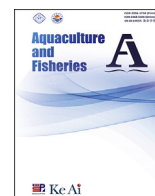


Contents lists available at [ScienceDirect](http://ScienceDirect)

## Aquaculture and Fisheries

journal homepage: [www.keaipublishing.com/en/journals/aquaculture-and-fisheries/](http://www.keaipublishing.com/en/journals/aquaculture-and-fisheries/)

## SuperSAGE digital expression analysis of differential growth rate in a European sea bass population

Bruno Louro<sup>a,\*</sup>, Rute S.T. Martins<sup>a</sup>, Patricia I.S. Pinto<sup>a</sup>, Richard Reinhardt<sup>b</sup>, Dirk-Jan de Koning<sup>c,d</sup>, Adelino V.M. Canario<sup>a</sup>, Deborah M. Power<sup>a</sup><sup>a</sup> CCMAR-Centre of Marine Sciences, University of Algarve, Campus de Gambelas, Faro 8005-139, Portugal<sup>b</sup> Max Planck Genome Centre, Carl-von-Linné-Weg 10, Köln 50829, Germany<sup>c</sup> The Roslin Institute and R(D)SVS, University of Edinburgh, Easter Bush, Midlothian EH25 9RG, Scotland UK<sup>d</sup> Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Box 7023, Uppsala 750 07, Sweden

## ARTICLE INFO

## Keywords:

European sea bass  
Specific growth rate  
Gene expression  
SuperSAGE  
Pathways

## ABSTRACT

One of the goals of the aquaculture industry is to understand and control growth associated traits through selective breeding. In the present study the molecular basis of growth heterogeneity in the European sea bass (*Dicentrarchus labrax*) was addressed. To establish growth heterogeneity in a group of hatchery bred sea bass individuals were tagged and their specific growth rates (SGR) determined at monthly intervals. Gene expression in the brain, liver and white muscle from fish with the most divergent sustained SGR (6 individuals of the first and last quartile) was assessed using SuperSAGE (Serial Analysis Gene Expression) combined with next generation SOLiD4 sequencing. A total of approx. 11 million edited tags (26 bp), on average 2 million tags per SAGE library, that represented 47,071 unique transcripts were identified. Comparison of transcripts in fish with high and low SGR yielded 344, 698 and 601 differently expressed tags (0.01% false discovery rate and 4-fold change) in brain, liver and muscle, respectively. The tags were mapped onto the sea bass genome and approximately one third of the tags could be assigned to annotated genes. Pathway enrichment analysis revealed in liver, muscle and brain intricate gene expression changes in endocrine regulatory pathways involved in growth, metabolic and the stress axis, underlying divergent SGR in sea bass.

## 1. Introduction

The aquaculture industry has the same general aspirations as terrestrial production systems, which is to enhance economically important traits (Canario et al, 2008). However, the advantages of selective breeding programs have been exploited for only relatively few aquaculture species and an optimistic estimate suggests that probably only 10% of aquaculture production is based on genetically improved stocks (Gjedrem & Baranski, 2009, p. 221). The salmonids, Atlantic salmon (*Salmo salar*) and rainbow trout (*Oncorhynchus mykiss*), are the main products of Northern European aquaculture and much of the production is based on genetically improved stocks (Gjedrem & Baranski, 2009, p. 221). In relatively few generations selection for improved growth yielded a 10%–14% size increase in Atlantic salmon (Gjøen & Bentsen, 1997) and 8%–13% in rainbow trout (Kause et al, 2005). However, with the more recent adoption of intensive aquaculture production systems, Southern European species lack large-scale formal selection programs to improve

economically important traits. This is the case of the European sea bass (*Dicentrarchus labrax*) with a production estimated at 126 thousand tonnes and a market value of 500 million Euros (FAO, 2012). This species has a molecular resource rich status and is a prime candidate for the development of a genetic selection program using molecular genetics (Canario et al, 2008). The use of molecular approaches in sea bass has led in a relatively short space of time to the establishment of hundreds of microsatellite markers and to the production of several linkage maps (Chistiakov et al, 2005, 2008) and candidate quantitative trait loci (QTL) (Chatziplis et al, 2007; Louro et al, 2016; Massault et al, 2010). Additional resources include a >12× coverage BAC-library (Whitaker, McAndrew, & Taggart, 2006), a radiation hybrid map (Guyon et al, 2010), over 30,000 expressed sequence tags (ESTs) (Louro et al, 2010), an oligonucleotide microarray (Ferraresso et al, 2010) and a well assembled and annotated genome (Tine et al, 2014).

While QTL mapping and a sequenced genome of an organism are important steps towards molecular assisted selection, still few causative

\* Corresponding author. CCMAR-Centre of Marine Sciences, University of Algarve, Campus de Gambelas, Faro 8005-139, Portugal.

E-mail address: [blouro@ualg.pt](mailto:blouro@ualg.pt) (B. Louro).<https://doi.org/10.1016/j.aaf.2018.03.001>

Received 19 November 2017; Received in revised form 1 March 2018; Accepted 5 March 2018

Available online 21 March 2018

2468-550X/© 2018 Shanghai Ocean University. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

genes behind the phenotype of interest have been identified. Even in terrestrial farm animals, few causative genes of QTLs have been identified (Mackay, Stone, & Ayroles, 2009; Rothschild & Plastow, 2008; Rothschild, Hu, & Jiang, 2007). Recent approaches to speed up identification of the genes associated with traits of interest have deployed a combination of QTL mapping and high throughput transcriptome analysis, also known as Systems Genomics or Genetical Genomics, when gene expression is considered the trait rather than the physical phenotype itself (Hubner et al, 2005; Schadt, Monks, & Friend, 2003; Wayne & McIntyre, 2002). Other strategies to define new QTL or loci for complex traits have used QTL analysis coupled to gene transcriptome analysis of individuals with distinct phenotype subtypes in a segregating population (Blum et al, 2010; Schadt et al., 2003).

In the present study the absence of individuals selected for improved growth phenotypes or with a known pedigree with a characterised QTL meant that genetical genomics was not feasible. The aim of the present study was therefore to identify differential gene expression profiles between high and low SGR fishes, i.e. that underlie divergent extreme growth phenotypes in the European sea bass. Furthermore, the existence of shared and tissue specific transcripts were analysed in three tissues selected for transcriptome analysis. The tissues were selected based on their relevance to the growth phenotype and included the liver (related to metabolism), muscle (body mass growth) and the brain (regulation). SuperSAGE (Serial Analysis Gene Expression) (Matsumura et al, 2010) was used to detect differential transcript expression between chronically slow and fast growers and cross referenced with the genes previously identified in candidate growth QTL identified in an independent mapping population (Massault et al, 2010). Pathway analysis of differentially expressed transcripts established the regulatory pathways most affected and gave insight into the tissue specific changes underlying divergent specific growth rate in sea bass.

## 2. Materials and methods

### 2.1. Growth trial

European sea bass juveniles obtained from a commercial fish farm (Viveiros Vila Nova, Vila Nova de Milfontes, Portugal) were grown at Ramalhe Marine station (University of Algarve, Faro, Portugal) in 1000 L tanks with running seawater at a density < 5 kg/m<sup>3</sup>, ambient temperature (16–22 °C) and natural photoperiod (from February to June, Latitude 37°). For the 70 days growth trial, two groups of 9 months old fish were created ( $n = 75$ ), one from below the first weight quartile (Q1 = 40 g) and the other from above the third weight quartile (Q3 = 53 g). The fish had no visible external abnormalities or lesions and were pit tagged before being randomly distributed between two cylindrical tanks (75 L). The tanks were in an open circuit and received a continuous flow-through of aerated sea water at 21–22 °C and were exposed to the natural photoperiod for June–September. Fish were fed to satiation with a continuous feeding regime and verifying periodically during the day that all food was consumed.

At monthly intervals fish were lightly anaesthetised (0.01% 2-phenoxyethanol) and weight and length measured. For final phenotyping and tissue collection all fish were sacrificed 24 h after last meal with an overdose of 2-phenoxyethanol. Whole brain (minus the pituitary gland), liver and white muscle were rapidly collected and immediately frozen in liquid nitrogen and stored at –80 °C until use. Specific growth rate (SGR) was calculated for each individual as

$$SGR = (\ln(W_f) - \ln(W_i)) * 100 / t$$

where  $\ln(W_f)$  is the natural logarithm of the final weight (g),  $\ln(W_i)$  is the natural logarithm of the initial weight (g), and  $t$  is the interval between measurements (days). The six fish with the most divergent SGR (i.e. persistent high or low growth rate) were selected for SAGE analysis.

