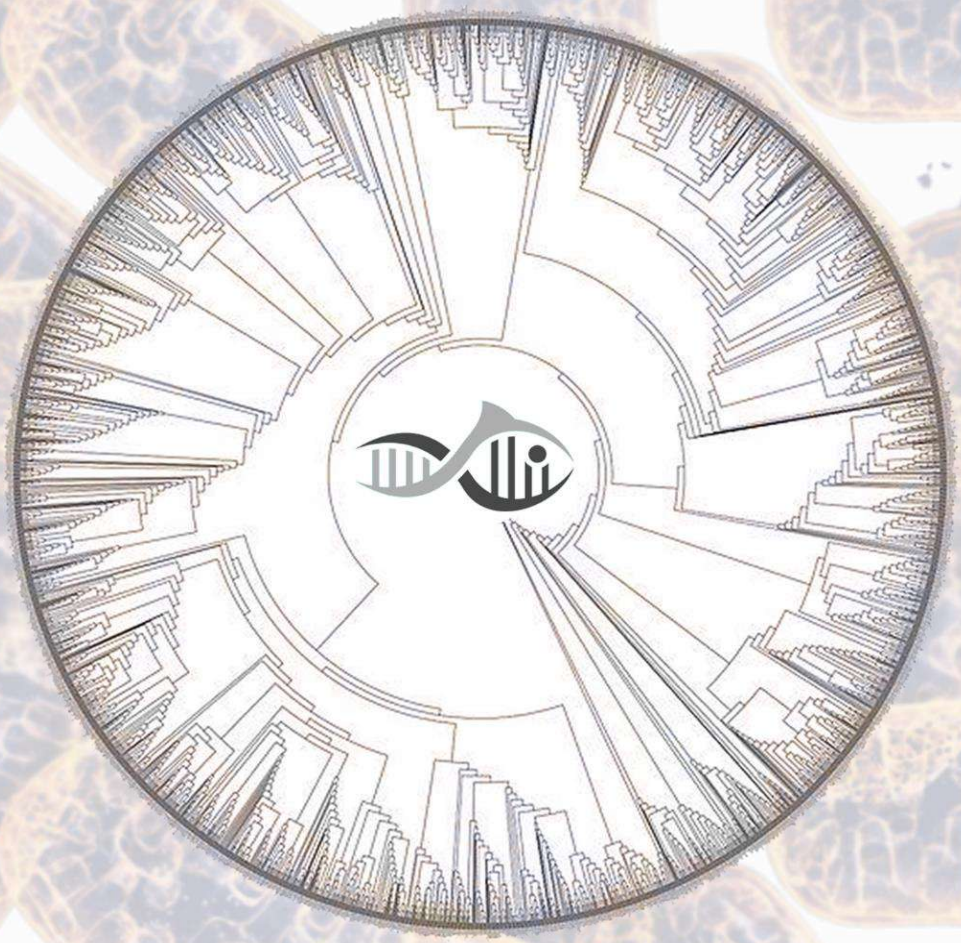
University of Thessaly

Department of Biochemistry and Biotechnology

PhD Thesis

Evolution of cellular energetics: comparative genomics and functional analysis

Andreas Tsipourlianos



2023

Evolution of cellular energetics: comparative genomics and functional analysis

Three-member Advisory Committee

Moutou Katerina, Professor, Vertebrate Biology, University of Thessaly (UTH)

Mamuris Zissis, Professor, Animal Population Genetics, University of Thessaly (UTH)

Psarra Anna-Maria, Associate Professor, Biochemistry, University of Thessaly (UTH)

Members of Seven-Member Examining Committee

Sarafidou Theologia, Assistant Professor, Animal Molecular Genetics, University of Thessaly (UTH)

Manousaki Tereza, Assistant Researcher, Genomics & Bioinformatics, Hellenic Centre for Marine Research (HCMR)

Power Deborah, Professor, Aquaculture and Biotechnology, Universidade do Algarve, Portugal (UAlg)

Cardoso João, Assistant professor, Comparative Genomics & Evolutionary Biology Universidade do Algarve, Portugal (UAlg)

Tsipourlianos Andreas

Evolution of cellular energetics: comparative genomics and functional analysis

Tsipourlianos Andreas

2023

Evolution of cellular energetics: comparative genomics and functional analysis

PhD thesis, University of Thessaly

Number of preliminary pages: 19

Total number of pages: 190

Number of Tables: 11

Number of Images: 38

Number of Literature references: 234

Abstract

Oxidative phosphorylation (OXPHOS) supplies over 90% of the cellular ATP requirements through the orchestrated function of five multiprotein complexes located in the inner mitochondrial membrane. The thesis explores how the evolutionary constraints imposed by the indispensable function of OXPHOS guide the outcome of Whole Genome Duplication (WGD) events. Two teleosts, the gilthead seabream (*Sparus aurata*, GSB) and the European seabass (*Dicentrarchus labrax*, ESB), which have undergone three rounds of WGDs were used as models during the highly plastic and energy-demanding period of early development. Fish ontogeny is a unique period in fish life marked with dramatic changes in morphology, physiology, metabolism and behavior. The tightly regulated landscape of cell divisions, migrations and differentiation driving these dramatic changes demands high energy supplies. The discovery of 24 and 22 OXPHOS gene families in GSB and ESB genome, respectively, and the subsequent phylogenetic analysis showed in most cases divergence from a common ancestor at the base of the teleost lineage, a process attributed to teleost-specific WGD. Overall results indicate that the WGD events have resulted in early retention of OXPHOS paralogue genes and subsequent species- or lineage-specific losses. OXPHOS paralogue gene expression levels were compared following RNA sequencing within and between distinct developmental stages in GSB and ESB. Different expression patterns between paralogs were revealed; some of them were in dosage balance, others were expressed only in particular stage(s) and a lot of them were differentially expressed between stages. The results were validated in a focused study comparing stage- and tissue-specific expression of OXPHOS paralogs in ESB. Differences in both the number and location of SNPs were revealed between paralogs, after merging the RNA sequencing data with whole genome sequencing data. Mutations were mapped mostly in the UTRs and very few in the CDS. The ratio of non-synonymous to synonymous substitutions when comparing the CDS variants revealed Purifying and Neutral Selection in action, safeguarding protein structural integrity and/or function. Overall, regulatory neo/subfunctionalization of OXPHOS paralogs appeared as the evolutionary mechanism behind the retention of the paralogs in GSB and ESB genomes in favor of ontogenetic plasticity. Indeed, OXPHOS paralogs exhibited differential thermal plasticity in GSB that became evident under high energy demanding circumstances.

Acknowledgements

This thesis is the outcome of a long trip that started earlier when I joined the laboratory of genetics, comparative and evolutionary biology as an undergraduate student and the existence of my PhD thesis proves that I liked it. First, I want to express gratitude to those who supported me throughout these years by simply being part of my life, shared in both my worries and joys, standing by me through moments of anxiety and celebration of achievements. A special thanks goes to my mother, who steadfastly stood beside me, supporting my decisions. Without her unwavering support and guidance during crucial moments, I may not had achieved anything close to what I have. A special thanks to my love Nantia Savva that keeps supporting and loving me even when I may not had been the easiest person to deal with. And a big thank you to my friends.

Me being a member of this lab was one of the defining moments of my adult life. While it's impossible to say everything, I wish to express in this space, that the individuals I've encountered in this scientific community possess not only remarkable expertise and talents, but also outstanding character. They exhibit a strong sense of support, care and a willingness to generously share their knowledge and experiences. This has contributed in creating an academic environment that aligns with my belief in the common essence of such collaborative and nurturing atmosphere. All those people shaped me in numerous ways and had a huge impact on my decision to start a PhD and try to follow a career in research.

I wish to express my gratitude to the one and only “my dear” Professor Katerina Moutou. A brilliant scientist and a charismatic teacher that made me really proud when she accepted to be my PhD supervisor. There is no point in describing the things that I learned and how she helped me; I learned a tone next to her and still learning, while I strongly believe that there is no limit to that. In the course of my PhD we shared endless and exhausting samplings (good memories), conferences and presentations, talks about science and not only. Thank you from my heart for all and I hope I was worthy of your trust and effort. A big thank you to Zissis Mamuris who, besides his contributions in my PhD, was my supervisor for both my bachelor's and master's thesis. I express my gratitude to him for his unwavering support and guidance throughout those years. I would also like to extend my thanks to Anna-Maria Psarra and Theologia Sarafidou. They not only

are members of my Advisory and Examining Committee, but also significantly contributed to my PhD journey, providing invaluable assistance and sharing their expertise from their respective disciplines. Through the course of my PhD I met a lot of scientists from different disciplines and affiliations and I learned a lot from them too. I want to express my gratitude to Deborah Power for being a member of my examining committee, for the always pleasant and informative conversations and even for giving me a push to actually start writing my thesis. I would also like to extend my thanks to João Cardoso for being in my Examining Committee; he helped me significantly at the beginning of my PhD, introduced me to phylogenetics and comparative genomics, and has his fair share to those good samplings' memories. I want to thank Tereza Manousaki for being in my Examining Committee, helping me to better understand genomics and providing me the tools to be more self-confident in front of the black screen of Unix environment. Lastly I want to thank Professor George Koumoundouros from the University of Crete and Research Director Nikos Papandroulakis from the Hellenic Centre for Marine Research (HCMR) that helped me understand better the physical subject of my PhD.

And now all those people that from different positions accompanied me in this journey: Kostas Stamatis who has an accurate solution for every problem; he helped me a lot and he played a significant role in teaching me and a lot other people how to be helpful and giving in a working environment, Eleni Galiopoulou, Dina Sarri and Maria Kalemkeridou for their help, support and all the good times shared at the lab, Lamprini Tzioga a great colleague and one of the most giving and kind persons that I know and of course a big thank you to the Assistant Professor Themis Giannoulis, a dear friend of mine from before I even thought about doing a PhD and a great colleague who introduced me to all of this; I have learned a lot working with him and he influenced me in numerus ways and because its impolite to only take and not give, I taught him how to be calm and to manage difficulties with composure, Rafael Angelakopoulos the next PhD candidate to the line; a good friend and a great colleague that we shared a lot and had the misfortune of being the one with expertise in Photoshop (thank you and sorry for that), Alexia Fytsili who I met during her bachelor thesis that contributed to my PhD and became one valuable colleague and our junior PhD candidate who helped me a lot (especially when we hunted wasps), and lastly Artemis Kotoula who had the same training as Themis on how to calmly manage

difficulties by observing me and contributed in my PhD through her bachelor thesis (and now is training to become top bioinformatitian).

Thank you all for your presence and support.

Andreas

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Cellular energetics and Oxidative Phosphorylation

Cellular energy production is based on the breakdown of carbohydrates, fats and proteins to release chemical energy that is ultimately captured in ATP molecules. Cellular catabolic processes are gradual and are orchestrated around the catabolism of glucose, fatty acids and amino acids that produce organic substrates for the citric acid (TCA) cycle. Entering the citric acid cycle, the organic substrates are successively oxidized, and the energy is stored in reduced coenzymes $FADH_2$ and NADH that act as electron carriers for the electron transport chain. Electron transport chain is fueling oxidative phosphorylation (OXPHOS), the main process for generating up to 90% of adenosine triphosphate (ATP) in the cell (Erecińska & Silver, 1989).

Oxidative phosphorylation is located within the mitochondria and ATP production through oxidative phosphorylation is ensured through the continuous entry into the mitochondria of respiratory substrates, ADP and Pi, and the output of ATP (Papa et al., 2012). The presence of oxidative phosphorylation in all higher animals and plants indicates its great importance for their existence, as well as that it has been a prerequisite for their evolutionary course (D. F. Wilson, 2017).

Oxidative phosphorylation involves five (5) protein complexes (**Figure 1.1**); four (4) of them build the respiratory chain (complexes I-IV), responsible for creating the proton gradient, while the last complex (complex V) is the ATP synthase, which utilizes the proton gradient to synthesize ATP. The OXPHOS complexes (**Figure 1.1**) based on HUGO Gene Nomenclature Committee (Seal et al., 2023) are:

- complex I: NADH:ubiquinone oxidoreductase
- complex II: Succinate dehydrogenase
- complex III: Ubiquinol-cytochrome c reductase
- complex IV: Cytochrome c oxidase
- complex V: ATP synthase

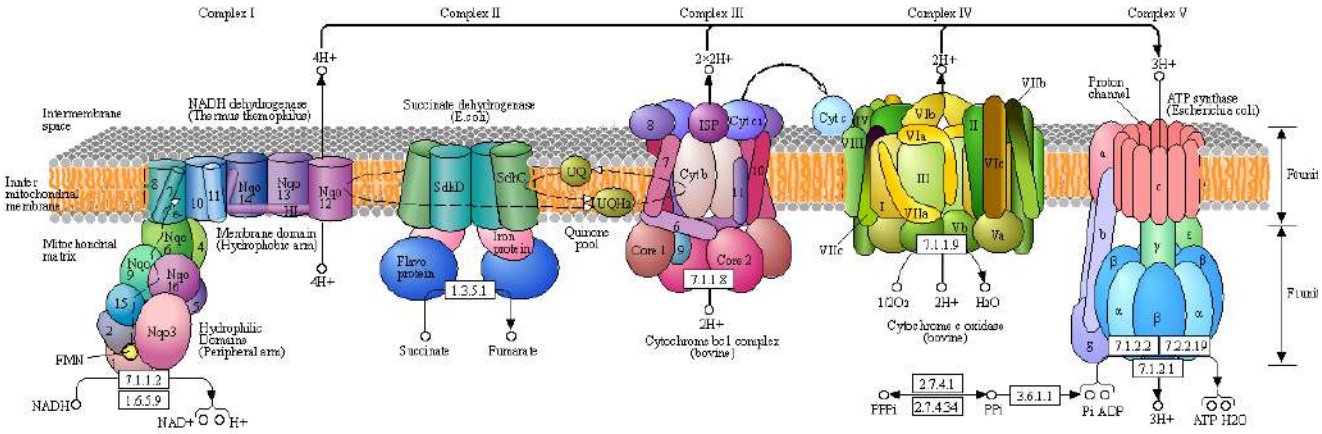


Figure 1.1. Oxidative phosphorylation involves five protein complexes located in the inner mitochondrial membrane, (Source: KEGG).

The appearance of aerobic respiration

Three billion years ago, the existence of life revolved around anaerobic microbes inhabiting an anoxic environment, and their metabolic processes involved a set of well-structured anaerobic pathways (Falkowski, 2006). The elevated levels of carbon dioxide in gaseous form (dissolving to form carbonic acid) would have resulted in an ocean with increased acidity. This inherent acidity transformed alkaline vents into chemiosmotic systems, creating a natural proton gradient across the inorganic membranes. This gradient, likely, generated a positive proton force, comparable in magnitude to the gradients of biological membranes (Lane et al., 2010). Mitchell (Mitchell, 1961) proposed the idea that the energy derived from the oxidation of NADH is utilized to generate a gradient of proton concentration across a membrane. This gradient, in turn, drives the synthesis of ATP through proteins initially labeled 'coupling factors', now commonly known as ATP synthases. The remarkable feature of chemiosmosis lies in the universal and fundamental nature of proton gradients across all life. When cyanobacteria emerged, along with photosynthesis, utilizing the energy of the sun and using water as a reductant to fuel metabolism, O_2 was produced as a metabolic by-product (Rasmussen et al., 2008) with life-changing consequences. Oxygen eventually overwhelmed the existing atmosphere driving the evolution of single-cell eukaryotes and eventually pushing the development of more complex life forms that populate

the planet today. The great oxidation event that took place 2.4 billion years ago, facilitated the emergence of eukaryotes (Erwin et al., 2011), whereas the Neoproterozoic oxidation event in the later Neoproterozoic Era has been associated with large and complex multicellular organisms (Sperling et al., 2015).

The respiratory chain

Keilin in 1966 (Keilin & Keilin, 1970) was the first to define the respiratory chain, following the discovery of cytochromes a,b,c, as redox carriers in aerobic organisms that sequentially link Wieland's activated dehydrogenase with Warburg's oxygen-activating enzyme (Papa et al., 2012). This chain is located in the inner mitochondrial membrane and uses the electrons of NADH and $FADH_2$, which are ultimately transferred to molecular oxygen. For each electron pair, the Gibbs free energy change is equal to $\Delta GO' = -52.5$ kcal/mol. Only with the use of redox carriers and the passage of electrons through them, can free energy become exploitable. The carriers are organized in four complexes in the inner membrane of the mitochondria, that make up the electron transport chain or otherwise the respiratory chain (Nolfi-Donegan et al., 2020). Initially, complex I and II receive electrons from the reducing equivalents (NADH, $FADH_2$) originating from the TCA cycle. Subsequently, the electrons are transferred to ubiquinone, a hydrophobic quinone, which can diffuse freely within the inner mitochondrial membrane and transfers electrons to complex III. From there, electron transfer to complex IV takes place via a membrane protein (cytochrome c) attached to the outer side of the inner membrane. After the electrons have been transferred to complex IV, oxygen is reduced to water. Parallel to transporting electrons, complexes I, III, and IV transport protons to the transmembrane space of the mitochondrion (**Figure 1.2**). For each pair of electrons transferred down the chain, a total of ten protons passes from the mitochondrial layer into the transmembrane space. Four of them come from complex I. Four protons are released into the transmembrane space by complex III; two of them are drawn from the mitochondrial matrix and two more protons come from coenzyme Q that performs the transfer from either complex I or II. Although complex IV pumps four protons, two of them will be used to reduce oxygen to water and the remaining two will pass to the

opposite side of the membrane. Fewer ATP molecules are formed from the oxidation of $FADH_2$ than from the oxidation of NADH as succinate-Q-couple reductase (complex II), unlike NADH-Q-couple oxidoreductase (complex I), does not pump protons from one side of the membrane to the other (Wikström et al., 2015). The resulting proton gradient across the inner membrane, from the respiratory chain (complex I-IV), will be later used by complex V to produce ATP (F. Zhang & Broughton, 2013).

The size, morphology and composition of the inner mitochondrial membrane is closely linked to oxidative phosphorylation. The inner membrane extends to contain numerous folds called mitochondrial cristae, the number of which is directly linked to oxidative phosphorylation (Cogliati et al., 2016). The inner membrane contains an apparently high number of proteins. More than 70% of these proteins play a role in oxidative phosphorylation and transport of metabolites between cytoplasm and mitochondria. One of the most important properties of the inner mitochondrial membrane is the fact that it is impermeable to most ions and small molecules. This makes it capable of creating the proton gradient that will drive oxidative phosphorylation to

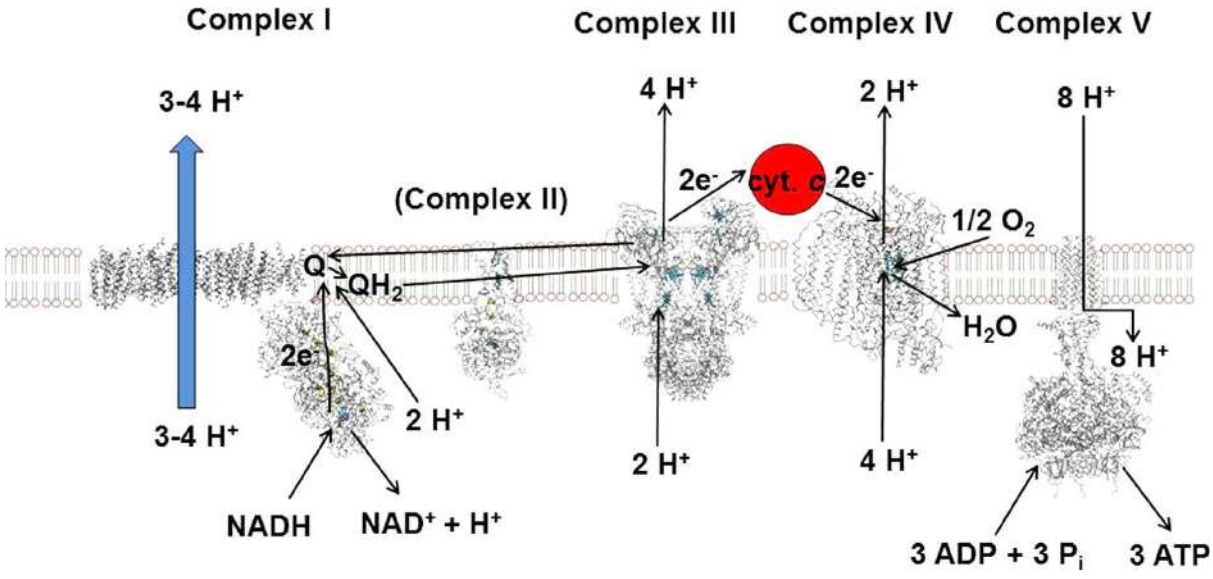


Figure 1.2 Schematic presentation of electron transport and proton pumping across the OXPHOS complexes, (Source: Wikström et al., 2015)

produce ATP. In contrast, the outer membrane contains a large number of porins that create freely permeable channels for ions and small, uncharged molecules (Bayrhuber et al., 2008).

The mitochondrial proteins are the products of two genomes, the mitochondrial and the nuclear genome. Mitochondria, besides their own genome, possess a translatory machinery and they produce a number of their proteins. Mitochondrial function and oxidative phosphorylation in particular, rely on the cooperation and combination of proteins encoded by the nucleus and proteins encoded by the mtDNA itself and translated within the organelle.

Metazoan mitochondria encode 13 proteins of the electron transport chain and the oxidative phosphorylation, as well as two rRNA molecules (16S and 12S) and 22 tRNA molecules, which are used in the translation of these 13 proteins (F. Zhang & Broughton, 2013). More than 95% of the total 1,500 mitochondrial proteins are coded by the nuclear genome and are translated in free ribosomes in the cytoplasm and enter the mitochondrion as complete polypeptide chains under the guidance of signal amino acid sequences. These proteins are involved in the oxidative phosphorylation and in mitochondrial metabolism along with the replication, transcription and translation of mitochondrial DNA (Fox, 2012).

Entry of proteins into the mitochondrion

The entry of the mitochondrial proteins translated in cytoplasmic ribosomes takes place through the protein complex of the outer mitochondrial membrane, TOM (**Figure 1.3**). This complex consists of the proteins Tom40, Tom20, Tom70 and Tom22. Tom20 receptor recognizes immature proteins, which have a mitochondrial localization signal at their amino-terminus, while Tom70 receptor recognizes precursor molecules of inner membrane metabolic carrier proteins, which contain many signaling motifs. Precursor molecules recognized by these two receptors are transferred to Tom22 and from there to Tom40, which creates oligomeric channels for their entry into the mitochondrial transmembrane space. At least two different mechanisms for protein transport through the TOM complex have been described. After proteins pass into the transmembrane space through the TOM complex, some are recognized by two different TIM

(presequence translocase of the inner membrane) complexes of the inner mitochondrial membrane, through which they pass to their final destinations. Proteins that will enter the matrix are recognized by TIM23, while those destined for the inner membrane are recognized by TIM22. In addition to TOM and TIM, the OXA protein is also located in the inner mitochondrial membrane, which imports proteins synthesized in the mitochondrial matrix into the inner membrane (Chacinska et al., 2009).

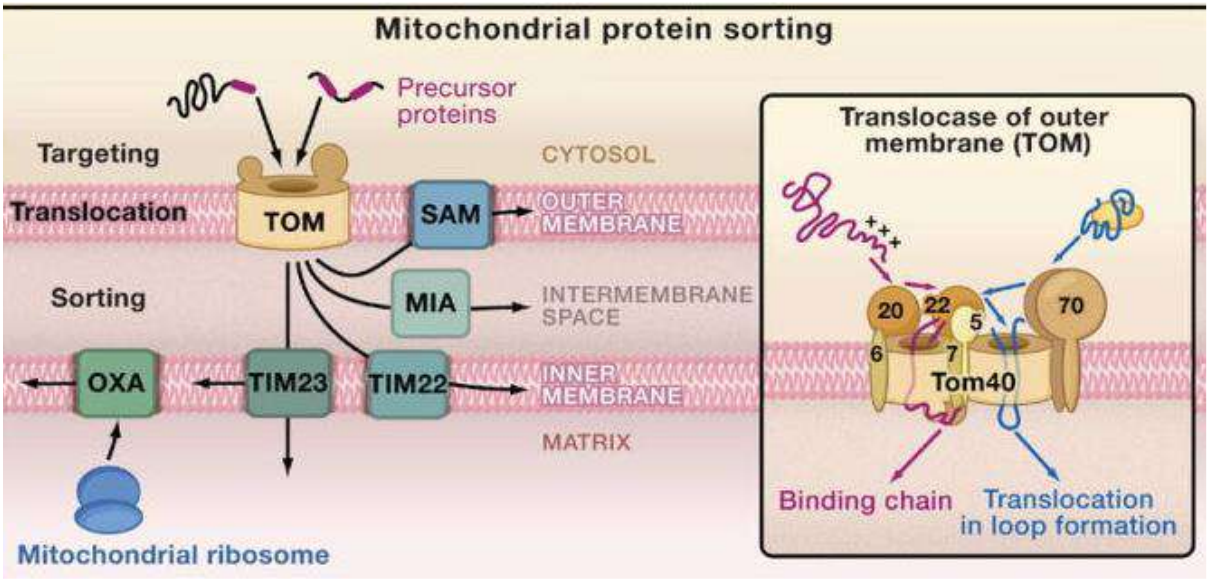


Figure 1.3 Sorting pathways of mitochondria. The translocase of the outer membrane (TOM complex) is the main entry gate into mitochondria. Subsequently, the precursor proteins follow different sorting pathways. MIA, mitochondrial intermembrane space assembly; OXA, insertase/export machinery of the inner membrane; SAM, sorting and assembly machinery; TIM22 complex, carrier translocase of the inner membrane; TIM23 complex, presequence translocase of the inner membrane. (Inset) The TOM complex consists of seven different subunits. The receptors Tom20, Tom22, and Tom70 recognize precursor proteins and transfer them to the central component, the channel-forming Tom40. Three small Tom proteins, Tom5, Tom6, and Tom7, are involved in the assembly and dynamics of the TOM complex. (Presequence-carrying precursor proteins, red; hydrophobic precursors with internal targeting signals, blue); (Source: Chacinska et al., 2009).

Mitochondrial localization signals

Depending on the final destination of the proteins within the mitochondria, different signal variations are found in their sequence, which are necessary not only for the passage into the interior, but also for targeting to their final destination. The signal sequences (**Figure 1.4**) are located at the amino terminus of precursor proteins, form amphipathic α-helices, and are usually 15 to 20 amino acids in length. Amphipathic α-helices are recognized by the Tom20 complex through the Tom20 motif, a hydrophobic surface of the protein (Saitoh et al., 2007). After entering the mitochondrion, most signals are proteolytically removed by the mitochondrial peptidase MPP (Mitochondrial Processing Peptidase) dimer. Proteins destined for the mitochondrial matrix contain only the signal of the amphipathic α helix. Proteins of the inner membrane, in addition to the α helix, contain after the cut site by the MPP another signal peptide, the sorting signal, which is not removed from the protein, but is used to attach the protein to the membrane. The signaling sequences of transmembrane proteins differ from those of inner membrane proteins in that their sorting signal is truncated. This incision is carried out by a different peptidase, IMP (Inner Membrane Peptidase) (Chacinska et al., 2009).

In addition to signal sequences that are excised, there are other types of targeting signals that are located within the peptide chains of mature mitochondrial proteins. β-Barrel proteins include a β signal (β signal) at their carboxy terminus, which is recognized by the outer membrane complex SAM (Sorting and Assembly Machinery), after passing through the TOM into the transmembrane space. Thus, SAM promotes β-barrels to the outer membrane (Wiedemann & Pfanner, 2017). In α-helices of the outer membrane, signals have been identified in the amino-terminus (Signal-anchor sequence), in the carboxy-terminus (C tail anchor) and inside the peptide sequence (Internal signal). Protein transporters of metabolic substances of the inner membrane may contain multiple internal signals (Multiple internal signals), while other types of inner membrane proteins contain after a hydrophobic sequence, a positively charged signal (Presequence-like internal signals). Finally, another signal that is not cut off from the peptide sequence directs the proteins to the transmembrane space. This signal contains a cysteine residue that interacts with the transmembrane space receptor MIA (Chacinska et al., 2009).

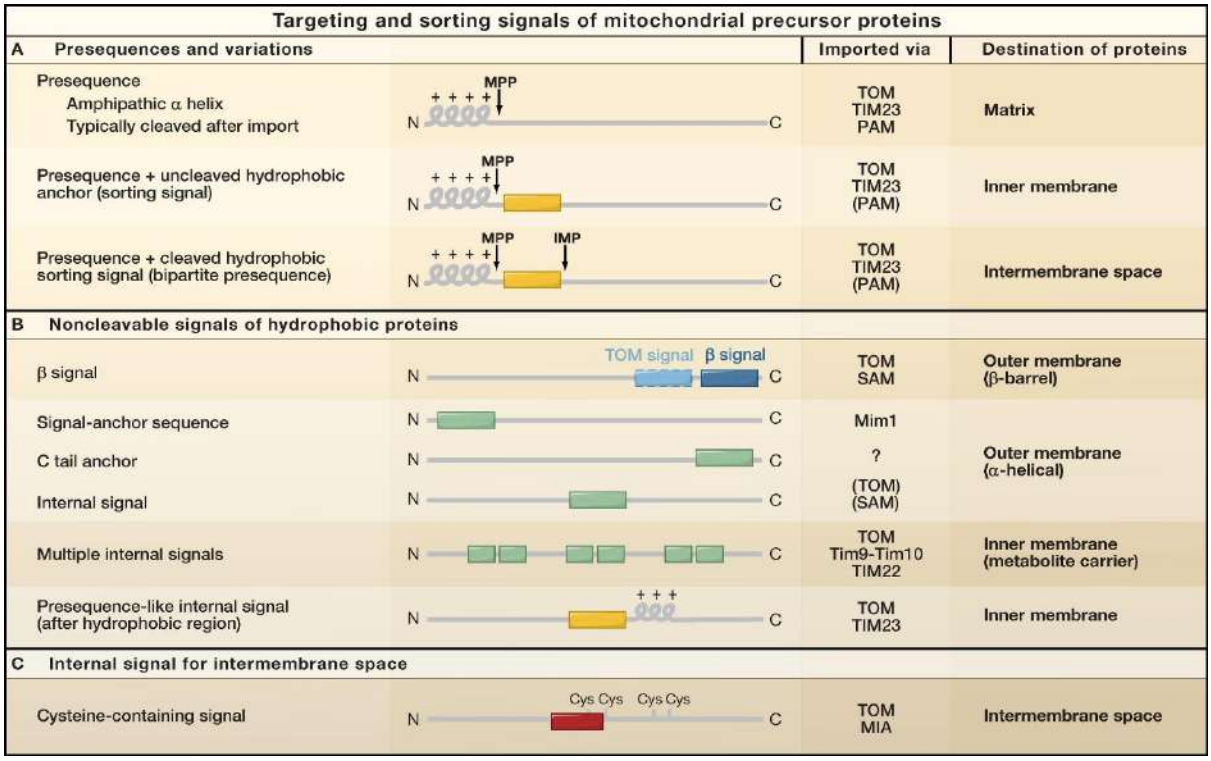


Figure 1.4. Targeting and Sorting Signals of Mitochondrial Precursor Proteins (Source: Chacinska et al., 2009).

Structure and function of OXPHOS complexes

Complex I (NADH:ubiquinone oxidoreductase)

Complex I (**Figure 1.5**) is one of the largest known membrane protein complexes and is the largest of the four in the respiratory chain. It is "L" shaped and consists of two elongated arms. The hydrophobic arm is embedded in the inner mitochondrial membrane and the hydrophilic arm protrudes from the matrix side. The complex is composed of 14 main subunits in bacteria, 7 of which are found in the hydrophilic arm and the rest in the hydrophobic arm. These subunits represent the minimal core of the complex required for electron transport and are all conserved in eukaryotes. In addition to the basic subunits, the complex contains a large number of auxiliary subunits, which can constitute up to 50% of the total mass of the complex, and their role is either

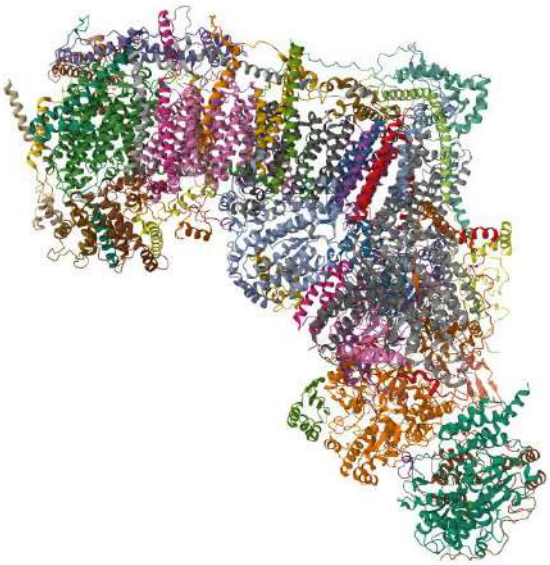


Figure 1.5. Cryo-EM structure of human respiratory complex I (PDB); (Source: Runyu Guo et al., 2017).

regulatory or they participate in the biogenesis of the complex (Berrisford et al., 2016; Janssen et al., 2006). NADH electrons enter the respiratory chain via the NADH-Q couple redox. First, NADH binds to the complex and a two-electron transfer occurs to the flavin mononucleotide (FMN) group of the complex, which is located on top of the hydrophilic arm, thus forming its

reduced form. From there the electrons are transferred across a series of 7 iron-sulfur complexes that extend at a distance of about 90 Å to the ubiquinone located in the binding pocket formed by the NU1M, NDUFS2, and NDUFS7 subunits at the base of the hydrophilic arm (Xu et al., 2020; Zickermann et al., 2009).

In the membrane arm then, four structural motifs are found and are likely to transport one proton each per catalytic cycle, from the mitochondrial matrix to the transmembrane space. The proton pumping mechanism is not only spatially separated but also occurs with a noticeable time difference. However, the exact mechanism of coupling between electron transport and proton pumping is still the least understood aspect of complex I (Wirth et al., 2016).

Electron transfer to Q leads to its formation, the negative charges of which interact electrostatically with negatively charged amino acids of the hydrophilic arm leading to changes that alter the helix structures and promote the pumping of 4 hydrogen ions. Finally, Q^{-2} accepts two protons from the matrix and QH_2 is formed leaving the enzyme complex for the Q pool. Overall, the complex is divided into three parts based on its functional characteristics; The hydrophilic arm accommodates the elements Q (Q module) and N (N module), which are two of the three functional parts of the complex. The N element consists of the subunits NDUFS1, NDUFV1 and NDYFV2 and is the point of entry of electrons from NADH into the respiratory chain. Element Q is made up of subunits NDUFS2, NDUFS3, NDUFS7, NDUFS8 and accepts electrons from iron sulfur complexes of element N19 and transfers them to ubiquinone. Finally, the remaining seven core subunits of the complex make up the P element, which includes the membrane arm and constitutes the proton pumping region of the complex (Wirth et al., 2016). The *de novo* synthesis of complex I occur through the assembly of intermediate complexes. Initially, the mtDNA-encoded NUM1 subunit joins with other subunits from either the nuclear or mitochondrial genome, such as NDUFS2 and NDUFS7, and form intermediate assembly complexes. NDUFS4 and NDUFS6 are also required for the assembly and stabilization of a part of complex I that contains a number of subunits, including NDUFV1, NDUFV2, and NDUFV3 (Fernandez-Vizarra & Zeviani, 2018).

Complex II (Succinate dehydrogenase)

Unlike the rest of the respiratory chain complexes, which display additional subunits in eukaryotes, complex II (**Figure 1.6**) has the same composition (Iwata et al., 1998) in mitochondria and prokaryotes. It consists of two hydrophilic proteins, a flavoprotein (Fp) containing a FAD cofactor and an iron-sulfur protein (Ip), and a mitochondrial membrane anchoring domain, which consists of two transmembrane proteins, CybL and CybS, with a heme group united. The membrane domain is less conserved than the catalytic subunits and may consist of one or two proteins with one, two, or no heme groups. The 68-kDa flavoprotein has a Rossmann fold and contains FAD cofactor, the electron acceptor for succinate. The 29-kDa Ip protein contains three iron-sulfur complexes, [2Fe-2S], [4Fe-4S], and [3Fe-4S], and forms the core of the complex. It binds to Fp on one side and to membrane anchor proteins on the other. All subunits of the complex are encoded by nuclear DNA (Rutter et al., 2010). The subunits are individually prepared for assembly, which is mediated by specific chaperone proteins. For the complex to be functional, all its subunits must be assembled (Fernandez-Vizarra & Zeviani, 2018).

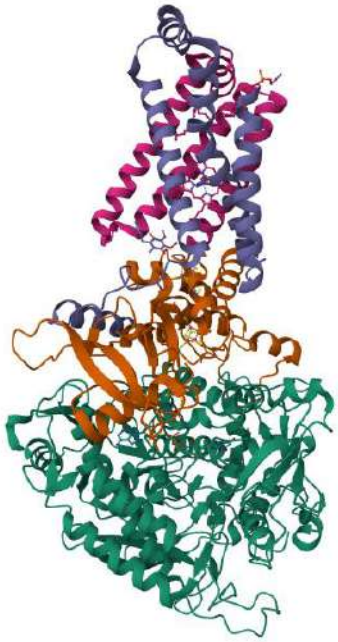


Figure 1.6. Cryo-EM structure of the human respiratory complex II (PDB); (Source: Zhanqiang Du et al., 2023).

Complex II is the entry point of $FADH_2$ into the respiratory chain. $FADH_2$ is produced in the citric acid cycle during the oxidation of succinate to fumarate, a reaction catalyzed by succinate dehydrogenase. Electrons from $FADH_2$ are transferred to iron sulfur centers and finally to ubiquinone. Complex II, in contrast to the other three complexes, is not a proton pump, therefore, during the transfer of electrons from $FADH_2$ to QH_2, no protons are transferred to the transmembrane space. As a result, less ATP is produced from the oxidation of $FADH_2$ compared

with the oxidation of NADH. Electrons from QH_2 are subsequently transferred to complex III (Rutter et al., 2010).

Complex III (Ubiquinol-cytochrome c reductase)

Complex III (**Figure 1.7**) is a symmetric dimer, with three subunits (cytochrome b, cytochrome c1, and the iron-sulfur 'Rieske' subunit) that are highly conserved from bacteria to mammals and form the core of the complex (Iwata et al., 1998). The conserved, catalytically active subunits are cytochrome b (cyt b), with two b-type heme molecules (low- and high-affinity bL and bH, respectively), and cytochrome c1 (cyt c1), with one c-type heme molecule. Cyt b is the only complex III protein encoded by the mitochondrial genome and harbors two distinct ubiquinone binding sites on opposite sides of the membrane (Q_0 and Q_i) (Xia et al., 1997). Also conserved is an iron sulfide protein, UQCRFS1 (ISP) with a 2Fe-2S center called the Rieske center (Tang et al., 2020). The one iron ion of the Rieske center forms a binding complex with two histidine amino acids thereby stabilizing the center in its open form. Consequently, its reduction potential increases so that it can more easily accept electrons from QH_2. In addition to the three catalytic subunits, there are another eight subunits that do not participate in electron transport or proton pumping. Their role concerns the assembly, stability and regulation of the activity of the complex (Iwata et al., 1998).

The ubiquinol molecules produced by complexes I and II transfer electrons to cytochrome c with the involvement of complex III. The mechanism initially involves the sequential binding of two QH_2 molecules to the complex. During this binding, two protons are released into the transmembrane space and two electrons. QH_2 binds to the first binding site of cytochrome b (Q0) and the electrons follow two different directions. In the first half of the cycle an electron is directed to the Rieske center, then to cytochrome c1 and finally to a cytochrome c molecule converting it to its reduced form. Subsequently, this reduced cytochrome c molecule can be removed from the enzyme continuing its course in the respiratory chain. The second electron is transferred via the heme group of cytochrome b to the second binding site (Q_i) and a semiquinone radical anion is formed. In the second half of the cycle a second QH_2 molecule

transfers one electron to another cytochrome c molecule and the other electron is used to reduce the hemiquinone to QH_2 with the uptake of two protons. This second electron transfer leads to the extraction of two protons from the matrix (Demin et al., 1998).

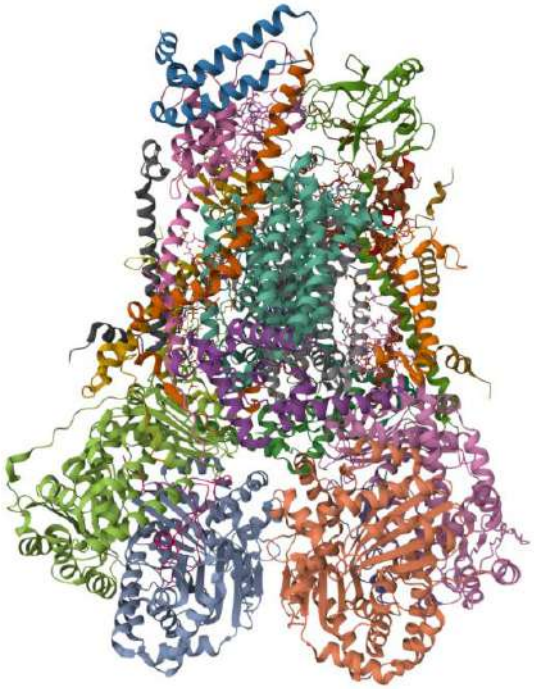


Figure 1.7. Cryo-EM structure of human respiratory complex III (PDB); (Source: Runyu Guo et al., 2017).

Assembly of complex III begins with the synthesis of cytochrome b (MT-CYB in humans) from the mitochondrial ribosome and its transport to the inner membrane, mediated by Cbp3/UQCC1 and Cbp6/UQCC2. Once the first structural subunits (UQCRB and UQCRQ) are incorporated, the UQCC1-UQCC2 factors dissociate and return to function as cytochrome b translational activators. CYC1, UQCRH, UQCRC2 and UQCRC1 subunits are then added to the complex and dimerization occurs. After dimerization, the UQCR10 subunit is added. The complex matures with the addition of the Rieske Fe-S protein (UQCRFS1) and the smaller subunit (UQCR11) to the dimeric precursor. After import into mitochondria, UQCRFS1 is bound and stabilized in the matrix by the factor

LYRM7. Incorporation of UQCRFS1 into precursor complex III is mediated by Bcs1. This step is considered critical for the catalytic activation of the complex (Fernandez-Vizarra & Zeviani, 2018).

Complex IV (Cytochrome c oxidase)

Complex IV (**Figure 1.8**) is composed of 14 subunits in mammals, three of which are encoded by the mitochondrial genome. These are subunits I, II, III and are also found in the bacterial enzyme. Subunit I, located mainly in the transmembrane region, consists of 12 transmembrane helices and a small extramembrane part, while containing three of the four cofactors of the complex, two heme groups (α, α3) and a CuB copper ion. The fourth cofactor consists of two CuA copper ions and is found in subunit II. This subunit is anchored to the membrane via two transmembrane helices and it also has a region in the transmembrane space to which cytochrome C binds. Subunit III has seven transmembrane helices, stabilizes the other two domains, and acts as an initial proton acceptor (Kadenbach et al., 1983; Zong et al., 2018).

In addition to the basic subunits, the complex also has auxiliary subunits that strongly suppress its activity and have a fundamental role in regulation. Subunit IV is the largest of the accessory subunits and binds ATP, being responsible for the ATP-dependent allosteric inhibition of the complex when the ATP/ADP ratio is high (Fontanesi et al., 2008; Tsukihara et al., 1996). Complex IV is the last proton-pumping complex of the respiratory chain and the one that carries out the transfer of electrons from reduced cytochrome c to molecular oxygen (final acceptor). The complete reduction of oxygen to H_2O requires four electrons, and the requirement for molecular oxygen in this reaction is what makes organisms aerobic. To carry out the reaction, four cytochrome c molecules are sequentially attached to the complex and each one transfers an electron. Initially, two molecules of reduced cytochrome c successively transfer two electrons, one to the α3 heme and one to the CuB copper ion. The two centers are reduced and are able to bind an O_2 molecule, which upon its binding forms a peroxide bridge (Hosler et al., 1993). Then, two more molecules of reduced cytochrome c are attached to the oxidase following the same path and combined with the pumping of two protons, the peroxide bridge is broken. Finally, the addition of two more protons leads to the release of two water molecules. All four protons of the

reaction originate from the mitochondrial matrix, thus contributing to the proton concentration gradient on both sides of the membrane (Wikström et al., 2018).

Assembly starts from subunit I (MT-CO1) which forms the central backbone around which the complex is to form. The assembly of the complex involves several chaperone proteins and assembly factors that contribute to its maturation and stabilization and are known as 'MITRAC'. First, COX4I1 and COX5A come together, and then the assembly factor CMC1 binds COA3 and COX14 and integrates them into the complex. COX10 and COX15 are required for heme biosynthesis, while COX11, COX17 and COX19 participate in the incorporation of the CuB center.

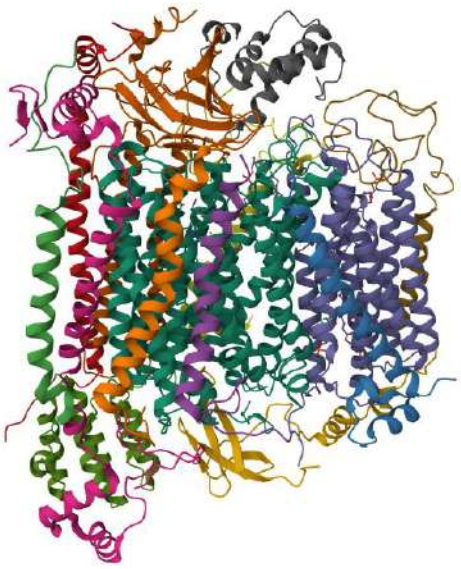


Figure 1.8. Structure of human cytochrome c oxidase (PDB); (Source: Zong et al., 2018).

In subunit II (MT-CO2) COX20, COX5B, COX6C, COX7C, COX8A and COX7B are assembled with the contribution of the assembly factor COX18. Incorporation of the CuA center into subunit II involves the copper binding factors COX17, SCO1 and SCO2 along with COA6 and COX16. Subunits I and II are then joined together, while lastly subunit III (MT-CO3) joins together with the protein NDUFA4, originally thought to be part of complex I (Fernandez-Vizarra & Zeviani, 2018).

ATP synthesis - Complex V ATP synthase

After the transfer of electrons from NADH and $FADH_2$ to molecular oxygen the next step in energy production is the synthesis of ATP. The energy stored in the proton gradient on both sides of the inner mitochondrial membrane is used to synthesize ATP from ADP. At this stage, the electrons in the matrix are fewer than those in the transmembrane space and therefore tend to move towards the matrix to equalize the concentrations on either side of the membrane. It is this flow of protons that triggers the ATP synthase complex (**Figure 1.9**) (Yoshida et al., 2001).

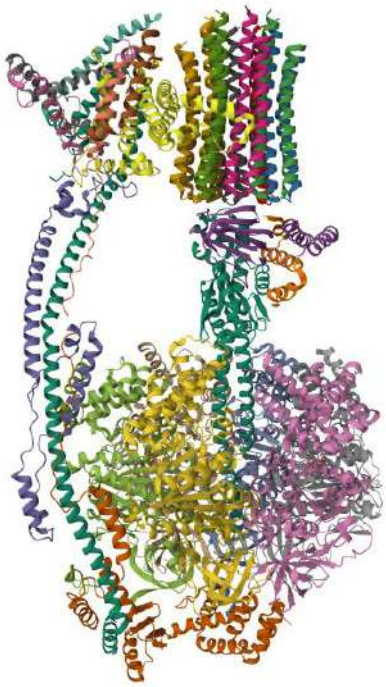


Figure 1.9. Human ATP synthase state open conformation (PDB); (Source: Lai, et al 2023).

ATP synthase is a large complex enzyme that is divided into two domains (F0 and F1). The F1 segment is the catalytic center, consisting of an alternating arrangement of three α and three β subunits, arranged around the γ subunit. The α and β subunits are homologous and have the same fold, but only the β subunit is catalytic. Apart from the three subunits, F1 has also subunits d and e, which participate in the connection with F0. The F0 segment is a hydrophobic region that spans the inner mitochondrial membrane. This part constitutes the proton channel and is composed of 8-14 subunits. In addition, there is subunit a, which is located outside the channel. The connection of the two parts (F1, F0) is carried out by the central stem of subunits γ, ε and externally through two subunits β and δ of the F1 part (Yoshida et al., 2001).

All subunits of the complex participate in ATP production. At rest, when there is no proton gradient, roughly equal amounts of ATP, ADP are bound to the catalytic center. Therefore, the role of the proton concentration gradient is not ATP synthesis but its release from the synthase. The enzyme can be divided into two functional components; the rotatable unit consisting of the

c-ring and the γε central stalk, and the static unit containing all other subunits. The catalytic centers of the enzyme are three, namely the three β subunits, which implement a different catalytic function at any time. The γ subunit located on the inner side of the three β subunits rotates, changing the conformation of the latter. In the L (loose) conformation it binds ADP and Pi, in the T (tight) conformation the β subunit shows a very high affinity for ATP resulting in its synthesis from ADP and Pi of the previous conformation. Finally, in the O (open) conformation the active center has the ability to bind and release adenine nucleotides (Dimroth, 2000).

The concentration gradient of protons and consequently their return to the mitochondrial matrix, creates the driving force for the rotation of the mobile unit which in turn induces ATP synthesis in the active center of the enzyme (Yoshida et al., 2001). Complex V is the smallest molecular motor known (Noji et al., 1997).

Building of supercomplexes

The complexes of the respiratory chain interact with each other to form higher order structures, which have been called supercomplexes. The resulting supercomplexes display various stoichiometries. Complexes I, III, IV are the ones that mainly participate in the creation of supercomplexes, while complexes II and V build oligomers without interaction with the other complexes.

The assembly of supercomplexes may be explained by two possible mechanisms. The first postulates that the individual complexes are fully assembled before joining to form the supercomplexes. This mode of action will allow the dynamic assembly-dissociation of individual complexes to adapt to different energy requirements. However, there is also evidence indicating the simultaneous assembly and formation of supercomplexes by the various individual enzymes before the completion of the formation of each individual complex. This hypothesis is supported by the fact that the maturation of complex I is not completed unless complexes III and IV are first bound to a precursor complex. Also, incompletely assembled complexes have been found in

tissues from patients carrying mutations in different structural subunits and assembly factors involved in the final steps of complex I and III assembly (Fernandez-Vizarra & Zeviani, 2018).

Gene and Genome duplication

Duplication of genes and whole genomes has been attributed an important role in evolution since the 1930s (Bridges, 1936; Ohno, 1970; Ohno et al., 1967; Stephens, 1951). In 1936, Bridges made the first observation of gene duplication following the duplication of a chromosomal region in a mutant of the fruit fly *Drosophila melanogaster*, which exhibited an extraordinary reduction in eye size, while Ohno in 1970 suggested that redundant genes are likely candidates for the provision of initial genetic material for the purpose of evolutionary innovation. Also, in his book "Evolution by Gene Duplication" he notes that "*if evolution were entirely dependent on natural selection, only numerous forms of bacteria would arise from one bacterium. The creation of vertebrates and eventually mammals from single-celled organisms would be completely impossible. Such great leaps in evolution require the creation of new gene loci with a previously nonexistent function. Only the gene that became redundant could escape the relentless pressure of natural selection. After it escaped, deleterious mutations accumulated in it to emerge as a new gene locus*" (Ohno, 1970). Thus, increasing the number of copies of a gene in a cell by gene duplication or by chromosome duplication, is an important mechanism by which new genetic material is produced during evolution. It can be defined as any repeat of a region of DNA that contains a gene. Gene duplications can arise as products of various types of errors during DNA replication and repair (J. Zhang, 2003).

Five molecular mechanisms by which DNA is duplicated have been identified (Hahn, 2009; W.-H. Li, 1997; Lynch, 2002):

1) Unequal crossing over is a reciprocal recombination between homologous chromosomes or sister chromatids (DNA double helix), in which no genetic material is lost. Unequal splicing changes the number of repeats affecting gene 'dosage'. It occurs during deceleration but also in

the resting interphase nucleus. It concerns areas that are the same or very similar to each other with a length that is not necessarily long. Interchromosomal recombination occurs between two short, repetitive DNA sequences located at either end of a gene. Once a gene is duplicated in this way, subsequent unequal crossovers can easily add more copies by matching a gene on one chromosome with its genetic copy on the second chromosome. It is the basic mechanism of gene duplication and creation of a whole group of highly related genes arranged in series, such as the globin gene family and other multigene families.

Non-allelic homologous recombination (NAHR), is the genesis of large rearrangements in homologous segments, with a frequency of 10^{-4} per generation. Rearrangements can be created as a result of gene recombination and gene copy number variations can occur (Hurles & Lupski, 2006).

2) DNA Transposition is the process by which segments of DNA can and do move to different regions of the genome. In the past, they were also called jumping genes, while the parts that move are called transposons. Thus, depending on the type of transposition, a segment of DNA is copied or cut from one region of the genome and attached to another with the help of special enzymes (transposase enzyme), while structural changes can occur. DNA transposition that takes part in gene duplication can be achieved through 1 or 2 major pathways (Hahn, 2009):

3) Retroposition. The gene copies resulting from this mechanism are the result of the reverse transcription of mRNA into cDNA, which enters a new position in the genome. By this mechanism the resulting paralogs lack multi-A-tails and introns. These transcripts carry non-coding sequences around them and are therefore less likely to be able to be expressed after duplication (Vinckenbosch et al., 2006).

4) Polyploidy. Whole genome duplications result in the creation of new copies of each gene in the genome. Although every gene is duplicated, only 10–30% of all genes are retained in the genome for a long time (Byrne & Wolfe, 2005; Maere et al., 2005; Paterson et al., 2006). The type or function of genes retained after polyploidy appears to differ from those duplicated by duplications involving smaller segments.

5) Segmental Duplications are segments of DNA, ranging in size from 1kb to 400kb and repeated in the genome. These sequences should have more than 90% similarity with respect to a reference sequence. They exist in various regions as a result of the duplication effect. Some understand the term segmental duplications to mean large-scale duplications involving more than one gene, or when referring to segmental duplications they refer to any duplication event that is not due to polyploidy or a duplication with nonsense sequences that are more than 90% similar (Bailey et al., 2001; Hahn, 2009).

Advantages of polyploidy

Evolutionary biologists have hypothesized that gene duplication has played an important role in the evolution of species (**Figure 1.10**), particularly in eukaryotes, whose genomes are characterized by the presence of numerous multigene families. By creating many copies of each gene, gene duplication allowed the evolution of new functions in proteins, playing an important role in adaptive evolution. Although polyploidy more often leads to an evolutionary dead end, it appears that polyploid organisms sometimes have advantages over their diploid relatives. In particular, some polyploid organisms are more resistant to changing environments, thus have a reduced risk of extinction (Van De Peer et al., 2021). The accelerated genomic and epigenetic changes that occur after whole-genome duplication may allow polyploids to adapt more quickly than diploids. In addition, polyploids have greater resistance to mutations, meaning that redundant gene copies can temporarily overshadow the effects of deleterious mutations in their paralog. Whole-genome duplication leads to rapid expansion of entire gene families, which can be maintained to evolve and produce genes with similar, but not identical, functions (Kondrashov, 2012). Thus, it is reasonable to assume that the duplication of the entire genome allows the optimization of already existing functions. Whole genome duplications have also been found in unicellular organisms. The first ancient genome duplication discovered in eukaryotes was in the yeast *Saccharomyces cerevisiae* (Wolfe & Shields, 1997). Because whole-genome duplications in plants and animals created particularly species-rich groups, for example nearly 30,000 species in teleosts and >350,000 species in angiosperms, polyploidy provides

opportunities for evolutionary success and facilitates the differentiation and speciation of organisms (**Figure 1.11**). In allopolyploidy and autopolyploidy, increased heterozygosity can lead to increased variation in gene expression, which can lead to increased robustness and faster adaptation to new conditions. Increased phenotypic variation and the consequences of heterosis, i.e. the increase in fitness that occurs in crosses between different populations, enable polyploid organisms to survive in environmental conditions unfavorable to their diploid ancestors (Glasauer & Neuhauss, 2014; Van De Peer et al., 2009).

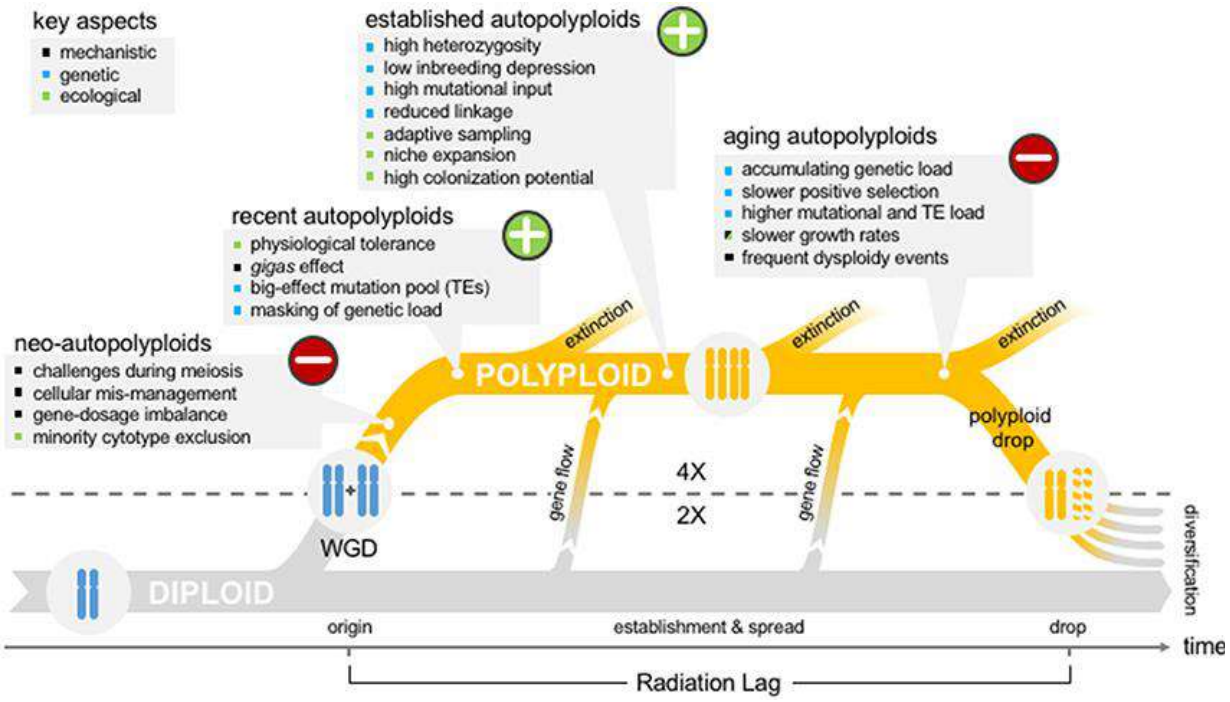


Figure 1.10. The temporal history of a diploid-autopolyploid system, along with the shifting benefits and dangers associated with WGD, which can contribute to the various stages of polyploid evolution; (Source: Baduel et al., 2018).

