Characterization and refinement of growth related quantitative trait loci in European sea bass (*Dicentrarchus labrax*) using a comparative approach

Bruno Louro[1,2], Heiner Kuhl[3], Mbaye Tine[4, a], Dirk-Jan de Koning[2, b], Costas Batargias[5], Filip A.M. Volckaert[6], Richard Reinhardt[4], Adelino V.M. Canario[1] and Deborah M. Power[1]

[1] Centre for Marine Sciences, CCMAR-CIMAR Laboratório Associado, Universidade do Algarve, Campus do Gambelas, 8005-139 Faro, Portugal; [2] The Roslin Institute and R(D)SVS, University of Edinburgh, The Roslin Building, Easter Bush, Midlothian, EH25 9RG, United Kingdom; [3] MPI Molecular Genetics, Ihnestrasse 63-73, D-14195 Berlin-Dahlem, Germany; [4] Max-Planck-Institut für Pflanzenzüchtungsforschung Carl-von-Linné-Weg 10, D-50829 Köln; [5] Laboratory of Applied Fish Genetics and Fish Breeding, Dept. of Aquaculture & Fisheries Management, Technological Educational Institute of Messolonghi, Nea Ktiria, 30200 Messolonghi, Greece; [6] Laboratory of Biodiversity and Evolutionary Genomics, University of Leuven, Ch. Deberiotstraat 32, B-3000 Leuven, Belgium

[a] Current address: The Molecular Zoology Laboratory, University of Johannesburg, Kingsway Campus, Auckland Park 2006, South Africa.

[b] Current address: Swedish University of Agricultural Sciences, Department of Animal Breeding and Genetics, 750 07 Uppsala, Sweden

Abbreviated title: QTL refinement

Corresponding author: Bruno Louro, Centre for Marine Sciences, CCMAR-CIMAR Laboratório Associado, Universidade do Algarve, Campus do Gambelas, 8005-139 Faro, Portugal, e-mail: blouro@ualg.pt

Abstract

The identification of genetic markers for traits of interest for aquaculture, such as growth, is an important step for the establishment of breeding programmes. As more genomic information becomes available the possibility of applying comparative genomics to identify and refine quantitative trait locus (QTLs) and potentially identify candidate genes responsible for the QTL effect may accelerate genetic improvement in established and new aquaculture species. Here we report such an approach on growth related traits in the European sea bass (*Dicentrarchus labrax*), an important species for European aquaculture. A genetic map was generated with markers targeted to previously identified QTL for growth which reduced distance and improved resolution in these regions. A total of 36 significant QTL were identified when morphometric traits were considered individually in maternal half sibs, paternal half sibs and sib pair analysis. Twenty seven new markers targeted to the growth QTLs, obtained by comparative mapping, reduced the average distance between markers from 23.4, 9.1, and 5.8 cM in the previous map to 3.4, 2.2, and 5.2 cM, on linkage group (LG) LG4, LG6 and LG15 respectively. Lists of genes embedded in the QTL - 591 genes in LG4, 234 genes in LG6 and 450 genes in LG15 - were obtained from the European sea bass genome. Comparative mapping revealed conserved gene synteny across teleost fishes. Functional protein association network analysis with the gene products of the 3 linkage groups revealed a large global association network including 42 gene products. Strikingly the association network was populated with genes of known biological importance for growth and body weight in terrestrial farm animals, such as elements of the signalling pathways for Jak-STAT, MAPK, adipocytokine and insulin, growth hormone, IGFI and II. This study demonstrates the feasibility of a comparative genomics

combined with functional gene annotation to refine the resolution of QTL and the establishment of hypothesis to accelerate discovery of putative responsible genes.

Keywords: comparative genomics; convergent mapping; fish; growth QTL; linkage map; marker assisted selection

1. Introduction

The impact of genomics on fish aquaculture remains limited probably as a consequence of the relatively recent adoption of intensive aquaculture, the diversity of species cultured and non-uniform nature of the industry (Power, et al., 2011). Nonetheless, successful breeding programs exist with the most advanced being for salmonids, such as Atlantic salmon (*Salmo salar*), and it has contributed to an increased production of global significance, with large and sustained genetic gain in economically important traits, such as increased resistance to infectious pancreatic necrosis trough marker assisted selection (MAS) implementation (Fjalestad, et al., 2003; Houston, et al., 2008). Salmon breeding companies have now the resources to start implementing genomic selection (GS), a high density SNP chip (Houston, et al., 2014) to initially genotype the reference population and confirm the expected potential of genetic gain shown in the simulation studies (Ødegård, et al., 2014)

The European sea bass (*Dicentrarchus labrax* L.; Dicentrarchidae), heretoforth sea bass, is a gonochoristic marine teleost fish, distributed in temperate European coastal areas of the Atlantic Ocean and Mediterranean Sea. Its intensive exploitation as an aquaculture species is relatively recent and production is concentrated predominantly in the Mediterranean basin. Total production was 112,183 tonnes in 2009, with a market value of 500 million Euros (FAO, 2010). Current developments of selective breeding of European sea bass are being performed in several private companies and respective results are confidential. On the other hand, ongoing Aquaexcel2020 European project with results open to the scientific and producer communities, aim to implement and characterize new breeding lines for a stable improvement of production traits.

Several growth related QTL were identified in two previous studies which analysed sea bass body weight, morphometric and stress response traits (Chatziplis, et al., 2007; Massault, et al., 2010). Heritability values for morphometric growth related traits in sea bass are high (0.52-0.68) (Saillant, et al., 2006; Volckaert, et al., 2012), suggesting that there is considerable potential for increased growth rates if such traits can be selected during breeding. High heritability also suggests that genetic

gain can be obtained by mass selection (Vandeputte, et al., 2009), contrary to others traits of interest which require to sacrifice the fish for phenotyping, such as fillet quality or disease resistance. Future MAS or GS will aim at those more appropriate traits with lower heritability or requirement to sacrifice the fish, but should also have a multiple trait selection approach. The selection index should include those traits and growth related traits to ensure positive correlations of genetic gain, since growth is still the most desired trait in the producer point of view. In vertebrates the major regulatory factors of growth are conserved in the growth hormone/insulin-like growth factor-I (GH/IGF-I) axis. The importance of these factors in growth physiology is evident from observations of vertebrates carrying mutations or that are transgenic (Du, et al., 1992; Hull, Harvey, 1999).

A recent approach in mammals to decipher QTL is to construct concordant QTL maps since orthologous genes are expected to have conserved function in biological and biochemical traits. Thus, those genes with a quantitative effect in one species may also be important in another species (Jiang, et al., 2007; Kim, et al., 2004; Rothschild, et al., 2007). The construction of concordant maps for teleosts may prove to be a useful approach in aquaculture where many different species of fish are cultivated but relatively few have extensive molecular resources. However, such an approach is likely to be more challenging than in terrestrial animals because of their richer evolutionary diversity (Canario, et al., 2008). Additional complexity in map construction may arise from the teleost specific whole genome duplication (Jaillon, et al., 2004) and selective gene loss and divergence of duplicate gene function which presumably may modify selective pressures (Brunet, et al., 2006), family specific partial chromosome duplications/rearrangements (Jaillon, et al., 2004), and differing evolutionary pressures on the genome arising from adaptation to a specific ecological niche (Hoegg, et al., 2004). Nonetheless there is evidence that the construction of concordant QTL maps and comparison of growth QTL genome regions at the taxonomic level of Family or Order in teleost fish may be feasible, as exemplified for the sea bass and gilthead sea bream (*Sparus auratus*) by Sarropoulou et al (2008).

Several recent studies have used a comparative approach to better characterize identified QTL in aquaculture species. Similar to mammals they revealed the existence of concordant chromosomal regions. For example, co-segregating QTL for body growth and sex determination in the gilthead sea bream were established using synteny between the genome of sea bass and Asian sea bass (*Lates calcarifer*) (Loukovitis, et al., 2011). A comparative approach between growth QTL identified in rainbow trout (*Oncorhynchus mykiss*), Arctic char (*Salvelinus alpinus*) and Atlantic salmon revealed several homologous linkage groups (Moghadam, et al., 2007; Reid, et al., 2005; Wringe, et al., 2010). Moreover, the comparison of a genomic region from the spotted green pufferfish (*Tetraodon nigroviridis*) microsyntenic with a growth related QTL from turbot (*Scophthalmus maximus*) revealed a number of genes with ontology classification (GO) for growth regulation and cell proliferation (Sanchez-Molano, et al., 2011).

In addition to the identified QTL, several genetic and genomic resources have also been developed in European sea bass. They include a medium dense linkage map with 190 microsatellites, 176 AFLP (amplified fragment length polymorphisms) and two single nucleotide polymorphisms (SNP) (Chistiakov, et al., 2008) to which a further 35 SNP were added (Souche, et al., 2012), a 30,000 plus EST collection representing almost 18.000 unigenes (Louro, et al., 2010), a comparative BAC (bacterial artificial chromosome) map (Kuhl, et al., 2010), a radiation hybrid (RH) map (Guyon, et al., 2010), and recently the fully assembled genome (Tine, et al., 2014). The availability and status of current molecular tools and genomic data for European sea bass gives an opportunity to apply genomics to assist selective breeding.

Taking into consideration the conservation of growth regulatory mechanisms between vertebrates and the synteny among related teleosts, here we take a comparative approach to narrow down the list of candidate genes located in growth QTL regions of the sea bass genome. Comparative genomics was used to link sea bass QTL and linkage map markers to the three-spined stickleback (*Gasterosteus aculeatus*) and green spotted puffer (*Tetraodon nigroviridis*) assembled genomes. To

increase the resolution of the sea bass QTL map and identify putative markers targeted to QTL regions, candidate genome regions were functionally annotated. A complimentary approach was to conduct pathway analysis using genes from sea bass QTL regions to assess whether these regions were enriched with genes of functional relevance for growth. These approaches were taken to fulfill the two main objectives of this study; first to increase the resolution of previously identified sea bass growth related QTLs and secondly to obtain a short-list of candidate genes.

2. Materials and Methods

2.1. Growth related genes bibliographic search

A bibliographic search was performed using the bibliosphere module of the Genomatix software suite (Genomatix Software GmbH, Munich, Germany). The human growth hormone (GH1) was used as keystone gene in the initial analysis of the bibliographic search of growth related genes. All co-citations were checked for functional relation in the hand-curated knowledgebase of Protein-Protein NetPro™ (http://www.interactions.net.in). Gene sequences were retrieved from the NCBI GenBank database, and those genes from species phylogenetically closer to sea bass were analysed further.

2.2. Comparative genomic maps of QTL regions

All genetic marker sequences from the sea bass linkage map and candidate gene protein sequences identified previously were queried using, respectively, BLASTN and TBLASTN algorithms (Altschul, et al., 1990), against the *G. aculeatus* and *T. nigroviridis* genome masked databases in Ensembl BLASTVIEW (http://www.ensembl.org/Multi/blastview). The BLASTN query was performed with distant homology parameters (-E:10 –B:100 –filter:dust –RepeatMasker –W:9 –M:1 –N:-1 –Q:2 –R:1)) for high sensitivity. The cut-off for positive hits was taken at E-value of $5e^{-3}$ for

at least one hit per chromosome. TBLASTN was used with near-exact matches parameters (-E:10^{-5} -:100 –filter:seg –W:4 –hitdist:40 –matrix:BLOSUM80 –Q:3 –R:3 –T:16) for low sensitivity. The annotated genomic region in both *G. aculeatus* and *T. nigroviridis* genomes which contained orthologs of sea bass candidate genes and linkage markers were used as input data in the MAPCHART software (Voorrips, 2002) to design the "candidate genes comparative map".

2.3. Identification of QTL-related genomic and mRNA sequences

The sequences of linkage markers from the sea bass linkage groups (1, 4, 6 and 15) positioned within the sea bass growth related QTL were queried against the *G. aculeatus* genome using ENSEMBL BLASTVIEW (www.ensembl.org). The BLASTN algorithm was used with distant homology parameters (-E:10 –B:100 –filter:dust –RepeatMasker –W:9 –M:1 –N:-1 –Q:2 –R:1)) for high sensitivity and a cut-off E-value of $5e^{-3}$. The MARTVIEW tool from the BIOMART system (www.biomart.org) was used to characterize the *G. aculeatus* genomic region orthologous to the sea bass QTL. This allowed the retrieval of the features as Ensembl gene ID plus the respective Gene Start (bp), Gene End (bp), and strand from the genomic region in question.

Genomic and mRNA databases were created with available sea bass genomic sequences (WGS and BAC ends) (Kuhl, et al., 2010) and expressed sequence tags (ESTs) (Louro, et al., 2010) using the formatDB executable within the NCBI blast-2.2.20 standalone pack (ftp://ftp.ncbi.nlm.nih.gov/blast/executables/release/2.2.20/). The sea bass genomic database was composed of 147,763 sequences and 157,675,311 letters. The *G. aculeatus* regions homologous to the flanking markers of sea bass QTL were queried against the sea bass genomic database using BLASTN (-e 10^{-5} -T F -v 10 -b 10) to identify the relevant genomic region in sea bass. All the identified genomic sequences were retrieved. If BAC ends from a unique sea bass BAC clone were both present in reverse orientation with a significant hit in the blast output, the genomic coordinates of the hits were annotated. The sequence of the *G. aculeatus* genomic region orthologous to the sea

bass QTL was used to search BAC end sequences and identify BAC clones covering the QTL and to design a BAC tiling map of the genomic region.