### 2.2. RNA extraction

Total RNA was extracted from whole brain, liver and white muscle using a total RNA purification kit and a Maxwell 16 MDXi robot (Promega, Madrid, Spain) following the recommended procedure. The concentration, quantity and integrity of total RNA in 300 µL of DEPC treated water were evaluated using a NanoDrop 1000 Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) and agarose (0.8%) gel electrophoresis. Pools of total RNA (10 µg) of each tissue were generated from the 6 selected high and low SGR individuals and treated with DNase using a Turbo DNA-free kit (Ambion, Huntington, UK) prior to SuperSAGE library production.

### 2.3. SuperSAGE library preparation

SuperSAGE was performed using a modification of the method described in Matsumura et al. (Matsumura et al, 2010), the tag flanking adapter and primer sequences were adjusted to make the method compatible with the SOLiD4 sequencing system (Life Technologies, Carlsbad, USA) (Matsumura et al., 2012). Adapter-A was common to all libraries and was obtained by annealing to the isolated RNA the two oligonucleotides 5'-TTCCTCATTCTCTCAAGCAGAAGACGGCATAACGAAATG AT ACGGGCAGCCACCGACAGGTCTAACGATGTACGCAGCAGCATG-3' and 5'-CTGCTGCGTACATCGTTAGACCTGTGCGTGGTCGCCGTATCATTTCGT ATGCCGTCTTCTGCTTGAGAGAATGAGGAA-amino-3' (Metabion, Martinsried, Germany). Adapter-B was indexed with library specific 6bp barcodes (XXXXXX) in the two oligonucleotides 5'-CCACTACGCCTCC GCTTTCCTCTCTATGGGAGTCGGTGATXXXXX-3' and 3'-NNXXXXXA TCACCGACTGCCCATAGAGAGGAAAGCGGAGGCGTAGTG GTT-amino-5' (Metabion).

SuperSAGE libraries for brain, liver and white muscle of high (H) and low (L) SGR fish were produced from double stranded cDNA (dscDNA) synthesised using total RNA as the template, random hexamers and biotinylated oligo-dT (Metabion). This was followed by of dscDNA digestion with *Nla*III (New England Biolabs, Ipswich, MA, USA) and capture of the biotinylated fragments with streptavidin-coated magnetic beads. Adapter-A was ligated to the biotinylated dscDNA biotinylated fragments and EcoP15I (New England Biolabs) restriction digestion performed. Adapter-B was ligated to the produced adapter-A-dscDNA tags. The construct was then amplified with phusion high-fidelity polymerase (New England Biolabs) and purified using an 8% non-denaturing polyacrylamide gel (Matsumura et al., 2012).

### 2.4. SOLiD4 sequencing, editing and quality control

SuperSAGE library quality was assessed using an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA). Emulsion PCR, slide preparation and sequencing runs were performed following the manufacturer's protocol (Applied Biosystems SOLiD Library Preparation Protocol, Life Technologies) and using the sequencing primer supplied 5'-CCACTACGCCTCCGTTTCTCTCTATGGG CAGTCGGTGAT-3'.

Primary data analysis including editing the 50 bp raw colour space sequences, quality control and library assignment. The sequencer output colour space fasta. csfasta and. qual files were converted to an integrated colour space Phred+33 Sanger fastq file using *bfast* 0.6.4e program (Homer, Merriman, & Nelson, 2009) script "solid2fastq". The pooled sequences were assigned to libraries and subsequently the 6 base pair barcode was removed with the perl script "fastx\_barcode\_splitter.pl", and "fastx\_trimmer" script available in the FASTX-Toolkit ([http://hannonlab.cshl.edu/fastx\\_toolkit/](http://hannonlab.cshl.edu/fastx_toolkit/)). Summary statistics of the sequence files for quality control was obtained using *FastQC* (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Quality screening at all steps of sequence editing was performed using FASTX-Toolkit scripts "fastx\_quality\_stats", "fasta\_clipping\_histogram.pl" and "fastq\_quality\_boxplot\_graph.sh". A second positive quality control was based on the identification of the 3'-adaptor (adaptor A) sequence added during

SuperSAGE library construction. Sequences that had a modified or absent 3'-adaptor sequence were eliminated from the analysis. Adaptor sequences were trimmed using *fastx\_trimmer* in the FASTX-Toolkit scripts. Taking advantage of the double encoded nature of colour space sequences, the positive identification of Adaptor A at the end of the sequence was used for quality control instead of the application cut off standard quality threshold of 20, and resulted in a higher yield of extracted tags with high confidence.

### 2.5. SuperSAGE tags counts and mapping

Tag counts and genomic mapping was performed with the command line version of *solid.sage.command.v109.pl* program (He, Liqun, Life Technologies, personal communication). The European sea bass draft genome sequence (ENA submission PRJEB5099) was edited to RefSeq fasta format and used to map the tag sequence data. The generated output mapping file for each library consisted of a list of tags, their frequency and their genomic coordinates, that were all concatenated into a matrix file.

### 2.6. Statistical analysis of differentially expressed tags

Differentially expressed tags in liver, brain and muscle from fast and slow growing European sea bass were identified by pairwise frequency comparisons (Fisher's exact test), followed by a false discovery rate test (FDR, proportion of false positives when a particular test is called significant). This was performed using the tag count matrix retrieved from the tag extraction and count steps as the input file for an in house developed R script (R v2.8.1) (supplied on request) that tested for the null hypothesis of independence of specific tags and library counts using an in-built Fisher's exact test. The *P*-values obtained from the Fisher's exact test were automatically used for *q*-value estimation for FDR using the available "q value" package (<http://cran.r-project.org/web/packages/qvalue/index.html>) (Storey & Tibshirani, 2003). No significance level (*P*-value) or minimum tag count cut-off was defined in the Fisher's exact test for differential expression in pairwise comparisons. A minimum tag count of five for at least one library was used as the cut-off to select tags for expression analysis. The cut-off for significance was set after evaluation of FDR results at a *q*-value  $< 1 \times 10^{-5}$  (0.01% FDR) and  $\log_2$  threshold of 2 for comparisons between library tag counts derived from high and low SGR fish.

### 2.7. SuperSAGE tags functional annotation

Mapped tags were annotated using their positional coordinates against the European sea bass GFF3 genomic gene annotation file via the "Operate on Genomic Intervals" tool available in the Galaxy server (Goecks et al, 2010) (<https://main.g2.bx.psu.edu/>). Significant tags from each library were analysed using the STRING v. 9.0 protein interaction database (Szklarczyk et al, 2011) (<http://string-db.org>). The interactions included direct (physical) and indirect (functional) associations derived from three sources: high-throughput experiments, co-expression, and previous knowledge. The interaction score for differential tags was established at a high confidence level ( $> 0.4$ ). Biological process and molecular function Gene Ontology (GO) enrichment analysis were also carried out using STRING v.9.0.

### 2.8. Ingenuity pathway analysis

A dataset file containing the gene identifiers,  $\log_2(B/S)$  and *q*-values was the input file for Ingenuity Pathway Analysis (IPA) (Ingenuity System Inc, USA) [<http://www.ingenuity.com/>]. The IPA "Core Analysis" function was used to interpret the experiment results in the context of biological function, gene networks and canonical pathways. Both  $\log_2$  (ratio) and *q*-values were defined as value parameters for the analysis. Canonical pathways and biological functions derived from IPA analysis

were sorted by significance using a right-tailed Fisher's Exact test *P*-value. The networks generated were also ordered by significance, using the score obtained with a modified Fisher Exact test with associated top biological functions.

## 3. Results

### 3.1. Growth experiment

Specific growth rates averaged 0.8% and 1.3% (min 0.70% - max 1.38%) for L and H SGR groups, respectively, during the first period (first 9 months) and 0.9% and 1.3% (min 0.67% - max 1.31%) during the second period (70 days growth trial). SGR was more homogenous in the selected high SGR than in the low SGR fish (supplementary file 1, figure S1).

### 3.2. Editing and quality control of the sequencing output

A total of 61,003,920 raw colour space reads was obtained and 39,805,391 reads were assigned to specific SuperSAGE libraries (supplementary file 2, table S1). Reads assigned to brain and liver libraries were similar and were approximately 6 million and 4 million, respectively. For the muscle libraries, group L had ca. 6 million reads and group H ca. 1.3 million reads. After quality editing and tag extraction the reads assigned to specific tissue libraries were more evenly spread between 1.2 and 3 million (supplementary file 2, table S1).

