The rainbow trout genome, an important landmark for aquaculture and genome evolution

2

Julien Bobe, Lucie Marandel**, Stéphane Panserat**, Pierre Boudinot[†], Camille Berthelot[‡,††], Edwige Quillet[§], Jean-Nicolas Volff[¶], Carine Genêt[§], Olivier Jaillon[††], Hugues Roest Crollius[‡‡], Yann Guiguen**

*INRA, Fish Physiology and Genomics, Rennes, France; **Nutrition, metabolism and aquaculture (NUMEA), INRA, Univ Pau & Pays Adour, Saint Pée sur Nivelle, France; [†]Molecular Virology and Immunology, INRA, Université Paris-Saclay, Jouy-en-Josas, France; [‡]European Molecular Biology Laboratory, European Bioinformatics Institute, Welcome Trust Genome Campus, Cambridge, United Kingdom; [§]Animal Genetics and Integrative Biology, INRA, AgroParisTech, Université Paris-Saclay, Jouy-en-Josas, France; [¶]Institute of Functional Genomics of Lyon (IGFL), Ecole Normale Supérieure de Lyon, CNRS, University of Lyon I, Lyon, France; [††]CEA-Genomic Institute, French National Sequencing Center, Evry Cedex, France; [‡‡]Ecole Normale Supérieure (ENS), ENS Biology Institute, IBENS, Paris, France

Rainbow trout

Rainbow trout and Salmonids commercial importance

Members of the salmonid family (Fig. 2.1) are present worldwide and many of them are species of major importance for aquaculture, wild stock fisheries or recreational sport fisheries. Their world aquaculture production was estimated over 3,000,000 metric tons (MT) with a global commercial value over US$ 17 billion in 2013, with Atlantic salmon being the most important farmed salmonid species with world aquaculture production accounting for 2,000,000 MT. Rainbow trout is a species native to western North America, but is considered as one of the most widely introduced fishes. According to FishBase (http://www.fishbase.org), rainbow trout is now recorded to be present in at least 70 different countries with a distribution in nearly all of the world's continents. Although its farming receives less attention than Atlantic salmon, the production data for rainbow trout are still quite impressive. Total world production reached 500,000 MT for the first time in 2001 and was recorded in 2013 as being over 800,000 MT with a total export value around US$ 3.5 billion. More than one third of this production currently comes from Asia (310,000 MT) with a significant production both in Iran (140,000 MT) and Turkey (130,000 MT), the remaining production being produced in America (220,000 MT) and Europe (280,000 MT). In Europe, Norway is the major producer of rainbow trout with 20% of the total European production

Genomics in Aquaculture. http://dx.doi.org/10.1016/B978-0-12-801418-9.00002-0

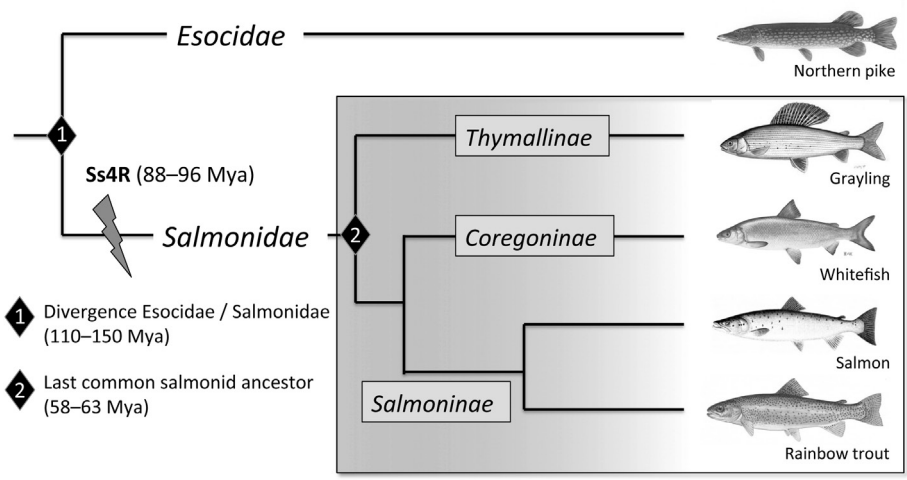

Figure 2.1 Salmonids evolution. The rainbow trout, *Oncorhynchus mykiss*, belongs to the salmonid (salmonidae) family that contains three sub-families, that is, the *Salmoninae* (mostly trouts, chars, and salmons), the *Thymallinae*, (graylings) and the *Coregoninae* (whitefish and ciscos). Salmonids diverged from their closest sister group, that is, *Esocidae* (pikes and mudminnows) 110–150 Mya (Near et al., 2012) and the last common ancestor of extant salmonids is dated around 58–63 Mya (Crête-Lafrenière et al., 2012). The specific whole genome duplication of salmonids (Ss4R) is dated between 88 Mya (Macqueen and Johnston, 2014) and 96 Mya (Berthelot et al., 2014).

(70,000 MT), followed by Italy (34,000 MT), France (29,000 MT), Russia (24,000), Spain (15,000 MT), United Kingdom (12,000 MT), and Finland (12,000 MT). More than 20 additional European countries share the remaining production. The majority of American production is localized in Chile (140,000 MT) and the production in the United States of America although rather modest (26,000 MT) was still worth $90 million in annual sales in 2013, and this does not include the industry aimed at restoration, conservation and recreation.

Rainbow trout: a freshwater laboratory rat

Beside its economic importance, rainbow trout has also been extensively used as a model of investigation in disciplines as diverse as cancer research, toxicology, comparative immunology, disease ecology, physiology, nutrition, and developmental and evolutionary biology (Thorgaard et al., 2002). *O. mykiss* is one the most studied fish species, with many experimental advantages including the availability of multiple natural populations (Behnke et al., 2002) and clonal lines (Komen and Thorgaard, 2007; Quillet et al., 2007), an easy breeding and the possibility to perform transgenesis (Thorgaard et al., 2002) and gene inactivation (Yano et al., 2014). Its relatively large body size compared to model fish like zebrafish or medaka makes rainbow trout a

well suited alternative model to carry out biochemical and molecular studies on specific tissues or cells that are much more difficult to analyze in smaller fish models (Thorgaard et al., 2002). In agreement with its major relevance to biomedical research and aquaculture, many genomic resources have been developed during the last two decades, including genetic and physical maps, BAC libraries and BAC end sequences, as well as numerous expressed sequence tags (see later sections for a synopsis of the genomic resources available in rainbow trout). Most of the needed presequencing phase genomic resources were then available when the first *O. mykiss* whole genome-sequencing project was initiated in 2010.

Rainbow trout life cycle

Rainbow trout is a native species from the Northwest cost of North America where its natural distribution ranges from Alaska up to the north of Mexico (Behnke et al., 2002). It is mostly a freshwater species that inhabits lakes and rivers although there are populations, called steelhead trout that live in the ocean and have an anadromous migration back to freshwater to spawn. Freshwater populations usually inhabit well-oxygenated rivers or lakes with preferundum temperatures ranging from 10 to 16°C. Reproduction is generally taking place during spring with spawning occurring in cold well-oxygenated shallow rivers with fine gravel bottoms. One single female rainbow trout can produce 1–8 thousands eggs (3.5–6 mm) which are laid in one or multiple batches and immediately fertilized by sperm from a single male. Eggs fall between gravel spaces and usually hatch 4–5 weeks after fertilization. After vitellus resorption young fry emerge from the gravel substrate and start to feed essentially on zoobenthos and zooplankton prey. At the adult stage rainbow trout is an opportunistic feeder that relies on many different type of food ranging from small insects to crayfish.

Genomic aspects

The genome of rainbow trout per se is extremely interesting from an evolutionary genomics point of view. First of all, its genome is of respectable size, that is, 2.4×10^9 bp (Young et al., 1998), that is, 6 times bigger than the genome of the pufferfish (Roest Crollius et al., 2000), and even larger than that of the zebrafish, that is, 1.7×10^9 bp (Howe et al., 2013). Second, the genome of the rainbow trout is complex, with a chromosome arm number (NF) equal to 104 and a variable diploid set of chromosomes ranging from 58 to 64 due to Robertsonian rearrangements (Thorgaard, 1976). One major reason for this complexity is an event of tetraploidization (whole genome duplication, WGD), which occurred in the salmonid lineage (the Ss4R: Salmonid-specific fourth round of WGD). The genomes of salmonids are then assumed to still be undergoing a process of rediploidization and this involves a differential evolution or loss of gene duplicates, a process called "divergent resolution" (Taylor et al., 2001) that has been proposed to reduce hybrid fitness and favor speciation. Such a phenomenon might have played an important role in the amazing richness of species observed in fish (Volff, 2005). Indeed, divergent resolution has been already reported in teleosts

subsequently to an ancestral genome duplication (third WGD or Teleost-specific third round of WGD = Ts3R) that took place between 225 and 330 million years ago (Mya) (Hurley et al., 2007; Near et al., 2012; Santini et al., 2009), but only very divergent species have been compared. Salmonid genomes provide a unique opportunity to analyze the differential evolution of tetraploidized genomes that diverged more recently (Allendorf and Thorgaard, 1984) and to assess the differential evolution of duplicated genes during early steps of rediploidization.

The rainbow trout genome: diversity, structure, organization

Salmonids combine both the essential characteristics of being very important commercial species and also species of considerable scientific importance with a key position in the ray-finned fish evolutionary tree (Fig. 2.1). These combined applied and scientific interests have deeply stimulated a well-structured international scientific community in developing a large number of genomics tools that are now available in many salmonid species.