The sea bass transcript database was formed of 17,716 sequences and 12,381139 letters. The identified and retrieved *G. aculeatus* genes within the syntenic regions with the markers flanking the sea bass QTL were blasted against the sea bass transcript database using BLASTN (-e 10^{-5} -T F -v 10 -b 10). This BLASTN output of the available sea bass transcript sequences was used to identify genes in the target QTL. The sequences of all the identified transcripts were retrieved and sequence ID annotated in a table containing the corresponding genes identified in the orthologous region of the *G. aculeatus* genome.

2.4. Marker development and genotyping

All the identified genomic sequences were scanned for microsatellites using the FASTPCR software repeat search (default parameters) (Kalendar, et al., 2009). Sequences containing microsatellites were BLAST queried against the whole *G. aculeatus* genome, in order to eliminate sequences with high similarity to multiple genomic regions. This approach ensured that duplicate genes present in multiple chromosomes were excluded from the analysis. The BLASTN algorithm was used to interrogate the target genomes (masked) database and distant homology parameters (-E:10 –B:100 –filter:dust –RepeatMasker –W:9 –M:1 –N:-1 –Q:2 –R:1)) were used to maximize sensitivity. The cut-off for a positive hit was taken at an E-value $5e^{-3}$. The sea bass genomic sequences that were restricted to sea bass QTL and contained microsatellites were selected for primer design. Criteria for selection included the presence of loci spread over as wide a range as possible of the target chromosome regions, potential for microsatellite development, potential for polymorphism, and primer design conditions.

PCR primers were designed to flank microsatellites of interest using PRIMER3 (http://frodo.wi.mit.edu/) (Rozen, Skaletsky, 2000). Compatibility among the primers designed was

checked using the FastPCR "primers list test" option (all primers against all primers). The forward primer was fluorescently labelled for detection of PCR products, and the protocol was carried out with a Qiagen Multiplex PCR kit (QIAGEN) using the conditions recommended by the manufacturer. Thirty three primer pairs were designed from the 50 loci chosen within QTL for marker development. Six of the primer pairs developed for markers were rejected because either they produced a PCR product of a significantly different size from that predicted (n=2), gave undefined alleles with stutter peaks (n=2), failed to amplify a product (n=1) or because the locus was homozygous (n=1). 27 markers were selected for development of three PCR multiplex amplification reactions.

The genomic panel was genotyped for 27 microsatellites with the three developed PCR multiplexes, using an ABI3730xl DNA analyser (Applied Biosystems) and an internal GENESCAN 500-LIZ size standard (Applied Biosystems) to determine allele sizes. Genotyping output was obtained using the GENEMAPPER v.4.0 software (Applied Biosystems). Markers sequence, allele variations and individual genotypes were submitted to the Probedb (http://www.ncbi.nlm.nih.gov/probe/) database with the same nomenclature used in this manuscript.

2.5. QTL mapping population

Massault et al (2010) detected two significant QTL for body weight in linkage group (LG) 4 and LG6 and six significant QTL for morphometric traits in LG1B, LG4, LG6, LG7, LG15 and LG23. Chatziplis et al (2007) using a different genomic panel described a growth QTL positioned in LG1. The sea bass genomic panel used in the present study for QTL targeted marker development and genotyping consisted of the same 5 full-sib families, comprised of 2 dams, 5 sires and their 8 months old offspring (n=576) of 40.6$\pm$14.21g body mass and 13.3$\pm$1.61 cm standard length previously described by Massault et al. (2010). The population structure (Table 6) of the genomic panel was selected from a larger pool of families (56 parents and 1151 progeny), identified by parentage

assignment (Volckaert, et al., 2012) to optimize power of QTL detection (Massault, et al., 2008). Individuals were phenotyped for 9 growth related traits: standard length (SL), head length (HL), body length (BL), pre anal length (PRAL), abdominal length (AL), post anal length (POAL), head depth (HD), body depth (BD) and body mass (BW); see detailed description in Massault et al (2010).

2.6. Genetic map

The QTL population linkage map was generated using software CRI-MAP 64-bits v. 2.503, a revised version of CRIMAP v. 2.4 (Green, et al., 1990) modified by Jill Maddox and Ian Evans (Jill Maddox, University of Melbourne, personal communication). Retrieved data from RESSPECIES (www.ResSpecies.org) (Law, Archibald, 2000) were checked for Mendelian errors prior to map construction. The data formatted in the ".gen" file consisted of 118 microsatellite marker genotypes of the 576 offspring from the five sires and two dams included in the pedigree. These 117 microsatellite marker genotypes consist of 27 new marker genotypes plus the previous 90 marker genotypes (87 markers in linkage) (Massault, et al., 2010). The CRI-MAP options "two-points", "build", "flips" and "fixed" were used to obtain the sex specific and averaged linkage group maps with a LOD threshold of three. Figures of linkage groups were designed using software MAPCHART version 2.2 (Voorrips, 2002).

2.7. QTL analysis

Two methods were used to detect QTL, half-sib regression analysis (HSr) (Knott, et al., 1996), and sib-pair regression analysis (SPr), both using the GRIDQTL web based interface (http://www.gridqtl.org.uk). Genotype, map and phenotype files were submitted to the half-sib and sib-pair portlets for the HSr and SPr analysis, respectively. One QTL and two QTL models per linkage group were tested at 1cM step regressions. Significant QTL were detected with standard F-test statistic as previously described (Knott, et al., 1996), in which genome wide permutation tests with 10,000 iterations were made to define the significance threshold. Bootstrap with 10,000

248 resamplings were made to define a 95% confidence interval of the detected QTL. In the SPr analysis
249 the residuals from the fixed effect analysis are used accordingly to Visscher & Hopper method
250 (Visscher, Hopper, 2001).

251 ### 2.8. QTL genes and annotation

252 The previous workflow carried out prior the availability of the sea bass genome assembly and
253 annotation, through a comparative genomics approach for marker development, allowed the
254 predictive characterization of the QTL regions using *G. aculeatus* (GAC) and *T. nigroviridis* (TNI)
255 ENSEMBL proteins. QTL genomic characterization was confirmed using the annotated sea bass
256 genome assembly (Tine, et al., 2014). Previously published and newly developed microsatellite
257 markers were queried via BLASTN against the recently sequenced sea bass genome in order to
258 establish a physical map.
259 A tabulated list of genes within the QTL confidence intervals was retrieved (Additional file 5) based
260 on the mapped microsatellite marker positions.The deduced gene products in the QTL confidence
261 intervals were mapped to KEGG pathways via orthology assignment using KAAS (KEGG
262 Automatic Annotation Server) (Moriya, et al., 2007) (http://www.genome.jp/tools/kaas). The
263 deduced protein sequences of genes in the QTL region were queried against a eukaryotic
264 representative gene data set using the bi-directional best hit assignment method. BRITE hierarchy
265 files were retrieved to visualise which sea bass genes are mapped to KEGG pathways using the
266 KegHier software (http://www.genome.jp/download). Gene orthology (KEGG Orthology, KO) and
267 KEGG pathway IDs were integrated with the tabulated list of genes.
268 The KOBAS v.2.0 (KEGG Orthology Based Annotation System) web server (Xie, et al., 2011)
269 (http://kobas.cbi.pku.edu.cn) was queried for annotation and identification of enriched pathways
270 within the genomic regions of interest (QTL confidence intervals). KO terms of the annotated genes
271 within the QTL were used to identify statistically enriched related pathways in the respective genome

region, with default cut-offs and default hypergeometric statistical test and Benjamini-Hochberg FDR correction, using a whole genome set as background.

The STRING v. 9.0 protein interaction database (Szklarczyk, et al., 2011) (http://string-db.org) was queried for gene network inferences within the genomic regions of interest (QTL confidence intervals). The interactions include direct (physical) and indirect (functional) associations derived from four sources: genomic context, high-throughput experiments, co-expression, and previous knowledge. The confidence of the interaction score was defined at maximum highest confidence interactions (>0.9) and the query was restricted to the population of genes in each of QTL confidence intervals.

To assess and validate the predictive power of the comparative approach interspecies synteny analysis was performed. Whole genome alignment files (.bed) of seven teleost reference genomes (*G. aculeatus, O. latipes, T. nigroviridis, T. rubripes, O. niloticus, G. morhua* and *D. rerio*) were retrieved from the sea bass genome browser (http://seabass.mpipz.de/). These whole genome alignments resulted from genomic pairwise alignments against sea bass LG4, LG6, and LG15 chromosomes and integration into blocks of shared collinearity. Circular genomic synteny representations were created using Circos (Krzywinski, et al., 2009).

3. Results

3.1. Genomic and genetic comparative maps

Sea bass 2nd generation linkage map *vs G. aculeatus* and *T. nigroviridis* genomes comparative maps allowed the relative linkage of the candidate genes, previously identified by bibliographic search, with the sea bass genetic markers and respective linkage groups. Inversely the comparative mapping between the linkage map of the QTL population (Massault, et al., 2010) allowed the relative linkage

between the identified QTL loci and *G. aculeatus* and *T. nigroviridis* physical genomes to be made (Figure 1). These comparative mappings were the scaffolds for new genetic marker development aimed to flank the previously identified QTLs in sea bass LG1, LG4, LG6 and LG15. Twenty eight 28 new microsatellites markers were designed based on these interspecies synteny blocks.

3.2. Genetic linkage map

A genetic map of sea bass representing 21 linkage groups (two or more markers in linkage) was made with 113 markers and incorporated the 27 new microsatellite markers developed in the current study (Table 1, Additional file 2). Only 1 of the 10 markers targeted a tentative growth QTL identified in LG1 (at present LG1a and LG1b) in a previous study with a different mapping population was linked (Chatziplis, et al., 2007) to LG1a. The remaining 17 new markers targeted the previously identified growth QTL on LG4 (n=8), LG6 (n=7) and LG15 (n=2) (Massault, et al., 2010) in the same QTL mapping population used in the present study (Additional file 2). In the targeted linkage groups, respectively, 3/8, 6/7 and 2/2 markers were in linkage in LG4, LG6 and LG15. The new markers which were not in linkage with the targeted LGs were dispersed among other LGs. The markers developed for LG4 which were not in linkage (5/8) were designed around a previously identified LG4 marker DLA0166 (Massault, et al., 2010), which in the present study failed to reach the LOD 3 threshold.

The inclusion of further markers in the European sea bass linkage map modified the position of previously mapped markers. Genotype data of markers DLA0251e, DLA0402 retrieved from ResSpecies database, which were used in the construction of the first sea bass linkage map, were homozygous, and consequently were excluded from the present map. No linkage was found for the markers DLA0245 and DLA0037, previously placed in LG9 and LG18, respectively in a different mapping population (Chistiakov, et al., 2005). LG16, LG21, and LG22 described in (Chistiakov, et

al., 2005) are not represented because no markers from those linkage groups were used to genotype the sea bass QTL mapping population in the present study.

The genetic map covers 644.2 cM in total length. The addition of the new markers to the previous map generated from the same mapping population led to an increase in linkage groups from 20 to 21, but the total length remained similar changing from 639 to 644 cM. This resulted in a considerable decrease in the averaged distance between markers, changing from 7.7 to 5.7 cM. The average distance between markers in LG4, LG6 and LG15 was reduced from 21.2, 6.2, and 4.0 cM to 8.3, 2.4 and 3.6 cM, respectively. Furthermore, in the specific region targeted in LG4, LG6 and LG15, the average distance between markers decreased from 23.4, 9.1, and 5.8 cM to 3.4, 2.2, and 5.2 cM, respectively.

3.3. QTL analysis

Significant QTL were detected for morphometric traits SL, HL, BL, PRAL, AL, POAL, HD, BD, and BW in both maternal and paternal half sibs (HS) and sib pairs (SP) regression analyses (Table 2; Figure 1). Results values of analyses for all linkage groups are described in additional file 3. The exception was trait HD in the paternal HS analysis which was only detected at a suggestive level (>5% chromosome-wide, CW). Significant QTL (>5% genome-wide, GW) were detected for trait body weight (BW) in both the maternal (MHS) and paternal (PHS) half sib analyses but was only at a suggestive level in the SP analysis (Table 2; Figure 1). The average information content (IC) of the linkage analysis of the sea bass genome was 0.81 for the MHS and 0.82 for the PHS. The average IC of each linkage group for both MHS and PHS analysis is indicated in Table 1. None of the significant or suggestive QTL in the linkage map fitted a two QTL model in any of the regression analysis (i.e. MHS, PHS and SP).