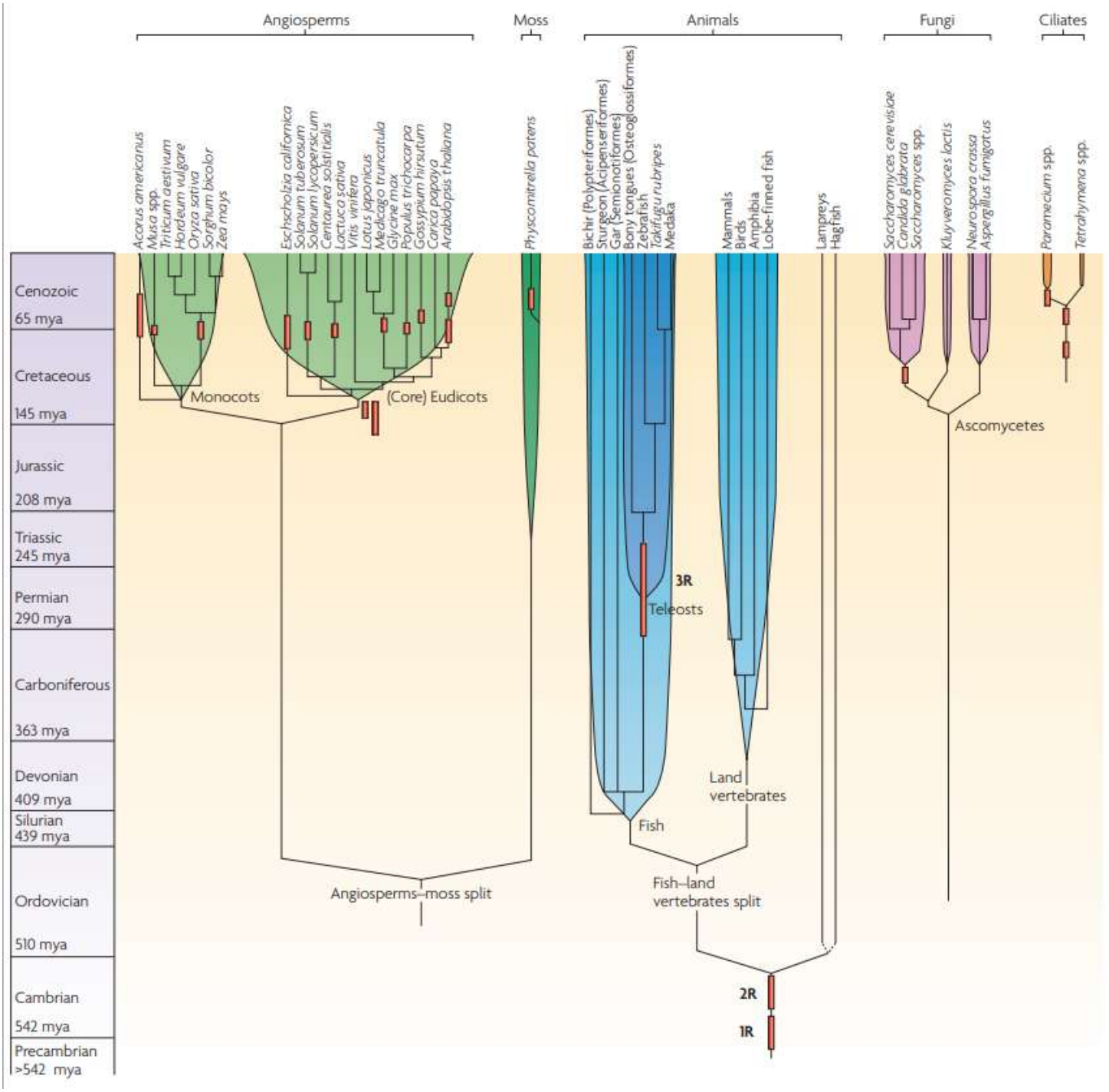


Figure 1.11. Whole-genome duplications across the phylogeny of eukaryotes; (Source: Van de Peer et al., 2009).

First and Second Rounds of Whole Genome Duplication in Vertebrates ancestor

Ohno in his famous book published in 1970 advanced the idea that Vertebrates underwent two rounds of duplication of their whole genome (WGD). The 1-2-4 rule is the most prevalent model to explain the evolution of gene families and of the genome in general. Based on this model, the two rounds of WGD occurred early in the evolution of Deuterostomes (Meyer & Van De Peer, 2005). Phylogenetic analyses of many duplicated genes in genomes, comparative analyses of their chromosomal locations, and evolutionary inferences of ancestral karyotypes confirm two rounds of WGD in the vertebrate ancestor (Dehal & Boore, 2005; Nakatani et al., 2007; Panopoulou & Poustka, 2005; Putnam et al., 2008).

Until recently, the most popular version of the hypothesis of two rounds of WGD was after the emergence of Cephalochordates and before the spread of Jaws. However, recent evidence has demonstrated that Urochordates are more closely related to Vertebrates than to Cephalochordates (Bourlat et al., 2006; Delsuc et al., 2006; Marlétaz et al., 2023) changing the modern version of the story to that the WGDs occurred after the appearance of the Urochordates. The Agnatha, represented by the Myxines and the Cyclostomes occupy an intermediate phylogenetic position between the Urochordates and the Jaws. The most recent versions of the first-round (1R) and second-round (2R) WGD hypotheses place 1R at the confluence of the common ancestor of the Jaws and Agnatha, while 2R at the ancestor of the Jaws (**Figure 1.12**) (Nakatani et al., 2021; Yu et al., 2023).

Third round of whole-genome duplication in Teleosts

The Teleosts (Teleostei) belong to the large homotaxy of the Osteichthyes, that includes all the bony fishes, and the superclass of Actinopterygii. Studies have shown that many fishes may have even more genes than humans, and research shows that an additional WGD occurred in Teleosts, i.e. the extension of the 1-2-4 rule to 1-2-4-8 (Aparicio et al., 2002; Meyer & Schartl, 1999; Wittbrodt et al., 1998). The conservation of large stretches of gene arrays within the teleost lineage and the detection of conserved syntenic regions (large chromosomal regions conserved

between organisms) in *Hox* gene clusters and other genes in teleosts, was seen as strong evidence for additional WGD in the Teleostei. From studies identifying duplicated genes in *Takifugu rubripes*, it was estimated that the third round (3R) of WGD occurred 350-320 million years ago (Christoffels et al., 2004; Vandepoele et al., 2004). The most convincing evidence for additional WGD in the teleosts came from comparative analysis between the *Tetraodon* and human genome sequences. Jailon and colleagues (Jailon et al.,2004) compared the distribution of genes in *Tetraodon* and human chromosomes and found that many groups of genes that remained linked, in the same order, in different human chromosomes (gene neighborhoods or otherwise synteny) were found in duplicate on two different chromosomes of *Tetraodon* (Glasauer & Neuhauss, 2014).

The 3R timing demonstrates that WGD in fish can be associated with both an increase in species number and biodiversity (**Figure 1.12**). The timing of 3R event, which follows the divergence of the non-Teleosteans, i.e. basal-early Actinopterygians consisting of 44 species belonging to 5 families, and preceding the origin of the Teleosts, provides further evidence for the connection between the 3R and the variety of organisms. The suborders Osteoglossomorpha (217 species) and Elopomorpha (37 species) representing the first generation to diverge after the WGD show a somewhat larger number of species, compared with the basal Actinopterygii (Hoegg et al., 2004). The hypothesis of a cause-and-effect relationship between the WGD that occurred in teleosts and their surprisingly species-rich clade, is supported by the fact that species-poor lineages, branched off from the original generation of Actinopterygii before the 3R WGD (Meyer & Van De Peer, 2005; Sato & Nishida, 2010).

Certain taxonomic groups within the teleost taxon are known to have undergone one or more rounds of independent WGD after 3R. Such groups are: Cyprinidae, Catostomidae, Cobitidae, Characiformes, Siluriformes, Salmoniformes, Diretmidae, Poeciliidae, and Channidae. Most of these additional duplications (Leggatt & Iwama, 2003) are thought to have occurred in certain genera of these groups except for the order Salmoniformes, in which an additional fourth round (4R) of WGD probably occurred in all species of the order (Allendorf & Thorgaard, 1984; Moghadam et al., 2005).

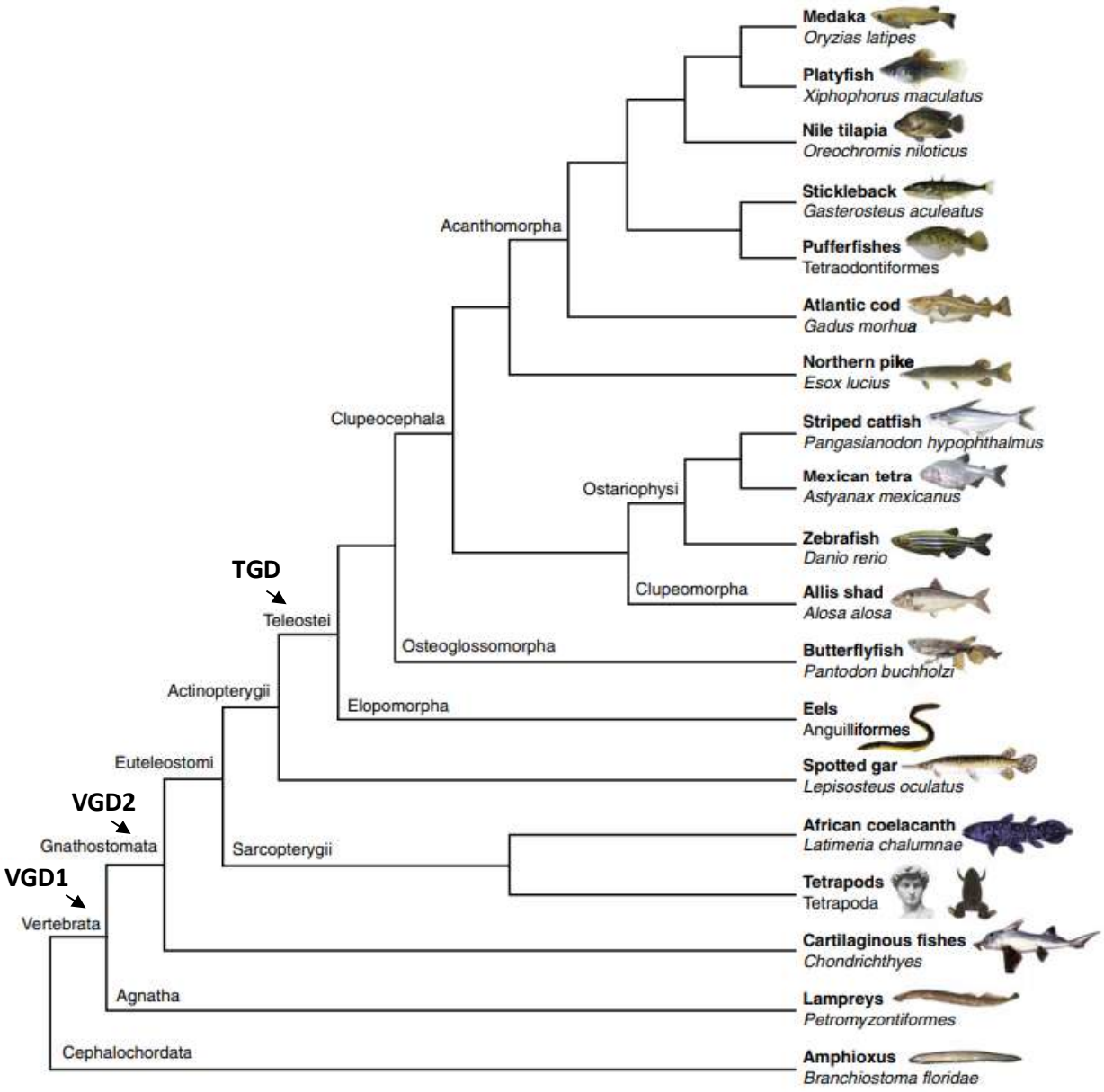


Figure 1.12. Cladogram showing phylogenetic relationships among vertebrate species. Tree topology was adopted from Near et al. (2012). VGD1/2: vertebrate genome duplication 1/2; TGD: teleost genome duplication; (Source: Braasch et al., 2014).

The fates of paralogous genes

Studies have shown that two evolutionary forces, namely *random genetic drift*, and *positive selection or directional selection*, act to determine the final fate of the gene pairs in terms of function (Conant & Wolfe, 2008; R. C. Moore & Purugganan, 2003) Genetic drift leads to the random increase or decrease of alleles from generation to generation, their random establishment or loss and it is more evident in smaller populations (Barton, 2008). Positive or directional selection favors one allele over another or favors increased values of some quantitative trait.

Multiple fates of a gene following its duplication have been observed (**Figure 1.13**). Maintenance of an unrelated duplicated gene through differentiation of ancestral function occurs due to genetic drift. On the other hand, the maintenance of genes either through acquisition of new function or divergence of function is driven by the selective advantage of the gene locus that has undergone duplication (J. Zhang, 2003). The exact mechanism of stabilization of duplicated gene loci depends on several factors, including the relative levels of neutral and deleterious mutations acting on the gene pair, selection coefficients, and in some cases the actual population size (Lynch & Force, 2000; R. C. Moore & Purugganan, 2003; Nowak et al., 1997; J. Zhang, 2003). Susumu Ohno (1970) pointed out three main results in the evolution of duplicated genes, which are the preservation of genes, the differentiation of their function, and the development of a new function (**Figure 1.13**). Other results are the loss of one of the two genes as well as the reciprocal loss of both copies with the formation of pseudogenes that usually occurs in the first million years after duplication, provided the gene is not under some evolutionary pressure.

Gene Conservation

Having duplicate genes can be advantageous simply because extra amounts of protein or RNA products are provided (Hurst & Smith, 1998; Ohno, 1970). In addition, there are two more possible mechanisms that could result in paralogous genes maintain the same function after duplication. The first is concerted evolution (Liao, 2008), in which members of a family preserve sequences and functions by repeated gene conversion; a specific type of homologous recombination between paralogs, and/or unequal crossing over. The second one is strong

purifying selection (Nei et al., 2000) against mutations that modify gene function, preventing duplicated genes from diverging.

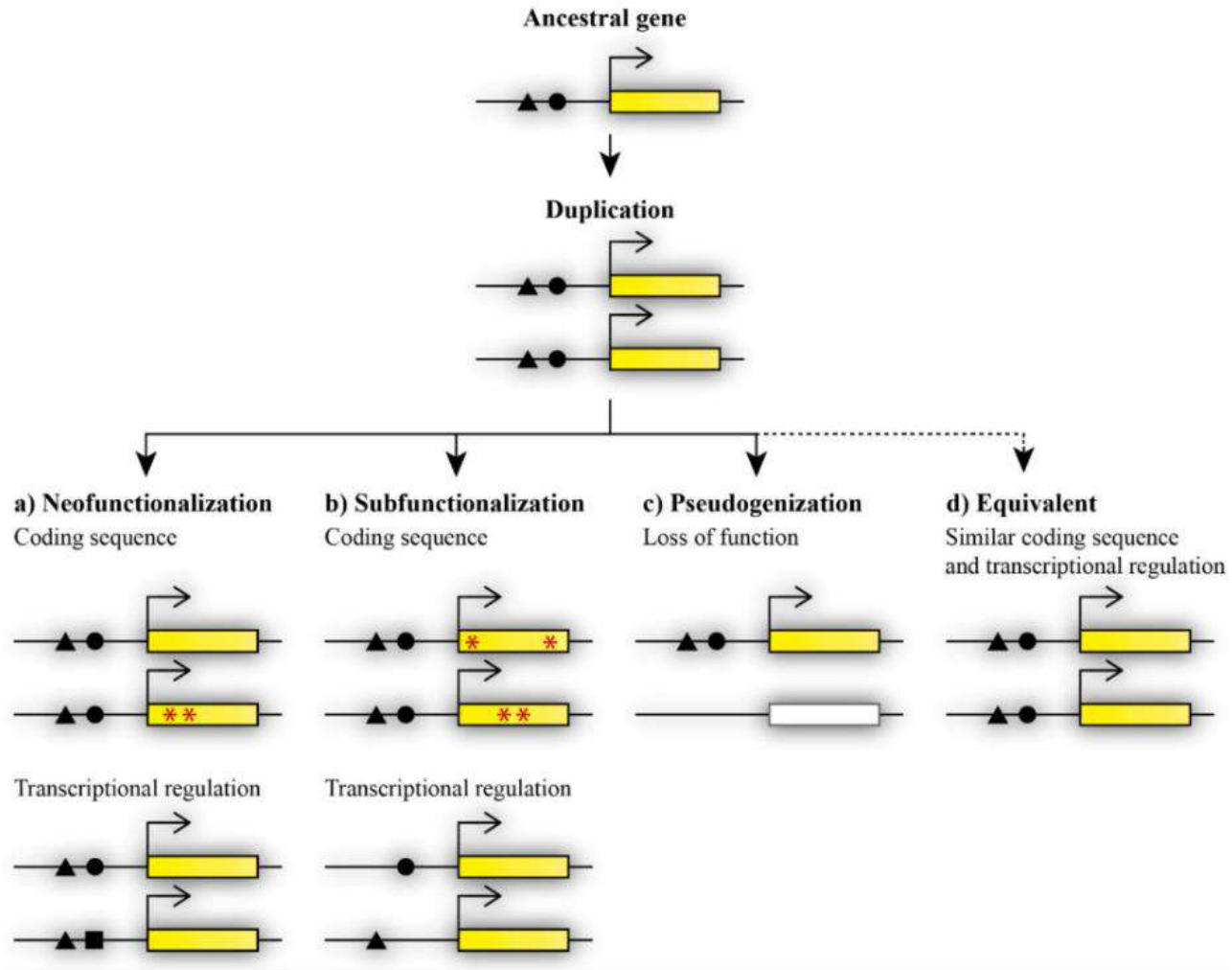


Figure 1.13. Possible fates of paralogue gene duplication. Duplicated genes can acquire divergent functions via (a) neofunctionalization, in which one duplicate retains the ancestral gene function while the other duplicate acquires a novel function, and the novel function is acquired via mutations in the gene coding sequence (indicated by red asterisks) or changes in transcriptional regulation (black square), or (b) subfunctionalization, in which mutations change the functions of both duplicates in a way that each paralog maintains a part of the ancestral function, and their functions are complementary. Other possible outcomes for duplicated genes include (c) loss of function during the pseudogenization process or, less frequently, (d) maintenance of equivalent functions; (Source: Drobek 2022).

New or differentiated function – loss of function

Alternatively, strong negative selection acts against mutations that alter gene function and may prevent the differentiation of genes that have arisen from duplication. Two copies can remain stable when they differ in some of their functions, where differentiation of their function (Subfunctionalization, SF) may have occurred, and each daughter gene adopts part of the functions of the ancestral gene (Liu et al., 2018). One form of differentiation of function that is potentially important in developmental evolution is the division of gene expression after duplication (Drobek, 2022). Many duplicated genes have been shown to evolve following this model. One of the most important results of gene duplication is the origin of new function. In many cases, however, a related function, rather than an entirely new function, evolves after the duplication event (Escriva et al., 2006).

The development of a new function (Neofunctionalization, NF) in duplicated genes requires varying numbers of amino acid substitutions. The development of a new function occurs when one of the two genes retains the functions of the ancestral gene, while the second is left free to develop some new functions. That is, it is an adaptive process and one of the gene copies must be mutated to develop a function that was not present in the ancestral gene (Force et al., 1999).

Another possible outcome is that one of the genes retains the functions of the ancestral gene, while the second is lost (Non-functionalization). This is also the most common outcome. In the first and second cases, all interactions of the ancestral gene are inherited intact in the pair, and depending on the extent of mutations that occur later, a number of interactions may be lost, while another number of interactions may appear. Analysis performed, showed that both variation of function and loss of function play a dominant role in the fate of duplicated genes (Chanderbali et al., 2017; Lynch & Conery, 2000; Sato & Nishida, 2010). Moreover, differentiation of function occurs rapidly after gene duplication whereas the model for the development of new function is a long process that continues long after duplication. Thus, the short-term maintenance of duplicated genes in the genome of organisms is mainly due to the model of functional differentiation, which is consistent with a much higher proportion of degenerative mutations than beneficial ones (Lynch & Force, 2000; Ohno, 1970; Rastogi & Liberles, 2005; J.

Zhang, 2003). Retention of duplicated genes in the genome and partial functional relaxation caused by loss of ancestral functions then provide the opportunity for beneficial mutations, which can lead to new functions, i.e. rapid SF combined with prolonged NF. The observations from a large percentage of duplicated genes support the above view and suggest the SNF (Subneofunctionalization) model, rapid differentiation of function accompanied by prolonged and significant development of a new function (He & Zhang, 2005).

Reciprocal gene loss

It is possible that differential gene duplication and pseudogenesis in geographically isolated populations causes reproductive isolation and speciation (Lynch & Conery, 2000) However, it was not until the late 1990s, when many genome sequences were determined and analyzed, that the prevalence and importance of gene duplication was clearly demonstrated. Through genome-wide analysis, population genetics, and molecular experimentation, rapid progress has been made in uncovering the mechanisms by which duplicated genes diverge in function and contribute to evolution (Lynch & Conery, 2000; J. Zhang, 2003). Whole genome duplications directly facilitate speciation through reciprocal gene loss. Reciprocal gene loss of the tens of thousands of genes and regulatory RNAs produced by genome duplication facilitates speciation (Lynch & Force, 2000). If mutual gene loss of duplicated genes continues with time, speciation events will occur. There is evidence to suggest that reciprocal gene loss occurred in the teleost lineage. It has been found that 8% of the gene loci of *Tetraodon nigroviridis* and *Danio rerio* have undergone reciprocal gene loss (Sémon & Wolfe, 2007).

Gene Balance Hypothesis

The Gene Balance Hypothesis predicts the fate of nuclear genes that have resulted from genome-wide and smaller-scale duplications based on the assumption that an imbalance in the concentration of protein subunits in a macromolecular complex or between proteins with opposing functions in a transcription or signaling network, can lead to either reduced fitness or lethality. Maintaining the correct protein and balance in the transcription process is critical to

maintaining normal function. For example, high concentration of a protein bound to the multi-protein complex would likely result in large negative pleiotropic effects. This hypothesis, supported by analysis of eukaryotic genomes, provides the basis for understanding copy conservation (Edger & Pires, 2009; Freeling, 2009), unlike the alternative hypotheses of Subfunctionalization and Neofunctionalization that do not make any predictions between gene function and post-duplication conservation frequency.

The gene balance hypothesis also provides a well-supported mechanism to explain the significant over-conservation of genes with specific functions. In Vertebrates the two rounds of WGD favor the conservation and expansion of genes, such as the number of *Hox* gene clusters, growth factors, insulin receptors and nuclear receptors. Growth and transcription regulators and signal transducers have been conserved in two copies after the 1R and 2R of WGD in vertebrates and from the 3R that followed in fish (Blomme et al., 2006; Brunet et al., 2006; Van De Peer et al., 2009). These effects of gene-genome duplications can be explained by the effect of gene dosage balance (Birchler et al., 2005; Papp et al., 2003; Van De Peer et al., 2009). After WGD, entire functional units are kept intrinsically duplicated through non-adaptive dosage balance effects and then adaptively evolve new functions resulting in an increase in morphological complexity (Freeling, 2009; Freeling & Thomas, 2006).

The Teleosts

Teleosts, (infraclass Teleostei), is a large and extremely diverse group of ray-finned fishes with a teleost-specific WGD or 3R occurring at the base of their lineage (**Figure 1.14**). Along with the chondrosteans and the holosteans, they are one of the three major subdivisions of the class Actinopterygii, the most advanced of the bony fishes. Teleosts range in size from tiny minnows that are less than one-third of an inch long when fully adult, to large marlins that exceed 3.4 meters in length and 550 kg in weight.

Different species and groups have widely varying life cycles and behaviors. Teleost fishes are adapted to habitats ranging from the cold Arctic and Antarctic oceans that remain colder than the freezing point of fresh water, to desert hot springs that reach temperatures over 38°C. In terms of physical habitat, the teleosts are adapted to a wide and interesting variety of conditions, from fast, rock-laden torrential mountain streams in the Himalayas to the lightless depths of ocean trenches 8,370 meters below the surface, where many species produce their own light (Facey et al., 2023). The diversity in habitat adaptation is ultimately reflected in life cycle, behavior, locomotion, anatomy, and reproduction.

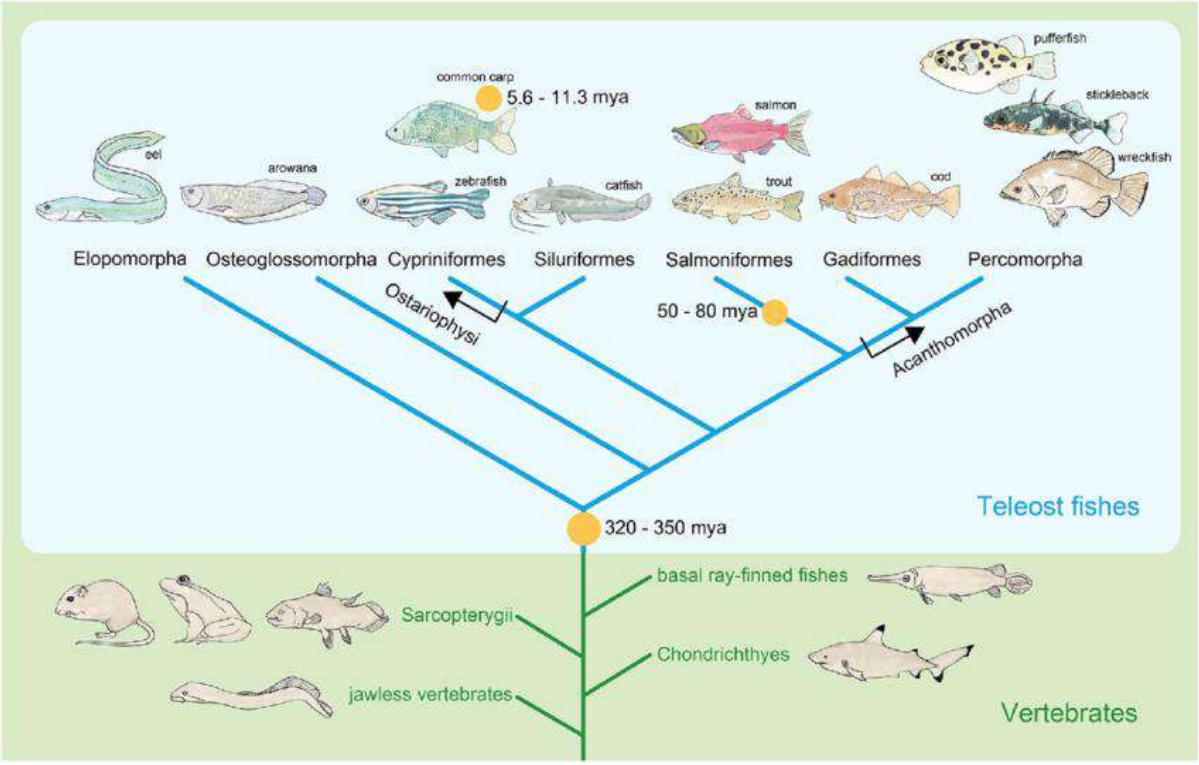


Figure 1.14. Simplified phylogeny of teleost fishes; (Source: Glasauer & Neuhauss, 2014).

The behavior of teleosts is as varied as their other attributes. Some oceanic fishes travel in close-ordered schools, seemingly responding to predation from larger fishes almost as a single organism. Larger, predatory fishes are usually solitary and hunt or wait for prey alone. Many marine shore and freshwater fishes establish territories during the breeding season, and some

may travel in relatively loose schools or shoals when not breeding. Many kinds of teleosts enter into symbiotic relationships with other species of fishes and organisms.

The structure of the tail and the efficiency of the swimming mechanism are the prime characters that distinguish teleosts from other, “lower,” fishes. The dorsal fin and the anal fin (a ventral median fin) are used partly to aid in stability and in turning and partly in forward locomotion. The paired pelvic or ventral fins and the paired pectoral fins behind the head are used to help stabilize the body and to turn the fish. Forward motion is provided by bending of the body and caudal fin; waves of muscular action pass from the head to the tail, pushing the sides of the body and tail against the water and forcing the fish forward. The common fusiform shape reduces the turbulence and drag of water flowing over the fish’s body, offering least resistance to the water. The head of the fish is adapted for feeding, breathing, and detecting prey and enemies, while it is relatively streamlined, offering as little resistance to the water as possible. The head is well formed to take in food and water. The fish forces the water in one direction, into the mouth, over the gills, and out the gill slits. Backflow is prevented by valves at the mouth and by the gill covers. The teleost head is equipped with eyes and organs for the sense of smell located in optimum spots for seeing and smelling food. At the same time, these organs offer little resistance to water flowing over the head (Facey et al., 2023).

The fish species in study: European seabass and gilthead seabream

The European seabass (*Dicentrarchus labrax*, **Figure 1.15**) and gilthead seabream (*Sparus aurata,* **Figure 1.16**) are teleosts that over the years have become the emblematic species of the Mediterranean marine fish farming and the top species produced in the EU waters. A comparative presentation of their taxonomy, biological characteristics and distribution is provided in **Table 1.1**.

Both species are marine fish commonly found in coastal areas, often in shallow waters, they are known to exhibit social behavior and often form schools or groups. Both species have carnivorous feeding habits, preying on smaller fish and various invertebrates, and they both are broadcast

spawners, releasing their eggs and sperm into the water column for fertilization, they have a wide tolerance for salinity levels, allowing them to adapt to different environmental conditions.

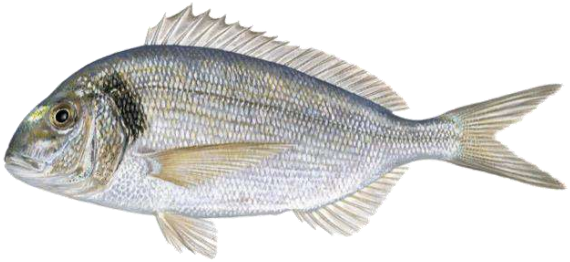


Figure 1.16. Gilthead Seabream (Sparus aurata)

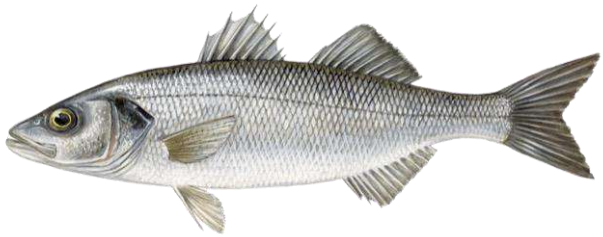


Figure 1.15. European seabass (Dicentrarchus labrax)

On the other hand, they belong to different teleost families; they differ in size, with European seabass generally growing larger compared with gilthead seabream; they have distinct body shapes and colorations as gilthead seabream has a deep-bodied form with a characteristic golden stripe on its head, whereas European seabass has a slender body with silver coloration; their preferred temperature ranges vary, with gilthead seabream preferring warmer waters compared with the European seabass; they exhibit different habitat preferences, with gilthead seabream often associated with rocky or seagrass-covered areas, whereas European seabass is commonly found in estuaries, lagoons, and rocky shores.

A wealth of genetic and genomic resources are available for both species; at NCBI there are 32,208 and 32,280 entries of gene sequences for gilthead seabream and European seabass respectively. In addition, there are reference genomes for both species at; for gilthead seabream there are three assemblies, fSpaAur1.1 alternate haplotype, ASM330901v1 and fSpaAur1.2 with the last two at chromosome level while the first one is at scaffold level; for European seabass there are also three assemblies, ASM18081v1, seabass_V1.0 and dlabrax2021 all at scaffold level.

Gilthead seabream and European seabass share common developmental trajectories with small differences, which are mainly associated with the temperatures they prefer according to the developmental stage they are in and the seasons they choose to reproduce (Zanuy et al., 2001). Their development can be distinguished in three main stages (**Figure 1.17**). First, the egg stage,

includes the developmental processes from fertilization to hatching. Second, is the larval stage which is the period between hatching and metamorphosis. The third stage is defined by the initiation of metamorphosis and the transition to juvenile stage until reaching reproductive maturity as an adult (Deplano et al., 1991).

Table 1.1. Main characteristics of European seabass and gilthead seabream

	Gilthead Seabream (*Sparus aurata*)	**European Seabass (*Dicentrarchus labrax*)**
Taxonomy	Teleost (*Sparidae*)	Teleost (*Moronidae*)
Ecology	Marine, warm-water species	Marine, warm-water species
Distribution	North-East Atlantic; Mediterranean and Black Seas	Eastern Atlantic- from the British Isles, the Strait of Gibraltar to Cape Verde and around the Canary Islands; Mediterranean Sea
Temperature Range	Eurythermal, 14 - 26°C	Eurythermal, 8 - 24°C
Salinity	Euryhaline, tolerate a wide range of salinities	Euryhaline, tolerate a wide range of salinities
Depth Range	1 - 150 m (generally 50m)	10 - 100 m
Habitat	Rocky, seagrass meadows, sandy grounds	Found in estuaries, lagoons, and rocky shores
Feeding	Carnivorous	Mainly Carnivorous, occasionally herbivorous
Behavior	Form schools and exhibit social behavior	Juveniles form schools, adults are less social
Sex	Protandrous hermaphrodite	Protandrous hermaphrodite
Reproductive Season	October to December	December to March (Mediterranean), June (Atlantic)
Breeding	Broadcast spawner, releasing eggs and sperm into the water	Broadcast spawner, releasing eggs and sperm into the water
Morphology	Oval body, with a characteristic golden stripe on the head	Slender, elongated bodies with silver coloration

Ontogeny can be separated into seven discrete phases (Pavlidis et al., 2011):

1) Embryos
2) Hatching
3) First feeding (pre-larvae): Mouth opens, developed pectoral fins, complete yolk sac absorption
4) Flexion: Completion of notochord flexion
5) Post flexion: First appearance of dorsal and anal fin bases. All fin rays started to form
6) All fins (end of metamorphosis): All fins have been established as in adults
7) Juveniles: Abundant scales all over the body

At hatching, young larvae are about 2.5mm in size, 0.3-0.4mg in weight and may be at a different stage of development depending on the size of the yolk sac. At this stage, larvae lack a functional mouth, eye coloration, differentiated fins, and feed exclusively on yolk (Mwaluma, J Strydom, N, 2014). Larvae open their mouths 5 days post hatching (DPH) and step into a transitional period when they still feed on yolk reserves and they try to feed on exogenous food. After the lecithotrophic phase is over, they feed exclusively on an exogenous source, which is plankton, and rapid changes in their skeletal elements occur.

The transition from the larva stage to the juvenile stage is broadly called metamorphosis. Metamorphosis encompasses several key elements: (i) a shift in habitat, (ii) a significant morphological transformation, (iii) a change in the adaptive landscape, and (iv) the acknowledgment that the pre-metamorphic stage occurs after the embryonic phase. Additionally, a majority of perspectives includes (v) a shift in feeding mode and (vi) typically signifies a transition from a pre-reproductive to a reproductive stage (Bishop et al., 2006). The modifications in internal morphology associated with the transition from larval to juvenile stages in fish seem to involve various studied tissues, including but not limited to changes in gastrointestinal tract structure, skin and pigmentation, as well as the skeletal framework

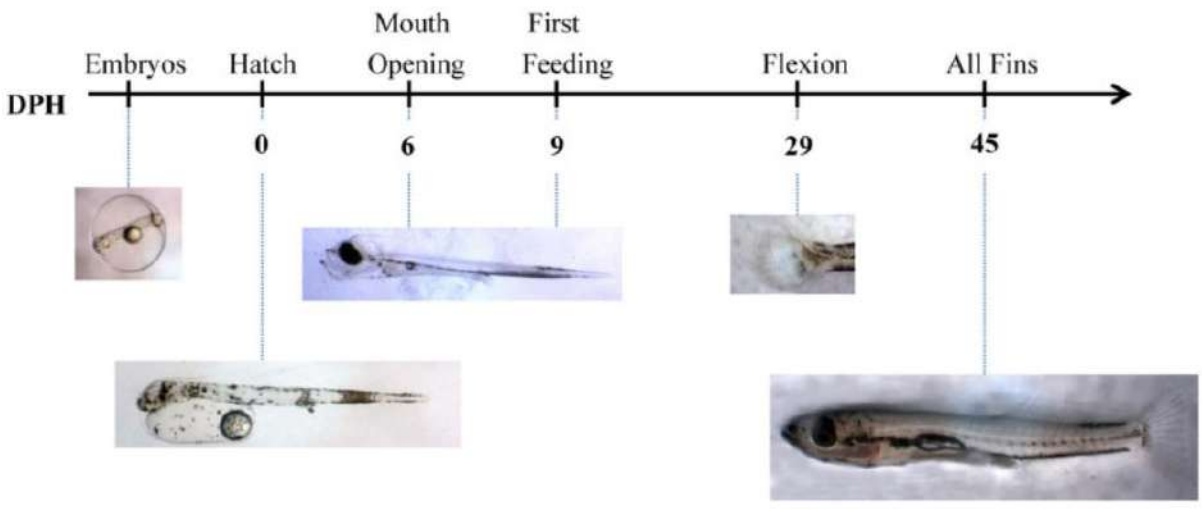


Figure 1.17. Ontogeny of the European sea bass. DPH: Days Post Hatch; (Source: Tsalafouta et al., 2014).

(Campinho et al., 2006; Power et al., 2008; Sæle et al., 2004). In essence, metamorphosis is characterized as a transformation in both form and function, enabling the organism to move between life stages and effectively adapt to a new habitat.

Aim of the thesis

Oxidative phosphorylation (OXPHOS) is the indispensable process that produces energy for life and it is unique in many ways; it is present at all levels of life regardless the complexity; it relies upon the cross-talk of two genomes, the nuclear and the mitochondrial, that differ in origin, ploidy level, rate of accumulating mutations, rate of evolution, location within the cell, transcriptional and translational apparatuses; it demands the collaborative operation of five protein complexes encoded by approximately 100 genes and targeted to their final location through highly regulated pathways. How the genes encoding for such a crucial and robust procedure have evolved over time?

This thesis explores the question by using two teleosts, the European seabass and gilthead seabream, as model organisms in a comparative approach that covers all the OXPHOS genes encoded by the nuclear DNA that has undergone three rounds of WGD. Both teleosts are important commercial species of different taxonomy, and with adequate genomic resources available to allow:

1. Mapping the nuclear genes encoding for OXPHOS proteins in genomes that have undergone three rounds of WGD and inferring the subcellular localization of the proteins based on signal sequences;
2. Tracing their evolutionary trajectories through extensive phylogenetic and syntenic analyses, and protein structure comparisons;
3. Mapping the level of gene polymorphism and establishing evidence about the selection forces driving their evolution;

4. Exploring functional characteristics of the OXPHOS gene families at tissue-specific level and during ontogeny to draw correlations between and within OXPHOS gene families, and with the changing energy demands during development;
5. Tracing thermal plasticity in their expression and the impact under challenging circumstances.

By addressing these objectives, the thesis investigates the pattern of OXPHOS gene fate after three rounds of WGD as a paradigm of how considerable restraints in biological function may determine the fate of duplicated genes. These restraints are expected to be reflected also on the level of genetic polymorphism and the selection forces in action.

Fish ontogeny is a unique and complex period in fish life marked with dramatic changes in morphology, physiology, metabolism and behavior. The tightly regulated landscape of cell divisions, migrations and differentiation driving these dramatic changes and the ultimate transition to the adult form through metamorphosis, demands and relies on high energy supplies. WGDs have provided a unique opportunity for increasing the number of genes that are coding for multiprotein complexes, and they are expected to have provided ground for regulatory plasticity during ontogeny.

For the first time, the total landscape of OXPHOS genes is considered in teleosts that, besides their commercial value, have emerged as the highly plastic and highly adaptive vertebrate models. As such, the thesis adds to the genetic/genomic resources available for European seabass and gilthead seabream contributing to the expansion of knowledge and resources on non-classical fish models, while it explores temperature as an epigenetic factor in response to rapidly warming climate.

Outline of the Thesis

This thesis consists of six (6) main chapters, four of which are dedicated to the presentation of the results produced.

In the first chapter, already presented, an Introduction to the scientific information used to develop the scientific questions of the thesis is provided. The presentation of information starts with the biological function in study and the structure of the proteins involved, to continue with the evolutionary mechanisms known to shape genes and genome and how these mechanisms have contributed to major evolutionary trajectories. The Introduction seeks to present the advantages of using teleosts as models to study gene evolution and the focal species in particular, before outlining the sequence of events during their development and the plasticity exhibited. The Introduction concludes with the aims of the thesis, the questions explored and the elements of novelty.

Chapters 2-5 are dedicated to the presentation of the results. Each chapter consists of a focused Introduction, detailed presentation of Materials and Methods used, the actual presentation and Discussion of the Results. Chapters 2 and 3 explore evolutionary perspectives of OXPHOS genes, while Chapters 4 and 5 touch on functional aspects of the OXPHOS paralogue genes.

Chapter 2 presents all the information on mapping the nuclear genes encoding for OXPHOS proteins in the genomes of gilthead seabream and European seabass, the phylogenetic and the syntenic analyses, and the comparison of the deduced protein structures to infer evolutionary trajectories. The chapter goes a step further to check the possible subcellular localization of the paralogue gene products based on the signal sequences they bear, as an indication of their functionality. The chapter concludes with suggested evolutionary trajectories of the retained OXPHOS paralogs.

In Chapter 3 the approach for and the results of mapping the polymorphism of paralogue genes are presented. SNPs are distinguished according to their localization (CDS and UTRs) and the result of substitution (synonymous, missense, splice region) before exploring the selection forces driving their evolution. The evidence produced added to the evolutionary trajectories drawn in Chapter 2.

The expression patterns of OXPHOS paralogs during gilthead seabream and European seabass ontogeny are presented in Chapter 4 to explore functionality and regulation of expression as evidence for the fate of paralogue genes. Whole transcriptome analysis revealed that OXPHOS paralogs are transcribed. Differential expressions during development and at tissue level indicated regulatory neo/subfunctionalization as plausible outcome.

Chapter 5 provides an insight into the epigenetic effects of early developmental temperature on the expression of OXPHOS paralogs in gilthead seabream under control and challenging circumstances, to confirm that paralogs are differentially regulated by temperature and add to the idea of regulatory neo/subfunctionalization.

Chapter 6 presents the overall conclusions of the thesis.

Chapter 2

Tracing the evolutionary trajectories of OXPHOS genes

Introduction

More than half of all living fish species belong to the diverse group of fishes known as teleosts. Teleosts have undergone three rounds of whole genome duplication that has produced new genetic material for developing new traits and adaptations. Indeed, this is manifested in the tremendous radiation that teleosts have undergone to produce a wide range of species with various morphologies, life histories, and ecologies, and it offers a rare chance to examine how genes and genomes have evolved in response to various selective pressures. This chapter focuses on mapping and tracing the origin and evolution of the paralogs of the nuclear genes encoding for OXPHOS proteins in two fish species, the European seabass and the gilthead seabream. OXPHOS is a fundamental biological process, and it is accepted that energy production and allocation and their regulation are critical for different aspects of fish life. Given the biological constraints, it is expected that OXPHOS gene number and structure is very well conserved. However, OXPHOS apparatus is the common product of the nuclear and the more plastic mitochondrial genome, and this interaction might impact gene structure and polymorphism. The comparative analysis in European seabass and gilthead seabream, two species that are phylogenetically close and share ecological adaptations, can help to draw the lineage-specific and species-specific landscape of OXPHOS paralogs. The full repertoire of the nuclear OXPHOS gene paralogs is described and mapped on the European seabass and gilthead seabream genomes, the origin of their paralogs was traced through phylogenetic analysis and functional characteristics of the genes were inferred with biostatistical tools.

Materials and methods

Database mining and sequence retrieval

In order to study gene evolution and potential divergent function of OXPHOS in gilthead seabream (*Sparus aurata*) and European seabass (*Dicentrarchus labrax*), the genome of the two fish species from Ensembl database Release 105 (Cunningham et al., 2022) was searched for

orthologs of Human OXPHOS genes based on HUGO Gene Nomenclature Committee (HGNC) (Tweedie et al., 2021) accessed in March 2022, using the human (*Homo sapiens*) sequences as bait and a reciprocal BLAST approach. In order to find potential OXPHOS gene families, the genomes of the two species were searched for paralogs using a reciprocal BLAST approach, using the previously found OXPHOS genes of each species as bait. Only the OXPHOS-related paralogs were taken into account for the rest of the analysis in cases where paralogs were found. While rebuilding the phylogeny of each gene family and in order to ascertain the evolutionary history and relationships of the OXPHOS gene families in both species under study, the genomes of numerous animal species were searched for orthologs and paralogs of the genes encoding for the five complexes of OXPHOS. The species that were included in the analysis were several teleost representatives, including Nile tilapia (*Oreochromis niloticus*), medaka (*Oryzias latipes*), amazon molly (*Poecilia Formosa*), platyfish (*Xiphophorus maculatus*), stickleback (*Gasterosteus aculeatus*), greater amberjack (*Seriola dumerili*), fugu (*Takifugu rubripes*) and tetraodon (*Tetraodon nigroviridis*). Searches also included the genomes of the jawless fish sea lamprey (*Petromyzon marinus*) and the cartilaginous fish elephant shark (*Callorhinchus milii*), as representatives of basal vertebrate lineages; the early diverging ray-finned fish, spotted gar (*Lepisosteus oculatus*); the basal Sarcopterygian representative, the African coelacanth (*Latimeria chalumnae*); two tetrapods, mouse (*Mus musculus*) and chicken (*Gallus gallus*); and finally amphioxus (*Branchiostoma floridae*) as a representative of early chordata. Using the BLAST algorithm (Altschul et al. 1990), homolog genes were identified in the Ensembl (Cunningham et al. 2022) and/or NCBI (Sayers et al. 2022) genomic databases (Ensembl Release 105) based on sequence similarity and e-value and, once again, only the OXPHOS related paralogs were taken into account. Their identities were then confirmed by reciprocal BLAST against the human, gilthead seabream and European seabass genomes.

Sequence alignments and phylogenetic analysis

Although peptide sequences are preferred for phylogenetic research in vast evolutionary range, the deduced protein sequences of the examined genes often yielded short sequences, which had

a negative impact on phylogenetic signal. For the majority of the reconstructed phylogenies, nucleotide coding sequences (CDS) were used. The MAFT (Katoh, 2002) algorithm was used to perform multiple sequence alignment. After manually editing for gaps in the alignment, the species from which low-quality or short-length sequences were acquired were removed from the study. The revised alignment was used as input to MEGAX (Kumar et al., 2018) in order to find the optimal nucleotide/amino acid substitution model that would describe the evolution of gene families most effectively. Lastly, Maximum Likelihood algorithm was used in the phylogenetic analysis. MEGAX was used to reconstruct the phylogenetic trees, and 100 bootstrap repetitions were performed to statistically analyze the sequence branches that were constructed. The generated topologies of the phylogenetic trees for each gene family were evaluated based on the three rounds of vertebrate genome duplications and the evolution of the species included.

Collinearity analysis and duplicate gene origin classification

The Multiple Collinearity Scan Toolkit (MCScanX) (Y. Wang et al., 2012) was used to examine the degree of collinearity between the genomes of the gilthead seabream, European seabass and human as well as between the gilthead seabream and seabass paralogous regions, respectively. Blastp searches were run using the command-line program BLAST+ (Camacho et al., 2009) with an e-value threshold set to 1×10^{-5} in order to find orthologous areas between the genomes of gilthead seabream, seabass and human as well as paralogous regions in both fish genomes. Ensembl annotation files for all species were parsed to produce the gff files required by MCScanX. With the MATCH SIZE option set to 2, which indicates that two genes are necessary to create a collinear block, MCScanX was used to find collinearity blocks among all three genomes. The same files were used to investigate the genes evolutionary origins, in the MCSCanX tool called duplicate gene classifier, which classified each duplicated gene in the gilthead seabream and European seabass according to their origin. Genes were categorized into whole-genome or segmental duplicates, which include anchor genes in collinear blocks, tandem duplicates, which are paralogs that are physically next to one another in the genome, proximal duplicates, which have multiple genes between them but are mapped in a nearby chromosomal region, and dispersed duplicates,

which do not fit into any of the above-mentioned categories. While running the duplicate gene classifier the MATCH SIZE and N PROXIMAL parameters, which specify the maximum number of genes between two proximal duplicates, were set to 2 and 10, respectively.

Prediction of subcellular localization

The prediction tools DeepLoc 2.0 (Thumuluri et al., 2022) and TargetP 2.0 (Emanuelsson et al., 2000) were used to determine the subcellular location of the gilthead seabream and seabass proteins generated by the paralog genes. The mitochondrial localization was confirmed using DeepLoc, and the presence of the mitochondrial transit peptide and the location of the cleavage site were predicted using TargetP.

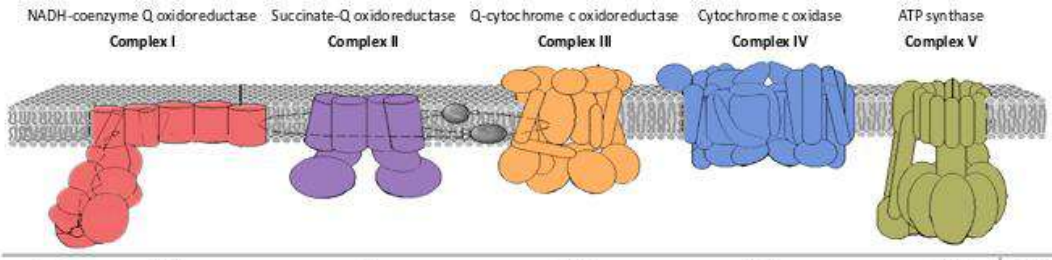
Functional conservation estimate

The InterPro 91.0 database (Paysan-Lafosse et al., 2023), accessed in November 2022, was utilized to investigate the degree of preserved function between the human and the homologs of either fish species. The protein sequences were entered into InterPro, and the outcomes of the fish paralogs and their human homologs were compared. On the basis of conserved domains or sites related to function as identified on their peptide sequences, the human and fish proteins were categorized into families and GO terms related to their biological processes and molecular functions. Lastly, CDS, UTRs, introns, and Pfam (Mistry et al., 2021) domains were obtained from Ensembl and plotted using GSDS 2.0 (Hu et al., 2015) for comparisons in order to estimate the degree of gene structure and protein domain conservation.

Results

OXPHOS gene families

Several OXPHOS gene families were discovered in both species after thorough searches using reciprocal BLAST to find OXPHOS orthologs between the human and fish genomes and within European seabass and gilthead seabream genome respectively, to find putative paralogs. Gene families were discovered in all five OXPHOS complexes, increasing the total number of OXPHOS genes to 105 and 103 for gilthead seabream and European seabass, respectively, in comparison with the 87 OXPHOS genes in the human genome (**Figure 2.1**). Spatial distribution and organization of OXPHOS genes across the genomes of the gilthead seabream and the European seabass are shown in **Figure 2.2** and **Figure 2.3**. Gene names were not consistently assigned to fish paralogs and/or human homologs based on Ensembl annotation, and some of them were designated as novel genes. The identified genes were renamed based on the gene name of their homolog at the human genome (best BLAST hit), adding at random an ascending number for each paralog to be more systematic and make it simpler to follow the gene families throughout the manuscript. The correspondence between the used gene name and Ensembl gene IDs is shown in **Table S2.1**.



	37	4	10	17	19	87
	40 (3)	5 (1)	13 (4)	21 (8)	26 (8)	105 (24)
	39 (3)	5 (1)	13 (4)	21 (7)	25 (7)	103 (22)

Figure 2.1. Number of OXPHOS genes per complex in the human, gilthead seabream and European seabass. In parenthesis, the number of gene families identified in the genomes of gilthead seabream and European seabass, respectively, are given.

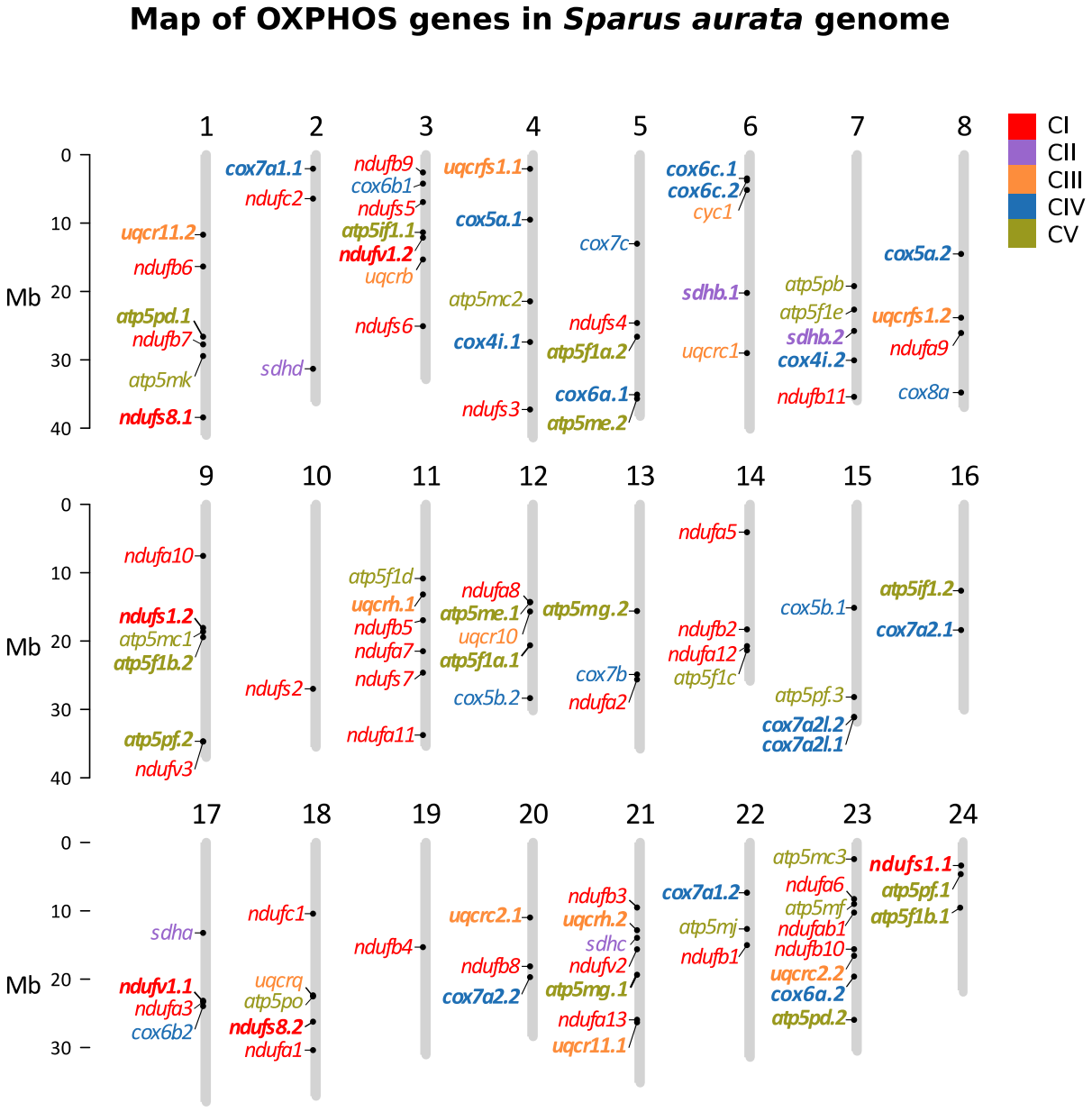


Figure 2.2. Mapping of OXPHOS genes onto the gilthead seabream genome (fSpaAur1.1) at the chromosome level was conducted using the R package chromPlot (Verdugo, 2017). Different colors were assigned to indicate the five OXPHOS complexes, and genes belonging to OXPHOS gene families are highlighted in bold.

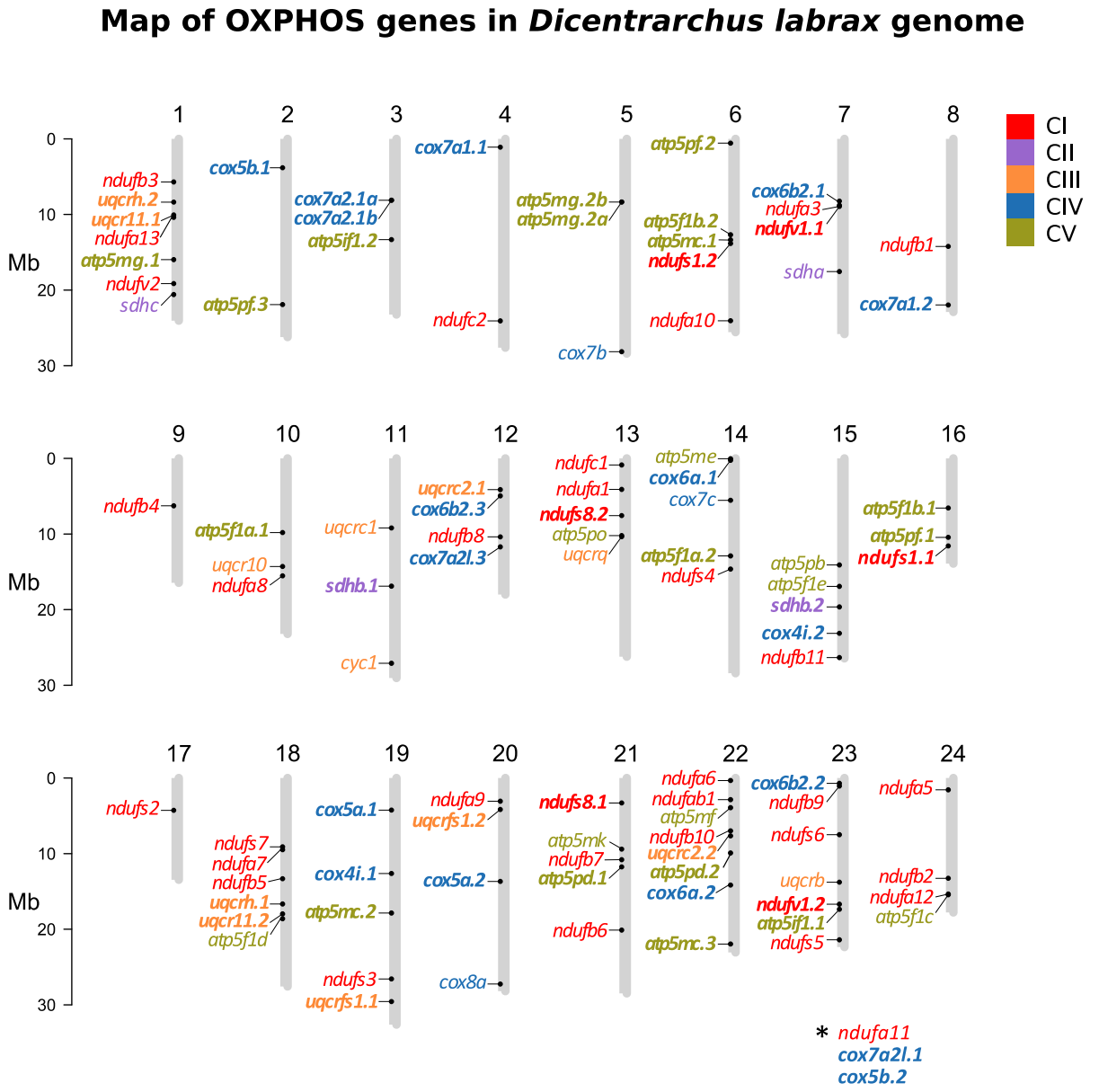


Figure 2.3. Mapping of OXPHOS genes onto the European seabass genome (seabass_V1.0) at the scaffold level was conducted using the R package chromPlot (Verdugo, 2017). Different colors were assigned to indicate the five OXPHOS complexes, and genes belonging to OXPHOS gene families are highlighted in bold. Asterisk represents identified unmapped genes

While the findings between the two species were relatively similar, some differences were observed. Twenty-four (24) OXPHOS gene families were identified in gilthead seabream and 22 in European seabass (**Table 2.1**) with *atp5me* and *cox6c* making the extra gilthead seabream families. Most shared gene families between the two species had the same gene number, except for *atp5mg*, *cox6b2*, *cox7a1* and *cox7a2l* gene families.

