were homogenized using a homogenizer with metal case with 6 mL TRI Reagent (Sigma-Aldrich) and it was cooled on the outside with liquid nitrogen to avoid overheating the sample, then 1 mL were separated to continue with the RNA isolation process. Samples were centrifuged with chloroform for phase separation (12,000 RCF, 15 min, 4 °C). The upper aqueous phase containing RNA was mixed with 75 % ethanol and transferred into the RNeasy spin column, where total RNA bound to a membrane and RW1 and RPE buffers were used to wash away contaminants. A DNase digestion step was added according to the manufacturer's protocol with the *RNase-Free DNase Set* (Qiagen). Purified RNA was eluted with 50 μL of RNase-free water in white muscle samples and in 35 μL of RNase-free water in vertebra and brain samples. RNA quality was checked by 1.4 % agarose electrophoresis and quantified by using NANODROP-8000 spectrophotometer (ThermoFisher) before sending them to the Sequencing Unit from the University of Malaga, Spain. There, RNA was analyzed prior to sequencing to assure that it met specific qualitative and quantitative standards. RNA quality was analyzed by lab-on-a-chip electrophoresis using the bioanalyzer 2100 kit (Agilent, Amstelveen, Netherlands). Quantification of samples was performed using a Qubit fluorimeter (LifeTechnologies). RNA aliquots were stored at -80 ºC for later processing.

In case of experiment 3, each sample tank-day, with at least 10 larvae, was pooled and extracted using the *MagMAXTM-96 RNA isolation kit* (Life Technologies, Carlsbad, CA, USA). The quality and integrity of RNA were checked with an Agilent Bioanalyzer 2100 (Agilent, Amstelveen, Netherlands) to verified the RNA Integrity Number were higher than 7.

2.8.2 RNA sequencing

In experiment 1, Illumina RNAseq libraries were prepared using the *Illumina TruSeq Stranded mRNA Library Preparation Kit* (NEB, USA), according to the manufacturer's instructions. Once mRNA libraries were prepared, sequencing was performed using *NextSeq 550* (Illumina), 150 cycles of 2 x 75 pb mRNA sequencing was performed.

In samples of experiment 3, Illumina RNAseq libraries were prepared from using the *Illumina NEBNextUltraTMRNALibrary Prep Kit* (NEB, USA), according to the manufacturer's instructions. All RNAseq libraries were sequenced on an Illumina *NovaSeq 6000* sequencing system in a 2 × 150 pb read format, according to the manufacturer's protocol. Raw sequenced data were deposited in the Sequence Read Archive (SRA) of the National Center for Biotechnology Information (SRA ACCESSIONS: SRR13136984-7005) under the Bioproject accession number PRJNA680766 (BioSample accession numbers: SAMN16913187-208).

2.8.3 Bioinformatic analysis

The reads resulting from the massive sequencing procedures must pass a sequence quality filter to eliminate short sequences or sequences with a high error rate. In addition, reads was mapped to a reference transcriptome or a correctly annotated genome to count the number of times each transcript is represented. Finally, this information is used for differential expression analysis between groups, functional analysis and variant calling analysis (Figure 5).

In experiment 1, Picasso supercomputer of University of Malaga was used for all bioinformatic tests, implementations and executions. It consisted of an *OpenSUSE LEAP 42.3* with *Slurm* queue system and Infiniband network (54/40 Gbps) containing 216 nodes with Intel E5-2670 2.6 GHz cores for a total of 3,456 cores and 22 TB of RAM. The required software was already installed.

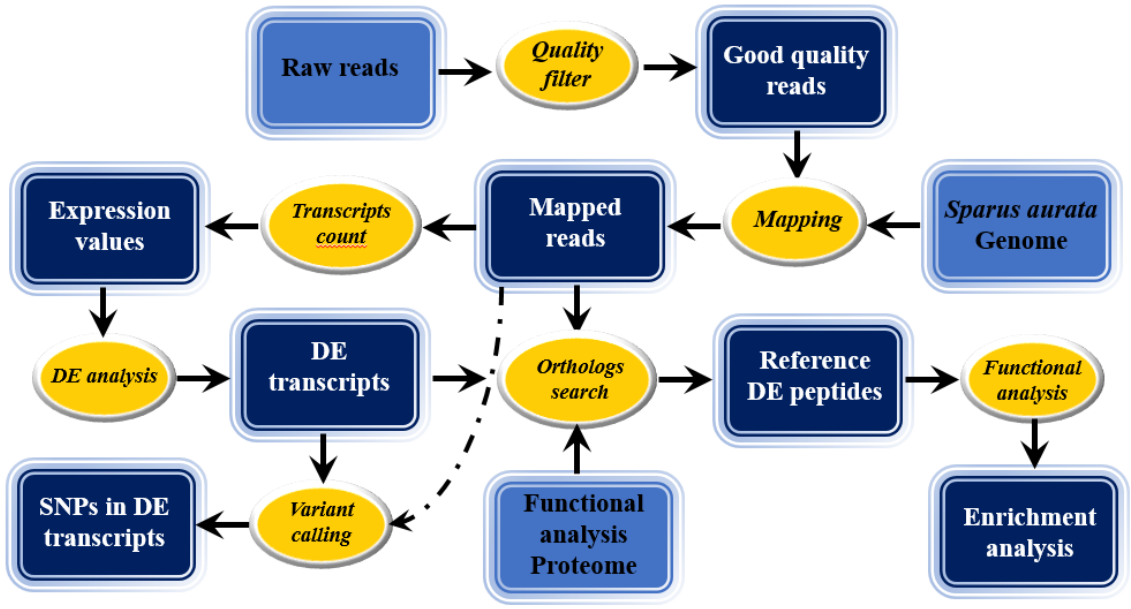


Figure 5. General workflow followed to carry out the processing and bioinformatics analysis of the output readings of the sequencer. Sequence filtering, its mapping in a reference genome, counting of expressed transcripts and later the different analyses (differential expression (DE), variant calling and functional enrichment analysis). In light blue is represented new data incomes, in yellow the process and in dark blue outputs or processed data.

2.8.3.1 Reads quality and genome mapping

In experiment 1, raw reads were filtered with *SeqTrim* using default parameters (Falgueras et al., 2010). Only good quality reads over 60 bp were mapped onto the transcriptome deduced from the haploid gilthead seabream genome assembly *fSpaAur1.1* (GCA_900880675.1), with *STAR* and using transcriptome mapping parameters (Dobin et al., 2013). Mapping was converted in expression values using *featureCounts* (Liao et al., 2014).

In experiment 3, quality analysis was performed with *FASTQC v0.11.7 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/)*, and libraries were filtered with *Prinseq* (Schmeider & Edwards, 2011) for reads reaching Q30 quality and less than 10 % of ambiguous bases in the sequence, represented as “Ns”. Q30 is considered a benchmark for quality in next-generation sequencing (meaning every 1,000 bp sequencing read may contain an error). A large percentage of N can be a sign of a low quality sequence and is used to infer the quality of reads if no quality scores are available. Libraries were mapped and annotated using *HISAT2 v2.05* (Kim et al., 2019) and the CSIC gilthead seabream draft genome as reference (Pérez-Sánchez et al., 2019). Read numbers mapped over each gene were counted

using *FeatureCounts* (Liao et al., 2014). To consider the effect of sequencing depth and gene lengths, gene expression levels were measured using reads per kilo base per million mapped reads (RPKM).

2.8.3.2 Differential expression analysis

In experiment 1, the differential expression analysis was performed using *DEGenes Hunter* (González et al., 2017), now available at *Bioconductor* as the *R library ExpHunterSuite* (Perkins et al., 2021). The libraries used were *edgeR* and *DEseq2* to generate files with quality control and expression data. Default thresholds for differential expression are absolute fold-change (FC) > 2 and false discovery rate (FDR) <0.05 in every algorithm (*edgeR* and *DESeq2*). Differential expression patterns were analyzed by family (gD1, gD2, gN1 or gN2 vs. gD or gN) and by condition (gD vs. gN). Some online tools were used to represent the data of differentially expressed genes (DEGs) in each comparison, volcano plot of the *Galaxy Platform* was used to represent the volcano plots (Afgan et al., 2016) and *Heatmapper server* to represent heatmaps (Babicki et al., 2016).

In experiment 3, DEGs were retrieved with normalized RPKM values using *DEseq2 v1.20.0* (Love et al., 2014) at an adjusted FDR of 0.05. To increase the number of DEG without loss of statistical robustness, supervised partial *least-squares discriminant analysis* (*PLS-DA*) and hierarchical clustering of samples were sequentially applied using *EZinfo v3.0* (Umetrics, Umea, Sweden) and R package *ggplot2*, respectively. The final list of genes contributing to group separation was determined by the minimum Variable Importance in the Projection (VIP) values (Li et al., 2012; Kieffer et al., 2016), driving the right clustering of all individuals in the heatmap analysis. To discard the possibility of over-fitting of supervised discriminant model, a validation test consisting in 500 random permutations was performed using *SIMCA-P+ v11.0* (Umetrics). The heatmap and clustering representations were constructed using *iDEP v0.91* (Ge et al., 2018), with the *average linkage method* and *Euclidean distance* as main parameters.

2.8.3.3 Gene enrichment analysis

Gilthead seabream is not included in most widely used algorithms for functional interpretations, so model organism orthologs was determined for each transcript as reference.

In experiment 1, zebrafish ortholog was determined for each differentially expressed transcript (DET) using the *Full-LengtherNext* algorithm annotation tool (instructions and off-line installation can be obtained in http://www.scbi.uma.es/site/scbi/downloads/313-full-lengthernext) and the current zebrafish proteome from UniProt (UP000000437). This annotation tool is relied in *BLAST+* (Camacho et al., 2019). Then, gene enrichment analysis was performed using a functional enrichment analysis web tool called *WebGestalt* (Liao et al., 2019) by an Over-Representation Analysis (ORA) of biological process gene ontology (GO-BP) with the total DET obtained from the family and condition comparisons (gD1 vs. gN + gD2 vs. gN + gD vs. gN1 + gD vs. gN2 + gD vs. gN) and by a Gene Set Enrichment Analysis (GSEA) with the condition comparison (gD vs. gN), all of them per tissue (brain, white muscle and vertebra). R package *apcluster* was used to cluster gene sets.

In experiment 3, genes with a VIP value > 1 (discriminant genes) were used in gene enrichment analysis. Discriminant genes were converted into their human equivalents and used to perform an ORA with GO-BP using the *ShinyGO v0.61* (Ge et al., 2020). Significantly enriched GO-BP categories were obtained after FDR correction with a cut-off of 0.05, and clustered to observe overlapping using the average linkage method and Euclidean distance of R function *hclust*.

2.8.3.4 Variant calling analysis

Variant calling analysis requires a large number of samples to guarantee adequate statistical power. So, this analysis was carried out in experiment 3, due to the pooled nature of the samples, with genes having a VIP value > 1 (discriminant genes). Alignment files generated in the mapping step were used for the calling of genetic variants, being the transcription start sites of discriminant genes in silico predicted within 5,000 bp upstream start codon (ATG) of each gene using *PROMOTER 2.0* (Knudsen, 1999). To minimize biased estimates of allele

frequencies, PCR duplicates were removed from the samples files using *SAMtools v1.6* (Li et al., 2009). The resultant files were then introduced in *Freebayes v1.0.2* (Garrison & Marth, 2012). Variants were considered in discriminant genes when (i) quality of the variant base-pair position was at least 30, (ii) at least 2 reads supported the alternate allele, and (iii) the variant allele frequency in the pool was above 0.05. The resulting variant calling files were formatted and allelic frequencies differences were calculated using the scripts *snp-frequency-diff.pl* and *fisher-test.pl* from *Popoolation2 v1.201* package (Kofler et al., 2011). Significant coding variants were annotated using *SnpEff v4.3* (Cingolani et al., 2012), and genetic architecture of discriminant genes with significant variants was represented using the *genemodel v1.1.0* R package. Variant physical position was located among the 24 super-scaffolds (SS) using *bedtools v2.29.2* (Quinlan & Hall, 2010), and clustering of higher variant density SS was performed using the average linkage method and Euclidean distance of R function *hclust*.

2.9 Statistical analysis

2.9.1 Breeders evaluation

All breeders data traits were tested for normality and homogeneity of variance by using a *General Linear Model analysis*, by *SPSS package, version 22.0* (SPSS Inc, Chicago, United States). Variance components for all traits were estimated by BLUP using the following mixed model:

$$y=X\beta+Zu+e$$

where y is the recorded data on the studied traits, β the fixed effects, u the random animal effect, and e the residual error. All models were resolved with the software package *VCE (v 6.0)* (Neumaier & Groeneveld, 1998; Groeneveld et al., 2010) and a complementary program to help for managing input data and automatize processes (*VCE_Executer, v3.0*) (León-Bernabeu et al., 2021).

2.9.2 Spawning quality

The data were tested for normality using the one sample *Kolmogorov–Smirnoff test*, as well as for homogeneity of variance using *Levene's test*. When a normal distribution and/or homogeneity of variance was not achieved, data were subjected to the *Kruskal–Wallis non-parametric test* (Zar, 1984). To study the variation of spawning quality traits throughout the whole spawning season, comparisons between types of crosses (gD vs. gN) and between experimental tanks (D1, D2, N1, N2) were carried out, using the day as a unit of measurement. The variation of spawning quality traits was also studied in short time periods, structuring the whole spawning season into twelve fortnights. Each spawning quality trait was compared between types of crosses (gD vs. gN) within each fortnight. Additionally, each trait was compared within types of crosses (gD or gN) between fortnights. The following general linear model was used:

$$Y_{ijk} = \mu + \alpha_i + \beta_{ij} + \varepsilon_{ijk}$$

Where, μ is the mean of the population, α_i is the effect of the genetic factor (types of crosses), β_{ij} is the effect of the tank factor within genetic factor, and ε_{ijk} is the residual error.

For the *oocyte yield* and *number of alive larvae* traits, within fortnights, the influence of genetic and tank factors was measured by the following statistical algorithm: execution of a standard regression analysis, and estimation of variance components of principal factors. Within fortnights, the same statistical algorithm was also applied to explain the *number of alive larvae* trait using as predictors the *oocyte yield, fertilization rate, viability rate, hatching rate* and *larvae survival rate* traits. Statistical analysis was performed using the *SPSS package, version 22.0* (SPSS Inc, Chicago, United States). All statistically significant comparisons of this work, include the correction factor of *Bonferroni's* to minimize the *Type I error* (Gordon et al., 2007). Thus, significant differences were considered from 0.05 as confident level.

2.9.3 Larvae deformity

The association between factors (mating, tank, dph, length) and deformity within experiment was analysed by a *Loglinear model* by using the *SPSS package, version 22.0* (SPSS Inc, Chicago, United States). Deformity factor was measured in each fish as presence (1) or absence (0) of the deformity. *Loglinear model* gives the significance of any deformity factor (its prevalence [*i*]) against any biological or functional factor (depending of data [*j*]), organized under a contingency table, through the normalized values or *Z-values*. *Normalized Z-values* for statistical significance of deformity in any family, breeder or mating ranged from > +1.96 (significant excess) to < −1.96 (significant deficit).

$$\mathrm{Ln} f_{ij} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij}$$

where $\ln f_{ij}$ is the expected frequency from observed frequency of each deformity (*i*) in each factor considered (*ij*); μ is the average value of expected frequencies logarithms, α_i is the effect of the deformity factor (*i*), β_j is the effect of the factor biological (*j*) and $\alpha\beta_{ij}$ is the effect due to interaction.

Chapter 3

Differential expression patterns in breeders of gilthead seabream (*Sparus aurata*) from divergent selection for the prevalence of skeletal deformities.

Álvaro **Lorenzo-Felipe**[1], Hyun Suk **Shin**[1], Sergio **León-Bernabeu**[2], Concha **Berbel**[3], Manuel **Manchado**[3], Rocío Bautista[4, 5], Gonzalo Claros[4, 5], María Jesús **Zamorano**[1] and Juan Manuel **Afonso-López**[1*]

[1] Aquaculture Research Group (GIA), Institute of Sustainable Aquaculture and Marine Ecosystems (IU-ECOAQUA), Universidad de las Palmas de Gran Canaria (ULPGC), Telde, Spain; [2] AQUANARIA S.L., San Bartolomé de Tirajana, Spain, [3] Center El Toruño, IFAPA, Consejería de Agricultura y Pesca, El Puerto de Santa María, Cádiz, Spain; [4]Department of Molecular Biology and Biochemistry, Faculty of Sciences, Campus de Teatinos s/n, University of Málaga, Málaga, Spain; [5] Plataforma Andaluza de Bioinformática, University of Málaga, Málaga, Spain.

In preparation

Abstract

Morphological abnormalities in farmed gilthead seabream (*Sparus aurata*) are an important problem that entails great economic losses. In this study, divergent selection for skeletal deformities was realized in order to describe differential expression patterns in brain, white muscle and vertebra as deformity target tissues of breeders. The mean estimated breeding values for the deformity in different family groups was 0.024 in genetically deformed (gD) families and -0,038 in genetically normal (gN) families. Brain and vertebrae seems to have similar transcriptome complexity with around 10,000 expressed transcripts, about double of transcripts than white muscle. After differential expression analysis a total of 59 differentially expressed transcripts (DET) were found between deformity condition levels, 2 of them in the three tissues, the coagulation factor XIII-A (*f13a1*) and *E3 ubiquitin-protein ligase like* (*TRIM39l*). At family level, a total of 702 DET were detected taking together the 3 tissues results. In terms of functional analysis, vertebra was the unique tissue showing enrichment categories between groups using Gene Ontology Biological Process database. For the first time, at least to our knowledge, an RNAseq study has been carried out in samples with divergent genetic predisposition to deformity, opening up numerous possibilities for future studies.

Keywords: Gilthead seabream, estimated breeding value, skeletal deformity, RNAseq, gene expression analysis, breeders.

3.1 Introduction

Gilthead seabream (*Sparus aurata*) is one of the most important aquaculture species in the Mediterranean countries with more than 95 % of the supply coming from aquaculture. It is currently produced in 20 countries and had a market of 278,199 t in 2020 and an estimated value at first selling of 1,251.9 M€ (APROMAR, 2021).

The presence of deformities is the second trait in economic importance for the industrial production sector of gilthead seabream, due to the devaluation of the market value that it causes in the product throughout the production chain, aggravated by its sale as whole fish (Bardon et al., 2009; Georgakopoulou et al., 2010; Boglione & Costa, 2011; Boglione et al., 2013a, 2013b). So, in this species, morphological quality traits have acquired the same importance as growth traits in European breeding programs, in terms of frequency (Janssen et al., 2017).

Deformities have a very high environmental component, but they also have a genetic component that can be used to improve breeders and stock quality permanently (Andrades et al., 1996; Afonso et al., 2000; Castro et al., 2008; Fernández et al., 2008; Boglione et al., 2013b). In gilthead seabream, a significant family association for a triple column deformity was found and more type of deformities was associated with their family in segregation studies (Afonso et al., 2000; Negrín-Báez et al., 2015b). Furthermore, diverse heritabilities for the presence of deformities (Astorga et al., 2004; Lee-Montero et al., 2015; Negrín-Báez et al., 2015b; García-Celdrán et al., 2015, 2016; Fragkoulis et al., 2018, 2020, 2021), QTLs (Negrín-Báez et al., 2015a, 2016a, 2016b) and differential expression patterns have been reported in this species (Riera-Heredia et al., 2019b).

Nowadays, with improvement of sequencing technologies there is an increasing number of genomic information from different species. Thus, Pauletto et al. assembled in 2018 the first complete genome of gilthead seabream, distributed in chromosomes and fully annotated (Pauletto et al., 2018).

This study is carried out within the framework of the national MORFOGEN Project, with the objective of use RNAseq technique to determine different gene expression patterns between animals with different genetic values for the presences of deformity. RNAseq is

cheaper than other techniques as WGS, and allows to map, annote and quantify at once gene expression in different samples (Piskol et al., 2013), benefiting of the available information of the genome to obtain information about the expressed genes (Manchado et al., 2016).

3.2 Material and methods

3.2.1 Biological Material and Experimental Design

A total of 5,281 gilthead seabream adults from the 3rd generation of the Spanish National Breeding Program (PROGENSA®-III) (León-Bernabeu et al., 2021) were visually assessed for the presence of deformity at commercial size (ATOL:0000087), in accordance with AquaExcel-ATOL (AQUAEXCEL, 2013). These animals come from a base population with a heritability for the deformity between 0.17-0.25 in adults (689 and 539 dph, respectively) and a genetic correlation between other stages of development higher than 83 % (Lee-Montero et al., 2015). The genotyping of all animals evaluated for morphology was also established through the molecular characterization of microsatellite markers (SMsa1), with *PCR* methodology, following the instructions of its authors (Lee-Montero et al., 2013). The genetic relationship between descendants (study population) and their parents (parental assignment) was reconstructed, comparing breeders' and descendants' genotypes by using *Vitassign* software (Vandeputte et al., 2006). Genetic parameters were estimated for both data matrix (phenotypic and genotypic information) by using the restricted maximum likelihood method with *VCE 6.0 software* (Neumaier & Groeneveld, 1998; Groeneveld et al., 2010) and a complementary program to help for managing input data and automatize processes (*VCE_Executer, v3.0*) (León-Bernabeu et al., 2021).

The EBVdef of the 5,281 evaluated gilthead seabream ranges between -0.044 for normality and +0.077 for deformity (Figure 6) distributed in 220 families, with an average of 24 descendants by family and standard deviation value of 39.9. The mean, min, and max EBVdef by family is represented in Figure 7. After discarding the dead and misshapen, a total of 12 breeders of 4 families (two gN families and two gD families) located in the Canary Islands region (AQUANARIA S.L. and FCPCT-ULPGC facilities) were selected on the basis of their EBVdef, considering that there were at least 3 full sibs per family with similar EBVdef to be

use as biological replicates. In this way, it was possible to establish two groups of breeders with opposite values in their EBVdef, which corresponds to a value of +0.024 in gD group (+0.023 in Family 1 and 0.025 in Family 2) and -0.038 in gN group (-0.033 in Family 1 and -0.043 in Family 2) (Table I). All these breeders was moved to the FCPCT-ULPGC facilities for conditioning.

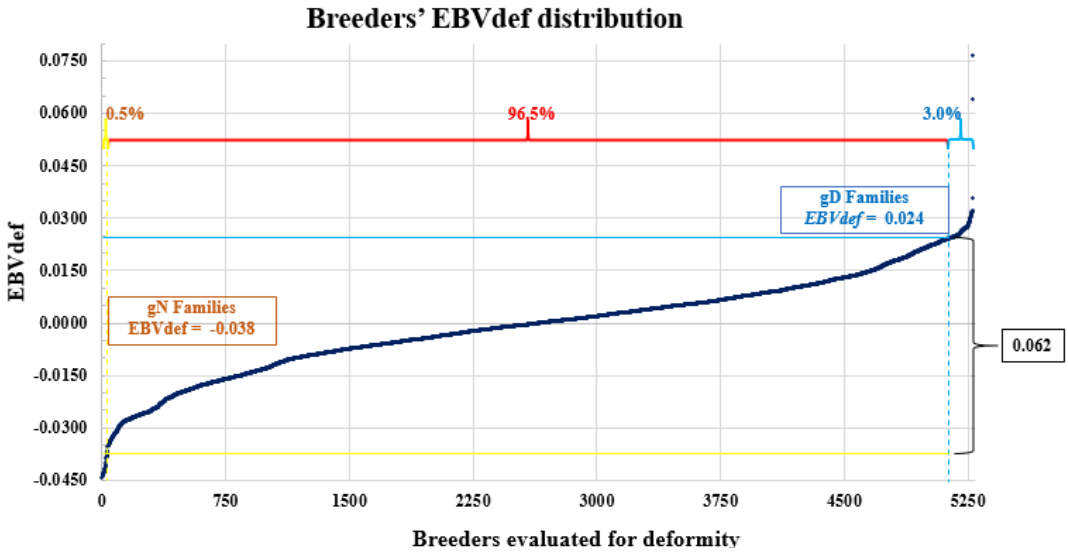


Figure 6. EBVdef' distribution of the 5,281 breeders evaluated for deformity. Yellow lines delimit the breeders with an EBVdef lower than the average of genetically normal selected families (gN Families) and blue lines the breeders with an EBVdef higher than the average of the genetically deformed selected families (gD Families). Between selected families, EBVdef distance is delimited in black and population distance in red.

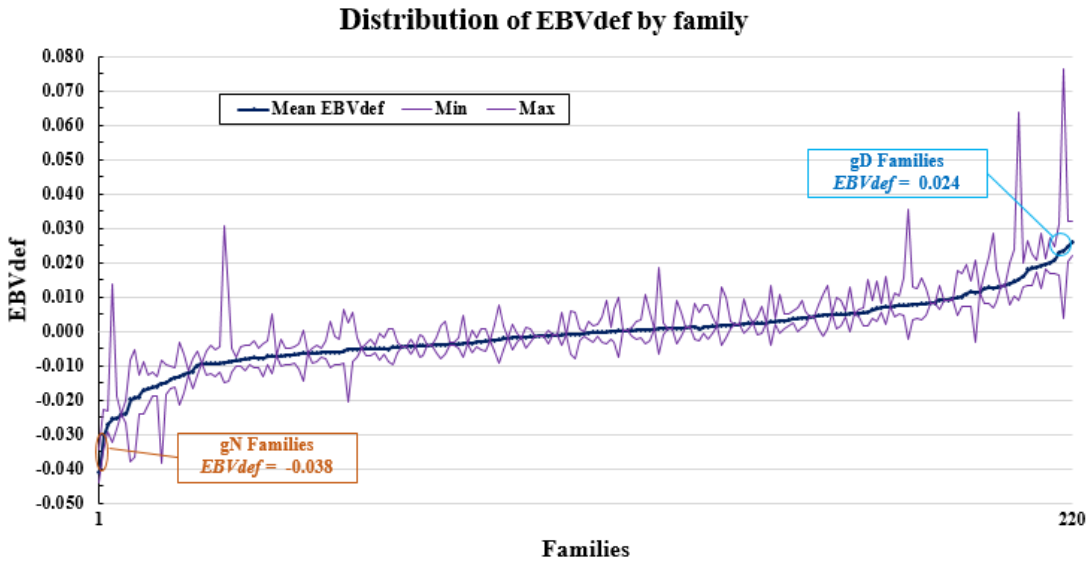


Figure 7. Distribution of the families' EBVdef in experiment 1. Mean, min and max EBVdef by Family are represented in dark blue and violet, respectively. The range where the two selected genetically normal families (gN Families) is marked in orange, and the two genetically deformed families (gD Families) in blue.

Table 1. Estimated breeding values for the deformity (EBVdef) of the 12 selected breeders in experiment 1. There are 3 animals per family as biological replicate (a, b, and c), two families per condition (Family 1 and Family 2) and two experimental conditions (genetically deformed (gD) and genetically Normal (gN))

Condition	Family	Animal	EBV_{def} Animal	EBV_{def} Family	EBV_{def} Condition
gD	Family 1	HD1a	0.022	0.023	0.024
		HD1b	0.024		
		HD1c	0.023		
	Family 2	HD2a	0.025	0.025	
		HD2b	0.025		
		HD2c	0.025		
gN	Family 1	LD1a	-0.033	-0.033	-0.038
		LD1b	-0.034		
		LD1c	-0.033		
	Family 2	LD2a	-0.044	-0.043	
		LD2b	-0.044		
		LD2c	-0.042		

3.2.2 Sampling procedures

Fish were fasted 24 h and anesthetized with clove oil (10 ppm clove oil:methanol, 1:1 volume in sea water). Then, each animal was slaughtered with excess of clove oil and certified by exsanguination, to collect white muscle, brain and vertebra samples for the RNAseq analysis. The cleaning procedure was applied between samples with propano AF and DEPC water. Tissue samples were collected from the same region in all animals. Vertebra samples were quickly frozen with liquid nitrogen and kept at −80 ◦C and white muscles and brain samples were conserved in RNA Later (Sigma-Aldrich) overnight at 4 °C, and kept at − 80 °C until RNA isolation.

3.2.3 RNAseq analysis

3.2.3.1 RNA isolation

Samples (60–70 mg) were extracted using the RNeasy Mini Kit (Qiagen). The homogenization process was carried out with Tissue Lyzer-II (Qiagen, Hilden, Germany) in 1 mL TRI Reagent (Sigma-Aldrich) in case of brain and white muscle samples. Vertebra samples were homogenized using a homogenizer with metal case with 6 mL TRI Reagent (Sigma-Aldrich) and it was cooled on the outside with liquid nitrogen to avoid overheating the sample, then 1 ml were separated to continue with the RNA isolation process following manufacters' protocol. A DNase digestion step was added with the RNase-Free DNase Set (QIAGEN). Purified RNA was eluted with 50 μL of RNase-free water in white muscle samples and in 35 μL of RNase-free water in vertebra and brain samples. RNA quality was checked by 1.4 % agarose electrophoresis and quantified by using NANODROP-8000 spectrophotometer (ThermoFisher) before being sent with dry ice to the Sequencing Unit from the University of Malaga, Spain. RNA quality was analyzed by lab-on-a-chip electrophoresis using the bioanalyzer 2100 kit (Agilent, Amstelveen, Netherlands). Quantification of samples was performed using a Qubit fluorimeter (LifeTechnologies). RNA aliquots were stored at -80 ºC for later processing.

3.2.3.1 RNA sequencing

Illumina RNAseq libraries were prepared using the *Illumina TruSeq Stranded mRNA Library Preparation Kit* (NEB, USA), according to the manufacturer's instructions. Once mRNA libraries were prepared, sequencing was performed using *NextSeq 550* (Illumina), 150 cycles of 2 x 75 pb mRNA sequencing was performed.

3.2.3.1 Bioinformatic analysis

Picasso supercomputer of University of Malaga was used for all bioinformatics tests, implementations and executions. A scheme of the workflow followed to obtain the different result from the raw reads is represented in Figure 8.

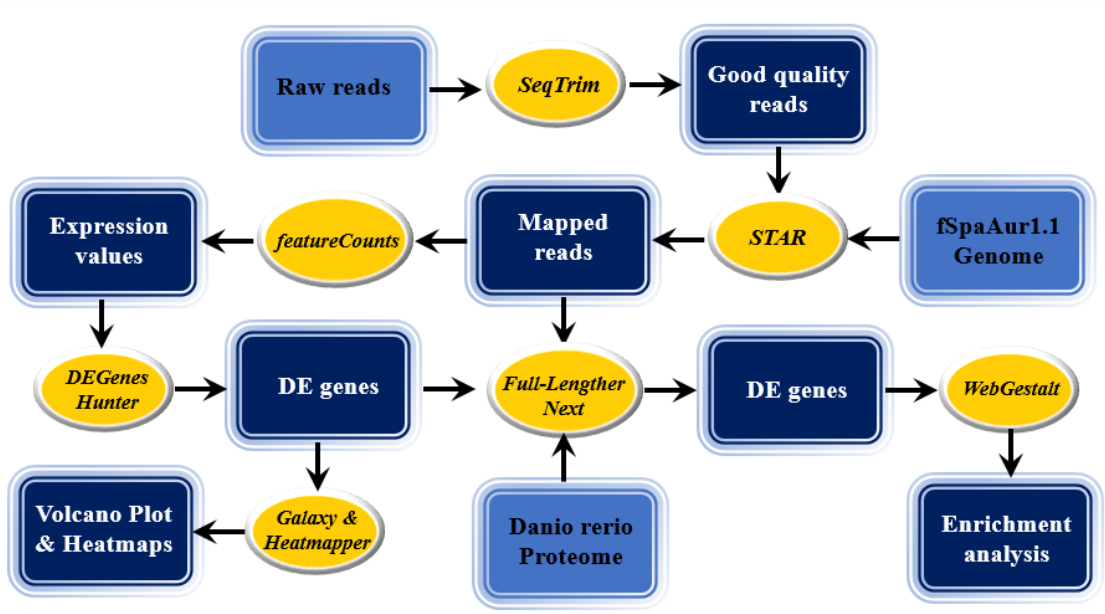


Figure 8. General workflow followed to carry out the processing and bioinformatics analysis of the output readings of the sequencer in experiment 1. Sequence filtering, its mapping in a reference genome, counting of expressed transcripts and later the different analyses (differential expression (DE) and functional enrichment analysis). In light blue is represented new data incomes, in yellow the softwares' names and in dark blue processed data.

3.2.3.1.1 Reads pre-processing and genome mapping

Raw reads were pre-processed using *SeqTrim* using default parameters (Falgueras et al., 2010). Then, reads were mapped onto the transcriptome deduced from the haploid gilthead seabream genome assembly *fSpaAur1.1* (GCA_900880675.1) using *STAR* (Dobin et al., 2013). Mapping was converted in expression values using *featureCounts* (Liao et al., 2014).