Eight of the 21 linkage groups, LG1b, LG2, LG4, LG6, LG7, LG15, LG20, and LG24 contained one or more morphometric QTL that passed the significance threshold (>5% GW). In general if one

morphometric QTL was detected at a significant level in a LG other morphometric QTL in the same LG were also detected at a significant or a suggestive level. This could be considered as an indirect indication of high genetic correlations between morphometric traits, as already have been shown by others (Chatziplis, et al., 2007; Volckaert, et al., 2012) (Table 2). Likewise, in general, significant and suggestive QTL for BW were highly concordant with significant morphometric QTL. If all morphometric traits are grouped as a single morphological trait (MORPH) following previous QTL nomenclature (Massault, et al., 2010) assigning the MORPH QTL peak location to be the strongest of the overlapping morphometric QTL detected, then the results of MHS, PHS and SP analysis gave a total of 14 significant MORPH QTL in the 8 LGs. However, when morphometric QTL traits were considered individually (SL, HL, BL, PRAL, AL, POAL, HD and BD) in MHS, PHS and SP analysis then a total of 36 significant QTL were identified (in bold in Table 2, trait column). The greatest MORPH QTL effect among the targeted LGs was for LG6 (27%), followed by LG4 (14%) and LG15 (10%). As previously reported by Massault et al. (2010) the present analysis of LG1a failed to identify a growth QTL and none of the markers developed were linked to this group. For this reason LG1a was excluded from further analysis.

In LG6, all 8 morphometric traits considered had a QTL peak between 10 and 15 cM. The most significant MORPH QTL in LG6 with MHS analysis had a QTL effect explaining up to 27% of the phenotypic variation (BL trait, F=20.4 with 5% GW threshold of 6.99) with a peak located at 15 cM and confidence interval spanning from 8 to 24 cM. SP analysis also gave a significant MORPH QTL in LG6 (trait POAL, F=97.22 with 5% GW threshold of 20.03) with a peak located at 14 cM and confidence interval spanning from 11 to 15 cM. No significant MORPH QTL was detected by PHS analysis, however, a suggestive MORPH QTL (SL, BL, AL, and POAL) was identified between 14and 23 cM and had a confidence intervals of 5-24 cM. Additional suggestive QTL were found by MHS analysis (HD, and BD, 1%CW, table 2) and SP analysis (HL, and BD, 1%CW, table 2).

In linkage group LG4, significant MORPH QTL were detected with all three regression analysis (MHS, PHS and SP). The peak of MHS QTL was at the beginning of the linkage group (1 cM), PHS QTL at 10 cM and SP QTL at the end of the LG (57 cM). Six MHS morphometric QTL (SL, HL, PRAL, AL, POAL, and HD) were detected with only one at a significant level (trait BD, F=7.15 with 5% GW threshold of 6.84), which explained up to 9% of the phenotypic variance (Table 2). In PHS analysis, 8 morphometric QTL were detected four of which were significant (SL, BL, PRAL, POAL, and BD, table 2). The highest QTL, trait BL (F=4.93 with 5% GW threshold of 4.30) explained up to 14% of the phenotypic variance. In the SP analysis, 8 morphometric QTL were detected, two (HL, and BD) at a significant level (Table 2). Confidence interval for the MSH MORPH QTL spanned from 0 to 17 cM and for PHS and SP analysis covered most of LG4.

Two significant BW QTL were found in LG4 and LG6 explaining approximately 14% and 12% of phenotypic variance in the PHS and MHS analysis, respectively. Moreover, the BW QTL peak (10 cM) in LG4 coincided with the MORPH QTL peak. The peak of the BW QTL in LG6 (15 cM) is in exactly the same position as the MHS and SP MORPH QTL. In general suggestive BW QTL in LG4 (MHS and SP) and LG6 (SP) were highly correlated with significant MORPH QTL.

In LG15 significant MORPH QTL were only detected with MHS regression analysis (Table 2). Four morphometric MHS QTL (HL, PRAL, HD and BD) were significant with the strongest trait, BD (F=7.87 with 5% GW threshold of 6.84), explaining up to 10% of the phenotypic variance. MORPH QTL were detected at a suggestive level with PHS and SP analysis (Table 2). Suggestive BW QTL detected with MHS, PHS and SP analysis were closely linked with the respective MORPH QTL.

MORPH QTL were also detected in LG1b, LG2, LG7, LG20 and LG24 but were not targeted for marker development (Table 2). In brief, LG1b PHS and SP analysis detected a MORPH QTL for the trait HL which explained up to 15% of the phenotypic variance in the PHS analysis. In LG2, a single significant QTL was detected by PHS analysis for trait AL, with an effect explaining up to 12% of the phenotypic variance. In LG7 both PHS and SP analysis gave significant MORPH QTL (SL, BL,

AL) at 30 cM and 28 cM, respectively which explained up to 18% of the phenotypic variance in the PHS analysis. In LG20, trait BD gave a significant QTL in MHS and SP analysis, which explained 9% of the phenotypic variance in in MHS analysis. In LG24, significant MORPH QTL were only detected with PHS analysis.

Suggestive MORPH and BW QTL (Additional file 4) were detected in other linkage groups (LG3, LG5, LG8, LG9, LG11, LG12, LG13, LG14, LG17, and LG19), and less overlap was found between QTL in those linkage groups.

3.4. QTL gene annotation and pathway analysis

Microsatellite markers in targeted linkage groups (LG4, LG6 and LG15) were physically mapped in the sea bass genomic assembly (Figure 1). The genomic regions containing microsatellites corresponding to growth QTL confidence intervals in LG4, LG6 and LG15, were extracted from the annotated sea bass genome assembly. A list of genes within the genomic regions of interest (QTL confidence intervals) was retrieved and corresponded to 518 (MHS MORPH), 594 (PHS MORPH & BW) and 149 (SP MORPH) genes in LG4, 528 (MHS MORPH & BW) and 233 (SP MORPH) genes in LG6 and 347 (MHS MORPH), and 110 (SP MORPH) genes in LG15 (Figure 3, Additional file 5). No significantly enriched KEGG ortholog (KO) reference pathways were identified in LG4 or LG15. One significantly enriched pathway was identified in LG6, which had a p-value of 0.04 and corresponded to “the axon guidance (ko04360) pathway” (Table 3). Of the 177 genes with KO annotation in LG6, 12 were present in the 127 genes comprising the axon guidance (ko04360) pathway (Table 3). Adipocytokine, MAPK, jak-STAT, and insulin KO reference signalling pathways, although not statistically significantly enriched for the traits in question, were the three most relevant for the trait KO pathways which appeared in the top pathways with their elements mapped with the genes present within the QTL confidence intervals (Table 4).

Using the maximum confidence interaction score stringent identification of interacting gene products from LG4, LG6 and LG15 were identified by STRING analysis (Table 5). The most significant association networks constructed included; 13 genes for LG4, 25 genes for LG6 and 15 genes for LG15 (confidence view, Figure 3a). A large global association network including 42 gene products arises when the gene products of the 3 linkage groups are analysed together (Figure 4b, Table 5). Strikingly the three association networks arising for LG4, LG6 and LG15 are populated with genes of biological importance for growth and body weight. For example, elements of the signalling pathways for Jak-STAT, insulin, MAPK and adipocytokine previously identified in the KO pathways analysis (Table 4). In addition, endocrine factors such as insulin, growth hormone, IGFI and II are also present in the association networks. The largest association network with the greatest number of gene products interlinked is obtained with the genes mapped in the LG4 confidence interval. In LG6 two main association networks were evident that merge via SMAD family member 3 (SMAD3) when the confidence threshold is lowered slightly (high confidence, >0.7). In LG15 a single large and highly significant association network is identified with myostatin and insulin growth factors binding proteins (IGFBP). Merging of the gene lists from the 3 LGs and using the highest confidence score for network analysis generates a single network with very strong associations (Figure 4b). The genes represented are good candidate genes for the morphometric growth-related QTL.

The synteny analysis results are concordant with the previous comparative mapping results that allowed the development of QTL flanking markers, candidate genes mapping and predictive QTL charactherization. In the overall synteny analysis shows that toward the tipping of the chromosomes the rate of synteny tends to diminish and blocks of synteny are smaller and with more mixed origin from different chromosomes (Figure 3). This might be the reason the why most new markers aimed for the tip of LG1a group ended up all linkage in LG25 (Additional file 2), as well as markers aimed to be in close linkage with marker DLA0166 also located at the tip of LG4.

4. Discussion

4.1. Delimiting the QTL

We have narrowed down the confidence interval of the QTL regions for growth in sea bass through a comparative approach. To that purpose we developed markers to increase the resolution within genomic regions flanking the QTL peaks in the sea bass genetic map. This was done by linking the genomic data and genetic information of sea bass with an assembled and annotated genome of a closely related model species. Although it did not lead to the identification of a specific growth related candidate gene it was possible, by taking advantage of demonstrated synteny in advanced teleosts (Sarropoulou, et al., 2007; Shimizu, Purugganan, 2005), to extract a list of candidate genes. Their functional analysis revealed to be consistent with growth related traits.

Identification of the causal gene for any trait is clearly a limiting step; even in the pig, in which over 1,675 candidate QTL regions have been identified over the past 15 years, only in a few cases has the causal mutation been identified (Rothschild, et al., 2007). Nevertheless, in the present study, the confidence interval of the growth QTL regions on sea bass on LG4, LG6 and LG15 were significantly reduced in comparison with previous studies (Chatziplis, et al., 2007; Massault, et al., 2010). Subsequent access to the sea bass draft genome confirmed the validity of the comparative mapping approach using genetic markers flanking the QTL and revealed that LG4, LG6 and LG15 contain approximately 871, 915 and 774 genes, respectively. A similar approach in rainbow trout using comparative genomic analysis of linkage groups containing growth QTL identified syntenic blocks in the zebrafish (*Danio rerio*), medaka (*Oryzias latipes*) and stickleback genomes with 61%, 56.5% and 53.6% gene homologies, respectively. This is high considering their evolutionary distance and differing ecological niche (Wringe, et al., 2010). A higher gene homology exists in our study, however, between sea bass and medaka and/or stickleback with large conserved blocks of synteny in the QTL interval regions (Additional file 6).

462 The largest effect BW and MORPH QTL co-localized in locus LG6 in sea bass which has a
463 homologs in rainbow trout RT-27q (Drew, et al., 2007; O'Malley, et al., 2003; Wringe, et al., 2010)
464 and Artic charr AC-4 (Kuttner, et al., 2011) linkage groups with detected BW and K (condition
465 factor) QTL. This genomic region was shown both in the Wringe et al. (2010) study and our own to
466 be homolog with stickleback chromosome group XIX and medaka chromosome 13. The results
467 support the strategy of using comparative genomics to accelerate the identification of genomic
468 regions with higher probability of having trait effect in particular in species which have undergone
469 little selection.

470 ### 4.2. Search for candidate genes

471 Minor divergence exists in the morphometric traits linked with the growth QTL between the three
472 LGs in sea bass. For example LG15 lacked the trait BW while LG4 and LG6 BW had a significant
473 association but the latter LG differed in relation to body length and body depth. Overall LG6 had the
474 greatest QTL effect of 26%. In the search for possible candidate genes the enrichment of QTL
475 regions with growth associated genes was assessed by KEGG and revealed pathway enrichment for
476 sugar, amino acid and nucleotide metabolism in LG4, PPAR signaling and focal adhesion in LG6,
477 and protein digestion, hedgehog signaling and ECM receptor interaction in LG15, although not
478 significantly enriched. Gene by gene analysis of the enriched pathway did not reveal any outstanding
479 candidates with a potential quantitative effect on growth.
480 An alternative approach to identify candidate growth genes is to select genes reported in the
481 literature to be causal or strong candidate genes for growth QTL in other vertebrates. Taking such an
482 approach, LG4 was found to contain the leptin receptor which has previously been associated with
483 backfat thickness in pigs (*Sus scrofa*) (Mackowski, et al., 2005; Munoz, et al., 2009), and insulin-like
484 growth factor binding protein-1 and -3, polymorphisms of which in chicken are associated with body
485 weight at 10 weeks (Ou, et al., 2009).

The sea bass LG6 has a high quantitative effect on growth and contains several genes, such as insulin-like growth factor 2 (IGF2), cathepsin D and leptin, with a proven causal effect on growth in other vertebrates. Of particular interest is IGF2 which in pigs is linked to a significant increase in muscle mass (30%) when a G>A substitution occurs in intron 3 (Van Laere, et al., 2003). Cursory analysis of LG15 immediately yielded an apparently strong candidate gene for growth, growth differentiation factor 8 (GDF 8) also known as myostatin. It is a negative regulator of muscle growth mutations which are responsible for double muscling in cattle breeds (Kambadur, et al., 1997; McPherron, Lee, 1997; Wiener, et al., 2002) and it has a causal effect on carcass yield in cattle (Sellick, et al., 2007) and sheep (Hadjipavlou, et al., 2008; Walling, et al., 2004). Myostatin is relatively well conserved from mammals to fish (Pie, Alvares, 2006) but the effect of knockdown or knock-out in teleosts is variable. In medaka only expression of a dominant negative forms causes double muscling (Sawatari, et al., 2010), but in zebrafish injection of antisense RNA-expressing vector to establish a myostatin gene knockdown is enough to cause double muscling phenotype (Lee, et al., 2009). The variable effect of myostatin knock-outs on teleost fish may be related to the presence of duplicate and even quadruplicate genes (2 in most osteichthyans and 4 in salmonids; Rodgers, Garikipati, 2008) as a consequence of teleost specific whole genome duplications (Jaillon, et al., 2004).