Table 2.1. OXPHOS gene families identified in the genomes of gilthead seabream and European seabass, respectively, and the number of genes in each family

Gene family	Gilthead seabream	European seabass
Complex I		
ndufs1	2	2
ndufs8	2	2
ndufv1	2	2
Complex II		
sdhb	2	2
Complex III		
uqcr11	2	2
uqcrc2	2	2
uqcrfs1	2	2
uqcrh	2	2
Complex IV		
cox4i	2	2
cox5a	2	2
cox5b	2	2
cox6a	2	2
cox6b2	2	3
cox6c	2	1
cox7a1	3	4
cox7a2l	3	2
Complex V		
atp5f1a	2	2
atp5f1b	2	2
atp5if1	2	2
atp5mc	3	3

atp5me	2	1
atp5mg	2	3
atp5pd	2	2
atp5pf	3	3

Evolutionary origin of paralogs

Different mechanisms, including small-scale duplications like single-gene or segmental duplications, and WGD, can lead to the emergence of duplicated genes. In contrast to humans and other non-teleost vertebrates, which have undergone only the first two rounds of vertebrate genome duplication, teleost evolution has resulted in a third round of WGD, the TSGD. Phylogenetic and synteny analyses were performed to determine the evolutionary origin of gilthead seabream and European seabass paralogs. Three main topologies for gene paralogs were identified by phylogenetic analysis of the 24 and 22 gene families of gilthead seabream and European seabass, respectively (**Figure 2.4**). Phylogenetic trees for each gene family can be found in **Figures S2.1-S2.22**.

Topology A indicates a WGD event, most likely the first or second round of vertebrate WGDs, as it demonstrates that both paralogs were descended from a duplication that occurred in the common ancestor of teleosts and non-teleost vertebrates. Topology B indicates a WGD event, most likely the TSGD, as both paralogs were descended from a duplication that occurred at the common ancestor of teleosts. Topology C shows that the grouping of paralogs is species- specific, which suggests either a within-species or a lineage-specific duplication that took place at the outer branches of teleosts and had no other species to be represented in the analysis. The *cox7a* gene family (Figure S1.14) was the only exception to the above, as the observed topology could not be fully explained by the vertebrate WGD events alone, indicating that there was at least one ancestral duplication that resulted in the *cox7a* and *cox7a2l* gene families.

Eighteen (18) and 17 paralog pairs in gilthead seabream and European seabass, respectively, followed topology B, which matches the expected topology of TSGD (**Table 2.2**). The observed topology in the case of the gene families *atp5f1b* (**Figure S2.16**) and *cox7a2l* (**Figure S2.14**) could be explained by TSGD followed by gene loss for a number of lineages; it could also be the

consequence of lineage-specific duplication that occurred at least in the common ancestor of gilthead seabream, European seabass, and greater amberjack. Six gene families have members that are following the topology A (**Table 2.2**) which is in accordance with one of the first two rounds of vertebrate genome duplications. From those, five families were shared between gilthead seabream and European seabass (*atp5if1, atp5mc, atp5pf, cox4i, cox7a1*) and one (*cox6b2*) was European seabass specific. Topology C was observed in two different gene families in each species. Species- or lineage-specific duplication event gave rise to members of *cox6c* and *cox7a2l* gene families in gilthead seabream and to *atp5mg* and *cox7a1* in European seabass. Lastly, the phylogenetic tree for *cox6a* gene family was not resolved.

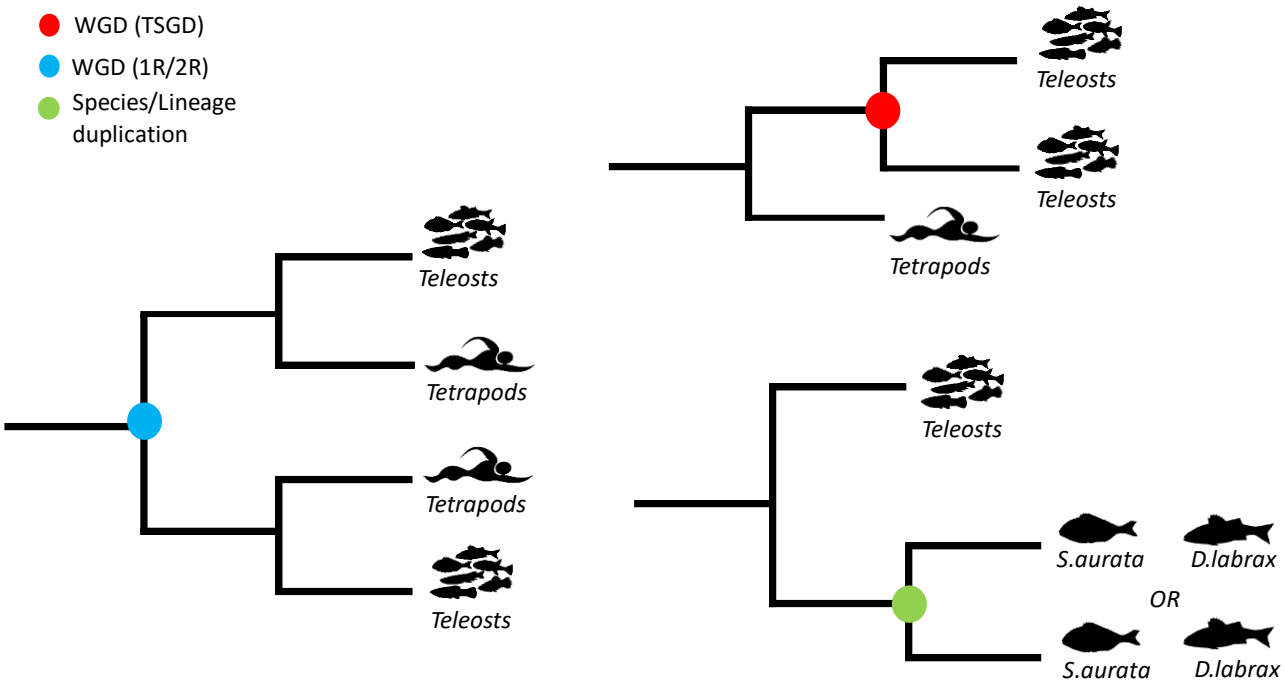


Figure 2.4. Tree main topologies were identified from the phylogenetic analysis. Circles represent duplication nodes and colors different duplication events. A) phylogenetic tree topology which is in accordance with 1R/2R; B) phylogenetic tree topology which is in accordance with TSGD; C) phylogenetic tree topology which is in accordance with species or linage specific duplication.

Table 2.2. Evolutionary origin of each paralog in gilthead seabream and European seabass based on phylogenetic analysis

Gene family	*Topology*	
	Sparus aurata	***Dicentrarchus labrax***
Complex I		
ndufs1	WGD (TSGD)	WGD (TSGD)
ndufs8	WGD (TSGD)	WGD (TSGD)
ndufv1	WGD (TSGD)	WGD (TSGD)
Complex II		
sdhb	WGD (TSGD)	WGD (TSGD)
Complex III		
uqcr11	WGD (TSGD)	WGD (TSGD)
uqcrc2	WGD (TSGD)	WGD (TSGD)
uqcrfs1	WGD (TSGD)	WGD (TSGD)
uqcrh	WGD (TSGD)	WGD (TSGD)
Complex IV		
cox4i	WGD	WGD
cox5a	WGD (TSGD)	WGD (TSGD)
cox5b	WGD (TSGD)	WGD (TSGD)
cox6a	Unresolved	Unresolved
cox6b2	WGD (TSGD)	WGD (TSGD), WGD
cox6c	species specific duplication	-
cox7a1	WGD	WGD, species specific duplication
cox7a2l	WGD (TSGD), species specific duplication	WGD (TSGD)
Complex V		
atp5f1a	WGD (TSGD)	WGD (TSGD)
atp5f1b	WGD (TSGD)	WGD (TSGD)
atp5if1	WGD	WGD
atp5mc	WGD	WGD
atp5me	WGD (TSGD)	-
atp5mg	WGD (TSGD)	WGD (TSGD), species specific duplication
atp5pd	WGD (TSGD)	WGD (TSGD)
atp5pf	WGD, WGD (TSGD)	WGD,WGD (TSGD)

Synteny analysis (**Figure 2.5**, **Figure 2.6**) classified OXPHOS paralogs into four categories (**Table 2.3**):

- **WGD/segmental**, indicating that the paralogs are members of collinear blocks;
- **proximal**, indicating that the paralogs are near each other but interrupted by several other genes;

- **tandem** that indicates that paralogs are next to each other; and
- **dispersed**, indicating that, while the genes are paralogs, they are not near each other on chromosomes nor show conserved synteny.

In gilthead seabream, the relationships between the members of 7 gene families were classified as WGD/segmental, 17 as dispersed and 2 as proximal, while in European seabass 7 were classified as WGD/segmental, 17 as dispersed, 1 as proximal and 1 as tandem (

Table 2.3). In both species, the families classified as proximal and/or tandem are the same that, based on phylogenetic analysis, were characterized as the outcome of species/ lineage specific duplications (Topology C). From the paralog pairs that were inferred as the possible result of TSGD (Topology B), the members of 5 gene families (*sdhb, uqcrc2, uqcr11, cox6b2* and *atp5f1a*) in gilthead seabream and 6 families (*sdhb, uqcrc2, uqcr11, cox6b2, atp5f1a and atp5f1b*) in European seabass were identified as members of collinear blocks, suggesting a conserved chromosomal environment for those genes. This segmental arrangement is expected after a WGD event and it is more probable to be the result of TSGD, which is the most recent such WGD event in the vertebrates.

On the basis of phylogenetic tree topology, members of the remaining gene families for both species that were inferred to be TSGD paralogs, were not identified in collinear blocks and were categorized as dispersed duplicates. Given that they were members of two district teleost clades with a well-supported duplication node at their bases, it was hypothesized that these paralogs are the result of TSGD followed by 1) significant gene losses to the paralogs chromosomal environment, 2) chromosomal rearrangements close to the paralog locations, or 3) translocations mediated by transposons or retrotransposons, at least for the one paralog. The majority of the members of the gene families resulting from one of the first rounds of vertebrate WGDs based on phylogenetic tree Topology A (**Figure 2.4**) were also characterized as dispersed, and their current location may be the result of the same mechanisms as described above, leading to low conservation of the chromosomal gene environment of those paralogs, too.

Table 2.3. Classification of paralogs for both species based on the level of collinearity between the paralogue pairs

Gene name	Collinearity	
	Sparus aurata	***Dicentrarchus labrax***
Complex I		
ndufs1	dispersed	dispersed
ndufs8	dispersed	dispersed
ndufv1	dispersed	dispersed
Complex II		
sdhb	WGD/segmental	WGD/segmental
Complex III		
uqcr11	WGD/segmental	WGD/segmental
uqcrc2	WGD/segmental	WGD/segmental
uqcrfs1	dispersed	dispersed
uqcrh	dispersed	dispersed
Complex IV		
cox4i	dispersed	dispersed
cox5a	dispersed	dispersed
cox5b	dispersed	dispersed
cox6a	dispersed	dispersed
cox6b2	WGD/segmental	WGD/segmental, dispersed
cox6c	proximal	-
cox7a1	dispersed, WGD/segmental	dispersed, tandem
cox7a2l	dispersed, proximal	dispersed
Complex V		
atp5f1a	WGD/segmental	WGD/segmental
atp5f1b	dispersed	WGD/segmental
atp5if1	dispersed	dispersed
atp5mc	WGD/segmental, dispersed	WGD/segmental, dispersed
atp5me	dispersed	-
atp5mg	dispersed	dispersed, proximal
atp5pd	dispersed	dispersed
atp5pf	dispersed	dispersed

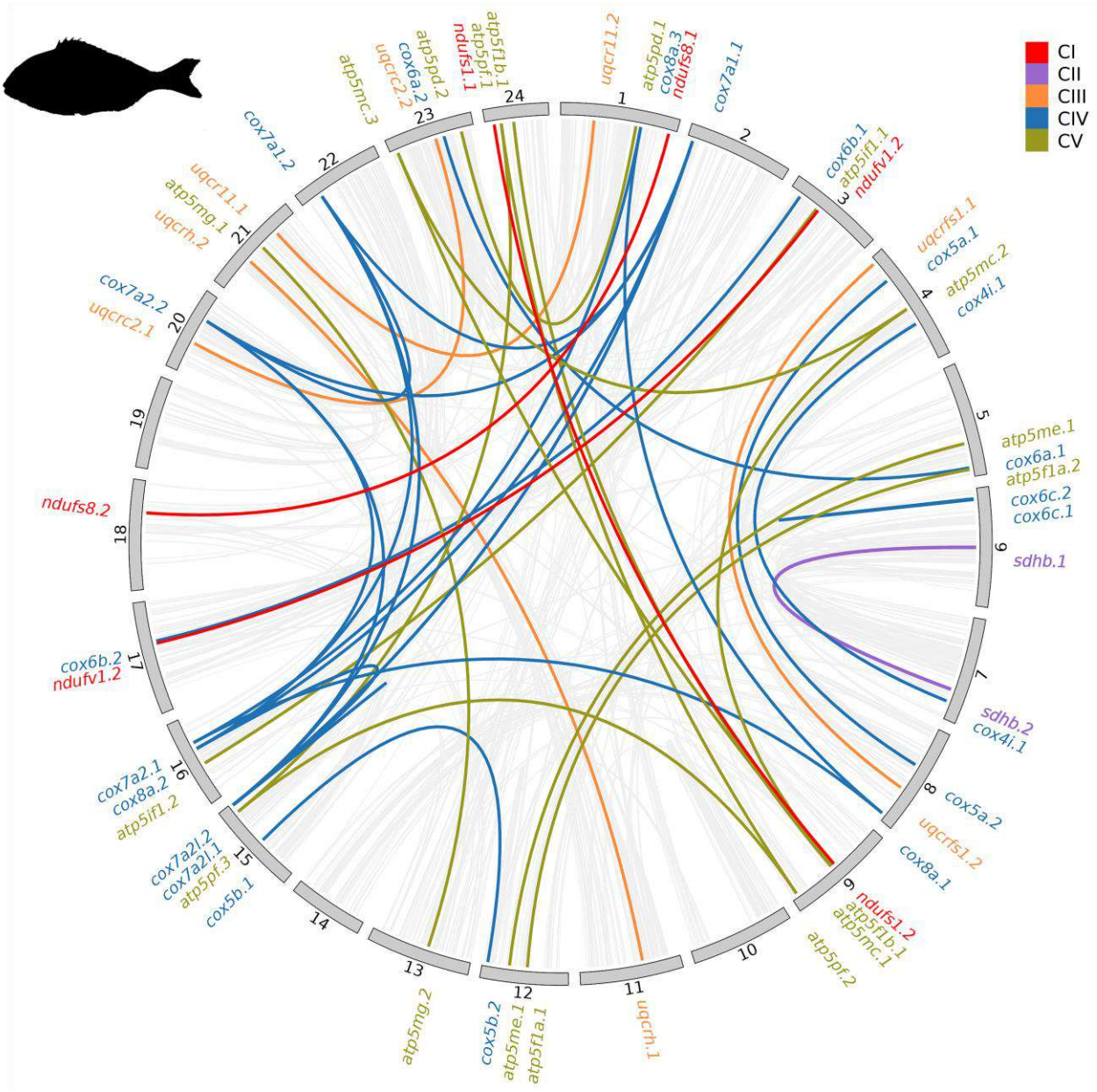


Figure 2.5. Circos plot depicting the paralogue relationships between gilthead seabream genes. Duplicated gene relationships between gilthead sea bream chromosomes are represented by inner lines. Different colors represent the five OXPHOS complexes.

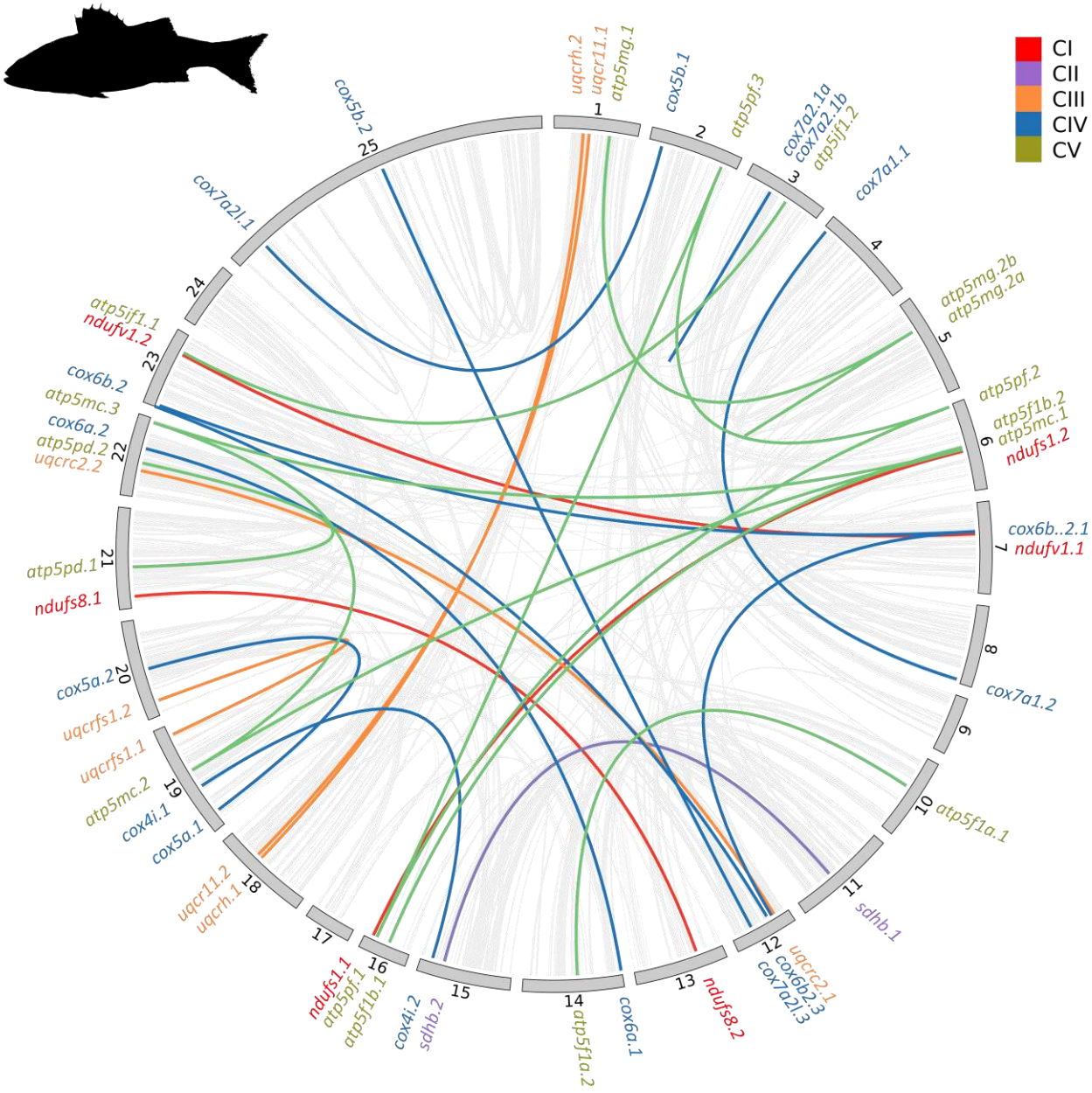


Figure 2.6. Circos plot depicting the paralogue relationships between European seabass genes. Duplicated gene relationships between European seabass chromosomes are represented by inner lines. Different colors represent the five OXPHOS complexes.

Functional characteristics of OXPHOS paralogs

The prediction *in silico* of the subcellular localization for OXPHOS paralogs in both species suggests mitochondria as the most probable localization for the majority of the genes (**Table 2.4**), with similar predictions for both species. In most of the cases the prediction yielded mitochondria and cytosol, the probability for cytosol localization was lower than for mitochondria. In both species all members of gene families *cox6b2* and *uqcrh* were predicted to be localized in cytosol, yet with a relatively low probability for that prediction. While this could be an indication of diverged evolution for those gene families it seems that this is not the case. When the same analysis was followed for the human orthologs of those genes, similar results were yielded as with the fish genes; predictions for both mitochondria and cytosol with a low probability score. Even in the cases without a clear prediction for mitochondria localization, that prediction had always higher probability compared with cytosol. Based on that and in respect to the level of genome and gene annotation of the human genome it is most probable that those genes are also localized in the mitochondria.

Based on InterPros predictions, functional analysis of OXPHOS gene families revealed that the majority of gilthead seabream and European seabass OXPHOS paralogs share protein family membership, molecular function, and biological processes with their human homolog for both species. The few exceptions to the above were the members of *uqcrc2* family in both species and *cox6c* family in gilthead seabream (**Table 2.5**).

The majority of OXPHOS paralogs had conserved gene structure (number of exons- introns; **Figures S2.23-S2.45**). Certain exceptions were recorded and even in those cases a high conservation of protein domains was observed; *ndufs1.2* in gilthead seabream missing the NADH-quinone oxidoreductase domain that exists at the human orthologe and *ndufs1.1* gilthead seabream paralog; *uqcrfs1.2* missing two domains, UCR_TM and Ubiq-Cytc-red_N, in both species studied, while those domains exist in its human ortholog and in the *uqcrfs1.1* paralog of gilthead seabream and European seabass. In addition, there was no prediction for a protein domain in Pfam for European seabass genes *cox7a2.1b* and *cox7a2l.1*.

Table 2.4. Predictions in silico of subcellular localization of the products of OXPHOS paralogs in gilthead seabream and European seabass. Most gene products were predicted to be targeted in mitochondrion. The exceptions are indicated in red

Sparus aurata			*Dicentrarchus labrax*		
Gene name	**Localization**	**Probability**	**Gene name**	**Localization**	**Probability**
Complex I					
ndufs1.1	Cytoplasm	0.7156	*ndufs1.1*	Mitochondrion	0.9113
ndufs1.2	Mitochondrion	0.9175	*ndufs1.2*	Mitochondrion	0.8681
ndufs8.1	Mitochondrion	0.9443	*ndufs8.1*	Mitochondrion	0.9514
ndufs8.2	Mitochondrion	0.9431	*ndufs8.2*	Mitochondrion	0.9398
ndufv1.1	Mitochondrion	0.8658	*ndufv1.1*	Mitochondrion	0.8198
ndufv1.2	Mitochondrion	0.9270	*ndufv1.2*	Mitochondrion	0.9207
Complex II					
sdhb.1	Mitochondrion	0.8927	*sdhb.1*	Mitochondrion	0.8516
sdhb.2	Mitochondrion	0.8736	*sdhb.2*	Mitochondrion	0.8858
Complex III					
uqcr11.1	Mitochondrion	0.9766	*uqcr11.1*	Mitochondrion	0.9752
uqcr11.2	Mitochondrion	0.9645	*uqcr11.2*	Mitochondrion	0.9690
uqcrc2.1	Mitochondrion	0.9682	*uqcrc2.1*	Mitochondrion	0.9669
uqcrc2.2	Mitochondrion	0.9359	*uqcrc2.2*	Mitochondrion	0.9302
uqcrfs1.1	Mitochondrion	0.8842	*uqcrfs1.1*	Mitochondrion	0.9105
uqcrfs1.2	Cytoplasm, Mitochondrion	0.6149, 0.6460	*uqcrfs1.2*	Mitochondrion	0.9415
uqcrh.1	Cytoplasm, Mitochondrion	0.5312, 0.6471	*uqcrh.1*	Cytoplasm, Mitochondrion	0.5497, 0.6586
uqcrh.2	Cytoplasm	0.5133	*uqcrh.2*	Cytoplasm	0.5530
Complex IV					
cox4i.1	Mitochondrion	0.9329	*cox4i.1*	Mitochondrion	0.9374
cox4i.2	Mitochondrion	0.9513	*cox4i.2*	Mitochondrion	0.9257
cox5a.1	Mitochondrion	0.9204	*cox5a.1*	Mitochondrion	0.9286
cox5a.2	Mitochondrion	0.9314	*cox5a.2*	Mitochondrion	0.9572
cox5b.1	Mitochondrion	0.9448	*cox5b.1*	Mitochondrion	0.9208
cox5b.2	Mitochondrion	0.9555	*cox5b.2*	Mitochondrion	0.9339
cox6a.1	Mitochondrion	0.9032	*cox6a.1*	Mitochondrion	0.892
cox6a.2	Mitochondrion	0.9095	*cox6a.2*	Mitochondrion	0.9241
cox6b2.1	Cytoplasm, Mitochondrion	0.5627, 0.6472	*cox6b2.1*	Cytoplasm	0.5835
cox6b2.2	Cytoplasm	0.5038	*cox6b2.2*	Cytoplasm	0.5590
			cox6b2.3	Cytoplasm	0.5590
cox6c.1	Mitochondrion	0.9246			
cox6c.2	Mitochondrion	0.9350			
cox7a1.1	Mitochondrion	0.9392	*cox7a1.1*	Mitochondrion	0.9376
cox7a1.2	Mitochondrion	0.9363	*cox7a1.2*	Mitochondrion	0.9401
cox7a2.1	Mitochondrion	0.9337	*cox7a2.1a*	Mitochondrion	0.9286

			cox7a2.1b	Mitochondrion	0.9008
cox7a2l.1	Mitochondrion	0.9299	*cox7a2l.1*	Mitochondrion	0.9190
cox7a2l.2	Mitochondrion	0.9299			
cox7a2l.3	Mitochondrion	0.9223	*cox7a2l.3*	Mitochondrion	0.9114
Complex V					
atp5f1a.1	Mitochondrion	0.8939	*atp5f1a.1*	Mitochondrion	0.8726
atp5f1a.2	Mitochondrion	0.6299	*atp5f1a.2*	Mitochondrion	0.8594
atp5f1b.1	Mitochondrion	0.8256	*atp5f1b.1*	Mitochondrion	0.7482
atp5f1b.2	Mitochondrion	0.8942	*atp5f1b.2*	Mitochondrion	0.8984
atp5if1.1	Mitochondrion	0.9714	*atp5if1.1*	Mitochondrion	0.9720
atp5if1.2	Mitochondrion	0.9747	*atp5if1.2*	Mitochondrion	0.9737
atp5mc.1	Mitochondrion	0.8216	*atp5mc.1*	Mitochondrion	0.8281
atp5mc.2	Mitochondrion	0.8046	*atp5mc.2*	Mitochondrion	0.7364
atp5mc.3	Mitochondrion	0.7420	*atp5mc.3*	Mitochondrion	0.7385
atp5me.1	Mitochondrion	0.8718			
atp5me.2	Mitochondrion	0.8841			
atp5mg.1	Mitochondrion	0.9382	*atp5mg.1*	Mitochondrion	0.9398
atp5mg.2	Mitochondrion	0.9532	*atp5mg.2a*	Mitochondrion	0.9354
			atp5mg.2b	Cytoplasm, Nucleus	0.5851, 0.5860
atp5pd.1	Mitochondrion	0.8477	*atp5pd.1*	Mitochondrion	0.9097
atp5pd.2	Mitochondrion	0.9526	*atp5pd.2*	Mitochondrion	0.9349
atp5pf.1	Mitochondrion	0.9568	*atp5pf.1*	Mitochondrion	0.9415
atp5pf.2	Mitochondrion	0.9124	*atp5pf.2*	Mitochondrion	0.9010
atp5pf.3	Cytoplasm	0.6992	*atp5pf.3*	Cytoplasm, Mitochondrion	0.5936, 0.6593

Table 2.5. Functional characteristics of OXPHOS paralogs in gilthead seabream and European seabass. The + symbol denotes conservation of each characteristic for each gene with its human ortholog

Sparus aurata				***Dicentrarchus labrax***			
Gene name	Protein family	Biological process	Molecular function	Gene name	Protein family	Biological process	Molecular function
atp5f1a.1	+	+	+	*atp5f1a.1*	+	+	+
atp5f1a.2	+	+	+	*atp5f1a.2*	+	+	+
atp5f1b.1	+	+	+	*atp5f1b.1*	+	+	+
atp5f1b.2	+	+	+	*atp5f1b.2*	+	+	+
atp5if1.1	+	+	+	*atp5if1.1*	+	+	+
atp5if1.2	+	+	+	*atp5if1.2*	+	+	+
atp5mc.1	+	+	+	*atp5mc.1*	+	+	+
atp5mc.2	+	+	+	*atp5mc.2*	+	+	+
atp5mc.3	+	+	+	*atp5mc.3*	+	+	+

atp5me.1	+	+	+				
atp5me.2	+	+	+				
atp5mg.1	+	+	+	*atp5mg.1*	+	+	+
atp5mg.2	+	+	+	*atp5mg.2a*	+	+	+
				atp5mg.2b	+	+	+
atp5pd.1	+	+	+	*atp5pd.1*	+	+	+
atp5pd.2	+	+	+	*atp5pd.2*	+	+	+
atp5pf.1	+	+	+	*atp5pf.1*	+	+	+
atp5pf.2	+	+	+	*atp5pf.2*	+	+	+
atp5pf.3	+	+	+	*atp5pf.3*	+	+	+
cox4i.1	+	+	+	*cox4i.1*	+	+	+
cox4i.2	+	+	+	*cox4i.2*	+	+	+
cox5a.1	+	+	+	*cox5a.1*	+	+	+
cox5a.2	+	+	+	*cox5a.2*	+	+	+
cox5b.1	+	+	+	*cox5b.1*	+	+	+
cox5b.2	+	+	+	*cox5b.2*	+	+	+
cox6a.1	+	+	+	*cox6a.1*	+	+	+
cox6a.2	+	+	+	*cox6a.2*	+	+	+
cox6b2.1	+	+	+	*cox6b2.1*	+	-	+
cox6b2.2	+	+	+	*cox6b2.2*	+	-	+
				cox6b2.3	+	+	+
cox6c.1	+	-	+				
cox6c.2	+	-	+				
cox7a1.1	+	+	+	*cox7a1.1*	+	+	+
cox7a1.2	+	+	+	*cox7a1.2*	+	+	+
cox7a2.1	+	+	+	*cox7a2.1a*	+	+	+
				cox7a2.1b	+	+	+
cox7a2l.1	+	+	+	*cox7a2l.1*	+	+	+
cox7a2l.2	+	+	+				
cox7a2l.3	+	+	+	*cox7a2l.3*	+	+	+
ndufs1.1	+	+	+	*ndufs1.1*	+	+	+
ndufs1.2	+	+	+	*ndufs1.2*	+	+	+
ndufs8.1	+	+	+	*ndufs8.1*	+	+	+
ndufs8.2	+	+	+	*ndufs8.2*	+	+	+
ndufv1.1	+	+	+	*ndufv1.1*	+	+	+
ndufv1.2	+	+	+	*ndufv1.2*	+	+	+
sdhb.1	+	+	+	*sdhb.1*	+	+	+
sdhb.2	+	+	+	*sdhb.2*	+	+	+
uqcr11.1	+	+	+	*uqcr11.1*	+	+	+
uqcr11.2	+	+	+	*uqcr11.2*	+	+	+
uqcrc2.1	+	-	-	*uqcrc2.1*	+	-	-

uqcrc2.2	+	-	-	*uqcrc2.2*	+	-	-
uqcrfs1.1	+	+	+	*uqcrfs1.1*	+	+	+
uqcrfs1.2	+	+	+	*uqcrfs1.2*	+	+	+
uqcrh.1	+	+	+	*uqcrh.1*	+	+	+
uqcrh.2	+	+	+	*uqcrh.2*	+	+	+

Discussion

It is evident that teleosts, a taxonomic group with nearly 30K species, have captured the interest of the scientific community only in the recent years despite their size. By using the search terms "oxidative phosphorylation" and "teleost", or "*Sparus aurata*" in a PubMed search for articles in November 2022, only 42 and 13 entries, respectively, were returned. The same search for European seabass returned just 2 articles. However, when the search terms were changed to mammals and human, more than 30K and 20K entries, respectively, were returned. The impressive volume of literature on mammals and humans is not surprising, considering that most of our knowledge about OXPHOS comes from studies in model organisms given the high interest about the implication of OXPHOS in cancer and other diseases. Another observation while searching the literature on teleosts was that the evolutionary aspects do not fall within the scope of studies on their biological processes including OXPHOS, which has resulted in overlooking the number of gene paralogs and their contribution to teleost biological diversity. This is the first study on gilthead seabream and European seabass that looked into OXPHOS as a whole and attempted to provide the complete landscape of gilthead seabream and European seabass OXPHOS genes.

There were 24 OXPHOS gene families identified in gilthead seabream and 22 in European seabass. Distinct potential origins for those gene families were discovered by phylogenetic and synteny analysis. The majority of paralogs in both species were outgrouped from the single gene of other non-teleost vertebrates and formed two distinct teleost clades. These results indicated that those OXPHOS paralogs diverged from a common ancestor located at the base of teleost lineage. Taking into consideration that those genes can only be identified in teleost species that underwent the

teleost-specific genome duplication (TSGD) (Meyer & Van De Peer, 2005; Voldoire et al., 2017), it is proposed that these genes originated from TSGD. The OXPHOS gene palette in teleosts appears to have been significantly impacted by TSGD. The early teleost genome preserved a sizable number of paralogs after the duplication, but afterwards there were species- and/or lineage-specific gene losses, perhaps reflecting the various evolutionary pressures or potential adaptations for the diverse teleost lineages (Brunet et al., 2006; Inoue et al., 2015). While the majority of gilthead seabream gene families appear to have descended from TSGD, the phylogenetic tree topologies for a number of gene families (e.g. *cox7a, cox4i, atp5mc1, atp5pf, and atp5if1*) suggest that one of the first two vertebrate WGDs rather than TSGD is the most likely origin of those paralogs. In those cases, two clades were again clearly separated from one another and were populated by teleost species genes, yet in at least one of those clades, non-teleost species were present, such as spotted gar or tetrapods representatives were grouped with teleosts. This topology may imply that those paralogs draw their evolutionary origin from one of the first two rounds of vertebrate WGDs, followed by the loss of one paralog from the linages of the non-teleost species that were included in the present analysis (Meyer & Schartl, 1999; Mu et al., 2013). Finally, two gene families were identified in each species that seem to be the result of species/lineage specific gene duplication based on the phylogenetic tree topology and synteny analysis.

Properties of the subunits coded by the identified gene families

Oxidative phosphorylation (OXPHOS), stands out as one of the most complex processes within the cell due to the intricate coordination of its five multiprotein complexes. Its complexity is evident not only in the specific functions of diverse proteins within these complexes, serving catalytic, structural, regulatory and assembly roles, but also in the collaborative formation of supercomplexes. Notably, the findings discussed in this chapter reveal that no specific pattern favoring a particular role was observed. Instead, gene families with retained paralogs appear to encompass all the mentioned functions of OXPHOS subunits. Some interesting observations about the roles of subunits from the identified gene families follow.

In complex I, which is the largest of all five complexes, three gene families (*ndufs1*, *ndufs8*, *ndufv1*) were identified. While the majority of complex I subunits are the supernumerary, mainly contributing to the assembly and the structure of the complex, none of those subunits seems to have retained paralogs. The subunits coded by the identified gene families belong to the seven core subunits of the complex encoded by the nuclear genome. Subunits *ndufs1* and *ndufs8* participate in the iron-sulfur component that is required in the electron transfer process, while *ndufs1* may form part of the active site crevice where NADH is oxidized. In addition, subunit *ndufv1* carries the NADH-binding site as well as flavin mononucleotide (FMN)- and Fe-S-biding sites (Wirth et al., 2016). Similarly, in complex II subunit *sdhb* is an iron sulfur protein one of the two hydrophilic catalytic subunits of the enzyme (Fernandez-Vizarra & Zeviani, 2018).

Complex III gene families encode for catalytic and structural subunits. Subunits *uqcrc11* and *uqcrh* are both members of the supernumerary subunits, they act as structural components of the enzyme, and they both play a role at the assembly. On the contrary subunit *uqcrc2,* besides its role in the assembly, exhibits peptidase activity along with *uqcrc1*. Subunit *uqcrfs1* is the last subunit to get incorporated in complex III, and it's a member of the catalytic core of the enzyme and essential for enzymatic activity (Fernandez-Vizarra & Zeviani, 2018).

In complex IV, there are eleven different proteins among the nuclear-encoded subunits, with six of them existing in isoforms in humans. The catalytic core of complex IV is constituted by three mitochondria-encoded subunits (COX1, COX2, COX3), while the nuclear-encoded subunits, although not essential for catalytic function, build the conserved catalytic core. These nuclear-encoded subunits may modulate COX activity under diverse conditions, and some of them have tissue-specific and development-specific isoforms, potentially allowing for the fine-tuning of COX function in individual tissues (Čunátová et al., 2020). Between those subunits there were 8 and 7 gene families identified for gilthead seabream and for European seabass respectively, adding to the complexity of teleosts complex IV gene pallet.

A number of gene families of complex V were identified in both species under study sharing between them catalytic structural and assembly roles. Two gene families, *atp5f1a* and *atp5f1b,* code for the water soluble catalytic domain F1 found in the mitochondrial matrix. In addition, the

atp5mc gene family codes for F0 domain of the complex where protons enter the matrix through the F0 pore and the F0 motor rotates anticlockwise to turn around F1, thereby driving ATP synthesis. Subunits d and F6 coded by the members of gene families *atp5pd* and *atp5pf* respectively, are part of the peripheral stalk, which connect F1 to F0 and keep the stators from spinning along with the rotor (Neupane et al., 2019; Signes & Fernandez-Vizarra, 2018). Furthermore, and probably the most interesting complex V subunits are e and g which are coded by the members of *atp5me* and *atp5mg*. These subunits play a role in the formation of this supramolecular structure. Deleting the genes responsible for these subunits doesn't impact the assembly or activity of the ATP synthase monomer but specifically influences its dimerization and oligomerization. (Habersetzer et al., 2013). The inner membrane of the mitochondrion folds inwards, forming the increasing the ATP-generating surface area and exerting an impact on ATP synthesis (Afzal et al., 2021). The process of ATP synthase dimerization and oligomerization plays a role in the creation of cristae, and subunits e and g from the mitochondrial ATP synthase are crucial in shaping the mitochondrial cristae morphology. The lack of subunits e and g, and consequently, ATP synthase supercomplexes, leads to alterations in mitochondrial ultrastructure (Habersetzer et al., 2013; Paumard et al., 2002). Furthermore, there is supporting evidence suggesting that the maturation of cristae, reliant on ATP synthase, represents a crucial developmental process essential for differentiation, independent of oxidative phosphorylation in *Drosophila* (Teixeira et al., 2015).

Although, as noted at the start of this section, there was no discernible pattern favoring a specific role, and the subunits coded by members of the identified gene families encompass for catalytic, structural, regulatory, and assembly functions, a pattern does emerge based on the molecular characteristics of some subunits. In complexes I, II, and III, the subunits coded for by the identified gene families were either iron-sulfur proteins (*ndufs1*, *ndufs8*, *sdhd*, and *uqcrfs1*) or featured an Fe-S-binding site (*ndufv1*). Fe–S clusters play an important role in the production of mitochondrial reactive oxygen species (ROS). Mitochondrial Fe–S clusters also play a critical role in the pathogenesis of disease (Read et al., 2021). While CIV is less prone to produce ROS (Zhao et al., 2019) there are a number of ROS production sites between the other three complexes of respiratory chain (Mazat et al., 2020). Complex I has two sites, I_F (FMN site) and I_Q (CoQ binding

site), complex II has one site the II_F, which is linked to succinate dehydrogenase, and the production of reactive oxygen species (ROS) by site II_F is insignificant under normal circumstances. Nonetheless, the elevated ROS levels observed in diseases associated with complex II mutations predominantly originate from site II_F. Finally, complex III has two ROS production sites the Qo and Qi. Overall, the identified OXPHOS gene families encode for subunits that participate in three of the total five ROS production sites of respiratory chain - *ndufv1* and *ndufs1* are part of the I_F, *sdhd* is part of the II_F and *uqcrfs1* is part of Qo. Moreover, while the majority of ROS are generated in the matrix, the Qo site is the only one which ROS are partly extruded in the intermembrane space (Mazat et al., 2020) thus in closer proximity to the cytosol. This highlights an additional crucial function of oxidative phosphorylation (OXPHOS) beyond energy generation: the production of reactive oxygen species (ROS). Apart from the adverse effects of ROS-induced oxidative stress on mitochondria and, consequently, the cell, ROS play a vital role as intracellular messengers in processes such as proliferation, migration, adhesion, and differentiation (Huo et al., 2009; Ray et al., 2012), where besides the implication of those in cancer they are the fundamental biological proses of development.

Evolutionary fate of OXPHOS paralogs

A number of different evolutionary events that took place at different points in the history of vertebrate evolution appear to have shaped OXPHOS in teleosts. For most of the cases, WGDs were the evolutionary origin of that extra genetic material that, after gene losses and gains based on species- or lineage-specific pressures and adaptations, have shaped the OXPHOS gene palette of teleosts as a whole. However, variations were spotted in the gene families and the number of members in each family between the various teleost lineages and even between gilthead seabream and European seabass, which can be considered closely related species. While the number of OXPHOS gene families is high in both species, the majority of OXPHOS genes come in a single copy, suggesting non-functionalization as the fate of the majority of the initially generated OXPHOS paralogs following the WGDs. For a few of the genes retained in the teleost genomes there was some evidence pointing towards neo functionalization as the possible

evolutionary fate. An additional subcellular localization besides mitochondria was identified *in silico* for those genes in both species, while missing protein domains were observed in some of them; both of those findings could be an indication of functional divergence. While the above is a possible scenario, taking into consideration that this evidence is produced by *in silico* analysis and that the genomes of those species are relatively new assemblies with automatic annotation, it is possible that those findings are based on sequencing and/or annotation artifacts. Nevertheless, these genes are good candidates for further investigation with experimental designs to verify or not their subcellular localization and their protein functions.

Based on the findings of this study indicating localization in the mitochondria, conservation of peptide structure as well as molecular and biological function conservation, it seems that the majority of OXPHOS paralogs in both species have retained their ancestral function. At this point is crucial to consider a well documented mechanism of paralog retention, the dosage effect. Should the products of genome duplication be subjected to selection for absolute gene product dosage, they will more likely be retained. This is an almost unique property of WGD duplicates where all the genes of metabolic pathways or multiprotein complexes are retained in a way that the overall stoichiometry is not disrupted (Birchler & Yang, 2022; Song et al., 2020). OXPHOS is clearly a biological function with those characteristics. Taking into consideration that a) synteny and phylogenetic analysis have shown that in both species the identified paralogs were originated by WGDs and most of them by TSGD, and b) the level of functional conservation between the paralogs in both species, the following potential evolutionary mechanism of OXPHOS paralogs for teleosts in respect to gilthead seabream and European seabass is proposed:

Originally after the WGD events the whole OXPHOS palette would be duplicated, and the paralogs would be retained respecting dosage balance. OXPHOS is a metabolic pathway implemented by five multiprotein complexes and it is a fundamentally crucial function for life. Assuming that changes in the original stoichiometry of OXPHOS could have catastrophic results, there would be evolutionary pressure for dosage balance retention at least early after the genome duplication event. The prolonged period of retention, on the other hand, would allow for the accumulation of substitutions, which upon release from dosage balance constraints, could lead subsequently to either neo-functionalization or sub-functionalization. While only a few

evidence provided from subcellular localization analysis and rather controversial at the moment, were found to support neo functionalization, the most probable mechanism for retention, when the original impact of dosage balance would decay over time, would be that of regulatory neo/subfunctionalization and a combination of gene loss and hypofunctionalization (Birchler & Yang, 2022).

Chapter 3

Genetic polymorphism and selection pressure in OXPHOS genes

Introduction

Paralogue genes can evolve independently through the accumulation of different mutations in each paralog, resulting in divergence in their sequences and functions. Gene mutations are changes in the DNA sequence of a gene that can occur spontaneously or as the result of environmental stimuli. Should these changes in DNA sequence provide a selective advantage to the organism, they can become fixed in the population through natural selection.

Following a gene duplication, the two resulting paralogous genes initially perform the same function. However, mutations can accumulate in each paralogous gene over time, resulting in functional divergence and ultimately dictating their fate. Mutation can cause gene paralogs to become nonfunctional by interrupting or deleting the ancestral function of one or both paralogs, resulting in loss of function, i.e. nonfunctionalization. There are various ways in which mutations can lead to loss of function (Behe, 2010). Frameshift mutations arise when a nucleotide is added or deleted from the DNA sequence, resulting in a shift in the reading frame. This can result in a nonfunctional shortened or altered protein. Nonsense mutations occur when a premature stop codon is introduced into the DNA sequence, resulting in a nonfunctional shortened or changed protein. Loss of regulatory components, such as enhancers or promoters, can result in reduced or no expression of one or both paralogs. Nonfunctionalization is an essential mechanism for the evolution of gene families that can result in reduced functional diversity and complexity.

Mutations can result in neofunctionalization of gene paralogs by introducing novel functions into one of the paralogs that were not present in the ancestral gene. Mutations may result in neofunctionalization (Teshima & Innan, 2008) of gene paralogs in numerous ways. Nonsynonymous mutations can affect the amino acid sequence of one of the paralogs, causing changes in the biochemical characteristics of the protein. Insertions or deletions can change the reading frame of a paralog, resulting in the synthesis of a protein with a different amino acid sequence. Gene fusion events can occur when two or more genes are joined together to form a single gene. Changes in gene regulation can alter the expression pattern of one of the paralogs. All the above result in the acquisition of a new function that the other paralog does not share.

Neofunctionalization is an important mechanism for the evolution of gene families that can result in enhanced functional diversity and complexity.

Mutations can also subfunctionalize gene paralogs by changing the expression patterns or biochemical properties of each paralog, resulting in complementary roles. There are various ways in which mutation might result in gene paralog subfunctionalization (Birchler & Yang, 2022). Regulatory mutations can change the expression patterns of each paralog, resulting in differential expression in different tissues or stages of development. Nonsynonymous mutations can affect the amino acid sequence of each paralog, altering its metabolic characteristics. As a result, the two paralogs may have different biochemical roles that are required for full expression of the ancestral gene function. Mutations can change the splicing patterns of each paralog, resulting in the formation of distinct protein isoforms. Overall, mutations have the potential to subfunctionalize gene paralogs by modifying their expression patterns, physicochemical characteristics, or splicing patterns. Subfunctionalization is a key strategy for gene family evolution that can lead to greater functional variety and complexity.

The different evolutionary trajectories are shaped by the different types of mutations and their results as well as the evolutionary pressures applied. Divergent, neutral, and purifying evolution will ultimately dictate the fate of the paralogs. When paralogs evolve to accomplish district roles, functional diversification occurs. (Beisswanger & Stephan, 2008). Neutral evolution happens when paralogs accrue alterations that do not influence their function or fitness. When paralogs retain their original function by minimizing the accumulation of harmful mutations, this is referred to as purifying (or negative) evolution. Negative selection is frequently at work in this type of evolution, which might result in subfunctionalization and/or the preservation of ancestral functions (Rodgers-Melnick et al., 2012). Selection pressures, functional redundancy, and gene dosage can all impact the specific type of evolution in paralogue genes.

Despite its intricacy, the OXPHOS pathway is extensively conserved throughout eukaryotic species. However, it is susceptible to gene mutations that can impair its activity. Mutations in mitochondrial and nuclear-encoded OXPHOS genes can cause OXPHOS malfunction, which can lead to a variety of illnesses and disorders, including mitochondrial disorders, neurological

diseases, and cancer. The integrity of oxidative phosphorylation (OXPHOS) and the associated metabolic processes are impacted by mito-nuclear interactions. Therefore, these interactions should undergo strong selection. In the absence of external factors, deleterious mutations in either the mitochondrial or nuclear genome should be rapidly eliminated to ensure normal function. Selection should promote combinations of alleles in the mitochondrial and nuclear genomes that work in synergy, leading to these combinations becoming more prevalent within populations over evolutionary periods. In this chapter the number and type of mutations of OXPHOS paralogs was investigated along with the types of selection acting on the polymorphic sites in gilthead seabream and seabass. Additionally, the ways of how the genetic variability of OXPHOS paralogs can dictate their evolutionary fate is discussed.

Materials and methods

SNP identification and selection analysis

To identify SNPs to the transcript sequences of gilthead seabream and European seabass OXPHOS paralogs, the transcriptome data for both species generated in the context of Chapter 4 were used, in addition to whole genome resequencing data from four individuals of gilthead seabream, one from France and three from Greece, and from five individuals of European seabass from Greece.

According to the manufacturer's instructions, PureLink Genomic DNA Mini Kit (Invitrogen, Catalogue number: K182002) was used for genomic DNA extraction. Agarose electrophoresis was used to determine the quality of the DNA, and the Qubit dsDNA BR Assay Kit was used to determine its amount (Invitrogen, Catalogue number: Q32850). The DNA samples were delivered to Novogene, where they underwent Illumina HiSeq 3000 sequencing at a mean sequencing coverage of 100x, assuming a uniform 25x coverage for each sample. FastQC (Andrews, 2010) was used to evaluate the quality of FASTQ files, and any sequences with PHRED scores lower than 30 were discarded. From the Ensembl Database, the reference genomes (fSpaAur1.1, INSDC Assembly GCA 900880675.1 and seabass_V1.0, INSDC Assembly GCA_000689215.1) were

obtained and indexed with bwa (H. Li & Durbin, 2009). Using bwa-mem (H. Li, 2013), pairs of sequences that satisfied the quality standards were mapped in the indexed genome. The SAM files of the alignments were converted to BAM files, sorted and indexed for the subsequent analysis using SAMtools (H. Li, 2011). Individual BAM files were merged using SAMtools. The merged BAM file was used for the variant calling using bcftools (H. Li, 2011). The quality filtering of SNPs was performed using bcftools with the following arguments, QUAL>20 and DP>10. The variant annotation and effect prediction was carried out using the tool SnpEff (Cingolani et al., 2012), with the gtf annotation file provided by Ensembl as input. FEL (Kosakovsky Pond et al., 2006) algorithm as included in Datamonkey (Delport et al., 2010) was used to estimate the rates of synonymous and non-synonymous substitutions and to identify polymorphic sites under positive or negative selection.

Results

The total number of SNPs was identified along with the position of each SNP (CDS, UTR's) and the type of mutation (synonymous, missense) for all OXPHOS genes in both species. Following the identification of the polymorphic sites at the coding sequences, selection analysis revealed the evolutionary mode in action for each position, again for all OXPHOS genes in both species. The total number of polymorphic sites, their location and the selection mode in action for both species is presented in **Table 3.1**, detailed catalogs of mutations per gene for the two species are provided at supplementary **Tables S3.1** and **S3.2**. No mutations were found at seven (7) genes (*ndufv1.2, uqcrfs1.2, cox6a.2, cox7a2.1, atp5mc.2, atp5me.1 and atp5mg.2)* and nine (9) genes *(ndufv1.1, uqcr11.2, uqcrfs1.1, uqcrh.1, cox7a2.1b, cox7a2l.1, atp5f1b.1, atp5pd.1, atp5mc.2)* in gilthead seabream and European seabass, respectively. As a result, three (3) gene families *(ndufv1, uqcrfs1 and cox7a2)* with no members bearing mutations were identified in common between both species. The majority of mutations in CDSs were synonymous and there was strong neutral or purifying selection in action upon those sites. Diversifying selection was observed only in four (4) genes of gilthead seabream (*ndufs1.1, atp5f1b.1, atp5pf.2, atp5pf.3*). Different rates of mutation accumulation were observed between paralogs (**Figure 3.1, Figure 3.2, Figure 3.3**);

in several gene families one paralog appeared to be highly conserved with just few or even no mutations identified on its sequence, whereas the other paralog was the one with higher rate of mutation accumulation (e.g. *ndufv1, uqcr11, uqcrfs1, atp5pd, atp5me*); on the other hand, the paralogs of several gene families appeared to accumulate mutations at the same rate (e.g. *uqcrc2, cox4i, cox5a, cox6c, atp5f1a, atp5if1*); lastly, five (5) gene families (*uqcrc2, cox4i*, *atp5f1a*, *atp5if1*, *atp5mc*) exhibited similar pattern of mutation accumulation between the paralogs between the two species.

Table 3.1. The number of SNPs identified in the 52 genes of gilthead seabream and the 50 genes of European seabass, respectively, under different selection mode, and at different gene domains

	Gilthead seabream	***European seabass***
Purifying	114	40
Diversifying	18	0
Neutral	129	73
Synonymous	182	96
Missense	99	30
5'UTR	42	15
3'UTR	148	60
Splice region	14	1

Discussion

The fact that the majority of identified mutations at coding region for the genes under study where synonymous substitutions, and that for 28 and 30 genes for gilthead seabream and European seabass respectively, no missense substitutions where observed may imply a selection pressure to preserve the gene and the structure and/or function of the encoded protein. It's worth noting that a considerable number of the identified mutations were found in the untranslated regions (UTRs) of the transcripts. UTRs are mRNA molecule sections that do not code for protein but exert regulatory functions on gene expression. UTR mutations can impact mRNA stability, translation efficiency, and interactions with regulatory molecules (Steri et al., 2018). The occurrence of mutations in UTRs may indicate that these areas are more prone to genetic alterations without significantly affecting protein function, although it may influence

gene regulation. In contrast, the number of mutations found in the coding sequence (CDS) may suggest that the protein-coding region has undergone some evolutionary changes. However, most of these mutations are synonymous substitutions that do not change the actual amino acid sequence of the protein. Selection analysis suggests that the mutations identified in the CDS are susceptible to either neutral or purifying selection. When mutations do not give a selective advantage or disadvantage to the organism, they are referred to as neutral selection, and they can accumulate over time. Purifying selection, on the other hand, implies that there is a selective pressure to keep the existing protein structure and function, removing mutations that would compromise the protein's normal activity (Soskine & Tawfik, 2010). Purifying selection in this situation suggests that the functional integrity of the protein is critical for the organism, and mutations that alter its function are likely to be harmful.

In conclusion, the location of mutations, the frequency of synonymous replacements, and the existence of neutral or purifying selection all imply that the gene and protein under study have experienced minimal evolutionary alterations. The detected mutations and selection patterns provide insights into the evolutionary constraints and that the duplicated genes have experienced selection pressure to preserve their functional integrity. Given the whole-genome duplication event, the duplicated genes may have undergone subfunctionalization or neofunctionalization processes. The observed conservation level and purifying selection in the duplicated genes (Nielsen et al., 2010), as well as the fact that these characteristics are similar across gene families, imply that subfunctionalization may be the most plausible scenario. The duplicated gene copies may have kept the ancestral functionalities and are susceptible to selective pressure to maintain their respective roles. However, further research, like functional analysis or expression studies, and a wider population screening would be required to offer more detailed insights into the functional divergence of these duplicated genes.

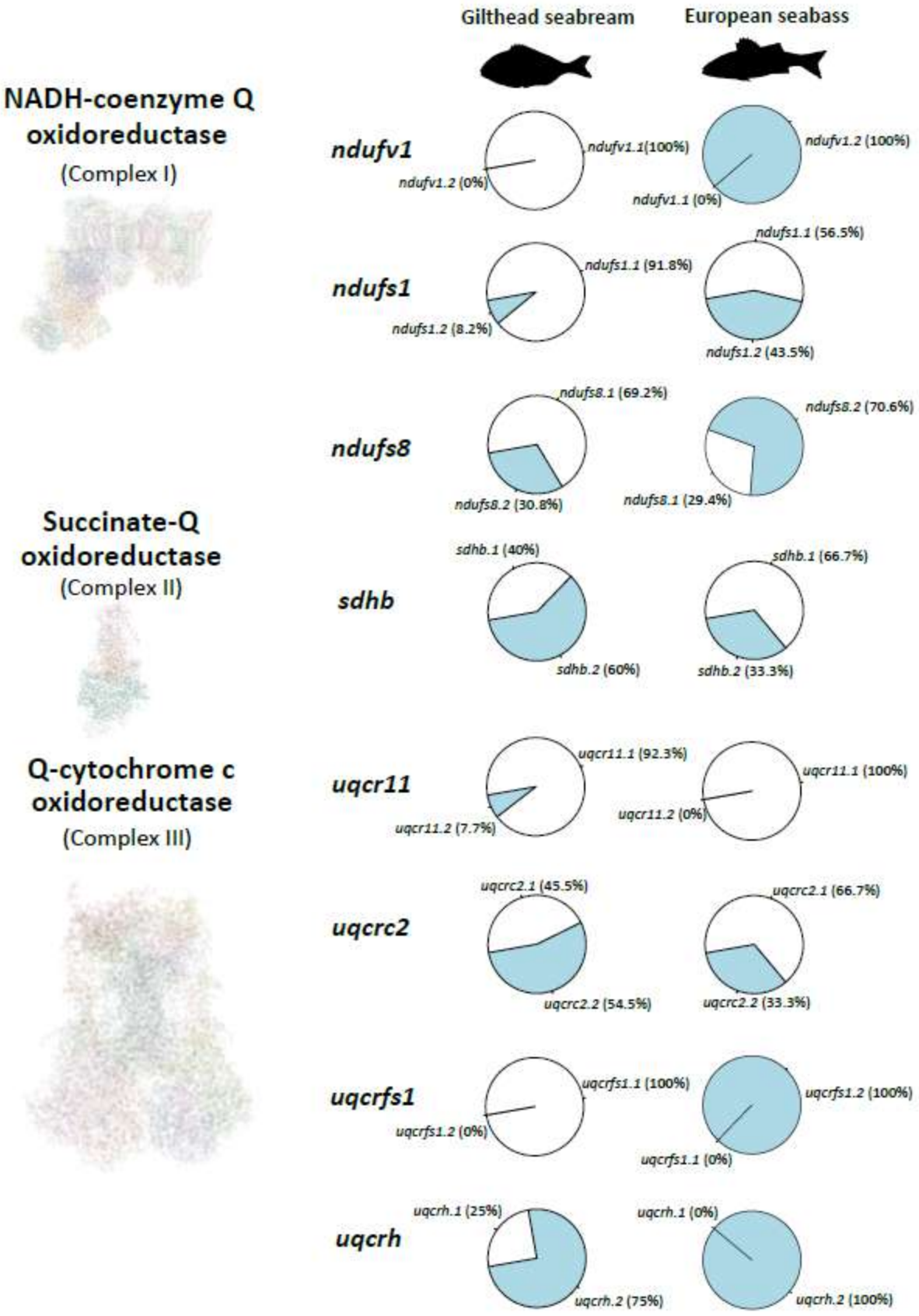


Figure 3.1. Distribution of mutation of paralogs per gene family for seabream (left) and seabass (right), for OXPHOS complexes I, II and III.