3.2.3.1.2 Differential expression analysis

The differential expression analysis was performed using *DEGenes Hunter* (González et al., 2017). The libraries used were *edgeR* and *DEseq2* to generate files with quality control and expression data. Default thresholds for differential expression were absolute FC > 2 and FDR < 0.05 in every algorithm (*edgeR* and *DESeq2*). Differential expression patterns were analyzed by family (gD1 vs. gN, gD2 vs. gN, gD vs. gN1 and gD vs. gN2) and by condition (gD vs. gN). Volcano plot online tools of the *Galaxy Platform* was used to represent expression data in volcano plots (Afgan et al., 2016) and *Heatmapper server* to represent DET expression data of all samples in heatmaps (Babicki et al., 2016).

3.2.3.1.2 Gene enrichment analysis

Gilthead seabream is not included in most widely used algorithms for functional interpreation, so the zebrafish ortholog was determined for each transcript using the *Full-LengtherNext* algorithm annotation tool (instructions and off-line installation can be obtained in http://www.scbi.uma.es/site/scbi/downloads/313-full-lengthernext) and the current zebrafish proteome from UniProt (UP000000437). This annotation tool is relied in *BLAST+* (Camacho et al., 2019). Then, gene enrichment analysis was performed using a functional enrichment analysis web tool *WebGestalt* (Liao et al., 2019) by an ORA of biological process using the GO-BP functional database with the total DET obtained from the family and condition comparisons (gD1 vs. gN + gD2 vs. gN + gD vs. gN1 + gD vs. gN2 + gD vs. gN) and by a GSEA with the condition comparison (gD vs. gN), all of them per tissue (brain, white muscle and vertebra). R package *apcluster* was used to cluster gene sets.

3.2.4 Statistical analysis

All breeders data traits were tested for normality and homogeneity of variance by using a *General Linear Model analysis*, by *SPSS package, version 22.0* (SPSS Inc, Chicago, United States). Variance components for all traits were estimated by BLUP using the following mixed model:

$$y=X\beta+Zu+e$$

where *y* is the recorded data recorded on the studied traits, *β* the fixed effects, *u* the random animal effect, and *e* the residual error. All models were resolved with the software package *VCE (v 6.0)* (Neumaier & Groeneveld, 1998; Groeneveld et al., 2010) and a complementary program to help for managing input data and automatize processes (*VCE_Executer, v3.0*) (León-Bernabeu et al., 2021).

3.3 Results

3.3.1 Reads Preprocessing

Only one of the 36 samples that were collected was lost during the sequencing process due to technical problems, corresponding to the gN2Va sample. The cDNA libraries of the other samples had an average size of fragment length between 200 bp and 500 bp, in addition to not presenting more than 5 % in sizes less than 200 bp. This resulted in an average of 34.2 M raw paired-end reads resulting from the sequencer.

3.3.2 Differential expression analysis

In terms of expression, a total of 10,506 transcript corresponding to 10,321 unique genes in brain, 5,525 transcripts corresponding to 5,493 unique genes in white muscle and 10,035 transcripts corresponding to 9,933 unique genes in vertebra, for a total of 60,650 transcript corresponding to 31.394 unique genes in *fSpaAur1.1* genome assembly.

In comparisons between EBVdef groups there were two DEG overexpressed in gD in the 3 tissues (*f13a1* and *TRIM39l*) and other two uncharacterized DEG, also overexpressed in gD group, in brain and vertebra (*Matn1l* and *LOC115582631*). Counts of the number of DEG in the analysis by EBVdef condition is represented in Table 2 and detailed below by tissue.

After carrying out the differential expression analysis between conditions (gD vs. gN), it was decided to study whether there were differences in expression at family level (gD1 vs. gN, gD2 vs. gN, gD vs. gN1 and gD vs. gN2), thus increasing the number of DET to be studied later in the functional analysis. Taking into account DEG in at least one of the family comparisons, the uncharacterized gene *LOC115582631* was overexpressed in the three tissues and another 2 new genes were overexpressed, one in gN (*entpd4l*) and another one in gD (*impdh2*). 1 DEG was overexpressed in gD group in brain and muscle (*LOC115590264*).

13 DEG were overexpressed in brain and vertebra, 10 in gD (*LOC115569868*, *smu1b*, *sparcl1*, *arf4al*, *fell*, *nlrp12l*, *dennd4b*, *slc25a19*, *hoxb6a* and *bhmtl*), and 3 in gN (*nudt4a*, *rrp36l* and *ch25h*). Finally, 5 DEG were overexpressed in muscle and vertebra, all of them in gD (*igfbp5a*, *yjefn3*, *tpm3*, *LOC115589942* and *LOC115577687*).

Table 2. Number of differentially expressed genes (DEG) by tissue and condition. There are 3 tissues (brain, muscle and vertebra) and two levels of one condition, genetically deformed group vs. genetically normal (gD vs. gN). The DEG are separated by those overexpressed in gD and gN. In addition, if genes appear in the two statistical analyses (DEseq2 and edgeR) are marked in blue, only DEseq2 in green and only edgeR in red.

Tissue	Nº DEG	Nº ↑gD	Nº ↑gN
Brain	18 3 2 ***23***	6 2 1 **9**	12 1 1 ***14***
Muscle	19 6 **25**	8 2 **10**	11 4 ***15***
Vertebra	5 12 **17**	4 6 **10**	1 6 *7*

3.3.2.1 Brain

There are a total of 23 DET in brain in the comparison between gD and gN groups, 18 as result of the two statistical analysis realized (*DEseq2* and *edgeR*), 3 only coming from de *DEseq2* analysis and 2 from the *edgeR* analysis. The data of the 23 DET is represented in Table 3 and the different Z score between samples should be see in the heatmap represented in Figure 9A. Of the 23 DET, 9 were overexpressed in gD group and 14 in gN group. The *edgeR* analysis was selected, for showing more restrictive results, to be represented in a volcano plot to show the differences of expression values between gD and gN groups in Figure 10A.

At family level, a total of 121 DET were detected in at least one of the 4 families and one of the two statistical analysis (*DEseq2* or *edgeR*), with 31, 43, 28 and 19 DET in families gD1, gD2, gN1 and gN2, respectively. Of the 121 DET, there are left 96, if transcripts that are repeated between the results of the different families are removed, and 78, counting only those that were not in the results of the comparison between gD and gN groups. Of this 78, 3 were differentially expressed in a gN and a gD family and the other 75 DET were distributed in 21, 30, 15 and 9 of gD1, gD2, gN1 and gN2, respectively.

Table 3. Information of the 23 differentially expressed transcripts (DET) in brain samples between genetically deformed and normal groups (gD vs. gN). The available information is Transcript ID, condition, lofFC and FDR of two statistical analysis (DEseq2 and edgeR), gene name, symbol and description.

Transcript ID	Condition	logFC (Deseq2|edgeR)	FDR (Deseq2-edgeR)	Gene	Symbol	Description
XR_003984196	gD vs. gN	(-9.76 | -9.52)	(0.000-0.002)	*LOC115582631*	*LOC115582631*	Uncharacterized LOC115582631
XM_030418694	gD vs. gN	(-8.87 | -9.42)	(0.003-0.000)	*LOC115582626*	*Matn1l*	Uncharacterized LOC115582626
XM_030418695	gD vs. gN	(-7.39 | -7.95)	(0.091-0.008)	*LOC115582627*	*LOC115582627*	RNA-binding protein 25-like
XM_030418686	gD vs. gN	(-6.52 | -6.23)	(0.043-0.024)	*LOC115582623*	*LOC115582623*	Uncharacterized LOC115582623
XM_030398122	gD vs. gN	(-3.42 | -3.38)	(0.012-0.046)	*f13a1*	*f13a1*	Coagulation factor XIII A chain
XM_030431575	gD vs. gN	(-2.66 | -2.65)	(0.007-0.028)	*LOC115590271*	*LOC115590271*	E3 ubiquitin-protein ligase TRIM39-like
XM_030429348	gD vs. gN	(-1.66 | -1.65)	(0.037-0.050)	*LOC115588637*	*ttnl*	Titin-like
XR_003981182	gD vs. gN	(-1.63 | -1.62)	(0.036-0.068)	*LOC115567383*	*LOC115567383*	Uncharacterized LOC115567383
XM_030422001	gD vs. gN	(-1.16 | -1.15)	(0.026-0.057)	*LOC115584527*	*smu1l*	WD40 repeat-containing protein SMU1-like
XM_030441183	gD vs. gN	(1.00 | 1.00)	(0.016-0.028)	*LOC115596284*	*LOC115596284*	Nuclear factor 7, ovary-like
XM_030428222	gD vs. gN	(1.06 | 1.06)	(0.022-0.046)	*LOC115587996*	*LOC115587996*	Coiled-coil domain-containing protein 8 homolog
XM_030431336	gD vs. gN	(1.14 | 1.14)	(0.000-0.000)	*LOC115590073*	*Havcr2h*	Hepatitis A virus cellular receptor 2 homolog
XM_030419906	gD vs. gN	(1.30 | 1.30)	(0.036-0.057)	*LOC115583256*	*dnase1l*	Deoxyribonuclease-1-like
XM_030424655	gD vs. gN	(1.35 | 1.35)	(0.028-0.046)	*LOC115585965*	*LOC115585965*	E3 ubiquitin-protein ligase TRIM47-like
XM_030416625	gD vs. gN	(1.43 | 1.43)	(0.004-0.008)	*gp1bb*	*gp1bb*	Glycoprotein Ib platelet subunit beta
XM_030440245	gD vs. gN	(1.50 | 1.50)	(0.000-0.000)	*LOC115595589*	*LOC115595589*	23 kDa integral membrane protein-like
XM_030429645	gD vs. gN	(1.55 | 1.54)	(0.028-0.044)	*LOC115588969*	*epdl1*	Ependymin-1-like
XM_030399416	gD vs. gN	(1.79 | 1.80)	(0.002-0.008)	*ch25h*	*ch25h*	Cholesterol 25-hydroxylase
XM_030416810	gD vs. gN	(1.83 | 1.83)	(0.000-0.000)	*LOC115581590*	*myo1ha*	Unconventional myosin-Ih-like
XM_030398382	gD vs. gN	(1.95 | 1.96)	(0.007-0.024)	*LOC115570106*	*Csta1Al*	Cystatin-A1-like
XM_030393540	gD vs. gN	(2.25 | 2.25)	(0.031-0.028)	*LOC115567190*	*hampl*	Uncharacterized LOC115567190
XM_030393120	gD vs. gN	(2.57 | 2.56)	(0.018-0.050)	*LOC115567004*	*hampl*	Hepcidin-like
XM_030398384	gD vs. gN	(3.07 | 3.06)	(0.000-0.002)	*LOC115570106*	*Csta1Al*	Cystatin-A1-like

3.3.2.2 White muscle

There are a total of 25 DET in muscle in the comparison between gD and gN groups, 19 as result of the two statistical analysis realized (*DEseq2* and *edgeR*) and 6 coming only from the *DEseq2* analysis. The data of the 25 DET is represented in Table 4 and the different Z score between samples should be see in the heatmap represented in Figure 9B. Of these 25 DET, 10 were overexpressed in gD group and 15 were overexpressed in gN group. The results of the DET by *edgeR* analysis between gD and gN groups were represented in a volcano plot in Figure 10B.

At family level, a total of 147 DET were detected in at least one of the 4 families and one of the two statistical analysis (*DEseq2* or *edgeR*), with 66, 29, 40 and 12 DET in families gD1, gD2, gN1 and gN2, respectively. Of the 147 DET, there are left 119, if transcripts that are repeated between the results of the different families are removed, and 97 counting only those that were not in the results of the comparison between gD and gN groups. Of this 97, 2 were

differentially expressed in a gN and a gD family and the other 95 DET were distributed in 52, 14, 26 and 3 of gD1, gD2, gN1 and gN2, respectively.

Table 4. Information of the 25 differentially expressed transcripts (DET) in white muscle samples between genetically deformed and normal groups (gD vs. gN). The available information is Transcript ID, condition, lofFC and FDR of two statistical analysis (DEseq2 and edgeR), gene name, symbol and description.

Transcript ID	Condition	logFC (Deseq2|edgeR)	FDR (Deseq2-edgeR)	Gene	Symbol	Description
XM_030431118	gD vs. gN	(-8.70 | -8.54)	(0.000-0.004)	*LOC115589942*	*LOC115589942*	Uncharacterized LOC115589942
XM_030398122	gD vs. gN	(-8.44 | -8.39)	(0.000-0.001)	*f13a1*	*f13a1*	Coagulation factor XIII A chain
XM_030431575	gD vs. gN	(-3.10 | -3.08)	(0.005-0.017)	*LOC115590271*	*LOC115590271*	E3 ubiquitin-protein ligase TRIM39-like
XM_030430209	gD vs. gN	(-2.76 | -2.75)	(0.000-0.000)	*LOC115589367*	*Gimap8l*	GTPase IMAP family member 8-like
XM_030440898	gD vs. gN	(-2.47 | -2.46)	(0.000-0.000)	*LOC115596101*	*cho-1l*	High-affinity choline transporter 1-like
XM_030406652	gD vs. gN	(-1.47 | -1.46)	(0.002-0.004)	*LOC115574953*	*LOC115574953*	Tryptase-like
XM_030412447	gD vs. gN	(-1.43 | -1.43)	(0.000-0.000)	*LOC115579000*	*pde5aa*	cGMP-specific 3',5'-cyclic phosphodiesterase-like
XM_030411820	gD vs. gN	(-1.33 | -1.32)	(0.018-0.032)	*LOC115578715*	*thsd7aa*	Thrombospondin type-1 domain-containing protein 7A-like
XM_030416800	gD vs. gN	(-1.25 | -1.26)	(0.025-0.054)	*selenop*	*selenop*	Selenoprotein P
XM_030392365	gD vs. gN	(-1.14 | -1.14)	(0.047-0.080)	*LOC115566466*	*adh1*	Alcohol dehydrogenase 1
XM_030441638	gD vs. gN	(1.06 | 1.07)	(0.018-0.032)	*phyhipl*	*phyhipl*	Phytanoyl-CoA 2-hydroxylase interacting protein like
XM_030418482	gD vs. gN	(1.12 | 1.12)	(0.036-0.062)	*LOC115582508*	*cho-1l*	High affinity choline transporter 1-like
XM_030425798	gD vs. gN	(1.13 | 1.13)	(0.012-0.017)	*chkb*	*chkb*	Choline kinase beta
XM_030398443	gD vs. gN	(1.22 | 1.22)	(0.001-0.002)	*gfod1*	*gfod1*	Glucose-fructose oxidoreductase domain containing 1
XM_030434733	gD vs. gN	(1.25 | 1.25)	(0.000-0.002)	*musk*	*musk*	Muscle associated receptor tyrosine kinase
XM_030405977	gD vs. gN	(1.38 | 1.39)	(0.008-0.014)	*LOC115574446*	*fosab*	Proto-oncogene c-Fos
XM_030413076	gD vs. gN	(1.51 | 1.51)	(0.009-0.016)	*angpt1*	*angpt1*	Angiopoietin 1
XM_030441054	gD vs. gN	(1.75 | 1.75)	(0.000-0.000)	*LOC115596203*	*ninl*	Ninein-like protein
XM_030426573	gD vs. gN	(1.75 | 1.74)	(0.000-0.000)	*LOC115587024*	*tmem189l*	Transmembrane protein 189-like
XM_030392730	gD vs. gN	(2.24 | 2.25)	(0.039-0.073)	*LOC115566755*	*etsrp*	Protein C-ets-1-like
XM_030426406	gD vs. gN	(2.41 | 2.41)	(0.009-0.017)	*fa2h*	*fa2h*	Fatty acid 2-hydroxylase
XM_030438702	gD vs. gN	(2.92 | 2.91)	(0.002-0.006)	*LOC115594560*	*myl10*	Myosin regulatory light chain 2B, cardiac muscle isoform-like
XM_030412881	gD vs. gN	(3.09 | 3.09)	(0.027-0.103)	*LOC115579373*	*tpm3*	Tropomyosin alpha-3 chain
XM_030431385	gD vs. gN	(3.75 | 3.72)	(0.018-0.062)	*ppargc1a*	*ppargc1a*	PPARG coactivator 1 alpha
XM_030395524	gD vs. gN	(4.03 | 4.02)	(0.005-0.029)	*LOC115568331*	*DDIT4L*	DNA damage-inducible transcript 4-like protein

3.3.2.3 Vertebra

There are a total of 17 DET in vertebra in the comparison between gD and gN groups, 5 as result of the two statistical analysis realized (*DEseq2* and *edgeR*) and 12 coming only from de *DEseq2* analysis. The data of the 17 DET is represented in Table 5 and the different Z score between samples should be see in the heatmap represented in Figure 9C. Of these 17 DET, 10 were overexpressed in gD group and 7 were overexpressed in gN group. The results of the DET by *edgeR* analysis between gD and gN groups were represented in a volcano plot in Figure 10C.

At family level, a total of 525 DET were detected in at least one of the 4 families and one of the two statistical analysis (*DEseq2* or *edgeR*), with 12, 408, 15 and 90 DET in families gD1, gD2, gN1 and gN2, respectively. Of the 525 DET, there are left 510, if transcripts that are repeated between the results of the different families are removed, and 492 counting only those

that were not in the results of the comparison between gD and gN groups. Of this 494, 5 were differentially expressed in a gN and a gD family and the other 489 DET were distributed in 9, 392, 7 and 81 of gD1, gD2, gN1 and gN2, respectively.

Table 5. Information of the 17 differentially expressed transcripts (DET) in vertebra samples between genetically deformed and normal groups (gD vs. gN). The available information is Transcript ID, condition, lofFC and FDR of two statistical analysis (DEseq2 and edgeR), gene name, symbol and description.

Transcript ID	Condition	logFC (Deseq2|edgeR)	FDR (Deseq2-edgeR)	Gene	Symbol	Description
XM_030398122	gD vs. gN	(-9.47 | -9.46)	(0.000-0.005)	*f13a1*	*f13a1*	Coagulation factor XIII A chain
XM_030418694	gD vs. gN	(-7.40 | -8.08)	(0.003-0.005)	*LOC115582626*	*Matn1l*	Uncharacterized LOC115582626
XR_003984196	gD vs. gN	(-6.82 | -6.77)	(0.000-0.070)	*LOC115582631*	*LOC115582631*	Uncharacterized LOC115582631
XM_030415961	gD vs. gN	(-6.74 | -6.53)	(0.000-0.000)	*LOC115581104*	*gcnt3l*	Beta-1,3-galactosyl-O-glycosyl-glycoprotein beta-1,6-N-acetylglucosaminyltransferase-like
XM_030412644	gD vs. gN	(-6.25 | -5.95)	(0.028-0.146)	*LOC115579134*	*LOC115579134*	Uncharacterized LOC115579134
XM_030398094	gD vs. gN	(-5.84 | -5.71)	(0.000-0.000)	*LOC115569888*	*nlrp12l*	NACHT, LRR and PYD domains-containing protein 12-like
XM_030412790	gD vs. gN	(-3.05 | -3.01)	(0.023-0.064)	*LOC115579252*	*cyp3a65*	Cytochrome P450 3A40-like
XM_030431304	gD vs. gN	(-2.96 | -2.95)	(0.037-0.131)	*LOC115590048*	*LOC115590048*	Uncharacterized LOC115590048
XM_030431575	gD vs. gN	(-2.28 | -2.28)	(0.018-0.064)	*LOC115590271*	*TRIM39-like*	E3 ubiquitin-protein ligase TRIM39-like
XM_030443781	gD vs. gN	(-2.26 | -2.24)	(0.034-0.073)	*LOC115597643*	*esrrd*	Estrogen-related receptor gamma-like
XM_030423926	gD vs. gN	(1.19 | 1.18)	(0.020-0.073)	*LOC115585506*	*samhd1*	Deoxynucleoside triphosphate triphosphohydrolase SAMHD1-like
XM_030409807	gD vs. gN	(1.35 | 1.34)	(0.027-0.073)	*LOC115577031*	*ghrl*	Growth hormone receptor-like
XM_030417431	gD vs. gN	(1.35 | 1.36)	(0.046-0.079)	*ccser1*	*ccser1*	Coiled-coil serine rich protein 1
XM_030424619	gD vs. gN	(1.60 | 1.61)	(0.048-0.073)	*LOC115585946*	*nudt4a*	Diphosphoinositol polyphosphate phosphohydrolase 3-beta-like
XM_030397710	gD vs. gN	(2.19 | 2.17)	(0.001-0.005)	*LOC115569678*	*lectinl*	Lectin-like
XM_030397091	gD vs. gN	(2.25 | 2.23)	(0.039-0.073)	*neil3*	*neil3*	Nei like DNA glycosylase 3
XM_030408032	gD vs. gN	(3.52 | 3.49)	(0.028-0.077)	*LOC115575757*	*LOC115575757*	Uncharacterized LOC115575757

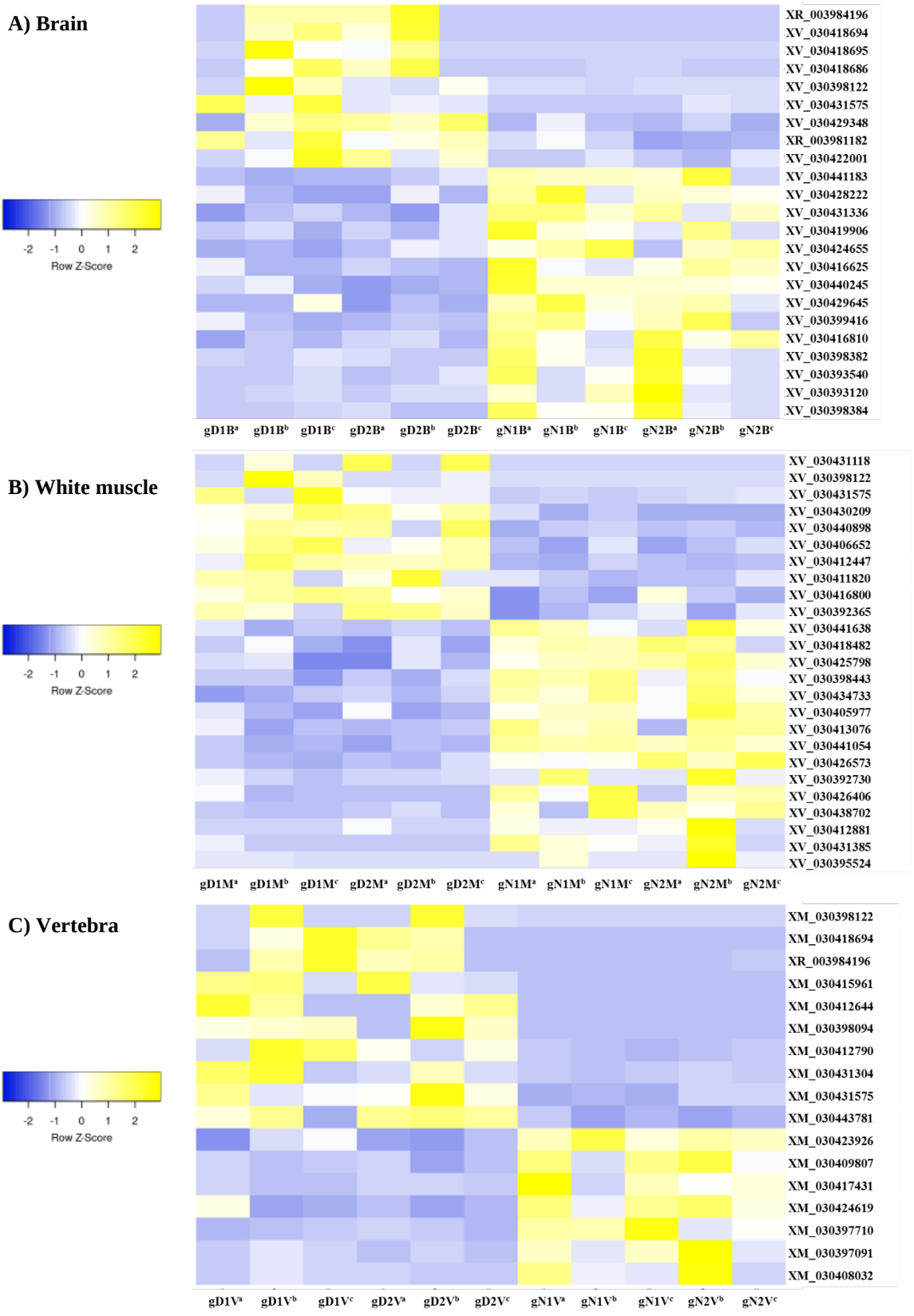
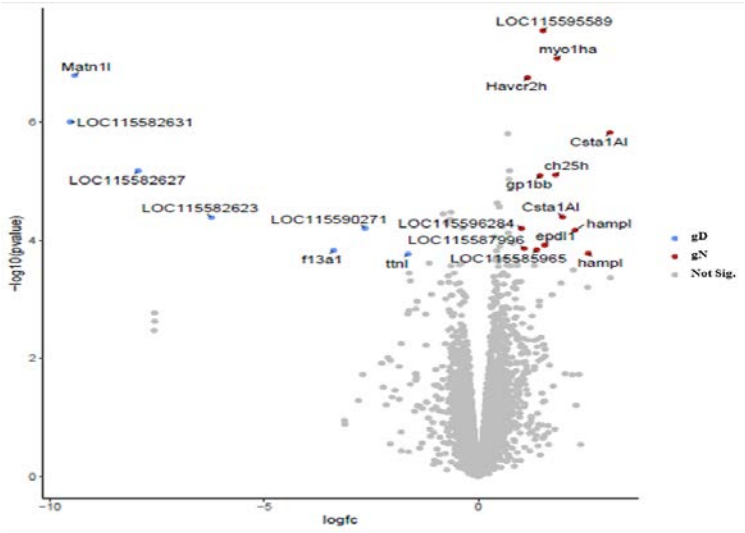
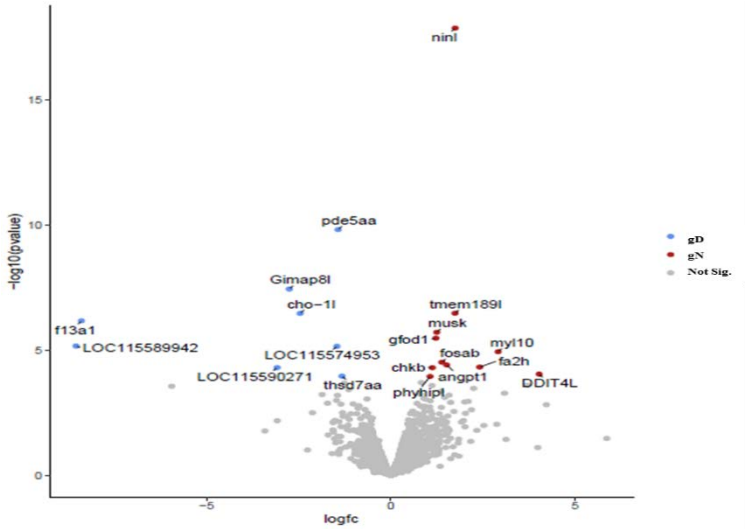


Figure 9. Heatmap representation with the expression values of the different samples by tissue using the condition comparison (gD vs gN). A) Brain, B) White muscle and C) Vertebra. In columns are represented the processed samples and in rows the DET. Colours range represent the row Z-score, being bright yellow more expression and dark blue less expression respect to the other samples.

A) Brain



B) White muscle



C) Vertebra

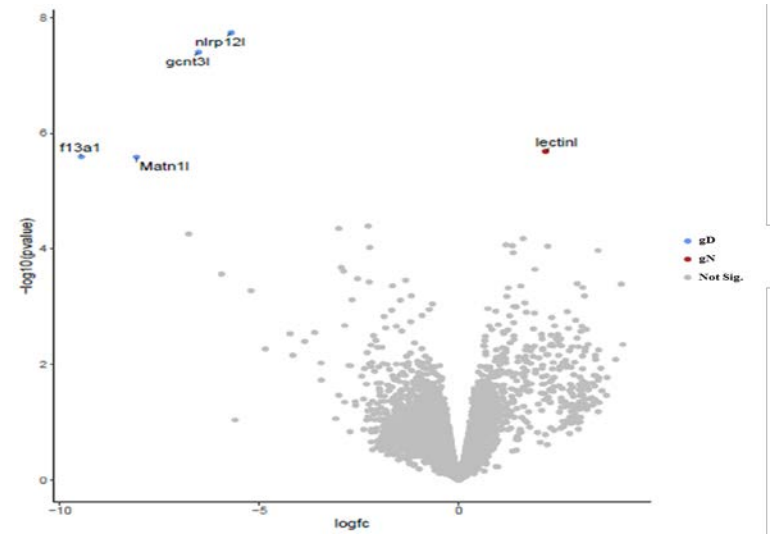


Figure 10. Expression values' volcano plots per genetic group using edgeR by tissue. Values with an FDR < 0.05 were coloured, in blue if it was overexpressen in gD group (logFC < -1) and in red if it was overexpressed in gN group (logFC > 1).

3.3.3 Gene enrichment analysis

Of the 60,650 transcripts available in the *fSpaAur1.1* gilthead seabream genome, 19,895 orthologs in zebrafish proteome were found. Of them, 18,812 were found in *Webgestalt* and 8,159 were annotated to the selected functional category. This last 8,159 set of peptides was used as reference set for ORA using the GO-BP functional database for the analysis in the three tissues (brain, white muscle and vertebra). None of the tissues reported enriched categories results with an FDR < 0.05, but the top 10 enriched categories was reported as a description of the biological process overrepresented by the DET. This results and GSEA are reported below by tissue.

3.3.3.1 Brain

A total of 101 DET were resulted of the different comparisons in brain (23 from gD vs. gN comparison and 78 from family comparisons), 75 with orthologs in zebrafish. Of them, 61 were found in *Webgestalt* and 28 were annotated to the GO-BP functional database. The top 10 enriched categories were represented in Figure 11A and were clustered in the next 4 non-redundant categories with the normalized enrichment score in brackets: *cellular modified amino acid catabolic process* (65), *glutamate metabolic process* (49), *alcohol metabolic process* (12) and *cellular nitrogen compound catabolic process* (8).

Of the 10,506 transcripts expressed in brain, 9,785 had orthologs in zebrafish and were used for GSEA. Of them 8,899 were unique peptides, 8,296 were found in *Webgestalt* and 3,720 were annotated to the GO-BP functional database. The top 10 enriched categories were represented in Figure 12A. The 10 enriched categories in gN group had a p-values < 0.05 and were clustered in the next 6 non-redundant categories with an normalized enrichment score between 1.7 and 1.8 (*reactive oxygen species metabolic process, immune effector process, extracellular structure organization, urogenital system development, regulation of hormone levels* and *response to xenobiotic stimulus*). On the other hand, only 4 of the 10 enriched categories in gD group had a p-value < 0.05 with an normalized enrichment score between -1.4 and -1.6 (*pattern specification process, sialylation, regulation of signaling receptor activity* and *positive regulation of biosynthetic process*).

3.3.3.2 White muscle

A total of 122 DET were resulted of the different comparisons in white muscle (25 from gD vs. gN comparison and 97 from family comparisons), 104 with orthologs in zebrafish. Of them, 98 were found in *Webgestalt* and 48 were annotated to the GO-BP functional database. The top 10 enriched categories were represented in Figure 11B and were clustered in the next 6 non-redundant categories with the normalized enrichment score in brackets: *isoprenoid catabolic process* (56), *carnitine metabolic process* (42), *vasculogenesis* (21), *ribosomal large subunit biogenesis* (14), *cellular modified amino acid metabolic process* (13) and *actomyosin structure organization* (9).

Of the 5,525 transcripts expressed in white muscle, 5,256 had orthologs in zebrafish and were used for GSEA. Of them 4,963 were unique peptides, 4,751 were found in *Webgestalt* and 2,205 were annotated to the GO-BP functional database. The top 10 enriched categories were represented in Figure 12B. 1 of the 10 enriched categories in gN group had a FDR < 0.05 and a normalized enrichment score of 1.9 (*response to abiotic stimulus*) and the others had p-values < 0.5 and were clustered in the next 4 non-redundant categories with an enrichment score between 1.7 and 1.5 (*organic acid transport*, *cardiovascular system* development, *supramolecular fiber organization* and *immune effector process*). On the other hand, only 1 of the 10 enriched categories in gD group had a p-value < 0.05 with an normalized enrichment score of -1.7 (*ossification*).

3.3.3.3 Vertebra

A total of 511 DET were resulted of the different comparisons in vertebra (17 from gD vs. gN comparison and 494 from family comparisons), 405 with orthologs in zebrafish. Of them, 377 were found in *Webgestalt* and 169 were annotated to the GO-BP functional database. The top 10 enriched categories were represented in Figure 11C and were clustered in the next 3 non-redundant categories with the enrichment score in brackets: *localization within membrane* (8), *neurotransmitter transport* (6) and *chemical synaptic transmission* (4).