Identification of causal genes for a given trait remains a key challenge in QTL analysis and is costly and laborious. Selection of candidate genes through knowledge about their biological function gleaned through bibliographic searches are one way of establishing gene lists but this approach is limited by the relatively small number of genes identified and validated, even in terrestrial farm animals. The development of concordant QTL maps offers an alternative approach but their validity and cross species applicability remains to be established. Yet another approach applicable to organisms with a sequenced genome, from which it is possible to draw up gene lists for a given QTL region, is to apply *in silico* methods such as STRING analysis to identify protein-protein interactions.

Such an approach permits genetic data to be integrated with the wealth of functional information in public domain databases on human and model species. In fact in the present study STRING queries yielded a list of highly interlinked candidate genes for growth related traits which will be exploited in future studies. Genes identified included growth hormone (Besson, et al., 2005; McCormack, et al., 2009), insulin (Edghill, et al., 2008), and leptin (Halaas, et al., 1995; Mammes, et al., 2000) which encode key hormones modulating growth, development, and energy metabolism. They have known allelic variations linked to growth and body weight-related traits variations and related diseases in mammals. Other genes in theses pathways, upstream or downstream to these key hormones, have known allelic variation linked to growth and body weight related traits such as IGF2 (Van Laere, et al., 2003), and mutations with major QTL effects on muscle growth in pig and sheep (Clop, et al., 2006). Overall the results point to a rapid means of identifying candidates for further investigation.

5. Conclusions

A comparative genomics approach, combined with functional gene annotation, provided an identification of candidate genes responsible for growth related QTL in sea bass. This study demonstrates the feasibility of such an approach to refine the resolution of QTL and the establishment of hypothesis to accelerate discovery of putative responsible genes. As production of genomic data is becoming more accessible, the implementation of this strategy will rapidly and efficiently provide the tools required for genetic selection in new candidate aquaculture species.

Acknowledgements

This study has been funded by the European Commission (project AQUAFIRST contract number FP6-STREP-2004-513692) and the Portuguese National Science Foundation (Pluriannual funds). DJK and acknowledge financial support from the BBSRC. BL acknowledges fellowships from FCT

533 (SFRH/BD/29171/2006, SFRH/BPD/89889/2012) and SABRETRAIN Marie Curie EST, and
534 mobility and resource exchange grants from the Aquagenome project (FP6 contract number 44481).

535 # References

536 Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search
537 tool. Journal of molecular biology. 215, 403-410.

538 Besson, A., Salemi, S., Deladoey, J., Vuissoz, J.M., Eble, A., Bidlingmaier, M., Burgi, S., Honegger,
539 U., Fluck, C., Mullis, P.E., 2005. Short stature caused by a biologically inactive mutant
540 growth hormone (GH-C53S). The Journal of clinical endocrinology and metabolism. 90,
541 2493-2499.

542 Brunet, F.G., Roest Crollius, H., Paris, M., Aury, J.M., Gibert, P., Jaillon, O., Laudet, V., Robinson-
543 Rechavi, M., 2006. Gene loss and evolutionary rates following whole-genome duplication in
544 teleost fishes. Molecular Biology and Evolution. 23, 1808-1816.

545 Canario, A.V.M., Bargelloni, L., Volckaert, F., Houston, R.D., Massault, C., Guiguen, Y., 2008.
546 Genomics Toolbox for Farmed Fish. Reviews in Fisheries Science. 16, 3-15.

547 Chatziplis, D., Batargias, C., Tsigenopoulos, C.S., Magoulas, A., Kollias, S., Kotoulas, G.,
548 Volckaert, F.A.M., Haley, C.S., 2007. Mapping quantitative trait loci in European sea bass
549 (*Dicentrarchus labrax*): The BASSMAP pilot study. Aquaculture. 272, S172-S182.

550 Chistiakov, D., Tsigenopoulos, C., Lagnel, J., Guo, Y., Hellemans, B., Haley, C., Volckaert, F.,
551 Kotoulas, G., 2008. A combined AFLP and microsatellite linkage map and pilot comparative
552 genomic analysis of European sea bass *Dicentrarchus labrax* L. Animal Genetics. 39, 623-
553 634.

554 Chistiakov, D.A., Hellemans, B., Haley, C.S., Law, A.S., Tsigenopoulos, C.S., Kotoulas, G.,
555 Bertotto, D., Libertini, A., Volckaert, F.A., 2005. A microsatellite linkage map of the
556 European sea bass *Dicentrarchus labrax* L. Genetics. 170, 1821-1826.

557 Clop, A., Marcq, F., Takeda, H., Pirottin, D., Tordoir, X., Bibe, B., Bouix, J., Caiment, F., Elsen,
558 J.M., Eychenne, F., Larzul, C., Laville, E., Meish, F., Milenkovic, D., Tobin, J., Charlier, C.,
559 Georges, M., 2006. A mutation creating a potential illegitimate microRNA target site in the
560 myostatin gene affects muscularity in sheep. Nature genetics. 38, 813-818.
561 Drew, R.E., Schwabl, H., Wheeler, P.A., Thorgaard, G.H., 2007. Detection of QTL influencing
562 cortisol levels in rainbow trout (*Oncorhynchus mykiss*). Aquaculture. 272, S183-S194.
563 Du, S.J., Gong, Z., Fletcher, G.L., Shears, M.A., King, M.J., Idler, D.R., Hew, C.L., 1992. Growth
564 enhancement in transgenic Atlantic salmon by the use of an "all fish" chimeric growth
565 hormone gene construct. Nat Biotech. 10, 176-181.
566 Edghill, E.L., Flanagan, S.E., Patch, A.M., Boustred, C., Parrish, A., Shields, B., Shepherd, M.H.,
567 Hussain, K., Kapoor, R.R., Malecki, M., MacDonald, M.J., Stoy, J., Steiner, D.F., Philipson,
568 L.H., Bell, G.I., Neonatal Diabetes International Collaborative, G., Hattersley, A.T., Ellard,
569 S., 2008. Insulin mutation screening in 1,044 patients with diabetes: mutations in the INS
570 gene are a common cause of neonatal diabetes but a rare cause of diabetes diagnosed in
571 childhood or adulthood. Diabetes. 57, 1034-1042.
572 FAO, 2010. The state of the world fisheries and aquaculture. Food and Agriculture Organization of
573 the United Nations, Rome.
574 Fjalestad, K.T., Moen, T., Gomez-Raya, L., 2003. Prospects for genetic technology in salmon
575 breeding programmes. Aquaculture Research. 34, 397-406.
576 Green, P., Falls, K., Crooks, S., 1990. Documentation for CRI-MAP, version 2.4. Washington
577 University, St. Louis, Missouri USA.
578 Guyon, R., Senger, F., Rakotomanga, M., Sadequi, N., Volckaert, F.A., Hitte, C., Galibert, F., 2010.
579 A radiation hybrid map of the European sea bass (*Dicentrarchus labrax*) based on 1581
580 markers: Synteny analysis with model fish genomes. Genomics. 96, 228-238.

581 Hadjipavlou, G., Matika, O., Clop, A., Bishop, S.C., 2008. Two single nucleotide polymorphisms in
582 the myostatin (GDF8) gene have significant association with muscle depth of commercial
583 Charollais sheep. Animal Genetics. 39, 346-353.
584 Halaas, J.L., Gajiwala, K.S., Maffei, M., Cohen, S.L., Chait, B.T., Rabinowitz, D., Lallone, R.L.,
585 Burley, S.K., Friedman, J.M., 1995. Weight-reducing effects of the plasma protein encoded
586 by the obese gene. Science. 269, 543-546.
587 Hoegg, S., Brinkmann, H., Taylor, J.S., Meyer, A., 2004. Phylogenetic timing of the fish-specific
588 genome duplication correlates with the diversification of teleost fish. Journal of molecular
589 evolution. 59, 190-203.
590 Houston, R., Taggart, J., Cezard, T., Bekaert, M., Lowe, N., Downing, A., Talbot, R., Bishop, S.,
591 Archibald, A., Bron, J., Penman, D., Davassi, A., Brew, F., Tinch, A., Gharbi, K., Hamilton,
592 A., 2014. Development and validation of a high density SNP genotyping array for Atlantic
593 salmon (Salmo salar). BMC Genomics. 15, 90.
594 Houston, R.D., Haley, C.S., Hamilton, A., Guy, D.R., Tinch, A.E., Taggart, J.B., McAndrew, B.J.,
595 Bishop, S.C., 2008. Major quantitative trait loci affect resistance to infectious pancreatic
596 necrosis in Atlantic salmon (*Salmo salar*). Genetics. 178, 1109-1115.
597 Hull, K.L., Harvey, S., 1999. Growth hormone resistance: clinical states and animal models. The
598 Journal of endocrinology. 163, 165-172.
599 Jaillon, O., Aury, J.M., Brunet, F., Petit, J.L., Stange-Thomann, N., Mauceli, E., Bouneau, L.,
600 Fischer, C., Ozouf-Costaz, C., Bernot, A., Nicaud, S., Jaffe, D., Fisher, S., Lutfalla, G.,
601 Dossat, C., Segurens, B., Dasilva, C., Salanoubat, M., Levy, M., Boudet, N., Castellano, S.,
602 Anthouard, V., Jubin, C., Castelli, V., Katinka, M., Vacherie, B., Biemont, C., Skalli, Z.,
603 Cattolico, L., Poulain, J., De Berardinis, V., Cruaud, C., Duprat, S., Brottier, P., Coutanceau,
604 J.P., Gouzy, J., Parra, G., Lardier, G., Chapple, C., McKernan, K.J., McEwan, P., Bosak, S.,
605 Kellis, M., Volff, J.N., Guigo, R., Zody, M.C., Mesirov, J., Lindblad-Toh, K., Birren, B.,

606 Nusbaum, C., Kahn, D., Robinson-Rechavi, M., Laudet, V., Schachter, V., Quetier, F.,
607 Saurin, W., Scarpelli, C., Wincker, P., Lander, E.S., Weissenbach, J., Roest Crollius, H.,
608 2004. Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early
609 vertebrate proto-karyotype. Nature. 431, 946-957.
610 Jiang, Z., Pappu, S.S., Rothschild, M.F., 2007. Hitting the jackpot twice: identifying and patenting
611 gene tests related to muscle lipid accumulation for meat quality in animals and type 2
612 diabetes/obesity in humans. Recent patents on DNA & gene sequences. 1, 100-111.
613 Kalendar, R., Lee, D., Schulman, A., 2009. FastPCR Software for PCR Primer and Probe Design and
614 Repeat Search. Genes, Genomes and Genomics. 3, 1-14.
615 Kambadur, R., Sharma, M., Smith, T.P., Bass, J.J., 1997. Mutations in myostatin (GDF8) in double-
616 muscled Belgian Blue and Piedmontese cattle. Genome Research. 7, 910-916.
617 Kim, K.S., Thomsen, H., Bastiaansen, J., Nguyen, N.T., Dekkers, J.C., Plastow, G.S., Rothschild,
618 M.F., 2004. Investigation of obesity candidate genes on porcine fat deposition quantitative
619 trait loci regions. Obesity research. 12, 1981-1994.
620 Knott, S.A., Elsen, J.M., Haley, C.S., 1996. Methods for multiple-marker mapping of quantitative
621 trait loci in half-sib populations. Theor Appl Genet. 93, 71-80.
622 Krzywinski, M., Schein, J., Birol, İ., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J., Marra,
623 M.A., 2009. Circos: An information aesthetic for comparative genomics. Genome Research.
624 19, 1639-1645.
625 Kuhl, H., Beck, A., Wozniak, G., Canario, A., Volckaert, F., Reinhardt, R., 2010. The European sea
626 bass Dicentrarchus labrax genome puzzle: comparative BAC-mapping and low coverage
627 shotgun sequencing. BMC Genomics. 11, 68.
628 Kuttner, E., Moghadam, H.K., Skulason, S., Danzmann, R.G., Ferguson, M.M., 2011. Genetic
629 architecture of body weight, condition factor and age of sexual maturation in Icelandic Arctic
630 charr (*Salvelinus alpinus*). Molecular genetics and genomics : MGG. 286, 67-79.