### 3.3. Differential gene expression, annotation and GO enrichment

Comparison of tag frequency from the L and H groups yielded 344, 698 and 601 differentially expressed in brain, liver and muscle, respectively (Table 1). Approximately 80% of the tags were assigned to genomic loci on the draft genome of the European sea bass (Table S1). In total 47,071 tag loci corresponding to unique transcripts were mapped and gene annotation retrieved for 15,584 tag loci, which corresponded to 9,066 genes. The differentially expressed tags between high SGR (B) and low SGR (S) corresponded to 322 genes in brain, 683 in liver and 596 in muscle, respectively. The greater number of tags identified relative to genes was the result of some tags assigning to different loci of the same gene, that presumably represent alternative transcripts and/or possible regulatory RNAs (Table 1) (detailed annotation and expression data in Supplementary File 2).

The top 10 gene ontology (GO) terms for biological process and molecular function characteristic of each tissue are listed in Tables 2 and 3, respectively. Although these included well known specific functions for each tissue, analysis of differentially expressed tags failed to produce significant enrichment of any specific GO terms presumably due to the relatively low number in the significant tags population for the stringent fold change threshold used.

Interestingly, 8% of the differentially expressed annotated transcripts between H and L were located within previously identified growth related QTLs (Massault et al, 2010) (Table 4). Thirty-five genes located within previously identified growth related QTLs with an identified physiological role in growth (Table 5) and may be candidates to explain the divergent growth phenotypes. Although these differentially expressed transcripts passed the FDR they did not pass the very stringent expression cut-off threshold of  $\log_2$  (B/S, minimum of 4-fold change in

**Table 1**  
Significant tag counts and descriptive statistics of the annotation.

Frequency	Brain	Liver	Muscle
Significant tags	344	698	601
Significant tags annotated	109 (32%)	233 (33%)	91 (15%)
Genes with significant tags	322	683	596
Genes with significant tags annotated	87 (27%)	218 (32%)	86 (14%)

**Table 2**

Top 10 gene ontology (GO) biological process enrichment. Enriched GO terms for the brain, liver and muscle identified transcripts are presented as GO term number (GO\_id), number of genes (#Genes), and false discovery rate (FDR) values.

GO_id	Biological process Term	#Genes	FDR
<b>Brain</b>			
GO:0006796	Phosphate-containing compound metabolic process	65	6.67E-6
GO:0006793	Phosphorus metabolic process	65	6.67E-6
GO:0051234	Establishment of localization	148	2.78E-5
GO:0072599	Establishment of protein localization in endoplasmic reticulum	19	2.78E-5
GO:0045047	Protein targeting to ER	19	2.78E-5
GO:0006605	Protein targeting	34	2.78E-5
GO:0006810	Transport	144	5.24E-5
GO:0022415	Viral reproductive process	24	5.66E-5
GO:0006614	SRP-dependent cotranslational protein targeting to membrane	18	6.86E-5
GO:0007411	Axon guidance	35	6.86E-5
<b>Liver</b>			
GO:0006957	Complement activation, alternative pathway	5	1.96E-2
GO:0042221	Response to chemical stimulus	65	7.85E-2
GO:0051259	Protein oligomerization	15	7.85E-2
GO:0007599	Hemostasis	21	7.85E-2
GO:0006066	Alcohol metabolic process	20	7.85E-2
GO:0051262	Protein tetramerization	8	1.17E-1
GO:0007596	Blood coagulation	20	1.21E-1
GO:0006949	Syncytium formation	4	1.21E-1
GO:0006958	Complement activation, classical pathway	7	1.21E-1
GO:0051289	Protein homotetramerization	6	1.54E-1
<b>Muscle</b>			
GO:0030049	Muscle filament sliding	12	9.81E-11
GO:0006414	Translational elongation	17	9.81E-11
GO:0006413	Translational initiation	18	2.57E-10
GO:0019083	Viral transcription	15	1.43E-9
GO:0006415	Translational termination	15	1.47E-9
GO:0030048	Actin filament-based movement	12	2.45E-9
GO:0019058	Viral infectious cycle	15	3.33E-9
GO:0043624	Cellular protein complex disassembly	15	8.42E-9
GO:0006614	SRP-dependent cotranslational protein targeting to membrane	15	8.42E-9
GO:0043241	Protein complex disassembly	15	8.42E-9

expression between fish with high or low SGR). As examples, insulin-like growth factor binding protein 5 (IGFBP5) maps to a LG15 growth QTL in European sea bass and is differentially expressed in the H and L groups in brain, liver and muscle. IGFBP5 in the H group had a lower expression in the brain (−0.66-fold) relative to the liver (2-fold) and muscle (1.65-fold). Insulin-like growth factor 2 (IGF2) falls within a LG6 growth QTL and had a lower expression in the liver of H fish (−1.75-fold).

### 3.4. Analysis of string protein networks

The largest interacting networks inferred from the differentially expressed tags were ribosomal proteins (Fig. 1) which interact (bind) to generate the functional ribosome. For example, the ribosomal binding protein complex was the largest transcript cluster found in the liver and albeit a smaller cluster was also enriched in brain and muscle. Overall, with the exception of the ribosomal clusters relatively few significant groups of interacting genes (protein products) were identified in the differentially expressed genes. Some common networks were found between muscle and liver that were linked to RNA transcription, for example polymerase RNAII (POLR2A), serine/arginine rich splicing factor 1 (SRSF1), pre-mRNA processing factor 8 homolog (PRPF8). In muscle a small group of non-ribosomal interacting products were identified, namely myosin heavy chain 3 (MYH3), troponin T (TNNT), heat

**Table 3**

Top 10 gene ontology (GO) - molecular function enrichment. Enriched GO terms for the transcripts identified in the brain, liver and muscle are presented as GO term number (GO\_id), number of genes (#Genes), and false discovery rate (FDR) values.

GO id	Molecular function Term	#Genes	FDR
<b>Brain</b>			
GO:0005515	Protein binding	284	5.66E-3
GO:0036094	Small molecule binding	132	6.63E-3
GO:0000166	Nucleotide binding	124	6.63E-3
GO:0005516	Calmodulin binding	18	2.47E-2
GO:0015085	Calcium ion transmembrane transporter activity	5	4.1E-2
GO:0035639	Purine ribonucleoside triphosphate binding	95	4.63E-2
GO:0008092	Cytoskeletal protein binding	37	6.87E-2
GO:0046875	Ephrin receptor binding	6	6.87E-2
GO:0032555	Purine ribonucleotide binding	94	6.91E-2
GO:0032553	Ribonucleotide binding	94	6.91E-2
<b>Liver</b>			
GO:0031701	Angiotensin receptor binding	3	4.53E-1
GO:0005254	Chloride channel activity	6	4.53E-1
GO:0005102	Receptor binding	30	4.53E-1
GO:0032403	Protein complex binding	13	4.53E-1
GO:0005253	Anion channel activity	6	4.53E-1
GO:0008395	Steroid hydroxylase activity	3	5.13E-1
GO:0030246	Carbohydrate binding	13	5.13E-1
GO:0016709	Oxidoreductase activity, acting on paired donors.	4	5.13E-1
GO:0050997	Quaternary ammonium group binding	3	5.31E-1
GO:0005324	Long-chain fatty acid transporter activity	2	5.31E-1
<b>Muscle</b>			
GO:0008307	Structural constituent of muscle	13	2.57E-11
GO:0005198	Structural molecule activity	30	1.88E-8
GO:0003779	Actin binding	19	3.63E-5
GO:0008092	Cytoskeletal protein binding	23	2.48E-4
GO:0008137	NADH dehydrogenase (ubiquinone) activity	7	2.48E-4
GO:0050136	NADH dehydrogenase (quinone) activity	7	2.48E-4
GO:0016655	Oxidoreductase activity, acting on NADH	7	1.92E-3
GO:0048407	Platelet-derived growth factor binding	4	2.5E-3
GO:0005515	Protein binding	111	2.5E-3
GO:0003729	mRNA binding	7	3.07E-2

shock protein 90 (HSP90AA1), peptidyl-prolyl *cis-trans* isomerase (cyclophilin A; PPIA). In the brain a network of energy producing enzymes (ENO1, ALDOB, GAPDHS) was prominent.

### 3.5. Pathway analysis

Significant canonical pathways identified by IPA in the brain were largely neurone specific (e.g. synaptic and gonadotropin releasing hormone signalling), in the liver they included acute phase response and synaptic response, and in muscle calcium and androgen signalling (Fig. 2).

Comparison of the canonical pathways of differentially expressed genes between tissues suggested common tissue wide biological regulation of some processes occurred in the H and L fish. As an example CREB signalling in the neurons canonical pathway was among the top five most significantly represented canonical pathways and was commonly affected in all tissues analysed (Fig. 3).