Of the 10,035 transcripts expressed in vertebra, 8,895 had orthologs in zebrafish and were used for GSEA. Of them 8,124 were unique peptides, 7,829 were found in *Webgestalt* and 3,564 were annotated to the GO-BP functional database. The top 10 enriched categories were represented in Figure 12C. All the enriched categories showed an FDR < 0.05 in gN group and were clustered in the next 3 non-redundant categories with an normalized enrichment score between 2.2 and 1.8 (*organelle fission, chromosome organization* and *reproduction*). Also in gD group, the 10 enriched categories showed FDR < 0.05 and were clustered in the next 4 non-redundant categories with an normalized enrichment score between -1.7 and -1.9 (*behavior, neuropeptide signaling pathway, receptor metabolic process* and *neurotransmitter transport*).

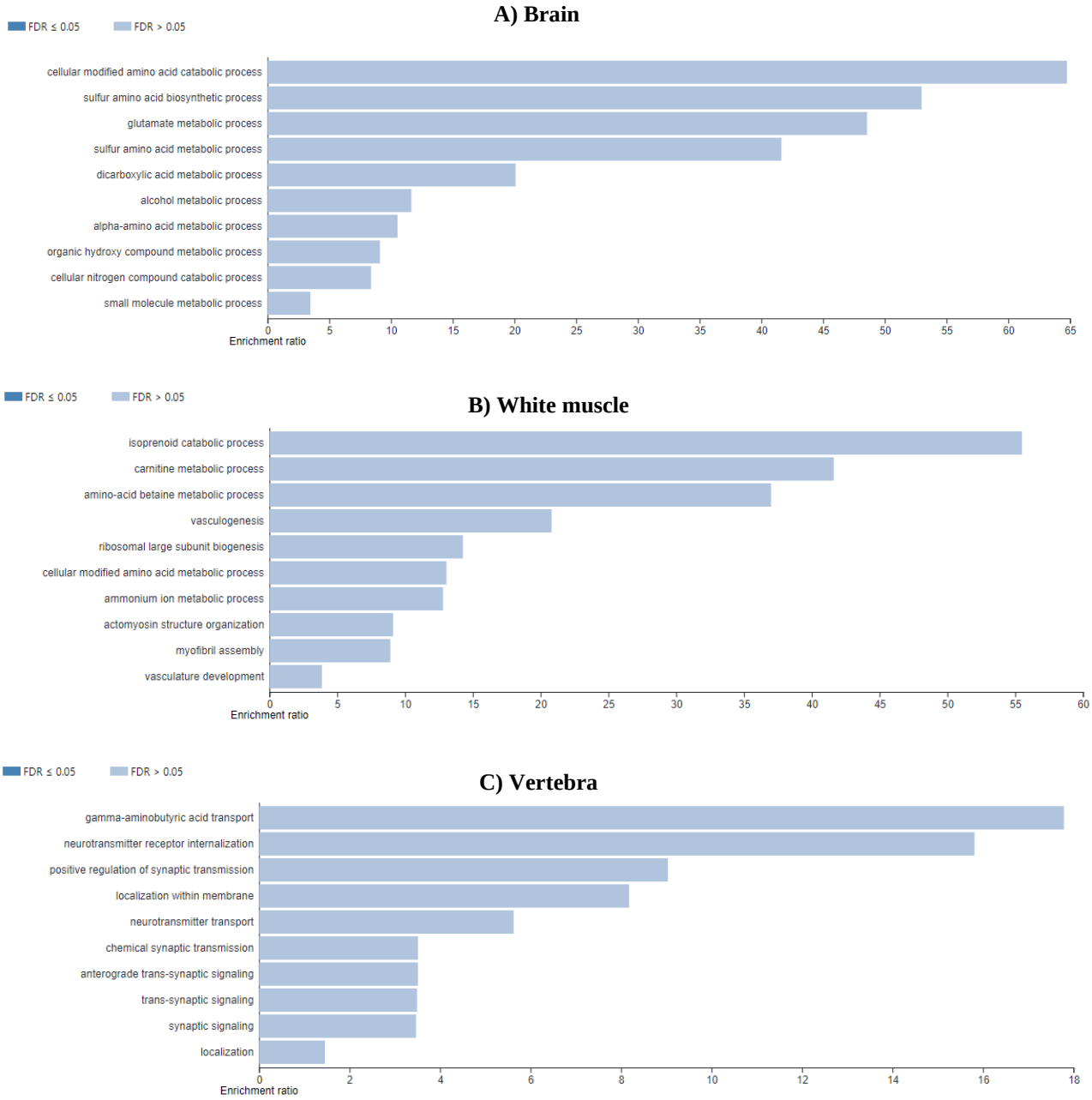


Figure 11. Top 10 GO-BP enriched categories sorted by enrichment ratio using ORA by tissue. The input were the addition of the differentially expressed transcript in the condition comparison (gD vs. gN) and the family comparisons (gD1 vs. gN + gD2 vs. gN + gD vs. gN1 + gD vs. gN2). A) Brain, B) White muscle and C) Vertebra. All the categories showed a p-Value < 0.05.

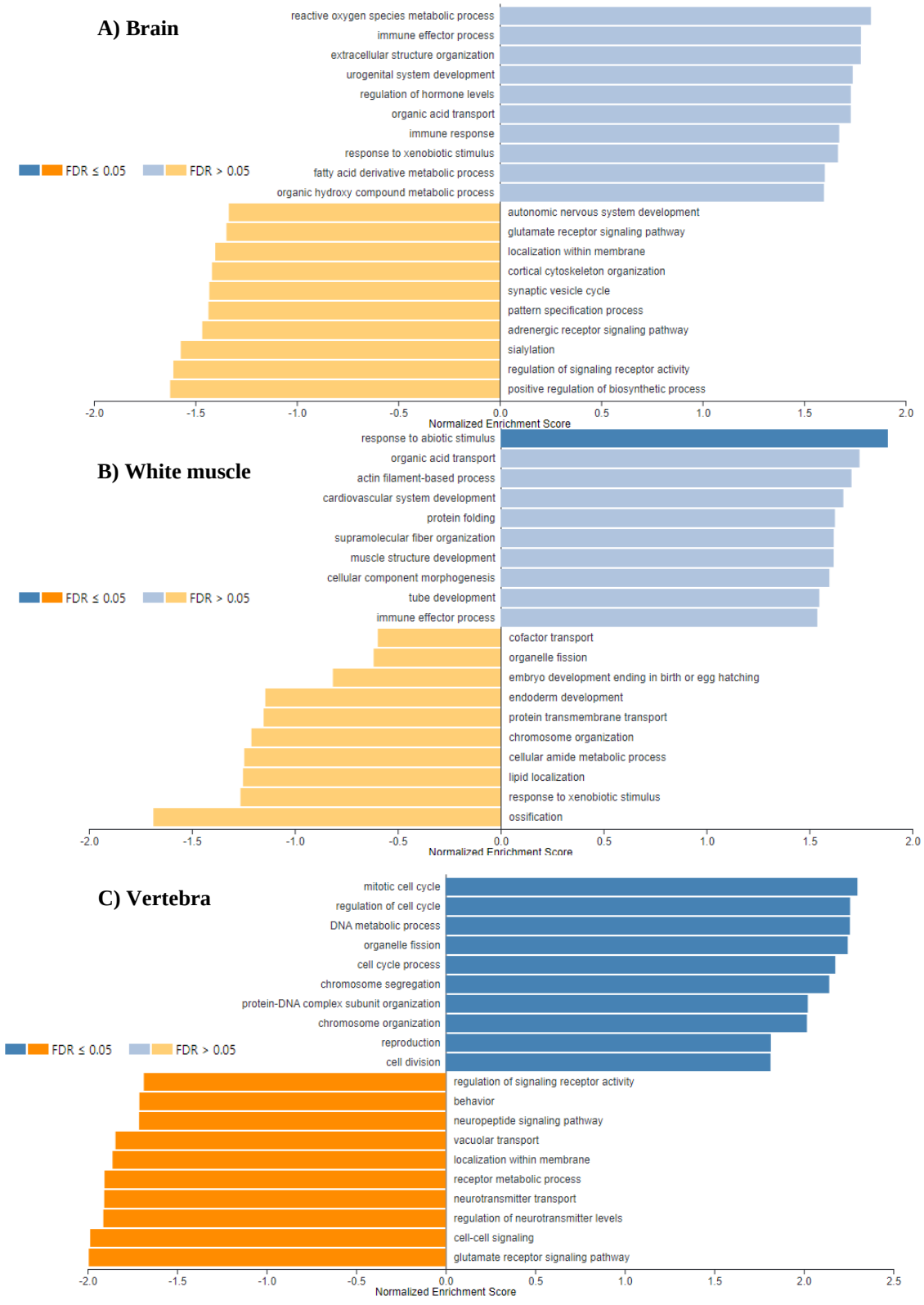


*Figure 12. Top 10 GO-BP enriched categories of gN and gD groups sorted by normalized enrichment score using GSEA by tissue. The input was the expression values of the condition comparison (gD vs. gN) per tissue. A) Brain, B) White muscle and C) Vertebra. Bright colours represent and FDR < 0.05, orange an overexpression in gD group and blue in gN group. * are used to mark categories with a p-value < 0.05.*

3.4 Discussion

In this study, the number of expressed transcripts varies greatly between tissues, having brain and vertebra higher numbers of transcripts than white muscle, 1.9 and 1.8 fold higher, respectively. Similar results have been reported in trout, having the brain between 1.4-3.9 times more transcripts than white muscle, but reporting less proportion of tissue specific genes (Salem et al., 2015). On the other hand, the same number of transcripts were expressed in brain and vertebra, suggesting similar transcriptome complexity, as it was reported in human (Jongeneel et al., 2005).

There are 2 DEG between the gD and gN groups (*f13a1* and *TRIM39l*). *f13a1* codify for the coagulation factor XIII-A with has a known role in blood clotting, but is likely involved in other aspects including bone formation angiogenesis. Thus, in mouse, the circulating factor is produced predominantly in bone marrow cells but is also produced in other tissues by monocytes, macrophages, chondrocytes, osteoblasts, osteocytes, and preadipocytes (Myneni et al., 2016). In these cell types, FXIII-A can regulate cell differentiation and extracellular matrix deposition (Myneni et al., 2016). But in seabream, this gene was reported to be overexpressed in cartilage, in a study comparing the transcriptional responsiveness of bone and skeletal muscle stimulated with a flooding dose of leucine on the development of skeletal deformities (Garcia de la serrana et al., 2016). On the other hand, *TRIM39l* is an *E3 ubiquitin-protein ligase like*, member of the tripartite motif family with a not well identified function, but it is related with immune system, cell homeostasis control, including mediating apoptosis, cell cycle, and cytostasis (Zhang et al., 2012; Wang et al., 2016).

There were a total of 59 DET in condition comparison (gD vs. gN), 702 taking into account the between families comparison. Of them, only an average between the three tissues of 33.4 % transcripts had a homolog in zebrafish and were annotated to the GO-BP functional database and 11.7 % are uncharacterized genes. This apparently high number of uncharacterized transcripts is lower than those reported in a study in the blackstripe livebearer (*Poeciliopsis prolifica*) reporting a percentage of 79.4 % of not identifiable highly expressed

transcripts (Jue et al., 2018). So, further research is needed in this field to really know the extent of the genes found in studies like these, using RNAseq tool.

Despite the lack of functional information for most of the genes, the vertebra was the only tissue in which differentially expressed transcripts showed functional enrichment by GSEA analysis with an FDR < 0.05 in GO-BP functional database. The categories enriched in vertebrae are mostly related to the *regulation of the cell cycle*, and *nervous and cell signaling*, specific categories compared with the categories variability showed in bone in other study in carp (Luo et al., 2021). On the other hand, muscle showed a functional category and brain none, with these statistical restrictions, being surprising that the only category that was overrepresented with a p-value < 0.05 in the gD group was *ossification*.

On the other hand, of the 702 DET found in family analysis, 69.6 % comes from vertebra tissue analysis, 16,6 % from muscle and 13.8 % from brain, suggesting that bone expression patterns plays an important role in the deformity ontogeny. Riera et al. 2019b reported in gilthead seabream using vertebra samples that the pattern of expression of genes related to *extracellular matrix maturation and mineralization* and *bone resorption* seems to be key to discern between normal and malformed fish in lordosis and LSK.

In conclusion, in this study a battery of 59 DET between gD and gN (702DET between family comparisons) using 3 different tissues between animal breeders with divergent genetic prevalence for the deformity is reported. These transcripts can be used to confirm their expression level in a larger number of individuals and families or even to study the existence of markers, such as SNPs, that can be used to identify deformed animals.

Chapter 4

The Effect of the Deformity Genetic Background of the Breeders on the Spawning Quality Seabream (*Sparus aurata L.*)

Álvaro **Lorenzo-Felipe**[1], Hyun Suk **Shin**[1], Sergio **León-Bernabeu**[2], Cathaysa **Pérez-García**[1], María Jesús **Zamorano**[1], Jaume **Pérez-Sánchez**[3] and Juan Manuel **Afonso-López**[1*]

[1] Aquaculture Research Group (GIA), Institute of Sustainable Aquaculture and Marine Ecosystems (IU-ECOAQUA), Universidad de las Palmas de Gran Canaria (ULPGC), Telde, Spain, [2] AQUANARIA S.L., San Bartolomé de Tirajana, Spain, [3] Nutrigenomics and Fish Growth Endocrinology, Institute of Aquaculture Torre de la Sal (IATS-CSIC), Castellón, Spain.

Published at *Frontiers in Marine Science* (https://doi.org/10.3389/fmars.2021.656901).

Impact factor 2020 (JCR): 4.912

Abstract

Fish egg quality is strongly related with the ability of the egg to be fertilized and develop a normal embryo with good survival and a lack of abnormalities. Large variations in the spawning quantity or quality impact directly in the competitiveness and sustainability of hatcheries, which create an overly large broodstock in order to satisfy the on-growing companies' demand for undeformed fry. The present study reports, for the first time in relation to gilthead seabream, the effect of the genetic background of breeders for presence or absence of deformity on their spawning quality and the importance of considering this when creating broodstock. The spawning quality of crosses of breeders with genetic background for presence or absence of deformity (EBVdef), were evaluated during a whole spawning season, through study of the following traits: oocyte yield, fertilization rate, viability rate, hatching rate, larval survival rate, fertilized eggs, viable eggs, hatched eggs, and number of alive larvae. Breeders with a genetic background for deformity and a normal phenotype had shorter spawning periods, lower oocyte yield and, consequently, produced a lower number of alive larvae. In these two traits, the genetic background of breeders was of greater importance during intermediate spawning periods, when spawning is generally considered optimal for the industry, while environmental factors were more important at the beginning and end of the spawning season. In conclusion, these results demonstrate the importance of controlling the breeders' genetics when creating broodstock.

Keywords: gilthead seabream, estimated breeding value, skeletal deformity, spawning quality, oocyte yield, viability rate, number of alive larvae

4.1 Introduction

The gilthead seabream (*Sparus aurata*) is a demersal and eurythermal species with a large reproductive capacity. It is found in subtropical coastal areas, from the Eastern Atlantic to the Mediterranean (Froese & Pauly, 2019). In the wild spawning starts in October in the Gulf of Cadiz or December-January in eastern Mediterranean and lasts for 3–4 months (Arias, 1980; Kissil et al., 2001). Every female releases over two million eggs per kilogram they weigh (APROMAR, 2017). In aquaculture, larval rearing of the gilthead seabream is carried out at 16–24 ºC (Tandler et al., 1989; Koumoundouros et al., 1997a,b; Shields, 2001), an optimal. Temperature for the hatching rate and larval survival rate at the yolk sac resorption stage (Polo et al., 1991). The wide natural distribution and environmental conditions experienced by the gilthead seabream have made it one of the most important species in the European marine fish farming and production is still expanding worldwide. In a European context, gilthead seabream was fifth in terms of annual production, at 91,964 t, and fourth in terms of economic value (APROMAR, 2020).

Breeding programs are essential tools that contribute to the long-term development of aquaculture in order to improve the quality of the farmed strains. Aquaculture production based on genetic selection schemes has increased from 1 % in 1997 to 80–83 % in 2017 (Gjedrem, 1997, 2004, 2012; Janssen et al., 2017). In gilthead seabream, genetically selected seed production only accounts 31–44 % of total fry in the European market (Chavanne et al., 2016), with an estimated increase in growth of 5–29 % (Knibb, 2000; Brown, 2003; Thorland et al., 2015). However, in this species quality of morphology and growth traits are deemed to have the same level of importance in the industry (Janssen et al., 2017) as the fish are sold whole and due to the loss of market value throughout the value chain (Bardon et al., 2009; Boglione & Costa, 2011; Boglione et al., 2013).

Skeletal deformity is the most important trait when defining the morphological quality of the fish, reducing the physiological ability of fish for correct development i.e., reducing their growth rate, increasing their mortality rate and significantly affecting their welfare (Andrades et al., 1996; Karahan et al., 2013; García-Celdrán et al., 2016). The presence of

deformities can affect up to 30 % of production, causing an annual loss of more than 50 M€ in the European aquaculture industry (Castro et al., 2008; Fernández et al., 2008; Haga et al., 2011). In gilthead seabream, deformity appears early in development and its prevalence increases with age (Lee-Montero et al., 2015).

Abiotic, biotic, xenobiotic, nutritional and environmental factors have been studied as the causes of deformities (Afonso & Roo, 2007), as well as genetic factors or interactions (Andrades et al., 1996; Afonso et al., 2000; Fernández et al., 2008). Although environment has a large effect on deformity, different families or cohorts have shown a variable response to specific environmental stressors, suggesting additive genetic variation in these traits (Kause et al., 2007; Lee-Montero et al., 2015; García-Celdrán et al., 2016), the presence of QTLs (Negrín-Báez et al., 2015a,b) or inbreeding depression as a consequence of the consanguinity level (Aulstad & Kittelsen, 1971; Kincaid, 1983). In gilthead seabream, segregation of a major gene (Astorga et al., 2004) and polygenic inheritance (Afonso et al., 2000; Lee-Montero et al., 2015) have been suggested as explanations for the incidence of skeletal deformities. Negrín-Báez et al. (2015b) reported a significant excess of descendants with severe deformities from directed crosses involving parents with the same deformities, suggesting the elimination of deformed fish from a breeding nucleus or the inclusion of this trait in breeding programs.

In sparid species, evidence of mineralization can be discerned from 30 dph (Socorro, 2006), where the temperature-dependent differential expression of genes associated with bone formation suggests the long-term regulation of osteogenesis (Riera-Heredia et al., 2018). In gilthead seabream, jaw and mouth bones are formed after the onset of exogenous feeding (Faustino & Power, 2005). In fact, additive components already appear from the larvae (Fragkoulis et al., 2018, 2020) and fingerling stages (García-Celdrán et al., 2016).

Therefore, a major issue in the aquaculture industry is the production of a large number of eggs through controlling gamete quality in order to satisfy market demand for seeds through breeding programs (Bobe & Labbé, 2010; Migaud et al., 2013). Fish egg quality, measured through traits such as fertilization rate, oocyte yield and alive offspring number, is considered

very important in accordance with European scientific criteria (AQUAEXCEL, 2013), including good survival and lack of abnormalities (Kjørsvik et al., 1990; Brooks et al., 1997). Gametes, both in terms of their quantity and quality, are closely associated with the competitiveness and sustainability of hatcheries and farms (Bromage et al., 1992; Migaud et al., 2013; Theodorou et al., 2016).

The spawning quality of breeders depends of many factors related with nutrition, genetics, environmental conditions and any stress (Brooks et al., 1997; Fernández-Palacios et al., 1997; Almansa et al., 1999; Bobe & Labbé, 2010; Scabini et al., 2011; Jerez et al., 2012; Migaud et al., 2013; Bobe, 2015). Many traits related to reproduction, survival, morphology and growth (Kjørsvik et al., 1990, 2003; Bromage et al., 1992; Fernández-Palacios et al., 1995; Brooks et al., 1997; Shields et al., 1997; Lahnsteiner & Patarnello, 2004, 2005; Lahnsteiner et al., 2008; Bobe & Labbé, 2010; Migaud et al., 2013; Bobe, 2015) have been proposed for the characterization of spawning quality in different species. At the industrial level it is important to start fattening a suitable quantity and quality of animals. Spawning quality can affect the success of juveniles in the completion of internal metamorphosis (Kjørsvik et al., 2003), survival, and the appearance of deformities (Bonnet et al., 2007).

The aims of this study were to research, for the first time in gilthead seabream, the additive genetic component effect of skeletal deformities on spawning quality and to establish the main traits for evaluating reproductive ability at an industrial level. For this purpose, estimated breeding values for the presence or absence of deformity when the breeders were a commercial size (EBV_{def}) was calculated and their reproductive ability traced throughout an entire spawning season.

4.2 Material and methods

All animal experiments were conducted in accordance with the European Union Directive (2010/63/EU) on the protection of animals for scientific purposes, at the Fundación Canaria Parque Científico Tecnológico (FCPCT) of the Universidad de Las Palmas de Gran Canaria (Canary Islands, Spain).

4.2.1 Biological Material and Experimental Design

A total of 4,108 gilthead seabream adults from the Canary Islands region and the 3rd generation of the Spanish National Breeding Program (PROGENSA®-III) (Lee-Montero et al., 2015) at AQUANARIA S.L. were visually asessed for the presence or absence of deformity at commercial size (ATOL:0000087), in accordance with AquaExcel-ATOL (AQUAEXCEL, 2013). The genotyping of all animals evaluated for morphology was also established through the molecular characterization of microsatellite markers (SMsa1), with PCR methodology, following the instructions of its authors (Lee-Montero et al., 2013). The genetic relationship between descendants (study population) and their parents (parental assignment) was reconstructed, comparing breeders' and descendants' genotypes by using *Vitassign* software (Vandeputte et al., 2006). Genetic parameters were estimated for both data matrix (phenotypic and genotypic information) by using the restricted maximum likelihood method with *VCE 6.0* software (Neumaier & Groeneveld, 1998; Groeneveld et al., 2010). A stock of breeders was then preselected on the basis of its EBVdef for conditioning and sexing before the spawning season, and moved to the FCPCT-ULPGC facilities. Fish with a deformed phenotype were discarded.

The EBVdef of the 4,108 gilthead seabream stock ranges between -0.10515 for normality and +0.07821 for deformity (Figure 13), with an average value of -0.01551 and a standard deviation value of 0.02681. After discarding the dead and misshapen, a total of 32 breeders were selected on the basis of their EBVdef, gender, weight and relationship coefficient. In this way, it was possible to establish two groups of breeders with opposite values in their EBVdef, which corresponds to a value of +0.05443 in the group of genetically deformed (gD) breeders, and -0.05964 in the group of genetically normal (gN) breeders. Between EBVdef values of gD and gN selected breeders, was contained almost 96% of the evaluated population (Figure 13).

Two different types of crosses, with different EBVdef, were established in 2 m^3 cylindroconical tanks, with duplicates to minimize the environmental factors. The first was a cross of gN fish (N1 and N2) and the second a cross of gD fish (D1 and D2), with the same

intensity of selection but reverse EBVdef (EBVdef = -0.05 vs. EBVdef = +0.05, respectively). The crosses were established by taking into account the EBVdef, relationship coefficient, biomass, and the weight ratio of males and females, so that the duplicates were similar and the differences between crosses were exclusively due to the EBVdef. The biomass sex-ratio (kg male/kg female) was similar between tanks; 1.43, 1.3, 1.18, and 1.1 for N1, N2, D1, and D2, respectively, with a mean of 1.28 kg male/kg female, as recommended by Fernández-Palacios et al. (1995, 1997). All breeders had normal phenotypes and the same age, between and within crosses, aiming to avoid the impact of females age on spawning quality (Jerez et al., 2012), by keeping the biomass sex-ratio constant. Thus, breeders' maturation and spawning occurred spontaneously under natural conditions. The coefficient of relationship, EBVdef, total biomass, male and female biomass, number of males and females were measured in breeders in the four experimental tanks (Figure 14). Tanks were supplied with 16 L/min seawater at 7.98 ± 0.10 pH and kept under natural light photoperiod (11–13 h light). Breeders were fed a commercial diet (Skretting ARC, Stavanger, Norway).

During the spawning season, the monthly average temperature decreased between December and February from 20.9 to 19.9 ºC, followed by an increase to 22.8 ºC in June. Food intake, expressed as a percentage of biomass, remained constant throughout the experiment with a range from 0.1 to 1.1 % and an average value of 0.4 % in all tanks.

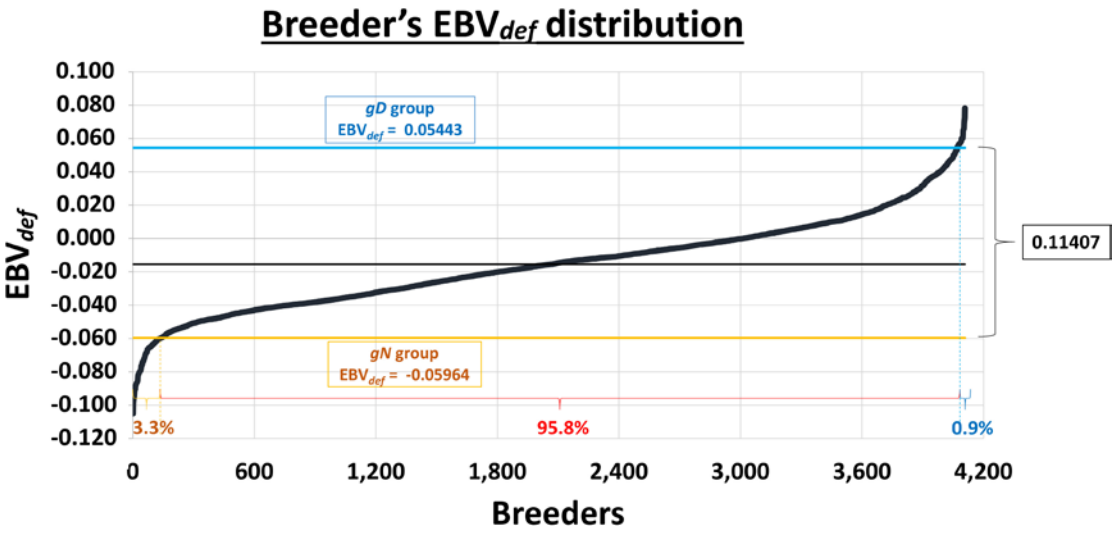
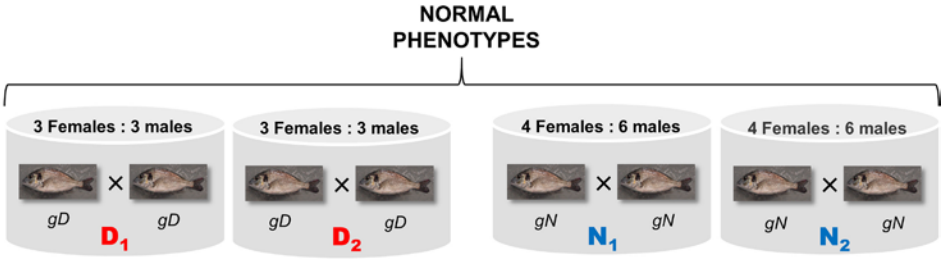


Figure 13. The distribution of 4,108 breeders from the 3rd generation of the Spanish National Breeding Program (PROGENSA®-III) according to their estimated breeding value for the presence or absence of deformity at commercial size (EBVdef). In yellow is the mean EBVdef of the genetically normal (gN) preselected breeders and the percentage of breeders with a lower value. In blue is the mean EBVdef of the genetically deformed (gD) preselected breeders and the percentage of breeders with a higher value. In red is the percentage of breeders with intermediate EBVdef.



a	0.2777	0.2829	0.2461	0.2532
EBV_{def}	+ 0.0553	+ 0.0536	- 0.0596	- 0.0596
Biomass	8,065 g	8,665 g	8,235 g	8,000 g
♀Weight	3,710 g	4,130 g	3,375 g	3,490 g
♂Weight	4,355 g	4,535 g	4,860 g	4,510 g

Figure 14. Experimental design representation that shows the two genetic groups (genetically deformed, gD; genetically normal, gN) with their duplicates (D1, D2 and N1, N2), the number of breeders per tank and their data: kinship (a), estimated breeding value for the presence or absence of deformity at commercial size (EBVdef), biomass, female weight (♀weight), and male weight (♂weight).

4.2.2 Spawning Quality Traits

During the spawning season, eggs were collected daily (20 h after release and fertilization) from each tank overflow in semi-submerged 500-µm mesh net egg collectors placed in a nearby incubation tank. The oocyte yield (ATOL:0001723), *fertilization rate* (ATOL:0001775), *viability rate* (ATOL:0001531), *hatching rate* (ATOL:0001531), *larvae*

survival rate at yolk sac resorption (ATOL:0001531) were measure daily in accordance with AquaExcel-ATOL (Kjørsvik et al., 1990; Fernández-Palacios et al., 1995, 1997; AQUAEXCEL, 2013).

Samples of eggs were collected and observed under a binocular stereo microscope. The *oocyte yield*, *fertilization rate* and *viability rate* were estimated volumetrically after samples were counted. In addition, two 96-well plates were used to hold 192 floating eggs (1 per well) with 200 mL filtered seawater and incubated at 19 ºC. After 24 and 96 h, the number of alive larvae was counted in order to calculate the hatching rate (*number of hatched larvae*/*number of fertilized eggs*) and the survival larval rate (%).

The *number of fertilized eggs*, *viable eggs*, *hatched eggs* and the *number of alive larvae* were also calculated as associated traits, since they represent the ultimate aim of the industry, a collection of quantitative and qualitative data that determines the final number of larvae available for rearing.

4.2.3 Statistical Analysis

The data were tested for normality using the one sample Kolmogorov–Smirnoff test, as well as for homogeneity of variance using Levene's test. When a normal distribution and/or homogeneity of variance was not achieved, data were subjected to the Kruskal–Wallis non-parametric test (Zar, 1984). To study the variation of spawning quality traits throughout the whole spawning season, comparisons between types of crosses (gD vs. gN) and between experimental tanks (D1, D2, N1, N2) were carried out, using the day as a unit of measurement (Table 6). The variation of spawning quality traits was also studied in short time periods, structuring the whole spawning season into twelve fortnights. Each spawning quality trait was compared between types of crosses (gD vs. gN) within each fortnight. Additionally, each trait was compared within types of crosses (gD or gN) between fortnights (Table 7). The following general linear model was used:

$$Y_{ijk} = \mu + \alpha_i + \beta_{ij} + \varepsilon_{ijk}$$

Where, μ is the mean of the population, α_i is the effect of the genetic factor (types of crosses), β_{ij} is the effect of the tank factor within genetic factor, and ε_{ijk} is the residual error.

For the *oocyte yield* and *number of alive larvae* traits, within fortnights, the influence of genetic and tank factors was measured by the following statistical algorithm: execution of a standard regression analysis, and estimation of variance components of principal factors. Within fortnights, the same statistical algorithm was also applied to explain the *number of alive larvae* trait using as predictors the *oocyte yield*, *fertilization rate*, *viability rate*, *hatching rate* and *larvae survival rate* traits. Statistical analysis was performed using the *SPSS package, version 22.0* (SPSS Inc, Chicago, United States). All statistically significant comparisons of this work, Tables 13, 14, include the correction factor of Bonferroni's to minimize the Type-I error (Gordon et al., 2007; see Supplementary Material for differences between Bonferroni and Benjamini-Hochberg tests). Thus, significant differences were considered from 0.05, as confident level.

Table 6. Mean values of spawning quality traits throughout the spawning season (oocyte yield, fertilization rate, viability rate, hatching rate, larvae survival rate, fertilized eggs, viable eggs, hatched eggs, number of alive larvae) per Tank (D1, D2, N1, N2) and per Genetic group (gD, gN). Results are expressed as means ± SEM. For each trait, values in the same row followed by different letters are significantly different (P < 0.05), values between genetic groups are expressed with capital letters, and values within genetic groups between tanks with lower-case letters.

Trait	Tank				Genetic group	
	D1	**D2**	**N1**	**N2**	***gD***	***gN***
Oocyte yield (10^3)	4.7 ± 0.4	8.3 ± 0.6	38.2 ± 1.0^{a}	30.4 ± 0.9^{b}	6.5 ± 0.4^{B}	34.3 ± 0.7^{A}
Fertilization rate (%)	98.4 ± 0.3^{a}	90.8 ± 1.2^{b}	96.1 ± 0.4	95.4 ± 0.6	94.1 ± 0.8	95.8 ± 0.3
Viability rate (%)	44.9 ± 3.5	39.9 ± 3.0	49.9 ± 1.2^{a}	40.9 ± 1.7^{b}	42.0 ± 2.3	45.3 ± 1.1
Hatching rate (%)	96.1 ± 0.3^{a}	84.9 ± 1.1^{b}	89.2 ± 1.2^{a}	87.6 ± 0.8^{b}	89.8 ± 0.8	88.4 ± 0.7
Larvae survival rate (%)	82.3 ± 1.5^{a}	70.2 ± 2.1^{b}	82.2 ± 0.9	79.3 ± 1.2	75.4 ± 1.4^{B}	80.8 ± 0.7^{A}
Fertilized eggs (10^3)	4.6 ± 0.4	7.4 ± 0.5	36.8 ± 1.0^{a}	29.2 ± 0.9^{b}	6.0 ± 0.3^{B}	33.0 ± 0.7^{A}
Viable eggs (10^3)	2.0 ± 0.2	3.6 ± 0.4	18.7 ± 0.7^{a}	13.6 ± 0.8^{b}	2.8 ± 0.2^{B}	16.1 ± 0.5^{A}
Hatched eggs (10^3)	1.9 ± 0.2	3.0 ± 0.3	16.7 ± 0.6^{a}	11.9 ± 0.7^{b}	2.5 ± 0.2^{B}	14.1 ± 0.4^{A}
Nº alive larvae (10^3)	1.6 ± 0.2	2.1 ± 0.2	13.7 ± 0.5^{a}	9.5 ± 0.5^{b}	2.1 ± 0.2^{B}	12.3 ± 0.4^{A}

Table 7. Mean values of spawning quality traits throughout the spawning season, by fortnight (oocyte yield, fertilization rate, viability rate, hatching rate, larval survival rate, fertilized eggs, viable eggs, hatched eggs, number of alive larvae) per Genetic groups (gD, gN). Results are expressed as means ± SEM. For each trait, values in the same column followed by different lower-case letters are significantly different (P < 0.05) between genetic groups' crosses during the same spawning period. Values in the same row followed by different capital letters are significantly different (P < 0.05).