631 Law, A.S., Archibald, A.L., 2000. Farm animal genome databases. Briefings in bioinformatics. 1,
632 151-160.
633 Lee, C.Y., Hu, S.Y., Gong, H.Y., Chen, M.H., Lu, J.K., Wu, J.L., 2009. Suppression of myostatin
634 with vector-based RNA interference causes a double-muscle effect in transgenic zebrafish.
635 Biochemical and biophysical research communications. 387, 766-771.
636 Loukovitis, D., Sarropoulou, E., Tsigenopoulos, C.S., Batargias, C., Magoulas, A., Apostolidis, A.P.,
637 Chatziplis, D., Kotoulas, G., 2011. *Quantitative trait loci* involved in sex determination and
638 body growth in the gilthead sea bream (*Sparus aurata* L.) through targeted genome scan.
639 PloS one. 6, e16599.
640 Louro, B., Passos, A.L., Souche, E.L., Tsigenopoulos, C., Beck, A., Lagnel, J., Bonhomme, F.,
641 Cancela, L., Cerda, J., Clark, M.S., Lubzens, E., Magoulas, A., Planas, J.V., Volckaert, F.A.,
642 Reinhardt, R., Canario, A.V., 2010. Gilthead sea bream (*Sparus auratus*) and European sea
643 bass (*Dicentrarchus labrax*) expressed sequence tags: Characterization, tissue-specific
644 expression and gene markers. Mar Genomics. 3, 179-191.
645 Mackowski, M., Szymoniak, K., Szydlowski, M., Kamyczek, M., Eckert, R., Rozycki, M.,
646 Switonski, M., 2005. Missense mutations in exon 4 of the porcine LEPR gene encoding
647 extracellular domain and their association with fatness traits. Animal Genetics. 36, 135-137.
648 Mammes, O., Betoulle, D., Aubert, R., Herbeth, B., Siest, G., Fumeron, F., 2000. Association of the
649 G-2548A polymorphism in the 5' region of the LEP gene with overweight. Annals of human
650 genetics. 64, 391-394.
651 Massault, C., Bovenhuis, H., Haley, C., de Koning, D.J., 2008. QTL mapping designs for
652 aquaculture. Aquaculture. 285, 23-29.
653 Massault, C., Hellemans, B., Louro, B., Batargias, C., Van Houdt, J.K., Canario, A., Volckaert, F.A.,
654 Bovenhuis, H., Haley, C., de Koning, D.J., 2010. QTL for body weight, morphometric traits

655 and stress response in European sea bass *Dicentrarchus labrax*. Animal Genetics. 41, 337-
656 345.
657 McCormack, B.L., Chase, C.C., Jr., Olson, T.A., Elsasser, T.H., Hammond, A.C., Welsh, T.H., Jr.,
658 Jiang, H., Randel, R.D., Okamura, C.A., Lucy, M.C., 2009. A miniature condition in
659 Brahman cattle is associated with a single nucleotide mutation within the growth hormone
660 gene. Domestic animal endocrinology. 37, 104-111.
661 McPherron, A.C., Lee, S.-J., 1997. Double muscling in cattle due to mutations in the myostatin gene.
662 Proceedings of the National Academy of Sciences of the United States of America. 94,
663 12457-12461.
664 Moghadam, H.K., Poissant, J., Fotherby, H., Haidle, L., Ferguson, M.M., Danzmann, R.G., 2007.
665 *Quantitative trait loci* for body weight, condition factor and age at sexual maturation in
666 Arctic charr (*Salvelinus alpinus*): comparative analysis with rainbow trout (*Oncorhynchus*
667 *mykiss*) and Atlantic salmon (*Salmo salar*). Molecular genetics and genomics : MGG. 277,
668 647-661.
669 Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A.C., Kanehisa, M., 2007. KAAS: an automatic
670 genome annotation and pathway reconstruction server. Nucleic acids research. 35, W182-
671 185.
672 Munoz, G., Ovilo, C., Silio, L., Tomas, A., Noguera, J.L., Rodriguez, M.C., 2009. Single- and joint-
673 population analyses of two experimental pig crosses to confirm quantitative trait loci on Sus
674 scrofa chromosome 6 and leptin receptor effects on fatness and growth traits. Journal of
675 animal science. 87, 459-468.
676 O'Malley, K.G., Sakamoto, T., Danzmann, R.G., Ferguson, M.M., 2003. Quantitative trait loci for
677 spawning date and body weight in rainbow trout: testing for conserved effects across
678 ancestrally duplicated chromosomes. The Journal of heredity. 94, 273-284.

679 Ødegård, J., Moen, T., Santi, N., Korsvoll, S.A., Kjøglum, S., Meuwissen, T.H.E., 2014. Genomic
680 prediction in an admixed population of Atlantic salmon (Salmo salar). Frontiers in Genetics.
681 5, 402.
682 Ou, J.T., Tang, S.Q., Sun, D.X., Zhang, Y., 2009. Polymorphisms of three neuroendocrine-correlated
683 genes associated with growth and reproductive traits in the chicken. Poultry science. 88, 722-
684 727.
685 Pie, M.R., Alvares, L.E., 2006. Evolution of myostatin in vertebrates: is there evidence for positive
686 selection? Molecular phylogenetics and evolution. 41, 730-734.
687 Power, D.M., Louro, B., Houston, R., Anjos, L., Cardoso, J.C.R., 2011. Genomic-proteomic research
688 in Sparidae and its application to genetic improvement, Sparidae. Wiley-Blackwell, pp. 359-
689 381.
690 Reid, D.P., Szanto, A., Glebe, B., Danzmann, R.G., Ferguson, M.M., 2005. QTL for body weight
691 and condition factor in Atlantic salmon (*Salmo salar*): comparative analysis with rainbow
692 trout (*Oncorhynchus mykiss*) and Arctic charr (*Salvelinus alpinus*). Heredity. 94, 166-172.
693 Rodgers, B.D., Garikipati, D.K., 2008. Clinical, agricultural, and evolutionary biology of myostatin:
694 a comparative review. Endocrine reviews. 29, 513-534.
695 Rothschild, M.F., Hu, Z.L., Jiang, Z., 2007. Advances in QTL mapping in pigs. International journal
696 of biological sciences. 3, 192-197.
697 Rozen, S., Skaletsky, H., 2000. Primer3 on the WWW for general users and for biologist
698 programmers. Methods in Molecular Biology. 132, 365-386.
699 Saillant, E., Dupont-Nivet, M., Haffray, P., Chatain, B., 2006. Estimates of heritability and
700 genotype–environment interactions for body weight in sea bass (*Dicentrarchus labrax* L.)
701 raised under communal rearing conditions. Aquaculture. 254, 139-147.

Sanchez-Molano, E., Cerna, A., Toro, M.A., Bouza, C., Hermida, M., Pardo, B.G., Cabaleiro, S., Fernandez, J., Martinez, P., 2011. Detection of growth-related QTL in turbot (*Scophthalmus maximus*). BMC Genomics. 12, 473.

Sarropoulou, E., Nousdili, D., Magoulas, A., Kotoulas, G., 2008. Linking the genomes of nonmodel teleosts through comparative genomics. Mar Biotechnol (NY). 10, 227-233.

Sarropoulou, E., Franch, R., Louro, B., Power, D.M., Bargelloni, L., Magoulas, A., Senger, F., Tsalavouta, M., Patarnello, T., Galibert, F., Kotoulas, G., Geisler, R., 2007. A gene-based radiation hybrid map of the gilthead sea bream *Sparus aurata* refines and exploits conserved synteny with *Tetraodon nigroviridis*. BMC Genomics. 8, 44.

Sawatari, E., Seki, R., Adachi, T., Hashimoto, H., Uji, S., Wakamatsu, Y., Nakata, T., Kinoshita, M., 2010. Overexpression of the dominant-negative form of myostatin results in doubling of muscle-fiber number in transgenic medaka (*Oryzias latipes*). Comparative biochemistry and physiology. Part A, Molecular & integrative physiology. 155, 183-189.

Sellick, G.S., Pitchford, W.S., Morris, C.A., Cullen, N.G., Crawford, A.M., Raadsma, H.W., Bottema, C.D., 2007. Effect of myostatin F94L on carcass yield in cattle. Animal Genetics. 38, 440-446.

Shimizu, K.K., Purugganan, M.D., 2005. Evolutionary and ecological genomics of *Arabidopsis*. Plant Physiology. 138, 578-584.

Souche, E.L., Hellemans, B., Van Houdt, J.K., Volckaert, F.A., 2012. Characterisation and validation of single-nucleotide polymorphism markers in expressed sequence tags of European sea bass. Molecular Ecology Resources. in press.

Szklarczyk, D., Franceschini, A., Kuhn, M., Simonovic, M., Roth, A., Minguez, P., Doerks, T., Stark, M., Muller, J., Bork, P., Jensen, L.J., von Mering, C., 2011. The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. Nucleic acids research. 39, D561-568.

727 Tine, M., Kuhl, H., Gagnaire, P.-A., Louro, B., Desmarais, E., Martins, R.S.T., Hecht, J., Knaust, F.,
728 Belkhir, K., Klages, S., Dieterich, R., Stueber, K., Piferrer, F., Guinand, B., Bierne, N.,
729 Volckaert, F.A.M., Bargelloni, L., Power, D.M., Bonhomme, F., Canario, A.V.M., Reinhardt,
730 R., 2014. European sea bass genome and its variation provide insights into adaptation to
731 euryhalinity and speciation. Nature communications. 5, 5770.
732 Van Laere, A.S., Nguyen, M., Braunschweig, M., Nezer, C., Collette, C., Moreau, L., Archibald,
733 A.L., Haley, C.S., Buys, N., Tally, M., Andersson, G., Georges, M., Andersson, L., 2003. A
734 regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. Nature.
735 425, 832-836.
736 Vandeputte, M., Dupont Nivet, M., Haffray, P., Chavanne, H., Cenadelli, S., Parati, K., Vidal, M.-O.,
737 Vergnet, A., Chatain, B., 2009. Response to domestication and selection for growth in the
738 European sea bass (Dicentrarchus labrax) in separate and mixed tanks. Aquaculture. 286, 20-
739 27.
740 Visscher, P.M., Hopper, J.L., 2001. Power of regression and maximum likelihood methods to map
741 QTL from sib-pair and DZ twin data. Annals of human genetics. 65, 583-601.
742 Volckaert, F., Hellemans, B., Batargias, C., Louro, B., Massault, C., Van Houdt, J., Haley, C., de
743 Koning, D., Canario, A., 2012. Heritability of morphometric traits and cortisol response to
744 confinement stress in European sea bass *Dicentrarchus labrax*. Genetics Selection Evolution.
745 44.
746 Voorrips, R.E., 2002. MapChart: software for the graphical presentation of linkage maps and QTLs.
747 The Journal of heredity. 93, 77-78.
748 Walling, G.A., Visscher, P.M., Wilson, A.D., McTeir, B.L., Simm, G., Bishop, S.C., 2004. Mapping
749 of quantitative trait loci for growth and carcass traits in commercial sheep populations.
750 Journal of animal science. 82, 2234-2245.

751 Wiener, P., Smith, J.A., Lewis, A.M., Woolliams, J.A., Williams, J.L., 2002. Muscle-related traits in
752 cattle: The role of the myostatin gene in the South Devon breed. Genetics Selection
753 Evolution. 34, 221-232.
754 Wringe, B.F., Devlin, R.H., Ferguson, M.M., Moghadam, H.K., Sakhrani, D., Danzmann, R.G.,
755 2010. Growth-related quantitative trait loci in domestic and wild rainbow trout
756 (*Oncorhynchus mykiss*). BMC genetics. 11, 63.
757 Xie, C., Mao, X., Huang, J., Ding, Y., Wu, J., Dong, S., Kong, L., Gao, G., Li, C.Y., Wei, L., 2011.
758 KOBAS 2.0: a web server for annotation and identification of enriched pathways and
759 diseases. Nucleic acids research. 39, W316-322.
760
761
762

763 **Figures**

764 **Figure 1. Sea bass LG4, LG6 and LG15 QTL comparative map workflow.**

765 The integration of all comparative maps starting with sea bass linkage map *vs G. aculeatus* and *T.*
766 *nigroviridis* genomes, *vs* sea bass published QTL map, *vs* present sea bass QTL map and ending in
767 sea bass genome. The banner is the resumed objective across the comparative workflow.
768 Sea bass 2nd generation linkage map groups are represented named "2nd Gen", *G. aculeatus* and *T.*
769 *nigroviridis* chromossomes are named "Ga" and "Tn", respectively. Published and present sea bass
770 QTL linkage groups are represented by "qtl" and "QTL" respectively; and sea bass chromosomes are
771 named "Chr LG". All sea bass LG4, LG6 and LG15 linkage groups (cM) and *T. nigroviridis*, *G.*
772 *aculeatus* and sea bass chromosomes (Mbp) are at a 1:1 scale, with the exception of bar breaks,
773 which each represent a 10 cM distance. The three analytical methods are colour coded: red, green
774 and blue are paternal (PHS), maternal (MHS) half sib and sib pair regression analysis, respectively.
775 Confidence intervals are represented with also colour coded bars at right of the respective linkage
776 group bar. The traits are MORPH for morphology and BWfor body weight, full QTL nomenclature
777 and significance levels are described in table 2. All QTLs are 5% genome-wide, with exception of
778 both underlined QTLs in LG15 that are 5% chromosome-wide. Sea bass orthologous markers to *G.*
779 *aculeatus* and *T. nigroviridis* loci are linked with solid lines. For sake of clarity not all orthology was
780 represented but all orthologous loci have the same nomenclature.
781

782 **Figure 2. Genetic linkage map of significant QTLs for BW and morphometric (MORPH)**
783 **traits.**

784 All represented QTLs are at least 5% genome-wide. The three analytical methods are colour coded:
785 red, green and blue are paternal (PHS), maternal (MHS) half sib and sib pair regression analysis,
786 respectively. Confidence intervals are represented with also colour coded bars at right of the
787 respective linkage group bar. The QTL designation describes method (MHS, or PHS, or SP), and

number of detected traits up to eight (n/8). The full list of QTLs detected and respective details is in table 2 and additional file 3.