## 4. Discussion

In the present study we hypothesised that there would be differentiation of transcript expression that underpin and are responsible for the trait of chronic fast and slow growth in non-selected sea bass. We further hypothesised that some of these transcripts may have a quantitative effect on growth traits. Cross-referencing the gene annotation of the differentially expressed genes in the fast and slow growing sea bass with



**Table 4**

Significantly different expressed genes with genomic position within the European sea bass LG4, LG6 and LG15 growth related QTL confidence intervals. The significance cut-off was set at a  $q$ -value  $< 1 \times 10^{-5}$  and  $\log_2$  ratio (H/L) threshold of 2 for pairwise comparisons of all tissues libraries.

Gene	Gene description	Tissue	Log2(H/L)	$q$ -value	Chr	Gene start	Gene end	Orient
ISLR2	Immunoglobulin superfamily L-rich repeat 2	Brain	2.9	1.1E-80	LG6	11934066	11936195	+
PDXK	Pyridoxal (pyridoxine – vitamin B6) kinase	Brain	-2.8	2.0E-07	LG15	8145231	8152657	+
RPL24	60S ribosomal protein L24	Brain	2.6	0.0E+00	LG15	16293533	16297163	-
FRYL	FRY-like	Liver	-3.6	2.4E-14	LG4	21320759	21361792	+
aldh9a1a	Aldehyde dehydrogenase 9 member A1A	Liver	4.3	2.5E-11	LG4	24572564	24574329	-
CRTC1	CREB regulated transcription coactivator 1	Liver	-4.6	6.4E-07	LG4	9361337	9374882	-
FETUB	Fetuin B	Liver	-4.4	3.9E-11	LG4	2784300	2789783	-
GPX4	Glutathione peroxidase 4	Liver	2.7	5.6E-81	LG4	19317875	19319599	+
CPA4	Carboxypeptidase A4	Liver	-5.4	3.4E-49	LG6	19837939	19844753	-
CAPRIN1	Cell cycle associated protein 1	Liver	-5.1	7.4E-10	LG6	13742751	13751644	-
MT	Metallothionein	Liver	4.9	1.9E-06	LG6	2014062	2014825	+
RPS13	Ribosomal protein S13	Liver	-6.3	0.0E+00	LG6	19159089	19162501	+
SLC45A3	Solute carrier family 45 member 3	Liver	-5.5	7.7E-13	LG6	2187105	2197339	+
SPIRE2	Spire homolog 2 (Drosophila)	Liver	4.7	1.8E-15	LG6	25656290	25692880	+
RORA	RAR-related orphan receptor A	Liver	-5.4	1.4E-12	LG6	11526244	11534924	+
HCE1	High choriolytic enzyme 1	liver	-2.7	4.8E-14	LG6	21266930	21271103	-
PPARA	Peroxisome proliferator-activated receptor $\alpha$	liver	3.2	8.1E-07	LG6	16681120	16697250	-
TPM1	Tropomyosin 1 (alpha)	liver	5.3	1.0E-08	LG6	11202329	11213230	-
myhz1.1	Myosin – heavy polypeptide 1 skeletal muscle	liver	-5.1	1.1E-09	LG6	22460588	22463113	+
PPIB	Peptidylprolyl isomerase B	liver	-2.6	2.1E-69	LG6	18436498	18440079	+
CXCR4	Chemokine (CXC motif) receptor 4	liver	4.9	7.7E-06	LG15	7892444	7893885	+
rpl8	60S ribosomal protein L8	liver	2.6	2.1E-38	LG15	19083529	19086264	-
HNRNPA3	Heterogeneous nuclear ribonucleoprotein A3	liver	-5.1	3.6E-09	LG15	13647136	13650264	+
EPB41L5	Erythrocyte membrane protein band 4.1 like 5	liver	-4.7	2.0E-07	LG15	6109897	6139853	+
Q8JIP9	Warm temperature acclimation related protein	liver	-4.9	0.0E+00	LG15	19754011	19758725	+
TRPM2	Transient receptor potential cation channel M2	liver	-2.7	7.8E-65	LG15	10987466	11005611	+
RPL24	60S ribosomal protein L24	liver	4.5	2.4E-298	LG15	16293533	16297163	-
SERBP1	SERPIN mRNA binding protein 1	muscle	-2.1	7.4E-41	LG4	19506224	19511352	+
TMEM38A	Transmembrane protein 38A	muscle	-3.3	3.2E-157	LG4	9330963	9339526	+
CCDC124	Coiled-coil domain containing 124	muscle	-3.2	3.5E-45	LG4	9570604	9572384	+
SPIRE2	Spire homolog 2 (Drosophila)	muscle	-4.3	5.9E-18	LG6	25656290	25692880	+
TPM1	Tropomyosin 1 (alpha)	muscle	-2.2	1.9E-266	LG6	11202329	11213230	-
NDUFA9	NADH dehydrogenase 1 $\alpha$ 9	muscle	-2.3	2.6E-39	LG6	3070462	3075461	+
ANO1	Anoctamin 1 $Ca^{2+}$ activated chloride channel	muscle	-2.6	3.6E-07	LG6	26811750	26854267	+
RPL24	60S ribosomal protein L24	muscle	-2.2	0.0E+00	LG15	16293533	16297163	-

**Table 5**

Candidate genes with genomic position within the European sea bass LG4, LG6 and LG15 growth related QTL confidence intervals. Significance cut-off was set at just for  $q$ -value  $< 1 \times 10^{-5}$  for all tissues libraries pairwise comparisons.

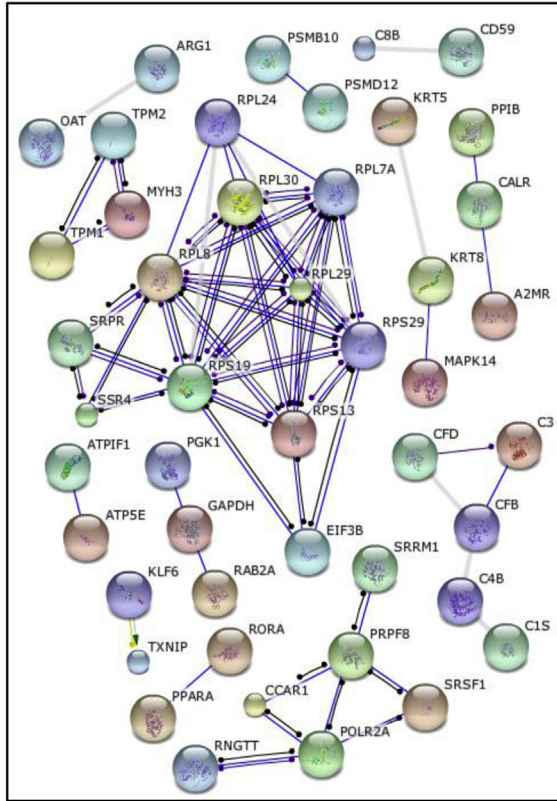
Gene	Gene description	Tissue	Log2(H/L)	$q$ -value	Chr	Gene start	Gene end	Orient
LEPR	Leptin receptor	brain	1.1	3.3E-11	LG4	6928664	6948623	+
TRHR	Similar thyrotropin-releasing hormone receptor	brain	0.4	6.1E-08	LG6	26373783	26391807	-
IGFBP5	Insulin-like growth factor binding protein 5	brain	-0.6	1.8E-05	LG15	3836279	3847001	+
INSR	Insulin receptor	liver	0.05	1.3E-05	LG4	14044023	14119027	-
IGF2	Insulin-like growth factor 2 (somatomedin A)	liver	-0.8	4.2E-25	LG6	5126328	5130860	+
IGFBP2	Insulin-like growth factor binding protein 2	liver	-0.8	2.0E-13	LG15	3856585	3874856	-
IGFBP5	Insulin-like growth factor binding protein 5	liver	1.0	2.3E-34	LG15	3836279	3847001	+
IGFBP3	Insulin-like growth factor binding protein 3	muscle	1.3	9.6E-11	LG4	1920612	1938180	+
IGFBP5	Insulin-like growth factor binding protein 5	muscle	0.7	3.4E-09	LG15	3836279	3847001	+

the genes populating previously identified growth QTL (Louro et al, 2016; Massault et al, 2010), it was possible to identify strong candidate genes for future studies. The biological role of those identified candidate genes listed in Table 4 and Table 5 is discussed further below.