Genomic resources preexisting before the first genome sequence in rainbow trout

The size of the rainbow trout genome is about 2.4×10^9 bp (Young et al., 1998) and its G + C content is 42% (Bernardi and Bernardi, 1990). One of the problems that plagued the initial sequencing and assembly of the zebrafish genome was the extent of polymorphism within and between the individuals of the pooled group that was used as the source of the DNA for the project. This problem has subsequently been overcome by using a single, doubled haploid fish that was produced by androgenesis. In rainbow trout, doubled haploid individuals were successfully obtained using gynogenesis or androgenesis (Chourrout, 1984; Parsons and Thorgaard, 1985; Quillet et al., 1991), and used to produce homozygous clonal lines (Scheerer et al., 1991; Quillet et al., 2007) that have been extensively characterized at genetic and phenotypic levels. Fully homozygous individuals were particularly desirable to develop genomic tools in rainbow trout, to deal with the high genome complexity that was anticipated because of the ancestral WGD in the species. The genomic DNA from a single doubled haploid fish (Swanson YY doubled haploid clonal line) was used to produce several rainbow trout bacterial artificial chromosomes (BAC) libraries. These BAC libraries have been fingerprinted to produce a first generation (Palti et al., 2011) and a second generation BAC physical map (Palti et al., 2012). Several linkage maps based on doubled haploid individuals and outbred populations have been constructed for rainbow trout using microsatellites and SNPs markers (Guyomard et al., 2012, 2006; Nichols et al., 2003; Rexroad et al., 2008; Sakamoto et al., 2000). A synthetic linkage map compiling previous ones was built (Guyomard et al., 2012) that comprised of 2,226 markers, with a medium density of 1 marker/cM.

In parallel, many different tissues and developmental stages cDNA libraries have been constructed and sequenced in the rainbow trout (Govoroun et al., 2006; Koop et al., 2008; Rexroad et al., 2003). As of Jun. 2015, 290,406 expressed sequence tags (ESTs) sequences are publicly available in NCBI GenBank and these ESTs have been the basis for several cDNA and oligos microarray platforms in rainbow trout (Canario et al., 2008). Sequences from both ends of 96,000 trout BACs were also available and these BAC-end sequences have been used to create a first trout repeat database (Genet et al., 2011), showing that repetitive DNA accounts for approximately 58% of the genome. These BAC-end sequences were screened for microsatellites and were used to produce the first and a second generation integrated maps (Palti et al., 2012, 2011).

Rainbow trout genome sequencing

The strategy developed for sequencing the rainbow trout genome was based on the NGS sequencing technologies available in 2010 at the beginning of this genome project. This genome was sequenced using a whole genome shotgun strategy (Berthelot et al., 2014) with genomic DNA from a unique doubled haploid (YY) rainbow trout male (Parsons and Thorgaard, 1985; Young et al., 1996). The availability of DNA from a totally homozygous individual has been a key resource that deeply facilitated the assembly of the rainbow trout genome. For the whole shotgun approach, single read, as well as 8 kb, 12 kb, and 20 kb mate-pairs libraries were sequenced using the 454 titanium technology up to a 20-fold sequence coverage with 454 Titanium and a 70-fold coverage using Solexa-Illumina sequences. Solexa-Illumina sequences were used to correct the contigs and scaffolds sequences resulting from the assembly of the 454 Titanium reads that are known to be error-rich especially in stretches of homopolymers (Gilles et al., 2011). This strategy was also complemented by a deep coverage of BAC-end sequences (Genet et al., 2011) that helped to scaffold contigs and to anchor the sequence onto the integrated maps. The total size of the resulting assembly of this first version of the rainbow trout genome (Berthelot et al., 2014) was 1.9 Gb with a scaffold N50 of 384 kb (half of the assembly is contained in 1,014 scaffolds longer than 384 kb). Using linkage and physical map information, scaffolds were anchored onto chromosomes at 898 distinct loci. Annotation of the genome sequence identified 46,585 protein-coding gene models with supporting protein evidence from other vertebrates and transcript evidence from 15 tissues of a doubled haploid rainbow trout adult obtained by RNA-seq. This high number of predicted protein-coding genes compared to other vertebrates is compatible with the recent Ss4R. 495 microRNA (miRNA) loci were also identified corresponding to 84 different families and 164 mature sequences. Transposable elements account for about 38% of the rainbow trout genome sequence.

Rainbow trout genome structure

The ancestral genome of teleost fish underwent a WGD event, termed here the teleost-specific third WGD (Ts3R) (Hurley et al., 2007; Near et al., 2012; Santini et al., 2009), approximately 300 million years ago (Mya). The signature of this Ts3R is still present in modern teleost genomes and comes in addition to two more ancestral WGD events

common to all bony vertebrates (vertebrate genome duplications 1 and 2, VGD1 and VGD2) (Dehal and Boore, 2005). While WGD events are rare within animal lineages, they represent important evolutionary landmarks from which some major lineages have diversified. Duplicated genomes eventually retain only a small proportion of duplicated genes, while seemingly redundant copies are inactivated in a poorly understood process termed gene fractionation (Langham et al., 2004). The genome structure of the rainbow trout is of particular interest in this context because salmonids have undergone an additional relatively recent WGD event (the salmonid-specific fourth WGD or Ss4R) that has been initially dated between 25 and 100 Mya (Allendorf and Thorgaard, 1984). Gene fractionation may thus be ongoing in that species, providing a unique opportunity to better understand the main determinants that influence this rediploidization process.

By reconstructing the Ss4R paralogous regions, the genome of rainbow trout allowed the reconstruction of the ancestral karyotype of salmonids before the Ss4R duplication. The modern rainbow trout genome is organized in 38 major pairs of duplicated regions, and 14 out of 30 chromosomes result from the fusion of two different post-Ss4R chromosomes, in agreement with the known paralogies inferred from the trout linkage maps (Guyomard et al., 2012). Other chromosomes are more complex mosaics of different post-Ss4R chromosomes, reflecting additional inter-chromosomal rearrangements that have occurred since the Ss4R event (Fig. 2.2). Numerous ohnologs (i.e., paralogous genes formed by a WGD event) were identified in these Ss4R paralogous regions. The computed distribution of silent nucleotide substitutions (dS) among pairs of Ss4R ohnologs identified in rainbow trout, and between Atlantic salmon and rainbow trout pairs of orthologs (representative of the speciation divergence time, i.e., around 30 Mya, between these two species) allowed to estimate the date of the Ss4R at 96 Mya, in the upper range of the 25–100 Mya from previous estimation (Allendorf and Thorgaard, 1984) but in the same range (88 Mya) as another recent estimation (Macqueen and Johnston, 2014). These results contrast with the age of the Salmonidae family that has been estimated 50–60 Mya (Crête-Lafrenière et al., 2012), suggesting that the Ss4R occurred long before (> 30 Mya) the last common ancestor of extant salmonids (Fig. 2.1). This is consistent with the WGD Radiation Lag-Time Model (Schranz et al., 2012) that has been proposed in plants in which significant lag-times are proposed to be needed between WGDs and the subsequent adaptive radiations that are often associated with these WGD events.

Within a few million years after a WGD, genome evolution is thought to involve the loss of one gene copy of most ohnologous gene pairs by gene fractionation. This process has never been documented at the whole genome scale in any vertebrate because all WGD studied to date are too ancient to capture such information. The rainbow trout genome enabled that analysis as high confidence paralogous regions can be easily characterized due to the more recent age of the Ss4R. Across the entire genome, around 50% of the Ss4R duplicated gene pairs have undergone gene fractionation and returned to a single copy state, while the remaining retained both ohnologs. Gene fractionation is thus a relatively slow process in the trout genome, since genes were inactivated at an average rate of approximately 170 genes per million years. Additionally, the majority of singletons can still be paired with clear paralogous sequences

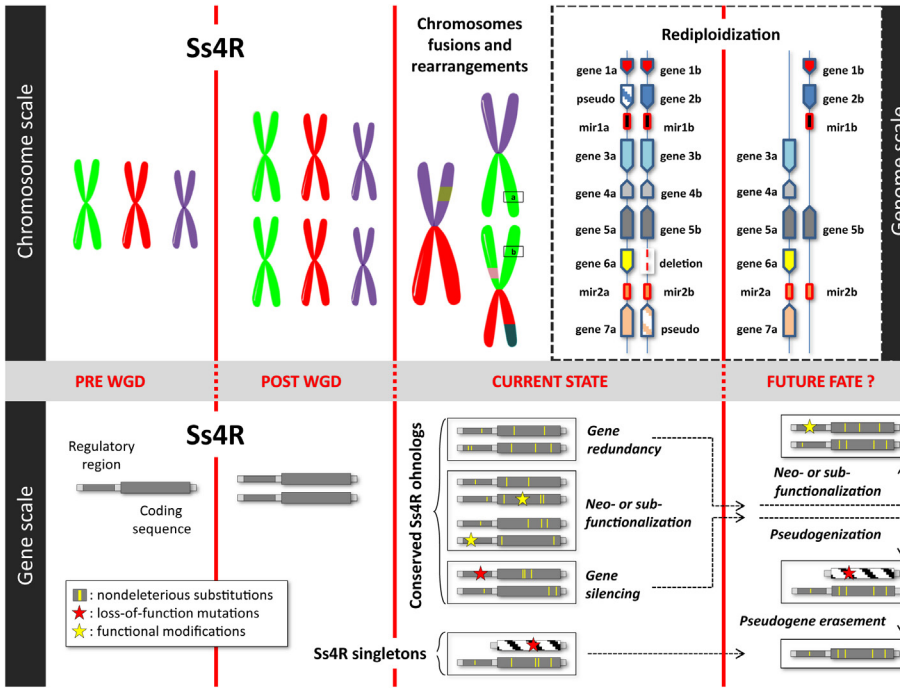

Figure 2.2 Schematic evolutionary history of the rainbow trout genome during the Ss4R WGD event. The Ss4R WGD resulted in a complete doubling of the genetic material, so every chromosome (and gene, associated with its regulatory regions) was present in two identical copies right after the Ss4R. The Ss4R WGD was followed by chromosome fusions and a few intrachromosomal rearrangements, bringing back the total number of chromosomes to 30, close to the preduplication karyotype. Rediploidization or gene fractionation, that is, the loss of duplicated genes after WGD by mutations and deletions to return to a mostly diploid state, is an ongoing process in the rainbow trout genome. At present, the duplicated regions within the genome (as exemplified by regions a and b in the boxed figure) are still highly collinear and retain both ohnologs for around 50% of the duplicated gene pairs. Ohnologs that are still present in both copies are slowly diverging at the sequence level, and may in the future acquire new or ancillary functions (neo- or sub-functionalization) or be lost via gene fractionation. Based on observations on older WGD events, only about 20–25% of duplicated genes are expected to be eventually retained in two copies (Howe et al., 2013).