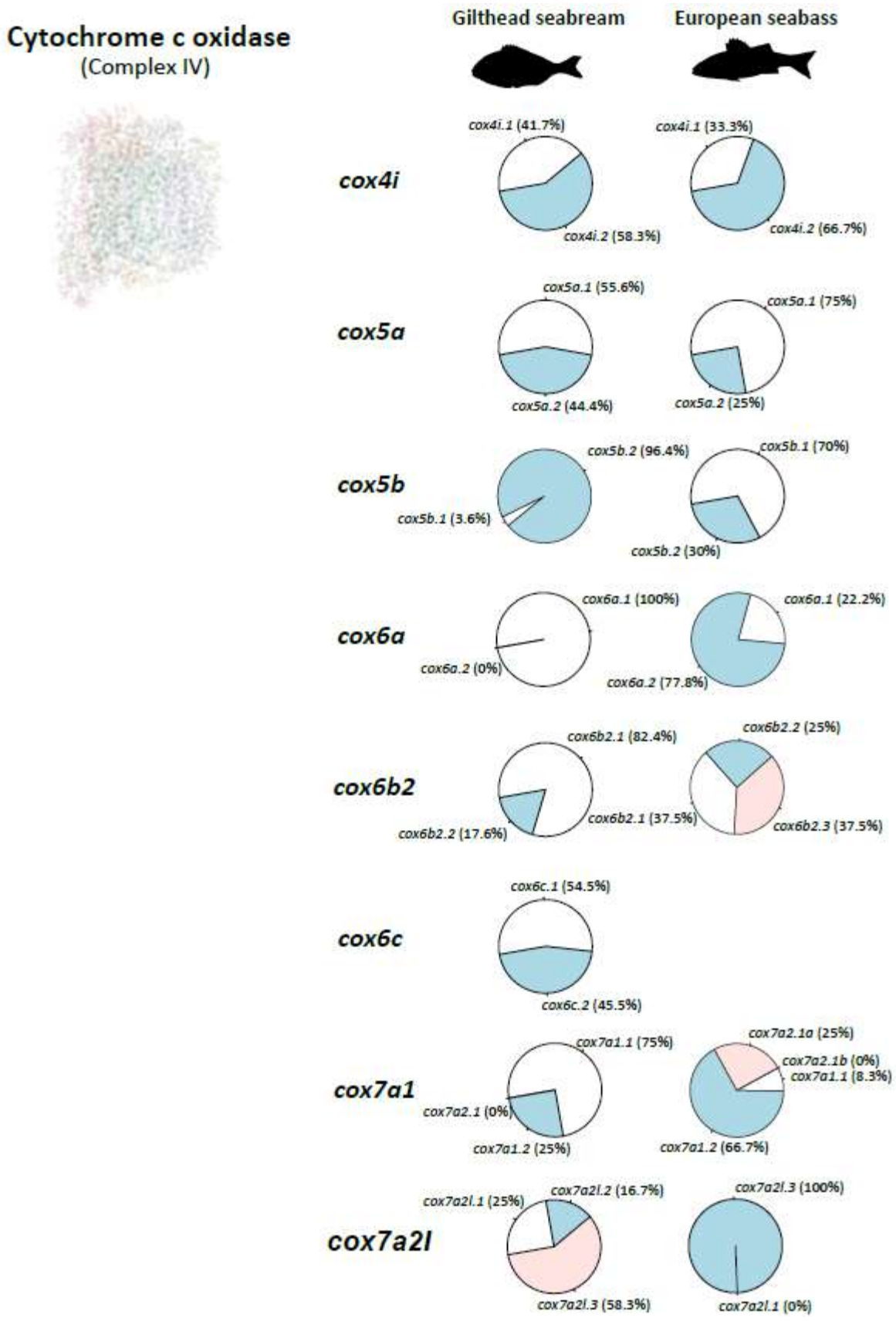


Figure 3.2. Distribution of mutation of paralogs per gene family for seabream (left) and seabass (right), for OXPHOS complex IV.

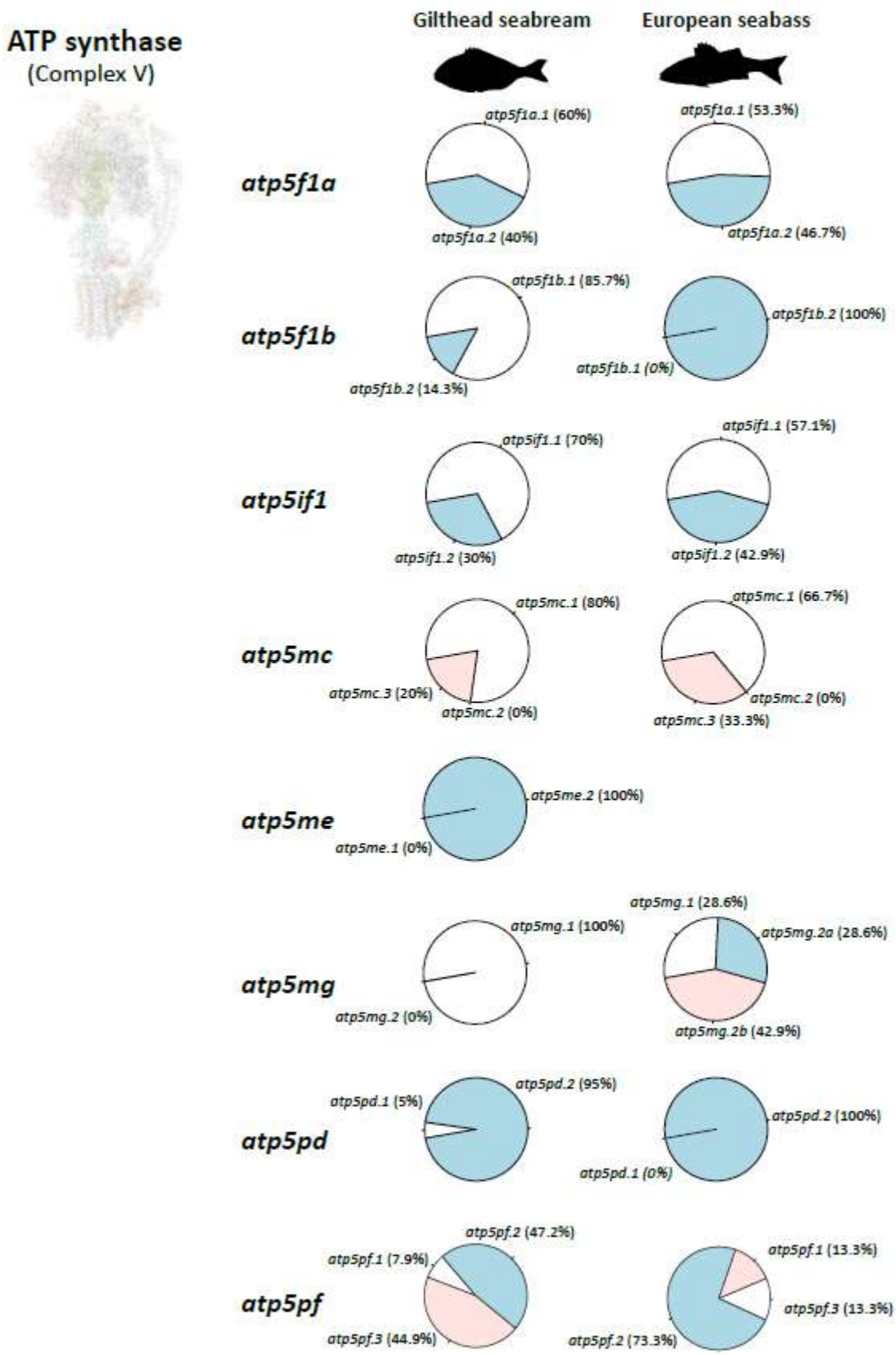


Figure 3.3. Distribution of mutation of paralogs per gene family for seabream (left) and seabass (right), for OXPHOS complex V.

Chapter 4

The fate of OXPHOS paralogue genes

Introduction

Following a gene duplication event in the genome, the duplicated genes can evolve independently through the accumulation of different mutations. Over time, the accumulation of mutations in the paralogs might result in the emergence of novel gene functions, or they can become non-functional or even disappear from the genome. Mutations can drive gene paralogs to become nonfunctional by interrupting or deleting the ancestral function of one or both paralogs, resulting in nonfunctionalization (Lynch & Force, 2000). Mutations can also result in neofunctionalization (Drobek, 2022) of gene paralogs, whereby one of the paralogs adopts new functions that were not performed by the ancestral gene. Finally, mutations may lead to the subfunctionalization (Liu et al., 2018) that entails complementary functionalities that are necessary for the ancestral gene function to be fully expressed.

The mechanisms underlying neofunctionalization and subfunctionalization can be coding and regulatory. Coding subfunctionalization and neofunctionalization refer to the development of unique protein functions for gene paralogs, whereas regulatory subfunctionalization and neofunctionalization refer to the evolution of diverse regulatory roles. Regulatory subfunctionalization occurs when the regulatory regions of the ancestral gene are divided between the two paralogs, with each paralog controlling a part of the expression pattern of the ancestor gene (Drobek, 2022). As a result, the two paralogs may exhibit tissue- or developmental-specific expression patterns. Neofunctionalization can be the result of the acquisition of new regulatory elements or the modification of existing ones, resulting in the emergence of a new distinct pattern of gene expression for one of the paralogs. It is widely accepted that coding and regulatory subfunctionalization and neofunctionalization are major mechanisms underlying the development of unique functions or roles for gene paralogs, either at the protein or regulatory level.

Chapter 2 provided evidence that several OXPHOS paralogs, products of WGD events, have survived in gilthead seabream and European seabass, and they encode proteins in all five OXPHOS complexes. *In silico* analysis indicated mitochondrial localization for the vast majority of paralogs. Chapter 3 mapped the genetic polymorphism of OXPHOS genes to reveal a) a high

frequency of synonymous replacements and neutral or purifying selection in action in CDS, indicative of evolutionary constraints and selection pressure to preserve the paralogs their functional integrity as components of OXPHOS machinery, b) a high number of SNPs mapped in the gene UTRs pointing to plasticity in gene regulation. This chapter moves forward to exploring the expression patterns of OXPHOS paralogue genes as a means to gain insight in their mode of functionality.

Wide transcriptome analysis was performed in distinct developmental stages in gilthead seabream and European seabass larvae reared in controlled environments for the purpose. Early development is regarded as the most crucial and challenging period in fish life. In fish, gene paralogs have been found to play important roles in various aspects of development. The development of the nervous system (Annona et al., 2022; Filippi et al., 2010), white muscle (Georgiou et al., 2016; Macqueen et al., 2010) and immune system (Liongue et al., 2012) are only a few areas of fish development where the redundancy and variety in gene function provided by the duplication of genes within the fish genome have considered crucial. Similarly, gene paralogs might exhibit a tissue-specific expression pattern. Several studies have reported differential expression of members of OXPHOS gene families in tissues at osmoregulation (Martos-Sitcha et al., 2016; Su et al., 2022), following dietary stimulus (Silva-Marrero et al., 2017; Song et al., 2020), at host-microbe interaction and infection (Lohman et al., 2017), hypoxia (Kocha et al., 2015; Martos-Sitcha et al., 2017), and temperature acclimation (Duggan et al., 2011) Following evidence provided by the transcriptome analysis, the expression pattern of paralogs of complex III in European seabass was analyzed in 21 tissues of adult individuals.

Materials and methods

Larvae rearing

The pseudo-green approach was used for rearing gilthead seabream and European seabass, in accordance with the accepted methodology for intense larval rearing (Papandroulakis et al., 2001). Larval rearing was performed at the Institute of Marine Biology, Biotechnology and

Aquaculture (HCMR). Recirculation systems with 500 L, black, cylindro-conical tanks (one tank per duplicate) and biological filters were utilized. Chillers and heaters were used to change the water's temperature to the required levels for the tests, and they were automatically controlled by electronic sensors (Eliwell, EU). From the egg stage (epiboly) until mid metamorphosis, the temperature was adjusted at 20 °C and 17ºC for gilthead seabream and European seabass, respectively. Larvae were successively fed on enriched (DHA Protein Selco, INVE) rotifers until they reached 6 mm total length (TL), enriched rotifers and *Artemia* nauplii (Easy DHA Selco, INVE) when they were between 6- and 10-mm TL and enriched *Artemia* nauplii after 10 mm TL. The rotifer concentration in the tanks was adjusted twice daily to 6–8 ind ml^{-1}. The concentration of *Artemia* nauplii was adjusted twice daily to 0.5–2 ind ml^{-1}. During the rotifer-feeding period, live microalgae *Chlorella* sp. were added thrice daily at a concentration of 6.5 ± 3. 10^5 cells ml^{-1}. Weaning to inert diets (O-range, INVE) started when fish were approximately 8 mm TL. During the rotifer-feeding period, the excess rotifers were removed from the culture medium by means of planktonic traps (50 µm mesh size), positioned at the water-outlet of tanks. During the first twenty days of feeding, air-blowing skimmers were installed to keep the water surface free from lipids and enhance swim bladder inflation. Water oxygen saturation was 5.8 ± 1.0 to 6.0 ± 0.9 mg/L^{-1}, pH 7.8 ± 0.3, salinity 35–36 ‰ and photoperiod 18L:6D. The water turnover rate ranged from 20% of the tank volume d^{-1} during the autotrophic phase to 70% h^{-1} at the end of larval rearing. Water was pumped from a deep (> 100 m) borehole.

Larvae sampling

The same protocol for larvae sampling was used for both species. A representative sample from each tank was collected, larvae were anaesthetized using 2-Phenoxyethanol (Sigma), rinsing with clean sea water and then placing in fixative, RNA later® solution; stored at 4°C overnight (to allow the solution to thoroughly penetrate the larvae), then at –20°C for longer-term storage. From both species samples were collected at three developmental stages, namely first feeding (FF), notochord flexion (FL), and mid metamorphosis (MM).

RNA extraction and transcriptome sequencing

Using the E.Z.N.A.® Total RNA Kit I, RNA was isolated from the total body of individuals at each developmental stage in both species. In order to have an adequate total RNA yield 3-5 individuals were pooled to generate a sample at FF, whereas individual RNA extractions were performed at FL and MM. Subsequently, twelve RNA samples were pooled for each developmental stage, and the pooled samples were sent to Novogene inc. (Europe) for evaluation. The RNA fragments were utilized to create cDNA, construct libraries and sequence the data on the HiSeq 2000 platform (Illumina, USA) using 150bp paired-end reads, producing about 30M reads per sample with over 97% of the reads passing Q20.

Bioinformatics analyses

Fastp (Chen et al., 2018) was used to initially process the raw data (raw reads) in FASTQ format). In this stage, reads with low quality and reads including adaptor and poly-N sequences were removed from the raw data to provide clean data (clean reads). The reference genome and gene annotation files were directly obtained from the Ensembl genome browser website. Using the HISAT2 program (Kim et al., 2019), paired-end clean reads were mapped to the reference genomes (fSpaAur1.1, INSDC Assembly GCA 900880675.1, July 2019 and seabass_V1.0, INSDC Assembly GCA_000689215.1, Apr 2014). Based on recommendations for datasets lacking biological replicates, edgeR (Robinson et al., 2010) was used to analyze the differential expression of two conditions. By utilizing the Benjamini and Hochberg approaches, the P values were modified. For significantly different expression, a corrected pvalue of 0.05 and a log2 (Fold Change) of 0.6 were established as the thresholds. Following that, data on the genes of interest were retrieved using R base commands (R Core Team, 2022; RStudio Team, 2022).

Validation of expression pattern of OXPHOS paralogs

Samples were collected from twenty-one tissues from four adult individuals (two males and two females) of European sea bass. The collected tissues were: brain (BR), pituitary gland (PIT), eye (E), gill (G), heart (H), spleen (SP), adipose tissue (AT), liver (L), stomach (S), posterior intestine (PI), pyloric caeca (PC), anterior intestine (AI), kidney (K) and blood (BL), white muscle (WM), red muscle (RM), bone (B), scales (SC), skin (SK), ovaries (OV), testis (TE).

Moreover, 123 samples (whole bodies) from two commercial hatcheries representing four early developmental stages were collected in RNAlater (Sigma) and kept at -20°C. From those, 30 samples were taken at the first feeding stage (FF), when the larvae first opened their mouths; 18 samples were taken at the time of flexion (FL), 20–25 dph; 35 samples were taken at the time of the end of larval rearing (ELR), 30-44 dph; and 40 samples were taken at the time of mid-metamorphosis, 50–57 dph (MM). Total RNA was extracted using the E.Z.N.A.® Total RNA Kit I from Omega, as described above, and any remaining genomic DNA was digested using the DNA-free™ DNA Removal Kit from Invitrogen. Using the High-Capacity cDNA Reverse Transcription Kit with RNase Inhibitor, 1 µg of DNAse-treated total RNA was reverse transcribed into cDNA for the cDNA synthesis (Applied Biosystems). The outcome of the cDNA synthesis was confirmed by PCR using specific primers (**Table 4.1**) for the *b-actin* gene and performing 35 cycles of 95°C for 15 s, 60°C for 20 s, and 72°C for 20 s, followed by 10 min at 72°C.

The distribution and levels of expression of the paralogue genes of the gene families *uqcrc2*, *uqcrfs1, uqcrh*, and *uqcr11* were examined in the European sea bass using real-time quantitative qRT-PCR. KAPA SYBR® FAST qPCR Master Mix (2X) Kit (Kapa Biosystems) and the proper concentration of unique set of primers for each transcript were used to produce reactions for a final volume of 20 ul (**Table 4.1**). Primer3 (v. 0.4.0) (Untergasser et al., 2012) and Beacon Designer™ software were used to design the primers, which were all made to span the transcript coding sequences. The following was the amplification cycle: 5 min at 95 °C, 40 cycles of 95 °C for 20 s and 60 °C for 20 s, followed by the dissociation curve step to confirm the amplification of a single product. Every reaction was carried in duplicate. To calculate amplification efficiency, a standard curve from pooled cDNA was built using dilutions of 1:5, 1:10, 1:20, 1:50, and 1:100.

Using the geNorm (Vandesompele et al., 2002) method, four housekeeping genes (*fau, rpl13a*, and *rpl19*) were evaluated, and the data was normalized using the geometric mean of the two most stably expressed housekeeping genes (*fau*, *rpl13a*).

Table 4.1. Set of primers and reaction efficiencies for the European seabass genes

Gene name	Gene id	FW	RV	efficiency %
uqcrc2.1	ENSDLAG00005014442	GGGAAGTGTCAGACCTCACG	CATGTGGTTGGGGCAGTACA	106.4
uqcrc2.2	ENSDLAG00005014897	AGAACTACTCCCCCGTGTCC	GCCAGTCGTAGCACATGAGA	98.9
uqcrfs1.1	ENSDLAG00005020221	TTGCCCACACAGACATTAGG	CTTCTGGATTCACTGCTCTCC	100.0
uqcrfs1.2	ENSDLAG00005021271	GGCAAAAACATGACCTTCAA	GCTCCCCCATATTCACAGAC	105.0
uqcrh.1	ENSDLAG00005003877	CTTCTGGAGCAGTGCGAGAC	CACTTACACAGTGGTCCCGAG	99.8
uqcrh.2	ENSDLAG00005009892	TACGAGCCAAGTGTGAGCAG	GTGCAGTCCTCCAAAGTGGT	103.4
uqcr11.1	ENSDLAG00005024816	GATCCCAACTTTGGCTGTATGG	TGCCACTGATGTAGGGAACGT	100.7
uqcr11.2	ENSDLAG00005014897	GGCGGCAAATACATTGCTA	ATCGCCAGTCTGTGAAGTGA	98.9
fau	ENSDLAG00005000664	CTTCGTGAATGTTGTGCC	ACTGATGGATGGTGATGA	108.0
rpl13a	ENSDLAG00005000624	GAAGGCATCAACATCTCC	CTCTGAAGTGGTAAGGTC	103.0
rpl19	ENSDLAG00005034639	ATCAGGTGCGGACAGTAT	ACACAGCGGAATCAATGG	101.0
actb	ENSDLAG00005028561	AAGCAGGAGTATGATGAGTC	GAAGTTGTTGGGCGTTTG	98.7

In order to satisfy the normative and/or homogeneous assumptions, gene expression data were square-root transformed. Using a two-sided t-test, differences in each gene's expression levels in the various tissues were analyzed. Differences were considered significant at $p<0.05$ level. All statistical analyses were performed using R packages for statistical analysis.

Results

Transcriptome analysis

The whole transcriptome sequencing yielded reads of high quality and abundance for all samples. The range of clean reads between the samples was from 29.6M to 43.7M. The accuracy of the sequencing was also high with the Q20 (%) be above 97% and Q30 (%) above 93% for all the

samples accompanied with low error rates (**Table 4.2**). In addition, the range of mapped reads between the samples was from 82.5% to 90.2% (**Table 4.3**). All the above are indicative of high sequencing quality and reliable data, suitable for gene expression analysis.

Table 4.2. Whole transcriptome data quality summary. The sample names denote the species (SA, gilthead seabream; DL, European seabass), the developmental stage (FF, First feeding; FL, flexion of the notochord; MM, mid-metamorphosis), and the rearing temperature (20°C or 17°C)

Sample name	Raw reads	Clean reads	Error rate (%)	Q20 (%)	Q30 (%)
SAFF20	30496169	29662496	0.03	97.69	93.57
SAFL20	35152628	34398391	0.03	97.77	93.77
SAMM20	36265851	35494824	0.03	97.79	93.83
DLFF17	30988451	30246645	0.03	97.79	93.69
DLFL17	33666672	32152948	0.03	97.34	92.61
DLMM17	45521814	43726235	0.03	97.83	93.89

OXPHOS gene expression in development

Transcriptome analysis in gilthead seabream revealed that all OXPHOS genes were expressed at least at one of the three developmental stages examined, with the only exception the gene *cox7a2l.1,* which was not detected at any stage. Distinct ontogenetic profiles were drawn for several genes (**Figure 4.1**), as a number of genes identified as DEGs (log2 fold change above 0.6) in the comparisons between the developmental stages (**Figure 4.1**, **Tables S4.1-S4.3**). Most of the DEGs were downregulated during the transition from FF to FL, while the opposite was observed for the transition from FL to MM. Additionally, a number of genes were expressed only at one stage, with traces or no expression in the other developmental stage(s), indicative of a stage-specific regulation of expression. Interestingly, 36 of the 52 seabream OXPHOS paralogs appeared to be DEGs, at least according to one stage-to-stage comparison, and were among the five and eight genes for gilthead seabream and European seabass respectively, that exhibited stage-specific expression patterns. Within each developmental stage, differences between the

expression levels of members of several gene families were observed, and those patterns were changing ontogenetically.

Table 4.3. Summary of transcriptome data mapping. The sample names denote the species (SA, gilthead seabream; DL, European seabass), the developmental stage (FF, first feeding; FL, flexion of the notochord; MM, mid-metamorphosis), and the rearing temperature (20°C or 17°C)

Sample name	SAFF20	SAFL20	SAMM20	DLFF17	DLFL17	DLMM17
Total reads	59324992	68796782	70989648	60493290	64305896	87452470
Total mapped reads	52664352	60961304	64058029	49959673	55817583	73381791
Uniquely mapped reads	50363173	58328096	60274717	48412541	54031293	69946459
Multiple mapped reads	2301179	2633208	3783312	1547132	1786290	3435332
Total mapping rate	88.77%	88.61%	90.24%	82.59%	86.80%	83.91%
Uniquely mapping rate	84.89%	84.78%	84.91%	80.03%	84.02%	79.98%
Multiple mapping rate	3.88%	3.83%	5.33%	2.56%	2.78%	3.93%

The ontogeny of OXPHOS genes in European seabass presented many similarities to gilthead seabream. The majority of genes were expressed at least in one stage with the only exception of the gene *cox7a2.1b* that no transcripts were traced in any stage. Overall, 37 of 50 OXPHOS paralogs appear to be DEGs (**Figure 4.1**, **Tables S4.4-S4.6**), at least according to one stage-to-stage comparison. Lower expression at FL stage compared with FF and MM was a widespread pattern.

In general, there were many similarities in the OXPHOS gene expression profiles between the two species. Within each developmental stage, genes were differentially expressed, indicative of different regulation mechanisms in place. The VENN diagrams revealed a big-scale transition in OXPHOS gene expression around FL stage, evident through the number of DEGS at the comparisons FF-FL and FL-MM (**Figure 4.1**). In gilthead seabream 34 and 82 DEGs were identified in FF-FL and FL-MM comparisons, respectively, as opposed to 20 in the FF-MM comparison. In European seabass, the differences in DEG number were even higher with 77 and 66 DEGs identified in FF-FL and FL-MM comparisons, respectively, as opposed to 25 in the FF-MM

comparison. This transition in expression around FL was the most prominent in the genes of complex I and complex III in European seabass (**Figure 4.1**).

Complex-specific differences were also observed; the genes of complex V exhibited the highest expression levels, followed by the genes of complex IV, and this observation was consistent in both species (**Figure 4.1**).

Expression profiles of OXPHOS paralogs

OXPHOS paralogs showed a variety of expression patterns between and within developmental stages (**Figure 4.2**). Three patterns were observed in the expression profiles of paralogs: 1) similar expression levels between paralogs, 2) one paralog is overexpressed compared with the other, and 3) one or more paralogs are expressed at a specific stage while the other is consistently expressed across stages.

A great example of the first pattern is the relative expression of complex III paralogs. The expression levels of the members of the gene families *uqcrh, uqcrfs1*, and *uqcr2* were nearly equal throughout the stages, and this pattern was consistent in both species. The fourth gene family, *uqcr11*, was the exception to the above; while the genes in European seabass share the same pattern with the other gene families, in gilthead seabream the *uqcr11.2* paralog was not expressed at all during FF and FL and it was only discovered at low levels at the stage of MM. The pattern of nearly balanced expression was also followed by the single gene family of complex II, *sdhb* in both species.

The expression profiles of the gene families of complex I make a good example of the second pattern observed. Complex I has three gene families, *ndufv1*, *ndufs1* and *ndufs8,* with two members each. The genes *ndufv1.2*, *ndufs1.1* and *ndufs8.1* were highly expressed across the stages while their paralogs were expressed at low levels. The paralogs of these families followed the same pattern in both species.

Finally, there were several genes that at least one paralog demonstrated stage specific expression pattern. In European seabass *cox4i.2* was moderately expressed only in MM, while the gilthead

seabream homolog was identified earlier in ontogenesis at the stage of FL and increased its levels at MM. In gilthead seabream, the gene *cox6a.2* was not expressed at all at the FF and FL stages and was moderately expressed at MM, whereas in European seabass the homolog was expressed in all three developmental stages but at very low levels. The *cox6b* gene family has only two members in gilthead seabream, which were expressed in all stages, whereas in European seabass one more paralog, the *cox6b.3*, was identified that was expressed only at the later stages of FL and MM.

Validation of expression patterns of OXPHOS genes

In order to verify the differential expression of paralogs between and within developmental stages, a new collection of European seabass larvae was created from commercial hatcheries to account for a wider genetic basis and rearing practices (Kourkouta et al., 2022) In order to verify the transition in expression around the FL stage, an additional developmental stage was sampled, that of end of larval rearing (ELR), when all fins have been developed. The paralogs of complex III were used for validation purposes given that complex III genes were significantly downregulated at FL in European seabass (**Figure 4.1**).

The expression of all gene families of complex III with paralogs in European seabass was confirmed in commercial larvae samples with real-time RT-PCR. Transcript abundance differentiated with gene family and developmental stage (**Figure 4.3**, **Figure 4.4**, **Figure 4.5**, **Figure 4.6**). Axis Y was plotted at the same scale across Figures 4.3-4.6 to allow direct comparison of expression levels of all paralogs measured. A systematic pattern was observed across all analysed genes; they were all expressed at FF and their expression levels dropped as development advanced at stages FL and ELR, before rising again at MM stage. In general, ELR was the stage with the lowest expression and significantly lower than in MM, with the exception of *uqcrc2* paralogs (**Figures 4.3-4.6**).

Tissue distribution of paralogue gene expression

Real-time RT-PCR analysis of relative gene expression revealed that complex III paralogs were expressed in all tissues studied, yet their transcript abundance was different (**Figure 4.7**, **Figure 4.8**, **Figure 4.9**, **Figure 4.10**). The calcified tissues (bone and scales), skin, spleen, and eye exhibited the lowest levels of gene expression for all paralogs, whereas the brain, red muscle, and gonads exhibited the highest.

The expression of the two *uqcrh* paralogs followed a consistent pattern, with the *uqcrh.1* paralog expressed at higher levels than *uqcrh.2* paralog in all tissues and that difference was statistically significant ($p < 0.05$) in the pituitary, gills, spleen, pyloric caeca, anterior and posterior intestines, liver, bone, and scales (**Figure 4.7**). When compared to *uqcr11.2*, *uqcr11.1* was typically the most abundant transcript, and this difference was statistically significant ($p < 0.05$) in the gills, heart, blood, spleen, kidney, stomach, pyloric caeca, posterior intestine, adipose tissue, bone, scales, and skin (**Figure 4.8**). The paralog *uqcrc2.2* was overexpressed in comparison with *uqcrc2.1* in most tissues of the European seabass, with the exception of the pyloric caeca, where *uqcrc2.1* was the prevalent isoform (**Figure 4.9**). Pituitary gland, spleen, pyloric caeca, adipose tissue, bone, and scales all showed statistically significant changes ($p < 0.05$). In most tissues, *uqcrfs1.1* was the highly expressed paralog, with the exception of the brain, heart, spleen, stomach, and red muscle, where *uqcrfs1.2* is more highly expressed. The expression of *uqcrfs1* paralogs was statistically significant different ($p < 0.05$) in the pituitary, heart, spleen, pyloric caeca, posterior intestine, scales, skin, and ovary (**Figure 4.10**).

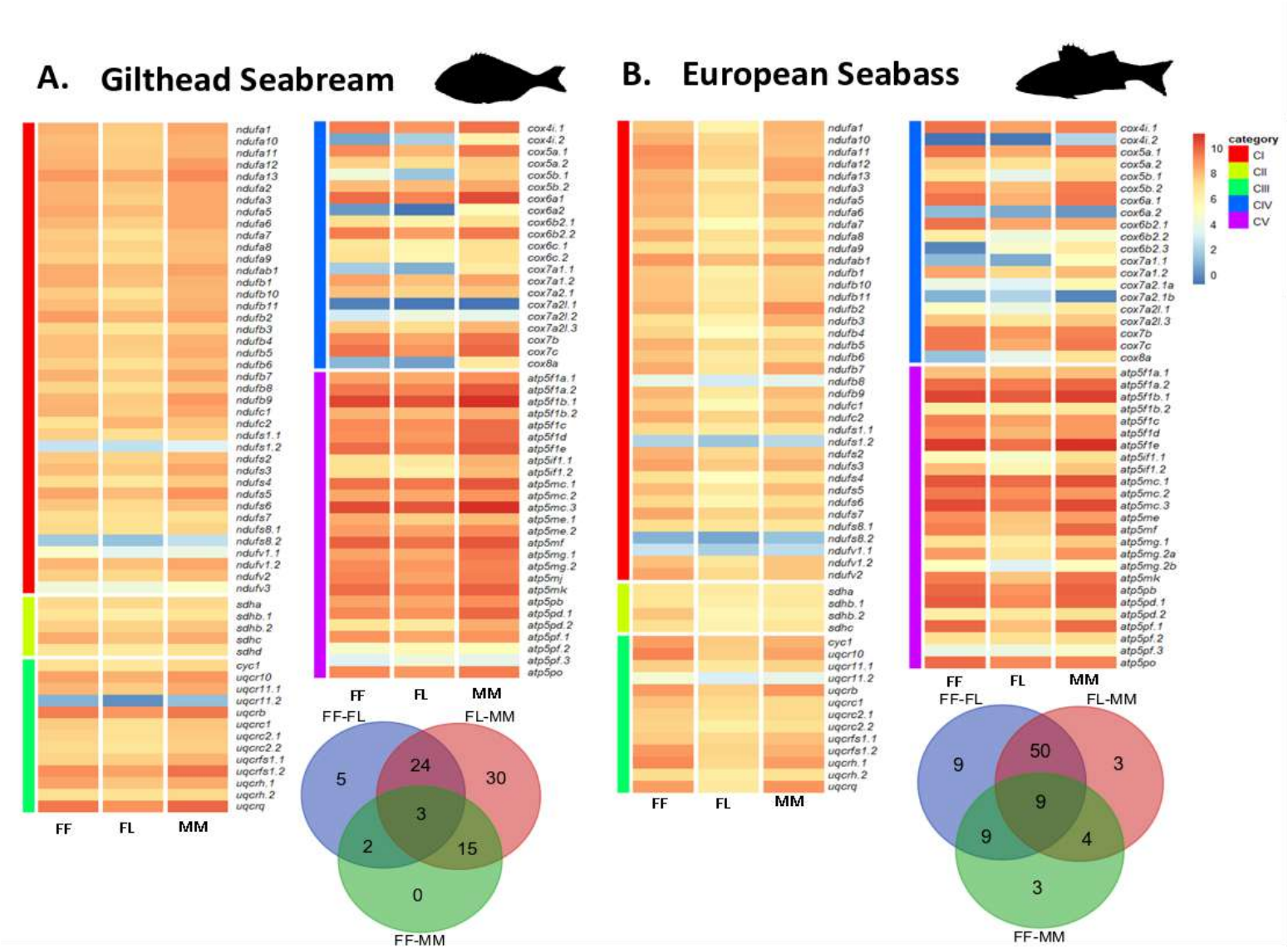


Figure 4.1. Heat map of gene expression of all OXPHOS genes and Venn diagrams for DEGs between the developmental stages for A) gilthead seabream and B) European seabass. Color bars on the left denote the OXPHOS complex the gene contributes to. FF, First Feeding; FL, Flexion of notochord; MM, Mid metamorphosis.

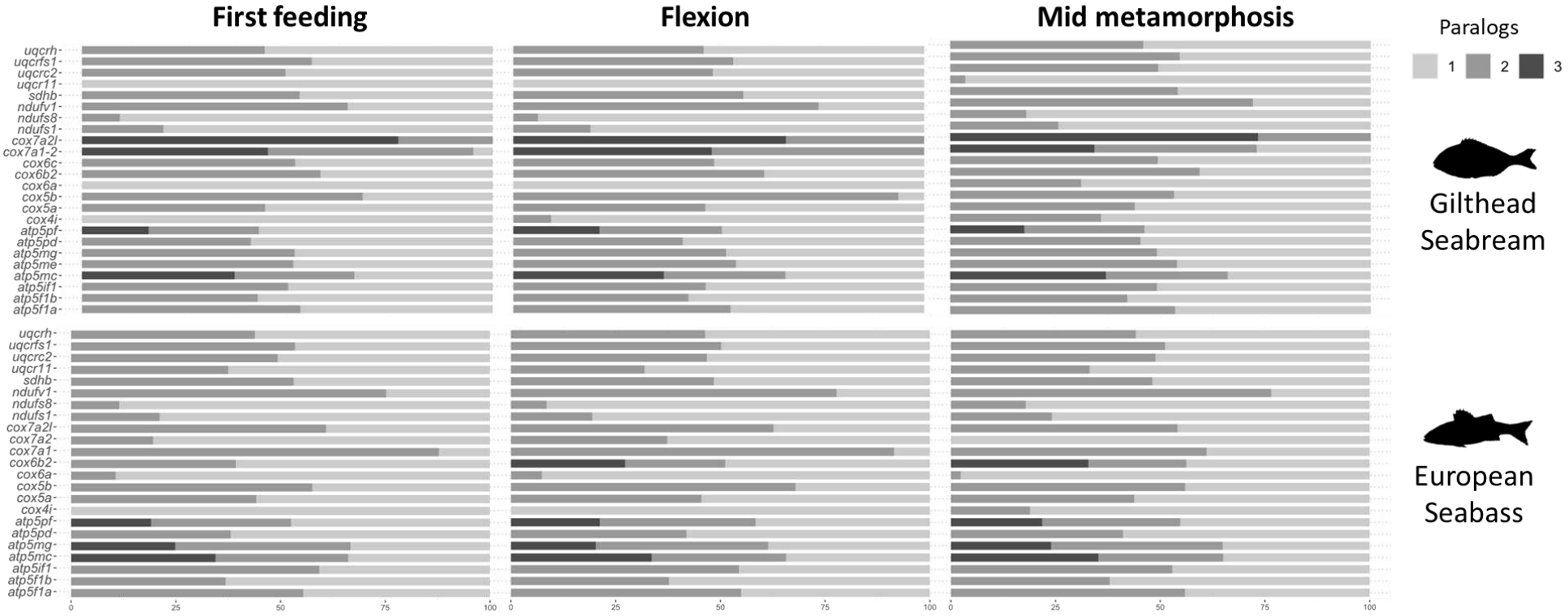


Figure 4.2. Relative expression levels of each paralog in each gene family expressed as percentage by developmental stage for gilthead seabream (upper panel) and European seabass (lower panel)

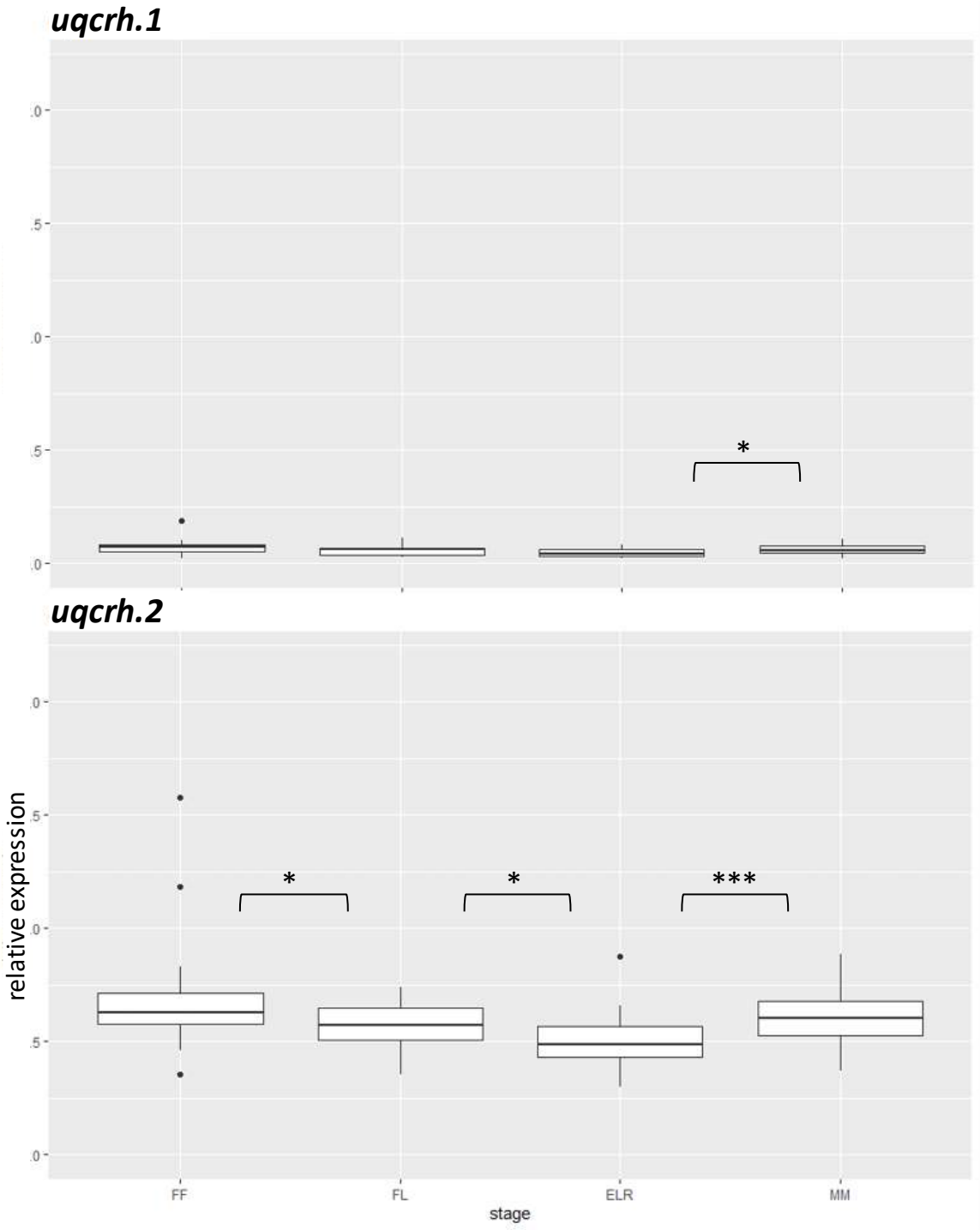


Figure 4.3. Expression levels of *uqcrh* paralogs in different developmental stages in European seabass. FF, First Feeding; FL, Flexion of notochord; ELR, end of larval rearing; MM, Mid-metamorphosis. Asterisks denote statistically significant differences between stages, *p < 0.05, **p < 0.01 and ***p < 0.001.

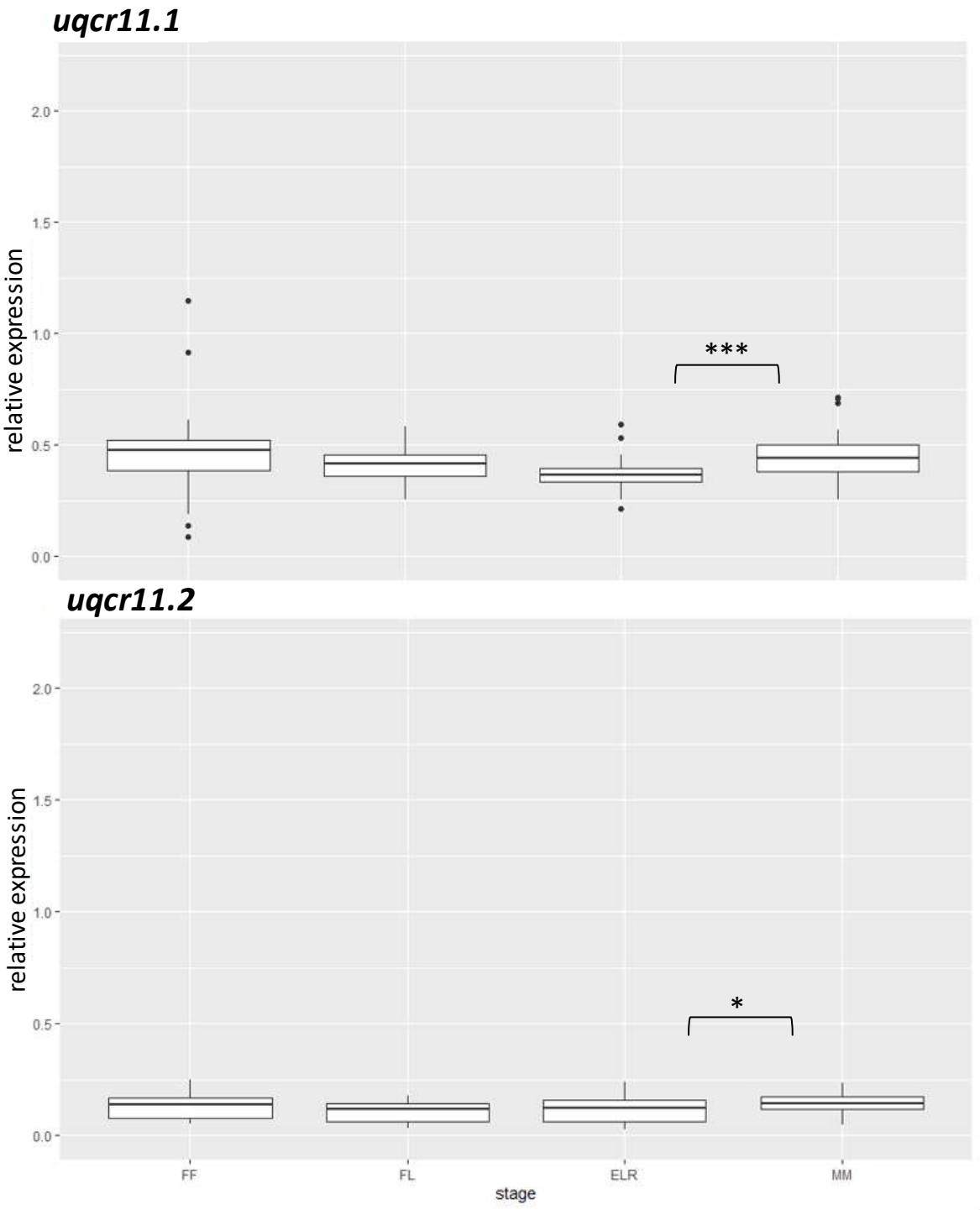


Figure 4.4. Expression levels of *uqcr11* paralogs in different developmental stages in European seabass. FF, First Feeding; FL, Flexion of notochord; ELR, end of larval rearing; MM, Mid-metamorphosis. Asterisks denote statistically significant differences between stages, *p < 0.05, **p < 0.01 and ***p < 0.001.

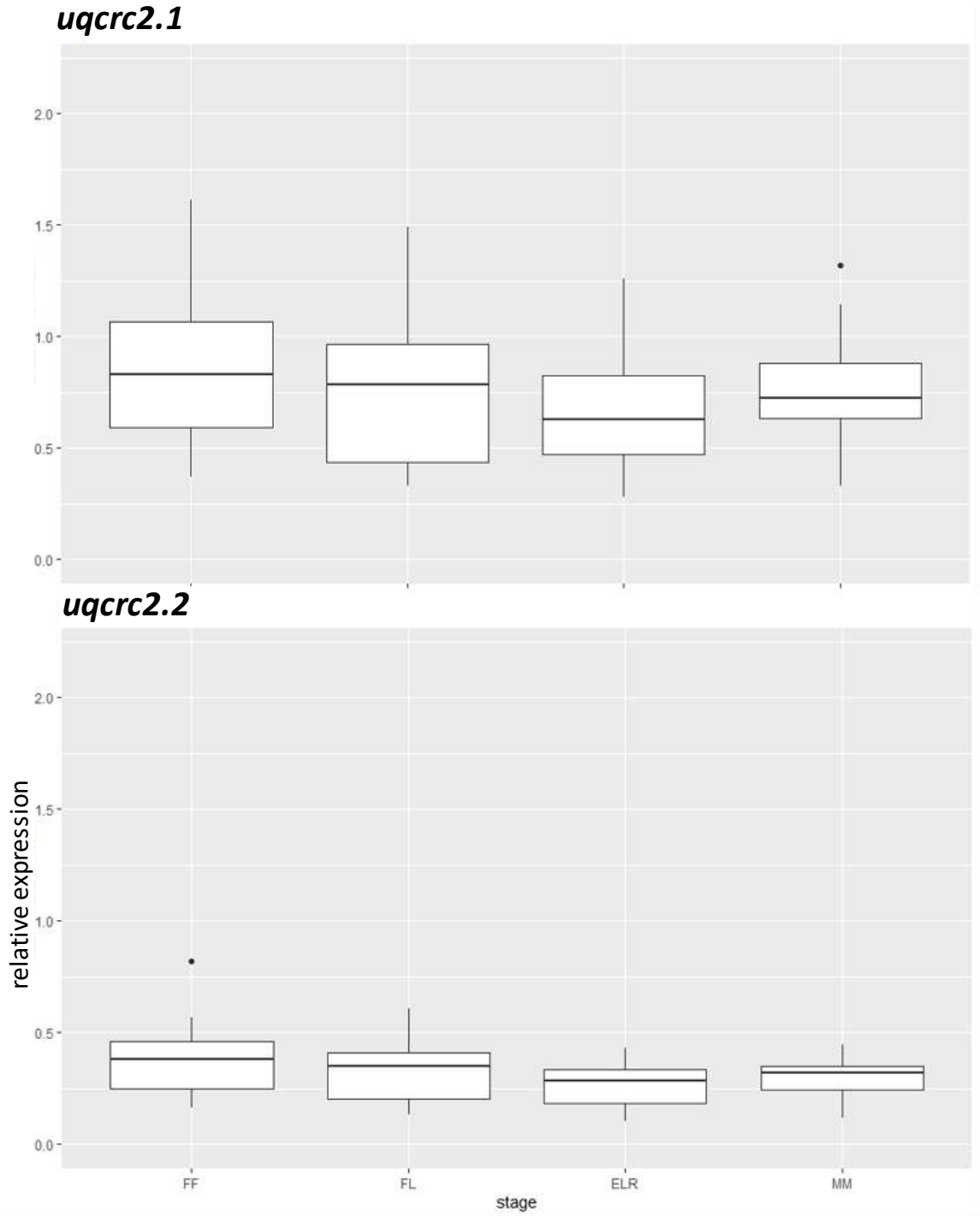


Figure 4.5. Expression levels of *uqcrc2* paralogs in different developmental stages in European seabass. FF, First Feeding; FL, Flexion of notochord; ELR, end of larval rearing; MM, Mid-metamorphosis. Asterisks denote statistically significant differences between stages, *p < 0.05, **p < 0.01 and ***p < 0.001.

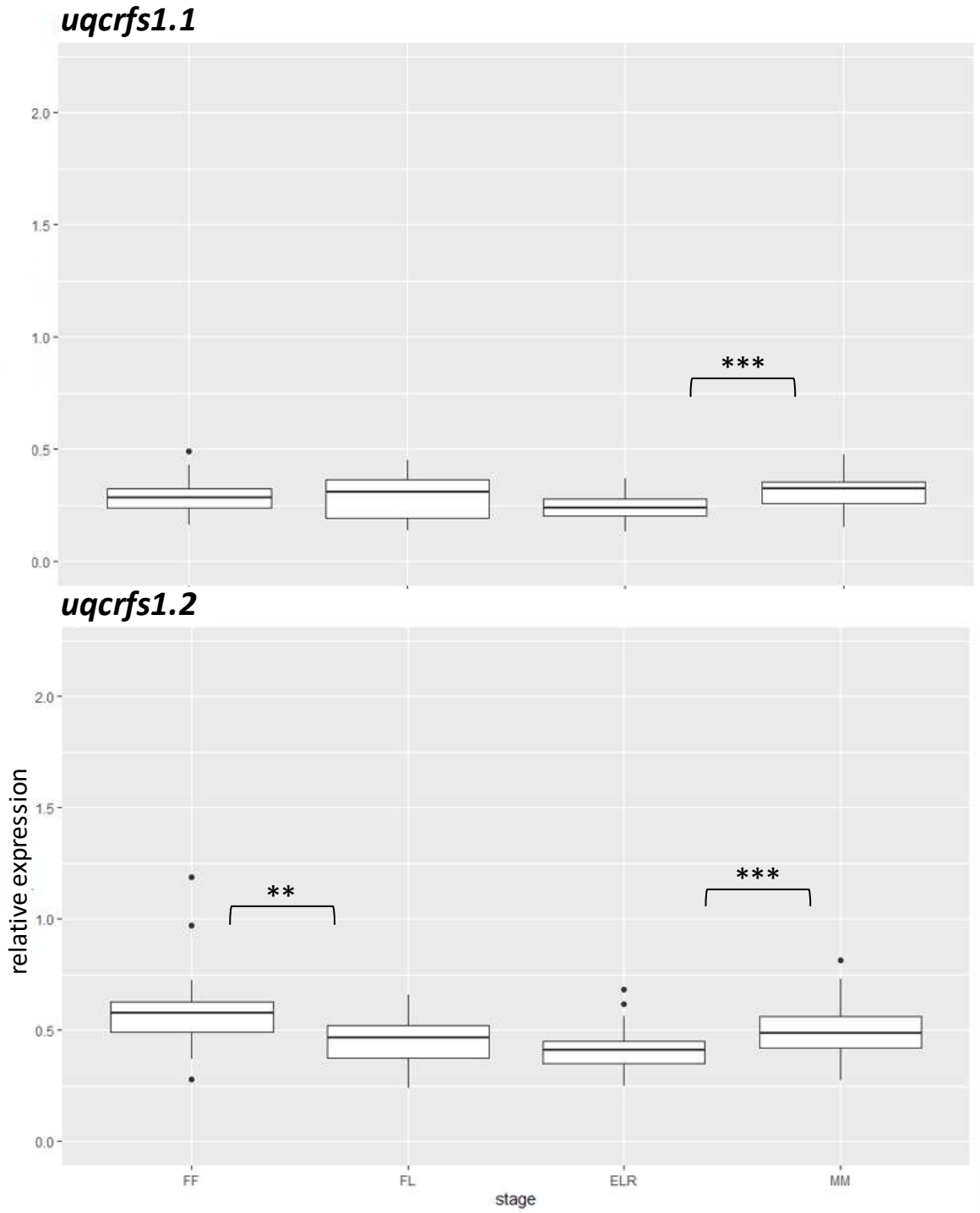


Figure 4.6. Expression levels of *uqcrfs1* paralogs in different developmental stages in European seabass. FF, First Feeding; FL, Flexion of notochord; ELR, end of larval rearing; MM, Mid-metamorphosis. Asterisks denote statistically significant differences between stages, *p < 0.05, **p < 0.01 and ***p < 0.001.

Expression of *uqcrh* paralogs

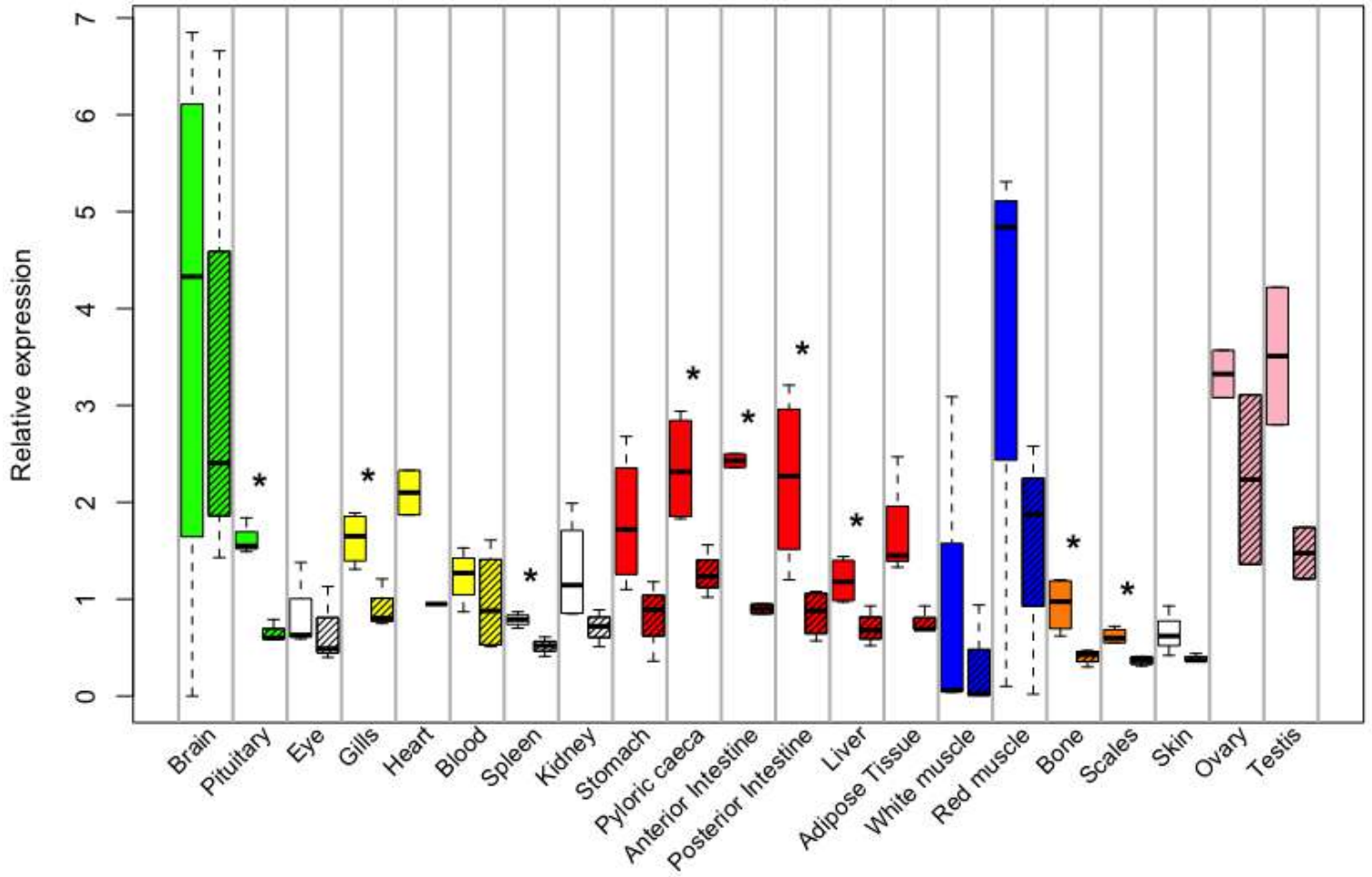


Figure 4.7. Expression levels of *uqcrh.1* (full) and *uqcrh.2* (shaded) paralogs in different tissues in European seabass. Asterisks denote statistically significant differences, *p < 0.05, **p < 0.01 and ***p < 0.001.