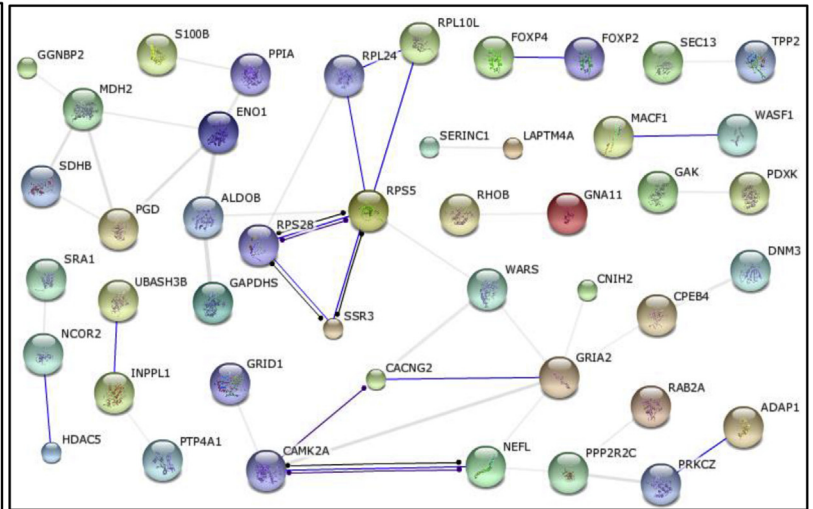
The SOLID4 SuperSAGE tag architecture required design and programming of a new editing and quality control pipeline integrating as a positive control the sequence of the forward and reverse tag, which increased the yield and confidence of tag extraction. In total it was possible to extract 10,732,697 SuperSAGE tags which represented 27% of the total raw colour space reads in the initial raw data. Relatively fewer differentially expressed transcripts were identified in the brain compared to the liver and white muscle. For differential tag/transcript expression a very stringent cut-off of 0.001% false discovery rate, and a  $\log_2$  ratio (B/S) threshold of 2 was chosen to adjust for individual variation resulting from sequencing of SuperSAGE libraries constructed with pools of RNA obtained from several individuals. At the time of the analysis, the incomplete nature of the gene annotation of the sea bass genome meant

that a relatively small proportion of all the tags generated were annotated (15%–33%). Nonetheless, comparison of the differentially expressed and annotated tags with the genes present in the previously identified growth QTL revealed considerable overlap. The most noticeable gene located within the identified growth related QTLs was IGF2, which has a known regulatory mutation that causes a major QTL effect on muscle growth in the pig (Van Laere et al, 2003). IGF2 is member of the insulin family of polypeptide growth factors, it is a key regulator of normal foetal development and growth (Cornish et al, 2007; Harris & Westwood, 2012) and related to intrauterine growth restriction diseases (Qiu et al, 2005). IGF2 is also known to be an imprinted gene with paternal inherited allele expression in both eutherians and marsupials, but reported bi-allelic expression in live bearing fishes (Lawton et al, 2005). In the current study although the expression variation did not pass the  $\log_2$  (H/L) expression threshold, it was more highly expressed in the L SGR group, and this reinforces the generally held idea that this hormone is less important in juvenile/adult growth regulation (Pierce et al, 2010a).

### Liver



### Brain



### Muscle

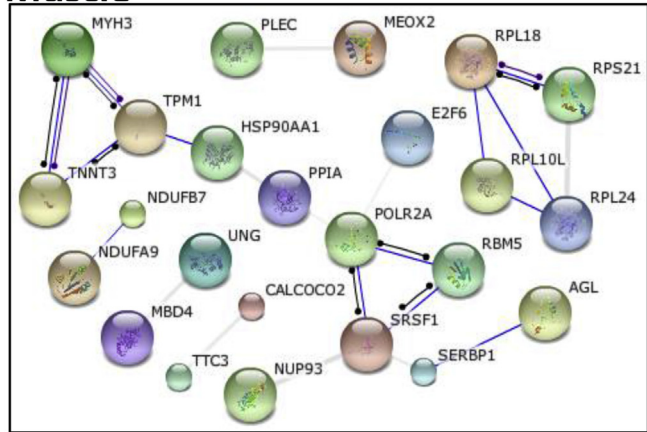


Fig. 1. STRING gene network inferences viewed in “actions view” in liver, brain and muscle.

There was no evidence that IGF1, the principal somatomedin acting as a growth factor in adult stages was differentially expressed between slow and fast growing fish (Picha et al., 2008). Surprisingly, hepatic IGF1 tags were of low abundance and gave a non-conclusive high FDR on comparison between the fast and slow growing fish. An explanation for the low abundance of IGF1 tags may be related to the relatively high abundance of insulin-like growth factor binding proteins (IGFBP) in all the tissues analysed. IGFBPs are also located within previously identified growth related QTLs in sea bass (Louro et al., 2016) namely, LG4, LG6 and LG15 and differential expression passed the confidence interval and achieved FDR significance. The IGFBP are multifunctional integrators which bind IGFs and in this way modulate their action and influence diverse physiological functions (Kelley et al., 2002). In the liver IGFBP 5 and 2 were identified and an SNP (single nucleotide polymorphism) in the latter binding protein has been associated with type 2 diabetes and triglyceride levels in humans (Harvard et al., 2007). In sea bass muscle IGFBP 5 and 3 were also present and in pigs an SNP in the former gene was associated with meat quality (Wang et al., 2010). Furthermore, the divergent expression of IGFBP5 in brain relative to liver and muscle in the present study, may indicate that QTLs can be tissue specific as a consequence of divergent gene regulatory processes such as gene methylation.

Pathway analysis (IPA) allowed the characterization and identification of the canonical pathways differentiated between L and H SGR groups. The CREB signalling in neurons canonical pathway was among the top five most significant pathways commonly affected in all three

tissue comparisons between slow and fast growing fish, suggesting a linked biological regulation. The cAMP/CREB signalling pathway has been strongly implicated in cell proliferation and survival, glucose homeostasis, spermatogenesis, circadian rhythms and synaptic plasticity that is associated with a variety of complex forms of memory including spatial and social learning (Mayr & Montminy, 2001; Shaywitz & Greenberg, 1999; Wen, Sakamoto, & Miller, 2010). It activates transcription of target genes in response to a diverse array of stimuli, including peptide hormones, growth factors, and neuronal activity, that activates a variety of protein kinases including protein kinase A (PKA), pp90 ribosomal S6 kinase (pp90RSK), and Ca<sup>2+</sup>/calmodulin-dependent protein kinases (CaMKs) (Shaywitz & Greenberg, 1999). In the comparisons between transcripts in the brain from fast growing and slow growing fish the abundance of several transcripts encoding proteins of the CREB signalling pathway were markedly modified. In particular, in the H group down regulation of the iGLUR and mGLUR group II/III was evident and the PIK3, cAMP and Ca<sup>2+</sup> signalling pathways appeared to also be down-regulated. The GLUR family are involved in activity dependent, long term synaptic strength and with the exception of GLUR1 their transcription was down regulated along with the down-stream transcripts for the proteins they signal through, PKA and cAMP (Barria et al., 1997). In contrast, transcripts encoding a number of elements of the Ras signalling pathway were up-regulated and included MEK, the extracellular kinase regulated (ERK) signalling pathway, which influences Elk a transcription factor involved in cell growth differentiation and survival (Besnard et al., 2011). Recently in an elegant series of