stemming from the Ss4R, although the latter are largely nonfunctional (pseudogenes). In contrast to this 50% retention rate of ohnologous protein-coding genes, we found that the post-Ss4R conservation of genes encoding microRNAs is nearly complete, with almost all microRNA ohnologs being conserved as duplicated Ss4R copies. This higher conservation does not seem to be simply the outcome of the shorter length of sequences coding for miRNA and might result from more subtle selective processes. Altogether, the high retention rate of ohnologous copies and of pseudogenes suggests that the fractionation process is largely incomplete and still ongoing in the trout

genome. At the genome organization level, the analysis of the Ss4R duplicated regions reveals a high colinearity between paralogous genomic sequences, consistent with a conserved order of ohnologs and no strong evidence of a clustering of singletons versus ohnologs throughout the genome. Together these observations show that gene fractionation does not involve many genomic rearrangements such as inversions or translocations that would modify the order of genes in the genome, or large deletions that would result in long clusters of singletons. The nucleotide sequence identity between paralogous genomic regions is still high (86%) and Ss4R ohnologous protein-coding sequences and microRNAs are also highly conserved at the sequence level with 92.9% amino-acid identity and 96.4% nucleotide identity, respectively. In addition, the identity between the Ss4R protein-coding singletons and their corresponding pseudogenes remains high (average amino acid identity 79.0%) suggesting that most of these gene inactivations took place recently.

A genome for genetic and functional investigations

Deciphering the genetic architecture of traits

Quantitative trait loci (QTLs) are genomic regions (loci) associated to the phenotypic variation of a trait. Typically, a QTL is linked to, or contains, gene(s) that control the target trait. Therefore, using the polymorphism (molecular or sequence information) associated to the QTL regions is a way to improve the efficiency of selective breeding by targeting the key regions that govern the variability of the target trait. Also, the identification of QTLs constitutes a step toward the molecular dissection of complex traits, the discovery of the genes that control functions of interest (causative genes) and their regulation.

Using the medium density genetic maps previously available (see earlier) and linkage association methodologies (LA) in family QTL designs, a number of QTL have been detected in rainbow trout for many traits, with a special focus on adaptation and disease resistance traits (Baerwald et al., 2011; Le-Bras et al., 2011; Hecht et al., 2012; Verrier et al., 2013; Liu et al., 2014; Quillet et al., 2014; Vallejo et al., 2014). However, the confidence intervals of QTL positions are usually wide, making their exploitation difficult.

Together with the rapid evolution of NGS technologies, high-density genotyping methods are now available. For instance, the restriction-site associated DNA tags (RAD-tags) sequencing (Etter et al., 2011) enables cost effective identification of thousands SNP. In rainbow trout, RAD-tags based linkage maps were produced resulting in an increased marker density (Miller et al., 2012). RAD-sequencing was also used in an extended SNP discovery study where nearly 145,000 SNPs were identified (Palti et al., 2014) allowing the development of high-density genotyping chips like the 57 K SNP chip recently developed for rainbow trout (Palti et al., 2015) that is now commercially available. In such studies, having a reference genome sequence as a template is of great interest to order the newly discovered SNPs into maps or build local haplotypes. These high-density genotyping tools allow a radical

change in QTL mapping methods and enable the switch to methods exploiting linkage disequilibrium (LD) and association analyses at the population level (LDLA, Genome Wide Association Study or GWAS) (Liu et al, 2015). This opens new areas for deciphering complex traits, and for the identification of genomic regions under selection and the future implementation of genomic selection for aquaculture breeding.

Functional insights in Rainbow trout nutrition

Aquafeeds have been developed over the past thirty years for rainbow trout in aquaculture based on the well-known nutrient requirements identified for this species. In the context of the development of sustainable aquaculture, it is imperative to develop new aquafeeds in order to replace dietary fish meal and fish oil (originating from threatened marine resources) by either plant ingredients or other ingredients (Naylor et al., 2009). This evolution of dietary composition necessitates an extensive knowledge of the intermediary metabolism mainly based on molecular approaches (Panserat et al., 2009; Panserat and Kaushik, 2010).

Selected examples of nutritional regulation of transcript levels

Nutrition and flesh quality: molecular LC-PUFA biosynthesis pathway is functional

One of the most important challenges after suppression of fish oil in aquafeed is to maintain high level of long-chain polyunsaturated fatty acids (LC-PUFAs, i.e., the docosahexaenoic acid—DHA—and the eicosapentaenoic acid—EPA) in the flesh (Monroig et al., 2013). Numerous studies in rainbow trout demonstrated that this species has the molecular capacities to produce these LC-PUFAs from precursors found in vegetable oils. Indeed, higher levels of mRNAs coding for desaturase and elongase enzymes were observed following suppression of fish oil in the diets (Seiliez et al., 2011). However, this induction is not sufficient to obtain the same level of muscle LC-PUFAs compared to fish fed with fish oil. LC-PUFA biosynthesis in rainbow trout seems to substrate limited, even when some fish lines presented higher levels of fatty biosynthesis molecular capacity (Kamalam et al., 2012).

Functional but atypically regulated glucose metabolism

Rainbow trout is well known to be a poor user of dietary carbohydrates, limiting the incorporation of starch (main component of rich-carbohydrate plants) in aquafeeds (Polakof et al., 2012). Recent studies prove that rainbow trout has the inducible capacities to catabolize dietary glucose as suggested by levels of transcripts for glucose transporter in muscle (Diaz et al., 2009), for glucokinase enzyme in liver (Panserat et al., 2014) and for miRNAs involved in regulation of insulin signaling and metabolism (Mennigen et al., 2014, 2012) following carbohydrate intake. By contrast, the molecular regulation of hepatic glucose production seems to be atypical (see further sections) and could be one of the reasons for the poor metabolic use of carbohydrate in trout.

Linking nutrition to metabolism: hormonal and nutrient sensor roles

The decrease of feed intake associated with a low level of plasmatic insulin/glucagon ratio (mainly due to the pursuit of glucagon secretion) are often observed in rainbow trout following the intake of new aquafeeds, having potential strong impacts on feed efficiency (Panserat and Medale, 2013). Recent studies about the nutritional regulation of nutrient sensing pathway in hypothalamus (center of the feed intake regulation) and in endocrine pancreas (center of insulin and glucagon secretion linked to the nutritional status) demonstrated the existence of glucose and lipid sensing mechanisms in these tissues (Caruso and Sheridan, 2011). Fine nutritional regulations of these pathways at a molecular level are in progress in rainbow trout fed alternative diets.

The 21st century or the "omics" boom in nutritional studies

In the 2000s, the omics approaches have been introduced for the nutritional studies in rainbow trout. In particular, the hepatic transcriptome analysis became a new tool to study nutrition in rainbow trout. Indeed, because the liver is the center of the intermediary metabolism, postprandial hepatic transcriptomics were analyzed after fish oil suppression (Kolditz et al., 2008; Panserat et al., 2008), fish oil replacement by vegetable oils (Panserat et al., 2008), total fish oil and fish meal replacement (Overturf et al., 2012) and micronutrient deficiencies (Olsvik et al., 2013). These data confirmed studies linked to the intermediary metabolism but opened also new hypothesis related to health (immunology), xenobiotic metabolism, proteolysis and cell division. Some of the transcriptomics studies were completed by proteomics analysis (Kolditz et al., 2008; Martin et al., 2003), which seemed to confirm the transcript analysis but were limited by the current technical inability to identify proteins in rainbow trout.

Nutritional regulation of hepatic glucose-6-phosphatase genes expression in rainbow trout: how the trout genome sequencing provides new answers to old questions

The rainbow trout is considered to be a glucose-intolerant species due mainly to persistent hyperglycemia after intake of carbohydrate-enriched meal(s) or glucose tolerance tests (Polakof et al., 2012). One hypothesis to explain such a phenotype is an insufficient inhibition of endogenous glucose production via gluconeogenesis pathway (Panserat et al., 2000). Hepatic glucose-6-phosphatase (G6pc), the last gluconeogenic enzyme catalyzing the hydrolysis of glucose-6-phosphate (G6P) in glucose, was proposed to be a major actor involved in blood glucose enrichment under hyperglycemic conditions [Type 2 diabetes (Rooney et al., 1993)] or after carbohydrate intake [in seabass (Viegas et al., 2015)] by establishing a futile glucose/G6P cycle together with glucokinase, the first glycolytic enzyme catalyzing the phosphorylation of glucose in G6P. A gene coding for G6pc in rainbow trout was for the first time partially sequenced in 2000 (Panserat et al., 2000) and the nutritional/hormonal regulation of its expression was then deeply investigated first by northern blot analysis (Panserat et al., 2002, 2000) progressively replaced by contemporary real-time