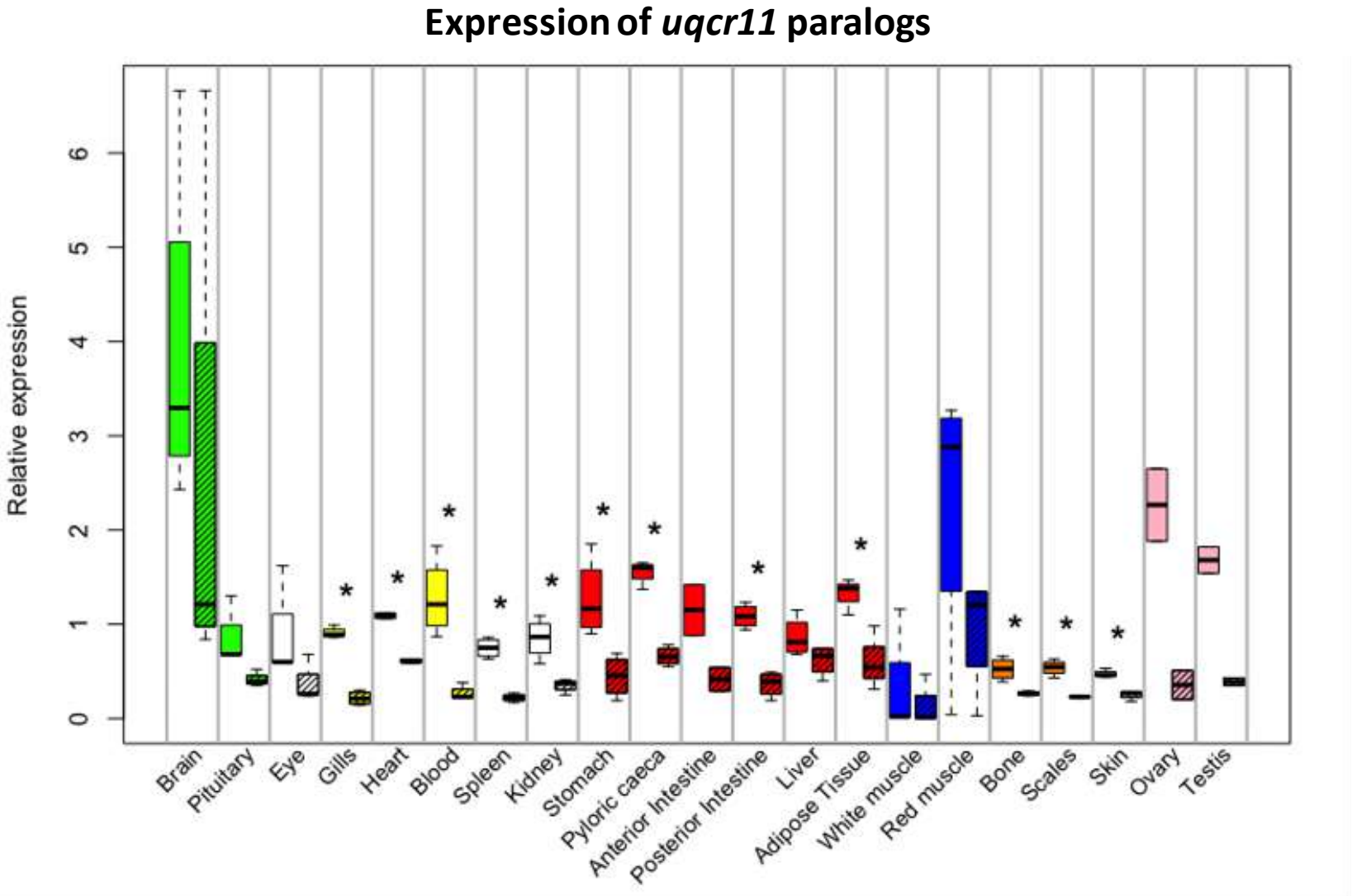


Figure 4.8. Expression levels of *uqcr11.1* (full) and *uqcr11.2* (shaded) paralogs in different tissues in European seabass. Asterisks denote statistically significant differences, *p < 0.05, **p < 0.01 and ***p < 0.001.

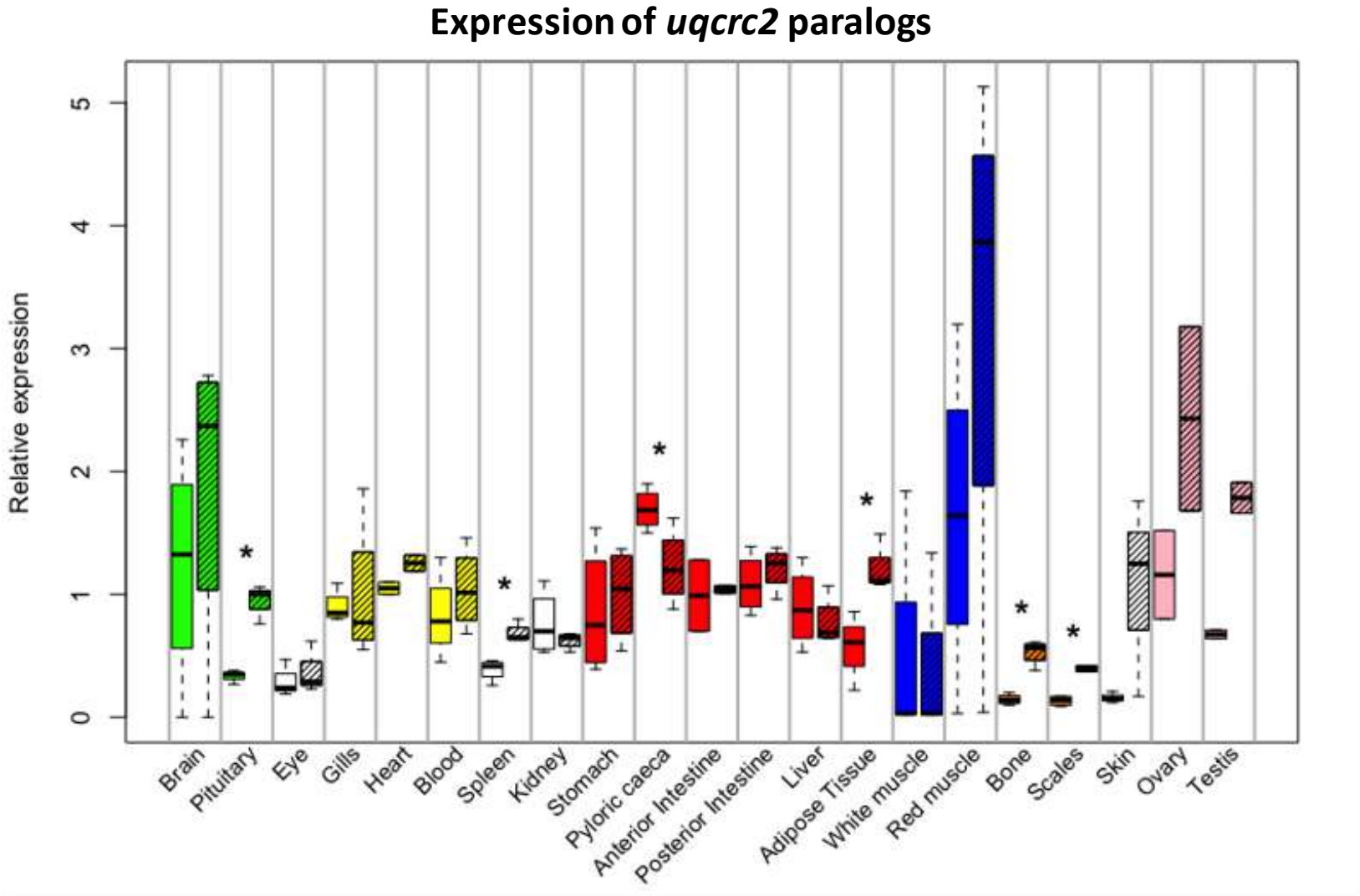


Figure 4.9. Expression levels of *uqcrc2.1* (full) and *uqcrc2.2* (shaded) paralogs in different tissues in European seabass. Asterisks denote statistically significant differences, *p < 0.05, **p < 0.01 and ***p < 0.001.

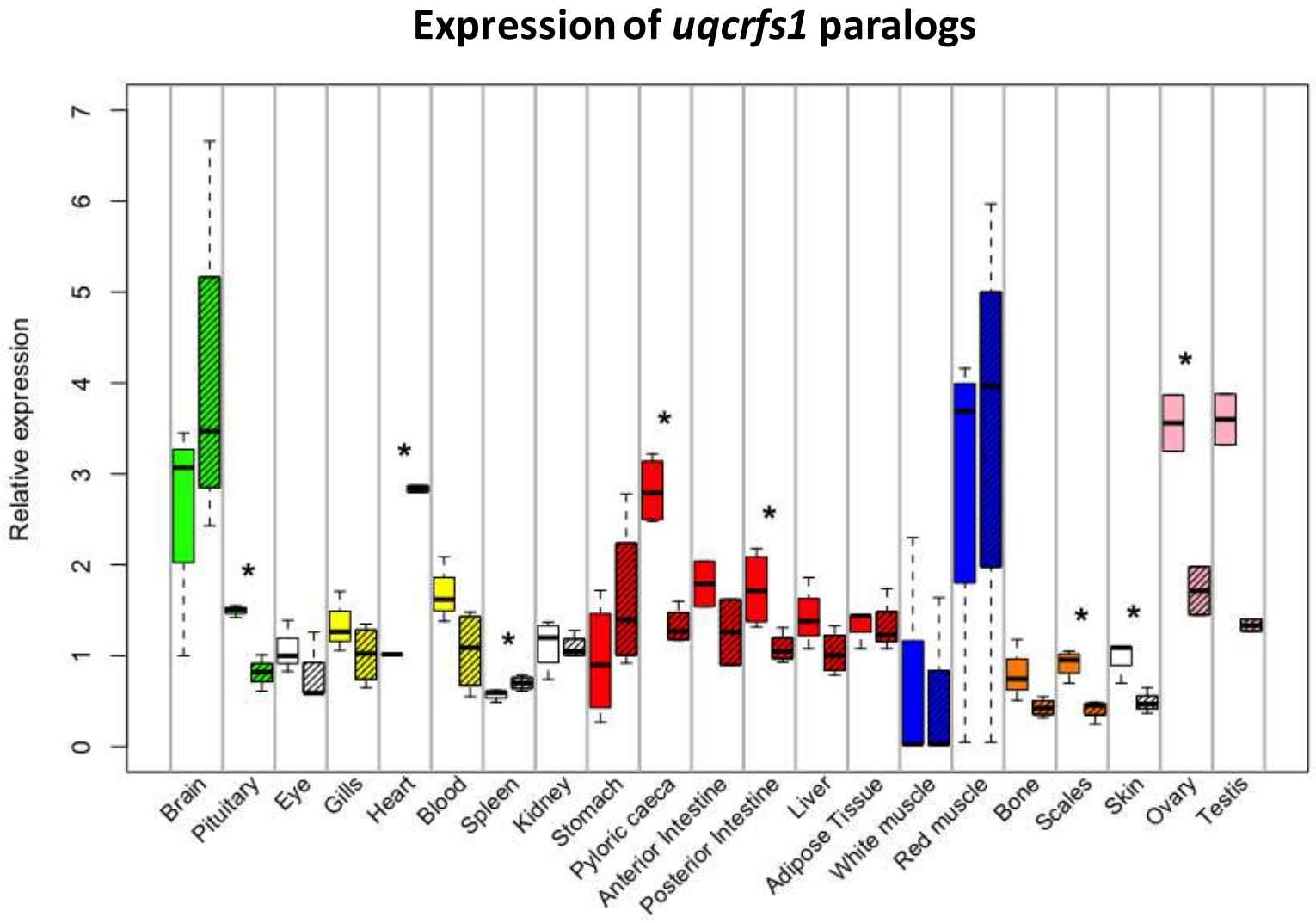


Figure 4.10. Expression levels of *uqcrfs1.1* (full) and *uqcrfs1.2* (shaded) paralogs in different tissues in European seabass. Asterisks denote statistically significant differences, *p < 0.05, **p < 0.01 and ***p < 0.001.

Discussion

The transcriptome analysis in both species provided evidence that the predicted genes identified in chapter one as OXPHOS paralogs are functional genes at least at the transcriptional level. In each species almost all genes were identified at least in one stage. The sole exceptions were the genes *cox7a2.1b* and *cox7a2l.1* for European seabass and gilthead seabream, respectively, which were represented by less than 20 reads and were therefore regarded as not expressed. However, the evidence produced here is not considered adequate to render those paralogs as non-functional, given that the transcriptome analysis covered only a fraction of the species life cycle. Thus, the appearance of those paralogs at later stages cannot be excluded.

OXPHOS complexes gene expression profile

Comparing the expression levels between the five complexes for all developmental stages different expression profiles are observed. A striking observation is that the genes encoding for the subunits of complex V exhibited the highest transcription rates, significantly higher than the other complexes. The second highest expression rate was observed for the genes of complex IV. It is known that the OXPHOS complexes are not present at equal rates but closer to 1.1/CI:1.3/CII:3/CIII:6.7/CIV:3.5/CV, with complexes IV and V present at the highest ratios as shown by (Hatefi, 1985; Schägger & Pfeiffer, 2001) in studies in bovine heart. Although transcriptome analysis did not provide data on protein stoichiometry, a trend of up regulation, mostly for complex V followed by complex IV, was observed in comparison with the other complexes across the developmental stages, and this observation is systematic in both species. While further investigation at the protein level is needed to answer if the observed pattern of gene regulation is reflecting the actual protein stoichiometry, the findings in both species support an unambiguous difference at the expression levels of the genes coding between the complexes of the respiratory chain (complex I-IV) and the ATP-synthase (complex V).

When examining the expression pattern of the members of each gene family between and within developmental stages (**Figure 4.2**), three patterns of relative expression were

observed: 1) similar expression levels between paralogs, 2) one paralog is overexpressed compared with the other, and 3) one or more paralogs are expressed at a specific stage while the other is consistently expressed across stages.

The observation of different patterns in the relative expression of the paralogs is indicative of different mechanisms shaping the fate of paralogs. Given that all paralogs encode for proteins that participate in big protein complexes, the need to maintain dosage balance and avoid imbalanced protein stoichiometry is expected to have driven evolutionary decisions right after the WGD event that created them (Almeida-Silva & Van De Peer, 2023; Ishikawa et al., 2020; Picolo et al., 2021).The first pattern that conforms with the dosage balance hypothesis, was the most often observed and the genes following this trend exhibited moderate to high expression across all stages. It is possible that these genes are absolutely essential for the overall stability, robustness and function of the complex that their copies have been retained right after the WGD event based on the dosage effect restrains, and dosage balance has generated a time-dependent selective barrier to subfunctionalization or to neofunctionalization (A. E. Wilson & Liberles, 2023).

Paralogs with differentiated relative expression (second pattern) may have been shaped through different pathways starting with the need to maintain dosage balance right after the WGD event. However, subsequent losses of genes or function of the proteins they interact with in their complex might have created the need for a new dosage status (Birchler & Veitia, 2021; Edger & Pires, 2009). The specific needs of the organism might have also added variation in expression patterns. A number of those genes are amongst the catalytic subunits of their corresponding complexes. A great example is the paralogs of complex I, where all three gene families follow this trend in both species. The recruitment of subunits encoded by different paralogs may have an impact on the efficiency of energy production or that the availability of that extra material may provide an advantage in rapid changes of the energy needs. The capacity for swift adaptation to new energy demand statuses provides an immense advantage to developing organisms as their organs differentiate and the processes of respiration, feeding and digestion, nutrient absorption and growth mature.

This hypothesis is reinforced by the observation that certain paralogs were transcribed only at particular developmental stage. There are numerus examples including myosins, growth factors, opsins and a lot more that stage specific isoforms are known to exist and gene families

with different members for embryonic, larvae, juvenile and adult forms (Davies et al., 2011; Georgiou et al., 2016; Gu et al., 2021). It seems that gene paralogs are a crucial feature of fish development and play a significant part in supplying the genetic variety and redundancy needed for the complicated developmental processes involved in the formation of varied fish body plans and adaptations.

While all those expression patterns are shared between OXPHOS paralogs, they are not uniformly divided between the five complexes. The paralogs with similar expression levels were mostly members of the gene families of complexes II, III, and V, whereas all the gene families of complex I followed the second pattern with one paralog identified as low expressed, and the majority of complex IV paralogs were following the stage-specific pattern. This observation adds to the evidence that different selective pressures may act on the different OXPHOS complexes driving differentiating adaptations.

The possibility OXPHOS paralogs are a tool to finetune energy production efficiency is also supported by the tissue-specific expression pattern of the paralogs. Using whole larvae samples does not allow any insight into tissue specific patterns. In whole-body samples where white muscle is overrepresented, the transcripts of a gene expressed specifically in another tissue would represent only a small fragment compared with the transcripts of white muscle, and they would be identified as low expressed in the whole-body sample. Studying the expression of complex III paralogs across 22 different tissues in European seabass made clear that tissue-specific expression is the rule for OXPHOS paralogs.

OXPHOS ontogenetic gene expression profile

The differential expression analysis revealed a number of OXPHOS genes that were identified as DEGs between the stage-to-stage comparisons. In both species the comparisons that exhibited the highest number of DEGs were the FF-FL and FL-MM, with the majority of OXPHOS genes identified as DEGs, while the comparison FF-MM had significantly less DEGs again in both species. The fact that numerus genes are identified as DEGs between the developmental stages but a lot of them with low log2 fold change (0.6 - 1), could be an indication of micro adjustments in OXPHOS regulation through the course of development. When examining the direction of regulation of the DEGs between the comparisons an

interesting pattern emerged; at the FF- FL comparison, 29 out of 34 DEGs in gilthead seabream, and 75 out of 77 in European seabass were down regulated, while in the FL- MM comparison, 71 out of 82 and 65 out of 66 DEGs, respectively, were up regulated.

This finding was confirmed through a different approach during the ontogeny of commercially- reared larvae of European seabass, where the expression patterns of the paralogs of complex III were explored using real-time RT-PCR in four different developmental stages. Between FF and FL, there was a tendency of decreasing expression levels that was followed by an increasing trend at the later stages between ELR and MM. It should be made clear that the pattern described above refers to four distinct stages of larval development and that it is possible that the change in gene expression levels from one stage to the next is not linear.

A possible explanation for that distinct and systematic expression pattern could lie with the specific needs and adaptations of fish larvae during development and the increments in fish white and red muscle development. In most fish species, skeletal muscle accounts for 40% to 60% of total body mass and it is the biggest growing tissue. Muscle fibers are classified as superficial (red), intermediate (pink), and deep (white). In European sea bass and gilthead seabream pink fibers appear first, relatively late in development at 80 days after hatching, while the myotome presumptive white and red fibers are present at hatching (Scapolo et al., 1988; Veggetti et al., 1990). The formation of skeletal muscle in teleost fish involves the formation of new fibers (hyperplasia) and subsequent enlargement of existing muscle fibers (hypertrophy). In contrast to mammalian muscle development, where hyperplasia is restricted to early developmental stages, in teleost species hyperplasia along with hypertrophic growth continue well after the post larva period. Early larvae are dependent on cutaneous respiration, and early embryonic and larval growth and development are primarily supported by aerobic metabolism (Finn & Fyhn, 2010; Pelster & Burggren, 1996,; Rombough, 1988). At the larval stages, both white and red fibers exhibit a high volume of mitochondria; in the case of Atlantic herring, the mitochondrial volume and density can reach up to 26% and 46%, respectively (Vieira & Johnston, 1992). At hatching, the white muscle has multiple layers of muscle fibers, whereas the red muscle consists of a single layer of fibers. The red fiber monolayer is initially found close to the lateral line, but by 10 dph, it has already extended both hypo- and epi-axially (Veggetti et al., 1990). Both red and white muscle fibers bind

antibodies for both fast and slow myosins during the time between hatching and 10 dph (Scapolo et al., 1988). At 28 dph, the presumptive red muscle layer extends further dorsoventrally, the original monolayer has more mitochondria and the regions closer to the lateral line begin to thicken by the addition of new layers of red fibers (Veggetti et al., 1990). The white fibers appear to bind only the fast myosin antibody, while the red fibers continue to react with both antibodies (fast, slow) during the same time period. There is a shift in the myosin isoforms for both white and red fibers (Scapolo et al., 1988). Diverse cyprinid species were found to exhibit similar patterns during the development of swimming muscles (El-Fiky et al., 1987). Between 30 and 40 dph, EI-Fiky and colleagues observed a switch from early (red layer) red muscle myosins to late (adult) isoforms, as well as a change in the proportion of red muscle to total muscle mass, which dropped from 12% at 3dph to 4% at 40dph. Evidently, the growth rates of white and red muscle are not equal, and the distinct processes involved in their development and growth, are in response to the specific requirements of fish larvae with regard to metabolic and developmental changes, as well as the challenge of new tasks such as swimming, exogenous feeding, and predator evasion. It is proposed that the expression pattern observed in OXPHOS complex III genes reflects changes in mitochondrial abundance, in relation to changes in the formation and expansion of red muscle fibers, as well as their contribution to the total mass of skeletal muscle.

This first known systematic recording of OXPHOS regulation in fish ontogeny made clear that OXPHOS is not a static process in early fish life, with a number of different patterns of OXPHOS regulation observed. The existence of OXPHOS paralogs seems to be adding to the complexity of OXPHOS regulation as a whole. The fact that the majority of the observed OXPHOS paralogs both in gilthead seabream and European seabass are the result of the TSGD (Chapter 2) is pointing towards the direction that the possible adaptations and evolutionary pressures, that play a role to their retention in fish genomes has to do with the challenges of aquatic life. In Chapter 2, it was proposed that the most probable evolutionary scenario for the retention of paralogs is regulatory neo/subfunctionalization. The findings of this chapter, i.e. a lot of paralogs were identified as DEGs between the developmental stages, the ratio between the paralogs in some cases was changing ontogenetically and some of them exhibited stage- and tissue- specific patterns, offer support to this scenario.

Chapter 5

OXPHOS paralogs are differentially regulated

Introduction

OXPHOS, acknowledged for its conserved role in cellular energetics and overall organismal fitness, reveals an additional level of complexity; while Chapter 2 contributed to the understanding of the complexity of OXPHOS at the genomic level, this was further underscored in Chapter 4, where evidence emerged of the differential expression of gene family members across diverse tissues and ontogenetic time points. Furthermore, chapter 4 pointed to the numerous studies revealing the distinct expression patterns between paralogs in response to various stimuli, such as osmoregulation, dietary changes, host-microbe interactions, infection, hypoxia, and temperature acclimation. Temperature is a key factor in aquatic environments, that significantly shapes fish phenotypes during early ontogeny, impacting both wild and reared stocks (Geladakis et al., 2022; Georgakopoulou et al., 2007; Reyes Corral & Aguirre, 2019). The growing concerns about global warming and fluctuations in seasonal temperatures have heightened the focus of scientific community on the thermally induced phenotypic plasticity of poikilothermic organisms, particularly fishes (Debes et al., 2021; Mitra et al., 2023; Vagner et al., 2019). The developmental temperature emerges as a crucial determinant of fish phenotypic plasticity influencing a variety of plastic traits such as, sex (Pavlidis et al., 2000), meristic characters (e.g., number of fin-rays or vertebrae) (Christou et al., 2018), ontogenetic scaling (Gibson & Johnston, 1995; Johnston, 2006), body shape (Georga & Koumoundouros, 2010; Turan, 2004), enzyme activity (Schnurr et al., 2013), muscle structure (Albokhadaim et al., 2007; Johnston, 2006; Johnston et al., 2009), cardiac morphology (Dimitriadi et al., 2018), lifespan (Lein et al., 1997), stress and immune response (Mateus et al., 2017), hypoxia tolerance (Zambonino-Infante et al., 2013) and thermal acclimation (Scott & Johnston, 2012). Research interest is concentrated on the embryonic and/or larval period up to metamorphosis, given the increased sensitivity of these stages to water temperature and the substantial phenotypic responses that they ensue. A collaborating research team at the University of Crete, Greece, conducted comparative research on the impact of developmental temperature (17, 20, or 23°C) during the brief embryonic and yolk-sac larval period on the later-stage swimming performance of metamorphosing gilthead seabream (Kourkouta et al., 2021). The study yielded several noteworthy findings. First, a significant decrease in the critical swimming speed was observed in metamorphosing larvae reared at 17°C. This decrease in swimming performance was found to be significantly

associated with fish body shape, with the 17°C group exhibiting a slender body shape, longer caudal peduncle, terminal mouth, and ventrally transposed pectoral fins compared to other groups. Second, an increase in the relative depth of the heart ventricle was recorded in fish reared at 17°C, and third, the incidence of caudal-fin abnormalities showed a significant decrease with the rise in developmental temperature. To delve deeper into the regulation of OXPHOS paralogs in fish, in this chapter the expression profiles of OXPXHOS complex I and III paralogs were studied in the fish derived from that same experiment run by the collaborating research team in Crete. The expression of complex I and complex III paralogs was compared in control and exercised gilthead seabream larvae reared at three different temperatures during the autotrophic stage. The paralog expression patterns are viewed in association with their swimming performance, a clearly high energy demanding process.

Materials and Methods

Experimental design and maintenance of fish populations

Nine groups of gilthead seabream were tested at three developmental temperatures (DT) 17, 20, or 23 °C in triplicate from the time when epiboly initially appeared until the end of the

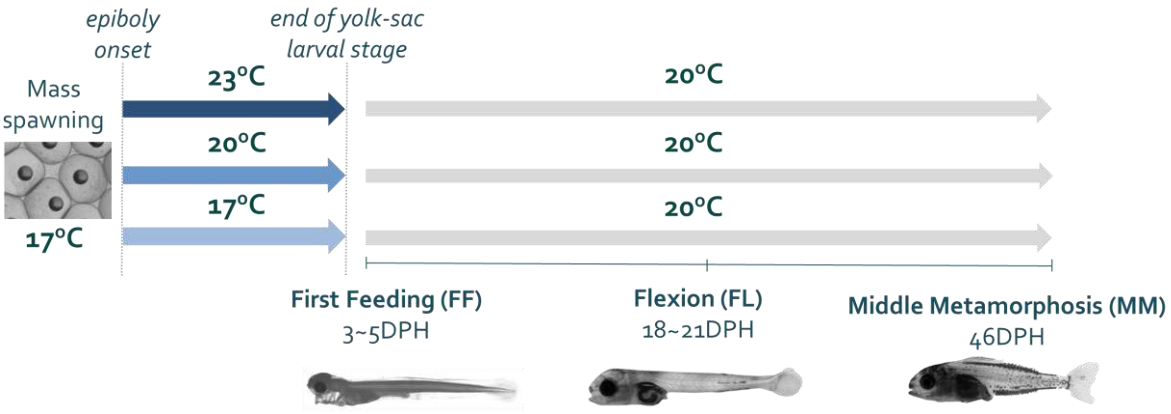


Figure 5.1. The rearing protocol to test the effects of temperature during gilthead sea bream egg rearing up to first feeding. The arrows indicate the thermal regimes applied at the start and end of the experiment and the gradual temperature change to the common regime at the end of the experiment (20 ºC).

yolk-sac larval stage and the start of exogenous feeding. Then, up until the end of the study, all groups were maintained under the same rearing conditions and temperature (20 °C) (**Figure 5.1**). The eggs were acclimated at rates of 0.5 and 0.2 degrees Celsius per hour from the spawning temperature (17 °C) to the test temperatures, and subsequently to the typical temperature for rearing larvae (20 °C). The same mass spawn of captive breeders produced all of the eggs.

The "pseudo-green water" technique was used for egg hatching and larval rearing (Papandroulakis et al., 2001). There were three water recirculation systems used, each with three 500 L tanks linked to biological filtration. Using chillers and heaters that were automatically managed by electronic sensors, water temperature was brought to the test levels. Until they reached 6 mm in total length (TL), larvae were fed enriched rotifers (DHA Protein Selco, INVE), followed by enriched rotifers and *Artemia* nauplii (Easy DHA Selco, INVE), and finally enriched *Artemia* nauplii after 10 mm TL.

Rotifer concentration in the tanks was raised to 6 to 8 indi ml^{-1} twice daily. *Artemia* nauplii concentration was adjusted twice daily to 0.5-2 ind ml^{-1}. Live *Chlorella minutissima* microalgae were added three times per day at a concentration of 6.5±3×105 cells mg/L^{-1} during the rotifer-feeding period. Fish were weaned to inert diets (Proton, Alfa, INVE) when they were about 8 mm TL. Rotifer overgrowth was removed from the culture medium during the rotifer-feeding period using planktonic traps (50 m mesh size), which were placed at the water-outlet of tanks. Between 7 and 25 dpf, air-blowing skimmers were placed to keep the water's surface clear of lipids and improve swim bladder inflation. Water oxygen saturation ranged from 5.8 to 6.0, pH was 7.8 to 0.3, salinity was 35 to 36, and photoperiod was 18L:6D. Larval rearing was performed at the Institute of Marine Biology, Biotechnology and Aquaculture (HCMR).

Swimming challenge

Swimming challenge was performed according to G. Koumoundouros et al., 2009a (Koumoundouros et al., 2009) using equipment designed for the purpose with a swimming channel that measured 70 cm in length, 10 cm in depth, and 5 cm in breadth. Through the use of external pumps and valves, the flow regime was altered. To measure the water speed in

the tunnel, a Valeport model 801 electromagnetic flow-meter was employed. In order to keep the laminar flow in the swimming channel and stop fish from escaping forward, a screen made of plastic straws was used. The water was kept at a constant 20 °C, 100% oxygen saturation, and 35 ‰ salinity.

Ten fish from each thermal experimental population were moved to a holding aquarium and starved for eighteen to twenty hours before the testing. Fish were placed in the swimming channel for 10 minutes at a water speed of 2 TL s^{-1} for the swimming test. The following time frame saw an increase in water velocity of 1 TL s^{-1} every 10 minutes. Fish were removed from the swimming channel when they became too tired to respond to visual or auditory cues coming from the side or back. Fish that were exhausted were given anesthetic (MS222, 100 mg/L-1), and samples were taken.

Gene expression analysis

A total of 68 samples of the caudal region from gilthead seabream juveniles were collected and stored in methanol at -20 °C until RNA extraction. From those samples 27 belonged to the 17°C group (6 control and 21 exercised), 20 to the 20°C group (4 control and 16 exercised) and 21 to the 23°C group (5 control and 16 exercised). Total RNA was extracted using E.Z.N.A.® Total RNA Kit I (Omega) and was DNAse treated with the DNA-free™ DNA Removal Kit (Invitrogen) to remove traces of genomic DNA. For the cDNA synthesis, 200 ng of DNAse treated total RNA was reverse transcribed to cDNA using the High-Capacity cDNA Reverse Transcription Kit with RNase Inhibitor (Applied Biosystems). cDNA was stored at -20°C until further use.

Real-time RT-PCR was performed in order to determine the expression levels of the 15 genes (**Table 5.1**). Twelve target genes are coding for subunits of the complexes I and III of the oxidative phosphorylation. Reactions were prepared for a 20 ul final volume using KAPA SYBR® FAST qPCR Master Mix (2X) Kit (Kapa Biosystems) and the appropriate concentration of each transcript specific set of primers. Primers were designed using the Primer3 (v.0.4.0) (Untergasser et al., 2012) and Beacon Designer ™ software; all primers were designed to span the coding sequences of the transcripts. Amplification cycle was: 5 min at 95 ºC, followed by 40 cycles of 95 ºC for 20 s and 60 ºC for 20 s, followed by the dissociation curve step to verify

for a single product amplification. Each reaction was performed in duplicate. A standard curve using a dilution series (1:5, 1:10, 1:20, 1:40, 1:100, 1:200) from pooled cDNA was constructed to estimate for amplification efficiency (**Table 5.1**). Three housekeeping genes (*ef1, rps18* and *rpl13*) were tested and assessed using the geNorm (Vandesompele et al., 2002) algorithm and the data was normalized using the geometric mean of the two most stably expressed housekeeping genes (*rps18* and *rpl13*).

Table 5.1. Set of primers and reaction efficiencies for OXPHOS paralogs in gilthead seabream

Gene name	Gene id	FW	RV	efficiency %
uqcrh.1	ENSSAUG00010009981	ATCCTCTTGATGCGATGCGA	TGGTCCCGAGCATGTAGGA	98
uqcrh.2	ENSSAUG00010022471	TGGTGGACCCTTTAGAATCG	TCACGTTGTGGAAGAGCTTG	98
uqcrc2.2	ENSSAUG00010022860	TTGCCAAAGCTACCACACAG	TCCGATACATTTCCCTCAGC	95
uqcrc2.1	ENSSAUG00010002780	CTTCAGAGGGTTTGCTGGAG	CTCATCCACGAAGGGTGTTT	97
uqcrfs1.2	ENSSAUG00010025510	GATATTGGGGAGCTGCGTGA	TGAAGCATCGTAGTGGGAGC	102
uqcrfs1.1	ENSSAUG00010000160	TCCGATTTGCCCACACAGAT	AGATTCGCTGCTGTCCTGAG	105
uqcr11.2	ENSSAUG00010016776	CTATTCTGAGGGCGTGGGTG	TGTACGGCACATAGTCGAGA	102
uqcr11.1	ENSSAUG00010018172	GCTCAGCAAAATAATCGGTCAGA	TGAAGTGGACCAGGGCTACA	103
ndufv1.2	ENSSAUG00010012863	TTCCTGCTGATGTGGGTGTG	CAGCTTTGTGCCGGAGTTTC	105
ndufv1.1	ENSSAUG00010003862	ATGAAGTCCATCACCGACCG	CGTGTTCTCCTCTGAAGCGT	98
ndufs1.1	ENSSAUG00010006080	CCCACTCTTCAACGCCAGAA	TCCCAGGTGGTCATACGAGT	104
ndufs1.2	ENSSAUG00010021399	ATGAGAGTCCTGCCTCGTCT	AGCATCTTCCCAGGTTGTGG	95
ef1	ENSSAUG00010026616	TCAAGGGATGGAAGGTTGAG	AGTTCCAATACCGCCGAT	99
rps18	ENSSAUG00010000811	AGGGTGTTGGCAGACGTTAC	GAGGACCTGGCTGTATTTGC	95
rpl13	ENSSAUG00010003114	TCTGGAGGACTGTCAGGGGCATGC	AGACGCACAATCTTAAGAGCAG	95

Statistical analysis

Gene expression data were square-root transformed to meet the assumptions of normality and/or homogeneity. In addition, the Shapiro-Wilk normality test was applied to the normalized gene expression data to check for actual normal distribution before the appropriate statistical tests were applied. Differences in the expression levels of each gene between the different DTs and control vs exercised groups were analyzed using two-sided t-test, Welch two sample t-test or its analogues Wilcoxon rank sum test. Differences were

considered significant at $p < 0.05$ level. All statistical analyses were performed using R (R Core Team, 2022; RStudio Team, 2022) packages for statistical analysis.

Results

The expression of OXPHOS genes of complex I (*ndufs11.1* and *ndufs1.2*) and complex III (*uqcrfs1.1, uqcrfs1.2, uqcrh.1, uqcrh.2, uqcrc2.1, uqcrc2.2, uqcr11.1, uqcr11.2*) was measured in non-exercised (control) and exercised fish (**Figure 5.2**). Genes encoding for OXPHOS proteins remained unaffected by DT at the control fish, with the only exception of *uqcr11a* that was significantly lower at 17°C compared to 20°C.

However, swimming led to significant increases in the expression of 9 genes (*ndufs1.1, ndufs1.2, uqcrfs1.1, uqcrfs1.2, uqcr11.2, uqcrh.1, uqcrh.2 uqcrc2.1, uqcrc2.2*) as compared with the control fish. These increases were very common in fish reared at 23°C, whereas none such increase was observed in fish reared at 17°C. As a result, DT governed differential responses in genes implicated both for complex I and complex III subunits; exercised fish reared at 20°C or/and 23°C exhibited significantly higher expression levels than exercised fish reared at 17°C (**Figure 5.2**).

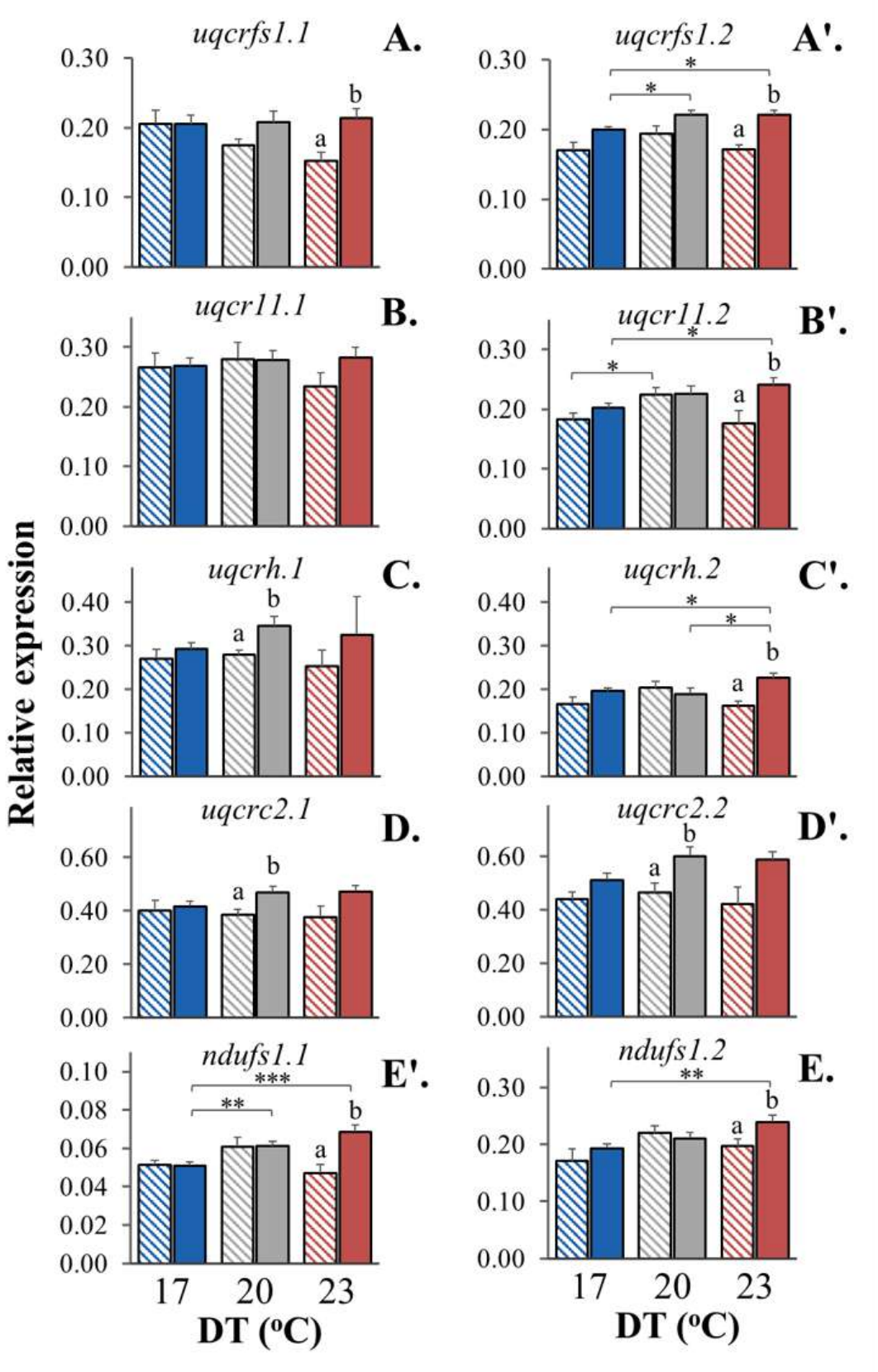


Figure 5.2. Relative expression of OXPHOS genes per DT in control (shaded) and exercised (full) fish. Asterisks represent statistical significant differences between the different DTs, *p < 0.05, **p < 0.01 and ***p < 0.001. letters represent statistically significant differences between the control and exercised fish.

Discussion

Numerous animals, including fish, have been used to study how exercise affects mitochondrial biogenesis and OXPHOS regulation and efficiency (Russell et al., 2005). Although the effects of exercise on those processes are obvious, the ways in which various elements such as the type and duration of the exercise (Martins et al., 2018), the nutrition intake (Perelló-Amorós et al., 2021), the general health of the exercised individual, the requirements of particular tissues (D. F. Wilson, 2017) and/or developmental stages, as well as environmental elements like temperature (Gerber et al., 2020) or oxygen availability (Y. Zhang et al., 2019) affect those processes can vary. In this chapter, it was demonstrated that while the various DTs seemed to have no effect on metamorphosing fish under normal conditions, the exercise challenge revealed a differential capacity to respond, and this capacity was shaped by early DT. DT changes that occurred within a brief window of time at the start of development were sufficient to affect the plasticity of the OXPHOS gene response later on, under highly energetically demanding circumstances like acute exercise. In accordance with other studies (Katakura et al., 2019; Maehr et al., 2013; T. Wang et al., 2018) that recorded the specific response of paralogs under different stimuli, no uniform response of the examined OXPHOS gene paralogs to DT and exercise was recorded, providing evidence that regulatory neo/subfunctionalization followed teleost-specific whole genome duplication. Two additional aspects of OXPHOS paralog plasticity became apparent by this chapter's findings. It was revealed that, while at an idle state there might be no differentiation in the expression levels between paralogs, when an energy challenging task was to be accomplished the transcription of one paralog was augmented. The retention of the paralogs and the way that they are regulated could indicate a preventing mechanism for aimless energy production, but at the same time provide the potential for quick adjustment when a high energy demanding response to a stimulus is developed. Taking into consideration that the original thermal stimulus was applied at early development and it could trigger a differential response later in fish life, there is a strong indication of epigenetic modification mechanism upon OXPHOS paralogs that may distinguish between paralogs. There are numerus studies (Anastasiadi et al., 2021; Jonsson & Jonsson, 2019; Rocha et al., 2015) underlining the contribution of epigenetic mechanisms in gene regulation of developing fish, while other studies have shown the epigenetic regulation on OXPHOS genes in hypoxia (Farhat et al.,

2022) and fish sperm quality (Merino et al., 2023); in that scope the investigation of the possible epigenetic alternations of OXPHOS gene paralogs in different aspect of fish life seems promising. The results of the collaborating team made clear that the lower temperature (17°C) had a negative impact on larvae physiology and morphology, and they correlated those findings with the poor swimming performance of that group. Interestingly, the same group of fish reared at 17^{o}C didn't manage to build a differential response to exercise (**Figure 5.2**). While the subject is really complex by itself and the results of this analysis are not sufficient to support any theory about the causality factors of those observations, it seems that the low DT was a detrimental factor shaping fish development, influencing larva physiology and morphology along with the regulation at the molecular level of a key metabolic function. In conclusion, the findings in this chapter offer additional evidence of the regulatory neo/subneofunctionalization fate of the retaining OXPHOS paralogs as an advantage under varying conditions. This advantage lies on the precise tuning of the OXPHOS complex, potentially impacting OXPHOS as a whole, to meet specific physiological, developmental, and behavioral requirements. It is evident that further exploration of the adaptability of the OXPHOS machinery and a deeper understanding of the regulatory mechanisms at play will contribute to illuminating cellular energetics in teleosts. In addition, understanding the adaptability of the OXPHOS machinery and the regulatory mechanisms involved could offer practical means to monitor fish performance in response to a changing environment.

Chapter 6

Conclusions

Studying oxidative phosphorylation (OXPHOS) evolution in teleosts, with a focus on gilthead seabream and European seabass, has revealed intriguing genomic and functional aspects of this critical cellular process. This study adds to our understanding of OXPHOS in teleosts, a taxonomic group with vast species diversity. The most important findings of the study are summarized below.

The entire gene palette of OXPHOS has been identified in two teleost species, namely the gilthead seabream and European seabass, both of significant economic and ecological interest. This study represents the first comprehensive exploration of OXPHOS evolution in teleosts with a complete mapping for the nuclear OXPHOS genes on chromosomes, their syntenic relationships and conservation of their deduced protein structure. The findings underline that gene duplication played a pivotal role in shaping OXPHOS in teleosts, with whole genome duplication (WGD) identified as the primary mechanism contributing to this. Among the three rounds of vertebrate WGDs, the most recent, the teleost specific genome duplication (TSGD,) emerged as the evolutionary origin of the majority of OXPHOS paralogs (Chapter 2). This extra complexity of OXPHOS in teleosts provides another view angle to our traditional comprehension, which has primarily relied on research conducted in unicellular organisms and mammals (Fernandez-Vizarra & Zeviani, 2018; Iwata et al., 1998, Rutter et al., 2010; Habersetzer et al., 2013; Liu & Barrientos, 2013; Mackler & Haynes, 1973). Besides the profound complexity introduced by the genetic redundancy, supported by the analysis of polymorphic sites, evolutionary modes at action, subcellular localization and conservation of gene structure and protein domains, OXPHOS genes are under strong evolutionary pressure to preserve the subunit function and structure safeguarding the function of OXPHOS machinery (Chapter 3). The evidence generated in Chapter 2 and Chapter 3 are supportive of regulatory neo/subfunctionalization as the most plausible fate of the retained paralogs, which would mean that the paralogs are transcribed, their transcription is differentially regulated, and the products of the transcripts are targeted in the mitochondria.

OXPHOS, despite its universal importance to life, is not a static biological function. Indeed, OXPHOS paralogs exhibited a variety at the expression levels and patterns throughout larval development indicative of a refined regulation. The study offers the first systematic exploration of OXPHOS genes expression in teleost ontogeny and provides insight into how expression patterns differentiate between the complexes and genes at different

developmental points (Chapter 4). This refined regulation was also evident in the tissue-specific expression profiles of the paralogs as well as the response to a challenge, suggesting that each paralog contributes to the fine tuning of OXPHOS to meet the needs of the organism as dictated by the physiological, morphological and behavioral characteristics of each developmental stage and following stimulus. Chapter 5 raised evidence that OXPHOS paralogs are also subject to differential epigenetic regulation, adding to the complexity of regulation of OXPHOS paralogs, and to the factors that may drive adaptive plasticity to the cellular energetics in teleosts.

The synthesis of findings from different approaches in this study clarified a notable duality. On one hand, there exists a significant pressure to maintain the functionality of OXPHOS machinery subunits, while on the other hand, there is a pronounced regulatory plasticity among OXPHOS paralogs. This points towards regulatory neo/subfunctionalization as the most probable mechanisms determining the fate of OXPHOS paralogs.

The proposed mechanism for the evolution of OXPHOS in teleosts initiates with the simultaneous duplication of all OXPHOS genes, in respect of dosage balance, an opportunity uniquely presented by whole genome duplication (WGD), in the common ancestor of teleosts. The diverse evolutionary trajectories and adaptations of teleost lineages and species have molded the OXPHOS machinery through lineage/species specific gene losses and duplications. When dosage constraints were gradually lifted over time, neo/subfunctionalization emerged as the predominant mechanism responsible for the retention of OXPHOS paralogs in teleost genomes.

Chapter 7

List of References

1. Afzal, N., Lederer, W. J., Jafri, M. S., & Mannella, C. A. (2021). Effect of crista morphology on mitochondrial ATP output: A computational study. *Current Research in Physiology*, *4*, 163–176. https://doi.org/10.1016/j.crphys.2021.03.005

2. Albokhadaim, I., Hammond, C. L., Ashton, C., Simbi, B. H., Bayol, S., Farrington, S., & Stickland, N. (2007). Larval programming of post-hatch muscle growth and activity in Atlantic salmon (*Salmo salar*). *Journal of Experimental Biology*, *210*(10), 1735–1741. https://doi.org/10.1242/jeb.003194

3. Allendorf, F. W., & Thorgaard, G. H. (1984). Tetraploidy and the Evolution of Salmonid Fishes. In B. J. Turner (Ed.), *Evolutionary Genetics of Fishes* (pp. 1–53). Springer US. https://doi.org/10.1007/978-1-4684-4652-4_1

4. Almeida-Silva, F., & Van De Peer, Y. (2023). Whole-genome Duplications and the Long-term Evolution of Gene Regulatory Networks in Angiosperms. *Molecular Biology and Evolution*, *40*(7), msad141. https://doi.org/10.1093/molbev/msad141

5. Anastasiadi, D., Shao, C., Chen, S., & Piferrer, F. (2021). Footprints of global change in marine life: Inferring past environment based on DNA methylation and gene expression marks. *Molecular Ecology*, *30*(3), 747–760. https://doi.org/10.1111/mec.15764

6. Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

7. Aparicio, S., Chapman, J., Stupka, E., Putnam, N., Chia, J., Dehal, P., Christoffels, A., Rash, S., Hoon, S., Smit, A., Gelpke, M. D. S., Roach, J., Oh, T., Ho, I. Y., Wong, M., Detter, C., Verhoef, F., Predki, P., Tay, A., … Brenner, S. (2002). Whole-Genome Shotgun Assembly and Analysis of the Genome of *Fugu rubripes*. *Science*, *297*(5585), 1301–1310. https://doi.org/10.1126/science.1072104

8. Baduel, P., Bray, S., Vallejo-Marin, M., Kolář, F., & Yant, L. (2018). The "Polyploid Hop": Shifting Challenges and Opportunities Over the Evolutionary Lifespan of Genome Duplications. In Frontiers in Ecology and Evolution (Vol. 6). Frontiers Media SA. https://doi.org/10.3389/fevo.2018.00117

9. Bailey, J. A., Yavor, A. M., Massa, H. F., Trask, B. J., & Eichler, E. E. (2001). Segmental Duplications: Organization and Impact Within the Current Human Genome Project Assembly. *Genome Research*, *11*(6), 1005–1017. https://doi.org/10.1101/gr.187101

10. Barton, N. H. (2008). The effect of a barrier to gene flow on patterns of geographic variation. *Genetics Research*, *90*(1), 139–149. https://doi.org/10.1017/S0016672307009081

11. Bayrhuber, M., Meins, T., Habeck, M., Becker, S., Giller, K., Villinger, S., Vonrhein, C., Griesinger, C., Zweckstetter, M., & Zeth, K. (2008). Structure of the human voltage-dependent

anion channel. *Proceedings of the National Academy of Sciences*, *105*(40), 15370–15375. https://doi.org/10.1073/pnas.0808115105

12. Behe, M. J. (2010). Experimental evolution, loss-of-function mutations, and "the first rule of adaptive evolution." *The Quarterly Review of Biology*, *85*(4), 419–445. https://doi.org/10.1086/656902

13. Beisswanger, S., & Stephan, W. (2008). Evidence that strong positive selection drives neofunctionalization in the tandemly duplicated *polyhomeotic* genes in *Drosophila*. *Proceedings of the National Academy of Sciences*, *105*(14), 5447–5452. https://doi.org/10.1073/pnas.0710892105

14. Berrisford, J. M., Baradaran, R., & Sazanov, L. A. (2016). Structure of bacterial respiratory complex I. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, *1857*(7), 892–901. https://doi.org/10.1016/j.bbabio.2016.01.012

15. Birchler, J. A., Riddle, N. C., Auger, D. L., & Veitia, R. A. (2005). Dosage balance in gene regulation: Biological implications. *Trends in Genetics*, *21*(4), 219–226. https://doi.org/10.1016/j.tig.2005.02.010

16. Birchler, J. A., & Veitia, R. A. (2021). One Hundred Years of Gene Balance: How Stoichiometric Issues Affect Gene Expression, Genome Evolution, and Quantitative Traits. *Cytogenetic and Genome Research*, *161*(10–11), 529–550. https://doi.org/10.1159/000519592

17. Birchler, J. A., & Yang, H. (2022). The multiple fates of gene duplications: Deletion, hypofunctionalization, subfunctionalization, neofunctionalization, dosage balance constraints, and neutral variation. *The Plant Cell*, *34*(7), 2466–2474. https://doi.org/10.1093/plcell/koac076

18. Bishop, C. D., Erezyilmaz, D. F., Flatt, T., Georgiou, C. D., Hadfield, M. G., Heyland, A., Hodin, J., Jacobs, M. W., Maslakova, S. A., Pires, A., Reitzel, A. M., Santagata, S., Tanaka, K., & Youson, J. H. (2006). What is metamorphosis? *Integrative and Comparative Biology*, *46*(6), 655–661. https://doi.org/10.1093/icb/icl004

19. Blomme, T., Vandepoele, K., De Bodt, S., Simillion, C., Maere, S., & Van De Peer, Y. (2006). [No title found]. *Genome Biology*, *7*(5), R43. https://doi.org/10.1186/gb-2006-7-5-r43

20. Bourlat, S. J., Juliusdottir, T., Lowe, C. J., Freeman, R., Aronowicz, J., Kirschner, M., Lander, E. S., Thorndyke, M., Nakano, H., Kohn, A. B., Heyland, A., Moroz, L. L., Copley, R. R., & Telford, M. J. (2006). Deuterostome phylogeny reveals monophyletic chordates and the new phylum Xenoturbellida. *Nature*, *444*(7115), 85–88. https://doi.org/10.1038/nature05241

21. Braasch, I., Guiguen, Y., Loker, R., Letaw, J. H., Ferrara, A., Bobe, J., & Postlethwait, J. H. (2014). Connectivity of vertebrate genomes: Paired-related homeobox (Prrx) genes in spotted gar,

basal teleosts, and tetrapods. In Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology (Vol. 163, pp. 24–36). Elsevier BV. https://doi.org/10.1016/j.cbpc.2014.01.005

22. Bridges, C. B. (1936). The Bar "Gene" a Duplication. *Science*, *83*(2148), 210–211. https://doi.org/10.1126/science.83.2148.210

23. Brunet, F. G., Crollius, H. R., Paris, M., Aury, J.-M., Gibert, P., Jaillon, O., Laudet, V., & Robinson-Rechavi, M. (2006). Gene Loss and Evolutionary Rates Following Whole-Genome Duplication in Teleost Fishes. *Molecular Biology and Evolution*, *23*(9), 1808–1816. https://doi.org/10.1093/molbev/msl049

24. Byrne, K. P., & Wolfe, K. H. (2005). The Yeast Gene Order Browser: Combining curated homology and syntenic context reveals gene fate in polyploid species. *Genome Research*, *15*(10), 1456–1461. https://doi.org/10.1101/gr.3672305

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

26. Campinho, M. A., Silva, N., Sweeney, G. E., & Power, D. M. (2006). Molecular, cellular and histological changes in skin from a larval to an adult phenotype during bony fish metamorphosis. *Cell and Tissue Research*, *327*(2), 267–284. https://doi.org/10.1007/s00441-006-0262-9

27. Chacinska, A., Koehler, C. M., Milenkovic, D., Lithgow, T., & Pfanner, N. (2009). Importing Mitochondrial Proteins: Machineries and Mechanisms. *Cell*, *138*(4), 628–644. https://doi.org/10.1016/j.cell.2009.08.005

28. Chanderbali, A. S., Berger, B. A., Howarth, D. G., Soltis, D. E., & Soltis, P. S. (2017). Evolution of floral diversity: Genomics, genes and *gamma*. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1713), 20150509. https://doi.org/10.1098/rstb.2015.0509

29. Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, *34*(17), i884–i890. https://doi.org/10.1093/bioinformatics/bty560

30. Christoffels, A., Koh, E. G. L., Chia, J., Brenner, S., Aparicio, S., & Venkatesh, B. (2004). Fugu Genome Analysis Provides Evidence for a Whole-Genome Duplication Early During the Evolution of Ray-Finned Fishes. *Molecular Biology and Evolution*, *21*(6), 1146–1151. https://doi.org/10.1093/molbev/msh114

31. Christou, M., Iliopoulou, M., Witten, P. E., & Koumoundouros, G. (2018). Segmentation pattern of zebrafish caudal fin is affected by developmental temperature and defined by

multiple fusions between segments. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, *330*(6–7), 330–340. https://doi.org/10.1002/jez.b.22825

32. Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., Land, S. J., Lu, X., & Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w ^{1118} ; iso-2; iso-3. *Fly*, *6*(2), 80–92. https://doi.org/10.4161/fly.19695

33. Cogliati, S., Enriquez, J. A., & Scorrano, L. (2016). Mitochondrial Cristae: Where Beauty Meets Functionality. *Trends in Biochemical Sciences*, *41*(3), 261–273. https://doi.org/10.1016/j.tibs.2016.01.001

34. Conant, G. C., & Wolfe, K. H. (2008). Turning a hobby into a job: How duplicated genes find new functions. *Nature Reviews Genetics*, *9*(12), 938–950. https://doi.org/10.1038/nrg2482

35. Čunátová, K., Pajuelo Reguera, D., Houštěk, J., Mráček, T., & Pecina, P. (2020). Role of cytochrome c oxidase nuclear-encoded subunits in health and disease. *Physiological Research*, 947–965. https://doi.org/10.33549/physiolres.934446

36. Cunningham, F., Allen, J. E., Allen, J., Alvarez-Jarreta, J., Amode, M. R., Armean, I. M., Austine-Orimoloye, O., Azov, A. G., Barnes, I., Bennett, R., Berry, A., Bhai, J., Bignell, A., Billis, K., Boddu, S., Brooks, L., Charkhchi, M., Cummins, C., Da Rin Fioretto, L., … Flicek, P. (2022). Ensembl 2022. *Nucleic Acids Research*, *50*(D1), D988–D995. https://doi.org/10.1093/nar/gkab1049

37. Davies, W. I. L., Zheng, L., Hughes, S., Tamai, T. K., Turton, M., Halford, S., Foster, R. G., Whitmore, D., & Hankins, M. W. (2011). Functional diversity of melanopsins and their global expression in the teleost retina. *Cellular and Molecular Life Sciences*, *68*(24), 4115–4132. https://doi.org/10.1007/s00018-011-0785-4

38. Debes, P. V., Solberg, M. F., Matre, I. H., Dyrhovden, L., & Glover, K. A. (2021). Genetic variation for upper thermal tolerance diminishes within and between populations with increasing acclimation temperature in Atlantic salmon. *Heredity*, *127*(5), 455–466. https://doi.org/10.1038/s41437-021-00469-y

39. Dehal, P., & Boore, J. L. (2005). Two Rounds of Whole Genome Duplication in the Ancestral Vertebrate. *PLoS Biology*, *3*(10), e314. https://doi.org/10.1371/journal.pbio.0030314

40. Delport, W., Poon, A. F. Y., Frost, S. D. W., & Kosakovsky Pond, S. L. (2010). Datamonkey 2010: A suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics*, *26*(19), 2455–2457. https://doi.org/10.1093/bioinformatics/btq429

41. Delsuc, F., Brinkmann, H., Chourrout, D., & Philippe, H. (2006). Tunicates and not cephalochordates are the closest living relatives of vertebrates. *Nature*, *439*(7079), 965–968. https://doi.org/10.1038/nature04336

42. Demin, O. V., Kholodenko, B. N., & Skulachev, V. P. (1998). A model of O·2 -generation in the complex III of the electron transport chain. In V. A. Saks, R. Ventura-Clapier, X. Leverve, A. Rossi, & M. Rigoulet (Eds.), *Bioenergetics of the Cell: Quantitative Aspects* (pp. 21–33). Springer US. https://doi.org/10.1007/978-1-4615-5653-4_3

43. Deplano, M., Diaz, J. P., Connes, R., Kentouri-Divanach, M., & Cavalier, F. (1991). Appearance of lipid-absorption capacities in larvae of the sea bassDicentrarchus labrax during transition to the exotrophic phase. *Marine Biology*, *108*(3), 361–371. https://doi.org/10.1007/BF01313645

44. Dimitriadi, A., Beis, D., Arvanitidis, C., Adriaens, D., & Koumoundouros, G. (2018). Developmental temperature has persistent, sexually dimorphic effects on zebrafish cardiac anatomy. *Scientific Reports*, *8*(1), 8125. https://doi.org/10.1038/s41598-018-25991-8

45. Dimroth, P. (2000). Operation of the F0 motor of the ATP synthase. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, *1458*(2–3), 374–386. https://doi.org/10.1016/S0005-2728(00)00088-8

46. Drobek, M. (2022). Paralogous Genes Involved in Embryonic Development: Lessons from the Eye and other Tissues. *Genes*, *13*(11), 2082. https://doi.org/10.3390/genes13112082

47. Du, Z., Zhou, X., Lai, Y., Xu, J., Zhang, Y., Zhou, S., Feng, Z., Yu, L., Tang, Y., Wang, W., Yu, L., Tian, C., Ran, T., Chen, H., Guddat, L. W., Liu, F., Gao, Y., Rao, Z., & Gong, H. (2023). Structure of the human respiratory complex II. Proceedings of the National Academy of Sciences of the United States of America, 120(18), e2216713120. https://doi.org/10.1073/pnas.2216713120

48. Duggan, A. T., Kocha, K. M., Monk, C. T., Bremer, K., & Moyes, C. D. (2011). Coordination of cytochrome *c* oxidase gene expression in the remodelling of skeletal muscle. *Journal of Experimental Biology*, *214*(11), 1880–1887. https://doi.org/10.1242/jeb.053322

49. Edger, P. P., & Pires, J. C. (2009). Gene and genome duplications: The impact of dosage-sensitivity on the fate of nuclear genes. *Chromosome Research*, *17*(5), 699–717. https://doi.org/10.1007/s10577-009-9055-9

50. El-Fiky, N., Hinterleitner, S., & Wieser, W. (1987). Differentiation of swimming muscles and gills, and development of anaerobic power in the larvae of cyprinid fish (Pisces, Teleostei). *Zoomorphology*, *107*(2), 126–132. https://doi.org/10.1007/BF00312122

51. Emanuelsson, O., Nielsen, H., Brunak, S., & Von Heijne, G. (2000). Predicting Subcellular Localization of Proteins Based on their N-terminal Amino Acid Sequence. *Journal of Molecular Biology*, *300*(4), 1005–1016. https://doi.org/10.1006/jmbi.2000.3903

52. Erecińska, M., & Silver, I. A. (1989). ATP and Brain Function. *Journal of Cerebral Blood Flow & Metabolism*, *9*(1), 2–19. https://doi.org/10.1038/jcbfm.1989.2

53. Erwin, D. H., Laflamme, M., Tweedt, S. M., Sperling, E. A., Pisani, D., & Peterson, K. J. (2011). The Cambrian Conundrum: Early Divergence and Later Ecological Success in the Early History of Animals. *Science*, *334*(6059), 1091–1097. https://doi.org/10.1126/science.1206375

54. Escriva, H., Bertrand, S., Germain, P., Robinson-Rechavi, M., Umbhauer, M., Cartry, J., Duffraisse, M., Holland, L., Gronemeyer, H., & Laudet, V. (2006). Neofunctionalization in Vertebrates: The Example of Retinoic Acid Receptors. *PLoS Genetics*, *2*(7), e102. https://doi.org/10.1371/journal.pgen.0020102

55. Facey, D. E., Bowen, B. W., Collette, B. B., & Helfman, G. S. (2023). *The diversity of fishes: Biology, evolution and ecology* (Third edition). Wiley.