Figure 3. Synteny of sea bass chromosomes LG4, LG6, and LG15 with teleost genomes.

Exterior tiles represent collinear blocks of the overall degree of synteny between the sea bass (*Dicentrarchus labrax*) LG4, LG6, LG15 chromosomes and seven other publicly available teleost genomes. From the inner to the outer layer: *G. aculeatus, O. latipes, T. nigroviridis, D. rerio, O.niloticus, T. rubripes* and *G. morhua*. The respective colour code is species and chromosomal specific. The central ring represents the sea bass chromosomes scaled in Mbp. The interior tiles represent the genomic span corresponding to the confidence interval of the identified QTL for BW and morphometric traits. Light and dark green are maternal half sib MORPH and BW QTL, respectively; light and dark red are paternal half sib MORPH and BW QTL, respectively; and blue is sib pair MORPH QTL.

Figure 4. String confidence view network display.

A) The gene networks identified for the growth QTL in LG4, LG6, and LG15 were constructed separately and the largest networks are shown. B) The gene networks for the growth QTL in LG4, LG6, and LG15 in A) are integrated. Genes with stronger associations (edges) are represented by thicker lines.

Tables

Table 1. Linkage groups (LG) of European sea bass used for QTL analysis.

Number (#) of total and new markers per LG, and their average number of alleles (avg # alleles); average information content for maternal (MHS avg IC) and paternal (PHS avg IC) half sib method analysis; and average additive effect for sib pair method analysis (SP avg AE). Shaded in grey are the LG targeted for new markers.

LG	Length (cM)	# markers	New # markers	Avg # alleles	MHS avg IC	PHS avg IC	SP avg AE
LG1a	55	6	1	5.3	0.76	0.81	1.34
LG1b	11	6	1	5.2	0.95	0.97	1.82
LG2	40	9	2	4	0.83	0.87	1.43
LG3	14	4	-	3.5	0.81	0.83	1.20
LG4	58	8	3	4.1	0.84	0.81	1.38
LG5	46	6	-	4.2	0.91	0.88	1.58
LG6	24	11	6	4.7	0.87	0.92	1.60
LG7	34	4	-	6	0.84	0.89	1.52
LG8	43	7	-	4.3	0.88	0.85	1.50
LG9	-	1	-	-	-	-	-
LG10	36	4	-	4.3	0.80	0.83	1.40
LG11	5	2	-	3.5	0.67	0.86	1.36
LG12	14	4	-	5.5	0.90	0.92	1.68
LG13	15	3	-	5	0.88	0.92	1.65
LG14	60	6	-	3.8	0.77	0.79	1.23
LG15	18	6	2	3.5	0.93	0.90	1.65
LG17	17	4	-	4.8	0.92	0.90	1.67
LG18	-	1	-	-	-	-	-
LG19	46	4	1	4.3	0.73	0.85	1.22
LG20	47	5	2	3.8	0.70	0.55	0.89
LG23	9	2	-	4.5	0.92	0.88	1.56
LG24	6	2	-	7	0.97	0.93	1.79
LG25	45	10	9	3.9	0.60	0.66	0.90

Table 2. Linkage groups with significant QTL.

Linkage groups with F value higher 5% than genome wide (GW) permutation threshold. Morphometric traits SL, standard length; HL, head length; BL, body length; PRAL, pre anal length; AL, abdominal length; POAL, post anal length; HD, head depth; BD Body depth and trait body weight (BW). QTL in bold are significant and QTL values showed are for the most significant trait effect which is underlined. QTL shaded in grey correspond to significant QTL represented in Figure 1. Method: maternal (MHS) and paternal (PHS) half sibs and sib pairs (SP). Thresholds: suggestive if above 5% chromosome wide (CW) permutations, significant if above 5% genome wide (GW) permutations thresholds.

Traits	**Method**	**Position (cM)**	**Linkage group**	**F**	**5% CW**	**5% GW**	**QTL effect (%)**	**ΔConf (cM)**
BL	MHS	7	LG1b	3.62	**3.59**	6.99	3.86	0-11
SL, **HL**, PRAL, AL, HD, BD	PHS	9	LG1b	5.02	**2.44**	**4.24**	14.50	0-11
SL, **HL**, BL, PRAL, AL, HD, BD	SP	10	LG1b	21.5	**8.43**	**20.95**	-	0-11
BW	PHS	11	LG1b	3.52	**2.47**	4.60	8.82	0-11
BW	SP	11	LG1b	11.97	**8.63**	21.83	-	9-11
SL, BL, PRAL, **AL**, POAL, HD, BD	PHS	21	LG2	4.36	**2.74**	**4.31**	12.21	0-35
BW	PHS	33	LG2	3.84	**2.76**	4.60	9.88	9-39
SL, HL, PRAL, AL, POAL, HD, **BD**	MHS	1	LG4	7.15	**3.90**	**6.84**	8.96	0-17
SL, HL, **BL**, **PRAL**, AL, **POAL**, HD, **BD**	PHS	10	LG4	4.93	**2.93**	**4.30**	14.20	0-56
SL, **HL**, BL, PRAL, AL, POAL, HD, **BD**	SP	57	LG4	22.19	**8.38**	**20.42**	-	8-57
BW	MHS	1	LG4	4.53	**3.98**	7.13	4.95	0-46
BW	PHS	10	LG4	5.04	**2.90**	**4.60**	13.92	0-53
BW	SP	57	LG4	20.04	**9.02**	21.83	-	12-57
SL, **HL**, **BL**, **PRAL**, **AL**, **POAL**, HD, BD	MHS	15	LG6	20.4	**3.81**	**6.99**	27.00	8-24
SL, BL, AL, POAL	PHS	18	LG6	3.61	**2.58**	4.20	9.53	5-24
SL, HL, **BL**, **PRAL**, **AL**, **POAL**, **HD**, BD	SP	14	LG6	97.22	**5.42**	**20.03**	-	11-15
BW	MHS	15	LG6	9.76	**3.51**	**7.13**	12.07	7-23
BW	SP	14	LG6	7.48	**5.37**	20.42	-	11-24

SL, HL, **BL**, **PRAL**, **AL**, POAL, HD, BD	PHS	30	LG7	5.91	**2.65**	**4.31**	17.58	0-34
SL, HL, **BL**, PRAL, **AL**, POAL, HD, BD	SP	28	LG7	22.13	**5.16**	**21.40**	-	0-33
BW	MHS	9	LG7	4.46	**3.52**	7.13	4.86	0-34
BW	PHS	34	LG7	4.46	**2.63**	4.60	11.97	0-34
BW	SP	25	LG7	13.34	**5.17**	21.83	-	0-34
SL, **HL**, BL, **PRAL**, AL, **HD**, **BD**	MHS	4	LG15	7.87	**3.47**	**6.84**	9.98	0-17
SL, HL, BL, PRAL, AL, POAL, HD, BD	PHS	8	LG15	3.89	**2.63**	4.20	10.55	0-17
SL, BL, PRAL, AL, HD	SP	11	LG15	7.87	**5.13**	20.58	-	2-12
BW	MHS	3	LG15	4.06	**3.37**	7.13	4.06	0-16
BW	PHS	7	LG15	3.79	**2.71**	4.60	3.79	0-17
BW	SP	10	LG15	5.15	**4.71**	21.83	-	1-12
SL, HL, BL, PRAL, AL, POAL, HD, **BD**	MHS	46	LG20	7.39	**3.83**	**6.84**	9.30	0-46
SL, HL, PRAL, AL, HD, **BD**	SP	30	LG20	22.10	**6.10**	**20.42**	-	2-46
BW	MHS	46	LG20	4.92	**3.87**	7.13	5.49	0-46
BW	SP	37	LG20	10.78	**6.18**	21.83	-	3-46
SL, HL, **BL**, PRAL, **AL**, POAL, HD, BD	PHS	0	LG24	5.73	**2.50**	**4.30**	16.98	0-6
SL, HL, BL, PRAL, AL, POAL, HD, BD	SP	4	LG24	20.12	**4.38**	21.40	-	0-6
BW	PHS	0	LG24	3.69	**2.45**	4.60	9.40	0-6
BW	SP	4	LG24	17.66	**4.70**	21.83	-	5-6

825

826

827 **Table 3. Top 5 enriched KEGG orthology (KO) pathways in QTL confidence intervals in LG4,**

828 **LG6, and LG15 of European sea bass.**

829 Sample number, number of input genes mapped / number of total input genes; Background number,

830 number of total pathway genes / number of total genes). Corrected P-value after false discovery rate

831 (FDR) test.

Linkage group	KEGG reference pathway (KO) Term	KO ID	Sample number	Background number	Corrected P-Value
LG4	Amino sugar and nucleotide sugar metabolism	ko00520	6 / 179	48 / 6148	0.25
	VEGF signaling pathway	ko04370	7 / 179	73 / 6148	0.25
	Cholinergic synapse	ko04725	9 / 179	112 / 6148	0.25
	Non-small cell lung cancer	ko05223	5 / 179	54 / 6148	0.34
	Fructose and mannose metabolism	ko00051	4 / 179	36 / 6148	0.34
LG6	Axon guidance	ko04360	12 / 177	127 / 6148	0.04
	PPAR signaling pathway	ko03320	7 / 177	69 / 6148	0.23
	Cholinergic synapse	ko04725	9 / 177	112 / 6148	0.23
	Primary bile acid biosynthesis	ko00120	3 / 177	16 / 6148	0.23
	Focal adhesion	ko04510	12 / 177	199 / 6148	0.23
LG15	Protein digestion and absorption	ko04974	7 / 165	79 / 6148	0.15
	Galactose metabolism	ko00052	4 / 165	27 / 6148	0.15
	Mucin type O-Glycan biosynthesis	ko00512	4 / 165	30 / 6148	0.15
	Hedgehog signaling pathway	ko04340	5 / 165	56 / 6148	0.25
	ECM-receptor interaction	ko04512	6 / 165	85 / 6148	0.36

832

833

834 **Table 4. European sea bass genes identified within QTL confidence intervals**

835 Gene names and KO reference identified as members of the adipocytokine, MAPK, jak-STAT, and
836 insulin signalling KEGG orthology reference pathways in linkage groups LG4, LG6 and LG15.

Gene name	KO Id	Sea bass gene	LG	Adipocytokine	Jak-STAT	Insulin	MAPK
NR2B3	K08526	DLA_IV_004730	LG4	x			
		DLA_VI_001610	LG6				
CPT1	K08765	DLA_VI_001620	LG6	x			
		DLA_VI_002600	LG6				
NR1C1	K07294	DLA_VI_006190	LG6	x			
JAK2	K04447	DLA_IV_002540	LG4	x	x		
LEPR	K05062	DLA_IV_001550	LG4	x	x		
JAK1	K11217	DLA_IV_001520	LG4		x		
TPOR	K05082	DLA_IV_001330	LG4		x		
IL12RB2	K05064	DLA_IV_006120	LG4		x		
		DLA_IV_006130	LG4				
PIAS	K04706	DLA_IV_005420	LG4		x		
		DLA_VI_004200	LG6				
TPO	K00431	DLA_VI_004520	LG6		x		
PIK3R	K02649	DLA_IV_002230	LG4		x	x	
PIK3C	K00922	DLA_IV_000980	LG4		x	x	
SHC	K06279	DLA_IV_004570	LG4			x	
INSR	K04527	DLA_IV_003780	LG4			x	
PRKAR	K04739	DLA_VI_005350	LG6			x	
PDE3	K13296	DLA_VI_005300	LG6			x	
PPP1R3	K07189	DLA_VI_006300	LG6			x	
IRS	K07187	DLA_XV_004720	LG15			x	
GSK3B	K03083	DLA_XV_004610	LG15			x	
MKNK	K04372	DLA_IV_001840	LG4			x	x
		DLA_IV_004870	LG4				
PKA	K04345	DLA_IV_003220	LG4			x	x
MAP2K2	K04369	DLA_IV_005430	LG4			x	x
KRAS	K07827	DLA_VI_002050	LG6			x	x
MAP2K1	K04368	DLA_VI_004100	LG6			x	x
HRAS	K02833	DLA_VI_001720	LG6			x	x
JUND	K04449	DLA_IV_002180	LG4				x
FGF	K04358	DLA_IV_005620	LG4				x
CACNA1E	K04852	DLA_IV_004010	LG4				x
PLA2G	K01047	DLA_IV_001700	LG4				x
DUSP	K04459	DLA_IV_000480	LG4				x
		DLA_VI_001510	LG6				
MAPK8IP2	K04435	DLA_VI_002650	LG6				x
RASGRF1	K04349	DLA_VI_003460	LG6				x
CACNA2D1	K04858	DLA_VI_002850	LG6				x
PTP	K04458	DLA_VI_001520	LG6				x
FLNA	K04437	DLA_VI_005400	LG6				x
MAP3K13	K04422	DLA_XV_002600	LG15				x