quantitative PCR analysis. *g6pc* mRNA level was thus shown to be up-regulated by high dietary lipids but was surprisingly not regulated by high carbohydrate diet intake containing digestible starch (Panserat et al., 2008). In vivo and in vitro analysis also demonstrated that *g6pc* mRNA level was positively correlated to protein (Kirchner et al., 2003a) and amino acids (AA) pool (Lansard et al., 2010) levels respectively. Moreover specific AA were involved in the regulation of *g6pc* gene such as alanine (Kirchner et al., 2003b), one gluconeogenic dispensable AA (G-DAA), or leucine and methionine (Lansard et al., 2011). Finally insulin regulation of this gene was investigated and demonstrated to have a down-regulation effect on *g6pc* mRNA level both in vivo (Polakof et al., 2009) and in vitro (Plagnes-Juan et al., 2008) and to be able to counteract the increase of mRNA level induced by glucose addition in the cell cultured medium or in trout fed a high carbohydrate diet. However, some of these expression results remained questioning when putting into perspective of enzyme activity results. For instance insulin seemed to act at the molecular level to down-regulate *g6pc* mRNA level but had no effect on G6pc hepatic and gut activity (Polakof et al., 2011). In the same way, *g6pc* mRNA level was down-regulated by only one G-DAA (alanine), whereas the enzyme activity was decreased in trout fed a three G-DAA substituted diets (alanine or aspartic acid or glutamic acid, (Kirchner et al., 2003b)). In addition, no clear reason could be given to explain why G6pc activity did not change when *g6pc* was repressed by a single meal with glucose (Panserat et al., 2000) until the identification of a second gene coding for G6pc in EST databases in 2008 (Plagnes-Juan et al., 2008). Indeed the analysis of this new gene brought new information to interpret data presented above and let sense that the complexity of rainbow trout genome has to be considered to better understand nutritional regulation. In particular, data showed that both *g6pc* genes were differentially regulated by nutritional status (Mennigen et al., 2013) but also displayed a contrasting regulation regarding the relative proportion of carbohydrate in the diet (Kamalam et al., 2012; Seiliez et al., 2011). With the recent sequencing of the rainbow trout genome (Berthelot et al., 2014) molecular comprehension of *g6pc*/G6pc nutritional regulation moved on to a next step. Actually in silico investigations revealed that in fact five paralogous genes coding for G6pc were retained in the rainbow trout genome after Ss4R and that two of them, never identified before, displayed an unexpected up-regulation of their mRNA levels in trout fed a high carbohydrate diet (Marandel et al., 2015). Overall enzyme activity measured in the liver of these fish suggested that proteins translated from the latter two genes might be involved in this activity. It was hypothesized that these two genes may contribute to the hyperglycemic phenotype by releasing glucose in blood via the establishment of a futile cycle together with glucokinase. Old results concerning *g6pc* genes regulation by insulin or G-DAA can be now reinterpreted in the light of this new discovery and by considering that *g6pc* paralogs are potentially differentially regulated but that they all may contribute to overall enzyme activity.

The example of *g6pc* genes strongly illustrates how advances in genomics can serve our understanding of nutritional regulation of metabolism and provide new answers to old questions. But as usual, close a window often opens a door and many questions related to the role of differential regulation of *g6pc* paralogous genes and about their related contribution to one or another phenotype arise.

Trout genomics for health

The rainbow trout immunome: the added value of a full repertoire of fish immunity genes

Toward the definition of a trout immunome

When the human genome was sequenced, a quasicomplete repertoire of proteins involved in immunity—the "Immunome"—became available (Ortutay and Vihinen, 2009). Lots of data were already available at the molecular, structural, and cellular levels, in both normal and diseased states regarding genes involved in mouse and human immunity; these data were integrated in databases and tools dedicated to immune genes and pathways. When the genome of another vertebrate is sequenced, its immunome can be annotated using three main approaches: (1) the functional knowledge available for immune genes of other species like human or mouse can be attached to their orthologs, considering that phylogenetic orthologs in different species often have similar roles, (2) datasets from high-throughput approaches of gene expression in infectious and other pathological states can identify genes modulated in diseases and during responses (3) at last, direct genomic and functional studies of genes or gene families can be undertaken from the newly available genome.

The annotation of the rainbow trout immunome is not an easy task and cannot be achieved by a direct transposition of the knowledge available in mouse and human databases; indeed, as mentioned above, two cycles of WGD occurred during the evolution of salmonids, providing a very large number of paralogous genes susceptible to sub functionalization. Additionally, several gene families such as chemokines and trims underwent considerable specific expansion in salmonids (Boudinot et al., 2011; Chen et al., 2013; Marancik et al., 2014). Hence, a well-assembled genome in which closely related duplicates are predicted constitutes a pivotal resource to define the immunome, especially from high-throughput transcriptomes.

An advanced trout immunome will provide a good basis for further developments such as the analysis of genetic polymorphism and epigenetic approaches, which will be instrumental for marker based selection programs.

The case of antigen specific receptors, Igs and TCRs

As new sequencing technologies gave access to the quasi-complete diversity of immunoglobulin (Ig) or T cell receptor (TCR) sequences expressed in a tissue—or even in a whole individual—studies of immune repertoires have greatly expanded in the last five years (Castro et al., 2013; Weinstein et al., 2009). These studies provide a detailed account of the adaptive immune responses against pathogens, and assess the status of the immune system in different pathologies. Such approaches have been developed in rainbow trout and zebrafish based on Ig and TCR sequences available in these species. The availability of a full trout genome sequence, and the potential possibility to obtain a detailed annotation of TCR and Ig loci in this species, will provide researchers with a much-empowered system to access the complete diversity of immune repertoires and responses. Additionally, the modelization approaches that apply to repertoire data using theoretical physics tools (Mora et al., 2010) can be developed only when the

sequences of all V, D, and J gene segments to recombine are available. The rainbow trout genome represents the first step—and a necessary resource—for the development of such approaches that will be instrumental for evaluation of future vaccines.

Interest of a genome assembly for linkage studies: gene clusters and linkage studies

The presence and functional importance of gene clusters involved in immunity remain open questions. Several selection pressures may account for such clusters, including local dynamics of gene duplication and coregulation of expression. The trout genome provides an interesting opportunity to address such issues, one of which refers to the history of the major histocompatibility complex (MHC). TCRs recognize antigens as small peptides (9–22 amino acids) bound to membrane proteins called MHC class I or class II molecules, that are expressed at the surface of "antigen presenting cells." The TCR recognizes both the MHC protein and the peptide antigen presented by the MHC molecule, a phenomenon known as "MHC restriction." The genetic "MHC" is a region of the genome defined by the presence of MHC class I and II molecules, and genes involved in the preparation and loading of the peptide antigen for MHC presentation. In fact, the MHC encodes many genes (more than 100 genes in human), approximately half of which are implicated in immunity (The_MHC_sequencing_consortium, 1999); thus the genes of this region bring an important contribution to the immunome and defense mechanisms. The origin and evolutionary pathway of the MHC remains incompletely understood. MHC is found in all jaw vertebrates from sharks to mammals. As mentioned earlier, it is now generally accepted that according to Ohno's theory (Ohno et al., 1968), two rounds of whole genome duplication occurred after the emergence of urochordates and before the radiation of jawed vertebrates, leading to four sets of paralogous regions. Such tetrads were identified for several key genetic regions, including the HOX and MHC complexes. Two additional tetrads were identified across vertebrate genomes, that apparently derived from neighboring regions on the same ancestral chromosome containing the "proto-MHC" of common ancestors of vertebrates and other phyla (Flajnik et al., 2012); such proto-MHC regions were, for example, identified in the placozoan *Trichoplax adhaerens*, which is considered as a basal bilaterian branch (Suurvali et al., 2014). Altogether, these three MHC related tetrads encode a large number of genes involved in immunity in vertebrates, including several major gene complexes such as the natural killer receptor complex and the leukocyte receptor complex; they also count many important immune gene families as B7 receptors, TRIMs, etc.

The MHC was conserved as a defined genomic entity across vertebrates from sharks to mammals, with the exception of fishes in which the regions encoding MHC class I and MHC class II are found on different chromosomes (Flajnik and Kasahara, 2010). The functional significance of this conservation, and of the different pattern observed in fish, remains elusive. Coregulation of many genes involved in immunity is a tempting hypothesis, but it has received little direct support to date and is challenged by the pattern observed in fish. Fishes represent interesting models to address such questions and get insights into the immunological impact of a "broken" MHC. Fish groups that have lost MHC class II genes and classical T helper responses—such as cods and

pipefishes (Star and Jentoft, 2012)—are good examples of potential repercussions; Salmonids, with additional cycles of genome duplication and ongoing rediploidization, represent another context of great interest to challenge and understand the fate of the MHC region and its paralogs or related regions.

The rainbow trout genome constitutes a great resource for fish and comparative immunologists. As a basis to define a complete immunome, it is a very important resource for future developments requiring monitoring of immune responses, including selection of more robust animals. Additionally, with the genome of the closely related Atlantic salmon, it also represents a unique opportunity to address generic and evolutionary questions for which teleosts harbor unique features. Finally, a good quality genome assembly will also be very important for future studies of the trout gut (or other) microbiota.

Rainbow trout reproduction with a genome sequence

A search in the Web of Science database (Science Citation Expanded 1975–present) yielded over 8,000 articles dealing with various aspects of rainbow trout reproduction. In contrast, a similar search for zebrafish—a widely used biological model—reproduction only yielded 3,500 references. This illustrates the importance of the scientific community working in the field of reproduction in rainbow trout and the potential impact of the rainbow trout genome and related genomic resources to study reproduction in this species. As for other disciplinary fields, the recent release of the rainbow trout genome now offers new possibilities to study gene synteny and to elucidate previously unresolved evolutionary history of genes playing important roles in reproduction as illustrated hereafter.

Sex hormone-binding globulins (Shbg) are carrier blood proteins involved in the transport of sex steroids in plasma and the regulation of their availability to target organs. Two genes—*shbga* and *shbgb*—exist in fish, *shbga* being the orthologs of the *SHBG* human gene. Shbgb has been found in rainbow trout and salmon but never described in any nonsalmonid teleost species to date. While this could suggest that *shbga* and *shbgb* originate from Ss4R WGD, the topology of the Shbg tree (Bobe et al., 2010) as well as recent analyses using the rainbow trout genome sequence suggest that this duplication is much more ancient than initially thought.