56. Falkowski, P. G. (2006). Tracing Oxygen's Imprint on Earth's Metabolic Evolution. *Science*, *311*(5768), 1724–1725. https://doi.org/10.1126/science.1125937

57. Farhat, E., Talarico, G. G. M., Grégoire, M., Weber, J.-M., & Mennigen, J. A. (2022). Epigenetic and post-transcriptional repression support metabolic suppression in chronically hypoxic goldfish. *Scientific Reports*, *12*(1), 5576. https://doi.org/10.1038/s41598-022-09374-8

58. Fernandez-Vizarra, E., & Zeviani, M. (2018). Mitochondrial complex III Rieske Fe-S protein processing and assembly. *Cell Cycle*, *17*(6), 681–687. https://doi.org/10.1080/15384101.2017.1417707

59. Finn, R. N., & Fyhn, H. J. (2010). Requirement for amino acids in ontogeny of fish. *Aquaculture Research*, *41*(5), 684–716. https://doi.org/10.1111/j.1365-2109.2009.02220.x

60. Fontanesi, F., Soto, I. C., & Barrientos, A. (2008). Cytochrome *c* oxidase biogenesis: New levels of regulation. *IUBMB Life*, *60*(9), 557–568. https://doi.org/10.1002/iub.86

61. Force, A., Lynch, M., Pickett, F. B., Amores, A., Yan, Y., & Postlethwait, J. (1999). Preservation of Duplicate Genes by Complementary, Degenerative Mutations. *Genetics*, *151*(4), 1531–1545. https://doi.org/10.1093/genetics/151.4.1531

62. Fox, T. D. (2012). Mitochondrial Protein Synthesis, Import, and Assembly. *Genetics*, *192*(4), 1203–1234. https://doi.org/10.1534/genetics.112.141267

63. Freeling, M. (2009). Bias in Plant Gene Content Following Different Sorts of Duplication: Tandem, Whole-Genome, Segmental, or by Transposition. *Annual Review of Plant Biology*, *60*(1), 433–453. https://doi.org/10.1146/annurev.arplant.043008.092122

64. Freeling, M., & Thomas, B. C. (2006). Gene-balanced duplications, like tetraploidy, provide predictable drive to increase morphological complexity. *Genome Research*, *16*(7), 805–814. https://doi.org/10.1101/gr.3681406

65. Geladakis, G., Kourkouta, C., Somarakis, S., & Koumoundouros, G. (2022). Developmental Temperature Shapes the Otolith Morphology of Metamorphosing and Juvenile Gilthead Seabream (Sparus aurata Linnaeus, 1758). *Fishes*, *7*(2), 82. https://doi.org/10.3390/fishes7020082

66. Georga, I., & Koumoundouros, G. (2010). Thermally induced plasticity of body shape in adult zebrafish *Danio rerio* (Hamilton, 1822). *Journal of Morphology*, *271*(11), 1319–1327. https://doi.org/10.1002/jmor.10874

67. Georgakopoulou, E., Sfakianakis, D. G., Kouttouki, S., Divanach, P., Kentouri, M., & Koumoundouros, G. (2007). The influence of temperature during early life on phenotypic expression at later ontogenetic stages in sea bass. *Journal of Fish Biology*, *70*(1), 278–291. https://doi.org/10.1111/j.1095-8649.2007.01305.x

68. Georgiou, S., Alami-Durante, H., Power, D. M., Sarropoulou, E., Mamuris, Z., & Moutou, K. A. (2016). Transient up- and down-regulation of expression of myosin light chain 2 and myostatin mRNA mark the changes from stratified hyperplasia to muscle fiber hypertrophy in larvae of gilthead sea bream (Sparus aurata L.). *Cell and Tissue Research*, *363*(2), 541–554. https://doi.org/10.1007/s00441-015-2254-0

69. Gerber, L., Clow, K. A., Mark, F. C., & Gamperl, A. K. (2020). Improved mitochondrial function in salmon (Salmo salar) following high temperature acclimation suggests that there are cracks in the proverbial 'ceiling.' *Scientific Reports*, *10*(1), 21636. https://doi.org/10.1038/s41598-020-78519-4

70. Gibson, S., & Johnston, I. A. (1995). Temperature and development in larvae of the turbot Scophthalmus maximus. *Marine Biology*, *124*(1), 17–25. https://doi.org/10.1007/BF00349142

71. Glasauer, S. M. K., & Neuhauss, S. C. F. (2014). Whole-genome duplication in teleost fishes and its evolutionary consequences. *Molecular Genetics and Genomics*, *289*(6), 1045–1060. https://doi.org/10.1007/s00438-014-0889-2

72. Gu, W., Yang, Y., Ning, C., Wang, Y., Hu, J., Zhang, M., Kuang, S., Sun, Y., Li, Y., Zhang, Y., Sun, J., Ying, D., & Xu, S. (2021). Identification and characteristics of insulin-like growth factor system in the brain, liver, and gonad during development of a seasonal breeding teleost, Pampus argenteus. *General and Comparative Endocrinology*, *300*, 113645. https://doi.org/10.1016/j.ygcen.2020.113645

73. Guo, R., Zong, S., Wu, M., Gu, J., & Yang, M. (2017). Architecture of Human Mitochondrial Respiratory Megacomplex I2III2IV2. Cell, 170(6), 1247–1257.e12. https://doi.org/10.1016/j.cell.2017.07.050

74. Habersetzer, J., Larrieu, I., Priault, M., Salin, B., Rossignol, R., Brèthes, D., & Paumard, P. (2013). Human F1F0 ATP Synthase, Mitochondrial Ultrastructure and OXPHOS Impairment: A (Super-)Complex Matter? *PLoS ONE*, *8*(10), e75429. https://doi.org/10.1371/journal.pone.0075429

75. Hahn, M. W. (2009). Distinguishing Among Evolutionary Models for the Maintenance of Gene Duplicates. *Journal of Heredity*, *100*(5), 605–617. https://doi.org/10.1093/jhered/esp047

76. Hatefi, Y. (1985). THE MITOCHONDRIAL ELECTRON TRANSPORT AND OXIDATIVE PHOSPHORYLATION SYSTEM. *Annual Review of Biochemistry*, *54*(1), 1015–1069. https://doi.org/10.1146/annurev.bi.54.070185.005055

77. He, X., & Zhang, J. (2005). Rapid Subfunctionalization Accompanied by Prolonged and Substantial Neofunctionalization in Duplicate Gene Evolution. *Genetics*, *169*(2), 1157–1164. https://doi.org/10.1534/genetics.104.037051

78. Hoegg, S., Brinkmann, H., Taylor, J. S., & Meyer, A. (2004). Phylogenetic Timing of the Fish-Specific Genome Duplication Correlates with the Diversification of Teleost Fish. *Journal of Molecular Evolution*, *59*(2), 190–203. https://doi.org/10.1007/s00239-004-2613-z

79. Hosler, J. P., Ferguson-Miller, S., Calhoun, M. W., Thomas, J. W., Hill, J., Lemieux, L., Ma, J., Georgiou, C., Fetter, J., Shapleigh, J., Tecklenburg, M. M. J., Babcock, G. T., & Gennis, R. B. (1993). Insight into the active-site structure and function of cytochrome oxidase by analysis of site-directed mutants of bacterial cytochromeaa 3 and cytochromebo. *Journal of Bioenergetics and Biomembranes*, *25*(2), 121–136. https://doi.org/10.1007/BF00762854

80. Hu, B., Jin, J., Guo, A.-Y., Zhang, H., Luo, J., & Gao, G. (2015). GSDS 2.0: An upgraded gene feature visualization server. *Bioinformatics*, *31*(8), 1296–1297. https://doi.org/10.1093/bioinformatics/btu817

81. Huo, Y., Qiu, W.-Y., Pan, Q., Yao, Y.-F., Xing, K., & Lou, M. F. (2009). Reactive oxygen species (ROS) are essential mediators in epidermal growth factor (EGF)-stimulated corneal epithelial cell proliferation, adhesion, migration, and wound healing. *Experimental Eye Research*, *89*(6), 876–886. https://doi.org/10.1016/j.exer.2009.07.012

82. Hurles, M. E., & Lupski, J. R. (2006). Recombination Hotspots in Nonallelic Homologous Recombination. In J. R. Lupski & P. Stankiewicz (Eds.), *Genomic Disorders* (pp. 341–355). Humana Press. https://doi.org/10.1007/978-1-59745-039-3_24

83. Hurst, L. D., & Smith, N. G. C. (1998). The evolution of concerted evolution. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *265*(1391), 121–127. https://doi.org/10.1098/rspb.1998.0272

84. Inoue, J., Sato, Y., Sinclair, R., Tsukamoto, K., & Nishida, M. (2015). Rapid genome reshaping by multiple-gene loss after whole-genome duplication in teleost fish suggested by mathematical modeling. *Proceedings of the National Academy of Sciences*, *112*(48), 14918–14923. https://doi.org/10.1073/pnas.1507669112

85. Ishikawa, K., Ishihara, A., & Moriya, H. (2020). Exploring the Complexity of Protein-Level Dosage Compensation that Fine-Tunes Stoichiometry of Multiprotein Complexes. *PLOS Genetics*, *16*(10), e1009091. https://doi.org/10.1371/journal.pgen.1009091

86. Iwata, S., Lee, J. W., Okada, K., Lee, J. K., Iwata, M., Rasmussen, B., Link, T. A., Ramaswamy, S., & Jap, B. K. (1998). Complete Structure of the 11-Subunit Bovine Mitochondrial Cytochrome bc $_1$ Complex. *Science*, *281*(5373), 64–71. https://doi.org/10.1126/science.281.5373.64

87. Janssen, R. J. R. J., Nijtmans, L. G., Heuvel, L. P. V. D., & Smeitink, J. A. M. (2006). Mitochondrial complex I: Structure, function and pathology. *Journal of Inherited Metabolic Disease*, *29*(4), 499–515. https://doi.org/10.1007/s10545-006-0362-4

88. Jaillon, O., Aury, J.-M., Brunet, F., Petit, J.-L., Stange-Thomann, N., Mauceli, E., Bouneau, L., Fischer, C., Ozouf-Costaz, C., Bernot, A., Nicaud, S., Jaffe, D., Fisher, S., Lutfalla, G., Dossat, C., Segurens, B., Dasilva, C., Salanoubat, M., Levy, M., … Roest Crollius, H. (2004). Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype. In Nature (Vol. 431, Issue 7011, pp. 946–957). Springer Science and Business Media LLC. https://doi.org/10.1038/nature03025

89. Johnston, I. A. (2006). Environment and plasticity of myogenesis in teleost fish. *Journal of Experimental Biology*, *209*(12), 2249–2264. https://doi.org/10.1242/jeb.02153

90. Johnston, I. A., Lee, H.-T., Macqueen, D. J., Paranthaman, K., Kawashima, C., Anwar, A., Kinghorn, J. R., & Dalmay, T. (2009). Embryonic temperature affects muscle fibre recruitment in adult zebrafish: Genome-wide changes in gene and microRNA expression associated with the transition from hyperplastic to hypertrophic growth phenotypes. *Journal of Experimental Biology*, *212*(12), 1781–1793. https://doi.org/10.1242/jeb.029918

91. Jonsson, B., & Jonsson, N. (2019). Phenotypic plasticity and epigenetics of fish: Embryo temperature affects later-developing life-history traits. *Aquatic Biology*, *28*, 21–32. https://doi.org/10.3354/ab00707

92. Kadenbach, B., Jarausch, J., Hartmann, R., & Merle, P. (1983). Separation of mammalian cytochrome c oxidase into 13 polypeptides by a sodium dodecyl sulfate-gel electrophoretic procedure. *Analytical Biochemistry*, *129*(2), 517–521. https://doi.org/10.1016/0003-2697(83)90586-9

93. Katakura, F., Nishiya, K., Wentzel, A. S., Hino, E., Miyamae, J., Okano, M., Wiegertjes, G. F., & Moritomo, T. (2019). Paralogs of Common Carp Granulocyte Colony-Stimulating Factor (G-CSF) Have Different Functions Regarding Development, Trafficking and Activation of Neutrophils. *Frontiers in Immunology*, *10*, 255. https://doi.org/10.3389/fimmu.2019.00255

94. Katoh, K. (2002). MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, *30*(14), 3059–3066. https://doi.org/10.1093/nar/gkf436

95. Keilin, D., & Keilin, J. (1970). *The history of cell respiration and cytochrome* (reprint). University Press.

96. Kim, D., Paggi, J. M., Park, C., Bennett, C., & Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology*, *37*(8), 907–915. https://doi.org/10.1038/s41587-019-0201-4

97. Kocha, K. M., Reilly, K., Porplycia, D. S. M., McDonald, J., Snider, T., & Moyes, C. D. (2015). Evolution of the oxygen sensitivity of cytochrome *c* oxidase subunit 4. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, *308*(4), R305–R320. https://doi.org/10.1152/ajpregu.00281.2014

98. Kondrashov, F. A. (2012). Gene duplication as a mechanism of genomic adaptation to a changing environment. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1749), 5048–5057. https://doi.org/10.1098/rspb.2012.1108

99. Kosakovsky Pond, S. L., Frost, S. D. W., Grossman, Z., Gravenor, M. B., Richman, D. D., & Brown, A. J. L. (2006). Adaptation to Different Human Populations by HIV-1 Revealed by Codon-Based Analyses. *PLoS Computational Biology*, *2*(6), e62. https://doi.org/10.1371/journal.pcbi.0020062

100. Koumoundouros, G., Ashton, C., Xenikoudakis, G., Giopanou, I., Georgakopoulou, E., & Stickland, N. (2009). Ontogenetic differentiation of swimming performance in Gilthead seabream (Sparus aurata, Linnaeus 1758) during metamorphosis. *Journal of Experimental Marine Biology and Ecology*, *370*(1–2), 75–81. https://doi.org/10.1016/j.jembe.2008.12.001

101. Kourkouta, C., Printzi, A., Geladakis, G., Mitrizakis, N., Papandroulakis, N., & Koumoundouros, G. (2021). Long lasting effects of early temperature exposure on the swimming performance and skeleton development of metamorphosing Gilthead seabream (Sparus aurata L.) larvae. *Scientific Reports*, *11*(1), 8787. https://doi.org/10.1038/s41598-021-88306-4

102. Kourkouta, C., Tsipourlianos, A., Power, D. M., Moutou, K. A., & Koumoundouros, G. (2022). Variability of key-performance-indicators in commercial gilthead seabream hatcheries. *Scientific Reports*, *12*(1), 17896. https://doi.org/10.1038/s41598-022-23008-z

103. Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution*, *35*(6), 1547–1549. https://doi.org/10.1093/molbev/msy096

104. Lai, Y., Zhang, Y., Zhou, S., Xu, J., Du, Z., Feng, Z., Yu, L., Zhao, Z., Wang, W., Tang, Y., Yang, X., Guddat, L. W., Liu, F., Gao, Y., Rao, Z., & Gong, H. (2023). Structure of the human ATP synthase. In Molecular Cell (Vol. 83, Issue 12, pp. 2137-2147.e4). Elsevier BV. https://doi.org/10.1016/j.molcel.2023.04.029

105. Lane, N., Allen, J. F., & Martin, W. (2010). How did LUCA make a living? Chemiosmosis in the origin of life. *BioEssays*, *32*(4), 271–280. https://doi.org/10.1002/bies.200900131

106. Leggatt, R. A., & Iwama, G. K. (2003). Occurrence of polyploidy in the fishes. *Reviews in Fish Biology and Fisheries*, *13*(3), 237–246. https://doi.org/10.1023/B:RFBF.0000033049.00668.fe

107. Lein, I., Holmefjord, I., & Rye, M. (1997). Effects of temperature on yolk sac larvae of Atlantic halibut (Hippoglossus hippoglossus L.). *Aquaculture*, *157*(1–2), 123–135. https://doi.org/10.1016/S0044-8486(97)00149-X

108. Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, *27*(21), 2987–2993. https://doi.org/10.1093/bioinformatics/btr509

109. Li, H. (2013). *Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM*. https://doi.org/10.48550/ARXIV.1303.3997

110. Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

111. Li, W.-H. (1997). *Molecular evolution*. Sinauer Associates.

112. Liao, D. (2008). Concerted Evolution. In John Wiley & Sons, Ltd (Ed.), *eLS* (1st ed.). Wiley. https://doi.org/10.1002/9780470015902.a0005132.pub2

113. Liongue, C., O'Sullivan, L. A., Trengove, M. C., & Ward, A. C. (2012). Evolution of JAK-STAT Pathway Components: Mechanisms and Role in Immune System Development. *PLoS ONE*, *7*(3), e32777. https://doi.org/10.1371/journal.pone.0032777

114. Liu, J., & Barrientos, A. (2013). Transcriptional Regulation of Yeast Oxidative Phosphorylation Hypoxic Genes by Oxidative Stress. In Antioxidants & Redox Signaling (Vol. 19, Issue 16, pp. 1916–1927). Mary Ann Liebert Inc. https://doi.org/10.1089/ars.2012.4589

115. Liu, L., Wu, Y., Liao, Z., Xiong, J., Wu, F., Xu, J., Lan, H., Tang, Q., Zhou, S., Liu, Y., & Lu, Y. (2018). Evolutionary conservation and functional divergence of the LFK gene family play important roles in the photoperiodic flowering pathway of land plants. *Heredity*, *120*(4), 310–328. https://doi.org/10.1038/s41437-017-0006-5

116. Lohman, B. K., Steinel, N. C., Weber, J. N., & Bolnick, D. I. (2017). Gene Expression Contributes to the Recent Evolution of Host Resistance in a Model Host Parasite System. *Frontiers in Immunology*, *8*, 1071. https://doi.org/10.3389/fimmu.2017.01071

117. Lynch, M. (2002). Gene Duplication and Evolution. *Science*, *297*(5583), 945–947. https://doi.org/10.1126/science.1075472

118. Lynch, M., & Conery, J. S. (2000). The Evolutionary Fate and Consequences of Duplicate Genes. *Science*, *290*(5494), 1151–1155. https://doi.org/10.1126/science.290.5494.1151

119. Lynch, M., & Force, A. (2000). The Probability of Duplicate Gene Preservation by Subfunctionalization. *Genetics*, *154*(1), 459–473. https://doi.org/10.1093/genetics/154.1.459

120. Mackler, B., & Haynes, B. (1973). Studies of oxidative phosphorylation in Saccharomyces cerevisiae and Saccharomyces carlsbergensis. In Biochimica et Biophysica Acta (BBA) - Bioenergetics (Vol. 292, Issue 1, pp. 88–91). Elsevier BV. https://doi.org/10.1016/0005-2728(73)90253-3

121. Macqueen, D. J., Kristjánsson, B. K., & Johnston, I. A. (2010). Salmonid genomes have a remarkably expanded *akirin* family, coexpressed with genes from conserved pathways governing skeletal muscle growth and catabolism. *Physiological Genomics*, *42*(1), 134–148. https://doi.org/10.1152/physiolgenomics.00045.2010

122. Maehr, T., Costa, M. M., González Vecino, J. L., Wadsworth, S., Martin, S. A. M., Wang, T., & Secombes, C. J. (2013). Transforming growth factor-β1b: A second TGF-β1 paralogue in the rainbow trout (Oncorhynchus mykiss) that has a lower constitutive expression but is more responsive to immune stimulation. *Fish & Shellfish Immunology*, *34*(2), 420–432. https://doi.org/10.1016/j.fsi.2012.11.011

123. Maere, S., De Bodt, S., Raes, J., Casneuf, T., Van Montagu, M., Kuiper, M., & Van De Peer, Y. (2005). Modeling gene and genome duplications in eukaryotes. *Proceedings of the National Academy of Sciences*, *102*(15), 5454–5459. https://doi.org/10.1073/pnas.0501102102

124. Marlétaz, F., Timoshevskaya, N., Timoshevskiy, V., Simakov, O., Parey, E., Gavriouchkina, D., Suzuki, M., Kubokawa, K., Brenner, S., Smith, J., & Rokhsar, D. S. (2023). *The hagfish genome and the evolution of vertebrates* [Preprint]. Genomics. https://doi.org/10.1101/2023.04.17.537254

125. Martins, E. L., Ricardo, J. C., de-Souza-Ferreira, E., Camacho-Pereira, J., Ramos-Filho, D., & Galina, A. (2018). Rapid regulation of substrate use for oxidative phosphorylation during a single session of high intensity interval or aerobic exercises in different rat skeletal muscles. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, *217*, 40–50. https://doi.org/10.1016/j.cbpb.2017.11.013

126. Martos-Sitcha, J. A., Bermejo-Nogales, A., Calduch-Giner, J. A., & Pérez-Sánchez, J. (2017). Gene expression profiling of whole blood cells supports a more efficient mitochondrial respiration in hypoxia-challenged gilthead sea bream (Sparus aurata). *Frontiers in Zoology*, *14*(1), 34. https://doi.org/10.1186/s12983-017-0220-2

127. Martos-Sitcha, J. A., Mancera, J. M., Calduch-Giner, J. A., Yúfera, M., Martínez-Rodríguez, G., & Pérez-Sánchez, J. (2016). Unraveling the Tissue-Specific Gene Signatures of Gilthead Sea Bream (Sparus aurata L.) after Hyper- and Hypo-Osmotic Challenges. *PLOS ONE*, *11*(2), e0148113. https://doi.org/10.1371/journal.pone.0148113

128. Mateus, A. P., Costa, R. A., Cardoso, J. C. R., Andree, K. B., Estévez, A., Gisbert, E., & Power, D. M. (2017). Thermal imprinting modifies adult stress and innate immune responsiveness in the teleost sea bream. *Journal of Endocrinology*, *233*(3), 381–394. https://doi.org/10.1530/JOE-16-0610

129. Mazat, J.-P., Devin, A., & Ransac, S. (2020). Modelling mitochondrial ROS production by the respiratory chain. *Cellular and Molecular Life Sciences*, *77*(3), 455–465. https://doi.org/10.1007/s00018-019-03381-1

130. Merino, O., Risopatrón, J., Valdebenito, I., Figueroa, E., & Farías, J. G. (2023). Effect of the temperature of activation medium on fish sperm quality: Impact on fertilization *in vitro* in aquaculture practice. *Reviews in Aquaculture*, *15*(2), 434–451. https://doi.org/10.1111/raq.12729

131. Meyer, A., & Schartl, M. (1999). Gene and genome duplications in vertebrates: The one-to-four (-to-eight in fish) rule and the evolution of novel gene functions. *Current Opinion in Cell Biology*, *11*(6), 699–704. https://doi.org/10.1016/S0955-0674(99)00039-3

132. Meyer, A., & Van De Peer, Y. (2005). From 2R to 3R: Evidence for a fish-specific genome duplication (FSGD). *BioEssays*, *27*(9), 937–945. https://doi.org/10.1002/bies.20293

133. Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L. L., Tosatto, S. C. E., Paladin, L., Raj, S., Richardson, L. J., Finn, R. D., & Bateman, A. (2021). Pfam: The protein families database in 2021. *Nucleic Acids Research*, *49*(D1), D412–D419. https://doi.org/10.1093/nar/gkaa913

134. Mitchell, P. (1961). Coupling of Phosphorylation to Electron and Hydrogen Transfer by a Chemi-Osmotic type of Mechanism. *Nature*, *191*(4784), 144–148. https://doi.org/10.1038/191144a0

135. Mitra, A., Abdel-Gawad, F. Kh., Bassem, S., Barua, P., Assisi, L., Parisi, C., Temraz, T. A., Vangone, R., Kajbaf, K., Kumar, V., & Guerriero, G. (2023). Climate Change and Reproductive Biocomplexity in Fishes: Innovative Management Approaches towards Sustainability of Fisheries and Aquaculture. *Water*, *15*(4), 725. https://doi.org/10.3390/w15040725

136. Moghadam, H. K., Ferguson, M. M., & Danzmann, R. G. (2005). Evidence for Hox Gene Duplication in Rainbow Trout (Oncorhynchus mykiss): A Tetraploid Model Species. *Journal of Molecular Evolution*, *61*(6), 804–818. https://doi.org/10.1007/s00239-004-0230-5

137. Moore, J. K., & Haber, J. E. (1996). Cell Cycle and Genetic Requirements of Two Pathways of Nonhomologous End-Joining Repair of Double-Strand Breaks in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology*, *16*(5), 2164–2173. https://doi.org/10.1128/MCB.16.5.2164

138. Moore, R. C., & Purugganan, M. D. (2003). The early stages of duplicate gene evolution. *Proceedings of the National Academy of Sciences*, *100*(26), 15682–15687. https://doi.org/10.1073/pnas.2535513100

139. Mu, W. J., Wen, H. S., Shi, D., & Yang, Y. P. (2013). Molecular cloning and expression analysis of estrogen receptor betas (ERβ1 and ERβ2) during gonad development in the Korean rockfish, Sebastes schlegeli. *Gene*, *523*(1), 39–49. https://doi.org/10.1016/j.gene.2013.03.109

140. Mwaluma, J Strydom, N. (2014). *A Guide to Commonly Occurring Larval Stages of Fishes in Kenyan Coastal Waters*. Western Indian Ocean Marine Science Association: Zanzibar.

141. Nakatani, Y., Shingate, P., Ravi, V., Pillai, N. E., Prasad, A., McLysaght, A., & Venkatesh, B. (2021). Reconstruction of proto-vertebrate, proto-cyclostome and proto-gnathostome genomes provides new insights into early vertebrate evolution. *Nature Communications*, *12*(1), 4489. https://doi.org/10.1038/s41467-021-24573-z

142. Nakatani, Y., Takeda, H., Kohara, Y., & Morishita, S. (2007). Reconstruction of the vertebrate ancestral genome reveals dynamic genome reorganization in early vertebrates. *Genome Research*, *17*(9), 1254–1265. https://doi.org/10.1101/gr.6316407

143. Nei, M., Rogozin, I. B., & Piontkivska, H. (2000). Purifying selection and birth-and-death evolution in the ubiquitin gene family. *Proceedings of the National Academy of Sciences*, *97*(20), 10866–10871. https://doi.org/10.1073/pnas.97.20.10866

144. Neupane, P., Bhuju, S., Thapa, N., & Bhattarai, H. K. (2019). ATP Synthase: Structure, Function and Inhibition. *Biomolecular Concepts*, *10*(1), 1–10. https://doi.org/10.1515/bmc-2019-0001

145. Nielsen, M. G., Gadagkar, S. R., & Gutzwiller, L. (2010). *RTeusebarcuh alritnicleevolution in insects: Gene duplication and subfunctionalization provide specialized isoforms in a functionally constrained gene family*.

146. Noji, H., Yasuda, R., Yoshida, M., & Kinosita, K. (1997). Direct observation of the rotation of F1-ATPase. *Nature*, *386*(6622), 299–302. https://doi.org/10.1038/386299a0

147. Nolfi-Donegan, D., Braganza, A., & Shiva, S. (2020). Mitochondrial electron transport chain: Oxidative phosphorylation, oxidant production, and methods of measurement. *Redox Biology*, *37*, 101674. https://doi.org/10.1016/j.redox.2020.101674

148. Nowak, M. A., Boerlijst, M. C., Cooke, J., & Smith, J. M. (1997). Evolution of genetic redundancy. *Nature*, *388*(6638), 167–171. https://doi.org/10.1038/40618

149. Ohno, S. (1970). *Evolution by Gene Duplication*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-86659-3

150. Ohno, S., Muramoto, J., Christian, L., & Atkin, N. B. (1967). Diploid-tetraploid relationship among old-world members of the fish family Cyprinidae. *Chromosoma*, *23*(1), 1–9. https://doi.org/10.1007/BF00293307

151. Panopoulou, G., & Poustka, A. (2005). Timing and mechanism of ancient vertebrate genome duplications – the adventure of a hypothesis. *Trends in Genetics*, *21*(10), 559–567. https://doi.org/10.1016/j.tig.2005.08.004

152. Papa, S., Martino, P. L., Capitanio, G., Gaballo, A., De Rasmo, D., Signorile, A., & Petruzzella, V. (2012). The Oxidative Phosphorylation System in Mammalian Mitochondria. In R. Scatena, P. Bottoni, & B. Giardina (Eds.), *Advances in Mitochondrial Medicine* (Vol. 942, pp. 3–37). Springer Netherlands. https://doi.org/10.1007/978-94-007-2869-1_1

153. Papandroulakis, N., Divanach, P., Anastasiadis, P., & Kentouri, M. (2001b). The pseudo-green water technique for intensive rearing of sea bream (Sparus aurata) larvae. *Aquaculture International*, *9*(3), 205–216. https://doi.org/10.1023/A:1016813623122

154. Papp, B., Pál, C., & Hurst, L. D. (2003). Dosage sensitivity and the evolution of gene families in yeast. *Nature*, *424*(6945), 194–197. https://doi.org/10.1038/nature01771

155. Paterson, A., Chapman, B., Kissinger, J., Bowers, J., Feltus, F., & Estill, J. (2006). Many gene and domain families have convergent fates following independent whole-genome duplication events in Arabidopsis, Oryza, Saccharomyces and Tetraodon. *Trends in Genetics*, *22*(11), 597–602. https://doi.org/10.1016/j.tig.2006.09.003

156. Paumard, P., Vaillier, J., Coulary, B., Schaeffer, J., Soubannier, V., Mueller, D. M., Brèthes, D., Di Rago, J.-P., & Velours, J. (2002). The ATP synthase is involved in generating mitochondrial

cristae morphology. *The EMBO Journal*, *21*(3), 221–230. https://doi.org/10.1093/emboj/21.3.221

157. Pavlidis, M., Karantzali, E., Fanouraki, E., Barsakis, C., Kollias, S., & Papandroulakis, N. (2011). Onset of the primary stress in European sea bass Dicentrarhus labrax, as indicated by whole body cortisol in relation to glucocorticoid receptor during early development. *Aquaculture*, *315*(1–2), 125–130. https://doi.org/10.1016/j.aquaculture.2010.09.013

158. Pavlidis, M., Koumoundouros, G., Sterioti, A., Somarakis, S., Divanach, P., & Kentouri, M. (2000). Evidence of temperature-dependent sex determination in the European sea bass (Dicentrarchus labrax L.). *Journal of Experimental Zoology*, *287*(3), 225–232. https://doi.org/10.1002/1097-010X(20000801)287:3<225::AID-JEZ4>3.0.CO;2-D

159. Paysan-Lafosse, T., Blum, M., Chuguransky, S., Grego, T., Pinto, B. L., Salazar, G. A., Bileschi, M. L., Bork, P., Bridge, A., Colwell, L., Gough, J., Haft, D. H., Letunić, I., Marchler-Bauer, A., Mi, H., Natale, D. A., Orengo, C. A., Pandurangan, A. P., Rivoire, C., … Bateman, A. (2023). InterPro in 2022. *Nucleic Acids Research*, *51*(D1), D418–D427. https://doi.org/10.1093/nar/gkac993

160. Pelster, B., & Burggren, W. W. (1996). Disruption of Hemoglobin Oxygen Transport Does Not Impact Oxygen-Dependent Physiological Processes in Developing Embryos of Zebra Fish (*Danio rerio*). *Circulation Research*, *79*(2), 358–362. https://doi.org/10.1161/01.RES.79.2.358

161. Perelló-Amorós, M., Fernández-Borràs, J., Sánchez-Moya, A., Vélez, E. J., García-Pérez, I., Gutiérrez, J., & Blasco, J. (2021). Mitochondrial Adaptation to Diet and Swimming Activity in Gilthead Seabream: Improved Nutritional Efficiency. *Frontiers in Physiology*, *12*, 678985. https://doi.org/10.3389/fphys.2021.678985

162. Picolo, F., Grandchamp, A., Piégu, B., Rolland, A. D., Veitia, R. A., & Monget, P. (2021). Genes Encoding Teleost Orthologs of Human Haploinsufficient and Monoallelically Expressed Genes Remain in Duplicate More Frequently Than the Whole Genome. *International Journal of Genomics*, *2021*, 1–6. https://doi.org/10.1155/2021/9028667

163. Power, D. M., Einarsdóttir, I. E., Pittman, K., Sweeney, G. E., Hildahl, J., Campinho, M. A., Silva, N., Sæle, Ø., Galay-Burgos, M., Småradóttir, H., & Björnsson, B. T. (2008). The Molecular and Endocrine Basis of Flatfish Metamorphosis. *Reviews in Fisheries Science*, *16*(sup1), 95–111. https://doi.org/10.1080/10641260802325377

164. Putnam, N. H., Butts, T., Ferrier, D. E. K., Furlong, R. F., Hellsten, U., Kawashima, T., Robinson-Rechavi, M., Shoguchi, E., Terry, A., Yu, J.-K., Benito-Gutiérrez, E., Dubchak, I., Garcia-Fernàndez, J., Gibson-Brown, J. J., Grigoriev, I. V., Horton, A. C., De Jong, P. J., Jurka, J., Kapitonov, V. V., … Rokhsar, D. S. (2008). The amphioxus genome and the evolution of the chordate karyotype. *Nature*, *453*(7198), 1064–1071. https://doi.org/10.1038/nature06967

165. R Core Team (2022) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.

166. Rasmussen, B., Fletcher, I. R., Brocks, J. J., & Kilburn, M. R. (2008). Reassessing the first appearance of eukaryotes and cyanobacteria. *Nature*, *455*(7216), 1101–1104. https://doi.org/10.1038/nature07381

167. Rastogi, S., & Liberles, D. A. (2005). Subfunctionalization of duplicated genes as a transition state to neofunctionalization. *BMC Evolutionary Biology*, *5*(1), 28. https://doi.org/10.1186/1471-2148-5-28

168. Ray, P. D., Huang, B.-W., & Tsuji, Y. (2012). Reactive oxygen species (ROS) homeostasis and redox regulation in cellular signaling. *Cellular Signalling*, *24*(5), 981–990. https://doi.org/10.1016/j.cellsig.2012.01.008

169. Read, A. D., Bentley, R. Et., Archer, S. L., & Dunham-Snary, K. J. (2021). Mitochondrial iron–sulfur clusters: Structure, function, and an emerging role in vascular biology. *Redox Biology*, *47*, 102164. https://doi.org/10.1016/j.redox.2021.102164

170. Reyes Corral, W. D., & Aguirre, W. E. (2019). Effects of temperature and water turbulence on vertebral number and body shape in Astyanax mexicanus (Teleostei: Characidae). *PLOS ONE*, *14*(7), e0219677. https://doi.org/10.1371/journal.pone.0219677

171. Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, *26*(1), 139–140. https://doi.org/10.1093/bioinformatics/btp616

172. Rocha, F., Dias, J., Engrola, S., Gavaia, P., Geurden, I., Dinis, M. T., & Panserat, S. (2015). Glucose metabolism and gene expression in juvenile zebrafish (*Danio rerio*) challenged with a high carbohydrate diet: Effects of an acute glucose stimulus during late embryonic life. *British Journal of Nutrition*, *113*(3), 403–413. https://doi.org/10.1017/S0007114514003869

173. Rodgers-Melnick, E., Mane, S. P., Dharmawardhana, P., Slavov, G. T., Crasta, O. R., Strauss, S. H., Brunner, A. M., & DiFazio, S. P. (2012). Contrasting patterns of evolution following whole genome versus tandem duplication events in *Populus*. *Genome Research*, *22*(1), 95–105. https://doi.org/10.1101/gr.125146.111

174. Rombough, P. J. (1988). Growth, aerobic metabolism, and dissolved oxygen requirements of embryos and alevins of steelhead, *Salmo gairdneri*. *Canadian Journal of Zoology*, *66*(3), 651–660. https://doi.org/10.1139/z88-097

175. RStudio Team (2022). RStudio: Integrated Development for R. RStudio. PBC.

176. Russell, A. P., Hesselink, M. K. C., Lo, S. K., & Schrauwen, P. (2005). Regulation of metabolic transcriptional co-activators and transcription factors with acute exercise. *The FASEB Journal*, *19*(8), 986–988. https://doi.org/10.1096/fj.04-3168fje

177. Rutter, J., Winge, D. R., & Schiffman, J. D. (2010). Succinate dehydrogenase – Assembly, regulation and role in human disease. *Mitochondrion*, *10*(4), 393–401. https://doi.org/10.1016/j.mito.2010.03.001

178. Sæle, Ø., Solbakken, J. S., Watanabe, K., Hamre, K., Power, D., & Pittman, K. (2004). Staging of Atlantic halibut (Hippoglossus hippoglossus L.) from first feeding through metamorphosis, including cranial ossification independent of eye migration. *Aquaculture*, *239*(1–4), 445–465. https://doi.org/10.1016/j.aquaculture.2004.05.025

179. Saitoh, T., Igura, M., Obita, T., Ose, T., Kojima, R., Maenaka, K., Endo, T., & Kohda, D. (2007). Tom20 recognizes mitochondrial presequences through dynamic equilibrium among multiple bound states. *The EMBO Journal*, *26*(22), 4777–4787. https://doi.org/10.1038/sj.emboj.7601888

180. Sato, Y., & Nishida, M. (2010). Teleost fish with specific genome duplication as unique models of vertebrate evolution. *Environmental Biology of Fishes*, *88*(2), 169–188. https://doi.org/10.1007/s10641-010-9628-7

181. Scapolo, P. A., Veggetti, A., Mascarello, F., & Romanello, M. G. (1988). Developmental transitions of myosin isoforms and organisation of the lateral muscle in the teleostDicentrarchus labrax (L.). *Anatomy and Embryology*, *178*(4), 287–295. https://doi.org/10.1007/BF00698660

182. Schägger, H., & Pfeiffer, K. (2001). The Ratio of Oxidative Phosphorylation Complexes I–V in Bovine Heart Mitochondria and the Composition of Respiratory Chain Supercomplexes. *Journal of Biological Chemistry*, *276*(41), 37861–37867. https://doi.org/10.1074/jbc.M106474200

183. Schnurr, M. E., Yin, Y., & Scott, G. R. (2013). Temperature during embryonic development has persistent effects on metabolic enzymes in the muscle of zebrafish. *Journal of Experimental Biology*, jeb.094037. https://doi.org/10.1242/jeb.094037

184. Scott, G. R., & Johnston, I. A. (2012). Temperature during embryonic development has persistent effects on thermal acclimation capacity in zebrafish. *Proceedings of the National Academy of Sciences*, *109*(35), 14247–14252. https://doi.org/10.1073/pnas.1205012109

185. Seal, R. L., Braschi, B., Gray, K., Jones, T. E. M., Tweedie, S., Haim-Vilmovsky, L., & Bruford, E. A. (2023). Genenames.org: The HGNC resources in 2023. *Nucleic Acids Research*, *51*(D1), D1003–D1009. https://doi.org/10.1093/nar/gkac888

186. Sémon, M., & Wolfe, K. H. (2007). Consequences of genome duplication. *Current Opinion in Genetics & Development*, *17*(6), 505–512. https://doi.org/10.1016/j.gde.2007.09.007

187. Signes, A., & Fernandez-Vizarra, E. (2018). Assembly of mammalian oxidative phosphorylation complexes I–V and supercomplexes. *Essays in Biochemistry*, *62*(3), 255–270. https://doi.org/10.1042/EBC20170098

188. Silva-Marrero, J. I., Sáez, A., Caballero-Solares, A., Viegas, I., Almajano, M. P., Fernández, F., Baanante, I. V., & Metón, I. (2017). A transcriptomic approach to study the effect of long-term starvation and diet composition on the expression of mitochondrial oxidative phosphorylation genes in gilthead sea bream (Sparus aurata). *BMC Genomics*, *18*(1), 768. https://doi.org/10.1186/s12864-017-4148-x

189. Song, M. J., Potter, B. I., Doyle, J. J., & Coate, J. E. (2020). Gene Balance Predicts Transcriptional Responses Immediately Following Ploidy Change in *Arabidopsis thaliana*. *The Plant Cell*, *32*(5), 1434–1448. https://doi.org/10.1105/tpc.19.00832

190. Soskine, M., & Tawfik, D. S. (2010). Mutational effects and the evolution of new protein functions. *Nature Reviews Genetics*, *11*(8), 572–582. https://doi.org/10.1038/nrg2808

191. Sperling, E. A., Knoll, A. H., & Girguis, P. R. (2015). The Ecological Physiology of Earth's Second Oxygen Revolution. *Annual Review of Ecology, Evolution, and Systematics*, *46*(1), 215–235. https://doi.org/10.1146/annurev-ecolsys-110512-135808

192. Stephens, S. G. (1951). Possible Significance of Duplication in Evolution. In *Advances in Genetics* (Vol. 4, pp. 247–265). Elsevier. https://doi.org/10.1016/S0065-2660(08)60237-0

193. Steri, M., Idda, M. L., Whalen, M. B., & Orrù, V. (2018). Genetic variants in mRNA untranslated regions. *WIREs RNA*, *9*(4), e1474. https://doi.org/10.1002/wrna.1474

194. Su, M., Liu, N., Zhang, Z., & Zhang, J. (2022). Osmoregulatory strategies of estuarine fish Scatophagus argus in response to environmental salinity changes. *BMC Genomics*, *23*(1), 545. https://doi.org/10.1186/s12864-022-08784-2

195. Tang, J. X., Thompson, K., Taylor, R. W., & Oláhová, M. (2020). Mitochondrial OXPHOS Biogenesis: Co-Regulation of Protein Synthesis, Import, and Assembly Pathways. *International Journal of Molecular Sciences*, *21*(11), 3820. https://doi.org/10.3390/ijms21113820

196. Teixeira, F. K., Sanchez, C. G., Hurd, T. R., Seifert, J. R. K., Czech, B., Preall, J. B., Hannon, G. J., & Lehmann, R. (2015). ATP synthase promotes germ cell differentiation independent of oxidative phosphorylation. *Nature Cell Biology*, *17*(5), 689–696. https://doi.org/10.1038/ncb3165

197. Teshima, K. M., & Innan, H. (2008). Neofunctionalization of Duplicated Genes Under the Pressure of Gene Conversion. *Genetics*, *178*(3), 1385–1398. https://doi.org/10.1534/genetics.107.082933

198. Thumuluri, V., Almagro Armenteros, J. J., Johansen, A. R., Nielsen, H., & Winther, O. (2022). DeepLoc 2.0: Multi-label subcellular localization prediction using protein language models. *Nucleic Acids Research*, *50*(W1), W228–W234. https://doi.org/10.1093/nar/gkac278

199. Tsalafouta, A., Papandroulakis, N., Gorissen, M., Katharios, P., Flik, G., & Pavlidis, M. (2014). Ontogenesis of the HPI axis and molecular regulation of the cortisol stress response during early development in Dicentrarchus labrax. In Scientific Reports (Vol. 4, Issue 1). Springer Science and Business Media LLC. https://doi.org/10.1038/srep05525

200. Tsukihara, T., Aoyama, H., Yamashita, E., Tomizaki, T., Yamaguchi, H., Shinzawa-Itoh, K., Nakashima, R., Yaono, R., & Yoshikawa, S. (1996). The Whole Structure of the 13-Subunit Oxidized Cytochrome c Oxidase at 2.8 Å. *Science*, *272*(5265), 1136–1144. https://doi.org/10.1126/science.272.5265.1136

201. Turan, C. (2004). Stock identification of Mediterranean horse mackerel (Trachurus mediterraneus) using morphometric and meristic characters. *ICES Journal of Marine Science*, *61*(5), 774–781. https://doi.org/10.1016/j.icesjms.2004.05.001

202. Tweedie, S., Braschi, B., Gray, K., Jones, T. E. M., Seal, R. L., Yates, B., & Bruford, E. A. (2021). Genenames.org: The HGNC and VGNC resources in 2021. *Nucleic Acids Research*, *49*(D1), D939–D946. https://doi.org/10.1093/nar/gkaa980

203. Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., & Rozen, S. G. (2012). Primer3—New capabilities and interfaces. *Nucleic Acids Research*, *40*(15), e115–e115. https://doi.org/10.1093/nar/gks596

204. Vagner, M., Zambonino-Infante, J.-L., & Mazurais, D. (2019). Fish facing global change: Are early stages the lifeline? *Marine Environmental Research*, *147*, 159–178. https://doi.org/10.1016/j.marenvres.2019.04.005

205. Van De Peer, Y., Ashman, T.-L., Soltis, P. S., & Soltis, D. E. (2021). Polyploidy: An evolutionary and ecological force in stressful times. *The Plant Cell*, *33*(1), 11–26. https://doi.org/10.1093/plcell/koaa015

206. Van De Peer, Y., Maere, S., & Meyer, A. (2009). The evolutionary significance of ancient genome duplications. *Nature Reviews Genetics*, *10*(10), 725–732. https://doi.org/10.1038/nrg2600

207. Vandepoele, K., De Vos, W., Taylor, J. S., Meyer, A., & Van De Peer, Y. (2004). Major events in the genome evolution of vertebrates: Paranome age and size differ considerably between

ray-finned fishes and land vertebrates. *Proceedings of the National Academy of Sciences*, *101*(6), 1638–1643. https://doi.org/10.1073/pnas.0307968100

208. Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., & Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes *Genome Biology*, *3*(7), research0034.1. https://doi.org/10.1186/gb-2002-3-7-research0034

209. Veggetti, A., Mascarello, F., Scapolo, P. A., & Rowlerson, A. (1990). Hyperplastic and hypertrophic growth of lateral muscle in Dicentrarchus labrax (L.): An ultrastructural and morphometric study. *Anatomy and Embryology*, *182*(1). https://doi.org/10.1007/BF00187522

210. Verdugo, R. A. (2017). chromPlot. Bioconductor. https://doi.org/10.18129/B9.BIOC.CHROMPLOT

211. Vieira, V. I. A., & Johnston, I. A. (1992). Influence of temperature on muscle-fibre development in larvae of the herring Clupea harengus. *Marine Biology*, *112*(2), 333–341. https://doi.org/10.1007/BF00702480

212. Vinckenbosch, N., Dupanloup, I., & Kaessmann, H. (2006). Evolutionary fate of retroposed gene copies in the human genome. *Proceedings of the National Academy of Sciences*, *103*(9), 3220–3225. https://doi.org/10.1073/pnas.0511307103

213. Voldoire, E., Brunet, F., Naville, M., Volff, J.-N., & Galiana, D. (2017). Expansion by whole genome duplication and evolution of the sox gene family in teleost fish. *PLOS ONE*, *12*(7), e0180936. https://doi.org/10.1371/journal.pone.0180936

214. Wang, T., Hu, Y., Wangkahart, E., Liu, F., Wang, A., Zahran, E., Maisey, K. R., Liu, M., Xu, Q., Imarai, M., & Secombes, C. J. (2018). Interleukin (IL)-2 Is a Key Regulator of T Helper 1 and T Helper 2 Cytokine Expression in Fish: Functional Characterization of Two Divergent IL2 Paralogs in Salmonids. *Frontiers in Immunology*, *9*, 1683. https://doi.org/10.3389/fimmu.2018.01683

215. Wang, Y., Tang, H., DeBarry, J. D., Tan, X., Li, J., Wang, X., Lee, T. -h., Jin, H., Marler, B., Guo, H., Kissinger, J. C., & Paterson, A. H. (2012). MCScanX: A toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Research*, *40*(7), e49–e49. https://doi.org/10.1093/nar/gkr1293

216. Wiedemann, N., & Pfanner, N. (2017). Mitochondrial Machineries for Protein Import and Assembly. *Annual Review of Biochemistry*, *86*(1), 685–714. https://doi.org/10.1146/annurev-biochem-060815-014352

217. Wikström, M., Krab, K., & Sharma, V. (2018). Oxygen Activation and Energy Conservation by Cytochrome *c* Oxidase. *Chemical Reviews*, *118*(5), 2469–2490. https://doi.org/10.1021/acs.chemrev.7b00664

218. Wikström, M., Sharma, V., Kaila, V. R. I., Hosler, J. P., & Hummer, G. (2015). New Perspectives on Proton Pumping in Cellular Respiration. *Chemical Reviews*, *115*(5), 2196–2221. https://doi.org/10.1021/cr500448t

219. Wilson, A. E., & Liberles, D. A. (2023). Dosage balance acts as a time-dependent selective barrier to subfunctionalization. *BMC Ecology and Evolution*, *23*(1), 14. https://doi.org/10.1186/s12862-023-02116-y

220. Wilson, D. F. (2017). Oxidative phosphorylation: Regulation and role in cellular and tissue metabolism. *The Journal of Physiology*, *595*(23), 7023–7038. https://doi.org/10.1113/JP273839

221. Wirth, C., Brandt, U., Hunte, C., & Zickermann, V. (2016). Structure and function of mitochondrial complex I. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, *1857*(7), 902–914. https://doi.org/10.1016/j.bbabio.2016.02.013

222. Wittbrodt, J., Meyer, A., & Schartl, M. (1998). More genes in fish? *BioEssays*, *20*(6), 511–515. https://doi.org/10.1002/(SICI)1521-1878(199806)20:6<511::AID-BIES10>3.0.CO;2-3

223. Wolfe, K. H., & Shields, D. C. (1997). Molecular evidence for an ancient duplication of the entire yeast genome. *Nature*, *387*(6634), 708–713. https://doi.org/10.1038/42711

224. Xia, D., Yu, C.-A., Kim, H., Xia, J.-Z., Kachurin, A. M., Zhang, L., Yu, L., & Deisenhofer, J. (1997). Crystal Structure of the Cytochrome bc $_1$ Complex from Bovine Heart Mitochondria. *Science*, *277*(5322), 60–66. https://doi.org/10.1126/science.277.5322.60

225. Xu, Y., Xue, D., Bankhead, A., & Neamati, N. (2020). Why All the Fuss about Oxidative Phosphorylation (OXPHOS)? *Journal of Medicinal Chemistry*, *63*(23), 14276–14307. https://doi.org/10.1021/acs.jmedchem.0c01013

226. Yoshida, M., Muneyuki, E., & Hisabori, T. (2001). ATP synthase—A marvellous rotary engine of the cell. *Nature Reviews Molecular Cell Biology*, *2*(9), 669–677. https://doi.org/10.1038/35089509

227. Yu, D., Ren, Y., Uesaka, M., Beavan, A. J. S., Muffato, M., Shen, J., Li, Y., Sato, I., Wan, W., Clark, J. W., Keating, J. N., Carlisle, E. M., Dearden, R. P., Giles, S., Randle, E., Sansom, R. S., Feuda, R., Fleming, J. F., Sugahara, F., … Pascual-Anaya, J. (2023). *Hagfish genome illuminates vertebrate whole genome duplications and their evolutionary consequences* [Preprint]. Genomics. https://doi.org/10.1101/2023.04.08.536076

228. Zambonino-Infante, J. L., Claireaux, G., Ernande, B., Jolivet, A., Quazuguel, P., Sévère, A., Huelvan, C., & Mazurais, D. (2013). Hypoxia tolerance of common sole juveniles depends on dietary regime and temperature at the larval stage: Evidence for environmental conditioning. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1758), 20123022. https://doi.org/10.1098/rspb.2012.3022

229. Zanuy, S., Carrillo, M., Felip, A., Rodrıguez, L., & Blazquez, M. (2001). *Genetic, hormonal and environmental approaches for the control of reproduction in the European sea bass žDicentrarchus labrax L./*.