MRAS	K07831	DLA_XV_001360	LG15	x
ZAK	K04424	DLA_XV_003630	LG15	x
ATF2	K04450	DLA_XV_003750	LG15	x

840 **Table 5. List of genes of European sea bass identified by STRING.**

841 List of genes from in the QTL confidence intervals from linkage groups LG4, LG6 and LG15,
842 identified by STRING confidence view network displayed in Figure 2.
843

Linkage group	Gene	protein
LG4	JAK1	janus kinase 1
	PLCG1	phospholipase C, gamma 1
	SHC2	SHC (Src homology 2 domain) transforming protein 2
	INSR	insulin receptor
	EPHB1	EPH receptor B1
	PIK3R2	phosphoinositide-3-kinase, regulatory subunit 2 (beta)
	LEPR	leptin receptor
	IGFBP1	insulin-like growth factor binding protein 1
	MPL	myeloproliferative leukemia virus oncogene
	JAK3	Janus kinase 3
	IRF4	interferon regulatory factor 4
	ZAP70	zeta-chain (TCR) associated protein kinase 70kDa
	PIK3CA	phosphoinositide-3-kinase, catalytic, alpha polypeptide
LG6	KCNQ1	potassium voltage-gated channel, KQT-like, member 1
	BHLHE41	basic helix-loop-helix family, member e41
	MCHR1	melanin-concentrating hormone receptor 1
	KRAS	v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog
	RORA	RAR-related orphan receptor A
	PPARA	peroxisome proliferator-activated receptor alpha
	CPT1A	carnitine palmitoyltransferase 1A (liver)
	ARNTL2	aryl hydrocarbon receptor nuclear translocator-like 2
	IGF1R	insulin-like growth factor 1 receptor
	MAP2K1	mitogen-activated protein kinase kinase 1
	HRAS	v-Ha-ras Harvey rat sarcoma viral oncogene homolog
	CPT1B	carnitine palmitoyltransferase 1B (muscle)
	LEP	leptin
	MORF4	mortality factor 4
	SMAD3	SMAD family member 3
	NR2C1	nuclear receptor subfamily 2, group C, member 1
	PNPLA2	patatin-like phospholipase domain containing 2
	IGF2	insulin-like growth factor 2 (somatomedin A)
	GNAI1	G protein, alpha inhibiting activity polypeptide 1
	RASSF7	ras association (RalGDS/AF-6) domain family member 7
	SIN3A	SIN3 homolog A, transcription regulator (yeast)
	MRPL23	mitochondrial ribosomal protein L23
	PTHLH	parathyroid hormone-like hormone
	RASGRF1	ras protein-specific guanine nucleotide-releasing factor 1
	CDKN1C	cyclin-dependent kinase inhibitor 1C (p57, Kip2)
LG15	COL4A1	collagen, type IV, alpha 1
	IGFBP2	insulin-like growth factor binding protein 2, 36kDa
	IGFBP5	insulin-like growth factor binding protein 5
	GSK3B	glycogen synthase kinase 3 beta

ARHGEF7	rho guanine nucleotide exchange factor (GEF) 7
CFLAR	CASP8 and FADD-like apoptosis regulator
FN1	fibronectin 1
COL4A2	collagen, type IV, alpha 2
EDNRB	endothelin receptor type B
ITGA6	integrin, alpha 6
CASP10	caspase 10, apoptosis-related cysteine peptidase
IRS2	insulin receptor substrate 2
APP	amyloid beta (A4) precursor protein
CASP8	caspase 8, apoptosis-related cysteine peptidase
LRP1B	low density lipoprotein-related protein 1B
MSTN	myostatin

844

845

Table 6. QTL population structure of European sea bass.

WTL population structure represents five full sib families, one dam mated with two sires and the other dam mated with three different sires, resulting in two dam half sib and five sire half sib families (Volckaert, et al., 2012).

Number of offspring	Sire1	Sire2	Sire3	Sire4	Sire5
Dam1	98	93			
Dam2			92	143	142
Dam half sib 1		191			
Dam half sib 2			377		

Figure 1

Comparative candidate genes | Target marker development | New QTL analysis | Validation via genomic mapping

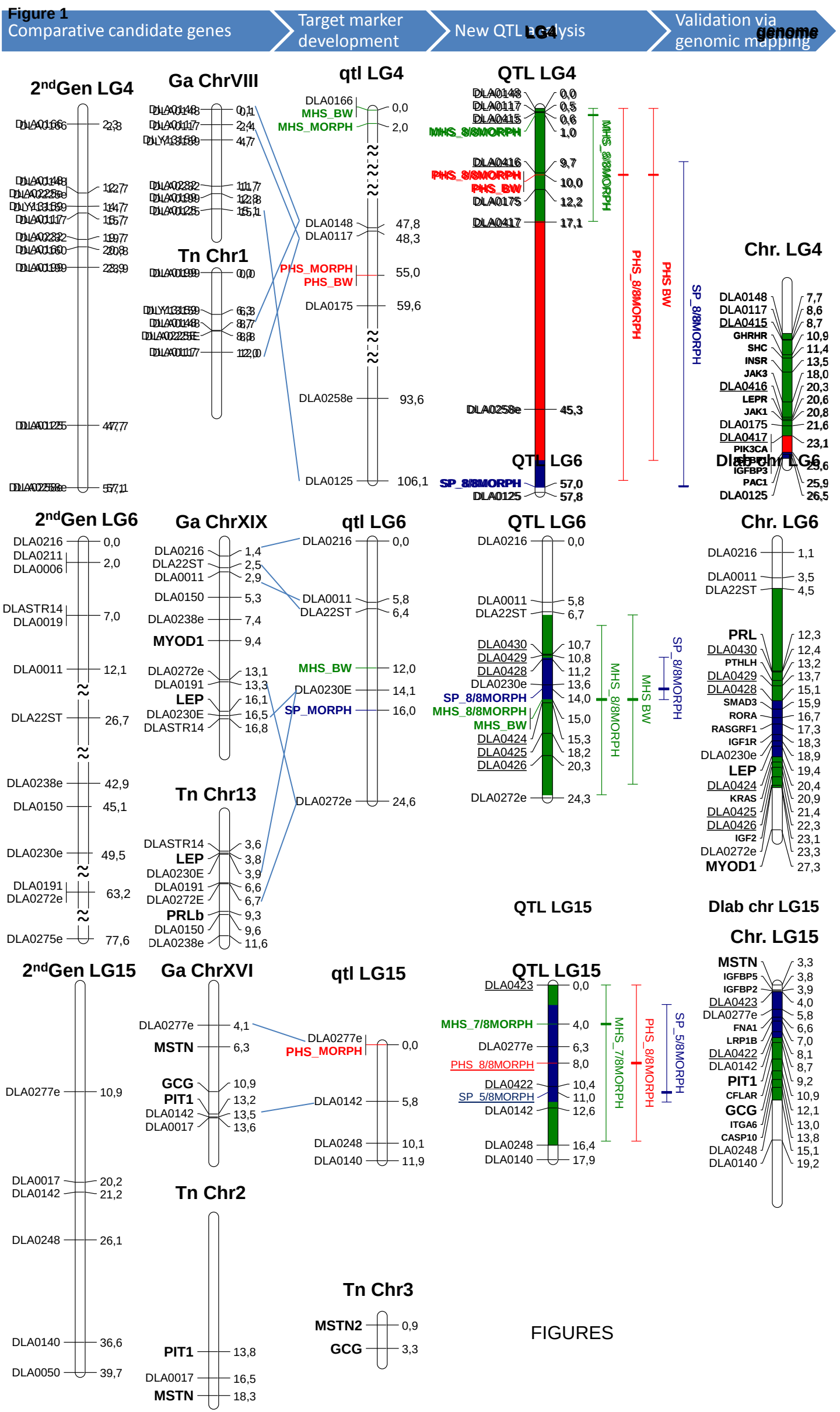


Figure 2

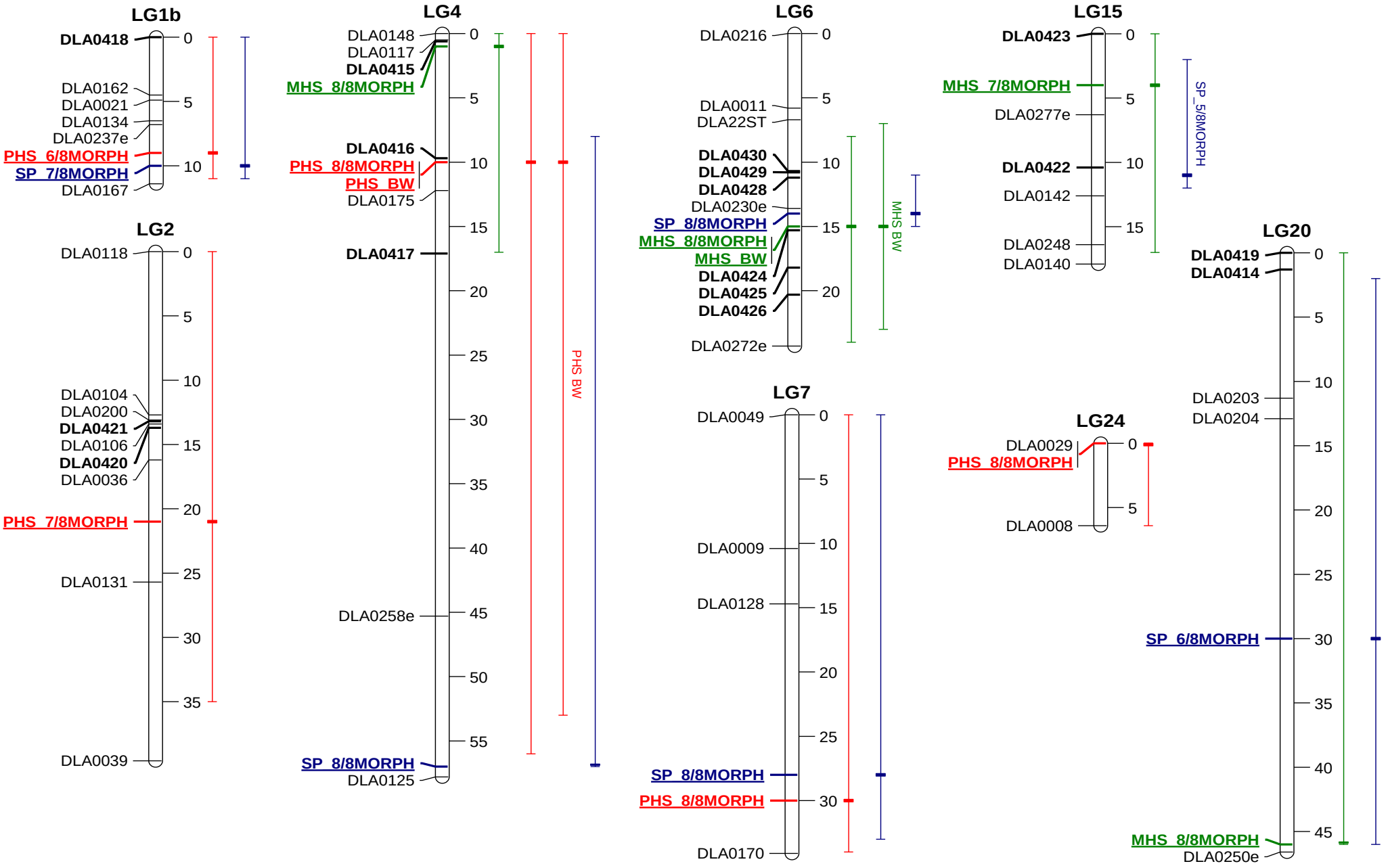

LG1b
DLA0418
DLA0162
DLA0021
DLA0134
DLA0237e
PHS_6/8MORPH
SP_7/8MORPH
DLA0167
LG2
DLA0118
DLA0104
DLA0200
DLA0421
DLA0106
DLA0420
DLA0036
PHS_7/8MORPH
DLA0131
DLA0039
LG4
DLA0148
DLA0117
DLA0415
MHS_8/8MORPH
DLA0416
PHS_8/8MORPH
PHS_BW
DLA0175
DLA0417
DLA0258e
SP_8/8MORPH
DLA0125
PHS BW
LG6
DLA0216
DLA0011
DLA22ST
DLA0430
DLA0429
DLA0428
DLA0230e
SP_8/8MORPH
MHS_8/8MORPH
MHS_BW
DLA0424
DLA0425
DLA0426
DLA0272e
MHS BW
LG7
DLA0049
DLA0009
DLA0128
SP_8/8MORPH
PHS_8/8MORPH
DLA0170
LG15
DLA0423
MHS_7/8MORPH
DLA0277e
DLA0422
DLA0142
DLA0248
DLA0140
SP_5/8MORPH
LG24
DLA0029
PHS_8/8MORPH
DLA0008
LG20
DLA0419
DLA0414
DLA0203
DLA0204
SP_6/8MORPH
MHS_8/8MORPH
DLA0250e