Trait	Genetic	14-31/12/2019		01-15/01/2019		16-31/01/2019		01-14/02/2019		15-28/02/2019		01-15/03/2019		16-31/03/2019		01-15/04/2019		16-30/04/2019		01-15/05/2019		16-31/05/2019		01-15/06/2019	
Oocyte yield (10^3)	*gD*	2.8 ± 1.3b	DE	11.5 ± 2.2b	ABC	13.2 ± 0.8b	A	11.8 ± 1.3b	AB	12.5 ± 0.9b	A	9.3 ± 0.7b	AB	7.4 ± 0.5b	ABC	5.0 ± 0.8b	BCD	2.8 ± 0.7b	CDE	0.1 ± 0.1b	E	0.0 ± 0.0b	E	0.0 ± 0.0b	E
	gN	39.1 ± 2.8a	AB	48.9 ± 2.3a	A	43.2 ± 1.7a	A	38.8 ± 1.5a	AB	37.8 ± 1.6a	AB	38.5 ± 1.0a	AB	36.6 ± 1.7a	AB	31.7 ± 1.1a	BC	26.5 ± 1.4a	CD	25.3 ± 1.4a	CD	23.8 ± 2.0a	CD	8.8 ± 2.0a	D
Fertilization rate (%)	*gD*	84.3 ± 3.9	B	86.3 ± 4.6b	AB	89.6 ± 3.3	AB	94.7 ± 1.4	AB	96.3 ± 1.0	A	95.3 ± 1.1	A	97.1 ± 0.6	A	96.8 ± 0.9	AB	97.9 ± 0.8	A	-		-		-	
	gN	88.7 ± 2.1	D	97.1 ± 0.4a	AB	96.7 ± 0.5	ABC	94.6 ± 2.0	ABC	97.3 ± 0.5	AB	96.9 ± 0.5	ABC	98.2 ± 0.3	A	98.2 ± 0.3	A	98.3 ± 0.2	A	97.3 ± 0.4	AB	95.0 ± 0.8	BCD	88.9 ± 2.0	CD
Viability rate (%)	*gD*	65.8 ± 6.3	ABC	68.9 ± 4.0	AB	78.6 ± 2.2a	A	50.9 ± 6.6	BC	22.7 ± 3.8b	CDE	42.1 ± 5.3b	BC	35.0 ± 4.7	BCD	11.7 ± 3.1b	DE	2.3 ± 0.6b	E	-		-		-	
	gN	50.1 ± 4.1	ABC	58.6 ± 4.2	AB	58.8 ± 1.8b	A	56.9 ± 2.0	AB	56.4 ± 1.6a	AB	55.1 ± 1.6a	AB	44.9 ± 2.4	BCD	33.7 ± 2.0a	ED	32.7 ± 3.2a	CDE	35.1 ± 2.8	CDE	24.0 ± 3.2	E	32.9 ± 7.2	CDE
Hatching rate (%)	*gD*	95.5 ± 2.0a		88.0 ± 3.0		93.3 ± 1.1a		87.6 ± 3.2		88.0 ± 1.8		88.1 ± 1.8		91.6 ± 1.6		92.0 ± 1.7		70.6 ± 10.2b		-		-		-	
	gN	80.4 ± 4.0b		80.1 ± 4.1		83.5 ± 3.0b		89.1 ± 1.2		85.7 ± 2.0		92.1 ± 0.9		92.7 ± 1.4		92.1 ± 0.9		91.6 ± 1.0a		92.9 ± 1.0		92.7 ± 0.8		94.9 ± 1.5	
Larval survival rate (%)	*gD*	93.6 ± 1.3a	A	83.7 ± 2.3	AB	89.8 ± 1.1a	A	79.9 ± 2.9	AB	64.3 ± 4.6b	B	70.3 ± 3.0b	B	69.1 ± 3.3b	B	59.0 ± 6.2b	B	39.6 ± 5.9b	B	-		-		-	
	gN	83.7 ± 1.5b	AB	83.9 ± 1.8	AB	85.9 ± 1.5b	A	85.6 ± 1.4	AB	77.6 ± 3.8a	AB	81.5 ± 2.4a	AB	80.7 ± 2.0a	AB	78.8 ± 2.6a	AB	72.2 ± 3.6a	B	75.4 ± 2.5	B	81.5 ± 1.8	AB	85.8 ± 2.4	AB
Fertilized eggs (10^3)	*gD*	2.3 ± 1.0b	DE	9.8 ± 1.9b	ABC	11.9 ± 0.8b	A	11.0 ± 1.1b	AB	12.1 ± 0.9b	A	8.8 ± 0.7b	AB	7.1 ± 0.5b	ABC	4.8 ± 0.8b	BCD	2.8 ± 0.6b	CDE	-		-		-	
	gN	35.5 ± 2.8a	AB	47.5 ± 2.2a	A	41.8 ± 1.7a	A	36.8 ± 1.6a	AB	36.9 ± 1.6a	AB	37.3 ± 1.0a	AB	36.0 ± 1.6a	AB	31.1 ± 1.0a	BC	26.1 ± 1.4a	CD	24.6 ± 1.3	CD	22.5 ± 1.8	CD	7.6 ± 1.7	D
Viable eggs (10^3)	*gD*	1.3 ± 0.6b	D	6.6 ± 1.3b	BC	9.6 ± 0.7b	A	5.6 ± 1.0b	AB	2.6 ± 0.4b	ABC	3.6 ± 0.5b	AB	2.2 ± 0.3b	BC	0.5 ± 0.2b	CD	0.0 ± 0.0b	D	-		-		-	
	gN	18.1 ± 2.0a	AB	26.8 ± 2.1a	A	24.6 ± 1.3a	A	21.5 ± 1.3a	A	20.8 ± 1.1a	A	20.5 ± 0.7a	A	16.2 ± 1.3a	ABC	10.3 ± 0.6a	CD	9.7 ± 1.2a	BCD	9.4 ± 1.1	CD	6.9 ± 1.2	D	3.2 ± 1.1	D
Hatched eggs (10^3)	*gD*	1.3 ± 0.6b	D	5.8 ± 1.1b	BC	8.9 ± 0.6b	A	4.8 ±0.9b	AB	2.2 ± 0.4b	AB	3.2 ± 0.5b	AB	2.0 ± 0.3b	BC	0.5 ± 0.1b	CD	0.0 ± 0b	D	-		-		-	
	gN	14.5 ± 1.6a	BCD	21.4 ± 1.6a	AB	20.4 ± 1.0a	A	19.3 ± 1.3a	AB	17.9 ± 1.0a	AB	18.9 ± 0.7a	AB	15.0 ± 1.1a	ABC	9.5 ± 0.6a	CD	9.0 ± 1.2a	CD	8.9 ± 1.0	CD	6.4 ± 1.1	D	3.0 ± 1.0	D
Nº alive larvae (10^3)	*gD*	1.2 ± 0.6b	D	5.0 ± 0.9b	BC	8.0 ± 0.6b	A	4.1 ± 0.7b	AB	1.8 ± 0.3b	AB	2.6 ± 0.4b	AB	1.6 ± 0.2b	BC	0.3 ± 0.1b	CD	0.0 ±0.0b	D	-		-		-	
	gN	11.7 ± 1.3a	BCD	17.1 ± 1.3a	AB	17.0 ± 0.8a	A	17.4 ± 1.2a	A	15.4 ± 0.9a	AB	17.4 ± 0.6a	A	13.8 ± 1.0a	ABC	8.7 ± 0.5a	CDE	8.4 ± 1.1a	CDE	8.3 ± 1.0	CDE	6.0 ± 1.0	DE	2.8 ± 0.9	E

4.3 Results

In all experimental tanks, breeders spawned daily between December 2018 and June 2019. The numbers of spawning days registered were 86, 116, 176, and 177 days for tanks D1, D2, N1, and N2, respectively. In Table 6, the mean values of all spawning quality traits are reported per tank and type of cross. In Table 7, the comparisons for all spawning quality traits are presented: within fortnight per types of crosses, and within type of cross per fortnights.

4.3.1 Oocyte Yield

Oocyte yield contribution among tanks, for both types of crosses, was generally higher and more reproducible in the middle of the spawning season and lower at either ends of it (Figure 15).

Across the whole spawning season the values between tanks were only significantly different within the gN group. Between genetic groups, the gN group reported a value 5.27 times higher than the gD group (Table 6). This significant difference was maintained across all fortnights (Table 7 and Figure 16A). Over the fortnights, within each genetic group both registered their best values in the first phase of the spawning season (fortnights two–five for gD, and fortnights one–seven for gN). Within fortnights *oocyte yield* was explained by genetic factor, from 29.1 % in fortnight twelve to 90.9 % in fortnight six (Figure 17A) and a mean value across fortnights of 59.6%. The tank factor explained less in the middle (0 % in fortnight six) and more at the beginning (44 % in fortnight two) and at the end (65 % in fortnight eleven) of the spawning season, with a mean value of 23.5 %.

4.3.2 Fertilization, Viability, and Hatching Rates

In the whole spawning season these three traits did not report statistical differences between genetic groups, while statistical differences were estimated within genetic groups between tanks (Table 6).

Within fortnights, statistical differences were detected between genetic groups in fortnight two for *fertilization rate*, in fortnights three, five, six, eight, and nine for *viability rate*, and in fortnights one, three, and nine for *hatching rate* (Table 7 and Figures 16B–D). The *fertilization rate* of gD was lower than for gN across the fortnights. The *viability rate* of gD

was higher than gN across fortnights at the beginning of the spawning season (fortnights one, two, and three) and lower than gN in the rest of the fortnights. The trend for *hatching rate* between gD and gN groups did not follow any regular pattern (Table 7).

4.3.3 Larval Survival Rate

For the whole spawning season, the gN group reported a value 7.2% higher than the gD group (Table 6). Within the same genetic group between tanks, only significant differences were found for the gD group.

Within fortnights, at the beginning of the spawning season (fortnights one-three), the gD group showed higher values than the gN group, a trend which was reversed for the rest of the spawning season (Table 7 and Figure 16E). Regarding the evolution of this trait across the fortnights, it was observed that the values of the gN crosses were generally constant throughout the spawning period, while in the gD crosses the values declined slightly at the end of the spawning period (Figure 16E).

4.3.4 Fertilized Eggs, Viable Eggs, Hatched Eggs, and Number of Alive Larvae

In the whole spawning season these traits all differed significantly between genetic groups, the gN group being almost 6 times higher than gD group (Table 6). Within the same genetic group between tanks there were significant differences only in the gN group. Within fortnights the gN group was superior throughout the whole spawning season (Table 7 and Figure 16F). Between fortnights, the gD and gN groups showed the same pattern in the distribution of values across all traits: low values at the beginning, high values in the middle and low values at the end (Table 7).

For the *number of alive larvae*, the explanation of its variability within fortnights due to genetic and tank factors is shown in Figure 17B. Thus, genetic factors were highest in the middle of the spawning season (from fortnights five to eight), ranging from 63 to 87 %. Conversely, the impact of the tank was high after the beginning of the spawning season, low in the middle and at its highest at the end (80–84 %). However, environmental factors (not

explained) were especially influential at the beginning (56–46 % in fortnights one to two, respectively) and at the end (66 % in fortnight twelve) of the spawning season.

Figure 17C shows the importance of the *oocyte yield*, *fertilization rate*, *viability rate*, *hatching rate*, and *larvae survival rate* on the *number of alive larvae*. It can be seen that *oocyte yield* and *viability rate* explain the majority of variance. The influence of *oocyte yield* was at its maximum between fortnights three and seven (70.8–89.1 %). Conversely, *viability rate* was at its maximum between fortnights nine and eleven (58–89.7 %). Taking all traits into account, the best linear model explained between 84.5 and 98.7 % of total variation of the *number of alive larvae*.

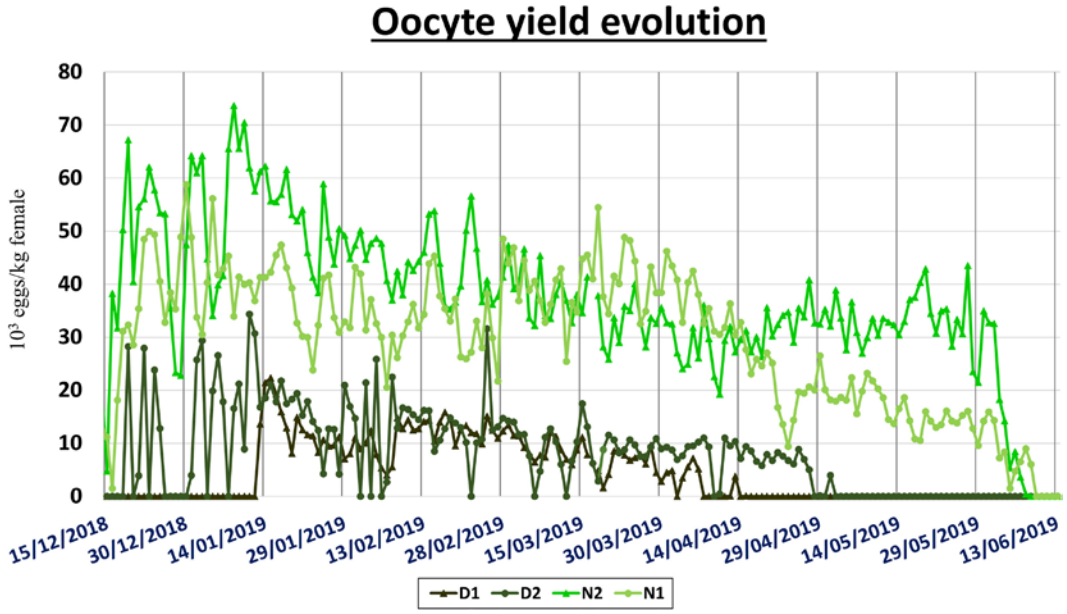


Figure 15. Evolution of the daily oocyte yield per tank (D1, D2, N1, N2) from the beginning to the end of the spawning season.

4.4 Discussion

Body malformation is one of the most important traits defining the quality of batches. Deformities have a negative effect on the profit of aquaculture companies, especially hatcheries, because on-growing companies do not accept deformed fry (Afonso & Roo, 2007). Most skeletal anomalies in marine aquaculture species occur during their embryonic, larval and metamorphosis stages (Koumoundouros, 2010), and poor spawning quality is often associated with the appearance of such deformities. In gilthead seabream, variation in the inbreeding of breeders produces differences in spawning quality (Astorga, 2005).

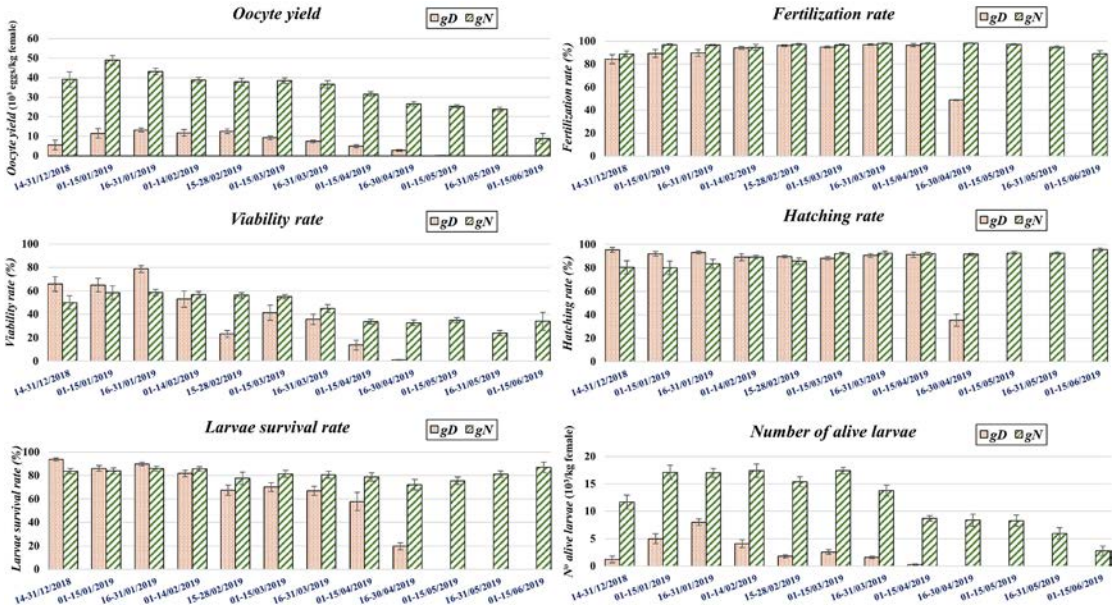


Figure 16. Evolution of the quality and quantity of spawning traits per genetic group (gD, gN) by fortnight, throughout the spawning season. Each capital letter corresponds to a trait: (A) oocyte yield, (B) fertilization rate, (C) viability rate, (D) hatching rate, (E) Larval survival rate, (F) Number of alive larvae.

In this study the effect of the EBVdef on spawning quality traits (*oocyte yield*, *fertilization rate*, *viability rate*, *hatching rate*, *larval survival rate*, *fertilized eggs*, *viable eggs*, *hatched eggs* and *number of alive larvae*) is reported for first time in gilthead seabream. Breeders started spawning naturally at the beginning of the species' spawning season in December-January and continued until April-June (Arias, 1980; Fernández-Palacios et al., 1995; Kissil et al., 2001; Ibarra-Zatarain & Duncan, 2015). The number of spawning days from all tanks ranged between 86 and 116 days for the gD group, and 176–177 for the gN group. These results agree with values reported by Jerez et al. (2012) (121–130 days of spawning). The difference in spawning days observed between both genetic groups was robust because it was reproduced in the same way within the genetic groups between tanks, suggesting at least a genetic effect based on the EBVdef.

The *oocyte yield* trait reported the greatest variability, showing differences between genetic groups (fed with commercial diets). In any event, their values ranged according to the locality and environmental conditions (Fernández-Palacios et al., 1995). While the *oocyte yield* of gN group was in line with Xu et al. (2019), *oocyle yield* of gD group was 5.3 times lower in comparison to the gN group for the same time period. These differences were kept within

fortnights. On the other hand, the gD group reported values lower than those reported by Jerez et al. (2012), and much lower than those reported by Xu et al. (2019), using diets with nutritional deficiencies. *Oocyte yield* trait was widely influenced by genetic factor throughout the whole spawning season, while environmental factors was only relevant at the beginning and end of the spawning season. In line with this, breeders with high genetic content in morphological deformity stimulate poor spawn quality through a lower level of *oocyte yield* and would be characterized for their EBVdef content.

Traits as *fertilization*, *viability*, and *hatching rates* are widely used for controlling spawning quality but in this study, they did not show significant differences between genetic groups, although viability rate and hatching rate estimates were, respectively, lower and similar to other studies (Fernández-Palacios et al., 1995, 1997; Scabini et al., 2011; Xu et al., 2019; Ferosekhan et al., 2020). The number of breeders and sex-ratio used in both types of crosses per tank (gN and gD groups) were in concordance with value used by Ibarra-Zatarain and Duncan (2015), in reproductive season of gilthead seabream. In fact, the fertilization rate registered was in line with values reported by other authors, also using mass spawning to study this species (Fernández-Palacios et al., 1995, 1997; Scabini et al., 2011; Jerez et al., 2012; Ibarra-Zatarain & Duncan, 2015), and higher than other research using mating with only one male or female (Gorshkov et al., 1997; Xu et al., 2019; Ferosekhan et al., 2020).

Larval survival rate reported in this study was within range for this species (Carrillo et al., 1989; Fernández-Palacios et al., 1995, 1997; Scabini et al., 2011; Xu et al., 2019; Ferosekhan et al., 2020) and the higher survival of the gN group compared to the gD group suggest a major susceptibility of the latter to environmental factors.

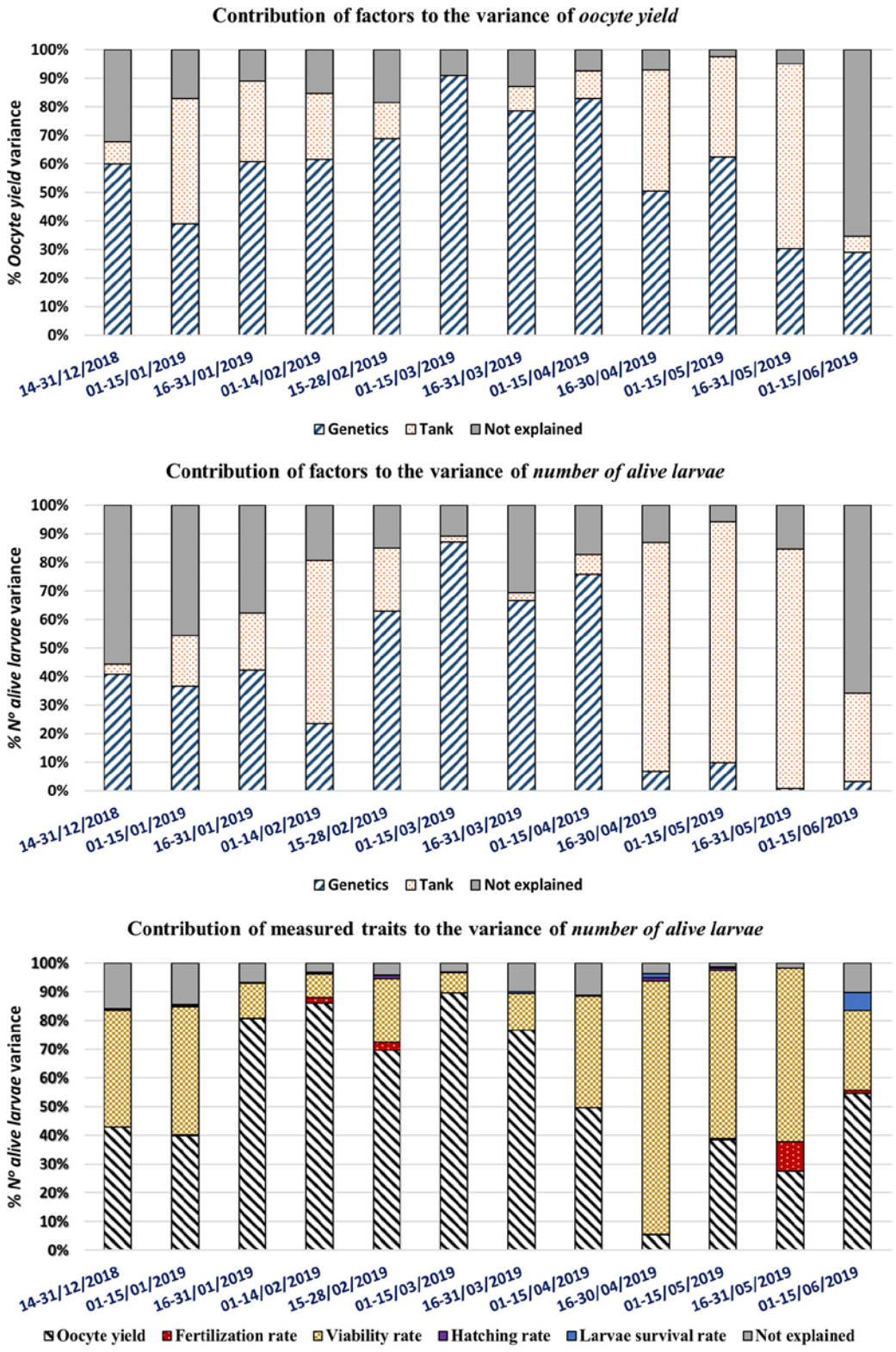


Figure 17. Histograms of the variance explained in the linear regression analysis by fortnight, throughout the spawning season. Each capital letter represents a different analysis: (A) the importance of environmental and genetic factors (Tank and Genetic, respectively) on the oocyte yield. (B) The importance of environmental and genetic factors (Tank and Genetic, respectively) on the number of alive larvae. (C) The importance of measured traits (oocyte yield, fertilization rate, viability rate, hatching rate, larval survival rate) on the number of alive larvae.

The large differences observed for *fertilized eggs*, *viable eggs*, *hatched eggs* and the *number of alive larvae* between genetic groups were similar along the whole spawning season, having significant differences between groups within fortnights (6 times higher in the gN group vs. the gD group). Moreover, between fortnights and independently of the genetic group, the same pattern in the distribution of values across these traits was reported: significantly low values at the beginning, high values in the middle, and low values at the end. *Oocyte yield* values described a distribution pattern practically equal but with differences of 5.3 times between groups, showing the high responsibility of it on all those traits, as expected.

The *number of alive larvae* is determinant for defining the effective production of hatcheries, in order to satisfy the on-growing companies' demand, which requires planning. It is therefore essential to know the spawning window within which the spawning quality is optimal. Genetic improvement is accumulative, permanent, and extendable to the whole production chain (Falconer & Mackay, 2001). In this study, the genetic factor showed a significant impact on the number of alive larvae (63–87 %), mainly in the middle of the spawning season, in line with the period when spawning is considered optimal at an industrial level (Navarro et al., 2009).

A linear regression modeling of spawning for number of alive larvae, revealed that *oocyte yield* and *viability rate* traits explained the majority of its variance, being maximum in the middle of the spawning season by the oocyte yield. These results are in concordance with the empiric suggestions reported by Fernández-Palacios et al. (1995).

In conclusion, this study demonstrates how the deformity genetic background of phenotypically normal breeders affects and explains their spawning quality, for the first time in gilthead seabream. Shorter spawning periods, lower *oocyte yield* and, consequently, a lower *number of alive larvae* were found among gD breeders. The *number of alive larvae*, the most determinant trait of spawning quality, peaked in the middle of the spawning season on the basis of *oocyte yield* and genetic factor, reflecting the importance of controlling the breeders' genetics and the utility of *oocyte yield* trait for tracking spawning quality. Conversely, the *viability rate*

and tank factor supported the number of alive larvae at the beginning and end of the spawning season. Given the results obtained, for oocyte yield trait and *number of alive larvae*, the gN breeders showed values 5.3 and 6 times higher than gD breeders, respectively. For future studies, it would be useful to estimate the differential gene expression patterns between breeders with different EBVdef values, as well as between their offspring, in order to discover the genetic structure involved in skeletal deformity determination, in gilthead seabream.

Chapter 5

Differential expression patterns in larvae of gilthead seabream (*Sparus aurata*) from divergent selection for the prevalence of skeletal deformities.

Álvaro **Lorenzo-Felipe**[1], Fernando **Naya Catalá**[3], Hyun Suk **Shin**[1], Josep **Calduch Giner**[3], Sergio **León-Bernabeu**[2], María Jesús **Zamorano**[1], Jaume **Pérez-Sánchez**[3] and Juan Manuel **Afonso-López**[1]

[1] Aquaculture Research Group (GIA), Institute of Sustainable Aquaculture and Marine Ecosystems (IU-ECOAQUA), Universidad de las Palmas de Gran Canaria (ULPGC), Telde, Spain, [2] AQUANARIA S.L., San Bartolomé de Tirajana, Spain, [3] Nutrigenomics and Fish Growth Endocrinology, Institute of Aquaculture Torre de la Sal (IATS-CSIC), Castellón, Spain.

In preparation

Abstract

The economic losses that the aquaculture industry has due to the presence of deformities is especially important in gilthead seabream due to it is sold as whole fish. Thus, this work proposed to verify the genetic architecture of morphological abnormalities trait using transcriptome tools in the offspring of two groups, genetically normal (gN) and deformed (gD), during different stages of larvae development. After the confirmation of deformity prevalence by osteology analysis, a total of 324 differentially expressed transcripts were found between groups, increased to 1,868 transcripts after a discriminant analysis approach that totally separated experimental groups at 35 dph. Gene ontology analysis highlighted that more than 50% of them participated in more than one of the over-represented categories. The major interaction was found between the *nervous system development* and *cellular developmental process* and between the processes of *organic substance transport* and the *regulation of signalling*. Finally after variant calling analysis, 64 discriminant and polymorphic genes between groups are proposed as candidates for future deformity studies. The results were consistent with previous deformity-associated studies, due to 8 genes presented overlapping to deformity QTL previously reported. This genes are *egfl7* (27 variants), *meis1* (13 variants), *supt20h* (5 variants), *kcnt1* (5 variants), *asrgl1* (5 variants), *mycp2* (2 variants), *psma2* (2 variants) and *palm* (1 variant).

Keywords: Gilthead seabream, estimated breeding value, skeletal deformity, RNAseq, gene expression analysis, larvae

5.1 Introduction

Gilthead seabream (*Sparus aurata* L.) is the most important marine fish farming of Mediterranean countries with a total production of 278,199 t in 2020 and a positives annual variation rates for production and economic value (7.2 % and 10.22 %, respectively) (APROMAR, 2021).

Despite the development and consolidation of the gilthead seabream aquaculture industry, it is estimated that the presence of deformities can affect up to 30 % of production, causing an annual loss of more than 50 M€ in the European aquaculture industry (Castro et al., 2008; Fernández et al., 2008; Haga et al., 2011). In this species, the most economically significant skeletal deformities include severe abnormalities of the opercular complex, neurocranium, vertebral column and appendicular skeleton abnormalities, given that fish are sold mainly as whole fish and customers rarely accept fish showing malformations (Koumoundouros et al., 1997; Boglione et al., 2001; Lee-Montero *et al.*, 2015; Boursiaki et al., 2019). In fish, the first stages of development, such as embryonic, larval and metamorphosis stages, are the most critical periods in which abnormal development can occur, giving rise to deformities that can worsen with age and define the quality of the final product (Koumoundouros, 2010). Most of skeletal deformities occurs at this periods, such as operculum and jaw deformities or column deviations (Boglione et al., 2013b). So, in marine aquaculture a prevalence of deformities around 20 % at the end of the hatchery phase is consider a good result (Boglione et al., 2013a).

Although there is a strong environmental component as a cause of the presence of deformities in fish, skeletal deformities have been associated to genetic basis as well (Andrades et al., 1996; Afonso et al., 2000; Castro et al., 2008; Fernández et al., 2008; Boglione et al., 2013b). So, in gilthead seabream, European breeding programs give the same importance to morphology and growth performance traits, in terms of frequency (Janssen et al., 2017).

In this species, different types of deformities had been associated with genetic bases and family association (Andrades et al., 1996; Afonso et al., 2000; Negrín-Báez et al., 2015b). Also, sustancial heritabilities for deformities and morphological traits has been described (normally, range between 0.15-0.56) (Astorga et al., 2004; Lee-Montero et al., 2015; Negrín-Báez et al., 2015b; García-Celdrán et al., 2015, 2016; Fragkoulis et al., 2018, 2020, 2021) and

high genetic correlation between different ages (0.83-0.99) (Lee-Montero et al., 2015). On the other hand some QTLs for different deformities (vertebral fusion, lordosis, jaw, LSK and operculum) has been reported in different studies (Negrín-Báez et al., 2015a, 2016a, 2016b)

In this study, in the framework of the Spanish National Breeding Program (PROGENSA®) and the European project PerformFISH, the EBVdef at commercial size of breeders of the 3rd generation of the breeding program has been estimated. Crosses have been forced to check whether this EBVdef also has the effect of causing a higher proportion of skeletal deformities in the offspring during larval development and to check in which window of development the deformities appear.

5.2 Materials and Methods

5.2.1 Experimental design and biological material

The breeders come from the 3rd generation of the National Spanish Breeding Program PROGENSA®-III (Lee-Montero et al., 2015; León-Bernabeu et al., 2021), from a base population with an estimated heritability for any type of deformity that range between 0.07-0.26 and a genetic correlation between ages between 0.83-0.99 (Lee-Montero et al., 2015). The presence of any deformity was visually assessed in 6,000 gilthead seabreams in accordance with AquaExcel-ATOL (ATOL: 0000087) (AQUAEXCEL, 2013). Also, caudal fin samples were collected for connecting information between animals through genealogical reconstruction using molecular tools. The parental assignment was established by using the microsatellite multiplex (SMAS1) (Lee-Montero et al., 2013), genotyping with the *GeneMapper* software (v5.0) (Applied Biosystems) and comparing the genotypes with the software *Vitassign*, to incorporate the coancestry data and to take care of the kinship (Vandeputte et al., 2006). Evaluation process was carried out using BLUP methodology with the software *VCE 6.0* (Groeneveld et al., 2010; Neumaier & Groeneveld, 1998) and a complementary program to help for managing input data and automatize processes (*VCE_Executer, v3.0*) (León-Bernabeu et al., 2021).