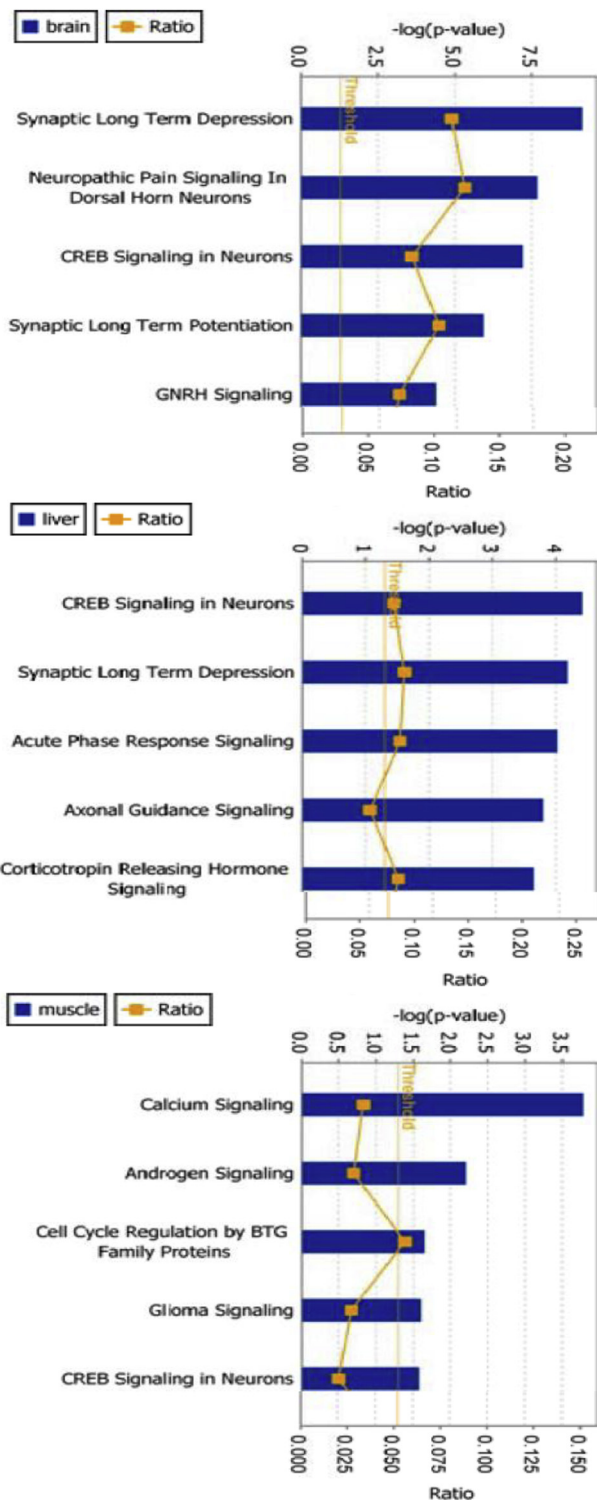


Fig. 2. Top five most significantly modified canonical pathways in all three tissue comparisons. The columns represent “-log (P-value)”, the orange points represent the ratio of the number of genes that meet the cut-off criteria, to the number of genes that make up the pathway.

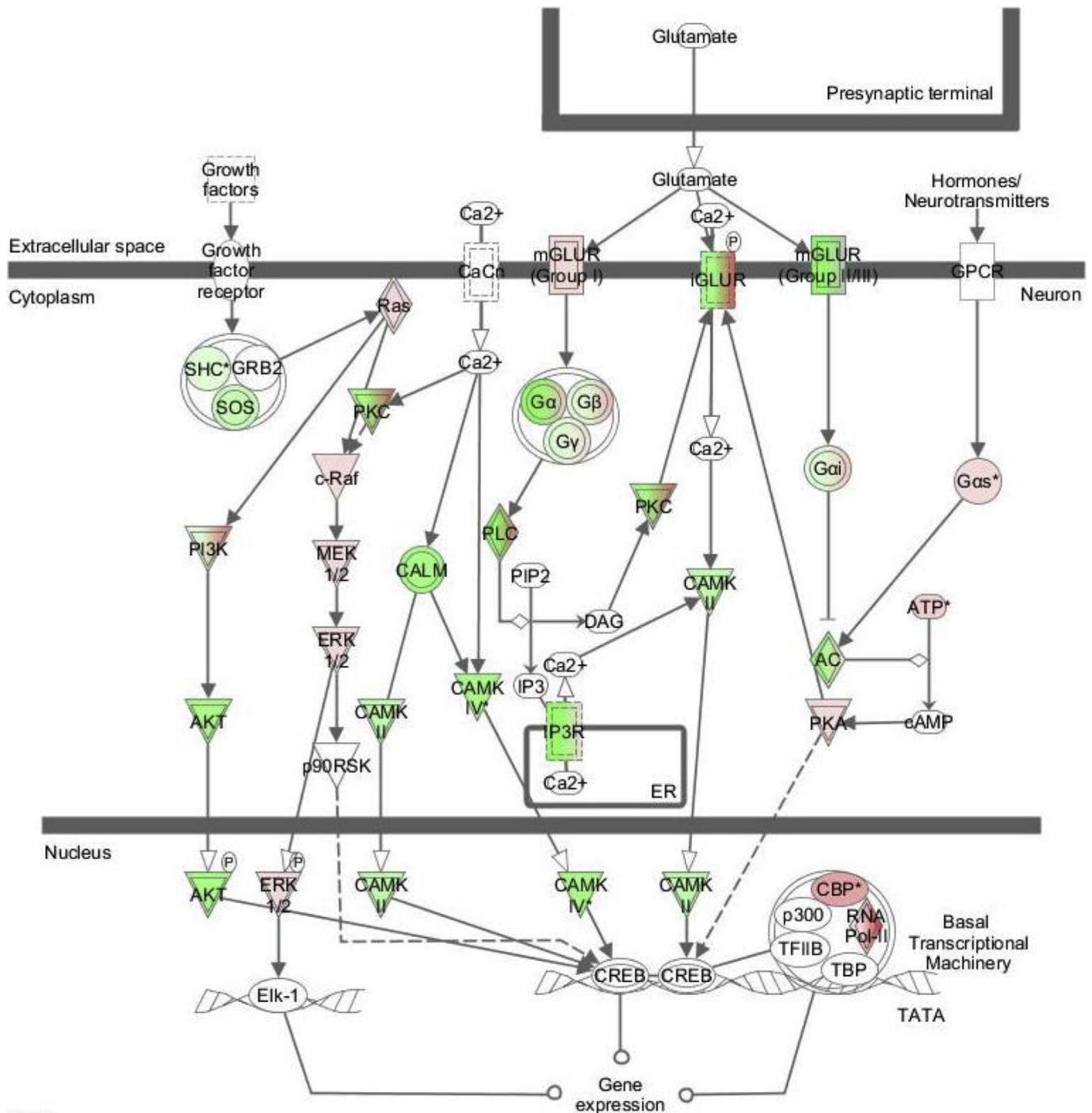
experiments ERK1/2 signalling was shown to be important for catch up growth in the zebrafish (Kamei et al, 2011) and the results of the present study revealing a significant modification in this pathway in fast growing sea bass raise the possibility that it might also be at the basis of divergent growth rates in older fish. It is apparent from this short appraisal that substantial modifications have occurred at the level of the brain between

fast and slow growing fish, although their functional significance remains to be established.

STRING analysis revealed a high representation of transcripts for ribosomal proteins and transcription factors in all tissues, which was perhaps unsurprisingly taking into consideration the abundance and constitutive character of transcription and translation. However, these transcripts are differentially expressed between fast and slow growing fish suggesting that there may be underlying differences in protein synthesis. Several endocrine pathways with anabolic actions were also differentially expressed between fast and slow growing sea bass and included elements and targets of the IGF and insulin axis. As in all vertebrates, many of the actions of GH in fish are mediated by insulin/insulin-like growth factor signalling (INS/IGF) [for review see Reincke, 2010]. The modification of insulin (INS) signalling revealed by the divergent expression between fast and slow growing fish may also be a factor explaining divergent growth. INS exerts its growth promoting action via similar mechanisms to IGF (Barbieri et al, 2003) and can also act by modifying IGF1 production by the liver when it synergizes with GH (Plisetskaya & Duan, 1994). The identification of differential expression of INS receptors in the liver suggested modifications in tissue responsiveness at this site between fast and slow growing sea bass. Within the INS/IGF axis our results interestingly highlight the divergence in IGF2 transcript abundance, which like IGF1 is an anabolic hormone (Pierce et al, 2010b), raising intriguing possibilities in relation to its cross-talk with INS receptors and its role in the sea bass with divergent growth in the present study. In the European seabass, the levels of both IGF1 and 2 mRNAs are regulated in the liver and muscle in response to fasting and refeeding protocols (Terova et al, 2007), suggesting they are involved in the regulation of energy balance in this species. The results presented herein further expand a candidate function of IGF2 in the regulation of growth in this species. In channel catfish (*Ictalurus punctatus*), IGF2 mRNA levels were greater in muscle (Peterson et al., 2004, 2008) and liver (Peterson et al., 2004) of fast growing fish compared to slow growing fish, contrary to our results with lower IGF2 expression levels in liver of H group. Such divergence might result from the regulatory balance of the myogenic effect of IGFs from other genes, namely the relationship between IGF and IGFBP and final effect on growth is fairly well established in vertebrates (Caruso & Sheridan, 2011; Fuentes et al, 2011; Kamei et al, 2011; Svanberg et al, 1996). In fish several studies with fasting and posterior refeeding challenges in Atlantic salmon (Bower et al, 2008) and fine flounder (Safian et al, 2012) suggest IGFBP4 and IGFBP5 as positive regulators of IGF role in compensatory growth, evidencing the weight of IGBPs effect upstream the IGF system and not solely pinpointing the IGF1 major role effect in the final muscle trait. The importance of the relationship between IGFs signalling, with hormone receptors and IGFBPs in the different tissues is also highlighted in the presented candidate genes listed in Table 5, as a result of the integrative approach of gene expression, pathway analysis and cross-mapping genes with previous growth related QTLs (Louro et al, 2016). In barramundi (*Lates calcarifer*), the IGF2 gene was also identified within a quantitative trait locus for body weight and length in LG10 (Wang et al, 2008).