230. Zhang, F., & Broughton, R. E. (2013). Mitochondrial–Nuclear Interactions: Compensatory Evolution or Variable Functional Constraint among Vertebrate Oxidative Phosphorylation Genes? *Genome Biology and Evolution*, *5*(10), 1781–1791. https://doi.org/10.1093/gbe/evt129

231. Zhang, J. (2003). Evolution by gene duplication: An update. *Trends in Ecology & Evolution*, *18*(6), 292–298. https://doi.org/10.1016/S0169-5347(03)00033-8

232. Zhang, Y., Gilbert, M. J. H., & Farrell, A. P. (2019). Finding the peak of dynamic oxygen uptake during fatiguing exercise in fish. *Journal of Experimental Biology*, jeb.196568. https://doi.org/10.1242/jeb.196568

233. Zhao, R., Jiang, S., Zhang, L., & Yu, Z. (2019). Mitochondrial electron transport chain, ROS generation and uncoupling (Review). *International Journal of Molecular Medicine*. https://doi.org/10.3892/ijmm.2019.4188

234. Zickermann, V., Kerscher, S., Zwicker, K., Tocilescu, M. A., Radermacher, M., & Brandt, U. (2009). Architecture of complex I and its implications for electron transfer and proton pumping. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, *1787*(6), 574–583. https://doi.org/10.1016/j.bbabio.2009.01.012

235. Zong, S., Wu, M., Gu, J., Liu, T., Guo, R., & Yang, M. (2018). Structure of the intact 14-subunit human cytochrome c oxidase. *Cell Research*, *28*(10), 1026–1034. https://doi.org/10.1038/s41422-018-0071-1

Chapter 8

Supplementary Material

Chapter 2

Table S2.1 Gene IDs and assigned names of OXPHOS genes of gilthead seabream and European seabass

Sparus aurata				*Dicentrarchus labrax*			
Gene ID	Gene name	Gene ID	Gene name	Gene ID	Gene name	Gene ID	Gene name
ENSSAUG00010001552	*atp5f1a.1*	ENSSAUG00010018219	*ndufa1*	ENSDLAG00005007330	*atp5f1a.1*	ENSDLAG00005015965	*ndufa1*
ENSSAUG00010008214	*atp5f1a.2*	ENSSAUG00010027521	*ndufa10*	ENSDLAG00005018532	*atp5f1a.2*	ENSDLAG00005008059	*ndufa10*
ENSSAUG00010007970	*atp5f1b.1*	ENSSAUG00010023201	*ndufa11*	ENSDLAG00005018338	*atp5f1b.1*	ENSDLAG00005021448	*ndufa11*
ENSSAUG00010026506	*atp5f1b.2*	ENSSAUG00010018631	*ndufa12*	ENSDLAG00005025099	*atp5f1b.2*	ENSDLAG00005024691	*ndufa12*
ENSSAUG00010022681	*atp5f1c*	ENSSAUG00010017787	*ndufa13*	ENSDLAG00005025077	*atp5f1c*	ENSDLAG00005024977	*ndufa13*
ENSSAUG00010005650	*atp5f1d*	ENSSAUG00010002150	*ndufa2*	ENSDLAG00005013665	*atp5f1d*	ENSDARG00000021984	*ndufa2*
ENSSAUG00010011548	*atp5f1e*	ENSSAUG00010020916	*ndufa3*	ENSDLAG00005016988	*atp5f1e*	ENSDLAG00005002836	*ndufa3*
ENSSAUG00010020382	*atp5if1.1*	ENSSAUG00010010015	*ndufa5*	ENSDLAG00005004647	*atp5if1.1*	ENSDLAG00005008847	*ndufa5*
ENSSAUG00010003151	*atp5if1.2*	ENSSAUG00010007020	*ndufa6*	ENSDLAG00005000620	*atp5if1.2*	ENSDLAG00005015610	*ndufa6*
ENSSAUG00010024766	*atp5mc.1*	ENSSAUG00010007465	*ndufa7*	ENSDLAG00005029965	*atp5mc.1*	ENSDLAG00005025335	*ndufa7*
ENSSAUG00010001003	*atp5mc.2*	ENSSAUG00010022401	*ndufa8*	ENSDLAG00005027449	*atp5mc.2*	ENSDLAG00005022483	*ndufa8*
ENSSAUG00010003390	*atp5mc.3*	ENSSAUG00010008168	*ndufa9*	ENSDLAG00005033252	*atp5mc.3*	ENSDLAG00005015966	*ndufa9*
ENSSAUG00010022955	*atp5me.1*	ENSSAUG00010021118	*ndufab1*	ENSDLAG00005030491	*atp5me*	ENSDLAG00005001413	*ndufab1*
ENSSAUG00010005603	*atp5me.2*	ENSSAUG00010022839	*ndufb1*	ENSDLAG00005007485	*atp5mf*	ENSDLAG00005001413	*ndufb1*
ENSSAUG00010015404	*atp5mf*	ENSSAUG00010023733	*ndufb10*	ENSDLAG00005019335	*atp5mg.1*	ENSDLAG00005011028	*ndufb10*
ENSSAUG00010011394	*atp5mg.1*	ENSSAUG00010012559	*ndufb11*	ENSDLAG00005006347	*atp5mg.2a*	ENSDLAG00005019348	*ndufb11*
ENSSAUG00010004998	*atp5mg.2*	ENSSAUG00010013498	*ndufb2*	ENSDLAG00005006335	*atp5mg.2b*	ENSDLAG00005019817	*ndufb2*
ENSSAUG00010012064	*atp5mj*	ENSSAUG00010025848	*ndufb3*	ENSDLAG00005024092	*atp5mk*	ENSDLAG00005024263	*ndufb3*
ENSSAUG00010005272	*atp5mk*	ENSSAUG00010015751	*ndufb4*	ENSDLAG00005021874	*atp5pb*	ENSDLAG00005018458	*ndufb4*
ENSSAUG00010001248	*atp5pb*	ENSSAUG00010027131	*ndufb5*	ENSDLAG00005003781	*atp5pd.1*	ENSDLAG00005026097	*ndufb5*
ENSSAUG00010003235	*atp5pd.1*	ENSSAUG00010018717	*ndufb6*	ENSDLAG00005021473	*atp5pd.2*	ENSDLAG00005008294	*ndufb6*
ENSSAUG00010027145	*atp5pd.2*	ENSSAUG00010018909	*ndufb7*	ENSDLAG00005015257	*atp5pf.1*	ENSDLAG00005020963	*ndufb7*
ENSSAUG00010017260	*atp5pf.1*	ENSSAUG00010023063	*ndufb8*	ENSDLAG00005000158	*atp5pf.2*	ENSDLAG00005022803	*ndufb8*
ENSSAUG00010001316	*atp5pf.2*	ENSSAUG00010000482	*ndufb9*	ENSDLAG00005015632	*atp5pf.3*	ENSDLAG00005022603	*ndufb9*
ENSSAUG00010012045	*atp5pf.3*	ENSSAUG00010012757	*ndufc1*	ENSDLAG00005017184	*atp5po*	ENSDLAG00005006810	*ndufc1*
ENSSAUG00010004698	*atp5po*	ENSSAUG00010000182	*ndufc2*	ENSDLAG00005016960	*cox4i.1*	ENSDLAG00005024193	*ndufc2*
ENSSAUG00010008794	*cox4i.1*	ENSSAUG00010006080	*ndufs1.1*	ENSDLAG00005020602	*cox4i.2*	ENSDLAG00005025864	*ndufs1.1*
ENSSAUG00010018002	*cox4i.2*	ENSSAUG00010021399	*ndufs1.2*	ENSDLAG00005023073	*cox5a.1*	ENSDLAG00005009594	*ndufs1.2*
ENSSAUG00010022937	*cox5a.1*	ENSSAUG00010004779	*ndufs2*	ENSDLAG00005014376	*cox5a.2*	ENSDLAG00005024309	*ndufs2*
ENSSAUG00010016180	*cox5a.2*	ENSSAUG00010004194	*ndufs3*	ENSDLAG00005009814	*cox5b.1*	ENSDLAG00005019820	*ndufs3*
ENSSAUG00010021488	*cox5b.1*	ENSSAUG00010021984	*ndufs4*	ENSDLAG00005014851	*cox5b.2*	ENSDLAG00005014329	*ndufs4*
ENSSAUG00010012698	*cox5b.2*	ENSSAUG00010020202	*ndufs5*	ENSDLAG00005011849	*cox6a.1*	ENSDLAG00005015202	*ndufs5*
ENSSAUG00010019745	*cox6a.1*	ENSSAUG00010008813	*ndufs6*	ENSDLAG00005010421	*cox6a.2*	ENSDLAG00005002523	*ndufs6*
ENSSAUG00010024406	*cox6a.2*	ENSSAUG00010010516	*ndufs7*	ENSDLAG00005001169	*cox6b2.1*	ENSDLAG00005022657	*ndufs7*
ENSSAUG00010020219	*cox6b2.1*	ENSSAUG00010003862	*ndufs8.1*	ENSDLAG00005008900	*cox6b2.2*	ENSDLAG00005012483	*ndufs8.1*
ENSSAUG00010014837	*cox6b2.2*	ENSSAUG00010009712	*ndufs8.2*	ENSDLAG00005001205	*cox6b2.3*	ENSDLAG00005015082	*ndufs8.2*

ENSSAUG00010005990	*cox6c.1*	ENSSAUG00010018928	*ndufv1.1*	ENSDLAG00005031075	*cox6c*	ENSDLAG00005004491	*ndufv1.1*
ENSSAUG00010005612	*cox6c.2*	ENSSAUG00010012863	*ndufv1.2*	ENSDLAG00005012733	*cox7a1.1*	ENSDLAG00005013355	*ndufv1.2*
ENSSAUG00010002730	*cox7a1.1*	ENSSAUG00010012337	*ndufv2*	ENSDLAG00005022694	*cox7a1.2*	ENSDLAG00005025947	*ndufv2*
ENSSAUG00010007755	*cox7a1.2*	ENSSAUG00010001379	*ndufv3*	ENSDLAG00005001713	*cox7a2.1a*	ENSDLAG00005001080	*sdha*
ENSSAUG00010024392	*cox7a2.1*	ENSSAUG00010014692	*sdha*	ENSDLAG00005001717	*cox7a2.1b*	ENSDLAG00005008624	*sdhb.1*
ENSSAUG00010023051	*cox7a2l.1*	ENSSAUG00010000613	*sdhb.1*	ENSDLAG00005007600	*cox7a2l.1*	ENSDLAG00005016007	*sdhb.2*
ENSSAUG00010023048	*cox7a2l.2*	ENSSAUG00010011676	*sdhb.2*	ENSDLAG00005023340	*cox7a2l.3*	ENSDLAG00005006491	*sdhc*
ENSSAUG00010026799	*cox7a2l.3*	ENSSAUG00010008991	*sdhc*	ENSDLAG00005023940	*cox7b*	ENSDLAG00005034190	*sdhd*
ENSSAUG00010006458	*cox7b*	ENSSAUG00010007110	*sdhd*	ENSDLAG00005025866	*cox7c*	ENSDLAG00005021401	*uqcr10*
ENSSAUG00010024581	*cox7c*	ENSSAUG00010021526	*uqcr10*	ENSDLAG00005009958	*cox8a*	ENSDLAG00005024816	*uqcr11.1*
ENSSAUG00010021868	*cox8a*	ENSSAUG00010018172	*uqcr11.1*	ENSDLAG00005010064	*cyc1*	ENSDLAG00005007897	*uqcr11.2*
ENSSAUG00010006809	*cyc1*	ENSSAUG00010016776	*uqcr11.2*			ENSDLAG00005019495	*uqcrb*
		ENSSAUG00010003983	*uqcrb*			ENSDLAG00005019752	*uqcrc1*
		ENSSAUG00010010721	*uqcrc1*			ENSDLAG00005014442	*uqcrc2.1*
		ENSSAUG00010002780	*uqcrc2.1*			ENSDLAG00005014897	*uqcrc2.2*
		ENSSAUG00010022860	*uqcrc2.2*			ENSDLAG00005020221	*uqcrfs1.1*
		ENSSAUG00010000160	*uqcrfs1.1*			ENSDLAG00005021271	*uqcrfs1.2*
		ENSSAUG00010025510	*uqcrfs1.2*			ENSDLAG00005003877	*uqcrh.1*
		ENSSAUG00010009981	*uqcrh.1*			ENSDLAG00005009892	*uqcrh.2*
		ENSSAUG00010022471	*uqcrh.2*			ENSDLAG00005001473	*uqcrq*
		ENSSAUG00010018590	*uqcrq*				

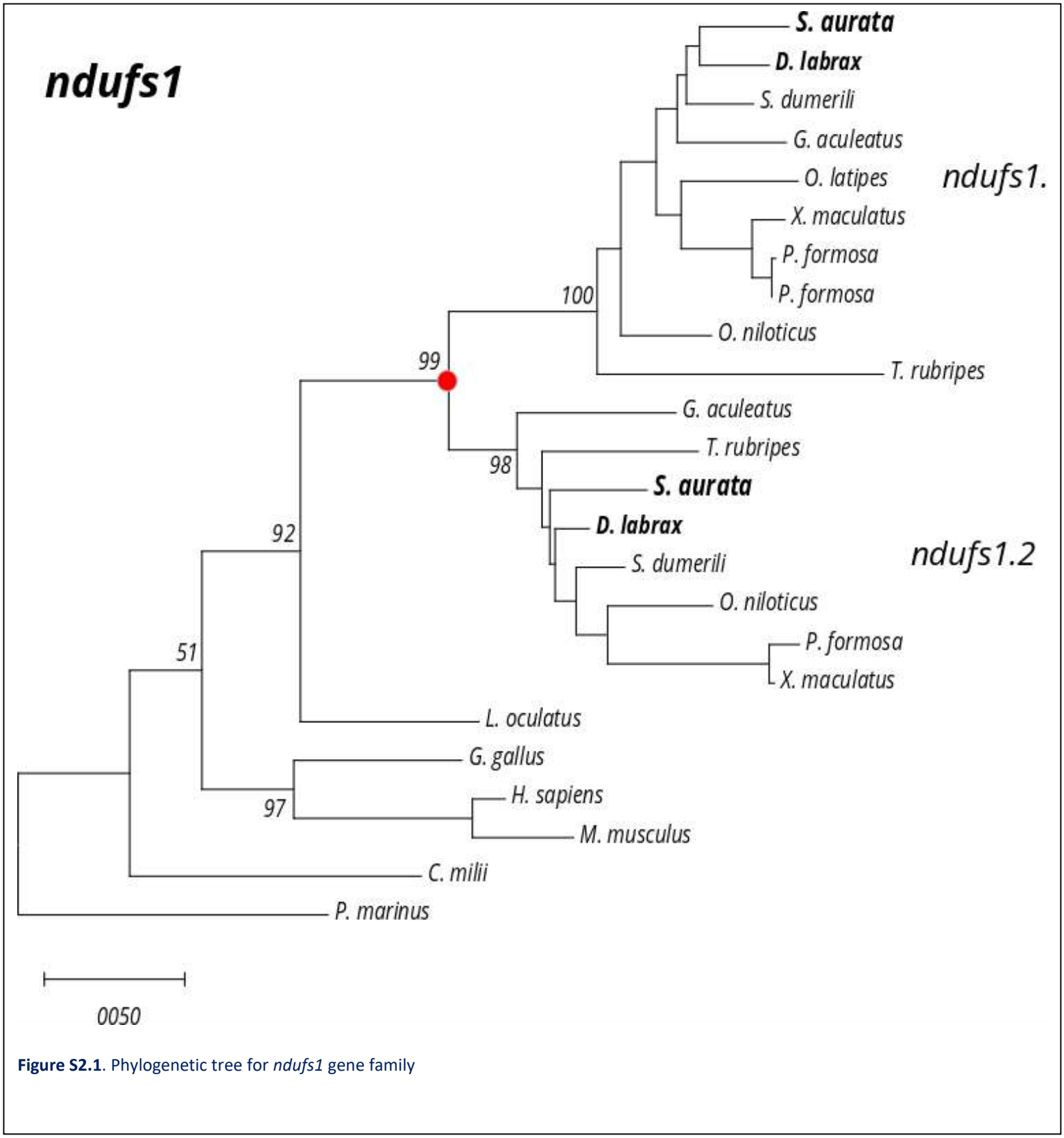


Figure S2.1. Phylogenetic tree for *ndufs1* gene family

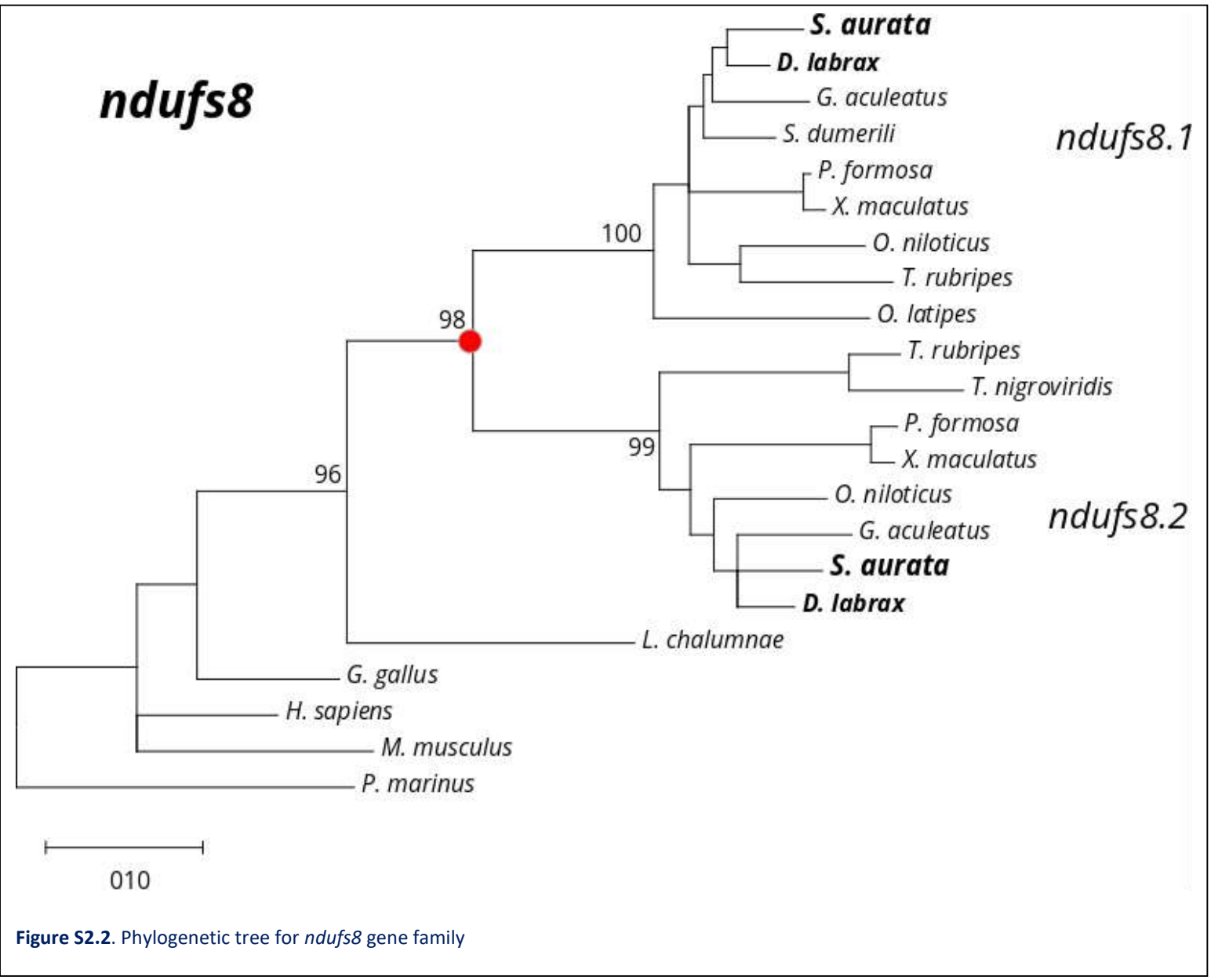


Figure S2.2. Phylogenetic tree for *ndufs8* gene family

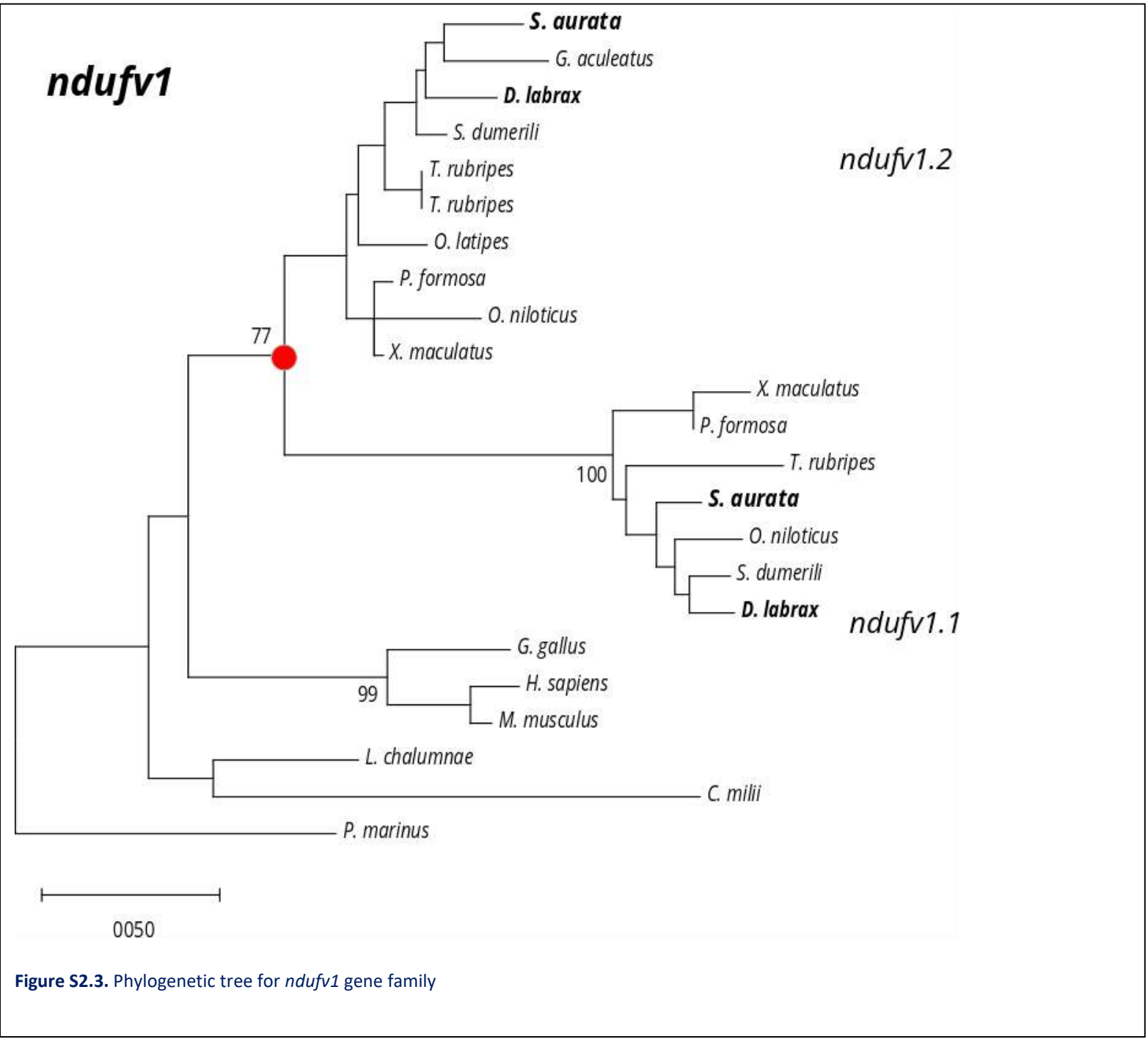


Figure S2.3. Phylogenetic tree for *ndufv1* gene family

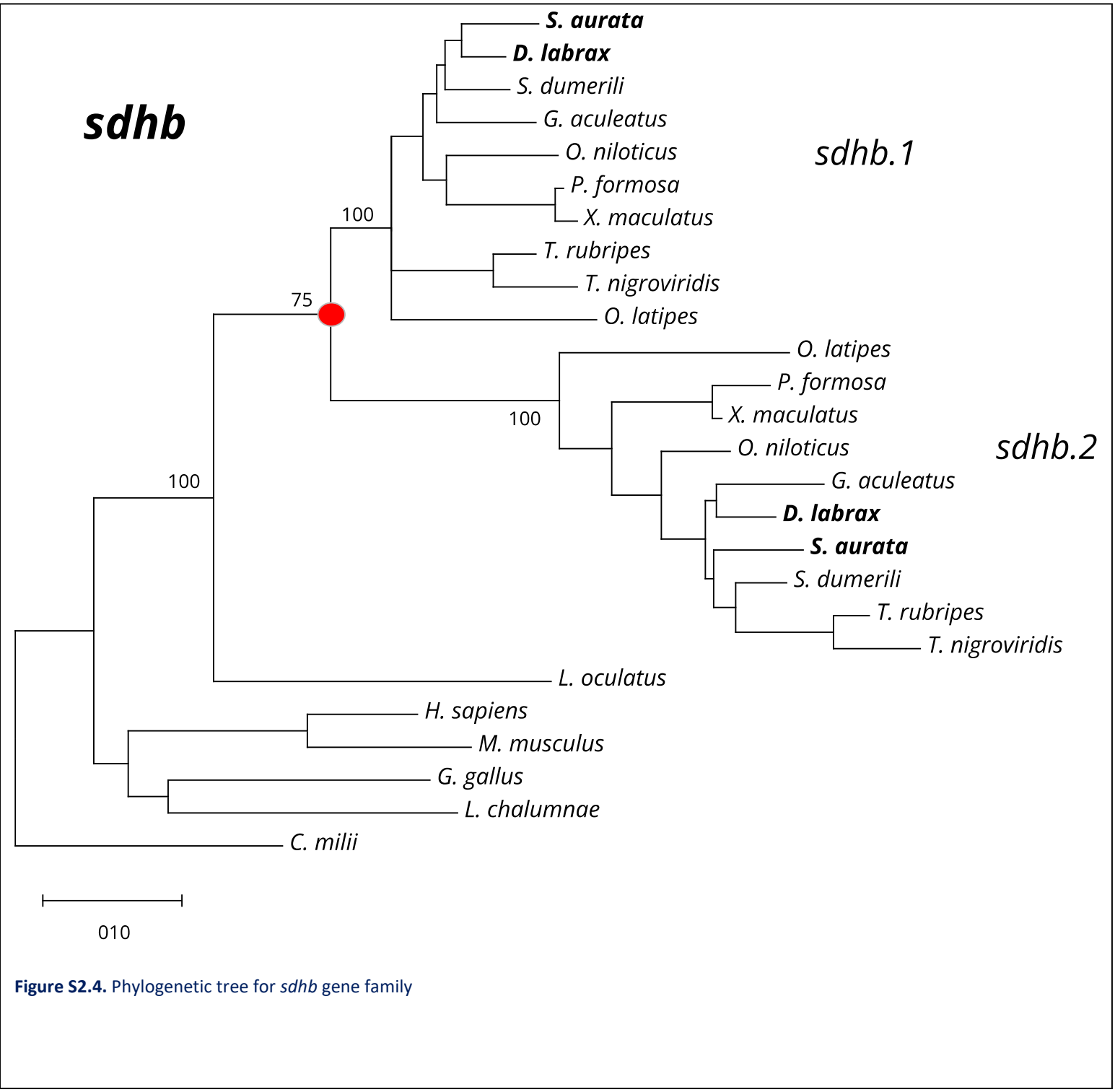


Figure S2.4. Phylogenetic tree for *sdhb* gene family

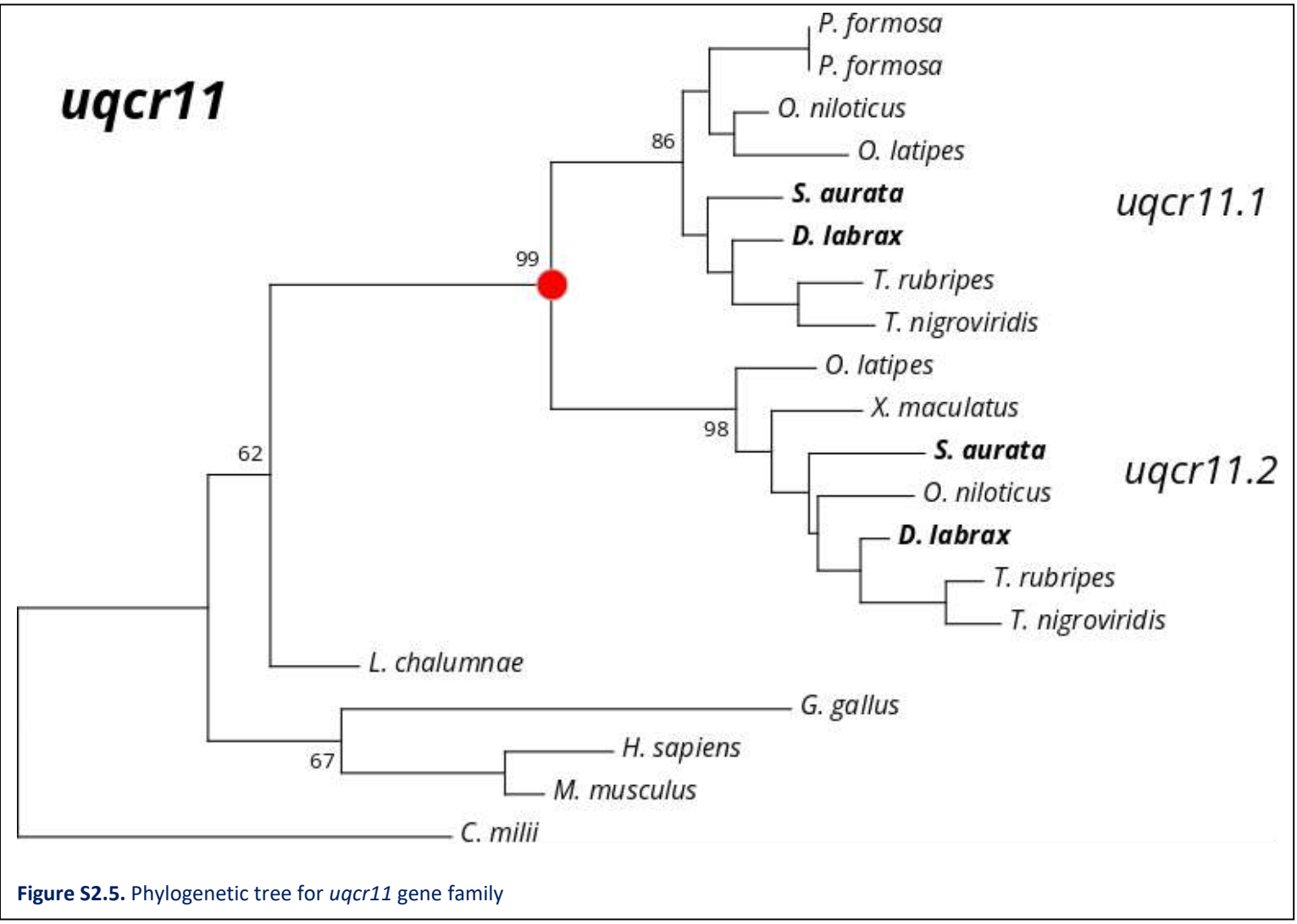


Figure S2.5. Phylogenetic tree for *uqcr11* gene family

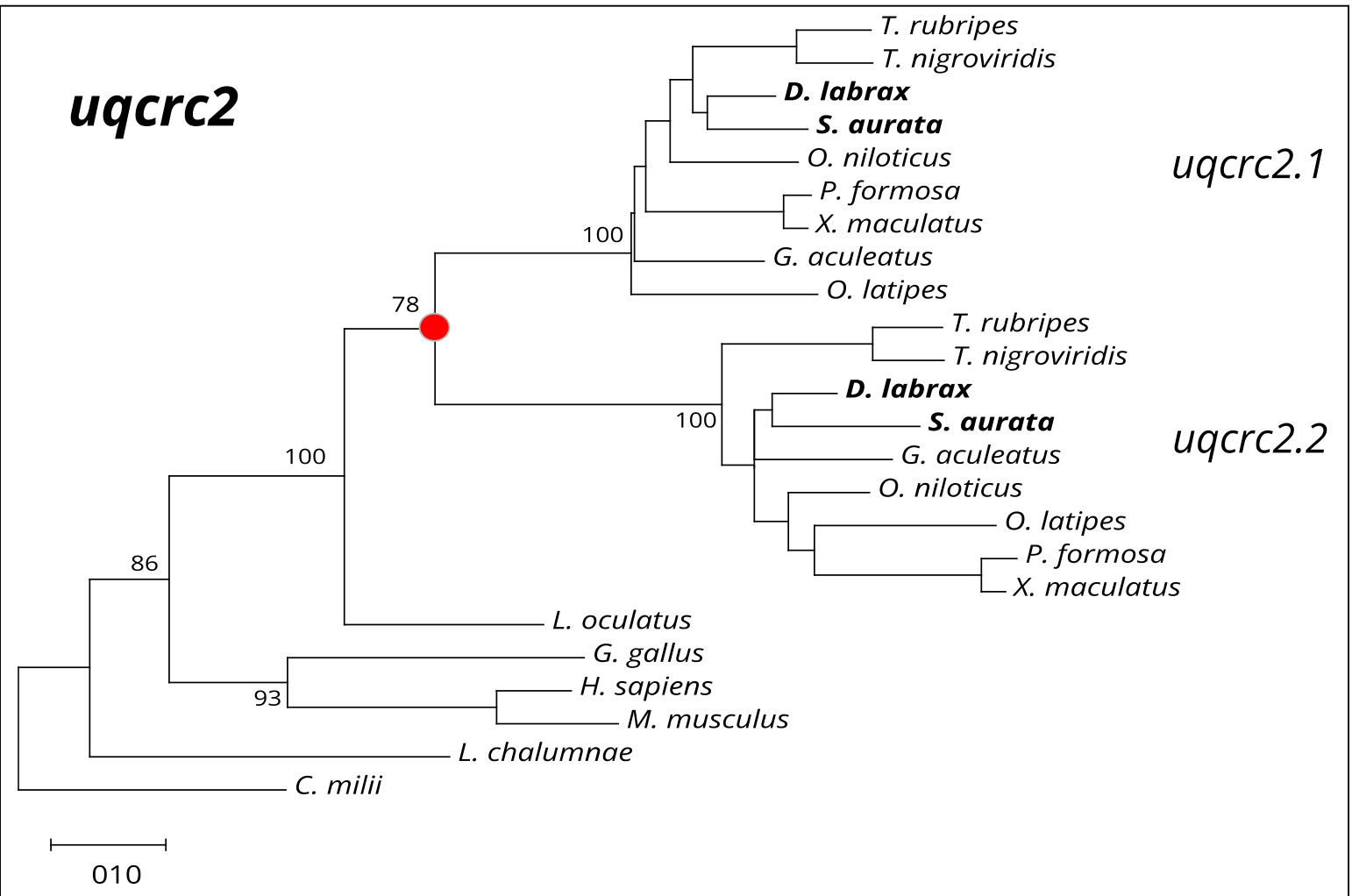


Figure S2.6. Phylogenetic tree for *uqcrc2* gene family

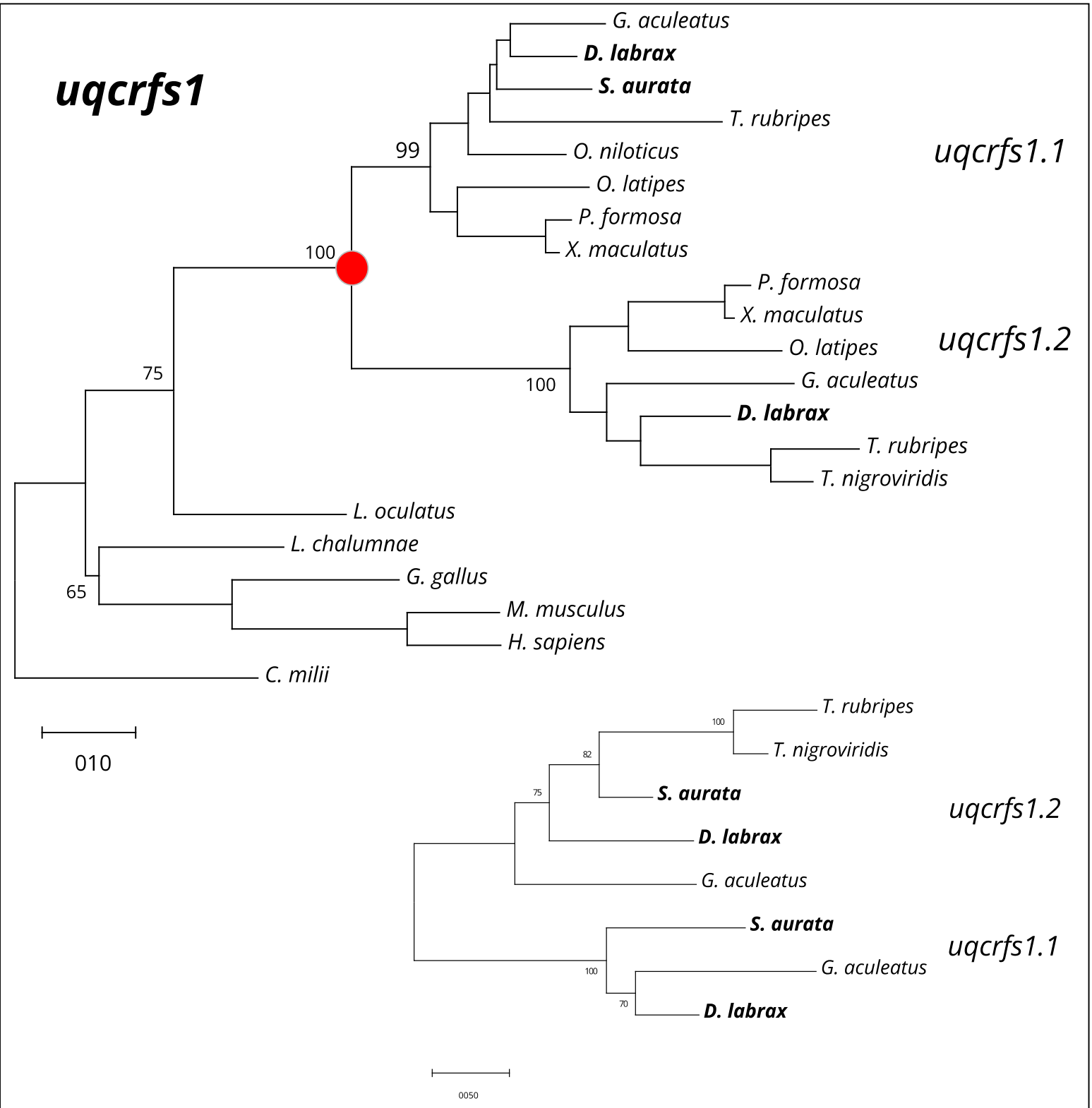


Figure S2.7. Phylogenetic tree for *uqcrfs1* gene family

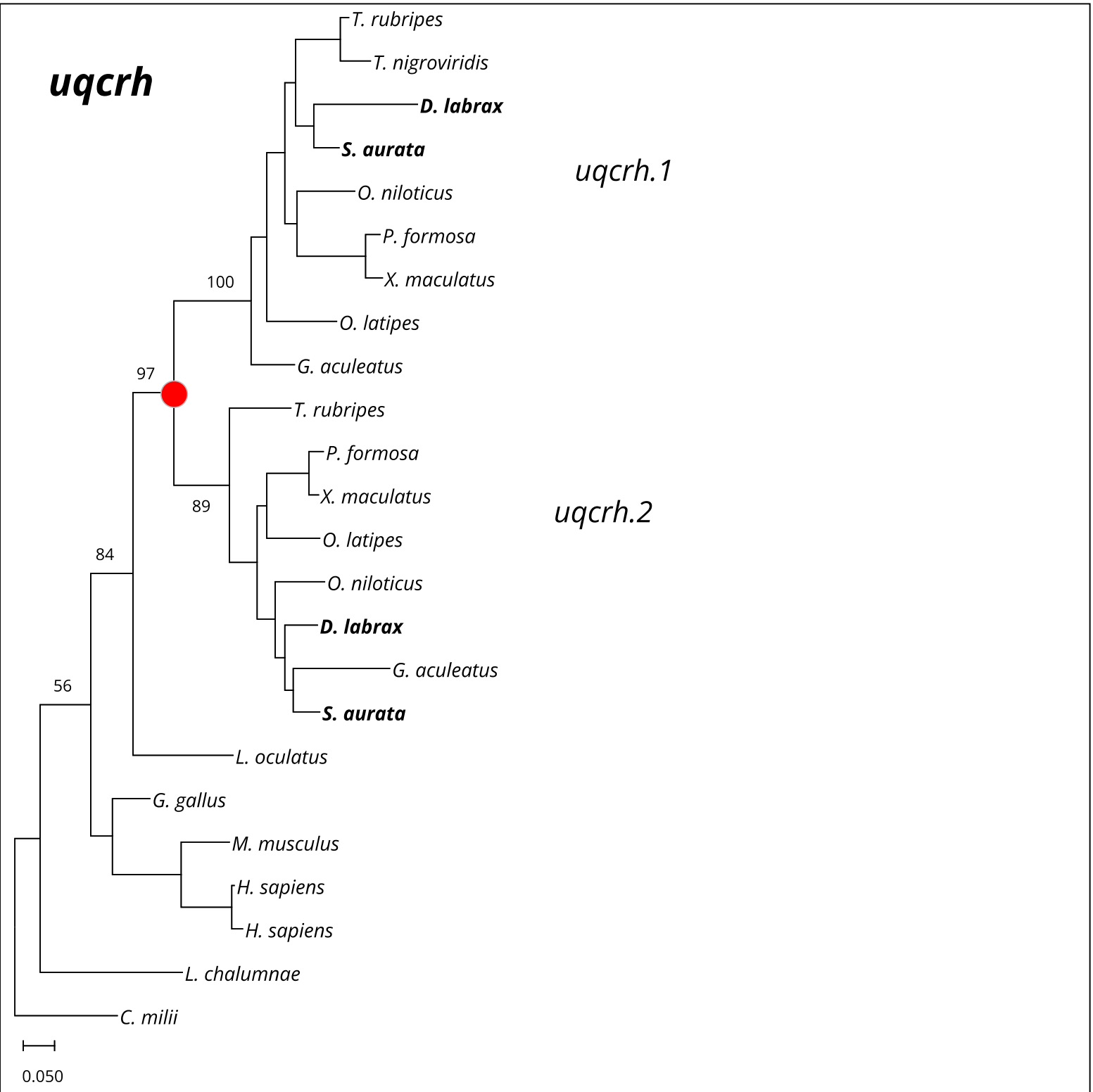


Figure S2.8. Phylogenetic tree for *uqcrh* gene family

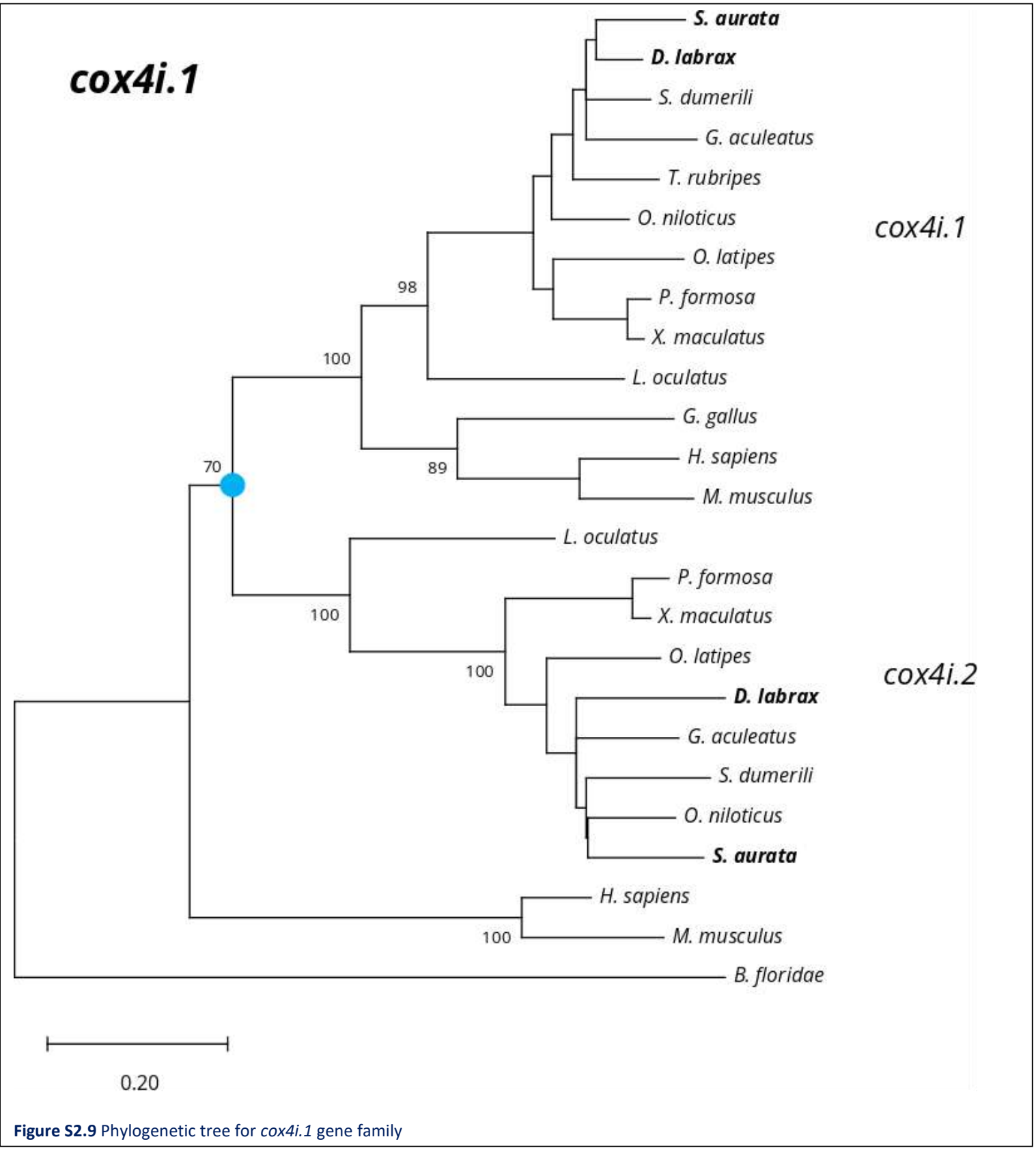


Figure S2.9 Phylogenetic tree for *cox4i.1* gene family

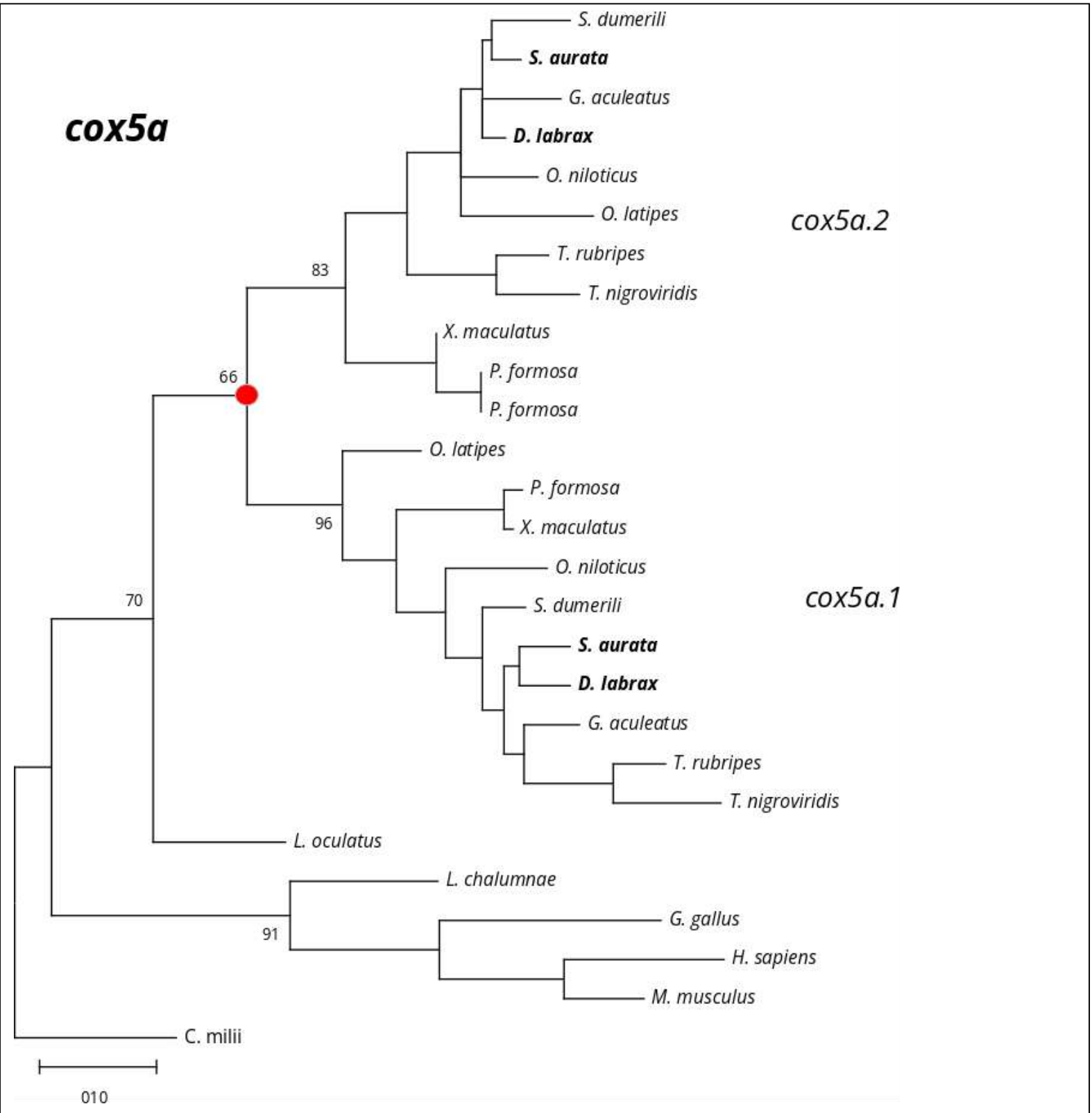


Figure S2.10. Phylogenetic tree for *cox5a* gene family

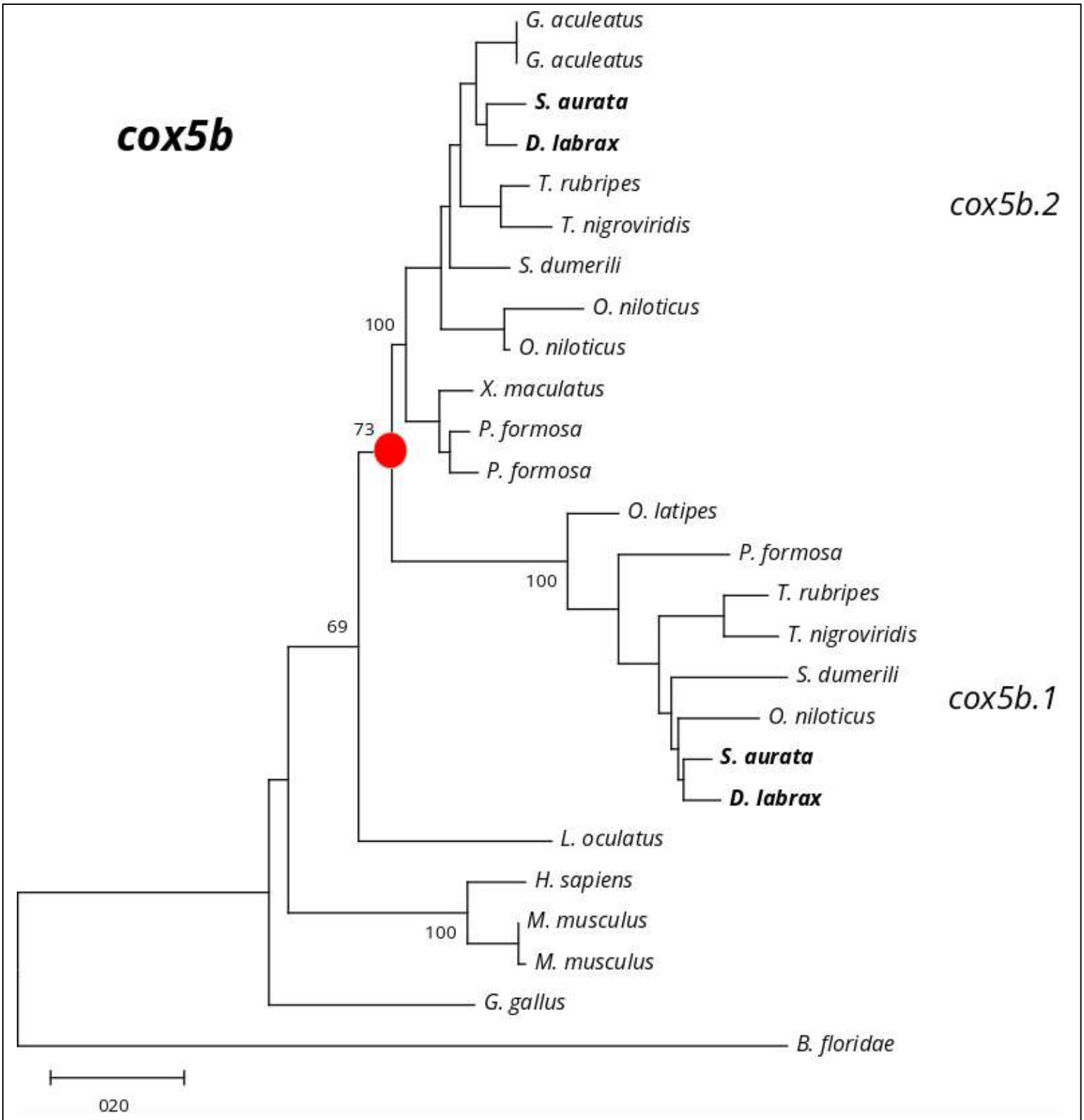


Figure S2.11. Phylogenetic tree for *cox5b* gene family

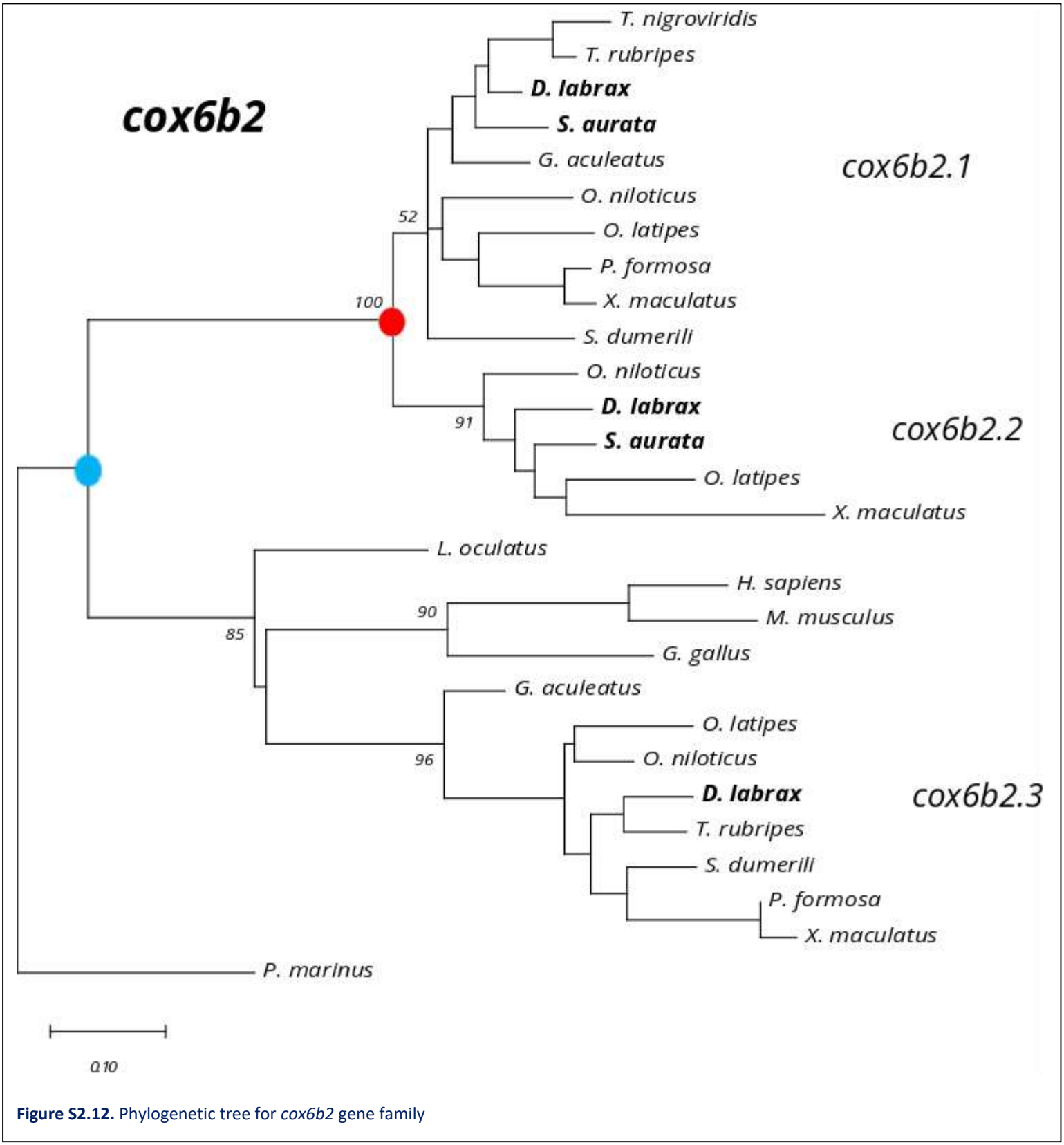


Figure S2.12. Phylogenetic tree for *cox6b2* gene family

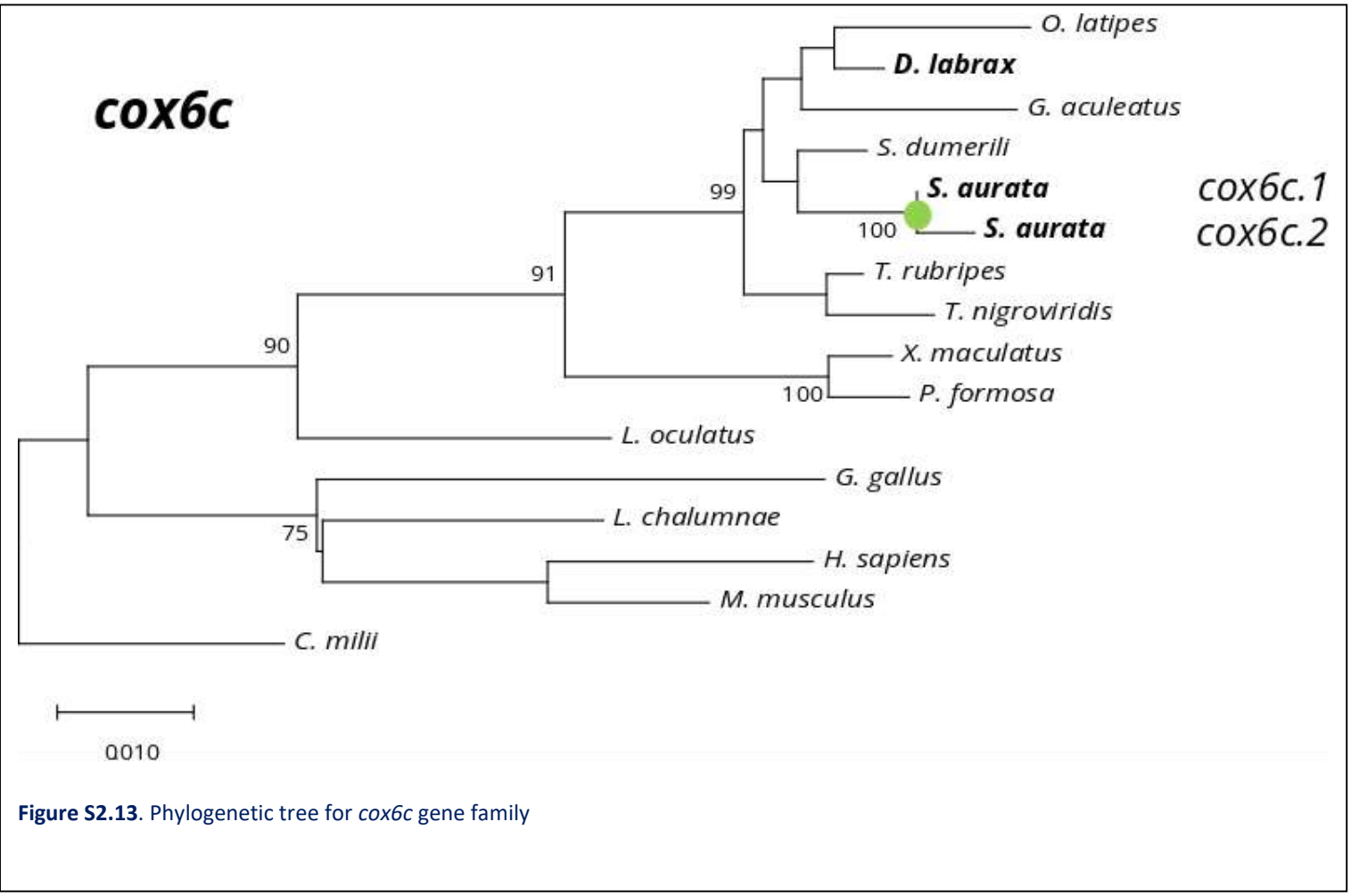


Figure S2.13. Phylogenetic tree for *cox6c* gene family

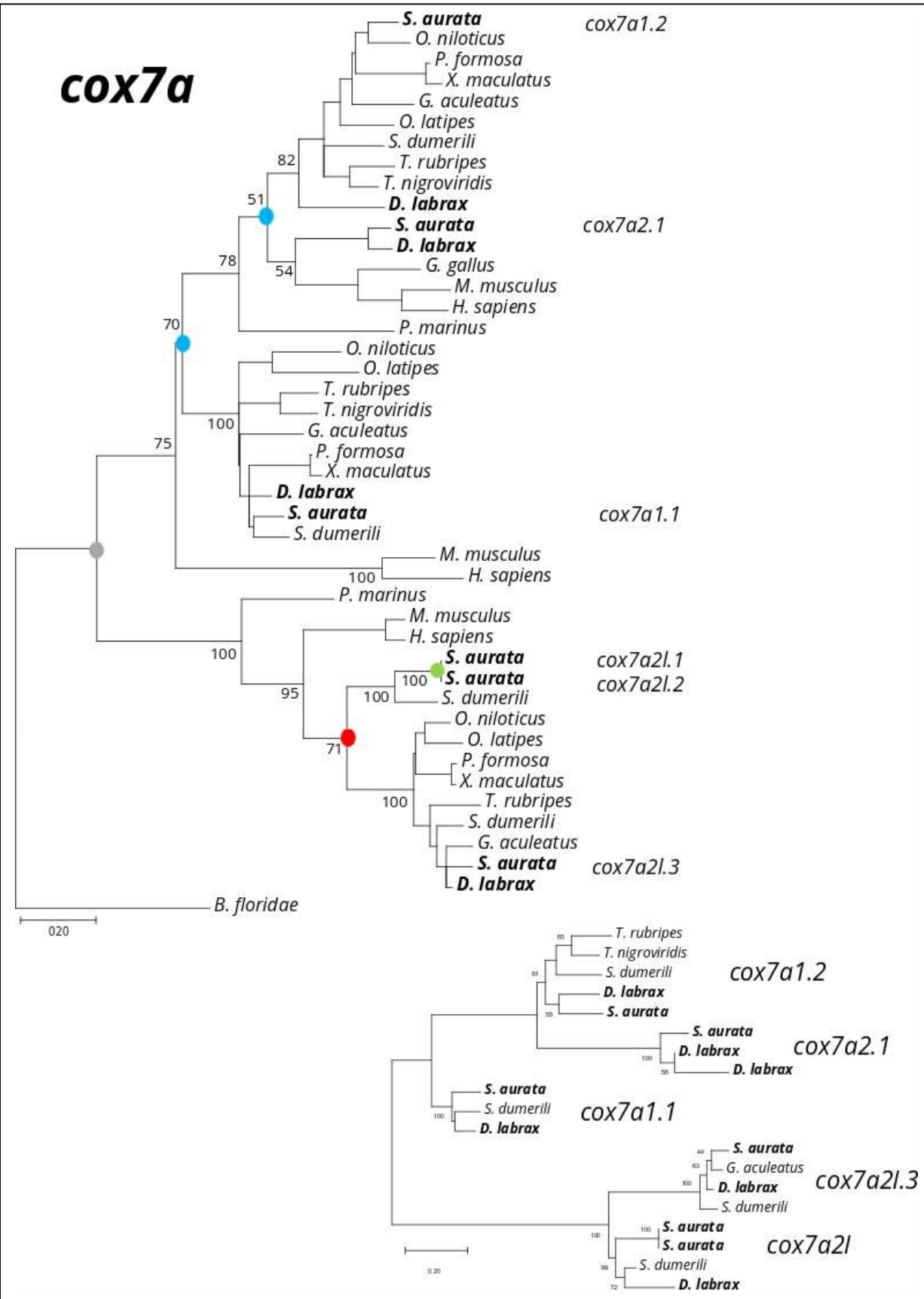


Figure S2.14. Phylogenetic tree for *cox7a* gene family

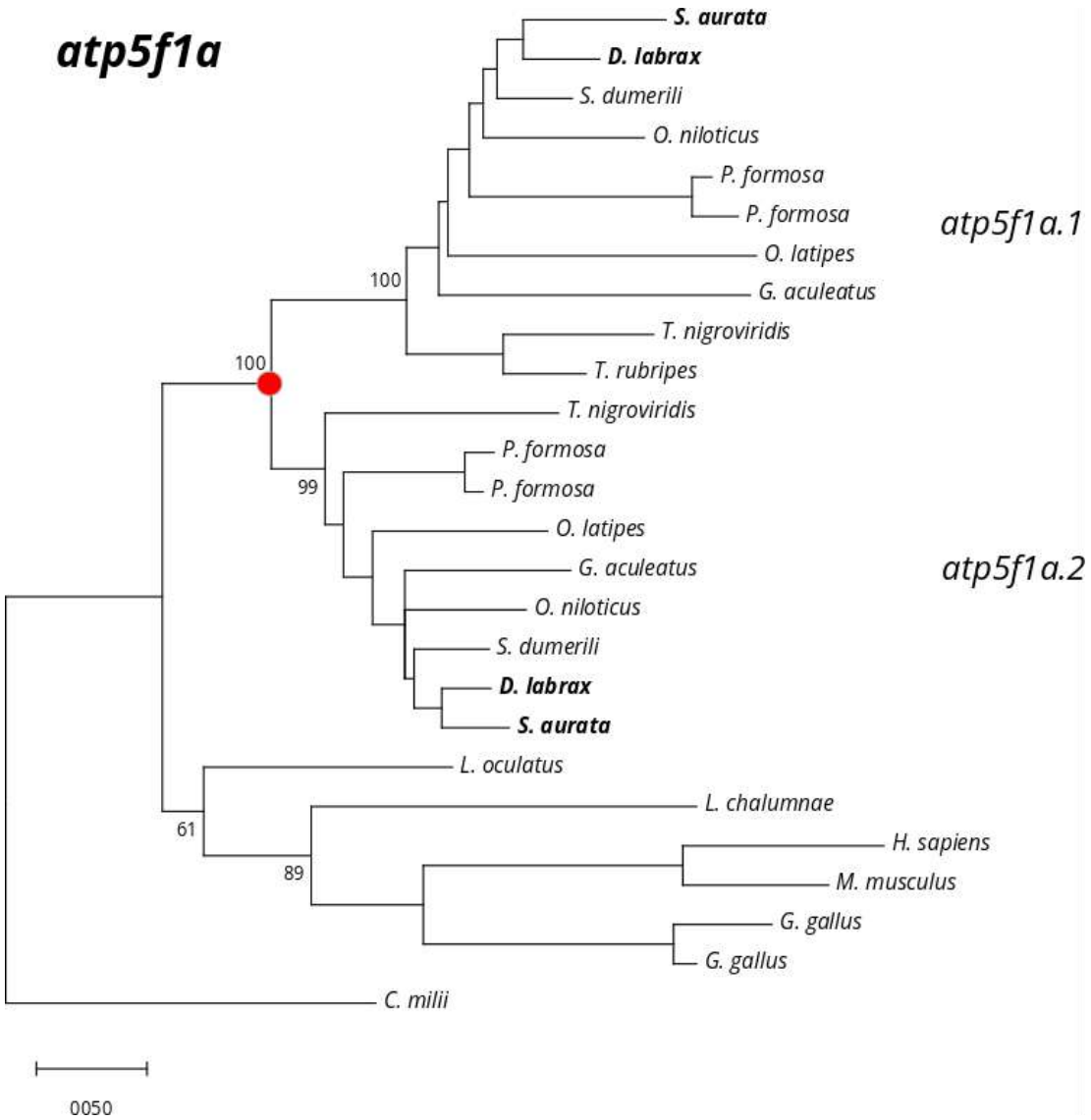


Figure S2.15. Phylogenetic tree for *atp5f1a* gene family.