Using this data two different mates were established, one with gN fish and another with gD fish and the highest EBVdef range between them but similar kindship, biomass and proportion of females and males, in weight (Figure 14). Phenotypically deformed fish were

discarded to simulate the real situation in the Hatchery, where these kind of fish are discarded as broodstock. Maturation and spawning occurred spontaneously under natural conditions in December. For each mate larvae were rearing until 73 days per triplicate to analysis the deformities in different tanks, minimizing the environmental factor.

5.2.2 Larval rearing

Larvae were reared per triplicate under controlled conditions using the protocols established in EcoAqua Institute facilities, after checking a good spawning quality (enough *oocyte yield*, a good *fertilization* and *viability rates* and a period with high *hatching* and *larvae survival rates*). *Viable eggs* were transferred to 500 L larval rearing tanks with an initial rearing density of approximately 100 eggs/L.

The larval rearing tanks had the water inlet from the bottom and the outlet at the top of the tank wall connected to a filter with a mesh size adaptable to the size of the larvae. They also have an aeration system in the center and an air blower skimmer. Temperatures and dissolved oxygen concentrations were measured daily, ranged between 19-22 °C and 6–12 ppm, respectively. The water renewal was adjusted from 10 to 40 % of the volume per h until the establishment of microdiet feeding (~46 dph). The 'green water' technique was applied during the period 3–28 dph by using a phytoplankton strain (*Nannochloropsis sp.*) (250 ± 100 × 103 cells/mL). Regarding feeding, larvae were fed twice a day with rotifers (*Brachionus plicatilis*) enriched with ORI-GREEN (Skretting). Rotifer density was progressively increased from 5 to 10 rotifers per mL. Artemia nauplii (EG, INVE, Belgium) were offered twice a day to larvae from 17 to 22 dph, in increasing density from 0.5 to 2 nauplii per mL, and 2 days enriched-metanauplii from 20 to 40 dph (1 to 5 metanauplii per mL). From 36 dph to the end of the experiment (73 dph), larvae were progressively weaned onto dry feed, with Gemma Micro (size range: 75 to 500 μm; Skretting, Spain) with manual feeding each hour from 08:00 to 16:00 and automatic feeding 3 times from 17:00 to 19:00, maintaining a feeding with enriched-metanauplii at 12:00.

5.2.3 Sampling procedures

The sampling was carried out daily from the egg phase until 35 dph and alternate days from 35 to 73 dph for osteology and RNAseq analysis. Samples were sacrificed by thermal shock by immersing them in ice, washed with DEPC-treated water and stored with 4 % formalin

buffered at 7.2 pH in a 1:10 ratio (sample:solvent) for osteology analysis and were conserved in RNAlater (Sigma-Aldrich) in a ratio 1:5 (sample:solvent), kept overnight at 4 ºC and stored at -80 ºC for RNAseq analysis. During the 'green water' period the larvae were collected by taking water samples at different points in the tank, since the larvae are homogeneously distributed. After this period, larvae were collected with the help of a net, quickly, during the feeding time, also taken the samples in different points of the tank and taken all the size represented in the tank to obtain a representative sample. For osteological analysis in each sample point (day and tank) a minimum of 10 larvae were taken, and 10 larvae or 100 mg of sample for RNAseq analysis.

5.2.4 Osteology analysis

For the examination of skeletal deformities, a mean of 22 formalin fixed specimens per condition (high and low EBVdef) and per sampling day from day 35 were stained for bone and cartilage using a double-staining protocol with some modifications (Potthoff, 1984; Taylor & Van Dyke, 1985). Specimens were washed in distilled water to eliminate formalin, followed by a bleaching step in peroxide (3 %) and KOH (0.5 %). Then, specimens were dehydrated through a graded series of ethanol (25–95 %) and stained with Alcian blue 8GX (0,02 % in 70 % alcohol and 30 % glacial acetic acid) and neutralized using a 1 % aqueous solution of KOH until skeletal elements were clearly visible. Later, series of trypsin digestion (35 % saturated sodium borate, 65 % distilled water and trypsin powder) and staining with Alizarin red S (stock solution: 1 % Alizarin red in 1 % KOH). Finally, samples were cleared by using trypsin and KOH (1 %) and preserved in glycerine with thymol. Staining time was variable and depended on the size of the specimen.

In accordance with the degree of mineralization, samples from 39 dph to 73 dph were used for the examination of severe skeletal deformities. In total a number of 808 specimens were evaluated, 453 of low EBVdef and 355 of high EBVdef. The analysis of deformities was based on the normal osteological phenotype of the species (Divanach et al., 1996; Faustino & Power, 1998; Fragkoulis et al., 2018; Koumoundouros et al., 1997a) and on the existing descriptions of skeletal deformities in marine fish larvae (Boglione et al., 2013a, 2013b, 2001; Boglione & Costa, 2011; Divanach et al., 1996).

Larvae samples were analysed under a light microscope (CX41, Olympus, Tokyo, Japan). Each section was individually analyzed (Skull, cranial and pre-haemal vertebrae, haemal and caudal, the pectoral, dorsal, anal and caudal fins).

5.2.5 Statistical analysis

The association between factors (mating, tank, dph, length) and deformity within experiment was analysed by a Loglinear Model by using the statistical software *SPSS* (*PASW Statistics* v18). Deformity factor was measured in each fish as presence (1) or absence (0) of the deformity. Loglinear model gives the significance of any deformity factor (its prevalence [i]) against any biological or functional factor (depending of data [j]), organized under a contingency table, through the normalized values or Z-values. Normalized Z-values for statistical significance of deformity in any family, breeder or mating ranged from > +1.96 (significant excess) to < −1.96 (significant deficit).

$$\mathrm{Ln} f_{ij} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij}$$

where $\ln f_{ij}$ is the expected frequency from observed frequency of each deformity (*i*) in each factor considered (*ij*); μ is the average value of expected frequencies logarithms, α_i is the effect of the deformity factor (*i*), β_j is the effect of the factor biological (*j*) and $\alpha\beta_{ij}$ is the effect due to these factors' interaction.

5.2.6 RNA sequencing analysis

5.2.6.1 RNA isolation

Sample were sent with dry ice to the Institute of Aquaculture Torre de la Sal (IATS-CSIC) for RNA isolation and sequencing analysis. Each sample tank-day, with at least 10 larvae, was pooled and extracted using the *MagMAXTM-96 RNA isolation kit* (Life Technologies, Carlsbad, CA, USA). The quality and integrity of isolated RNA were checked with an Agilent Bioanalyzer 2100 (Agilent, Amstelveen, Netherlands) varying the RNA Integrity Number values between 7 and 10.

5.2.6.1 RNA sequencing

Illumina RNAseq libraries were prepared from using the Illumina *NEBNextUltra^{TM}RNALibrary* Prep Kit (NEB, USA), according to the manufacturer's instructions. All RNAseq libraries were sequenced on an Illumina NovaSeq 6000 sequencing

system in a 2 × 150 pb read format, according to the manufacturer's protocol. Raw sequenced data were deposited in the Sequence Read Archive (SRA) of the National Center for Biotechnology Information (SRA ACCESSIONS: SRR13136984-7005) under the Bioproject accession number PRJNA680766 (BioSample accession numbers: SAMN16913187-208).

5.2.6.1 Bioinformatic analysis

A scheme of the workflow followed to obtain the different result from the raw reads is represented in Figure 18. Briefly, two phases were necessary, one for preprocessing the readings and another for analysis. In the first, good quality readings were filtered and mapped to the genome, and their expression level counted. In the second, differential expression analyses, functional analysis, variant calling and physical distribution analysis were carried out.

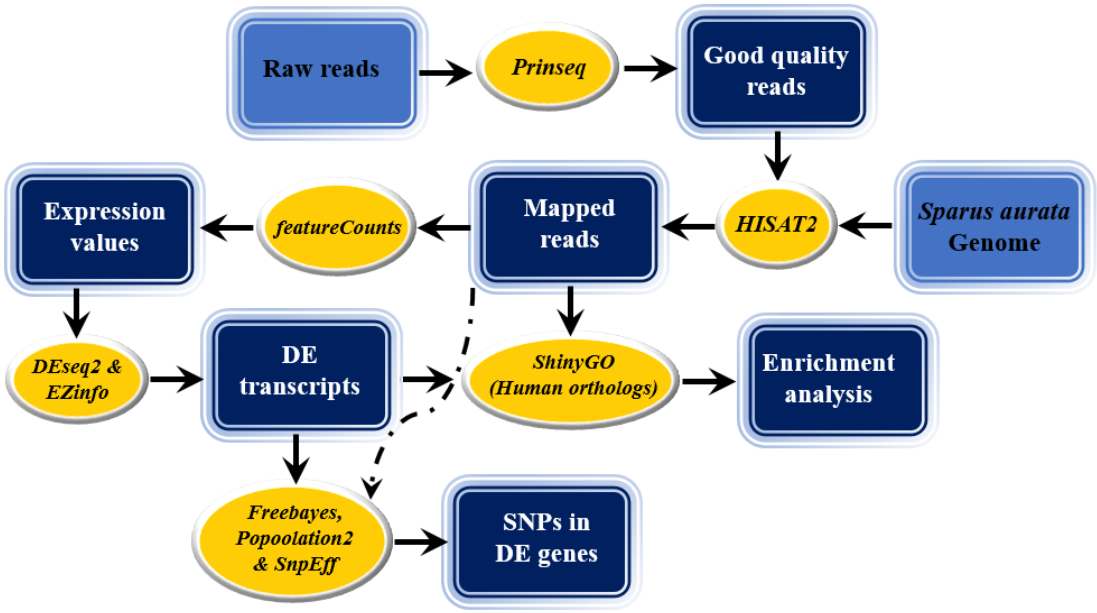


Figure 18. General workflow followed to carry out the processing and bioinformatics analysis of the output readings of the sequencer in experiment 3. Sequence filtering, its mapping in a reference genome, counting of expressed transcripts and later the different analyses (differential expression (DE), functional enrichment and variant calling analysis). In light blue is represented new data incomes, in yellow the softwares' names and in dark blue processed data.

5.2.6.1.1 Reads pre-processing and genome mapping

Quality analysis was performed with *FASTQC v0.11.7 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/)*, and libraries were filtered with *Prinseq* (Schmeider & Edwards, 2011) for reads reaching Q30 quality and less than 10 % of

ambiguous bases in the sequence. Then, reads were mapped and annotated using *HISAT2 v2.05* (Kim et al., 2019) using the CSIC gilthead seabream draft genome as reference (Pérez-Sánchez et al., 2019). Read numbers mapped over each gene were counted using *FeatureCounts* (Liao et al., 2014). To consider the effect of sequencing depth and gene lengths, gene expression levels were measured using RPKM.

5.2.6.1.2 Differential expression analysis

DEG were retrieved with normalized RPKM values using *DEseq2 v1.20.0* (Love et al., 2014) at an adjusted FDR < 0.05. To increase the number of DEG without loss of statistical robustness, supervised PLS-DA and hierarchical clustering of samples were sequentially applied using *EZinfo v3.0* (Umetrics, Umea, Sweden) and R package *ggplot2*, respectively. The final list of genes contributing to group separation (gD35 vs. gN35) was determined by the minimum VIP values (Li et al., 2012; Kieffer et al., 2016), driving the right clustering of all individuals in the heatmap analysis. To discard the possibility of over-fitting of supervised discriminant model, a validation test consisting in 500 random permutations was performed using *SIMCA-P+ v11.0* (Umetrics). The heatmap and clustering representations were constructed using *iDEP v0.91* (Ge et al., 2018), with the average linkage method and Euclidean distance as main parameters. Genes with a VIP > 1 (discriminant genes) were listed and used in three different procedures, including gene enrichment analysis, SS localization and variant calling analysis.

5.2.6.1.2 Gene enrichment analysis

Discriminant genes were converted into their human equivalents, and used to perform Biological Process Gene Ontology (GO-BP) enrichment analysis using the *ShinyGO v0.61* (Ge et al., 2020). Significantly enriched GO categories were obtained after FDR correction with a cut-off of 0.05, and clustered to observe overlapping using the average linkage method and Euclidean distance of R function *hclust*. Finally, alignment files generated in the mapping step were used for the calling of genetic variants, being the transcription start sites of discriminant genes in silico predicted within 5,000 bp upstream start codon (ATG) of each gene using *PROMOTER 2.0* (Knudsen, 1999).

5.2.6.1.3 Variant calling analysis

To minimize biased estimates of allele frequencies, PCR duplicates were removed from the samples files using *SAMtools v1.6* (Li et al., 2009). The resultant files were then introduced in *Freebayes v1.0.2* (Garrison and Marth, 2012). Considering the pooled nature of the samples, variants were considered in discriminant genes when (i) quality of the variant base-pair position was at least 30, (ii) at least 2 reads supported the alternate allele, and (iii) the variant allele frequency in the pool was above 0.05. The resulting variant calling files were formatted and allelic frequencies differences were calculated using the scripts *snp-frequency-diff.pl* and *fisher-test.pl* from *Popoolation2 v1.201* package (Kofler et al., 2011). Significant coding variants were annotated using *SnpEff v4.3* (Cingolani et al., 2012), and genetic architecture of discriminant genes with significant variants was represented using the *genemodel v1.1.0* R package. Variant physical position was located among the 24 SS using *bedtools v2.29.2* (Quinlan and Hall, 2010), and clustering of higher variant density SS was performed using the average linkage method and Euclidean distance of R function *hclust*.

5.3 Results

5.3.1 Osteology analysis and association study

In the osteological analysis, all severe deformities were classified as kyphosis, lordosis, partial and total vertebral body fusion, vertebral anomaly (shape anomaly, ossification ridges, marked reduction in length or elongation, intervertebral bony plate), anomalous maxillary and/or pre-maxillary, anomalous dentary and other cephalic anomalies (Figure 19).

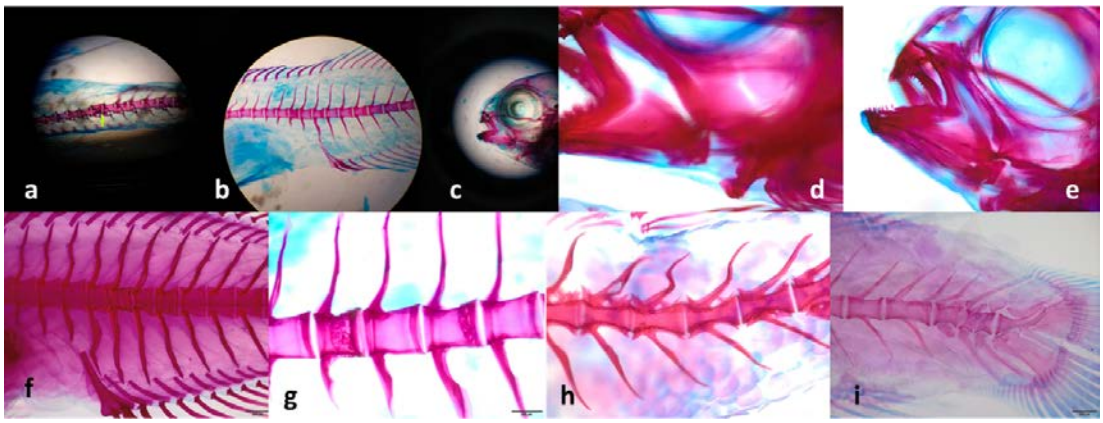


Figure 19. Osteological deformities in gilthead seabream larvae. a, Vertebral anomaly. b, Vertebral anomaly. c, Anomalous maxillary and/or pre-maxillary. d, Cephalic anomalies. e, Cephalic anomalies. f, Vertebral anomaly. g, Vertebral anomaly. h, Lordosis. i, Partial vertebral fusion, total vertebral body fusion and vertebral anomaly.

The association analysis between the presence-absence of severe deformities per type of genetic cross is reported in Table 8. *Normal phenotype* was significantly associated with gN cross, reporting an excess of larvae with this phenotype ($Z = + 4{,}12$, $p < 0.05$). At the same time, *deformity phenotype* was significantly associated with gD cross, reporting an excess of larvae with this phenotype ($Z = + 4.12$, $p < 0.05$)

Table 8. Z-values between morphological phenotype (normal, deformity) against type of genetic cross (gD and gN) in the spawning season as a whole.

	TYPE OF CROSS	
PHENOTYPE	**gN**	**gD**
Normal	4,109396	-4,109396
Deformity	-4,109396	4,109396

In Table 9, the interaction between *deformity* or *normal phenotypes* per dph time is shown, where can be appreciated a significant excess of *deformity phenotype* between 67–71 dph ($Z = + 1.99$ / $+ 2.96$, p-value < 0.05), within gD. While, in those same 67–71 dph exists a significant defect of *normal phenotype* (p-value < 0.05). It is repeated in the global analysis but including a significant excess of *deformity phenotype* at 72 dph, and a significant excess of *normal phenotype* at 41 dph ($Z = + 2.12$, $p < 0.05$). In gN cross, the distribution between *deformity* and *normal phenotypes* and dph time was not significant.

Table 9. Z-values between morphological phenotype (normal, deformity) against 35–73 dph, globally and per type of cross (gD and gN).

GLOBAL

	BOTH CROSSES																				
	35	39	41	43	45	46	47	49	51	53	55	57	59	61	63	65	67	69	71	72	73
Normal	1,167	,893	2,120	-,001	1,626	-,294	-,991	,171	,314	,216	,529	-,149	-1,182	1,514	-,618	-,456	-2,382	-1,996	-2,830	-2,434	-,249
Deformity	-1,167	-,893	-2,120	,001	-1,626	,294	,991	-,171	-,314	-,216	-,529	,149	1,182	-1,514	,618	,456	2,382	1,996	2,830	2,434	,249

NxN

	gN																				
	35	39	41	43	45	46	47	49	51	53	55	57	59	61	63	65	67	69	71	72	73
Normal	1,066	,336	1,957	-,305	,668	-,812	-,128	-,877	1,387	1,090	-,115	-,784	-1,257	1,209	,088	,272	-,471	,088	,088	-,812	-,471
Deformity	-1,066	-,336	-1,957	,305	-,668	,812	,128	,877	-1,387	-1,090	,115	,784	1,257	-1,209	-,088	-,272	,471	-,088	-,088	,812	,471

DxD

	gD																				
	35	39	41	43	45	46	47	49	51	53	55	57	59	61	63	65	67	69	71	72	73
Normal	-,509	1,383	1,299	,861	1,784	,352	-,823	1,630	-,498	-,355	1,017	,978	,122	1,157	-,355	-,355	-1,986	-2,702	-2,955	-1,618	-,509
Deformity	,509	-1,383	-1,299	-,861	-1,784	-,352	,823	-1,630	,498	,355	-1,017	-,978	-,122	-1,157	,355	,355	1,986	2,702	2,955	1,618	,509

The projection of *deformity phenotype* frequency distribution between *genetic crosses* (gN and gD) per dph time, is shown in Figure 20. A significant change of tendency can be observed between both types of genetic crosses, reporting an early intersection in development (43 dph).

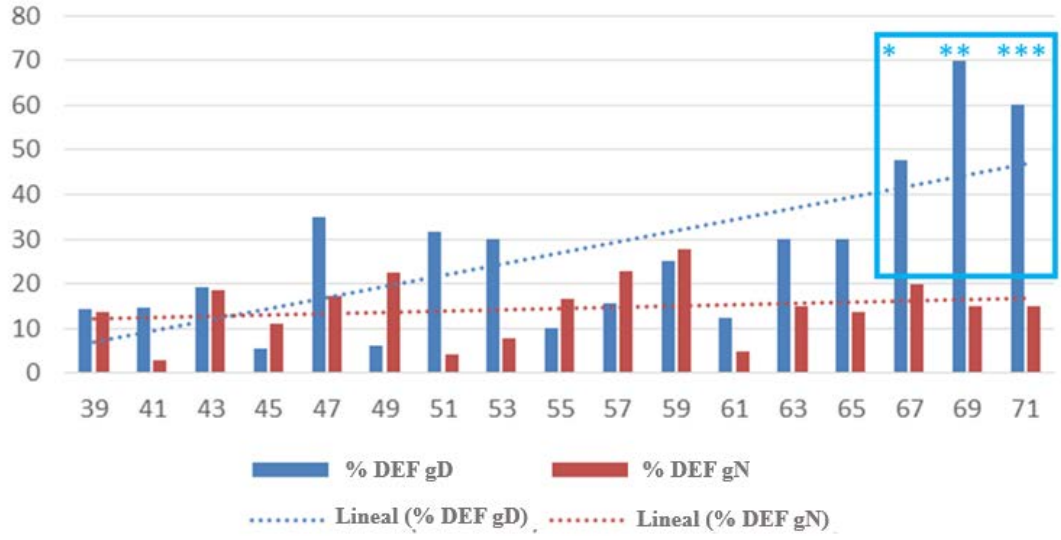


Figure 20. Deformity phenotype frequency distribution (%) versus time (dph).

The association between *deformity phenotype* and dph time was confirmed in different replicate (tank within cross) (Table 10). Only significant excesses of *deformity phenotype* occurred at 67 dph (Z = + 2.53, $p < 0.05$) and at 71 dph (Z = + 3.18, $p < 0.05$) within gD crosses.

Table 10. Z-values between morphological phenotype (normal, deformity) against 35–71 dph, in tank (1 and 2) within type of cross (gD and gN).

Severe deformity (dph/tank)

	gN1																		
	35	39	41	43	45	46	47	49	51	53	55	57	59	61	63	65	67	69	71
Normal	-,760	,876	1,456	,756	,162	-,760	-1,042	-1,042	1,252	,772	,506	-,711	-1,434	,433	-1,042	-,326	-1,042	,433	1,123
Deformity	,760	-,876	-1,456	-,756	-,162	,760	1,042	1,042	-1,252	-,772	-,506	,711	1,434	-,433	1,042	,326	1,042	-,433	-1,123

	gN2																		
	35	39	41	43	45	46	47	49	51	53	55	57	59	61	63	65	67	69	71
Normal	-,748	,457	1,138	,457	-,295	-,748	1,065	,457	-,748	1,138	-1,004	-,295	-,295	1,138	1,138	,681	,457	-,295	-1,004
Deformity	,748	-,457	-1,138	-,457	,295	,748	-1,065	-,457	,748	-1,138	1,004	,295	,295	-1,138	-1,138	-,681	-,457	,295	1,004

	gD1																		
	35	39	41	43	45	46	47	49	51	53	55	57	59	61	63	65	67	69	71
Normal	-,130	,769	1,737	1,502	1,353	-,130	-1,521	1,149	-,076	,185	-,130	,001	,185	,426	-,400	-,400	**-2,528**	-1,521	-,972
Deformity	,130	-,769	-1,737	-1,502	-1,353	,130	1,521	-1,149	,076	-,185	,130	-,001	-,185	-,426	,400	,400	**2,528**	1,521	,972

	gD2																		
	35	39	41	43	45	46	47	49	51	53	55	57	59	61	63	65	67	69	71
Normal	-,877	,962	,177	-,629	1,029	-,629	,962	,962	-,629	-,629	,177	1,090	,177	,962	,177	,177	,962	-,877	**-3,176**
Deformity	,877	-,962	-,177	,629	-1,029	,629	-,962	-,962	,629	,629	-,177	-1,090	-,177	-,962	-,177	-,177	-,962	,877	**3,176**

5.3.2 Differential expression analysis

Approximately 1,152 M paired-end reads were obtained from 22 pooled sequenced samples (1 gN5, 4 gD5, 10 gN35 and 8 gD35), with an average of ~52 M reads per sample. After trimming and quality filtering, remaining reads ranged between 48-67 M reads in all samples. Up to 89 % of these pre-processed reads were mapped against the gilthead seabream reference genome, which showed a total of 42,852 expressed transcripts, corresponding to 18,577 unique gene descriptions (87.42 % of total predicted unique transcripts). Of these, 17,998 unique gene descriptions were expressed at 5 dph and 17,042 at 35 dph.

5 dph samples from different genetically crosses were taken as a whole in the same category as an initial reference point (gN5 & gD5 = g5). Then comparisons between 5 and 35 dph (g5 vs. gN35 & g5 vs. gD35) and within genetically crosses in 35 dph (gN35 vs. gD35) were made, obtaining 3,817, 2.718 and 324 DET, respectively (corresponding to 2,456, 1,843 and 260 unique gene descriptions, respectively) (Figure 21). Altogether, 4,678 DET (3,302 unique descriptions) in at least one comparison were obtained and used in PLS-DA. The two first components of the resulted model explained 83 % of total variance at 35 dph and clearly separated gN35 and gD35 groups (Figure 22A). Differences in the model were driven by a total of 1,868 discriminant transcripts (VIP > 1; 1,398 unique gene descriptions). The classifier performance of PLS-DA model was validated by 500-model permutation tests and hierarchical clustering of samples and discriminant transcripts shown in Figure 22B.

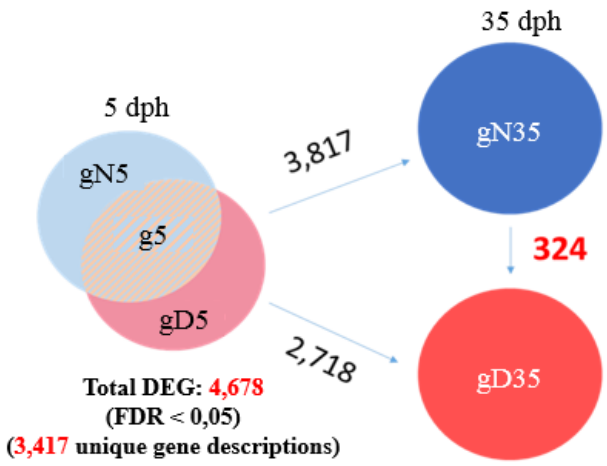


Figure 21. Total DEG results from the differential expression analysis performed in larvae using Deseq2. Circles represent conditions (gD35 and gN35), merged circle represents two conditions taken as one (gN5 + gD5); arrows represent comparisons and the number of DEG resulted.

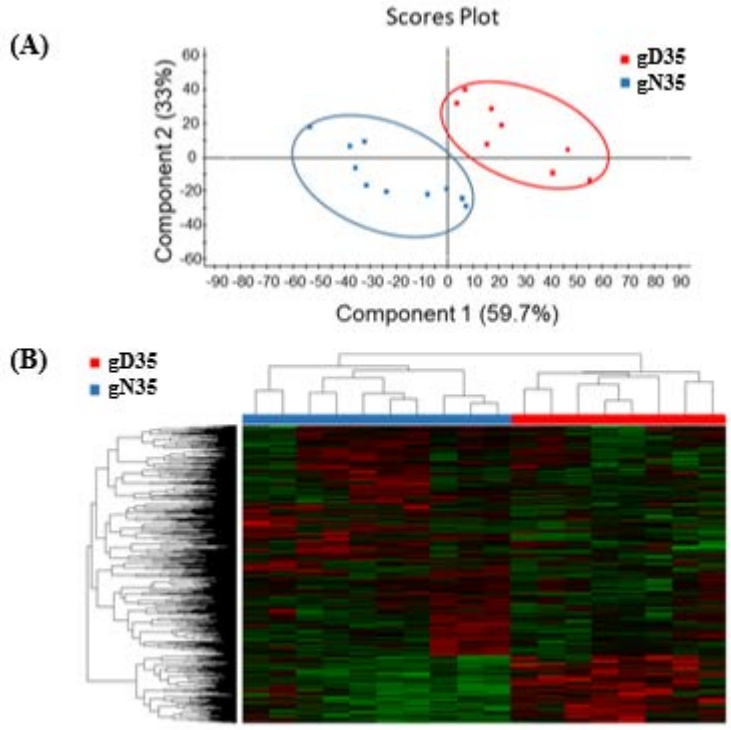


Figure 22. PLS-DA and clustering analyses that separate the two genetic conditions of deformity at 35 dph (gN35 in blue and gD35 in red). A) PLS-DA model using total DEG as input; B) Heatmap with the expression values of the 1,868 discriminant transcripts (rows) in the 35 dph samples (columns). Dendrograms shows the hierarchical clustering.

5.3.3 Gene enrichment analysis

GO-BP enrichment analysis disclosed 6 different enriched GO-BP categories, allocating 317 discriminant transcripts (Figure 23A) that corresponded to 305 unique gene descriptions. The most significant categories were *nervous system development* (GO:0007399) with 129 human equivalents, representing the 5.2% of genes of this category in the GO-BP database; *biological adhesion* (GO:0022610; 51 human equivalents, 3.2% total); *regulation of signaling* (GO:0023051; 101 human equivalents, 2.6% total); *cellular developmental process* (GO:0048869; 112 human equivalents, 2.3% total); *organic substance transport* (GO:0071702; 69 human equivalents, 2.2% total); and *small molecule metabolic process* (GO:0044281; 36 human equivalents, 1.4% total).

The clustering of the enriched GO-BP categories showed two different nodes of associations (Figure 23B). The first one included 117 discriminant transcripts corresponding to the categories of *nervous system development*, *cellular developmental process*, and *biological*

adhesion. The second cluster (120 discriminant transcripts) was composed by the categories of *regulation of signaling*, *organic substance transport* and *small molecule metabolic process*. Between them, the higher degree of overlap were found between *nervous system development* and *cellular developmental process* (83 discriminant transcripts), and between *regulation of signaling* and *organic substance transport* (85 discriminant transcripts).

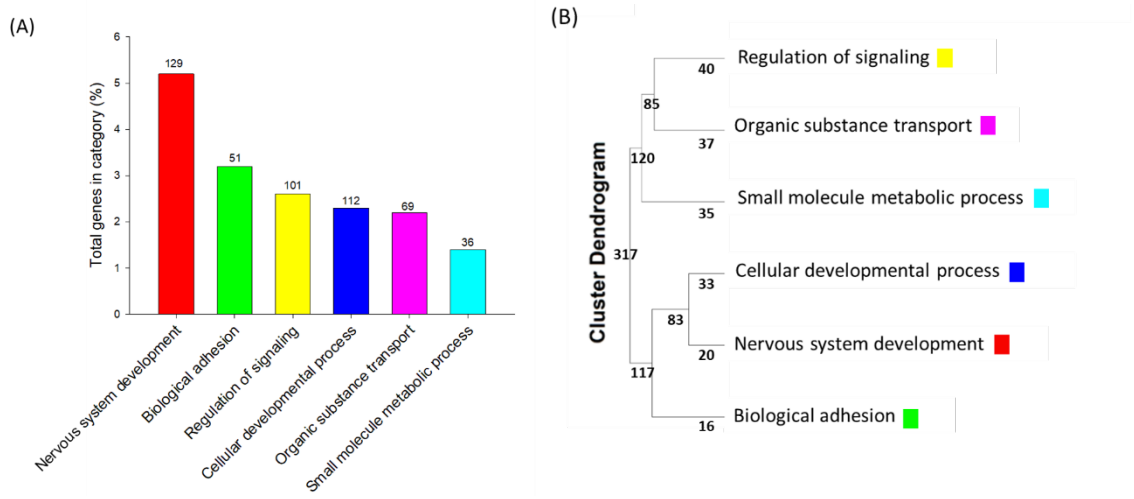


Figure 23. Gene enrichment categories results using GO-BP database. A) Bar graphic with the 6 GO-BP enriched categories (FDR < 0.05) and the percentage of genes in the categories (the number of genes is also located above each bar). B) Cluster dendrogram that shows the number of discriminant genes in common between 6 GO-BP enriched categories.

5.3.4 Variant calling and variants physical distribution

A total of 973 discriminant transcripts (52.2%) were located along the 24 SS sequences constituting the virtual karyotype of gilthead seabream genome (https://seabreamdb.nutrigroup-iats.org) (Figure 24). The SS17 showed the highest number of discriminant transcripts (68), followed by SS5 (62), SS21 (58), SS1 (56) and SS20 (56). After scaffold-length normalization, the SS20 showed the highest number of discriminant transcripts per Mbp (2.5), followed by SS12 (1.9), SS17 (1.69), SS7 (1.66), and SS21 (1.6). This top five super-scaffolds concentrated 27.8 % of DET.

A total of 205,375 bi-allelic variants were called in discriminant genes using both gN and gD reads. When allelic frequencies were estimated and tested (FDR < 0.05), significant differences between gN35 and gD35 groups were found for 832 variants, all coming from alternate alleles of the deformity group. According to the type of mutation (Figure 25A), 736 (88.5%) significant variants were classified as SNPs, 54 as complex (6.4%), 32 as insertion (3.8%) and 8 as deletions (1.3%). Giving its position in the genomic region (Figure 25B), 304 (36.34%) significant variants were located in 5' untranslated regions (5'UTR), 271 (32.37%) in intronic regions, and 262 (31.29%) in coding exons sequences. The effect of the mutations in amino acid sequence was determined and a total of 154 variants (18.5%) were classified as missense, 105 (12.6%) were classified as synonymous, and 3 (0.3%) were considered as non-sense mutations.