Gonadal soma-derived growth factor (Gsdf) is a recently identified member of the TGF-beta superfamily (Sawatari et al., 2007) that is believed to play an important role for proliferation of primordial germ cells and spermatogonia. It is also spatially and temporally correlated with early testicular differentiation in medaka (Shibata et al., 2010). Within medaka gonads, *gsdf* gene expression is restricted to Sertoli and granulosa cells. The fate of *gsdf* after Ss4R WGD remains however uncharacterized but recent preliminary analyses suggest that Ss4R yielded two *gsdf* genes in rainbow trout with possible different expression patterns (Fig. 2.3). The rainbow trout genome will allow thorough characterization of *gsdf* ohnologs including any gene silencing or sub-/neo-functionalization after duplication.

With the genome sequence of rainbow trout, the use of genome editing will also become possible without relying on existing resources (e.g., BAC libraries) to target

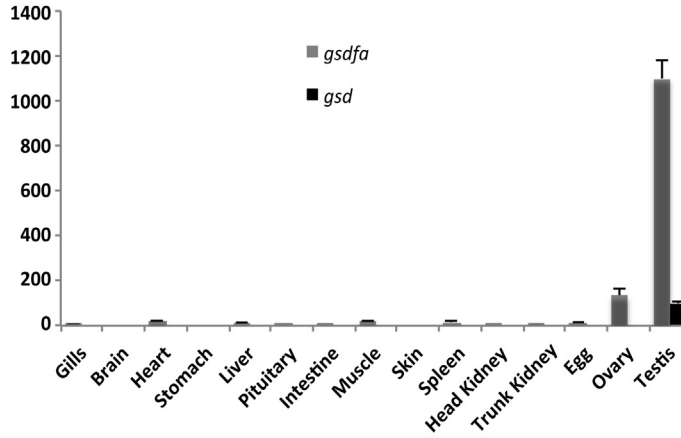

Figure 2.3 Example of evolution of gene expression after Ss4R gene duplication. Expression levels of *gsdf* Ss4R ohnologs in different rainbow trout tissues, showing that one of the two *gsdf* ohnolog is predominantly expressed even if the expression profiles of these two genes are still identical. mRNA levels (mean, SD) were measured in tissues originating from three individuals. Expression was arbitrarily set to 100 for *gsdfb* expression in testis.

important genes and describe the phenotypes associated with corresponding knockouts. While rainbow trout has a relatively long lifecycle, in comparison to other fish model species (e.g., zebrafish, medaka), gene knockout is sometimes necessary to decipher the function of salmonid specific genes such as the sex determining gene *sdY* (Yano et al., 2012). Because of its long-term use as a model for reproduction, rainbow trout could be an interesting model to perform genome editing in a well-characterized teleost model exhibiting group-synchronous oogenetic development. A recent study showed that the promising and now widely used Crispr/CAS9 technology can successfully be used in salmonids (Edvardsen et al., 2014).

In summary, the recent publication of the rainbow trout genome offers new possibilities in the field of reproductive biology with major outcomes related to the control of sex-ratio, puberty, fecundity, and sterility.

Future directions and concluding remarks: resources are still expending

Next generation sequencing (NGS) is still a rapidly evolving field, and second-generation of NGS technologies are now available with the emergence of longer-read length technologies and higher volumes of produced sequence information. Using NGS in rainbow trout, transcriptome resources have been extended by numerous studies that have sequenced cDNAs in different tissues, stages of development, or conditions of treatments (Ali et al., 2014; Liu et al., 2014; Marancik et al., 2014; Salem et al., 2015). In addition to cDNA transcripts, deep microRNA sequencing has also been carried

out in rainbow trout (Farlora et al., 2015; Juanchich et al., 2013; Ma et al., 2015; Salem et al., 2010) leading to detailed catalogs of microRNA repertoires. With new NGS technologies resequencing of the whole genome of new rainbow trout individuals from differing genetic origins, including doubled haploids individuals, will also be possible at limited cost. Therefore, new markers and new polymorphisms will be discovered, that will facilitate further QTL detection and refine functional analyses.

A new updated version of the rainbow trout genome has been announced to be released in 2016 based on recent NGS improvements of the Illumina sequencing technique allowing the production of 2 × 250 bp reads in large amounts (USDA initiative, Unpublished). The total size of the resulting assembly of this updated version of the rainbow trout genome is announced to be 2,17 Gb with a scaffold N50 of 1,700 kb. All these quickly growing resources will now have to be integrated to produce reference comprehensive maps of all functional elements in the rainbow trout genomes for instance through international initiative like the Functional Annotation of Animal Genomes (Andersson et al., 2015), which is the equivalent of the ENCODE Human genome. Altogether these joined efforts will lead to a better understanding of the rainbow trout genome allowing the use of these integrated resources for whole genome genetic selection, investigation of epigenetic regulations on a genome-scale (Baerwald et al., 2015) and all kinds of whole genome analysis.

References

Ali, A., Rexroad, C.E., Thorgaard, G.H., Yao, J., Salem, M., 2014. Characterization of the rainbow trout spleen transcriptome and identification of immune-related genes. Front. Genet. 5, 348.

Allendorf, F.W., Thorgaard, G.H., 1984. Tetraploidy and the evolution of salmonid fishes. In: Turner, B.J. (Ed.), Evolutionary Genetics of Fishes, Monographs in Evolutionary Biology. Springer, US, pp. 1–53.

Andersson, L., Archibald, A.L., Bottema, C.D., Brauning, R., Burgess, S.C., Burt, D.W., Casas, E., Cheng, H.H., Clarke, L., Couldrey, C., Dalrymple, B.P., Elsik, C.G., Foissac, S., Giuffra, E., Groenen, M.A., Hayes, B.J., Huang, L.S., Khatib, H., Kijas, J.W., Kim, H., Lunney, J.K., McCarthy, F.M., McEwan, J.C., Moore, S., Nanduri, B., Notredame, C., Palti, Y., Plastow, G.S., Reecy, J.M., Rohrer, G.A., Sarropoulou, E., Schmidt, C.J., Silverstein, J., Tellam, R.L., Tixier-Boichard, M., Tosser-Klopp, G., Tuggle, C.K., Vilkki, J., White, S.N., Zhao, S., Zhou, H., FAANG Consortium, 2015. Coordinated international action to accelerate genome-to-phenome with FAANG, the functional annotation of animal genomes project. Genome Biol. 16, 57.

Baerwald, M.R., Meek, M.H., Stephens, M.R., Nagarajan, R.P., Goodbla, A.M., Tomalty, K.M.H., Thorgaard, G.H., May, B., Nichols, K.M., 2015. Migration-related phenotypic divergence is associated with epigenetic modifications in rainbow trout. Mol. Ecol. 25 (8), 1785–1800.

Baerwald, M.R., Petersen, J.L., Hedrick, R.P., Schisler, G.J., May, B., 2011. A major effect quantitative trait locus for whirling disease resistance identified in rainbow trout (*Oncorhynchus mykiss*). Heredity 106, 920–926.

Behnke, R.J., Tomelleri, J., Proebstel, D.S., 2002. Trout and Salmon of North America, first ed. Chanticleer Press, Free Press, New York.

Bernardi, G., Bernardi, G., 1990. Compositional transitions in the nuclear genomes of cold-blooded vertebrates. J. Mol. Evol. 31, 282–293.

Berthelot, C., Brunet, F., Chalopin, D., Juanchich, A., Bernard, M., Noël, B., Bento, P., Da Silva, C., Labadie, K., Alberti, A., Aury, J.-M., Louis, A., Dehais, P., Bardou, P., Montfort, J., Klopp, C., Cabau, C., Gaspin, C., Thorgaard, G.H., Boussaha, M., Quillet, E., Guyomard, R., Galiana, D., Bobe, J., Volff, J.-N., Genêt, C., Wincker, P., Jaillon, O., Roest Crollius, H., Guiguen, Y., 2014. The rainbow trout genome provides novel insights into evolution after whole-genome duplication in vertebrates. Nat. Commun. 5, 3657.

Bobe, J., Guiguen, Y., Fostier, A., 2010. Diversity and biological significance of sex hormone-binding globulin in fish, an evolutionary perspective. Mol. Cell. Endocrinol. 316, 66–78.

Boudinot, P., van der Aa, L.M., Jouneau, L., Pasquier, L. Du, Pontarotti, P., Briolat, V., Benmansour, A., Levraud, J.P., 2011. Origin and evolution of TRIM proteins: new insights from the complete TRIM repertoire of zebrafish and pufferfish. PLoS One 6, e22022.

Canario, A.V.M., Bargelloni, L., Volckaert, F., Houston, R.D., Massault, C., Guiguen, Y., 2008. Genomics toolbox for farmed fish. Rev. Fish. Sci. 16, 3–15.

Caruso, M.A., Sheridan, M.A., 2011. New insights into the signaling system and function of insulin in fish. Gen. Comp. Endocrinol. 173, 227–247.

Castro, R., Jouneau, L., Pham, H.P., Bouchez, O., Giudicelli, V., Lefranc, M.P., Quillet, E., Benmansour, A., Cazals, F., Six, A., Fillatreau, S., Sunyer, O., Boudinot, P., 2013. Teleost fish mount complex clonal IgM and IgT responses in spleen upon systemic viral infection. PLoS Pathog. 9, e1003098.

Chen, J., Xu, Q., Wang, T., Collet, B., Corripio-Miyar, Y., Bird, S., Xie, P., Nie, P., Secombes, C.J., Zou, J., 2013. Phylogenetic analysis of vertebrate CXC chemokines reveals novel lineage specific groups in teleost fish. Dev. Comp. Immunol. 41, 137–152.