In our study higher expression of cAMP/CREB path via G-protein Gn-alpha (Gαs) and protein kinase-A (PKA) occurred in the fish with a higher SGR and this pathway was coupled to G-protein receptors (GPCR) of several hormones and neurotransmitters. In contrast, Ca<sup>2+</sup>/calmodulin-dependent protein kinase (CaMKs) the alternative signalling pathway for GPCRs via Ca<sup>2+</sup>/calmodulin was reduced. Specifically, in muscle this Ca<sup>2+</sup> GPCR signalling pathway was up-regulated and it is part of a signalling cascade with an important role in cellular growth and development. Transcript encoding proteins involved in CREB regulation varied between tissues, indicating as expected tissue specific regulation. The results regarding tissue specificity was the complementation of tissue specific pathways that when interlinked via hormonal stimuli create a higher physiological response such as sexual differentiation. GNRH, corticotropin releasing hormone (CRH) and androgen signalling pathways were among the top five enriched pathways in brain, liver and





**Fig. 3.** The CREB signalling pathway, highlighting the differentially expressed transcripts between low (L) and high (H) SGR sea bass when differential transcripts from all tissue analysed were pooled. Red shading indicates higher expression values in (H) fast growing fish, green shading indicates higher expression (L) in slow growing fish and white is not significantly different. The intensity of the colour denotes the intensity of the response as assessed by transcript abundance. Expression results are shown for the sum/overlap of differential transcripts represented in all three tissues. Note transcripts of the G-protein, Ca<sup>2+</sup> and PI3K intracellular signalling pathways appear to be down-regulated. In contrast the ERK signalling pathway was up-regulated through GLUR (group I) and presumably acts via Elk-1 (no information) to bring about its effect. Notably in the top five canonical pathways, in addition to the CREB pathway, a number of complimentary pathways involved in sexual differentiation were observed. The pathways identified included gonadotropin releasing hormone (GNRH) in the brain, corticotropin releasing hormone (CRH) in the liver and androgen signalling pathways in the muscle (Fig. S2). The results indicate slow or fast growth was the outcome of the integrated tissue specific responses that interlink to create a whole body physiological response such as sexually dimorphic growth.

muscle, respectively. The CRH network was most populated with transcripts and was down-regulated in the fast growing fish. The CRH signalling pathway is principally recognised for its role in stress (Brewer et al, 2003), although it has also been implicated in reproduction (Leatherland, Li, & Barkataki, 2010). The down regulation of the CRH

receptor and elements of this signalling pathway in the liver of large fish was intriguing and suggested possibly a lower stress response. It will be important to assess plasma indicators of the stress axis to determine if the smaller fish have a more active stress-axis and if this contributed to their lower growth rate. Alternatively, and taking together the three principal



pathways that intersect the three tissues, which were linked with reproduction they may suggest growth divergence as a consequence of sex. The fish used in the experiment had probably already undergone sexual differentiation, but it was not possible to visually check sexual gender during the sampling due to the small size of the fish. As such, most gonads were sampled for future histological studies which might confirm this gender bias to explore the possibility that the bigger fish might be females.

## 5. Conclusion

The present study has demonstrated the utility of high throughput methods for revealing transcriptional changes that underlie divergent growth phenotypes in the sea bass. The three tissues tested, the liver, muscle and brain suffered different changes suggesting tissue specific programs operate, although intersecting pathways common to the three tissues were identified suggesting an integrated response. The results suggested that changes in endocrine regulatory pathways involving the IGF, insulin and stress axis may be at the basis of the divergent growth observed in our study. By merging transcriptomics with QTL analysis it has been possible to identify a suite of candidate genes which may explain the divergent growth phenotypes. The SAGE method deployed coupled to the SOLiD4 sequencing platform permitted identification of SNPs, which should contribute to future studies aimed at identifying the causative mutations explaining QTL. Phenotype characterization needs to be extended to other markers, such as plasma biochemical parameters and to test experimentally the link between pathway/transcript changes and physiology.

## Author's contributions

BL planned the work, carried out the fish growth experiment, all laboratorial practical work, all bioinformatics analysis and wrote the manuscript. RSTM and PISP assisted in the SAGE libraries construction, RR carried out the sequencing production and provided genome data, DJ de K provided funds and participated in planning the bioinformatics analysis. AVMC and DMP provided funds, planned the work, contributed to data analysis and wrote the manuscript.

## Acknowledgements

The authors acknowledge funding by the European Commission of the European Union through the Network of Excellence Marine Genomics Europe (contract GOCE-CT-2004-505403) and by the Portuguese Foundation for Science and Technology (FCT) through project CMAR/Multi/04326/2013 and grants to BL (SFRH/BD/29171/2006), PISP (SFRH/BPD/25247/2005) and RSTM (SFRH/BPD/66742/2009). BL benefited from a SABRETRAIN Marie Curie EST fellowship. The authors thank Nádia Silva, Rita Costa, Pedro Guerreiro and Ângela Ramos for experimental sampling and Dr Hideo Matsumura for advices on the SuperSAGE – SOLiD adaptors and protocol.