The alternate alleles overlapped with 64 genes that were part of the list of discriminant transcripts in the PLS-DA (Table 11). The average number of alternative alleles per gene was 13, ranging from 1 to 30. Among this final list of genes, the most abundant enriched GO-BP category was *cellular developmental process* (28 genes), followed by *nervous system development* (23), *regulation of signalling* (18), *organic substance transport* (13), *biological adhesion* (9), and *small molecule metabolic process* (8).

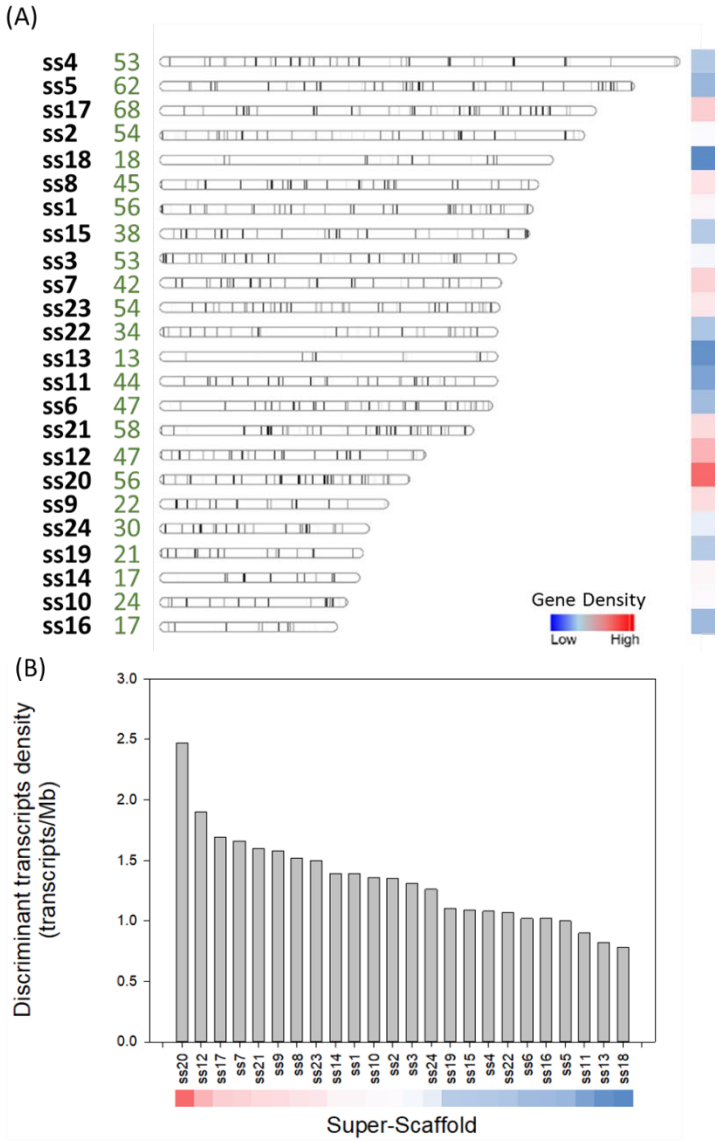


Figure 24. Discriminant transcripts' distribution in larvae between genetic groups. (A) Number and position of discriminant transcripts in SS; (B) Discriminant transcripts density (transcripts/Mb) per SS.

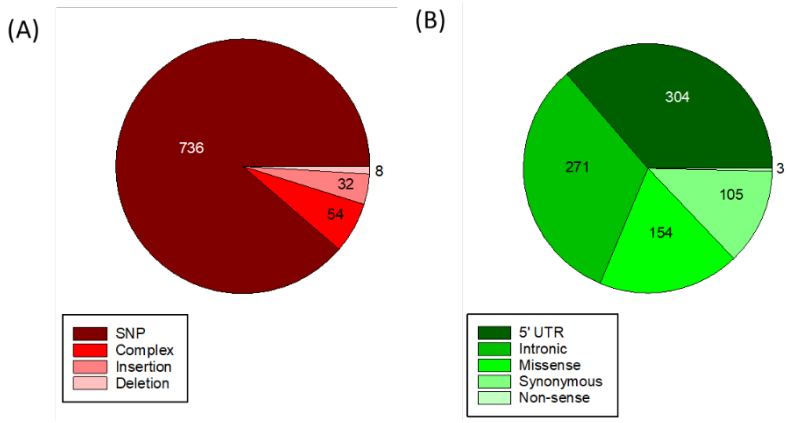


Figure 25. Classification of the genetic variants of the discriminant transcripts between genetic groups in larvae. (A) According to the type of mutation (SNP, Complex, Insertion or Deletion); (B) According to the genomic localization and effect over protein sequence (5'UTR, Intronic, Missense, Synonymous or Non-sense).

Table 11. List of 64 differentially expressed genes carrying differential variants related with deformity trait ordered by number of variants. In red are coloured the genes that coincide with the position of deformity QTLs described in gilthead seabream.

Gene	Number of variants (%)	GO-BP		Gene	Number of variants (%)	GO-BP		Gene	Number of variants (%)	GO-BP
egfl7	27 (3.24)	2,3,4		***gpat3***	18 (2.16)	3		***btn2a2***	10 (1.20)	2,3
cacna2d2	22 (2.64)	1,4,5		***rpl39***	18 (2.16)	5		***flnb***	10 (1.20)	4
hspb6	16 (1.92)	4		***chl1***	17 (2.04)	1,2,4		***hsd17b8***	10 (1.20)	6
pabpn1	12 (1.44)	5		***kctd6***	17 (2.04)	3		***kidins220***	10 (1.20)	1,3,4
f11	9 (1.08)	4		***cuzd1***	15 (1.80)	2		***kdr***	9 (1.08)	2,3,4
oxr1	8 (0.96)	1		***ndufs3***	15 (1.80)	1,3		***atp6ap1***	9 (1.08)	3,4
scn1b	8 (0.96)	1,2,4		***nubpl***	15 (1.80)	4		***chrnd***	9 (1.08)	1
psma2	2 (0.24)	3		***pigh***	15 (1.80)	6		***dclk1***	9 (1.08)	1,4
slc13a3	30 (3.60)	5		***rpl12***	15 (1.80)	5		***kcnj15***	9 (1.08)	5
fcgrt	23 (2.76)	5		***stx7***	15 (1.80)	1,5		***rpl14***	9 (1.08)	5
whamm	20 (2.40)	2		***gapdh***	14 (1.68)	1,5		***tbcd***	9 (1.08)	1,2,4
neto2	16 (1.92)	1,3		***llgl2***	14 (1.68)	3		***lrrc7***	8 (0.96)	1,2,3
ss18l2	11 (1.32)	1,4		***rgl3***	14 (1.68)	3		***aco2***	8 (0.96)	4
ctsd	9 (1.08)	5		***snap25***	14 (1.68)	1,4,5		***popdc2***	8 (0.96)	4
tkt	9 (1.08)	6		***got1***	13 (1.56)	3		***ptprz1***	8 (0.96)	1,4
cyp1a1	8 (0.96)	4		***meis1***	13 (1.56)	1,4		***asrgl1***	5 (0.60)	6
man2b1	28 (3.36)	5		***apopt1***	12 (1.44)	3		***kcnt1***	5 (0.60)	5
add2	24 (2.88)	1,2		***evc***	12 (1.44)	3		***supt20h***	5 (0.60)	4,6
cand1	24 (2.88)	4		***magi1***	12 (1.44)	2		***mycbp2***	2 (0.24)	1,4
elovl4	23 (2.76)	6		***rad9b***	12 (1.44)	3		***palm***	1 (0.12)	1,3,4
syne2	21 (2.52)	1,4		***gfi1***	10 (1.20)	1,3,4				
sesn1	20 (2.40)	3		***agxt2***	10 (1.20)	1				

5.4 Discussion

The most important economically significant skeletal deformities include the opercular complex, neurocranium, vertebral column and appendicular skeleton abnormalities in *Sparus aurata* (Koumoundouros et al., 1997; Boglione et al., 2001; Lee-Montero et al., 2015; Boursiaki et al., 2019). Being the most critical periods in fish the first stages of development, such as embryonic, larval and metamorphosis stages in which abnormal development can occur (Koumoundouros, 2010). Genetic architecture of morphological abnormalities trait and the analysis of skeletal deformities has been developed in this work.

According to the osteology analysis four development windows were pointed as relevant in the appearance of deformities and two of them were selected for RNAseq analysis. The osteology analysis showed a differential amount of deformities between 67-71 dph in gD cross. At this time, the metamorphosis is completely finished and it is considered the end of the hatchery stage in the industry (Faustino y Power, 1998; Sitjà-Bobadilla et al., 2005). A significant change of tendency can be observed between deformity phenotype frequency and both types of genetic crosses, reporting an early intersection after the late metamorphosis at 43 dph, where weaning is normally carried out (Mhalhel et al., 2020). Also, the complete evaluation of the entire skeleton could be done from 35 dph. In sparid species, evidence of mineralization can be discerned from 30 dph (Socorro, 2006) and 34 dph is consider the earliest moment to determine complete mineralisation in seabream larvae (Izquierdo et al., 2013). Finally, in this species, jaw and mouth bones are formed after the onset of exogenous feeding (Faustino & Power, 2005), thus, 5 dph was selected as initial reference point, prior to swim bladder inflation. In concordance with our results, samples from 5 and 35 dph development windows were considered for *RNAseq* analysis.

The use of pools of whole larvae families represented a cost-effective strategy which enabled the identification of several differentially expressed genes, genomic regions, pathways, and genetic variants associated to the deformity trait in this important cultured species. The conducted approach in this study combined typical differential expression tests with a discriminant analysis capable of increase the number of discriminant genes driving the separation of groups at 35 dph. These genes were located in all 24 gilthead seabream's SS.

Nervous system development is the most over-represented functional category within the set of differentially expressed genes. Gene ontology analysis of discriminant genes between gN35 and gD35 highlighted that more than 50 % of them participated in more than one of the over-represented processes. The major interaction was found between the processes of *nervous system development* and *cellular developmental process*, which hypothesized that changes in genes involved in neural differentiation processes are related to the deformity trait. An important interaction was also found between the processes of *organic substance transport* and

the *regulation of signalling*, which suggested that altered conformation of membrane proteins or solute carriers can cause changes in the signal transduction of the deformed fish cells.

Variant calling showed that all differential variants had a higher alternative allele proportion in deformity samples. This pointed out to the fact that deformed individuals are more prone to have their genomic sequence modified than normal individuals. Variant calling landscape also displayed an equivalence of differential variant proportion associated to 5' UTR, intronic and codifying sequences.

In total, 64 genes were determined as discriminant, enriched and polymorphic at the same time, which strongly correlate genomic regions with differentially expressed genes, metabolic routes and genetic variants related to the deformity, becoming this trait a complex hub with diverse levels. Indeed, these results were consistent with previous deformity-associated studies, where a total of 22 QTL were linked to this complex trait (Negrín-Báez *et al.*, 2015b, 2016a, b). The alignment of these QTL markers displayed 33 overlapping gilthead seabream genes. From these, 8 genes were up-regulated in the stringent FDR comparison between gD35 (log2FC > 2.5) against gN35 group. Furthermore, the 8 genes were found as enriched discriminant genes, carrying all of them at least 1 differential variant associated to deformity. The list of genes includes *egfl7* (27 variants), *meis1* (13 variants), *supt20h* (5 variants), *kcnt1* (5 variants), *asrgl1* (5 variants), *mycp2* (2 variants), *psma2* (2 variants) and *palm* (1 variant) overlapped with QTLSK3, QTLOP2, QTLSK14, QTLSK3, QTLOP4, QTLJW1, QTLOP4 and QTLORD2, respectively (Negrín-Báez *et al.*, 2015b, 2016a, b).

In conclusion, this study has generated a list of genes proposed as useful to discriminate the presence or absence of deformity, contributing with new criteria to the *European Certificate of Juvenile Quality* (ECJQ, a proposal to the sea bream aquaculture industry).

Chapter 6

Conclusions

1. The amplitude of the divergent selection according to the EBVdef was an effective method to investigate differential expression patterns between genetically deformed and genetically normal breeders and larvae.

2. A set of 59 differentially expressed transcripts between genetically and normal groups (702 between family comparisons) was determined from 3 target tissues between the selected breeders with divergent estimated breeding value for deformity.

3. Genetically deformed breeders showed significantly shorter spawning periods, lower oocyte yield and, consequently, lower number of alive larvae than genetically normal breeders.

4. The number of alive larvae in the middle of the spawning period is closely related with the oocyte yield and genetic factor.

5. The offspring coming from genetically deformed breeders showed a significantly higher incidence of deformity in the developmental window between 67-71 days post hatching than those coming from genetically normal breeders through osteological analysis.

6. A set of 324 differentially expressed transcripts between genetically deformed and normal larvae, coming from selected mates with divergent estimated breeding values for deformity, was determined at the beginning of mineralisation process.

7. Nervous system development is the most over-represented functional category within the set of differentially expressed genes between genetically deformed and normal larvae.

8. A number of 64 discriminant genes between genetically deformed and normal larvae showed polymorphic sequences engaging a total of 832 variants.

9. A list of 8 polymorphic and discriminant genes, between genetically deformed and normal larvae, was overlapped with deformity QTL genome regions, as previously reported.

These results provide a large number of candidate genes that can be used to develop precise molecular tools applicable in the industrial sector to decrease the prevalence of deformities. These tools could be implemented in breeding programs and contribute to improving the competitiveness of the gilthead seabream production.

Chapter 7

Resumen ampliado en español

7.1 Importancia económica del cultivo de la dorada

La dorada es una de las especies acuícolas más importantes de los países del Mediterráneo, con un mercado de 240.786 toneladas en 2019 y una demanda que ha crecido un 27,3 % en los últimos 5 años (APROMAR, 2020, 2019, 2018, 2017, 2016). En los últimos años, más del 95 % de la oferta de dorada procede de la acuicultura, habiéndose mantenido constante los volúmenes de captura de pesquería de esta especie entre 6,000-9,000 toneladas anuales. Esto supone que la demanda del mercado debe ser satisfecha a partir de la producción acuícola, cuya producción en el mediterráneo en 2019 supuso en valor económico 1.135,8 M€ (APROMAR, 2020).

Actualmente existe producción acuícola de dorada en 20 países, y si nos centramos en los datos de la UE, la producción de dorada en 2018 ocupó la tercera posición de pescado de crianza, después del salmón y la trucha arcoíris. En cuanto a valor económico, ocupa la 4º posición con una cuantía de 434,1 M€, habiendo sido superada durante el último lustro por la lubina (APROMAR, 2020, 2019, 2018, 2017, 2016).

La producción total de alevines de dorada en Europa (incluyendo Turquía) se mantiene en un valor aproximado de 700 millones de unidades desde el año 2016, un 20-25 % mayor que la cantidad producida el lustro anterior (2011-2015). El precio medio estimado de compraventa en España de 0,22 €/unidad, a peso equivalente de 2 g/unidad (APROMAR 2020), lo que supondría un mercado de 154 M€ si el precio en el resto de países es similar.

7.2 Especie en estudio

La dorada, *Sparus aurata* (Linnaeus, 1758), es un teleósteo marino perteneciente a la clase *Actinopterygii*, orden *Perciformes*, familia *Sparidae*:

Reino: *Animalia*

Filo: *Chordata*

Subfilo: *Vertebrata*

Clase: *Actinopterygii*

Infraclase: *Teleostei*

Superorden: *Acanthopterygii*

Orden: *Perciformes*

Familia: *Sparidae*

Género: *Sparus*

Especie: *aurata*

Posee el cuerpo de forma ovalada, profundo y comprimido. Tiene el perfil de la cabeza curvo, los ojos pequeños y la boca baja, muy levemente oblicua y con labios gruesos. Exhibe en la primera hilera de dientes entre cuatro y seis dientes de tipo canino, seguidos por entre 2 y 4 hileras de dientes cada vez más romos o molares. Posee fuertes mandíbulas y los dientes de las filas externas son más fuertes que los interiores. Luce un color gris plateado y una gran mancha negra en el origen de la línea lateral que se extiende sobre el margen superior del opérculo. Además, posee una banda frontal dorada característica entre los ojos, bordeada por dos áreas oscuras; líneas longitudinales oscuras sobre los costados del cuerpo; una banda oscura sobre la aleta dorsal; la horquilla y las puntas de la aleta caudal bordeadas con negro (Figura 1).

Se distribuye geográficamente (Figura 2) en zonas costeras subtropicales del Mediterráneo, Mar Negro y Atlántico oriental, desde las costas de Reino Unido y del Mar del Norte hasta las costas de Senegal (Froese & Pauly, 2019). Debido a sus hábitos eurihalinos y euritermos, la dorada se encuentra tanto en ambientes marinos como en lagunas costeras y estuarios, además de haber sido introducida en otras regiones por la acuicultura. Habita entre las algas de zonas rocosas y superficies arenosas poco profundas, generalmente desde el nivel del mar hasta los 50 m de profundidad, pudiendo llegar a los 150 m. Nace en alta mar y los juveniles migran a aguas costeras poco profundas durante la primavera y regresan a aguas más

profundas a finales de otoño, momento en el que los adultos se reproducen (Moretti et al., 1999).

Esta especie es hermafrodita protándrica, alcanza la madurez sexual como macho a los 2 años de edad y cambia a hembra a los 2–3 años condicionado por factores sociales y hormonales. Las hembras desovan en tandas intermitentes, pudiendo poner 20.000–80.000 huevos cada día por un período de hasta 4 meses. Su estación reproductiva comienza cuando los días son cortos y las horas de luz decrecientes, generalmente de octubre a diciembre en el Mediterráneo occidental (Arias, 1980) y entre noviembre y febrero en el Mediterráneo oriental (Ben-Tuvia, 1979). Los huevos son planctónicos, esféricos y transparentes, de aproximadamente 1 mm de diámetro y con una o varias gotas lipídicas en su interior (Moretti *et al*., 1999).

7.3 Repercusión industrial de las deformidades morfológicas en piscicultura

En la actualidad es frecuente encontrar individuos que presenten deformidades morfológicas durante el cultivo de las principales especies acuícolas con importancia comercial. Estas deformidades causan grandes consecuencias económicas a las empresas de acuicultura al incidir en su valor de mercado y aumentar los costes de producción. En estudios previos se ha estimado que en Europa la pérdida económica debida a las deformidades esqueléticas es de al menos 50 M€ anuales (Haga et al., 2011). Además, pese a que la incidencia de deformidades esqueléticas es especie dependiente, se han reportado incidencias de entre el 50-100 % de la producción en especies como pargo japonés (*Pagrus major*), dorada (*Sparus aurata*), lubina (*Dicentrarchus labrax*), bacalao del Atlántico (*Gadus morhua*), mero (*Epinephelus marginatus y Epinephelus septemfasciatus*), lenguado japonés (*Paralichthys olivaceus*), medregal (*Seriola*), lenguado senegalés (*Solea senegalensis*), salmón del Atlántico (*Salmo salar*) y trucha arcoíris (*Oncorhynchus mykiss*) (Haga et al., 2011).

Las pérdidas económicas producidas por la aparición de deformidades ocurren en dos niveles, a nivel de producción de alevines y de engorde de juveniles. La aparición de animales deformes en las *hatcheries* reduce la tasa de supervivencia y de crecimiento de las larvas deformes (Fernández et al., 2008). Esto obliga a las mismas a introducir sobrecostes sobre los animales sanos al tener que sobredimensionar su producción y/o realizar procedimientos de

cribado manuales para eliminar los animales deformes, ya que las empresas de engorde no aceptan lotes con prevalencias de deformes mayores al 5 % (Afonso & Roo, 2007), teniendo en cuenta que en este estadío algunos peces deformes pueden pasar desapercibidos por observación externa. A nivel de empresas de engorde de juveniles, los peces deformes deben descartarse o venderse por debajo del nivel de mercado, lo que supone un gasto adicional, además del producido por su menor rendimiento productivo (menor capacidad natatoria, ratio de conversión y de crecimiento y tasa de mortalidad y susceptibilidad al estrés y afección por patógenos) (Andrades et al., 1996; Boglione et al., 2001; Fernández et al., 2008; Karahan et al., 2013). Por otro lado, la presencia de deformidades genera un impacto negativo en el bienestar animal y en la percepción que tiene el consumidor sobre los productos acuícolas (Komen et al., 2002). Por ello, la presencia de deformidades es el segundo carácter en importancia económica para el sector de producción industrial de dorada, debido a la devaluación del valor de mercado que provoca en el producto a lo largo de la cadena de producción, agravado por su venta como pez entero (Bardon et al., 2009; Georgakopoulou et al., 2010; Boglione & Costa, 2011; Boglione et al., 2013a, 2013b). Todo esto ha influido en que los caracteres de calidad morfológica hayan adquirido la misma importancia que los caracteres de crecimiento en los programas de mejora genética europeos, en términos de frecuencia (Janssen et al., 2017).

7.4 Principales tipos de deformidades morfológicas y su importancia en acuicultura

Las anomalías morfológicas son desviaciones de la morfología del pez respecto a un estándar de calidad. La prevalencia de estas anomalías en lotes industriales está asociada a la intensificación del cultivo de especies piscícolas de interés comercial y afectan a su calidad integral (Divanach et al., 1996; Afonso & Roo, 2007). Las principales anomalías que tienen importancia industrial son aquellas que afectan la forma del cuerpo del pez, ya que la forma es el primer criterio visual de evaluación de la calidad de una especie, especialmente en aquellas que se comercializan como pez entero (Divanach et al., 1996; Afonso & Roo, 2007). De manera general en piscifactorías, los animales que son reconocidos por tener una forma corporal anómala tienen deformidades vertebrales (Gjedrem & Baranski, 2009). Las deformidades en la espina dorsal, junto con deformidades neurocraneales y en el esqueleto apendicular, se clasifican dentro del grupo de deformidades esqueléticas y son la principal causa de variación

de la forma del pez (Afonso & Roo, 2007). Las anomalías morfológicas pueden clasificarse en malformaciones de las escamas y problemas de pigmentación, además de en deformidades esqueléticas, pero estas anomalías tienen generalmente menor impacto económico, particularmente en esta especie de estudio.

7.5 Etiología de las deformidades morfológicas en peces

La causa de la mayoría de las deformidades no se conoce con exactitud y en la mayoría de las especies aparecen durante el desarrollo embriológico, larvario o comienzo de la etapa adulta (Boglione et al. 2009; Loizides et al. 2014; Azevedo, et al. 2017; Fragkoulis et al. 2017; Thuong et al. 2017). De manera general el aumento de prevalencia de deformidades se asocia a condiciones abióticas desfavorables, como manejo excesivo, presencia de xenobióticos o desbalances nutricionales, y condiciones bióticas como parásitos. Sin embargo, aunque las deformidades tienen un fuerte componente ambiental, existen diversos estudios que sugieren un componente genético suficiente como para ser explotado en los programas de mejora.

9.5.1 Factores no genéticos

Existen numerosos estudios que asocian condiciones desfavorables de manejo o de cultivo con una mayor incidencia de animales deformes en dorada y en otras especies. Así pues, se han relacionado las altas densidades de cultivo en dorada con la aparición de una mayor proporción de anomalías caudales y deformidades vertebrales en general a nivel industrial (Koumoundouros et al., 1997; Prestinicola et al., 2013). Chaptain y Ounais-Guschemann publicaron la aparición de un 90 % de prevalencia de lordosis en larvas de dorada ocasionada por la no inflación de la vejiga debido a capas de aceite en la superficie del agua o a altas condiciones hidrodinámicas (Chatain & Ounais-Guschemann, 1990; Chatain, 1994). Pese a que la lordosis puede asociarse a problemas de inflación de vejiga, también puede desarrollarse con vejigas natatorias completamente funcionales (Andrades, 1996). La aparición de deformidades debidas a altas corrientes de agua ha sido descrita por otros autores en otras especies cercanas como lubina y dorada roja (Chatain, 1994; Divanach et al., 1997; Kihara et al., 2002). Otros factores como la temperatura también pueden aumentar la prevalencia de deformidades en dorada, así como la tasa de mortalidad durante estadíos larvarios tempranos (Polo, Yufera & Pascual, 1991; Georgakopoulou et al., 2010). Esto ocurre además en otros espáridos como la

breca (Sfakianakis et al., 2004) y especies como la lubina (Koumoundouros et al., 2001; Abdel et al., 2004; Sfakianakis et al., 2006; Georgakopoulou et al., 2007).

En cuanto a xenobióticos, se ha reportado un incremento en la incidencia de deformidades en peces expuestos de manera prolongada a pesticidas (Chun et al., 1981; Whittle et al. 1992). Del mismo modo existen reportes de mayor incidencia de deformidades en peces procedentes de aguas contaminadas con efluentes agrícolas, industriales y urbanos (Slooff, 1982; Browder et al., 1993), así como con residuos de selenio (Lemly, 1993).

La prevalencia de deformidades en peces está asociada también a desbalances nutricionales durante el desarrollo larvario (Cahu et al., 2003b; Boglione et al., 2013b). Así pues, un exceso de vitamina A es causante de una mayor proporción de deformidades en cultivo de dorada (Fernández et al., 2008) y otras especies piscícolas (Haga et al., 2011). Cahu y cols. (2003a) reportaron que una dieta con valores inadecuados de fosfolípidos tiene como resultado mayor prevalencia de deformidades de boca y de columna en lubina. De igual forma, utilizar niveles elevados de ácido docosahexaenoico en la dieta logró reducir un 50 % la incidencia de deformidades esqueléticas en pargo rojo (Izquierdo et al., 2010).

Como efectos bióticos que pueden desencadenar la aparición de deformidades en peces, Lom y cols. han descrito que la parasitación por la familia myxosporea puede causar anomalías esqueléticas por la aglomeración ectópica a lo largo de los tejidos conectivos del canal central que rodean el nervio espinal (Lom *et al.,* 1991). Del mismo modo, se han asociado la aparición de anomalías de mandíbula con parásitos en la boca en fletán del atlántico (Morrison & MacDonald, 1995).

9.5.2 Factores genéticos

Las deformidades esqueléticas también se han asociado con una base genética. Así, se ha reportado que la depresión consanguínea, que es consecuencia del nivel de consanguinidad incide en la presencia de malformaciones. En trucha arcoíris, Austal y Kittelsen (1971) encontraron una prevalencia de alevines deformes 38 veces mayor en familias de hermanos carnales (Aulstad & Kittelsen, 1971). Por otro lado, Kincaid (1983) detectó una prevalencia de deformidades un 37,6% respecto a familias de individuos no emparentados (Kincaid, 1983).

En dorada, Andrades y cols. (1996) propusieron un origen genético para explicar la prevalencia de lordosis en una población de cultivo en dorada. Afonso y cols. (2000), describieron por primera vez la asociación estadística de una deformidad compleja a una familia concreta perteneciente a un programa de mejora genética. Esta anomalía es una triple malformación esquelética (lordosis, escoliosis y cifosis (LSK)), repetida desde la cabeza hasta la cola. Astorga y cols. (2004) propusieron la segregación de un gen mayor al considerar todas las deformidades como un solo rasgo (presencia-ausencia de alguna deformidad), tras obtener estimas de heredabilidad muy elevadas tanto mediante un modelo animal como con análisis bayesiano (0.78 y 0.85, respectivamente). Posteriormente, Lee-Montero y cols. (2015) estimaron heredabilidades medio-bajas (0.07-0.26) para presencia-ausencia de cualquier deformidad desde alevines hasta adultos, pero con altas correlaciones genéticas entre ellas (0.83-0.99), utilizando un modelo animal. Al mismo tiempo, Negrín-Báez y cols. (2015b) reportaron un número significativamente mayor de descendientes con deformidad de opérculo, lordosis y fusión vertebral provenientes de cruces dirigidos de padres con las mismas deformidades, con estimas de heredabilidades de entre 0.34-0.46 para estas tres deformidades. García-Celdrán y cols. (2015) confirmaron estas altas heredabilidades para deformidades de columna (0.38-0.56) y de opérculo (0.43-0.46) mediante análisis bayesiano a los 163 y 690 días después de la eclosión, respectivamente, y posteriormente mediante un modelo animal a 163 días después de la eclosión, con heredabilidades de 0.53 y 0.37 para lordosis y deformidad de opérculo, respectivamente (García-Celdrán et al., 2016). En este último estudio se determinó también una heredabilidad media para la no inflación de la vejiga natatoria (0.36) y una correlación genética positiva de la misma con la lordosis (García-Celdrán et al., 2016). También en larvas de 39 días posteriores a la eclosión Fragkoulis y cols. (2018) estimaron una heredabilidad media para la fusión de maxilares con premaxilares (0.22) y para anomalías en la aleta caudal (0.15-0.27) (Fragkoulis et al., 2020). En estudios en adultos, Fragkoulis y cols. (2021) estudiaron la variabilidad de la forma corporal de juveniles de dorada de 155 días hasta los 589 días de edad, determinando la existencia de heredabilidades sustanciales con valores entre 0,17-0,51 y una correlación genética entre las distintas edades de 0,22-0,76.

7.6 Evolución de los programas de mejora genética en dorada

El cultivo de dorada comenzó en la década de los 70, a partir del engorde de larvas capturadas del medio (Janssen et al., 2017), acelerándose el desarrollo de esta industria en 1990 gracias a la mejora de las técnicas de reproducción (Divanach & Kentouri, 2000). Los primeros ensayos de mejora genética en la especie llegaron a mediados de 1990 (Knibb et al., 1997, 1998) y el primer programa de mejora genética a nivel industrial comenzó iniciada la década 2000 (Thorland et al., 2007; Chavanne et al., 2016; Janssen et al., 2017). El principal objetivo de los programas de mejora es aumentar la rentabilidad y sostenibilidad de las empresas productivas, pero manteniendo la variabilidad genética de los stocks de cultivo (Martinez, 2007). Tradicionalmente el principal carácter de interés industrial en los programas de mejora ha sido el crecimiento, pero con el tiempo se han sumado caracteres que han ido ganando importancia industrial como la deformidad o la resistencia a patógenos (Chavanne et al. 2016; Janssen et al. 2017; Vandeputte et al., 2019; Boudry et al., 2021).

En la actualidad existen empresas que llevan a cabo programas de mejora genética de dorada en Francia, Grecia, Italia, España, Croacia, Israel y Turquía (Boudry et al., 2021). La mayoría de ellas utilizan la selección masal como método de selección, aunque algunas utilizan métodos de selección familiar, incluyendo metodología BLUP mediante pedigrí molecular o crecimiento de las familias por separado (Boudry et al., 2021). El uso de marcadores moleculares, como los microsatélites y los SNP, permite mejorar la eficiencia de los métodos de selección, gracias a la asociación por desequilibro de ligamiento de estos marcadores situados en regiones concretas del genoma con el fenotipo de caracteres cuantitativos de interés, determinando así un QTL (Pérez-Enciso & Toro, 2007; Yue, 2014).

Los enfoques de detección de QTLs tradicionales permiten encontrar aquellos QTLs con un efecto medio-alto sobre el carácter de interés, pero no aquellos con un bajo efecto (Yue et al., 2014). Esto permite realizar selección asistida por marcadores (MAS), mejorando la precisión de la selección y la ganancia genética, especialmente en caracteres difíciles de medir o de baja heredabilidad, como la deformidad (Meuwissen, 2003; Martinez, 2007; Yue et al., 2014; Boudry et al., 2021). En la actualidad la mejora de las técnicas de secuenciación masiva y el abaratamiento de estas ha permitido la disposición de un gran número de marcadores

moleculares, así como el genoma y transcriptoma de muchas de las especies de cultivo. Esto ha contribuido al desarrollo de técnicas como el RNAseq o la selección genómica.

El RNAseq es una técnica que permite secuenciar el RNA total de una muestra, mapearlo, anotarlo y cuantificarlo de una vez (Saroglia & Liu, 2012; Qian et al., 2014). En acuicultura de peces esta técnica ha sido utilizada para ensamblar nuevos transcriptomas, anotar genomas, detectar SNPs y realizar análisis de expresión (Qian et al., 2014). El RNAseq es una tecnología más económica que la secuenciación completa del DNA (WGS) y pese a que se limita a las regiones codificantes del genoma, puede ser utilizado para detectar de manera fiable variantes genómicas como SNPs (Piskol et al., 2013). Existen marcadores que están asociados a loci genómicos que regulan los niveles de expresión de determinados genes, estas regiones son conocidas como eQTLs (Saroglia & Liu, 2012). El análisis de eQTLs de caracteres complejos permite estudiar los procesos moleculares que subyacen a la patobiología (Ogura et al., 2015).