Chourrout, D., 1984. Pressure-induced retention of second polar body and suppression of first cleavage in rainbow trout: production of all-triploids, all-tetraploids, and heterozygous and homozygous diploid gynogenetics. Aquaculture 36, 111–126.

Crête-Lafrenière, A., Weir, L.K., Bernatchez, L., 2012. Framing the Salmonidae family phylogenetic portrait: a more complete picture from increased taxon sampling. PloS One 7, e46662.

Dehal, P., Boore, J.L., 2005. Two rounds of whole genome duplication in the ancestral vertebrate. PLoS Biol. 3, e314.

Diaz, M., Vraskou, Y., Gutierrez, J., Planas, J.V., 2009. Expression of rainbow trout glucose transporters GLUT1 and GLUT4 during in vitro muscle cell differentiation and regulation by insulin and IGF-I. Am. J. Physiol. Regul. Integr. Comp. Physiol. 296, R794–R800.

Edvardsen, R.B., Leininger, S., Kleppe, L., Skaftnesmo, K.O., Wargelius, A., 2014. Targeted mutagenesis in Atlantic salmon (*Salmo salar* L.) using the CRISPR/Cas9 system induces complete knockout individuals in the F0 generation. PloS One 9, e108622.

Etter, P.D., Bassham, S., Hohenlohe, P.A., Johnson, E.A., Cresko, W.A., 2011. SNP discovery and genotyping for evolutionary genetics using RAD sequencing. Methods Mol. Biol. Clifton NJ 772, 157–178.

Farlora, R., Valenzuela-Miranda, D., Alarcón-Matus, P., Gallardo-Escárate, C., 2015. Identification of microRNAs associated with sexual maturity in rainbow trout brain and testis through small RNA deep sequencing. Mol. Reprod. Dev. 82 (9), 651–662.

Flajnik, M.F., Kasahara, M., 2010. Origin and evolution of the adaptive immune system: genetic events and selective pressures. Nat. Rev. Genet. 11, 47–59.

Flajnik, M.F., Tlapakova, T., Criscitiello, M.F., Krylov, V., Ohta, Y., 2012. Evolution of the B7 family: co-evolution of B7H6 and NKp30, identification of a new B7 family member, B7H7, and of B7's historical relationship with the MHC. Immunogenetics 64, 571–590.

Genet, C., Dehais, P., Palti, Y., Gao, G., Gavory, F., Wincker, P., Quillet, E., Boussaha, M., 2011. Analysis of BAC-end sequences in rainbow trout: content characterization and assessment of synteny between trout and other fish genomes. BMC Genomics 12, 314.

Gilles, A., Meglécz, E., Pech, N., Ferreira, S., Malausa, T., Martin, J.-F., 2011. Accuracy and quality assessment of 454 GS-FLX Titanium pyrosequencing. BMC Genomics 12, 245.

Govoroun, M., Le Gac, F., Guiguen, Y., 2006. Generation of a large scale repertoire of expressed sequence tags (ESTs) from normalised rainbow trout cDNA libraries. BMC Genomics 7, 196.

Guyomard, R., Mauger, S., Tabet-Canale, K., Martineau, S., Genet, C., Krieg, F., Quillet, E., 2006. A type I and type II microsatellite linkage map of rainbow trout (*Oncorhynchus mykiss*) with presumptive coverage of all chromosome arms. BMC Genomics 7, 302.

Guyomard, R., Boussaha, M., Krieg, F., Hervet, C., Quillet, E., 2012. A synthetic rainbow trout linkage map provides new insights into the salmonid whole genome duplication and the conservation of synteny among teleosts. BMC Genet. 13, 15.

Hecht, B.C., Thrower, F.P., Hale, M.C., Miller, M.R., Nichols, K.M., 2012. Genetic architecture of migration-related traits in rainbow and steelhead trout, *Oncorhynchus mykiss*. Genes Genom. Genet. 2, 1113–1127.

Howe, K., Clark, M.D., Torroja, C.F., Torrance, J., Berthelot, C., Muffato, M., Collins, J.E., Humphray, S., McLaren, K., Matthews, L., McLaren, S., Sealy, I., Caccamo, M., Churcher, C., Scott, C., Barrett, J.C., Koch, R., Rauch, G.-J., White, S., Chow, W., Kilian, B., Quintais, L.T., Guerra-Assunção, J.A., Zhou, Y., Gu, Y., Yen, J., Vogel, J.-H., Eyre, T., Redmond, S., Banerjee, R., Chi, J., Fu, B., Langley, E., Maguire, S.F., Laird, G.K., Lloyd, D., Kenyon, E., Donaldson, S., Sehra, H., Almeida-King, J., Loveland, J., Trevanion, S., Jones, M., Quail, M., Willey, D., Hunt, A., Burton, J., Sims, S., McLay, K., Plumb, B., Davis, J., Clee, C., Oliver, K., Clark, R., Riddle, C., Elliot, D., Eliott, D., Threadgold, G., Harden, G., Ware, D., Begum, S., Mortimore, B., Mortimer, B., Kerry, G., Heath, P., Phillimore, B., Tracey, A., Corby, N., Dunn, M., Johnson, C., Wood, J., Clark, S., Pelan, S., Griffiths, G., Smith, M., Glithero, R., Howden, P., Barker, N., Lloyd, C., Stevens, C., Harley, J., Holt, K., Panagiotidis, G., Lovell, J., Beasley, H., Henderson, C., Gordon, D., Auger, K., Wright, D., Collins, J., Raisen, C., Dyer, L., Leung, K., Robertson, L., Ambridge, K., Leongamornlert, D., McGuire, S., Gilderthorp, R., Griffiths, C., Manthravadi, D., Nichol, S., Barker, G., Whitehead, S., Kay, M., Brown, J., Murnane, C., Gray, E., Humphries, M., Sycamore, N., Barker, D., Saunders, D., Wallis, J., Babbage, A., Hammond, S., Mashreghi-Mohammadi, M., Barr, L., Martin, S., Wray, P., Ellington, A., Matthews, N., Ellwood, M., Woodmansey, R., Clark, G., Cooper, J.D., Cooper, J., Tromans, A., Grafham, D., Skuce, C., Pandian, R., Andrews, R., Harrison, E., Kimberley, A., Garnett, J., Fosker, N., Hall, R., Garner, P., Kelly, D., Bird, C., Palmer, S., Gehring, I., Berger, A., Dooley, C.M., Ersan-Ürün, Z., Eser, C., Geiger, H., Geisler, M., Karotki, L., Kirn, A., Konantz, J., Konantz, M., Oberländer, M., Rudolph-Geiger, S., Teucke, M., Lanz, C., Raddatz, G., Osoegawa, K., Zhu, B., Rapp, A., Widaa, S., Langford, C., Yang, F., Schuster, S.C., Carter, N.P., Harrow, J., Ning, Z., Herrero, J., Searle, S.M.J., Enright, A., Geisler, R., Plasterk, R.H.A., Lee, C., Westerfield, M., de Jong, P.J., Zon, L.I., Postlethwait, J.H., Nüsslein-Volhard, C., Hubbard, T.J.P., Roest Crollius, H., Rogers, J., Stemple, D.L., 2013. The zebrafish reference genome sequence and its relationship to the human genome. Nature 496, 498–503.

Hurley, I.A., Mueller, R.L., Dunn, K.A., Schmidt, E.J., Friedman, M., Ho, R.K., Prince, V.E., Yang, Z., Thomas, M.G., Coates, M.I., 2007. A new time-scale for ray-finned fish evolution. Proc. Biol. Sci. 274, 489–498.

Juanchich, A., Le Cam, A., Montfort, J., Guiguen, Y., Bobe, J., 2013. Identification of differentially expressed miRNAs and their potential targets during fish ovarian development. Biol. Reprod. 88, 128.

Kamalam, B.S., Medale, F., Kaushik, S., Polakof, S., Skiba-Cassy, S., Panserat, S., 2012. Regulation of metabolism by dietary carbohydrates in two lines of rainbow trout divergently selected for muscle fat content. J. Exp. Biol. 215, 2567–2578.

Kirchner, S., Kaushik, S., Panserat, S., 2003a. Low protein intake is associated with reduced hepatic gluconeogenic enzyme expression in rainbow trout (*Oncorhynchus mykiss*). J. Nutr. 133, 2561–2564.

Kirchner, S., Kaushik, S., Panserat, S., 2003b. Effect of partial substitution of dietary protein by a single gluconeogenic dispensable amino acid on hepatic glucose metabolism in rainbow trout (*Oncorhynchus mykiss*). Comp. Biochem. Physiol. Mol. Integr. Physiol. 134, 337–347.

Kolditz, C.I., Paboeuf, G., Borthaire, M., Esquerre, D., SanCristobal, M., Lefevre, F., Medale, F., 2008. Changes induced by dietary energy intake and divergent selection for muscle fat content in rainbow trout (*Oncorhynchus mykiss*), assessed by transcriptome and proteome analysis of the liver. BMC Genomics 9, 506.

Komen, H., Thorgaard, G.H., 2007. Androgenesis, gynogenesis and the production of clones in fishes: a review. Aquaculture 269, 150–173.

Koop, B.F., von Schalburg, K.R., Leong, J., Walker, N., Lieph, R., Cooper, G.A., Robb, A., Beetz-Sargent, M., Holt, R.A., Moore, R., Brahmbhatt, S., Rosner, J., Rexroad, C.E., McGowan, C.R., Davidson, W.S., 2008. A salmonid EST genomic study: genes, duplications, phylogeny and microarrays. BMC Genomics 9, 545.

Langham, R.J., Walsh, J., Dunn, M., Ko, C., Goff, S.A., Freeling, M., 2004. Genomic duplication, fractionation and the origin of regulatory novelty. Genetics 166, 935–945.