## Appendix A. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.aaf.2018.03.001>.

## References

- Barbieri, M., et al. (2003). Insulin/IGF-I-signaling pathway: An evolutionarily conserved mechanism of longevity from yeast to humans. *American Journal of Physiology - Endocrinology And Metabolism*, 285(5), E1064–E1071.
- Barria, A., et al. (1997). Regulatory phosphorylation of ampa-type glutamate receptors by cam-kii during long-term potentiation. *Science*, 276(5321), 2042–2045.
- Besnard, a., et al. (2011). Elk-1 a transcription factor with multiple facets in the brain. *Frontiers in Neuroscience*, 5.
- Blum, Y., et al. (2010). A factor model to analyze heterogeneity in gene expression. *BMC Bioinformatics*, 11.
- Bower, N. I., et al. (2008). Switching to fast growth: The insulin-like growth factor (IGF) system in skeletal muscle of atlantic salmon. *Journal of Experimental Biology*, 211(Pt 24), 3859–3870.
- Brewer, J. A., et al. (2003). Dissecting adrenal and behavioral responses to stress by targeted gene inactivation in mice. *Stress: The International Journal on the Biology of Stress*, 6(2), 121–125.
- Canario, A. V. M., et al. (2008). Genomics toolbox for farmed fish. *Reviews in Fisheries Science*, 16(1 supp 1), 3–15.
- Caruso, M. A., & Sheridan, M. A. (2011). New insights into the signaling system and function of insulin in fish. *General and Comparative Endocrinology*, 173(2), 227–247.
- Chatziplis, D., et al. (2007). Mapping quantitative trait loci in European sea bass (*Dicentrarchus labrax*): The BASSMAP pilot study. *Aquaculture*, 272(Supplement 1), S172–S182.
- Chistiakov, D. A., et al. (2005). A microsatellite linkage map of the European sea bass *Dicentrarchus labrax* L. *Genetics*, 170(4), 1821–1826.
- Chistiakov, D. A., et al. (2008). A combined AFLP and microsatellite linkage map and pilot comparative genomic analysis of European sea bass *Dicentrarchus labrax* L. *Animal Genetics*, 39(6), 623–634.
- Cornish, J., et al. (2007). Preptin, another peptide product of the pancreatic  $\beta$ -cell, is osteogenic in vitro and in vivo. *American Journal of Physiology - Endocrinology And Metabolism*, 292(1), E117–E122.
- FAO. (2012). *The state of the world fisheries and aquaculture*. Rome: Food and Agriculture Organization of the United Nations.
- Ferraresso, S., et al. (2010). Development of an oligo DNA microarray for the European sea bass and its application to expression profiling of jaw deformity. *BMC Genomics*, 11, 354.
- Fuentes, E. N., et al. (2011). IGF-I/PI3K/Akt and IGF-I/MAPK/ERK pathways in vivo in skeletal muscle are regulated by nutrition and contribute to somatic growth in the fine flounder. *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology*, 300(6), R1532–R1542.
- Gjedrem, T., & Baranski, M. (2009). *Selective breeding in aquaculture: An introduction. Reviews: Methods and technologies in fish biology and fisheries*. Dordrecht ; New York: Springer.
- Gjoen, H. M., & Bentsen, H. B. (1997). Past, present, and future of genetic improvement in salmon aquaculture. *ICES Journal of Marine Science: Journal du Conseil*, 54(6), 1009–1014.
- Goecks, J., et al. (2010). Galaxy: A comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biology*, 11(8), R86.
- Guyon, R., et al. (2010). A radiation hybrid map of the European sea bass (*Dicentrarchus labrax*) based on 1581 markers: Synteny analysis with model fish genomes. *Genomics*, 96(4), 228–238.
- Harris, L. K., & Westwood, M. (2012). Biology and significance of signalling pathways activated by IGF-II. *Growth Factors*, 30(1), 1–12.
- Harvard, D. G. I. o. B. I. o., et al. (2007). Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science*, 316(5829), 1331–1336.
- Homer, N., Merriman, B., & Nelson, S. F. (2009). BFAST: An alignment tool for large scale genome resequencing. *PLoS One*, 4(11), e7767.
- Hubner, N., et al. (2005). Integrated transcriptional profiling and linkage analysis for identification of genes underlying disease. *Nature Genetics*, 37(3), 243–253.
- Kamei, H., et al. (2011). Role of IGF signaling in catch-up growth and accelerated temporal development in zebrafish embryos in response to oxygen availability. *Development*, 138(4), 777–786.
- Kause, A., et al. (2005). Genetic trends in growth, sexual maturity and skeletal deformations, and rate of inbreeding in a breeding programme for rainbow trout (*Oncorhynchus mykiss*). *Aquaculture*, 247(1–4), 177–187.
- Kelley, K., et al. (2002). Comparative endocrinology of the insulin-like growth factor-binding protein. *Journal of Endocrinology*, 175(1), 3–18.
- Lawton, B. R., et al. (2005). Allelic expression of IGF2 in live-bearing, matrotrophic fishes. *Development Genes and Evolution*, 215(4), 207–212.
- Leatherland, J. F., Li, M., & Barkataki, S. (2010). Stressors, glucocorticoids and ovarian function in teleosts. *Journal of Fish Biology*, 76(1), 86.
- Louro, B., et al. (2010). Gilthead sea bream (*Sparus auratus*) and European sea bass (*Dicentrarchus labrax*) expressed sequence tags: Characterization, tissue-specific expression and gene markers. *Mar Genomics*, 3(3–4), 179–191.
- Louro, B., et al. (2016). Characterization and refinement of growth related quantitative trait loci in European sea bass (*Dicentrarchus labrax*) using a comparative approach. *Aquaculture*, 455, 8–21.
- Mackay, T. F., Stone, E. A., & Ayroles, J. F. (2009). The genetics of quantitative traits: Challenges and prospects. *Nature Reviews Genetics*, 10(8), 565–577.
- Massault, C., et al. (2010). QTL for body weight, morphometric traits and stress response in European sea bass *Dicentrarchus labrax*. *Animal Genetics*, 41(4), 337–345.
- Matsumura, H., et al. (2010). High-throughput SuperSAGE for digital gene expression analysis of multiple samples using next generation sequencing. *PLoS One*, 5(8), e12010.
- Matsumura, H., et al. (2012). SuperSAGE: Powerful serial analysis of gene expression. In H. Jin, & W. Gassmann (Eds.), *RNA abundance Analysis: Methods and protocols* (pp. 1–17). Totowa, NJ: Humana Press.
- Mayr, B., & Montminy, M. (2001). Transcriptional regulation by the phosphorylation-dependent factor CREB. *Nature Reviews Molecular Cell Biology*, 2(8), 599–609.
- Peterson, B. C., et al. (2008). Endocrine responses of fast- and slow-growing families of channel catfish. *North American Journal of Aquaculture*, 70(2), 240–250.
- Peterson, B. C., Waldbieser, G. C., & Bilodeau, L. (2004). IGF-I and IGF-II mRNA expression in slow and fast growing families of USDA 103 channel catfish (*Ictalurus punctatus*). *Comparative Biochemistry and Physiology a-Molecular & Integrative Physiology*, 139(3), 317–323.

- Picha, M. E., et al. (2008). *Endocrine biomarkers of growth and applications to aquaculture: A minireview of growth hormone, insulin-like growth factor (IGF)-I, and igf-binding proteins as potential growth indicators in fish* (vol. 70, p. 16). Bethesda, MD, ETATS-UNIS: American Fisheries Society. <https://www.tandfonline.com/doi/abs/10.1577/A07-038.1>.
- Pierce, A. L., et al. (2010). Metabolic hormones regulate basal and growth hormone-dependent igf2 mRNA level in primary cultured coho salmon hepatocytes: Effects of insulin, glucagon, dexamethasone, and triiodothyronine. *Journal of Endocrinology*, 204(3), 331–339.
- Pierce, A., et al. (2010). Metabolic hormones regulate basal and growth hormone-dependent igf2 mRNA level in primary cultured coho salmon hepatocytes: Effects of insulin, glucagon, dexamethasone, and triiodothyronine. 204(3), 331–339.
- Plisetskaya, E. M., & Duan, C. (1994). Insulin and insulin-like growth factor I in coho salmon *Oncorhynchus kisutch* injected with streptozotocin. *American Journal of Physiology*, 267(5 Pt 2), R1408–R1412.
- Qiu, Q., et al. (2005). Role of pro-IGF-II processing by proprotein convertase 4 in human placental development. *Proceedings of the National Academy of Sciences of the United States of America*, 102(31), 11047–11052.
- Reinecke, M. (2010). Influences of the environment on the endocrine and paracrine fish growth hormone-insulin-like growth factor-I system. *Journal of Fish Biology*, 76(6), 1233–1254.
- Rothschild, M. F., Hu, Z. L., & Jiang, Z. (2007). Advances in QTL mapping in pigs. *International Journal of Biological Sciences*, 3(3), 192–197.
- Rothschild, M. F., & Plastow, G. S. (2008). Impact of genomics on animal agriculture and opportunities for animal health. *Trends Biotechnol*, 26(1), 21–25.
- Safian, D., et al. (2012). Dynamic transcriptional regulation of autocrine/paracrine igfbp1, 2, 3, 4, 5, and 6 in the skeletal muscle of the fine flounder during different nutritional statuses. *Journal of Endocrinology*, 214(1), 95–108.
- Schadt, E. E., Monks, S. A., & Friend, S. H. (2003). A new paradigm for drug discovery: Integrating clinical, genetic, genomic and molecular phenotype data to identify drug targets. *Biochemical Society Transactions*, 31(2), 437–443.
- Shaywitz, A. J., & Greenberg, M. E. (1999). Creb: A stimulus-induced transcription factor Activated by A Diverse array of extracellular signals. *Annual Review of Biochemistry*, 68(1), 821–861.
- Storey, J. D., & Tibshirani, R. (2003). Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences*, 100(16), 9440–9445.
- Svanberg, E., et al. (1996). Role of insulin and IGF-I in activation of muscle protein synthesis after oral feeding. *American Journal of Physiology - Endocrinology And Metabolism*, 270(4), E614–E620.
- Szklarczyk, D., et al. (2011). The STRING database in 2011: Functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Research*, 39(suppl 1), D561–D568.
- Terova, G., et al. (2007). Cloning and expression analysis of insulin-like growth factor I and II in liver and muscle of sea bass (*Dicentrarchus labrax*, L.) during long-term fasting and refeeding. *Journal of Fish Biology*, 70, 219–233.
- Tine, M., et al. (2014). European sea bass genome and its variation provide insights into adaptation to euryhalinity and speciation. *Nature Communications*, 5.
- Van Laere, A.-S., et al. (2003). A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. *Nature*, 425(6960), 832–836.
- Wang, C. M., et al. (2008). Identification and verification of QTL associated with growth traits in two genetic backgrounds of Barramundi (*Lates calcarifer*). *Animal Genetics*, 39(1), 34–39.
- Wang, W., et al. (2010). Association of porcine IGF binding Protein-5 gene with meat quality. *Biochemical Genetics*, 48(3), 257–265.
- Wayne, M. L., & McIntyre, L. M. (2002). Combining mapping and arraying: An approach to candidate gene identification. *Proc Natl Acad Sci U S A*, 99(23), 14903–14906.
- Wen, A. Y., Sakamoto, K. M., & Miller, L. S. (2010). The role of the transcription factor CREB in immune function. *The Journal of Immunology*, 185(11), 6413–6419.
- Whitaker, H. A., McAndrew, B. J., & Taggart, J. B. (2006). Construction and characterization of a BAC library for the European sea bass *Dicentrarchus labrax*. *Animal Genetics*, 37(5), 526.