La selección genómica (GS) fue propuesta por Meuwissen y cols. en 2001 y dicta que con un número suficientemente elevado de marcadores moleculares repartidos por todo el genoma (GW), la mayoría de los loci de caracteres cuantitativos estarían en fuerte desequilibrio de ligamiento con al menos un marcador, lo que permitiría sumar los efectos de cada uno de ellos para predecir el comportamiento de dicho carácter (Meuwissen et al., 2001). De esta forma, se puede resumir que mediante la selección BLUP podemos calcular el potencial genético de los reproductores a través de la información de sus parientes, la selección MAS utiliza marcadores moleculares asociados a QTLs con un efecto medio-alto en el carácter de interés y la GS permite predecir el potencial genético a través de un modelo realizado utilizando la suma de efectos de los GW (Figura 3).

Pese a que se ha demostrado que la selección MAS y GS es potencialmente eficaz en experimentos controlados, no son métodos de selección extendidos en un contexto industrial en acuicultura, probablemente debido a la falta de estudios acerca del balance de sobrecoste y beneficios extra que supone implementar estas técnicas (Yue et al., 2014; Boudry et al., 2021).

7.7 Marcadores moleculares y recursos genómicos disponibles en dorada

La importancia industrial de la dorada ha hecho que sea una especie sobre la que se centran numerosos estudios, lo que ha permitido que con el tiempo existan numerosos recursos genómicos y marcadores moleculares para esta especie (Manchado et al., 2016). A continuación se detallan algunos de estos recursos genómicos y marcadores, asociados con la deformidad.

Uno de los primeros intentos de generar recursos genómicos en esta especie se realizó en el marco del proyecto BRIDGEMAP (2000–2004), financiado con fondos comunitarios europeos (Manchado et al., 2016). En 2005, Sarropoulou y cols. construyeron el primer microarray de cDNA con 10.176 clones de una biblioteca de cDNA procedente de un mix de estadíos embrionarios y larvario (Sarropoulou et al., 2005). En 2006, Senger y cols. construyeron un mapa híbrido de radiación preliminar. Ese mismo año fue reportado por Franch y cols. (2006) el primer mapa de ligamiento, con 26 grupos de ligamiento (LG), realizado a partir de 204 marcadores microsatélites. Por otro lado, numerosos estudios centraron sus esfuerzos en obtener ESTs de distintos tejidos y momentos del desarrollo para poder estudiar no sólo marcadores en el genoma, sino descubrir qué genes se expresan. Ya en 2010 distintos autores reportaron transcriptomas adicionales para muestras de músculo esquelético, tejidos esqueléticos, intestino, sangre y riñón de juveniles de dorada, lo que dio como resultado un transcriptoma global de 113.927 contigs, de los cuales aproximadamente el 50 % fueron anotados (García de la Serrana et al., 2012; Calduch-Giner et al., 2013; Vieira et al., 2013; Manchado et al., 2016). A finales de esta misma década, se comenzó a disponer de genomas de esta especie, suficientemente completos como para ser ensamblados formando cromosomas. Dray y cols. reportaron en 2016 el genoma mitocondrial y posteriormente, dentro del marco de los proyectos AquaTrace (2012-2016) y FISHBOOST (2014-2019), Pauletto y cols. (2018) lograron el genoma completo. En 2019, Pérez-Sánchez y cols., dentro del marco del proyecto AQUAEXEL2020 (2015-2020), reportaron también otro ensamblaje completo del genoma de dorada. Toda esta información genómica hace que se esté muy cerca de obtener un transcriptoma y genoma estable que pueda ser utilizado para realizar análisis RNAseq robustos y una anotación estandarizada que permita comparar con otras especies y llevar a cabo un correcto análisis funcional. Aunque el RNAseq no requiere un conocimiento *a priori* del genoma de la especie, claramente puede beneficiarse de la abundante información contenida en

los genomas al reducir su complejidad de manera práctica (Manchado et al., 2016).

Cabe destacar que la sucesión de marcadores moleculares reportados en esta especie y el avance de las técnicas moleculares y genómicas hizo que en dorada comenzaran a realizarse estudios de QTLs para distintos caracteres, incluyéndose caracteres morfológicos y deformidades esqueléticas. Boulton y cols. encontraron un QTL significativo que explica el 13-23 % de la varianza fenotípica de su población para rasgos morfométricos (Boulton et al., 2011). Entre 2015 y 2016 Negrín-Báez y cols. realizaron distintos experimentos en los que reportaron varios QTLs asociados a distintas deformidades esqueléticas. En 2015, encontraron tres QTL significativos que explicaban 1.6-11.4 % de la varianza, uno para fusión de vertebras (QTLFV3), uno para lordosis (QTLOR1) y otro para deformidad mandibular (QTLJW1), este último localizado muy próximo a 2 de los microsatélites utilizados (CId-26-H, CId-03-F) (Negrín-Báez et al., 2015a). En 2016, realizaron un análisis de QTLs para la deformidad LSK mediante una regresión lineal y un modelo mixto lineal, detectando cuatro QTLs (QTLSK3, QTLSK6, QTLSK12, QTLSK14) significativos a nivel de genoma, que mostraron un gran efecto en ambos métodos (>35 %), explicando conjuntamente el 65-94 % de la varianza y encontrando además 2 marcadores microsatélites (DId-03-T and Bt-14-F) fuertemente ligado a esta deformidad (Negrín-Báez et al., 2016b). Ese mismo año, mediante metodología de regresión lineal, detectaron 2 QTLs significativos (QTLOP1 and QTLOP2) para la deformidad del opérculo, con un efecto entorno al 27 % (Negrín-Báez et al., 2016a).

El estudio en cultivos primarios de células madre mesenquimales (MSC) derivadas de células óseas de dorada, ha permitido caracterizar las redes de genes que regulan el proceso de la osteogénesis en distintas condiciones de temperatura (Riera-Heredia et al., 2018) y tratados con ácidos grasos (EPA, DHA, LA y ALA) (Riera-Heredia et al., 2019a). Por otro lado, mediante la comparación de tejidos de peces deformes y peces normales, se determinó que el patrón de expresión de genes relacionados con la maduración y mineralización de la matriz extracelular ósea (*col1a1, op, ocn, mgp* y *tnap*) así como los implicados en la reabsorción ósea (*ctsk* y *mmp9),* parecen ser clave para discernir entre peces normales y peces con deformidades de columna (lordosis y LSK), sin encontrarse un patrón diferencial en deformidades de mandíbula y opérculo (Riera-Heredia et al., 2019b).

7.8 Objetivos

El transcurso de generaciones de producción piscícola ha logrado la domesticación de diversas especies de cultivo mediante el estudio de su biología y la mejora de las metodologías para su cultivo y reproducción. Este hecho ha permitido que la oferta de estas especies de acuicultura se haya aproximado a los niveles de demanda regular de pescado del mercado, lo que ha conseguido que el precio de estos productos se mantenga más o menos constante a lo largo del año. Antes de esta situación, la estrategia seguida en los sistemas productivos era la intensificación de los cultivos y la selección a través de caracteres de crecimiento para aumentar la producción y satisfacer la demanda del mercado. Pero con el tiempo, el incremento de la competitividad del sector y la madurez del mercado ha hecho que el precio de estos productos se haya estabilizado e incluso decaído, obligando al sector industrial a bajar los costes de producción o dar valor añadido a sus productos para mantener la rentabilidad.

El sector de la dorada es un claro ejemplo de este acontecimiento. En esta especie ha ganado interés industrial la mejora de la calidad del producto final y la reducción de costes de producción, por lo que reducir el número de animales deformes en los cultivos es de vital importancia.

Con estos antecedentes, el objetivo de este estudio es descubrir la estructura genética subyacente en la prevalencia de deformidades esqueléticas en dorada de acuicultura y describir marcadores génicos que puedan ser utilizados para evaluar el potencial genético de los animales utilizados como reproductores y reducir la incidencia de deformidades de origen genético en la producción.

Para lograr este objetivo general se establecen los objetivos específicos siguientes:

-Analizar el valor genético aditivo para la prevalencia de deformidad en un gran número de animales procedentes de lotes industriales, garantizando un número de recombinantes suficiente para dilucidar el mecanismo genético a través de la comparación entre animales con distintos valores de cría para la deformidad (EBVdef).

-Determinar la existencia de un patrón de expresión diferencial en distintos tejidos procedentes de familias con EBVdef distantes, utilizando animales reproductores con elevada

predisposición genética para la deformidad y con elevada predisposición genética para la normalidad.

-Establecer lotes de reproductores fenotípicamente normales y EBVdef distantes (uno con un EBVdef elevado y positivo y otro con un EBVdef elevado y negativo) con el objeto de analizar la calidad de la puesta de los lotes de reproductores formados y realizar el cultivo larvario de sus descendientes para determinar la existencia de diferencias en la prevalencia de la deformidad y el momento del desarrollo en el que se origina.

-Determinar la existencia de un patrón de expresión diferencial y diferencias de secuencia entre los genes de ambos grupos (gD y gN) en diferentes estadíos del desarrollo larvario.

7.9 Resúmenes de los estudios

Las prevalencia de anomalías morfológicas en especies acuícolas es un problema que conlleva grandes pérdidas económicas en el sector industrial, esto es especialmente importante en la dorada de cultivo (*Sparus aurata*), debido a que se comercializa entera.

En el primer experimento se realizó una selección divergente en base al antecedente genético de deformidades esqueléticas con el fin de describir los patrones de expresión diferencial en reproductores, utilizando cerebro, músculo blanco y vértebras, como tejidos diana de la deformidad. Los valores mejorantes medios de deformidad en diferentes grupos familiares fueron 0,024 en familias genéticamente deformes (gD) y -0,038 en familias genéticamente normales (gN). Los transcriptomas de cerebro y vértebra muestran una complejidad similar, con unos 10.000 transcritos expresados, mientras que el transcriptoma de músculo blanco posee la mitad de transcritos expresados. Tras el análisis de expresión diferencial, se encontraron un total de 59 transcritos diferencialmente expresados (DET) entre los dos niveles genéticos de deformidad, 2 de los cuales se encontraban diferencialmente expresados en los tres tejidos diana, el *factor de coagulación XIII-A* (*f13a1*) y *gen similar a E3 ubiquitina-proteína ligasa* (*TRIM39l*). A nivel familiar se detectaron un total de 702 DET tomando en conjunto los resultados de los 3 tejidos. En términos de análisis funcional, la vértebra fue el único tejido que mostró un enriquecimiento en categorías funcionales entre los dos grupos, utilizando como base de datos de referencia la ontología génica de procesos

biológicos. Por primera vez, se ha llevado a cabo un estudio de RNAseq en muestras de animales seleccionados de manera divergente en base a su predisposición genética a la deformidad, lo que abre numerosas posibilidades para futuros estudios.

En el segundo estudio, se estudia por primera vez en esta especie, como los antecedentes genéticos de presencia de deformidad afecta en la calidad de la puesta de reproductores, poniendo de manifiesto la importancia de controlar este valor genético aditivo a la hora de establecer lotes de reproductores. La calidad de los huevos de peces se define como la capacidad que tienen para ser fertilizados y desarrollar embriones con buena supervivencia y sin anomalías. Las grandes variaciones en la cantidad y calidad del desove existentes en especies marinas impactan directamente en la competitividad y la sostenibilidad de los criaderos, que sobredimensionan sus lotes de reproductores para garantizar la producción de alevines, sin animales deformes, que demandan las empresas de engorde. En el experimento, se evalúa la calidad de la puesta de lotes de reproductores genéticamente deformes y normales durante una temporada de puesta completa mediante el control de los siguientes parámetros: rendimiento de ovocitos, tasa de fertilización, tasa de viabilidad, tasa de eclosión, tasa de supervivencia larvaria, número de huevos fertilizados, número de huevos viables, número de huevos eclosionados y número de larvas vivas. Aun teniendo un fenotipo normal, los reproductores con antecedentes genéticos de deformidad tuvieron períodos de desove más cortos, menor rendimiento de ovocitos y, en consecuencia, produjeron un menor número de larvas vivas. En estos dos parámetros, el antecedente genético de los reproductores fue de mayor importancia durante los períodos intermedios de puesta, momento en el que generalmente la puesta es considerada óptima en cultivos industriales, mientras que los factores ambientales fueron más importantes al principio y al final de la temporada de puesta.

En el tercer estudio, se utiliza la descendencia de grupos de dos grupos de reproductores seleccionados de manera divergente en base a su potencial genético de deformidad con el objetivo de esclarecer la arquitectura genética subyacente en el componente genético de la deformidad. Para ello, se tomaron muestras durante los cultivos larvarios de ambos grupos, se confirmó la prevalencia de animales deformes en una ventana del desarrollo larvario entre grupos mediante análisis osteológico y se utilizaron herramientas transcriptómicas para

determinar la existencia de patrones diferenciales de expresión y variaciones entre sus secuencias. Como resultado se encontró, en larvas de 35 días post eclosión, un total de 324 transcritos diferencialmente expresados entre grupos genéticos, número que aumentó a 1.868 después de un procedimiento de análisis discriminante que separó por completo los grupos experimentales. Además, un análisis de ontología génica muestra 6 categorías funcionales sobrerrepresentadas en estos transcritos discriminantes, destacando que más del 50 % participan en más de una de las categorías funcionales. La principal interacción se encontró entre el desarrollo del sistema nervioso y el proceso de desarrollo celular y entre los procesos de transporte de sustancias orgánicas y la regulación de la señalización. Finalmente, después de un análisis de variantes, se proponen 64 genes discriminantes y polimórficos entre grupos como candidatos para futuros estudios de deformidad. Estos resultados fueron consistentes con estudios previos de deformidad en esta especie, ya que 8 de los genes propuestos se encuentran en regiones del genoma compatibles con QTL de deformidad previamente reportados. Estos genes son *egfl7* (27 variantes), *meis1* (13 variantes), s*upt20h* (5 variantes), *kcnt1* (5 variantes), *asrgl1* (5 variantes), *mycp2* (2 variantes), *psma2* (2 variantes) y *palm* (1 variante).

7.10 Conclusiones

1. La amplitud de la selección divergente atendiendo al valor mejorante para la deformidad fue un método efectivo para investigar patrones de expresión diferencial en reproductores y larvas genéticamente deformes y genéticamente normales.

2. Se determinaron un total de 59 transcritos diferencialmente expresados entre grupos genéticamente deformes y normales (702 transcritos teniendo en cuenta las comparaciones entre familias) a partir de 3 tejidos diana en reproductores seleccionados de manera divergente según su valor mejorante para la deformidad.

3. Los reproductores genéticamente deformes mostraron periodos de desove significativamente más cortos, con un menor rendimiento de ovocitos y, en consecuencia, un menor número de larvas vivas que los reproductores genéticamente normales.

4. El número de larvas vivas en la mitad del período de puesta está estrechamente relacionado con la producción de ovocitos y el factor genético de sus reproductores.

5. Las larvas provenientes de reproductores genéticamente deformes mostraron una incidencia significativamente mayor de animales deformes en la ventana de desarrollo comprendido entre 67-71 días post eclosión respecto a las provenientes de reproductores genéticamente normales, a través del análisis osteológico.

6. Se determinó un conjunto de 324 transcritos diferencialmente expresados entre larvas genéticamente deformes y normales, provenientes de cruces seleccionados de manera divergente en base a su valor mejorante para la deformidad, al inicio del proceso de mineralización.

7. El desarrollo del sistema nervioso es la categoría funcional más sobre-representada dentro del conjunto de genes diferencialmente expresados entre larvas genéticamente deformes y normales.

8. Un número total de 64 genes discriminantes entre larvas genéticamente deformes y normales mostraron secuencias polimórficas incluyendo un total de 832 variantes.

9. Una lista de 8 genes polimórficos y discriminantes entre larvas genéticamente deformes y normales, co-localizan en regiones del genoma asociadas a QTL para deformidades, descritos anteriormente.

Estos resultados proporcionan un gran número de genes candidatos que pueden utilizarse para desarrollar herramientas moleculares precisas aplicables en el sector industrial para disminuir la prevalencia de deformidades. Estas herramientas podrían implementarse en programas de mejora genética y contribuir a la mejora de la competitividad del sector de la dorada.

Reference

Abdel, I., Abellán, E., López-Albors, O., Valdés, P., Nortes, M.J., García-Alcázar, A., 2004. Abnormalities in the juvenile stage of sea bass (Dicentrarchus labrax L.) reared at different temperatures: types, prevalence and effect on growth. Aquac. Int. 12, 523-538. https://doi.org/10.1007/s10499-004-0349-9

Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Čech, M., Chilton, J., Clements, D., Coraor, N., Eberhard, C., Grüning, B., Guerler, A., Hillman-Jackson, J., Von Kuster, G., Rasche, E., Soranzo, N., Turaga, N., Taylor, J., Nekrutenko, A., Goecks, J., 2016. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. Nucleic Acids Res. 44, W3-W10. https://doi.org/10.1093/nar/gkw343

Afonso, J.M., Montero, D., Robaina, L., Astorga, N., Ginés, R., 2000. Association of a lordosis-scoliosis-kyphosis deformity in gilthead seabream (Sparus aurata) with family structure, Fish Physiology and Biochemistry.

Afonso, J.M., Roo, J., 2007. ANOMALÍAS EN PECES CULTIVADOS: MORFOLÓGICAS HEREDABILIDAD Y SELECCIÓN, en: Espinosa de los Monteros, J. (Ed.), Genética y Genómica en Acuicultura. pp. 213-230.

Almansa, E., Pérez, M.J., Cejas, J.R., Badía, P., Villamandos, J.E., Lorenzo, A., 1999. Influence of broodstock gilthead seabream (Sparus aurata L.) dietary fatty acids on egg quality and egg fatty acid composition throughout the spawning season. Aquaculture 170, 323-336. https://doi.org/10.1016/S0044-8486(98)00415-3

Andrades, J.A., Becerra, J., Fernádez-Llebrez, P., 1996. Skeletal deformities in larval, juvenile and adult stages of cultured gilthead sea bream (Sparus aurata L.), Aquaculture.

APROMAR, 2021. La acuicultura en España 2021. Asociación Empresarial de Productores de Cultivos Marinos de España. 91.

APROMAR, 2020. La acuicultura en España 2020. Asociación Empresarial de Productores de Cultivos Marinos de España. 95.

APROMAR, 2019. La acuicultura en España 2019. Asociación Empresarial de Productores de Cultivos Marinos de España. 91.

APROMAR, 2018. La acuicultura en España 2018. Asociación Empresarial de Productores de Cultivos Marinos de España. 91.

APROMAR, 2017. La acuicultura en España 2017. Asociación Empresarial de Productores de Cultivos Marinos de España. 93.

APROMAR, 2016. La acuicultura en España 2016. Asociación Empresarial de Productores de Cultivos Marinos de España. 89.

AQUAEXCEL, 2013. Deliverable 3.2. Best practices & cross-applicability of methods to measure

phenotypes. Proj. number 262336 106.

Arias, A., 1980. Crecimiento, régimen alimentario y reproducción de la dorada (Sparus aurata L.) y del robalo (Dicentrarchus labrax L.) en los esteros de Cádiz. Invest. Pesq. 44, 59-83.

Astorga, N., 2005. Estudio genético de las deformidades esqueléticas de columna en dorada (Sparus aurata L.) en condiciones de cultivo intensivo. Universidad de las Palmas de Gran Canaria.

Astorga, N., Zamorano, M.J., Toro, M.A., García Cortes, L.A., Montero, D., Afonso, J.M., 2004. Heredabilidad del carácter presencial o ausencia de deformaciones esqueléticas en dorada (Sparus aurata L.). ITEA 100A, 256-260.

Aulstad, D., Kittelsen, A., 1971. Abnormal Body Curvatures of Rainbow Trout (Salmo gairdneri) Inbred Fry. J. Fish. Res. Board Canada 28, 1918-1920. https://doi.org/https://doi.org/10.1139/f71-290

Azevedo, A.M., Losada, A.P., Ferreiro, I., Riaza, A., Vázquez, S., Quiroga, M.I., 2017. New insight on vertebral anomalies in cultured Senegalese sole (Solea senegalensis, Kaup) at early stages of development. J. Fish Dis. 40, 987-1000. https://doi.org/10.1111/jfd.12575

Babicki, S., Arndt, D., Marcu, A., Liang, Y., Grant, J.R., Maciejewski, A., Wishart, D.S., 2016. Heatmapper: web-enabled heat mapping for all. Nucleic Acids Res. 44, W147-W153. https://doi.org/10.1093/nar/gkw419

Bardon, A., Vandeputte, M., Dupont-Nivet, 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, 194-201. https://doi.org/10.1016/j.aquaculture.2009.06.018

Ben-Tuvia, A., 1979. Studies of the population and fisheries of Sparus aurata L. in the Bardawil Lagoon, eastern Mediterranean. Investig. Pesq. 43, 43-67.

Bobe, J., 2015. Egg quality in fish: Present and future challenges. Anim. Front. 5, 66-72. https://doi.org/10.2527/af.2015-0010

Bobe, J., Labbé, C., 2010. Egg and sperm quality in fish. Gen. Comp. Endocrinol. 165, 535-548. https://doi.org/10.1016/j.ygcen.2009.02.011

Boglione, C., Costa, C., 2011. Skeletal Deformities and Juvenile Quality. Sparidae, Wiley Online Books. https://doi.org/doi:10.1002/9781444392210.ch8

Boglione, C., Gagliardi, F., Scardi, M., Cataudella, S., 2001. Skeletal descriptors and quality assessment in larvae and post-larvae of wild-caught and hatchery-reared gilthead sea bream (Sparus aurata L. 1758). Aquaculture 192, 1-22. https://doi.org/10.1016/S0044-8486(00)00446-4

Boglione, C., Gavaia, P., Koumoundouros, G., Gisbert, E., Moren, M., Fontagné, S., Witten, P.E., 2013a. Skeletal anomalies in reared European fish larvae and juveniles. Part 1: Normal and anomalous skeletogenic processes. Rev. Aquac. https://doi.org/10.1111/raq.12015

Boglione, C., Gisbert, E., Gavaia, P., Witten, P.E., Moren, M., Fontagné, S., Koumoundouros, G., 2013b. Skeletal anomalies in reared European fish larvae and juveniles. Part 2: Main typologies, occurrences and causative factors. Rev. Aquac. 5. https://doi.org/10.1111/raq.12016

Boglione, C., Marino, G., Giganti, M., Longobardi, A., De Marzi, P., Cataudella, S., 2009. Skeletal anomalies in dusky grouper Epinephelus marginatus (Lowe 1834) juveniles reared with different methodologies and larval densities. Aquaculture 291, 48-60. https://doi.org/10.1016/j.aquaculture.2009.02.041

Bonnet, E., Fostier, A., Bobe, J., 2007. Characterization of rainbow trout egg quality: A case study using four different breeding protocols, with emphasis on the incidence of embryonic malformations. Theriogenology 67, 786-794. https://doi.org/10.1016/j.theriogenology.2006.10.008

Boudry, P., Allal, F., Aslam, M.L., Bargelloni, L., Bean, T.P., Brard-Fudulea, S., Brieuc, M.S.O., Calboli, F.C.F., Gilbey, J., Haffray, P., Lamy, J.B., Morvezen, R., Purcell, C., Prodöhl, P.A., Vandeputte, M., Waldbieser, G.C., Sonesson, A.K., Houston, R.D., 2021. Current status and potential of genomic selection to improve selective breeding in the main aquaculture species of International Council for the Exploration of the Sea (ICES) member countries. Aquac. Reports 20. https://doi.org/10.1016/j.aqrep.2021.100700

Boulton, K., Massault, C., Houston, R.D., de Koning, D.J., Haley, C.S., Bovenhuis, H., Batargias, C., Canario, A.V.M., Kotoulas, G., Tsigenopoulos, C.S., 2011. QTL affecting morphometric traits and stress response in the gilthead seabream (Sparus aurata). Aquaculture 319, 58-66. https://doi.org/10.1016/j.aquaculture.2011.06.044

Boursiaki, V., Theochari, C., Zaoutsos, S.P., Mente, E., Vafidis, D., Apostologamvrou, C., Berillis, P., 2019. Skeletal deformity of scoliosis in gilthead seabreams (Sparus aurata): Association with changes to calcium-phosphor hydroxyapatite salts and collagen fibers. Water (Switzerland) 11. https://doi.org/10.3390/w11020257

Bromage, N., Jones, J., Randall, C., Thrush, M., Davies, B., Springate, J., Duston, J., Barker, G., 1992. Broodstock management, fecundity, egg quality and the timing of egg production in the rainbow trout (Oncorhynchus mykiss). Aquaculture 100, 141-166. https://doi.org/10.1016/0044-8486(92)90355-O

Brooks, S., Tyler, C.R., Sumpter, J.P., 1997. Egg quality in fish: What makes a good egg? Rev. Fish Biol. Fish. 7, 387-416. https://doi.org/10.1023/A:1018400130692

Browder, J.A., McClellan, D.B., Harper, D.E., Kandrashoff, M.L.G., Kandrashoff, W., 1993. A major developmental defect observed in several Biscayne Bay, Florida, fish species. Environ. Biol. Fishes 37, 181-188. https://doi.org/10.1007/BF00000593

Brown, R.C., 2003. Genetic Management and Selective Breeding in Farmed Populations of Gilthead Seabream (Sparus aurata).

Cahu, C.L., Infante, J.L.Z., Barbosa, V., 2003a. Effect of dietary phospholipid level and phospholipid:neutral lipid value on the development of sea bass (Dicentrarchus labrax)

larvae fed a compound diet . Br. J. Nutr. 90, 21-28. https://doi.org/10.1079/bjn2003880

Cahu, C.L., Zambonino Infante, J., Takeuchi, T., 2003b. Nutritional components affecting quality in marine fish larval development. Aquaculture 227, 254-258.

Calduch-Giner, J.A., Bermejo-Nogales, A., Benedito-Palos, L., Estensoro, I., Ballester-Lozano, G., Sitjà-Bobadilla, A., 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. BMC Genomics 14. https://doi.org/10.1186/1471-2164-14-178

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L., 2009. BLAST+: Architecture and applications. BMC Bioinformatics 10, 1-9. https://doi.org/10.1186/1471-2105-10-421

Carrillo, M., Bromage, N., Zanuy, S., Serrano, R., Prat, F., 1989. The effect of modifications in photoperiod on spawning time, ovarian development and egg quality in the sea bass (Dicentrarchus labrax L.). Aquaculture 81, 351-365. https://doi.org/10.1016/0044-8486(89)90159-2

Castro, J., Pino-Querido, A., Hermida, M., Chavarrías, D., Romero, R., García-Cortés, L.A., Toro, M.A., Martínez, P., 2008. Heritability of skeleton abnormalities (lordosis, lack of operculum) in gilthead seabream (Sparus aurata) supported by microsatellite family data. Aquaculture 279, 18-22. https://doi.org/10.1016/j.aquaculture.2008.04.023

Chatain, B., 1994. Abnormal swimbladder development and lordosis in sea bass (Dicentrarchus labrax) and sea bream (Sparus auratus). Aquaculture. https://doi.org/10.1016/0044-8486(94)90301-8

Chatain, B., Ounais-Guschemann, N., 1990. Improved rate of initial swim bladder inflation in intensively reared Sparus auratus. Aquaculture 84, 345-353. https://doi.org/10.1016/0044-8486(90)90099-9

Chavanne, H., Janssen, K., Hofherr, J., Contini, F., Haffray, P., Aquatrace Consortium, Komen, H., Nielsen, E.E., Bargelloni, L., 2016. A comprehensive survey on selective breeding programs and seed market in the European aquaculture fish industry. Aquac. Int. 24, 1287-1307. https://doi.org/10.1007/s10499-016-9985-0

Chun, S., Park, C., Lee, C., 1981. Acute toxicity of certain pesticides to teleost, with special respect to TLM and vertebral abnormality. Publ. Inst. Mar. Sci. Nat. Fish. Univ. Busan 13, 43-48.

Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., Ruden, D.M., 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin). 6, 80-92. https://doi.org/10.4161/fly.19695

Divanach, P., Boglione, C., Menu, B., Koumoundouros, G., Kentouri, M., Cataudella, S., 1996. ABNORMALITIES IN FINFISH MARICULTURE: AN OVERVIEW OF THE PROBLEM, CAUSES AND SOLUTIONS.

Divanach, P., Kentouri, M., 2000. Hatchery techniques for specific diversification in Mediterranean finfish larviculture. Cah. Options Méditerranéennes 47, 75-87.

Divanach, P., Papandroulakis, N., Anastasiadis, P., Koumoundouros, G., Kentouri, M., 1997. Effect of water currents on the development of skeletal deformities in sea bass (Dicentrarchus labtax L.) with functional swimbladder during postlarval and nursery phase. Aquaculture 156, 145-155. https://doi.org/10.1016/S0044-8486(97)00072-0

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., Gingeras, T.R., 2013. STAR: Ultrafast universal RNA-seq aligner. Bioinformatics 29, 15-21. https://doi.org/10.1093/bioinformatics/bts635

Dominguez, D., Castro, P., Lall, S., Montero, D., Zamorano, M.J., Fontanillas, R., Izquierdo, M., 2022. Effects of Menadione Sodium Bisulphite (Vitamin K3) Supplementation of the Diets Based on Plant Feed Ingredients on Growth and Bone Health of Gilthead Seabream (Sparus aurata) Fingerlings. Aquac. Nutr. 2022, 1-8. https://doi.org/10.1155/2022/1613030

Dray, L., Neuhof, M., Diamant, A., Huchon, D., 2016. The complete mitochondrial genome of the gilthead seabream Sparus aurata L. (Sparidae). Mitochondrial DNA 27, 781-782. https://doi.org/10.3109/19401736.2014.928861

Falconer, D.S., Mackay, T.F.C., 2001. Introducción a la Genética Cuantitativa, Zaragoza: Acribia, 469.

Falgueras, J., Lara, A.J., Fernández-Pozo, N., Cantón, F.R., Pérez-Trabado, G., Claros, M.G., 2010. SeqTrim: A high-throughput pipeline for pre-processing any type of sequence read. BMC Bioinformatics 11. https://doi.org/10.1186/1471-2105-11-38

Faustino, M., Power, D.M., 2005. Osteologic development of the viscerocranial skeleton in sea bream: alternative ossification strategies in teleost fish. J. Fish Biol. 58, 537–572. https://doi.org/10.1111/j.1095-8649.2001.tb02272.x

Faustino, M., Power, D.M., 1998. Development of osteological structures in the sea bream: vertebral column and caudal fin complex. J. Fish Biol. 52, 11-22. https://doi.org/10.1111/j.1095-8649.1998.tb01548.x

Fernandez-Palacios, H., Izquierdo, M.S., Robaina, L., Valencia, A., Salhi, M., Montero, D., 1997. The effect of dietary protein and lipid from squid and fish meals on egg quality of broodstock for gilthead seabream (Sparus aurata), Aquaculture.

Fernandez-Palacios, H., Izquierdo, M.S., Robaina, L., Valencia, A., Salhi, M., Vergara, J.M., 1995. Effect of n-3 HUFA level in broodstock diets on egg quality of gilthead sea bream (Sparus aurata L.), Aquaculture.

Fernández, I., Hontoria, F., Ortiz-Delgado, J.B., Kotzamanis, Y., Estévez, A., Zambonino-Infante, J.L., Gisbert, E., 2008. Larval performance and skeletal deformities in farmed gilthead sea bream (Sparus aurata) fed with graded levels of Vitamin A enriched rotifers (Brachionus plicatilis). Aquaculture 283, 102-115. https://doi.org/10.1016/j.aquaculture.2008.06.037

Ferosekhan, S., Xu, H., Turkmen, S., Gómez, A., Afonso, J.M., Fontanillas, R., Rosenlund, G., Kaushik, S., Izquierdo, M.S., 2020. Reproductive performance of gilthead seabream (Sparus aurata) broodstock showing different expression of fatty acyl desaturase 2 and fed two dietary fatty acid profiles. Sci. Rep. 10. https://doi.org/10.1038/s41598-020-72166-5

Fragkoulis, S., Batargias, C., Kolios, P., Koumoundouros, G., 2018. Genetic parameters of the upper-jaw abnormalities in Gilthead seabream Sparus aurata. Aquaculture 497, 226-233. https://doi.org/10.1016/j.aquaculture.2018.07.071

Fragkoulis, S., Economou, I., Moukas, G., Koumoundouros, G., Batargias, C., 2020. Caudal fin abnormalities in Gilthead seabream (Sparus aurata L.) have a strong genetic variance component. J. Fish Dis. 43, 825-828. https://doi.org/10.1111/jfd.13180

Fragkoulis, S., Kerasovitis, D., Batargias, C., Koumoundouros, G., 2021. Body-shape trajectories and their genetic variance component in Gilthead seabream (Sparus aurata L.). Sci. Rep. 11, 1-11. https://doi.org/10.1038/s41598-021-95726-9

Fragkoulis, S., Paliogiannis, H., Kokkinias, P., Chiers, K., Adriaens, D., Koumoundouros, G., 2017. Saddleback syndrome in European sea bass Dicentrarchus labrax (Linnaeus, 1758): anatomy, ontogeny and correlation with lateral-line, anal and pelvic fin abnormalities. J. Fish Dis. 40, 83-95. https://doi.org/10.1111/jfd.12494

Franch, R., Louro, B., Tsalavouta, M., Chatziplis, 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, 851-861. https://doi.org/10.1534/genetics.106.059014

Froese, R., Pauly, D., 2019. Editors. FishBase. Publication., World Wide Web electronic Www.fishbase.org, (12/2019) [WWW Document].