Figure 3

Click here to download high resolution image

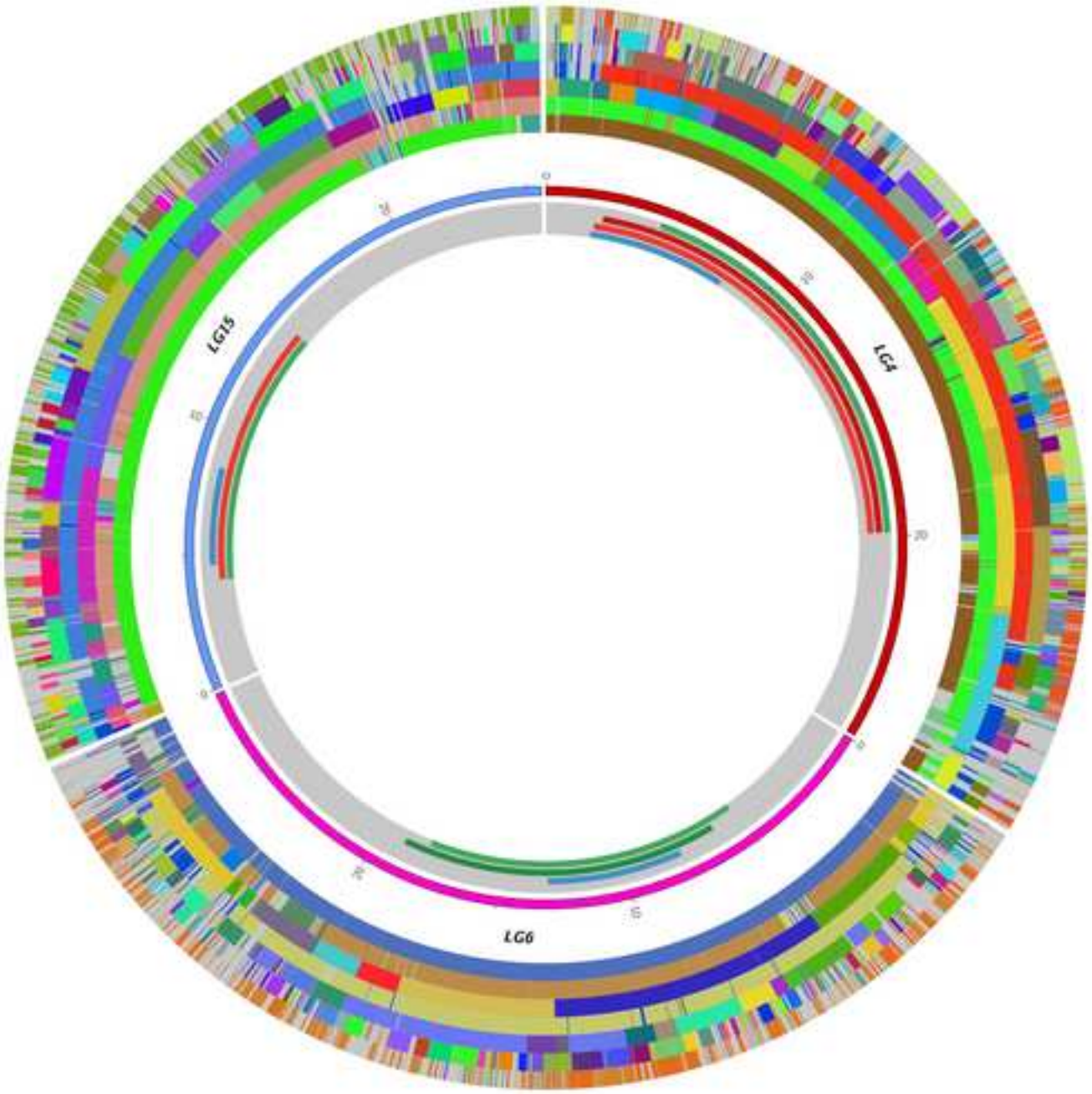


Figure 4a

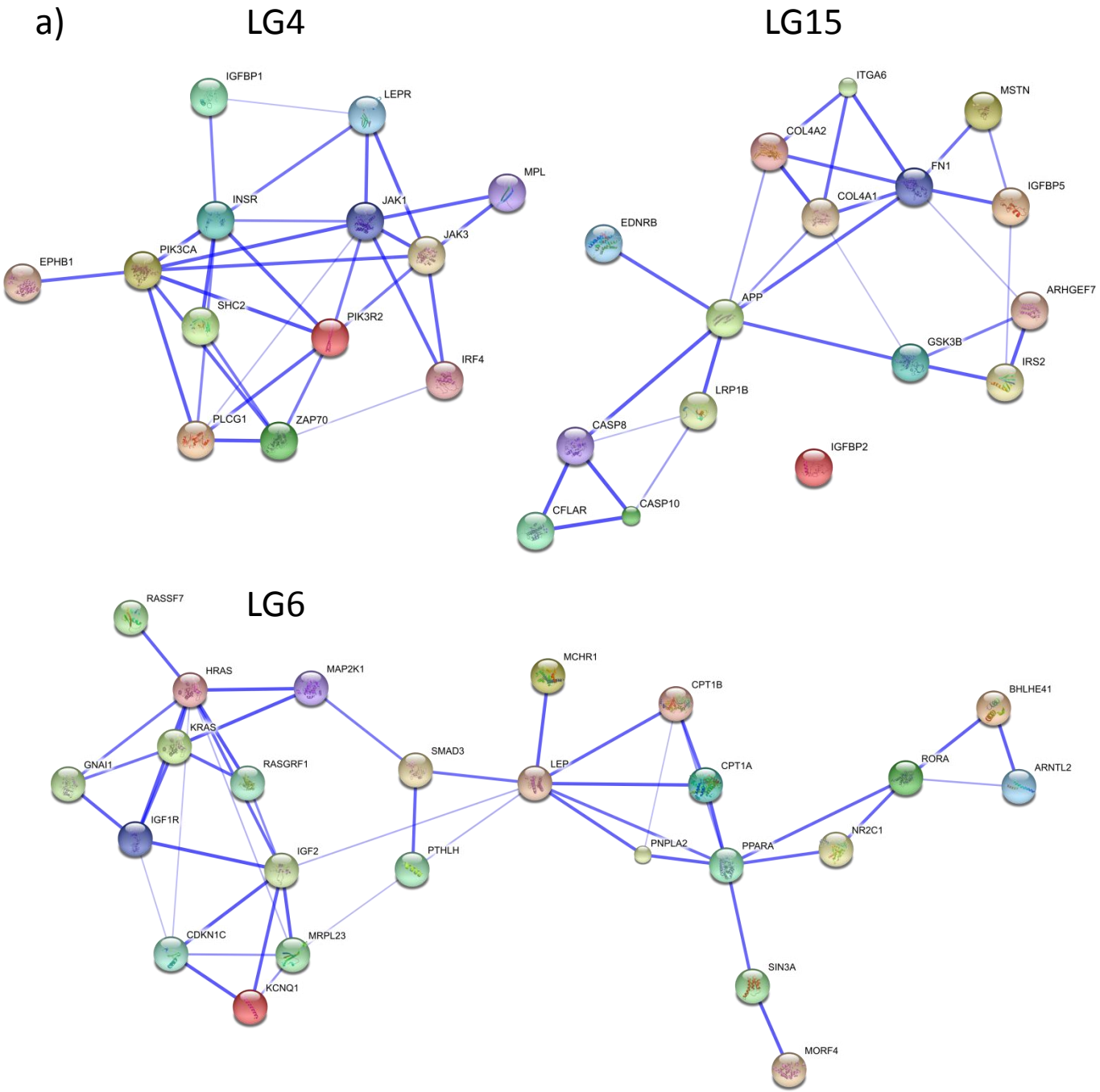
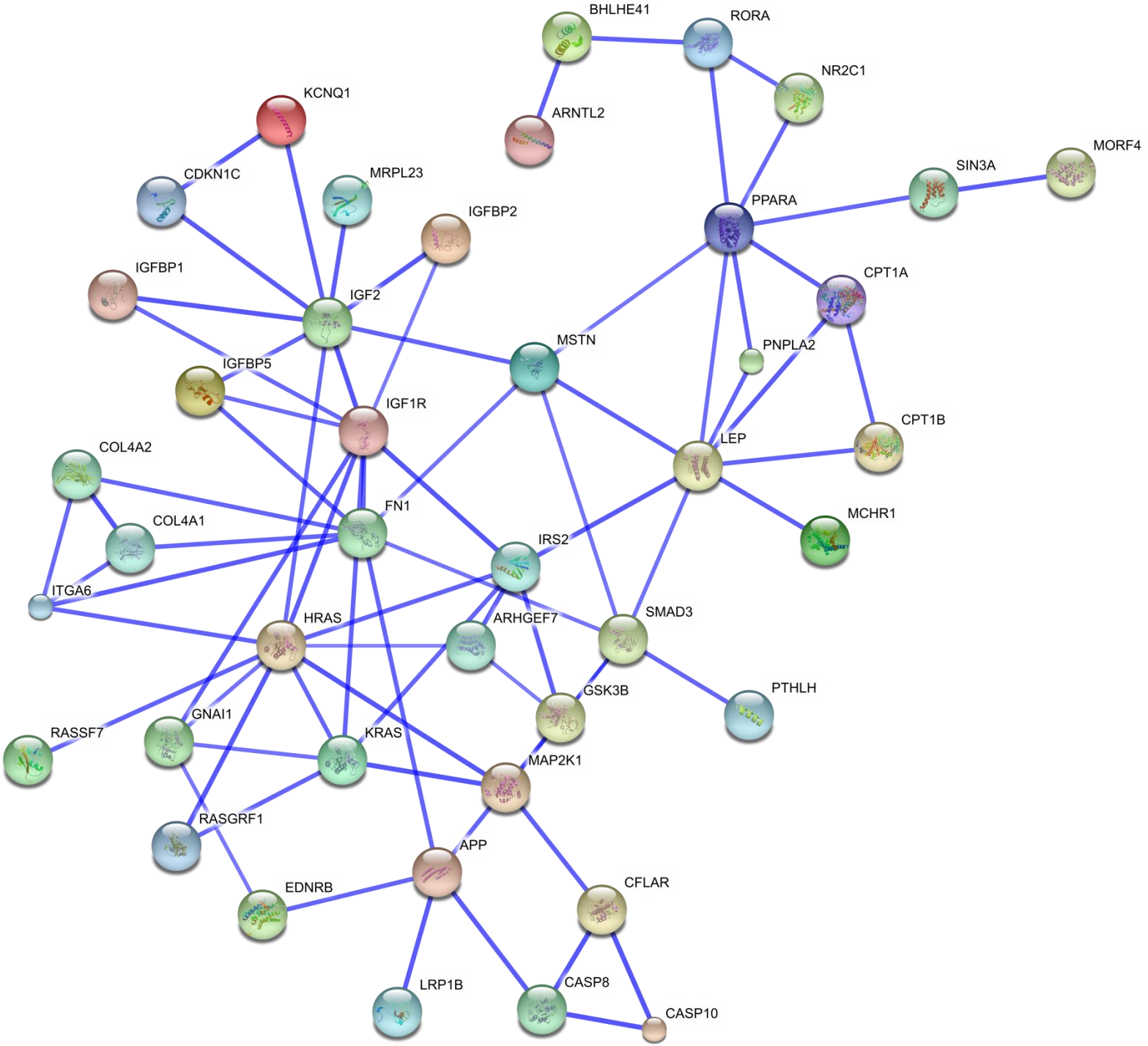

a)
LG4
IGFBP1
LEPR
MPL
INSR
JAK1
JAK3
EPHB1
PIK3CA
SHC2
PIK3R2
IRF4
PLCG1
ZAP70
LG15
ITGA6
MSTN
COL4A2
FN1
IGFBP5
COL4A1
EDNRB
ARHGEF7
APP
GSK3B
IRS2
LRP1B
CASP8
IGFBP2
CFLAR
CASP10
LG6
RASSF7
HRAS
MAP2K1
MCHR1
CPT1B
BHLHE41
KRAS
SMAD3
LEP
CPT1A
RORA
ARNTL2
GNAI1
RASGRF1
IGF1R
PNPLA2
PPARA
NR2C1
IGF2
PTHLH
CDKN1C
MRPL23
SIN3A
KCNQ1
MORF4

Figure 4b

b)


BHLHE41
RORA
NR2C1
KCNQ1
ARNTL2
MORF4
SIN3A
CDKN1C
MRPL23
PPARA
IGFBP2
IGFBP1
CPT1A
IGF2
MSTN
PNPLA2
IGFBP5
IGF1R
CPT1B
COL4A2
LEP
FN1
MCHR1
COL4A1
IRS2
ITGA6
HRAS
ARHGEF7
SMAD3
PTHLH
GSK3B
GNAI1
RASSF7
KRAS
MAP2K1
RASGRF1
APP
EDNRB
CFLAR
LRP1B
CASP8
CASP10