Lansard, M., Panserat, S., Plagnes-Juan, E., Seiliez, I., Skiba-Cassy, S., 2010. Integration of insulin and amino acid signals that regulate hepatic metabolism-related gene expression in rainbow trout: role of TOR. Amino Acids 39, 801–810.

Lansard, M., Panserat, S., Plagnes-Juan, E., Dias, K., Seiliez, I., Skiba-Cassy, S., 2011. L-leucine, L-methionine, and L-lysine are involved in the regulation of intermediary metabolism-related gene expression in rainbow trout hepatocytes. J. Nutr. 141, 75–80.

Le-Bras, Y., Dechamp, N., Krieg, F., Filangi, O., Guyomard, R., Boussaha, M., Bovenhuis, H., Pottinger, T.G., Prunet, P., Le-Roy, P., Quillet, E., 2011. Detection of QTL with effects on osmoregulation capacities in the rainbow trout (*Oncorhynchus mykiss*). BMC Genet. 12, 46.

Liu, S., Gao, G., Palti, Y., Cleveland, B.M., Weber, G.M., Rexroad, C.E., 2014. RNA-seq analysis of early hepatic response to handling and confinement stress in rainbow trout. PloS One 9, e88492.

Liu, S.X., Vallejo, R.L., Gao, G.T., Palti, Y., Weber, G.M., Hernandez, A., Rexroad, C.E., 2015. Identification of Single-Nucleotide Polymorphism Markers Associated with Cortisol Response to Crowding in Rainbow Trout. Mar. Biotechnol. 17 (3), 328–337.

Ma, H., Weber, G.M., Hostuttler, M.A., Wei, H., Wang, L., Yao, J., 2015. MicroRNA expression profiles from eggs of different qualities associated with post-ovulatory ageing in rainbow trout (*Oncorhynchus mykiss*). BMC Genomics 16, 201.

Macqueen, D.J., Johnston, I.A., 2014. A well-constrained estimate for the timing of the salmonid whole genome duplication reveals major decoupling from species diversification. Proc. Biol. Sci. 281, 20132881.

Marancik, D., Gao, G., Paneru, B., Ma, H., Hernandez, A.G., Salem, M., Yao, J., Palti, Y., Wiens, G.D., 2014. Whole-body transcriptome of selectively bred, resistant-, control-, and susceptible-line rainbow trout following experimental challenge with *Flavobacterium psychrophilum*. Front Genet. 5, 453.

Marandel, L., Seiliez, I., Veron, V., Skiba-Cassy, S., Panserat, S., 2015. New insights into the nutritional regulation of gluconeogenesis in carnivorous rainbow trout (*Oncorhynchus mykiss*): a gene duplication trail. Physiol. Genomics 47 (7), 253–263.

Martin, S.A., Vilhelmsson, O., Medale, F., Watt, P., Kaushik, S., Houlihan, D.F., 2003. Proteomic sensitivity to dietary manipulations in rainbow trout. Biochim. Biophys. Acta 1651, 17–29.

Mennigen, J.A., Panserat, S., Larquier, M., Plagnes-Juan, E., Medale, F., Seiliez, I., Skiba-Cassy, S., 2012. Postprandial regulation of hepatic microRNAs predicted to target the insulin pathway in rainbow trout. PLoS One 7, e38604.

Mennigen, J.A., Skiba-Cassy, S., Panserat, S., 2013. Ontogenetic expression of metabolic genes and microRNAs in rainbow trout alevins during the transition from the endogenous to the exogenous feeding period. J. Exp. Biol. 216, 1597–1608.

Mennigen, J.A., Martyniuk, C.J., Seiliez, I., Panserat, S., Skiba-Cassy, S., 2014. Metabolic consequences of microRNA-122 inhibition in rainbow trout, *Oncorhynchus mykiss*. BMC Genomics 15, 70.

Miller, M.R., Brunelli, J.P., Wheeler, P.A., Liu, S., Rexroad, C.E., Palti, Y., Doe, C.Q., Thorgaard, G.H., 2012. A conserved haplotype controls parallel adaptation in geographically distant salmonid populations. Mol. Ecol. 21, 237–249.

Monroig, O., Tocher, D.R., Navarro, J.C., 2013. Biosynthesis of polyunsaturated fatty acids in marine invertebrates: recent advances in molecular mechanisms. Mar. Drugs 11, 3998–4018.

Mora, T., Walczak, A.M., Bialek, W., Callan, C.G., 2010. Maximum entropy models for antibody diversity. Proc. Natl. Acad. Sci. USA 107, 5405–5410.

Naylor, R.L., Hardy, R.W., Bureau, D.P., Chiu, A., Elliott, M., Farrell, A.P., Forster, I., Gatlin, D.M., Goldburg, R.J., Hua, K., Nichols, P.D., 2009. Feeding aquaculture in an era of finite resources. Proc. Natl. Acad. Sci. USA 106, 15103–15110.

Near, T.J., Eytan, R.I., Dornburg, A., Kuhn, K.L., Moore, J.A., Davis, M.P., Wainwright, P.C., Friedman, M., Smith, W.L., 2012. Resolution of ray-finned fish phylogeny and timing of diversification. Proc. Natl. Acad. Sci. USA 109, 13698–13703.

Nichols, K.M., Young, W.P., Danzmann, R.G., Robison, B.D., Rexroad, C., Noakes, M., Phillips, R.B., Bentzen, P., Spies, I., Knudsen, K., Allendorf, F.W., Cunningham, B.M., Brunelli, J., Zhang, H., Ristow, S., Drew, R., Brown, K.H., Wheeler, P.A., Thorgaard, G.H., 2003. A consolidated linkage map for rainbow trout (*Oncorhynchus mykiss*). Anim. Genet. 34, 102–115.

Ohno, S., Wolf, U., Atkin, N.B., 1968. Evolution from fish to mammals by gene duplication. Hereditas 59, 169–187.

Olsvik, P.A., Hemre, G.I., Waagbo, R., 2013. Exploring early micronutrient deficiencies in rainbow Trout (*Oncorhynchus mykiss*) by next-generation sequencing technology--from black box to functional genomics. PLoS One 8, e69461.

Ortutay, C., Vihinen, M., 2009. Immunome knowledge base (IKB): an integrated service for immunome research. BMC Immunol. 10, 3.

Overturf, K., Vallejo, R.L., Palti, Y., Barrows, F.T., Parsons, J.E., 2012. Microarray analysis of differential utilization of plant-based diets by rainbow trout. Aquac. Int. 20, 213–232.

Palti, Y., Genet, C., Luo, M.-C., Charlet, A., Gao, G., Hu, Y., Castaño-Sánchez, C., Tabet-Canale, K., Krieg, F., Yao, J., Vallejo, R.L., Rexroad, C.E., 2011. A first generation integrated map of the rainbow trout genome. BMC Genomics 12, 180.

Palti, Y., Genet, C., Gao, G., Hu, Y., You, F.M., Boussaha, M., Rexroad, C.E., Luo, M.-C., 2012. A second generation integrated map of the rainbow trout (*Oncorhynchus mykiss*) genome: analysis of conserved synteny with model fish genomes. Mar. Biotechnol. 14, 343–357.

Palti, Y., Gao, G., Miller, M.R., Vallejo, R.L., Wheeler, P.A., Quillet, E., Yao, J., Thorgaard, G.H., Salem, M., Rexroad, C.E., 2014. A resource of single-nucleotide polymorphisms for rainbow trout generated by restriction-site associated DNA sequencing of doubled haploids. Mol. Ecol. Resour. 14, 588–596.

Palti, Y., Gao, G., Liu, S., Kent, M.P., Lien, S., Miller, M.R., Rexroad, C.E., Moen, T., 2015. The development and characterization of a 57K single nucleotide polymorphism array for rainbow trout. Mol. Ecol. Resour. 15, 662–672.

Panserat, S., Kaushik, S.J., 2010. Regulation of gene expression by nutritional factors in fish. Aquac. Res. 41, 751–762.

Panserat, S.K.S., Medale, F., 2013. Rainbow trout as a model for nutrition and nutrient metabolism studies. Trout Physiol. Conserv, 8, -Nova Sci. Publ., Hauppauge, USA, p. 22.

Panserat, S., Medale, F., Blin, C., Breque, J., Vachot, C., Plagnes-Juan, E., Gomes, E., Krishnamoorthy, R., Kaushik, S., 2000. Hepatic glucokinase is induced by dietary carbohydrates in rainbow trout, gilthead seabream, and common carp. Am. J. Physiol. Regul. Integr. Comp. Physiol. 278, R1164–R1170.

Panserat, S., Perrin, A., Kaushik, S., 2002. High dietary lipids induce liver glucose-6-phosphatase expression in rainbow trout (*Oncorhynchus mykiss*). J. Nutr. 132, 137–141.

Panserat, S., Ducasse-Cabanot, S., Plagnes-Juan, E., Srivastava, P.P., Kolditz, C., Piumi, F., Esquerre, D., Kaushik, S., 2008. Dietary fat level modifies the expression of hepatic genes in juvenile rainbow trout (*Oncorhynchus mykiss*) as revealed by microarray analysis. Aquaculture 275, 235–241.

Panserat, S., Hortopan, G.A., Plagnes-Juan, E., Kolditz, C., Lansard, M., Skiba-Cassy, S., Esquerre, D., Geurden, I., Medale, F., Kaushik, S., Corraze, G., 2009. Differential gene expression after total replacement of dietary fish meal and fish oil by plant products in rainbow trout (*Oncorhynchus mykiss*) liver. Aquaculture 294, 123–131.

Panserat, S., Rideau, N., Polakof, S., 2014. Nutritional regulation of glucokinase: a cross-species story. Nutr. Res. Rev. 27, 21–47.