García-Celdrán, M., Cutáková, Z., Ramis, G., Estévez, A., Manchado, M., Navarro, A., María-Dolores, E., Peñalver, J., Sánchez, J.A., Armero, E., 2016. Estimates of heritabilities and genetic correlations of skeletal deformities and uninflated swimbladder in a reared gilthead sea bream (Sparus aurata L.) juvenile population sourced from three broodstocks along the Spanish coasts. Aquaculture 464, 601-608. https://doi.org/10.1016/j.aquaculture.2016.08.004

García-Celdrán, M., Ramis, G., Manchado, M., Estévez, A., Afonso, J.M., María-Dolores, E., Peñalver, J., Armero, E., 2015. Estimates of heritabilities and genetic correlations of growth and external skeletal deformities at different ages in a reared gilthead sea bream (Sparus aurata L.) population sourced from three broodstocks along the Spanish coasts. Aquaculture 445, 33-41. https://doi.org/10.1016/j.aquaculture.2015.04.006

Garcia de la serrana, D., Mareco, E.A., LA Vieira, V., Power, D.M., Johnston, I.A., 2016. Comparison of the transcriptional responses of skeletal muscle and bone to a flooding dose of leucine in the gilthead sea bream (Sparus aurata). Comp. Biochem. Physiol. Part - B Biochem. Mol. Biol. 199, 50-57. https://doi.org/10.1016/j.cbpb.2016.02.003

Garrison, E., Marth, G., 2012. Haplotype-based variant detection from short-read sequencing 1-9.

Ge, S.X., Jung, D., Jung, D., Yao, R., 2020. ShinyGO: A graphical gene-set enrichment tool for animals and plants. Bioinformatics 36, 2628-2629. https://doi.org/10.1093/bioinformatics/btz931

Ge, S.X., Son, E.W., Yao, R., 2018. iDEP: An integrated web application for differential expression and pathway analysis of RNA-Seq data. BMC Bioinformatics 19, 1-24. https://doi.org/10.1186/s12859-018-2486-6

Georgakopoulou, E., Angelopoulou, A., Kaspiris, P., Divanach, P., Koumoundouros, G., 2007. Temperature effects on cranial deformities in European sea bass, Dicentrarchus labrax (L.). J. Appl. Ichthyol. 23, 99-103. https://doi.org/10.1111/j.1439-0426.2006.00810.x

Georgakopoulou, E., Katharios, P., Divanach, P., Koumoundouros, G., 2010. Effect of temperature on the development of skeletal deformities in Gilthead seabream (Sparus aurata Linnaeus, 1758). Aquaculture 308, 13-19. https://doi.org/10.1016/j.aquaculture.2010.08.006

Gjedrem, T., 2012. Genetic improvement for the development of efficient global aquaculture: A personal opinion review. Aquaculture 344-349, 12-22. https://doi.org/10.1016/j.aquaculture.2012.03.003

Gjedrem, T., 2004. Status for breeding programs in aquaculture, en: Fish Breeder's Roundtable 2004. Häholmen, Norway.

Gjedrem, T., 1997. Selective breeding to improve aquaculture production. World Aquac. Rouge 33-46.

Gjedrem, T., Baranski, M., 2009. Selective Breeding in Aquaculture: An Introduction., en: Reviews: Methods and Technologies in Fish Biology and Fisheries. Springer, Dordrecht, Netherlands, p. 221.

González Gayte, I., Bautista Moreno, R., Seoane Zonjic, P., Claros, M.G., 2017. DEgenes Hunter - A Flexible R Pipeline for Automated RNA-seq Studies in Organisms without Reference Genome. Genomics Comput. Biol. 3, 31. https://doi.org/10.18547/gcb.2017.vol3.iss3.e31

Gordon, A., Glazko, G., Qiu, X., Yakovlev, A., 2007. Control of the mean number of false discoveries, Bonferroni and stability of multiple testing. Ann. Appl. Stat. 1, 179-190. https://doi.org/10.1214/07-aoas102

Gorshkov, S., Gordin, H., Gorshkova, G., Knibb, W., 1997. Reproductive constraints for family selection of the gilthead seabream (Sparus aurata L.). Isr. J. Aquac. - Bamidgeh 49, 124-134.

Groeneveld, E., Kovaˇc, M., Mielenz, N., 2010. VCE User's Guide and Reference Manual Version 6.0. (ftp://ftp.tzv.fal.de/pub/vce6/doc/vce6-manual-3.1-A4.pdf).

Haga, Y., Du, S.J., Satoh, S., Kotani, T., Fushimi, H., Takeuchi, T., 2011. Analysis of the

mechanism of skeletal deformity in fish larvae using a vitamin A-induced bone deformity model. Aquaculture 315, 26-33. https://doi.org/10.1016/j.aquaculture.2010.11.026

Ibarra-Zatarain, Z., Duncan, N., 2015. Mating behaviour and gamete release in gilthead seabream (Sparus aurata, Linnaeus 1758) held in captivity. Spanish J. Agric. Res. 13, 11. https://doi.org/10.5424/sjar/2015131-6750

Izquierdo, M.S., Scolamacchia, M., Betancor, M., Roo, J., Caballero, M.J., Terova, G., Witten, P.E., 2013. Effects of dietary DHA and α-tocopherol on bone development, early mineralisation and oxidative stress in Sparus aurata (Linnaeus, 1758) larvae. Br. J. Nutr. 109, 1796-1805. https://doi.org/10.1017/S0007114512003935

Izquierdo, M.S., Socorro, J., Roo Filgueira, F.J., 2010. Studies on the appearance of skeletal anomalies in red porgy: effect of culture intensiveness, feeding habits and nutritional quality of live preys.

Janssen, K., Chavanne, H., Berentsen, P., Komen, H., 2017. Impact of selective breeding on European aquaculture. Aquaculture 472, 8-16. https://doi.org/10.1016/j.aquaculture.2016.03.012

Jerez, S., Rodríguez, C., Cejas, J.R., Martín, M. V., Bolaños, A., 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. https://doi.org/10.1016/j.aquaculture.2012.04.018

Jongeneel, C.V., Delorenzi, M., Iseli, C., Zhou, D., Haudenschild, C.D., Khrebtukova, I., Kuznetsov, D., Stevenson, B.J., Strausberg, R.L., Simpson, A.J.G., Vasicek, T.J., 2005. An atlas of human gene expression from massively parallel signature sequencing (MPSS). Genome Res. 15, 1007-1014. https://doi.org/10.1101/gr.4041005

Jue, N.K., Foley, R.J., Reznick, D.N., O'Neill, R.J., O'Neill, M.J., 2018. Tissue-specific transcriptome for Poeciliopsis prolifica reveals evidence for genetic adaptation related to the evolution of a placental fish. G3 Genes, Genomes, Genet. 8, 2181-2192. https://doi.org/10.1534/g3.118.200270

Karahan, B., Chatain, B., Chavanne, H., Vergnet, A., Bardon-Albaret, A., Haffray, P., Dupont-Nivet, M., Vandeputte, M., 2013. Heritabilities and correlations of deformities and growth-related traits in the European sea bass (Dicentrarchus labrax, L) in four different sites. Aquac. Res. 44. https://doi.org/10.1111/j.1365-2109.2011.03082.x

Kause, A., Ritola, O., Paananen, T., 2007. Changes in the expression of genetic characteristics across cohorts in skeletal deformations of farmed salmonids. Genet. Sel. Evol. 39, 529-543. https://doi.org/10.1051/gse:2007019

Kieffer, D.A., Piccolo, B.D., Vaziri, N.D., Liu, S., Lau, W.L., Khazaeli, M., Nazertehrani, S., Moore, M.E., Marco, M.L., Martin, R.J., Adams, S.H., 2016. Resistant starch alters gut microbiome and metabolomic profiles concurrent with amelioration of chronic kidney disease in rats. Am. J. Physiol. - Ren. Physiol. 310, F857-F871. https://doi.org/10.1152/ajprenal.00513.2015

Kihara, M., Ogata, S., Kawano, N., Kubota, I., Yamaguchi, R., 2002. Lordosis induction in juvenile red sea bream, Pagrus major, by high swimming activity. Aquaculture 212, 149-158. https://doi.org/10.1016/S0044-8486(01)00871-7

Kim, D., Paggi, J.M., Park, C., Bennett, C., Salzberg, S.L., 2019. Graph-Based Genome Alignment and Genotyping with HISAT2 and HISAT-genotype. Nat. Biotechnol. 37, 907-915. https://doi.org/10.1038/s41587-019-0201-4.

Kincaid, H.L., 1983. INBREEDING IN FISH POPULATIONS USED FOR AQUACULTURE. Aquaculture 33, 215-227.

Kissil, G.W., Lupatsch, I., Elizur, A., Zohar, Y., 2001. Long photoperiod delayed spawning and increased somatic growth in gilthead seabream (Sparus aurata), Aquaculture.

Kjørsvik, E., Hoehne-Reitan, K., Reitan, K.I., 2003. Egg and larval quality criteria as predictive measures for juvenile production in turbot (Scophthalmus maximus L.). Aquaculture 227, 9-20. https://doi.org/10.1016/S0044-8486(03)00492-7

Kjørsvik, E., Mangor-Jensen, A., Holmefjord, I., 1990. Egg Quality in Fishes, en: Blaxter, J.H.S., Southward, A.J. (Eds.), Advances in Marine Biology. Academic Press, pp. 71-113. https://doi.org/https://doi.org/10.1016/S0065-2881(08)60199-6

Knibb, W., 2000. Genetic improvement of marine fish – which method for industry? Aquac. Res. 31, 11-23.

Knibb, W., Gorshkova, G., Gorshkov, S., 1998. Selection and crossbreeding in Mediterranean cultured marine fish.

Knibb, W., Gorshkova, G., Gorshkov, S., 1997. Selection for growth in the gilthead seabream, Sparus auratai L. Isr. J. Aquac. 49, 57-66.

Knudsen, S., 1999. Promoter2.0: For the recognition of PolII promoter sequences. Bioinformatics 15, 356-361. https://doi.org/10.1093/bioinformatics/15.5.356

Kofler, R., Orozco-terWengel, P., de Maio, N., Pandey, R.V., Nolte, V., Futschik, A., Kosiol, C., Schlötterer, C., 2011. Popoolation: A toolbox for population genetic analysis of next generation sequencing data from pooled individuals. PLoS One 6. https://doi.org/10.1371/journal.pone.0015925

Komen, H., Haffray, P., Kaushik, S., New, M., Olesen, I., Liinamo, A.E., 2002. Defining breeding goals for future sustainable aquaculture. World Aquac. 33.

Koumoundouros, G., 2010. Morpho-anatomical abnormalities in Mediterranean marine aquaculture. Recent Adv. Aquac. Res. 661, 125-148.

Koumoundouros, G., Divanach, P., Anezaki, L., Kentouri, M., 2001. Temperature-induced ontogenetic plasticity in sea bass (Dicentrarchus labrax). Mar. Biol. 139, 817-830. https://doi.org/10.1007/s002270100635

Koumoundouros, G., Gagliardi, F., Divanach, P., Boglione, C., Cataudella, S., Kentouri, M., 1997a. Normal and abnormal osteological development of caudal fin in Sparus aurata L.

fry. Aquaculture 149, 215-226.

Koumoundouros, G., Oran, G., Divanach, P., Stefanakis, S., Kentouri, M., 1997b. The opercular complex deformity in intensive gilthead sea bream (Sparus aurata L.) larviculture. Moment of apparition and description, Aquaculture.

Lahnsteiner, F., Giménez, G., Estévez, A., 2008. Egg quality determination based on the shape of the lipid vesicle in common dentex, Dentex dentex. Aquac. Res. 39, 144-149. https://doi.org/10.1111/j.1365-2109.2007.01865.x

Lahnsteiner, F., Patarnello, P., 2005. The shape of the lipid vesicle is a potential marker for egg quality determination in the gilthead seabream, Sparus aurata, and in the sharpsnout seabream, Diplodus puntazzo. Aquaculture 246, 423-435. https://doi.org/10.1016/j.aquaculture.2005.01.018

Lahnsteiner, F., Patarnello, P., 2004. Egg quality determination in the gilthead seabream, Sparus aurata, with biochemical parameters. Aquaculture 237, 443-459. https://doi.org/10.1016/j.aquaculture.2004.04.017

Lee-Montero, I., Navarro, A., Borrell, Y., García-Celdrán, M., Martín, N., Negrín-Báez, D., Blanco, G., Armero, E., Berbel, C., Zamorano, M.J., Sánchez, J.J., Estévez, A., Ramis, G., Manchado, M., Afonso, J.M., 2013. Development of the first standardised panel of two new microsatellite multiplex PCRs for gilthead seabream (Sparus aurata L.). Anim. Genet. 44, 533-546. https://doi.org/10.1111/age.12037

Lee-Montero, I., Navarro, A., Negrín-Báez, D., Zamorano, M.J., Berbel, C., Sánchez, J.A., García-Celdran, M., Manchado, M., Estévez, A., Armero, E., Afonso, J.M., 2015. Genetic parameters and genotype-environment interactions for skeleton deformities and growth traits at different ages on gilthead seabream (Sparus aurata L.) in four Spanish regions. Anim. Genet. 46, 164-174. https://doi.org/10.1111/age.12258

Lemly, A.D., 1993. Teratogenic Effects of Selenium in Natural Populations of Fresh Water Fish. Ecotoxicol. Environ. Saf. https://doi.org/10.1006/eesa.1993.1049

León-Bernabeu, S., Shin, H.S., Lorenzo-Felipe, Á., García-Pérez, C., Berbel, C., Elalfy, I.S., Armero, E., Pérez-Sánchez, J., Arizcun, M., Zamorano, M.J., Manchado, M., Afonso, J.M., 2021. Genetic parameter estimations of new traits of morphological quality on gilthead seabream (Sparus aurata) by using IMAFISH_ML software. Aquac. Reports 21. https://doi.org/10.1016/j.aqrep.2021.100883

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078-2079. https://doi.org/10.1093/bioinformatics/btp352

Li, H., Ma, M.L., Luo, S., Zhang, R.M., Han, P., Hu, W., 2012. Metabolic responses to ethanol in Saccharomyces cerevisiae using a gas chromatography tandem mass spectrometry-based metabolomics approach. Int. J. Biochem. Cell Biol. 44, 1087-1096. https://doi.org/10.1016/j.biocel.2012.03.017

Liao, Y., Smyth, G.K., Shi, W., 2014. FeatureCounts: An efficient general purpose program for

assigning sequence reads to genomic features. Bioinformatics 30, 923-930. https://doi.org/10.1093/bioinformatics/btt656

Liao, Y., Wang, J., Jaehnig, E.J., Shi, Z., Zhang, B., 2019. WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. Nucleic Acids Res. 47, W199-W205. https://doi.org/10.1093/nar/gkz401

Loizides, M., Georgiou, A.N., Somarakis, S., Witten, P.E., Koumoundouros, G., 2014. A new type of lordosis and vertebral body compression in Gilthead sea bream, Sparus aurata L.: Aetiology, anatomy and consequences for survival. J. Fish Dis. 37, 949-957. https://doi.org/10.1111/jfd.12189

Lom, J., Pike, A.W., Dykova, I., 1991. Myxobolus sandrae, the agent of vertebral column deformities of perch Perca fluviatilis in northeast Scotland. Dis. Aquat. Organ. 12, 49-53.

Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 1-21. https://doi.org/10.1186/s13059-014-0550-8

Manchado, M., Planas, J. V., Cousin, X., Rebordinos, L., Claros, M.G., 2016. Current status in other finfish species. Genomics Aquac. 195-221. https://doi.org/10.1016/b978-0-12-801418-9.00008-1

Martinez, V., 2007. Marker-assisted selection in fish and shellfish breeding schemes, en: Current status and future perspectives in crops, livestock, forestry and fish. FAO, Rome, pp. 329-362.

Meuwissen, T., 2003. Genomic Selection : the Future of Marker Assisted Selection and Animal Breeding, en: Marker-Assisted Selection: A Fast Track to Increase Genetic Gain in Plant and Animal Breeding? Session II: MAS in Animals. FAO, Turin, Italy, pp. 54-59.

Meuwissen, T.H.E., Hayes, B.J., Goddard, M.E., 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157, 1819-1829. https://doi.org/10.1093/genetics/157.4.1819

Mhalhel, K., Germanà, A., Abbate, F., Guerrera, M.C., Levanti, M., Laurà, R., Montalbano, G., 2020. The effect of orally supplemented melatonin on larval performance and skeletal deformities in farmed gilthead seabream (Sparus aurata). Int. J. Mol. Sci. 21, 1-18. https://doi.org/10.3390/ijms21249597

Migaud, H., Bell, G., Cabrita, E., McAndrew, B.J., Davie, A., Bobe, J., Herráez, M.P., Carrillo, M., 2013. Gamete quality and broodstock management in temperate fish. Rev. Aquac. https://doi.org/10.1111/raq.12025

Moretti, A., Fernandez-Criado, M.P., Cittolin, G., Guidastri, R., 1999. Manual on Hatchery Production of Seabass and Gilthead Seabream. Rome, FAO Volume 1, 194.

Morrison, C.M., MacDonald, C.A., 1995. Normal and abnormal jaw development of the yolk-sac larva of Atlantic halibut Hippoglossus hippoglossus. Dis. Aquat. Organ. 22, 173-184. https://doi.org/10.3354/dao022173

Myneni, V.D., Mousa, A., Kaartinen, M.T., 2016. Factor XIII-A transglutaminase deficient mice show signs of metabolically healthy obesity on high fat diet. Sci. Rep. 6, 1-11. https://doi.org/10.1038/srep35574

Navarro, A., Oliva, V., Zamorano, M.J., Ginés, R., Izquierdo, M.S., Astorga, N., Afonso, J.M., 2006. Evaluation of PIT system as a method to tag fingerlings of gilthead seabream (Sparus auratus L.): Effects on growth, mortality and tag loss. Aquaculture 257, 309-315. https://doi.org/10.1016/j.aquaculture.2006.02.072

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. https://doi.org/10.1016/j.aquaculture.2008.12.024

Negrín-Báez, D., Navarro, A., Afonso, J.M., Ginés, R., Zamorano, M.J., 2015a. Detection of QTL associated with three skeletal deformities in gilthead seabream (Sparus aurata L.): Lordosis, vertebral fusion and jaw abnormality. Aquaculture 448, 123-127. https://doi.org/10.1016/j.aquaculture.2015.05.025

Negrín-Báez, D., Navarro, A., Afonso, J.M., Toro, M.A., Zamorano, M.J., 2016a. Quantitative trait loci for a neurocranium deformity, lack of operculum, in gilthead seabream (Sparus aurata L.). Anim. Genet. 47, 230-233. https://doi.org/10.1111/age.12397

Negrín-Báez, D., Navarro, A., Lee-Montero, I., Soula, M., Afonso, J.M., Zamorano, M.J., 2015b. Inheritance of skeletal deformities in gilthead seabream (Sparus aurata) –lack of operculum, lordosis, vertebral fusion and LSK complex. J. Anim. Sci. 93, 53-61. https://doi.org/10.2527/jas.2014-7968

Negrín-Báez, D., Navarro, A., Rodríguez-Ramilo, S.T., Afonso, J.M., Zamorano, M.J., 2016b. Identification of Quantitative Trait Loci Associated with the Skeletal Deformity LSK complex in Gilthead Seabream (Sparus aurata L.). Mar. Biotechnol. 18, 98-106. https://doi.org/10.1007/s10126-015-9671-7

Neumaier, A., Groeneveld, E., 1998. Restricted maximum likelihood estimation of covariances in sparse linear models. Genet. Sel. Evol. 30, 3-26. https://doi.org/10.1051/gse:19980101

Ogura, Y., Kou, I., Miura, S., Takahashi, A., Xu, L., Takeda, K., Takahashi, Y., Kono, K., Kawakami, N., Uno, K., Ito, M., Minami, S., Yonezawa, I., Yanagida, H., Taneichi, H., Zhu, Z., Tsuji, T., Suzuki, T., Sudo, H., Kotani, T., Watanabe, K., Hosogane, N., Okada, E., Iida, A., Nakajima, M., Sudo, A., Chiba, K., Hiraki, Y., Toyama, Y., Qiu, Y., Shukunami, C., Kamatani, Y., Kubo, M., Matsumoto, M., Ikegawa, S., 2015. A Functional SNP in BNC2 Is Associated with Adolescent Idiopathic Scoliosis. Am. J. Hum. Genet. 97, 337-342. https://doi.org/10.1016/j.ajhg.2015.06.012

Pauletto, M., Manousaki, T., Ferraresso, S., Babbucci, M., Tsakogiannis, A., Louro, B., Vitulo, N., Quoc, V.H., Carraro, R., Bertotto, D., Franch, R., Maroso, F., Aslam, M.L., Sonesson, A.K., Simionati, B., Malacrida, G., Cestaro, A., Caberlotto, S., Sarropoulou, E., Mylonas, C.C., Power, D.M., Patarnello, T., Canario, A.V.M., Tsigenopoulos, C., Bargelloni, L., 2018.

Genomic analysis of Sparus aurata reveals the evolutionary dynamics of sex-biased genes in a sequential hermaphrodite fish. Commun. Biol. 1. https://doi.org/10.1038/s42003-018-0122-7

Pérez-Enciso, M., Toro, M.A., 2007. Loclizacion De Genes Y Seleccion Mediante Marcadores Moleculares, en: Genética y Genómica en Acuicultura. Madrid, pp. 361-400.

Pérez-Sánchez, J., Naya-Català, F., Soriano, B., Piazzon, M.C., Hafez, A., Gabaldón, T., Llorens, C., Sitjà-Bobadilla, A., Calduch-Giner, J.A., 2019. Genome Sequencing and Transcriptome Analysis Reveal Recent Species-Specific Gene Duplications in the Plastic Gilthead Sea Bream (Sparus aurata). Front. Mar. Sci. 6, 1-18. https://doi.org/10.3389/fmars.2019.00760

Perkins, J., Seoane Zonjic, P., Moreno Jabato, F., Córdoba Caballero, J Rojano Rivera, E., Bautista Moreno, R., Claros, M., Gonzalez Gayte, I., García Ranea, J., 2021. ExpHunterSuite: Package For The Comprehensive Analysis Of Transcriptomic Data. R package version 1.2.0.

Piskol, R., Ramaswami, G., Li, J.B., 2013. Reliable identification of genomic variants from RNA-seq data. Am. J. Hum. Genet. 93, 641-651. https://doi.org/10.1016/j.ajhg.2013.08.008

Polo, A., Yúfera, M., Pascual, E., 1991. Effects of temperature on egg and larval development of Sparus aurata L. Aquaculture 92, 367-375. https://doi.org/https://doi.org/10.1016/0044-8486(91)90042-6

Potthoff, T., 1984. Clearing and Staining Techniques, en: Ontogeny and systematics of fishes : based on an international symposium dedicated to the memory of Elbert Halvor Ahlstrom / sponsored by the National Marine Fisheries Service, National Oceanic and Atmospheric Administration, United States Dept. of Commer. pp. 35-37. https://doi.org/10.5962/bhl.title.4434

Prestinicola, L., Boglione, C., Makridis, P., Spanò, A., Rimatori, V., Palamara, E., Scardi, M., Cataudella, S., 2013. Environmental Conditioning of Skeletal Anomalies Typology and Frequency in Gilthead Seabream (Sparus aurata L., 1758) Juveniles. PLoS One 8. https://doi.org/10.1371/journal.pone.0055736

Qian, X., Ba, Y., Zhuang, Q., Zhong, G., 2014. RNA-seq technology and its application in fish transcriptomics. Omi. A J. Integr. Biol. 18, 98-110. https://doi.org/10.1089/omi.2013.0110

Quinlan, A.R., Hall, I.M., 2010. BEDTools: A flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841-842. https://doi.org/10.1093/bioinformatics/btq033

Riera-Heredia, N., Lutfi, E., Gutiérrez, J., Navarro, I., Capilla, E., 2019a. Fatty acids from fish or vegetable oils promote the adipogenic fate of mesenchymal stem cells derived from gilthead sea bream bone potentially through different pathways. PLoS One 14, 1-21. https://doi.org/10.1371/journal.pone.0215926

Riera-Heredia, N., Martins, R., Mateus, A.P., Costa, R.A., Gisbert, E., Navarro, I., Gutiérrez, J., Power, D.M., Capilla, E., 2018. Temperature responsiveness of gilthead sea bream bone;

an in vitro and in vivo approach. Sci. Rep. 8, 1-14. https://doi.org/10.1038/s41598-018-29570-9

Riera-Heredia, N., Vélez, E.J., Gutiérrez, J., Navarro, I., Capilla, E., 2019b. Gene expression analyses in malformed skeletal structures of gilthead sea bream (Sparus aurata). J. Fish Dis. 42, 1169-1180. https://doi.org/10.1111/jfd.13019

Salem, M., Paneru, B., Al-Tobasei, R., Abdouni, F., Thorgaard, G.H., Rexroad, C.E., Yao, J., 2015. Transcriptome assembly, gene annotation and tissue gene expression atlas of the rainbow trout. PLoS One 10, 1-27. https://doi.org/10.1371/journal.pone.0121778

Saroglia, M., Liu, Z.J., 2012. Functional Genomics in Aquaculture, Functional Genomics in Aquaculture. https://doi.org/10.1002/9781118350041

Sarropoulou, E., Kotoulas, G., Power, D.M., Geisler, R., 2005. 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. https://doi.org/10.1152/physiolgenomics.00139.2005

Scabini, V., Fernandez-Palacios, H., Robaina, L., Kalinowski, T., Izquierdo, M.S., 2011. Reproductive performance of gilthead seabream (Sparus aurata L., 1758) fed two combined levels of carotenoids from paprika oleoresin and essential fatty acids. Aquac. Nutr. 17, 304-312. https://doi.org/10.1111/j.1365-2095.2010.00766.x

Schmieder, R., Edwards, R., 2011. Quality control and preprocessing of metagenomic datasets. Bioinformatics 27, 863-864. https://doi.org/10.1093/bioinformatics/btr026

Senger, F., Priat, C., Hitte, C., Sarropoulou, E., Franch, R., Geisler, R., Bargelloni, L., Power, D., Galibert, F., 2006. The first radiation hybrid map of a perch-like fish: The gilthead seabream (Sparus aurata L). Genomics 87, 793-800. https://doi.org/10.1016/j.ygeno.2005.11.019

Sfakianakis, D.G., Georgakopoulou, E., Papadakis, I.E., Divanach, P., Kentouri, M., Koumoundouros, G., 2006. Environmental determinants of haemal lordosis in European sea bass, Dicentrarchus labrax (Linnaeus, 1758). Aquaculture 254, 54-64. https://doi.org/10.1016/j.aquaculture.2005.10.028

Sfakianakis, D.G., Koumoundouros, G., Divanach, P., Kentouri, M., 2004. Osteological development of the vertebral column and of the fins in Pagellus erythrinus (L. 1758). Temperature effect on the developmental plasticity and morpho-anatomical abnormalities. Aquaculture 232, 407-424. https://doi.org/10.1016/j.aquaculture.2003.08.014

Shields, R.J., 2001. Larviculture of marine finfish in Europe. Aquaculture 200, 55-88.

Shields, R.J., Brown, N.P., Bromage, N.R., 1997. Blastomere morphology as a predictive measure of fish egg viability. Aquaculture 155, 1-12. https://doi.org/10.1016/S0044-8486(97)00105-1

Sitjà-Bobadilla, A., Padrós, F., Aguilera, C., Alvarez-Pellitero, P., 2005. Epidemiology of

Cryptosporidium molnari in Spanish gilthead sea bream (Sparus aurata L.) and European sea bass (Dicentrarchus labrax L.) cultures: From hatchery to market size. Appl. Environ. Microbiol. 71, 131-139. https://doi.org/10.1128/AEM.71.1.131-139.2005

Slooff, W., 1982. Skeletal anomalies in fish from polluted surface waters. Aquat. Toxicol. 2, 157-173. https://doi.org/https://doi.org/10.1016/0166-445X(82)90013-3

Socorro, C.J.A., 2006. Estudio Comparado del Desarrollo Embrionario y Larvario del Bocinegro (Pagrus pagrus) y de la Sama de Pluma (Dentex gibbosus). University of Las Palmas de Gran Canaria, Spain.

Tandler, A., Har'el, M., Wilks, M., Levinson, A., Brickell, L., Christie, S., Avital, E., Barr, Y., 1989. Effect of environmental temperature on survival, growth and population structure in the mass rearing of the gilthead seabream, Sparus aurata. Aquaculture 78, 277-284. https://doi.org/https://doi.org/10.1016/0044-8486(89)90105-1

Taylor, W., Van Dyke, G., 1985. Revised procedures for staining and clearing small fishes and other vertebrates for bone and cartilage study. Cybium (Paris) 9, 107-119.

Theodorou, J.A., Perdikaris, C., Venou, B., 2016. Origin of broodstock and effects on the deformities of gilthead sea bream (Sparus aurata L. 1758) in a Mediterranean commercial hatchery. Int. Aquat. Res. 8, 275-282. https://doi.org/10.1007/s40071-016-0140-3

Thorland, I., Kottaras, L., Refstie, T., Dimitroglou, A., Papaharisis, L., Rye, M., 2015. Response to Selection for Harvest Weight in a Family Based Selection Program of Gilthead Seabream (Sparus aurata), en: XII international symposium on genetics in aquaculture, Santiago de Compostela, Spain.

Thorland, I., Papaioannou, N., Kottaras, L., Refstie, T., Papasolomontos, S., Rye, M., 2007. Family based selection for production traits in gilthead seabream (Sparus aurata) and European sea bass (Dicentrarchus labrax) in Greece. Aquaculture 272, S313-S314. https://doi.org/10.1016/j.aquaculture.2007.07.192

Thuong, N.P., Verstraeten, B., Kegel, B.D., Christiaens, J., Wolf, T.D., Sorgeloos, P., Bonte, D., Adriaens, D., 2017. Ontogenesis of opercular deformities in gilthead sea bream Sparus aurata: a histological description. J. Fish Biol. 91, 1419-1434. https://doi.org/10.1111/jfb.13460

Vandeputte, M., Gagnaire, P.A., Allal, F., 2019. The European sea bass: a key marine fish model in the wild and in aquaculture. Anim. Genet. 50, 195-206. https://doi.org/10.1111/age.12779

Vandeputte, M., Mauger, S., Dupont-Nivet, M., 2006. An evaluation of allowing for mismatches as a way to manage genotyping errors in parentage assignment by exclusion. Mol. Ecol. Notes 6, 265-267. https://doi.org/10.1111/j.1471-8286.2005.01167.x

Vieira, F.A., Thorne, M.A.S., Stueber, K., Darias, M., Reinhardt, R., Clark, M.S., Gisbert, E., Power, D.M., 2013. Comparative analysis of a teleost skeleton transcriptome provides

insight into its regulation. Gen. Comp. Endocrinol. 191, 45-58. https://doi.org/10.1016/j.ygcen.2013.05.025

Wang, W., Huang, Y., Yu, Y., Yang, Y., Xu, M., Chen, X., Ni, S., Qin, Q., Huang, X., 2016. Fish TRIM39 regulates cell cycle progression and exerts its antiviral function against iridovirus and nodavirus. Fish Shellfish Immunol. 50, 1-10. https://doi.org/10.1016/j.fsi.2016.01.016

Whittle, D.M., Sergeant, D.B., Huestis, S.Y., Hyatt, W.H., 1992. Foodchain accumulation of PCDD and PCDF isomers in the Great Lakes aquatic community. Chemosphere 25, 181-184. https://doi.org/https://doi.org/10.1016/0045-6535(92)90508-O

Xu, H., Turkmen, S., Rimoldi, S., Terova, G., Zamorano, M.J., Afonso, J.M., Sarih, S., Fernandez-Palacios, H., Izquierdo, M.S., 2019. Nutritional intervention through dietary vegetable proteins and lipids to gilthead sea bream (Sparus aurata) broodstock affects the offspring utilization of a low fishmeal/fish oil diet. Aquaculture 513. https://doi.org/10.1016/j.aquaculture.2019.734402

Yue, G.H., 2014. Recent advances of genome mapping and marker-assisted selection in aquaculture. Fish Fish. 15, 376-396. https://doi.org/10.1111/faf.12020

Zar, J.H., 1984. Biostatistical Analysis, 3rd ed. ed. Upper Saddle River, NJ.

Zhang, L., Mei, Y., Fu, N.Y., Guan, L., Xie, W., Liu, H.H., Yu, C.D., Yin, Z., Yu, V.C., You, H., 2012. TRIM39 regulates cell cycle progression and DNA damage responses via stabilizing p21. Proc. Natl. Acad. Sci. U. S. A. 109, 20937-20942. https://doi.org/10.1073/pnas.1214156110