Parsons, J., Thorgaard, G., 1985. Production of androgenetic diploid rainbow-trout. J. Hered. 76, 177–181.

Plagnes-Juan, E., Lansard, M., Seiliez, I., Medale, F., Corraze, G., Kaushik, S., Panserat, S., Skiba-Cassy, S., 2008. Insulin regulates the expression of several metabolism-related genes in the liver and primary hepatocytes of rainbow trout (*Oncorhynchus mykiss*). J. Exp. Biol. 211, 2510–2518.

Polakof, S., Skiba-Cassy, S., Panserat, S., 2009. Glucose homeostasis is impaired by a paradoxical interaction between metformin and insulin in carnivorous rainbow trout. Am. J. Physiol. Regul. Integr. Comp. Physiol. 297, R1769–R1776.

Polakof, S., Moon, T.W., Aguirre, P., Skiba-Cassy, S., Panserat, S., 2011. Glucose homeostasis in rainbow trout fed a high-carbohydrate diet: metformin and insulin interact in a tissue-dependent manner. Am. J. Physiol. Regul. Integr. Comp. Physiol. 300, R166–R174.

Polakof, S., Panserat, S., Soengas, J.L., Moon, T.W., 2012. Glucose metabolism in fish: a review. J. Comp. Physiol. B 182, 1015–1045.

Quillet, E., Garcia, P., Guyomard, R., 1991. Analysis of the production of all homozygous lines of rainbow trout by gynogenesis. J. Exp. Zool. 257, 367–374.

Quillet, E., Dorson, M., Le-Guillou, S., Benmansour, A., Boudinot, P., 2007. Wide range of suceptibility to rhabdoviruses in homozygous clones of rainbow trout, Fish Shellfish. Immunol. 22, 510–519.

Quillet, E., Krieg, F., Dechamp, N., Hervet, C., Bérard, A., Le-Roy, P., Guyomard, R., Prunet, P., Pottinger, T.G., 2014. Quantitative trait loci for magnitude of the plasma cortisol response to confinement in rainbow trout. Anim. Genet. 45, 223–234.

Rexroad, C.E., Lee, Y., Keele, J.W., Karamycheva, S., Brown, G., Koop, B., Gahr, S.A., Palti, Y., Quackenbush, J., 2003. Sequence analysis of a rainbow trout cDNA library and creation of a gene index. Cytogenet. Genome Res. 102, 347–354.

Rexroad, C.E., Palti, Y., Gahr, S.A., Vallejo, R.L., 2008. A second generation genetic map for rainbow trout (*Oncorhynchus mykiss*). BMC Genet. 9, 74.

Roest Crollius, H., Jaillon, O., Bernot, A., Dasilva, C., Bouneau, L., Fischer, C., Fizames, C., Wincker, P., Brottier, P., Quétier, F., Saurin, W., Weissenbach, J., 2000. Estimate of human gene number provided by genome-wide analysis using Tetraodon nigroviridis DNA sequence. Nat. Genet. 25, 235–238.

Rooney, D.P., Neely, R.D., Beatty, O., Bell, N.P., Sheridan, B., Atkinson, A.B., Trimble, E.R., Bell, P.M., 1993. Contribution of glucose/glucose 6-phosphate cycle activity to insulin resistance in type 2 (non-insulin-dependent) diabetes mellitus. Diabetologia 36, 106–112.

Sakamoto, T., Danzmann, R.G., Gharbi, K., Howard, P., Ozaki, A., Khoo, S.K., Woram, R.A., Okamoto, N., Ferguson, M.M., Holm, L.E., Guyomard, R., Hoyheim, B., 2000. A microsatellite linkage map of rainbow trout (*Oncorhynchus mykiss*) characterized by large sex-specific differences in recombination rates. Genetics 155, 1331–1345.

Salem, M., Xiao, C., Womack, J., Rexroad, C.E., Yao, J., 2010. A microRNA repertoire for functional genome research in rainbow trout (*Oncorhynchus mykiss*). Mar. Biotechnol. 12, 410–429.

Salem, M., Paneru, B., Al-Tobasei, R., Abdouni, F., Thorgaard, G.H., Rexroad, C.E., Yao, J., 2015. Transcriptome assembly, gene annotation and tissue gene expression atlas of the rainbow trout. PloS One 10, e0121778.

Santini, F., Harmon, L.J., Carnevale, G., Alfaro, M.E., 2009. Did genome duplication drive the origin of teleosts? A comparative study of diversification in ray-finned fishes. BMC Evol. Biol. 9, 194.

Sawatari, E., Shikina, S., Takeuchi, T., Yoshizaki, G., 2007. A novel transforming growth factor-beta superfamily member expressed in gonadal somatic cells enhances primordial germ cell and spermatogonial proliferation in rainbow trout (*Oncorhynchus mykiss*). Dev. Biol. 301, 266–275.

Scheerer, P.D., Thorgaard, G.H., Allendorf, F.W., 1991. Genetic analysis of androgenetic rainbow trout. J. Exp. Zool. 260, 382–390.

Schranz, M.E., Mohammadin, S., Edger, P.P., 2012. Ancient whole genome duplications, novelty and diversification: the WGD radiation lag-time model. Curr. Opin. Plant Biol. 15, 147–153.

Seiliez, I., Panserat, S., Lansard, M., Polakof, S., Plagnes-Juan, E., Surget, A., Dias, K., Larquier, M., Kaushik, S., Skiba-Cassy, S., 2011. Dietary carbohydrate-to-protein ratio affects TOR signaling and metabolism-related gene expression in the liver and muscle of rainbow trout after a single meal. Am. J. Physiol. Regul. Integr. Comp. Physiol. 300, R733–R743.

Shibata, Y., Paul-Prasanth, B., Suzuki, A., Usami, T., Nakamoto, M., Matsuda, M., Nagahama, Y., 2010. Expression of gonadal soma derived factor (GSDF) is spatially and temporally correlated with early testicular differentiation in medaka. Gene Expr. Patterns GEP 10, 283–289.

Star, B., Jentoft, S., 2012. Why does the immune system of Atlantic cod lack MHC II? BioEssays News Rev. Mol. Cell. Dev. Biol. 34, 648–651.

Suurvali, J., Jouneau, L., Thepot, D., Grusea, S., Pontarotti, P., Pasquier, L. Du, Ruutel Boudinot, S., Boudinot, P., 2014. The proto-MHC of placozoans, a region specialized in cellular stress and ubiquitination/proteasome pathways. J. Immunol. 193, 2891–2901.

Taylor, J.S., Van de Peer, Y., Meyer, A., 2001. Genome duplication, divergent resolution and speciation. Trends Genet. 17, 299–301.

The_MHC_sequencing_consortium, 1999. Complete sequence and gene map of a human major histocompatibility complex. Nature 401, 921–923.

Thorgaard, G.H., 1976. Robertsonian polymorphism and constitutive heterochromatin distribution in chromosomes of the rainbow trout (*Salmo gairdneri*). Cytogenet. Cell Genet. 17, 174–184.

Thorgaard, G.H., Bailey, G.S., Williams, D., Buhler, D.R., Kaattari, S.L., Ristow, S.S., Hansen, J.D., Winton, J.R., Bartholomew, J.L., Nagler, J.J., Walsh, P.J., Vijayan, M.M., Devlin, R.H., Hardy, R.W., Overturf, K.E., Young, W.P., Robison, B.D., Rexroad, C., Palti, Y., 2002. Status and opportunities for genomics research with rainbow trout. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 133, 609–646.

Vallejo, R.L., Palti, Y., Liu, S.X., Evenhuis, J.P., Gao, G.T., Rexroad, C.E., Wiens, G.D., 2014. Detection of QTL in Rainbow Trout Affecting Survival When Challenged with *Flavobacterium psychrophilum*. Mar. Biotechnol. 16 (3), 349–360.

Verrier, E.R., Dorson, M., Mauger, S., Torhy, C., Ciobotaru, C., Hervet, C., Dechamp, N., Genêt, C., Boudinot, P., Quillet, E., 2013. Resistance to a rhabdovirus (VHSV) in rainbow trout: identification of a major QTL related to innate mechanisms. PLoS ONE 8 (2), e55302.

Viegas, I., Rito, J., Jarak, I., Leston, S., Caballero-Solares, A., Metón, I., Pardal, M.A., Baanante, I.V., Jones, J.G., 2015. Contribution of dietary starch to hepatic and systemic carbohydrate fluxes in European seabass (*Dicentrarchus labrax* L.). Br. J. Nutr. 113, 1345–1354.

Volff, J.-N., 2005. Genome evolution and biodiversity in teleost fish. Heredity 94, 280–294.

Weinstein, J.A., Jiang, N., Fisher, D.S., Quake, S.R.S., White, III, R.A., 2009. High-throughput sequencing of the zebrafish antibody repertoire. Science 324, 807–811.

Yano, A., Guyomard, R., Nicol, B., Jouanno, E., Quillet, E., Klopp, C., Cabau, C., Bouchez, O., Fostier, A., Guiguen, Y., 2012. An immune-related gene evolved into the master sex-determining gene in rainbow trout, Oncorhynchus mykiss. Curr. Biol. 22, 1423–1428.

Yano, A., Nicol, B., Jouanno, E., Guiguen, Y., 2014. Heritable targeted inactivation of the rainbow trout (*Oncorhynchus mykiss*) master sex-determining gene using zinc-finger nucleases. Mar. Biotechnol. 16, 243–250.

Young, W.P., Wheeler, P.A., Fields, R.D., Thorgaard, G.H., 1996. DNA fingerprinting confirms isogenicity of androgenetically derived rainbow trout lines. J. Hered. 87, 77–80.

Young, W.P., Wheeler, P.A., Coryell, V.H., Keim, P., Thorgaard, G.H., 1998. A detailed linkage map of rainbow trout produced using doubled haploids. Genetics 148, 839–850.