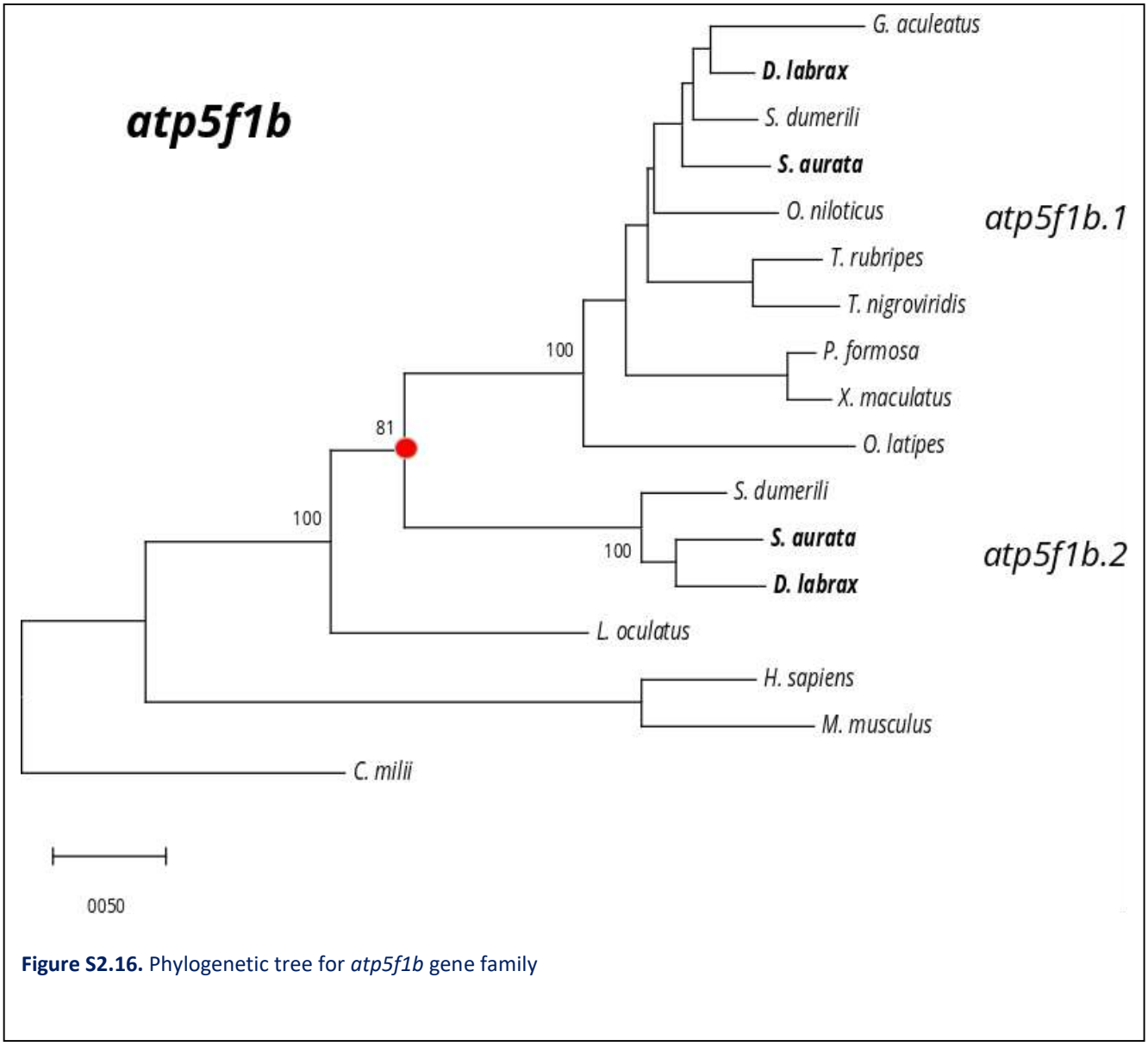


Figure S2.16. Phylogenetic tree for *atp5f1b* gene family

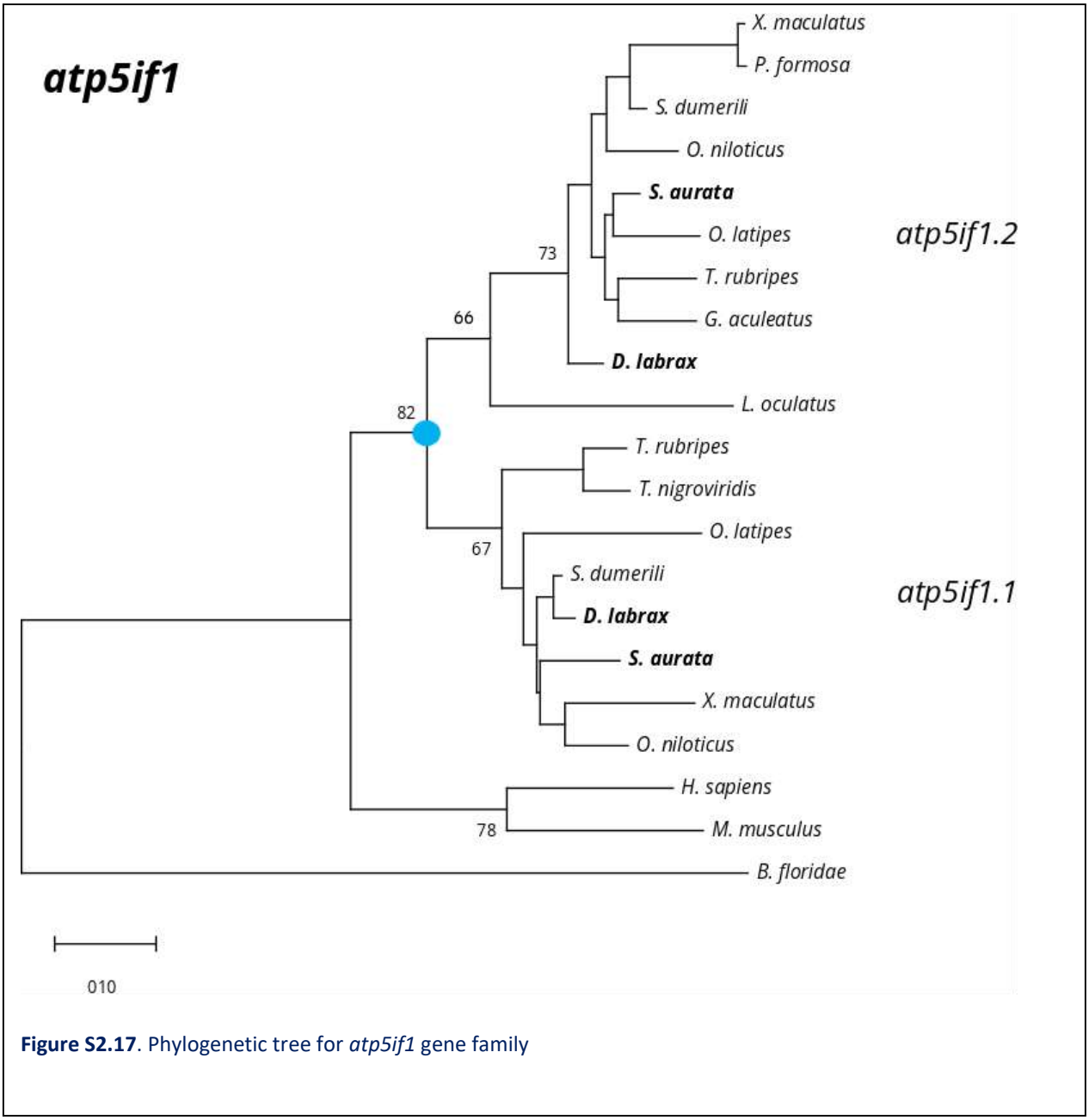


Figure S2.17. Phylogenetic tree for *atp5if1* gene family

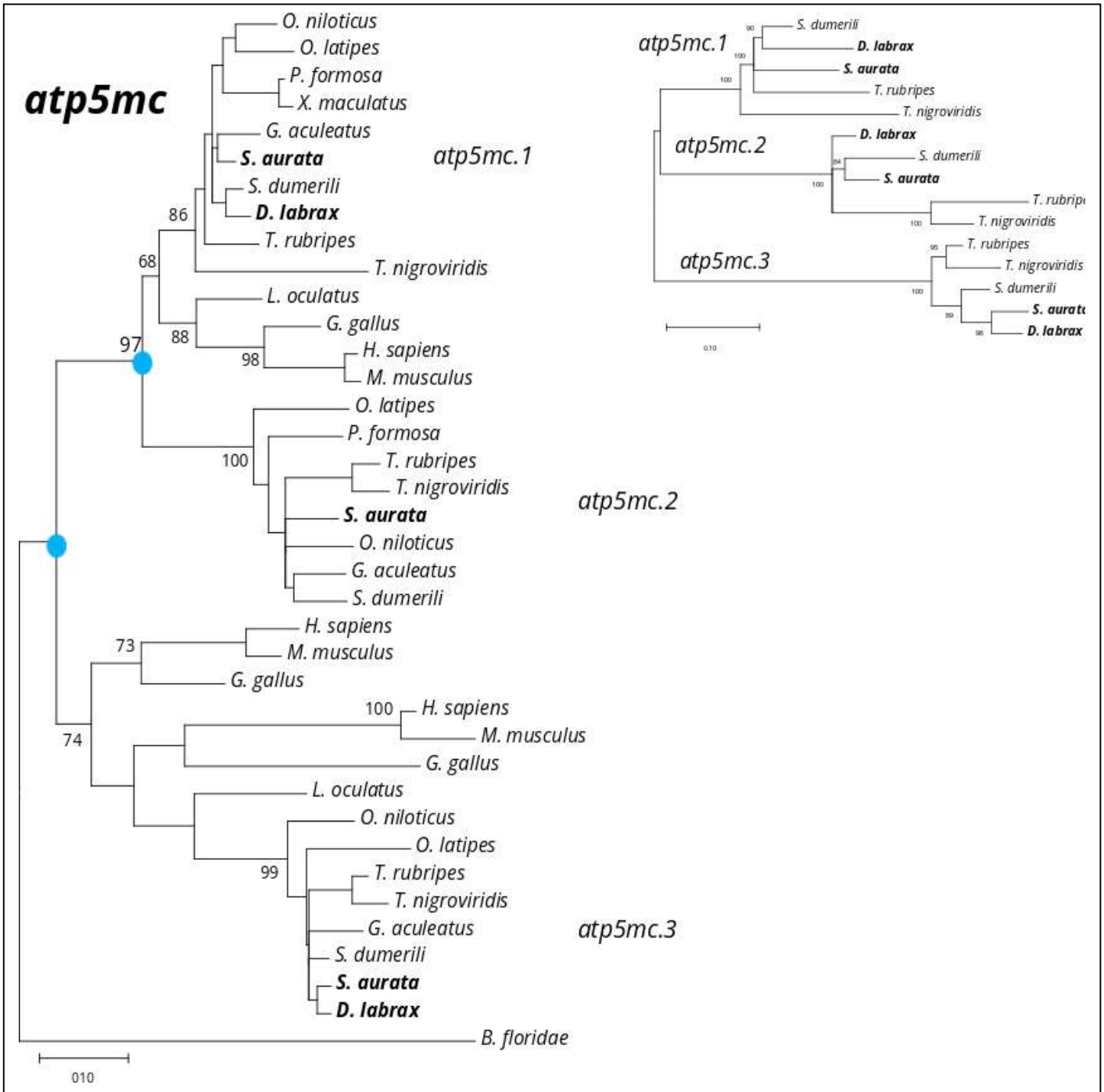


Figure S2.18. Phylogenetic tree for *atp5mc* gene family

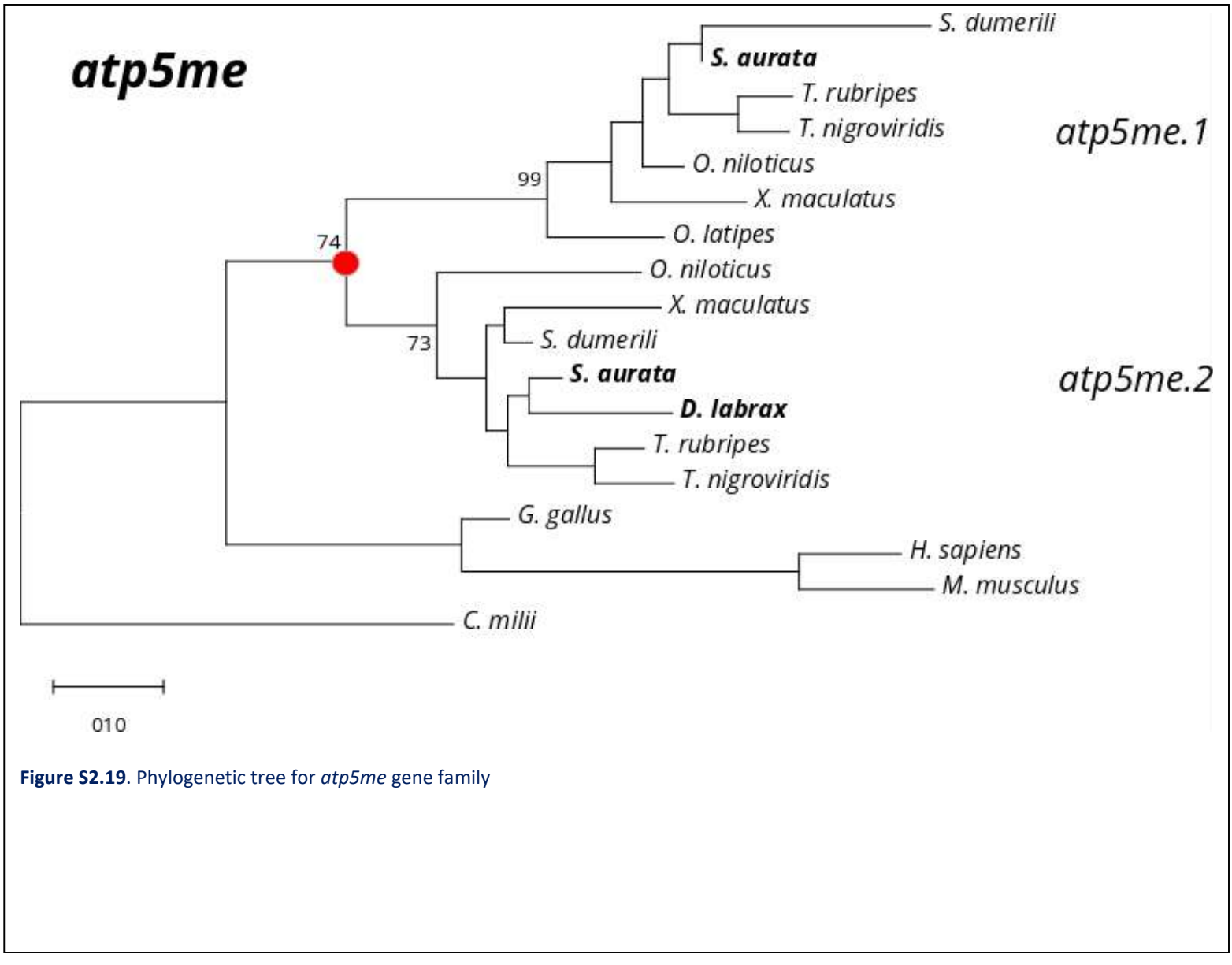


Figure S2.19. Phylogenetic tree for *atp5me* gene family

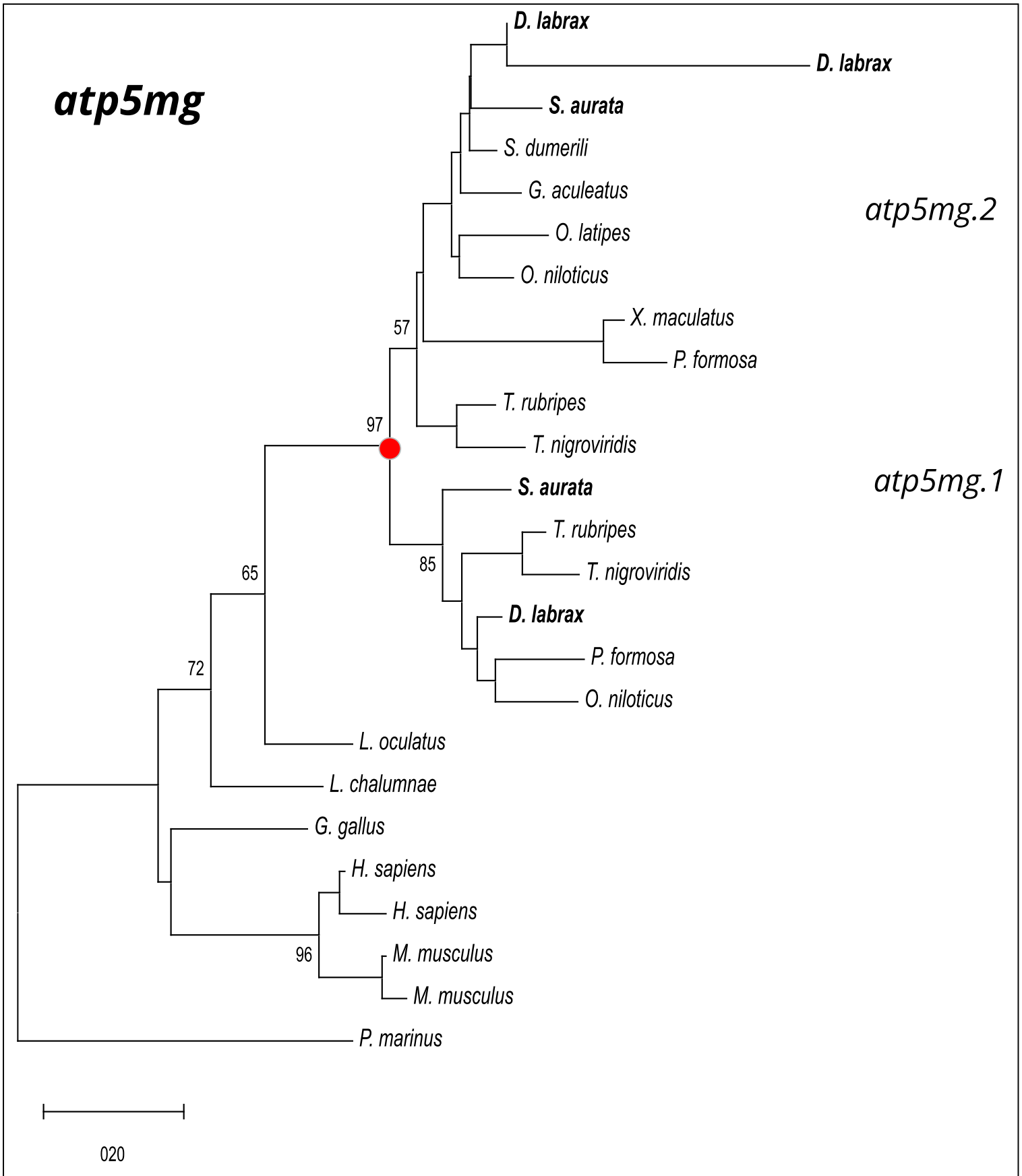


Figure S2.20. Phylogenetic tree for *atp5mg* gene family

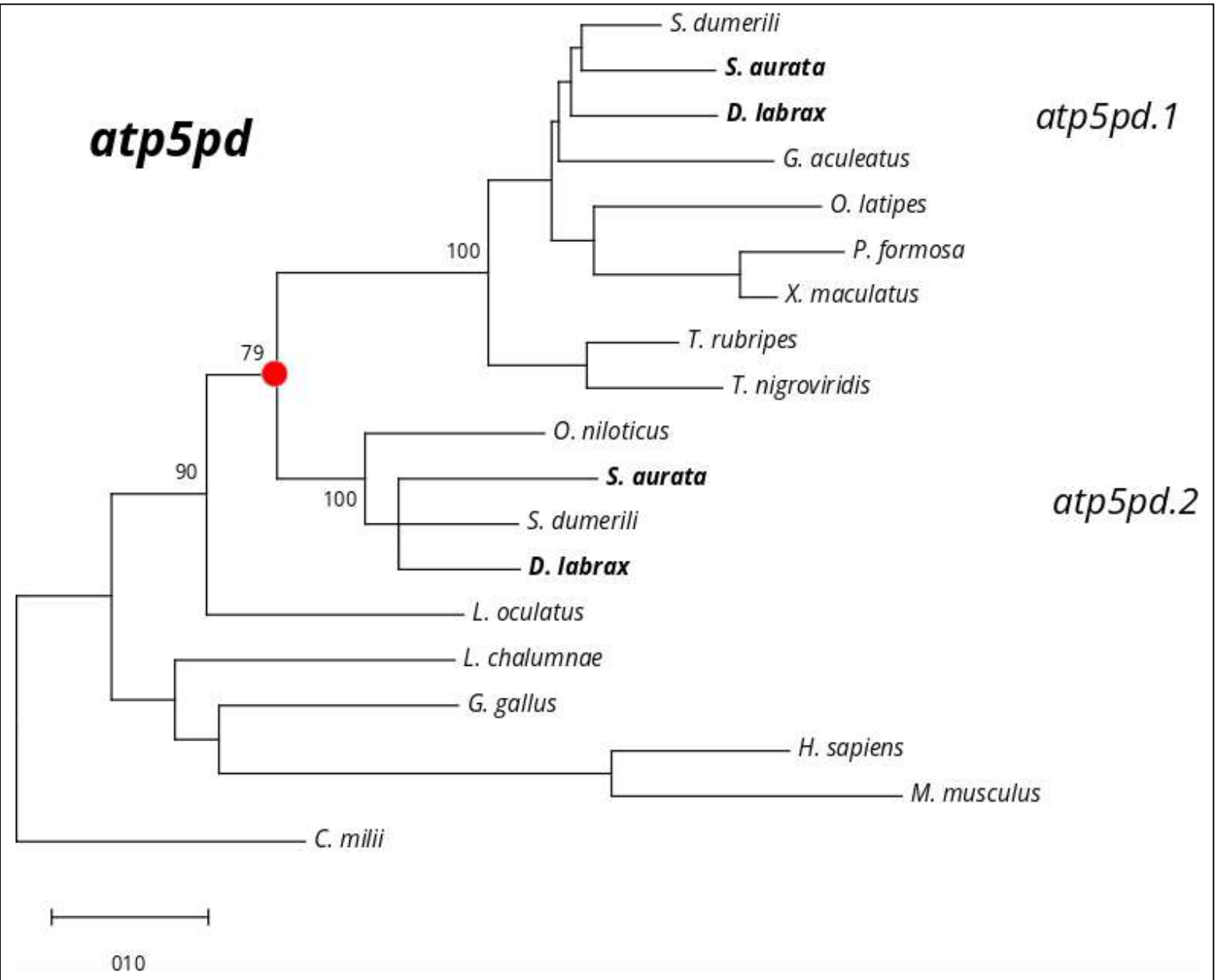


Figure S2.21. Phylogenetic tree for *atp5pd* gene family

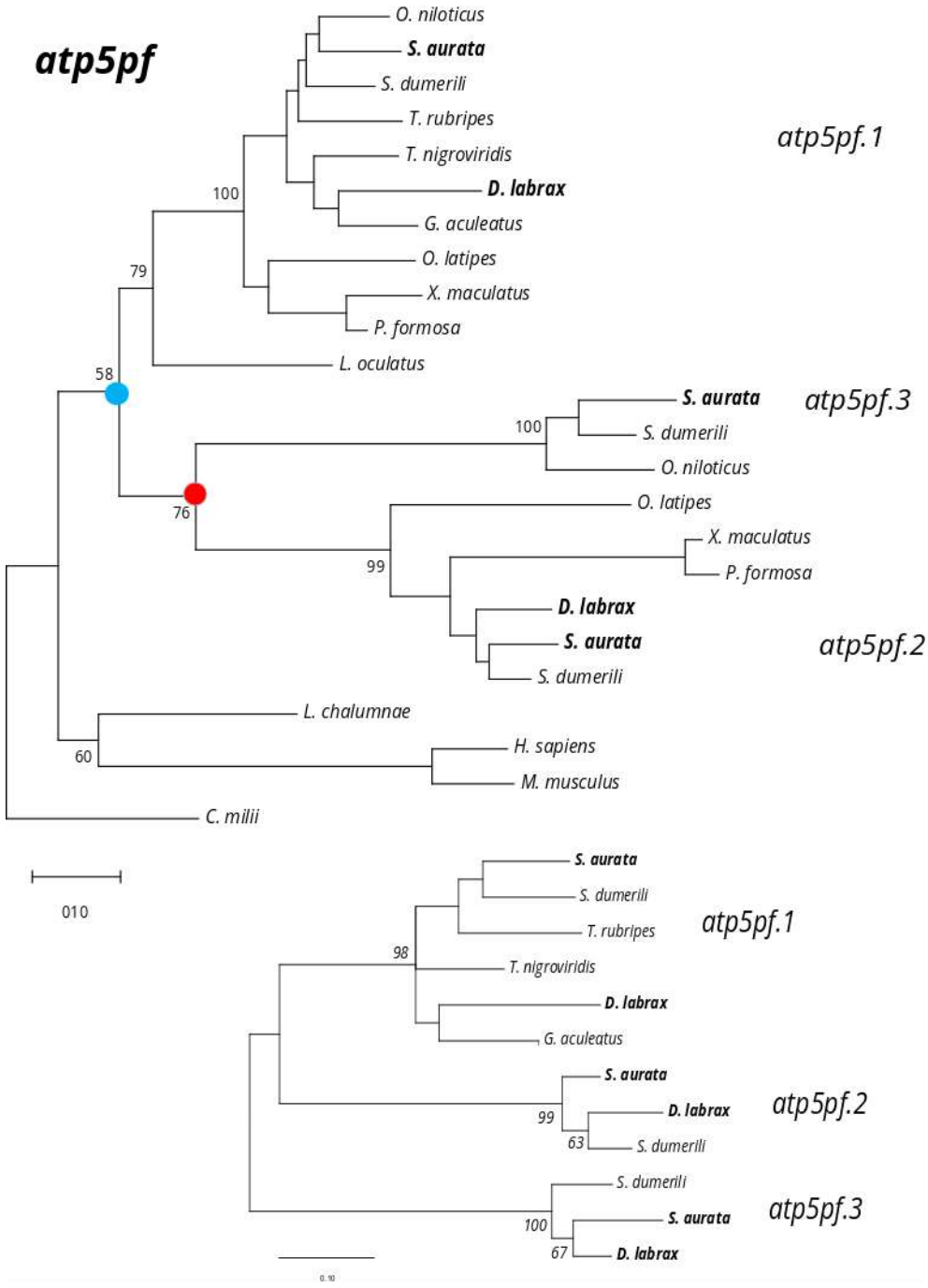


Figure S2.22. Phylogenetic tree for *atp5pf* gene family

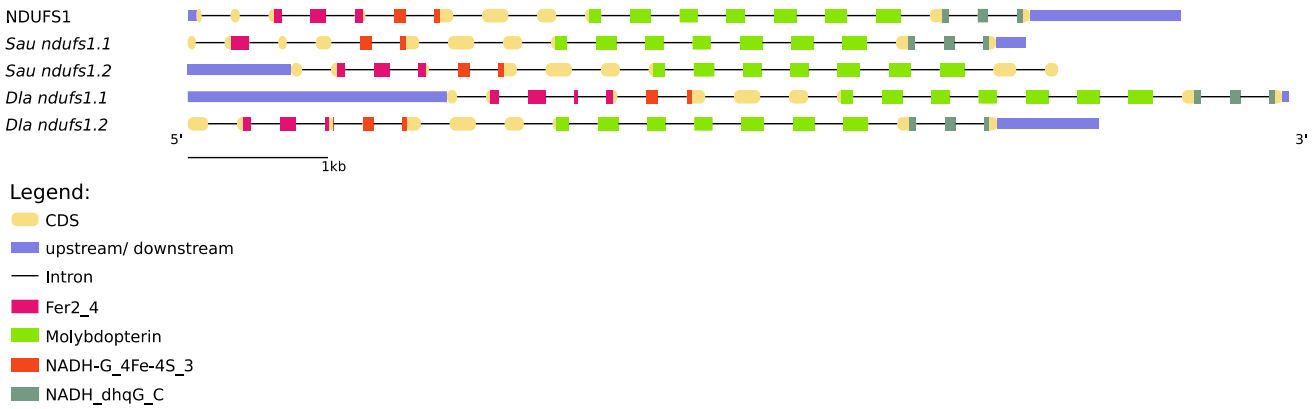


Figure S2.23. Gene structure and protein domains for *ndufs1* gene family in gilthead seabream (Sau), European seabass (Dla) and human (capital letters).

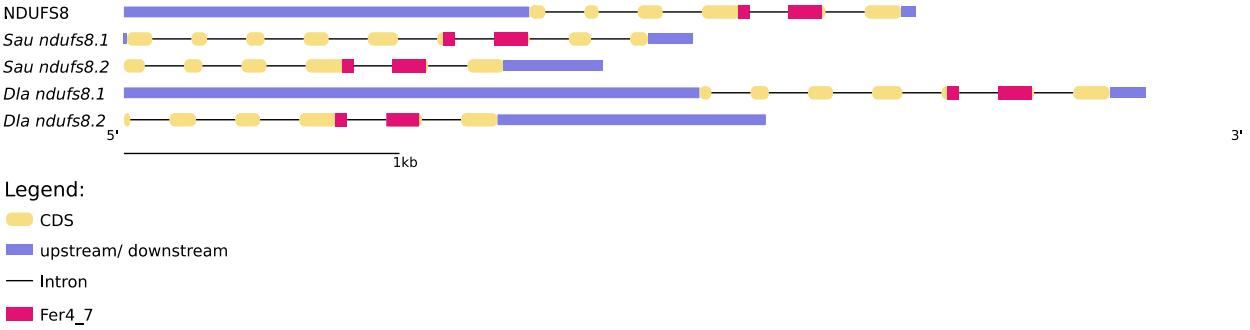


Figure S2.24. Gene structure and protein domains for *ndufs8* gene family in gilthead seabream (Sau), European seabass (Dla) and human (capital letters).

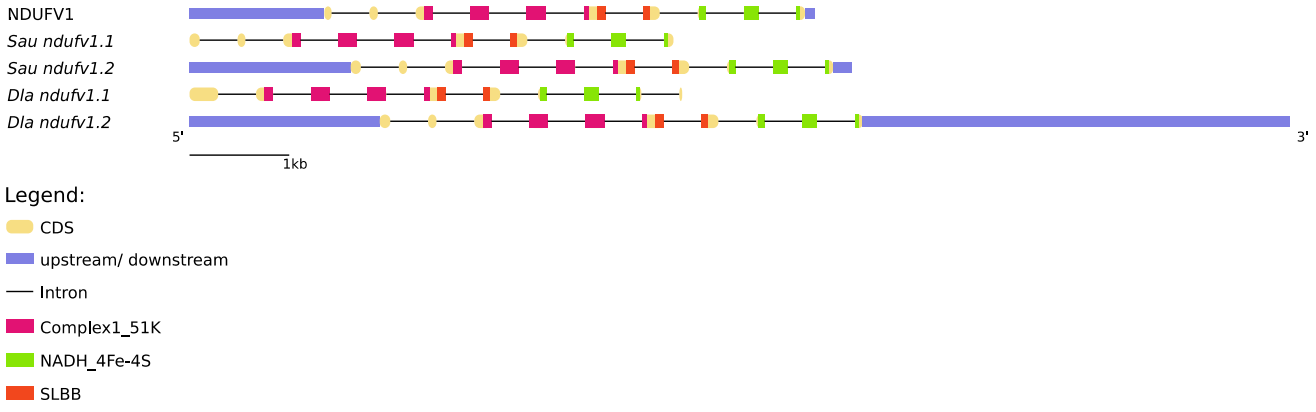


Figure S2.25. Gene structure and protein domains for *ndufv1* gene family in gilthead seabream (Sau), European seabass (Dla) and human (capital letters).

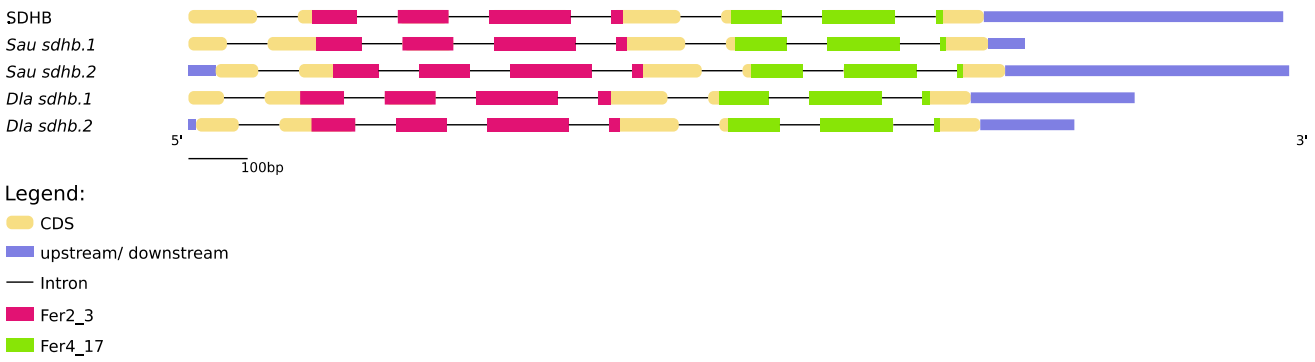


Figure S2.26. Gene structure and protein domains for *sdhb* gene family in gilthead seabream (Sau), European seabass (Dla) and human (capital letters).

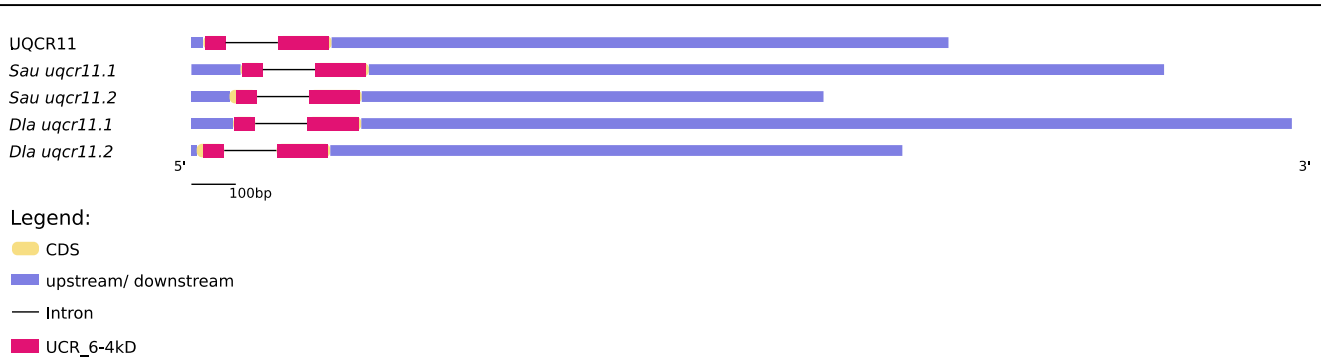


Figure S2.27. Gene structure and protein domains for *uqcr11* gene family in gilthead seabream (Sau), European seabass (Dla) and human (capital letters).

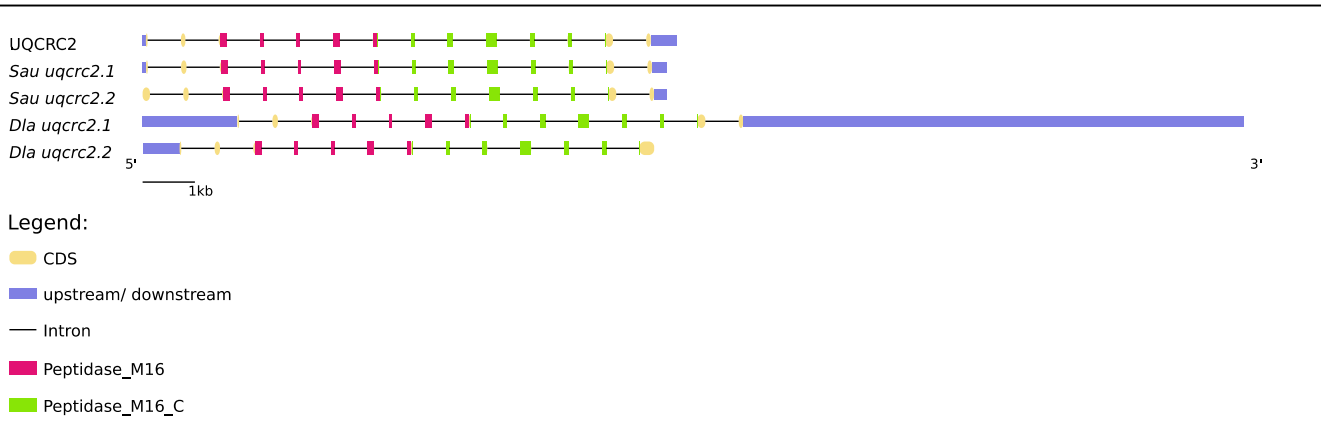


Figure S2.28. Gene structure and protein domains for *uqcrc2* gene family in gilthead seabream (Sau), European seabass (Dla) and human (capital letters).

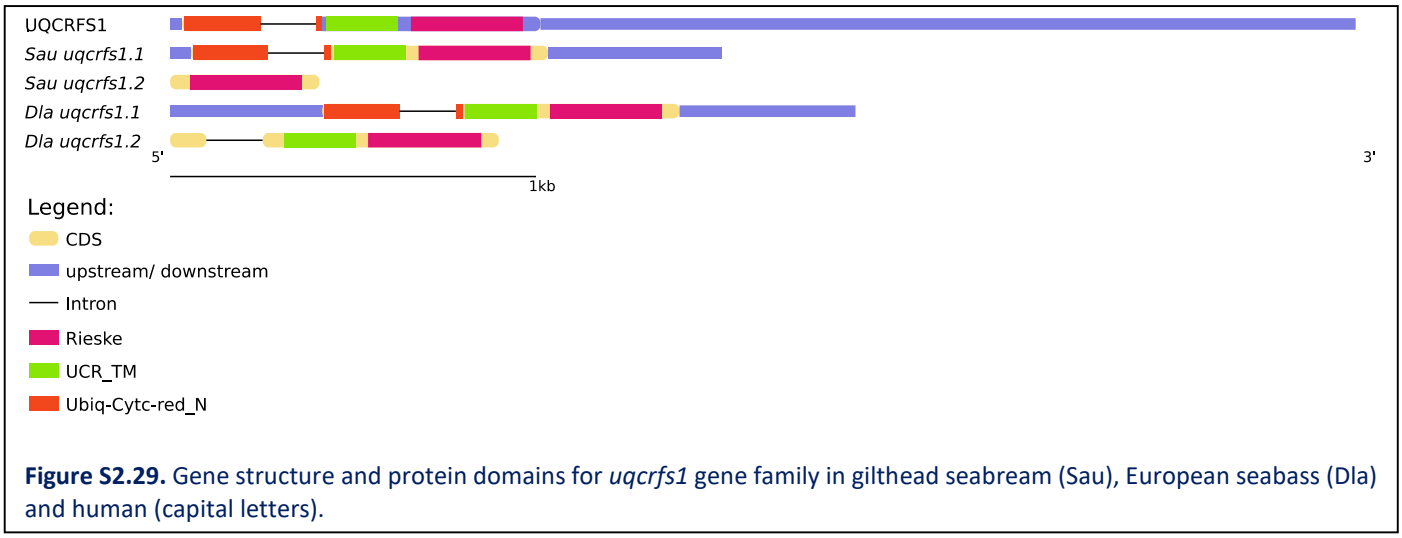


Figure S2.29. Gene structure and protein domains for *uqcrfs1* gene family in gilthead seabream (Sau), European seabass (Dla) and human (capital letters).

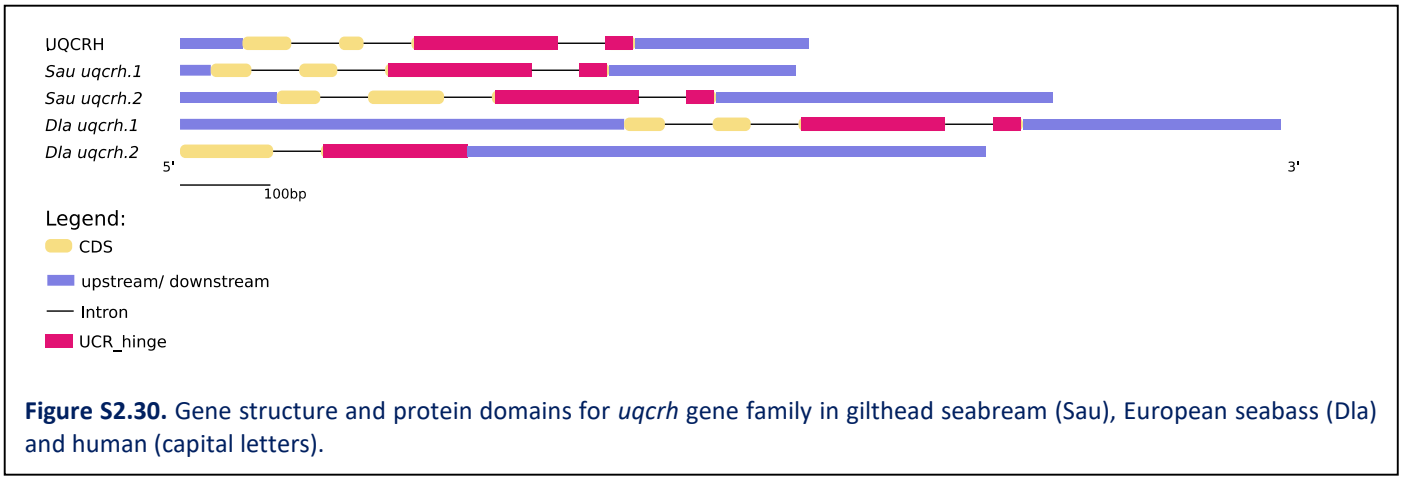


Figure S2.30. Gene structure and protein domains for *uqcrh* gene family in gilthead seabream (Sau), European seabass (Dla) and human (capital letters).

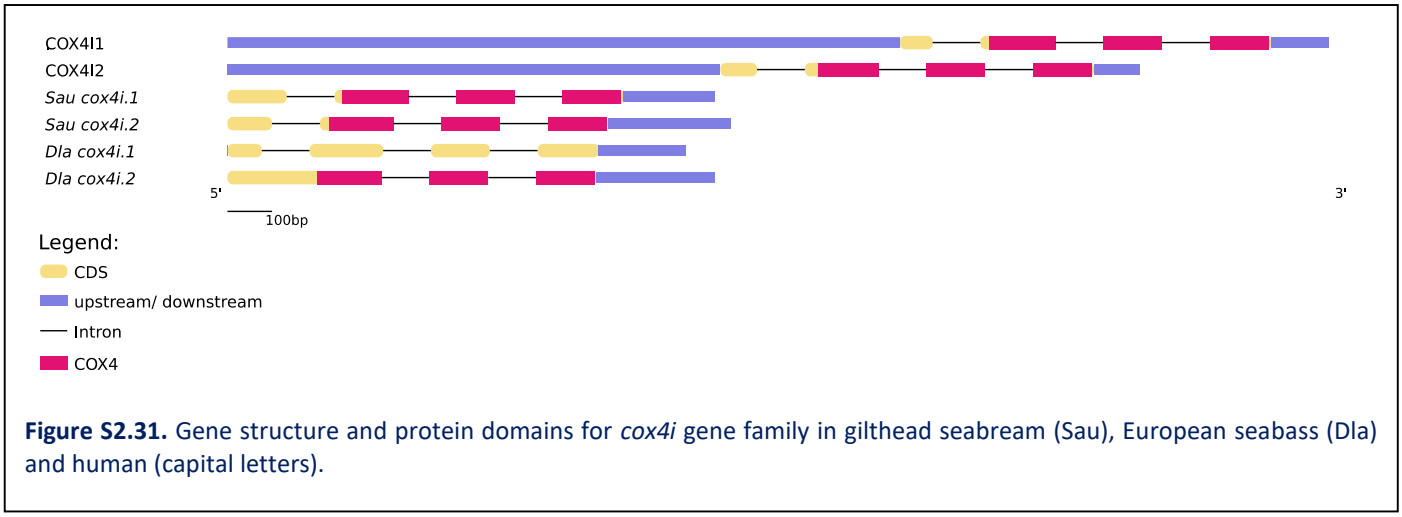


Figure S2.31. Gene structure and protein domains for *cox4i* gene family in gilthead seabream (Sau), European seabass (Dla) and human (capital letters).

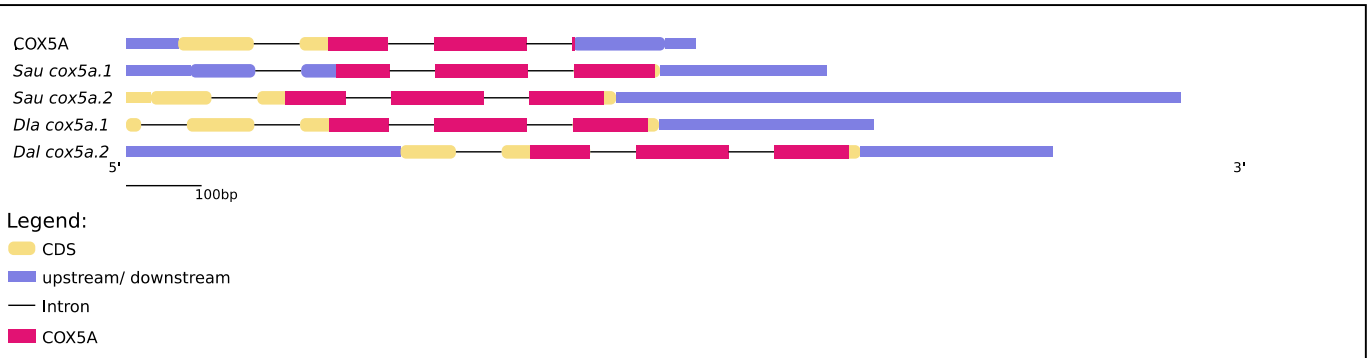


Figure S2.32. Gene structure and protein domains for *cox5a* gene family in gilthead seabream (Sau), European seabass (Dla) and human (capital letters).

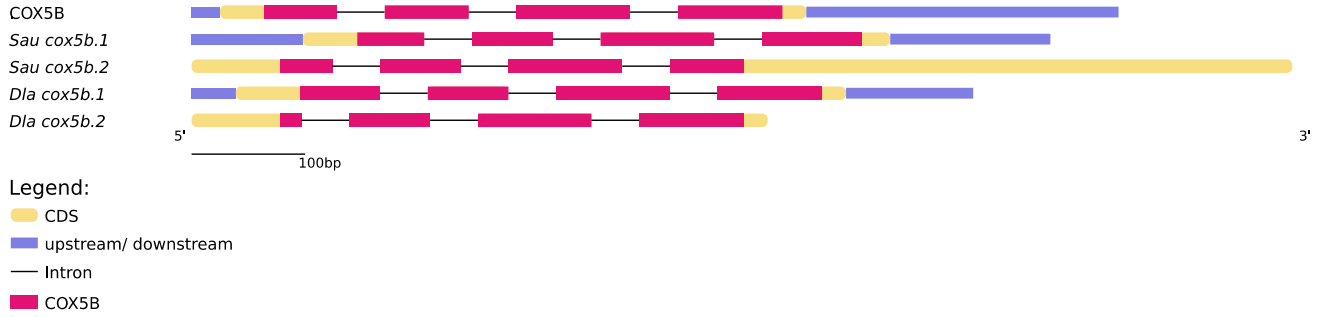


Figure S2.33. Gene structure and protein domains for *cox5b* gene family in gilthead seabream (Sau), European seabass (Dla) and human (capital letters).

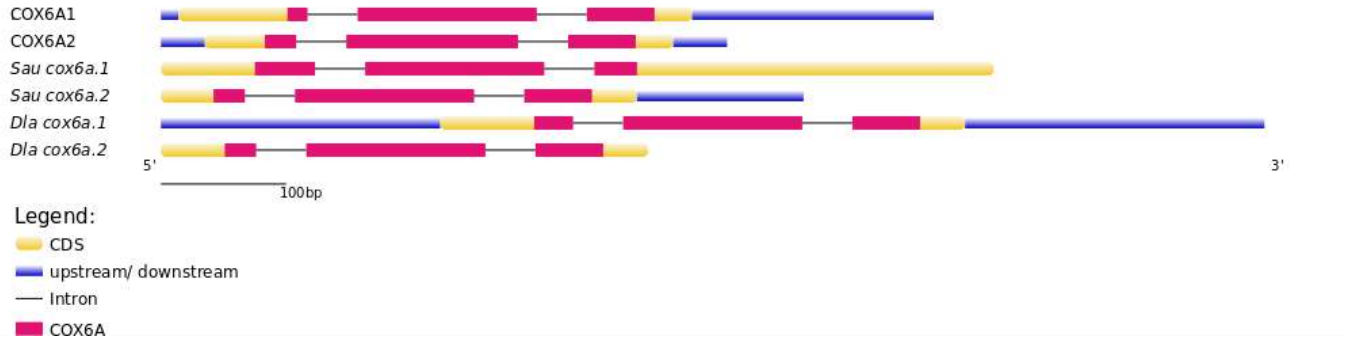


Figure S2.34. Gene structure and protein domains for *cox6a* gene family in gilthead seabream (Sau), European seabass (Dla) and human (capital letters).

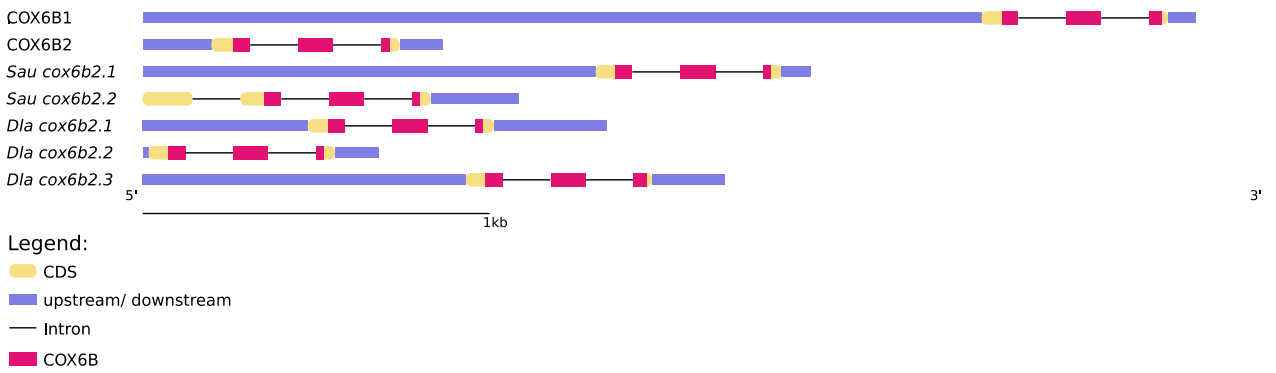


Figure S2.35. Gene structure and protein domains for *cox6b* gene family in gilthead seabream (Sau), European seabass (Dla) and human (capital letters).

Figure S2.36. Gene structure and protein domains for *cox6c* gene family in gilthead seabream (Sau), European seabass (Dla) and human (capital letters).

Figure S2.37. Gene structure and protein domains for *cox7a* gene family in gilthead seabream (Sau), European seabass (Dla) and human (capital letters).

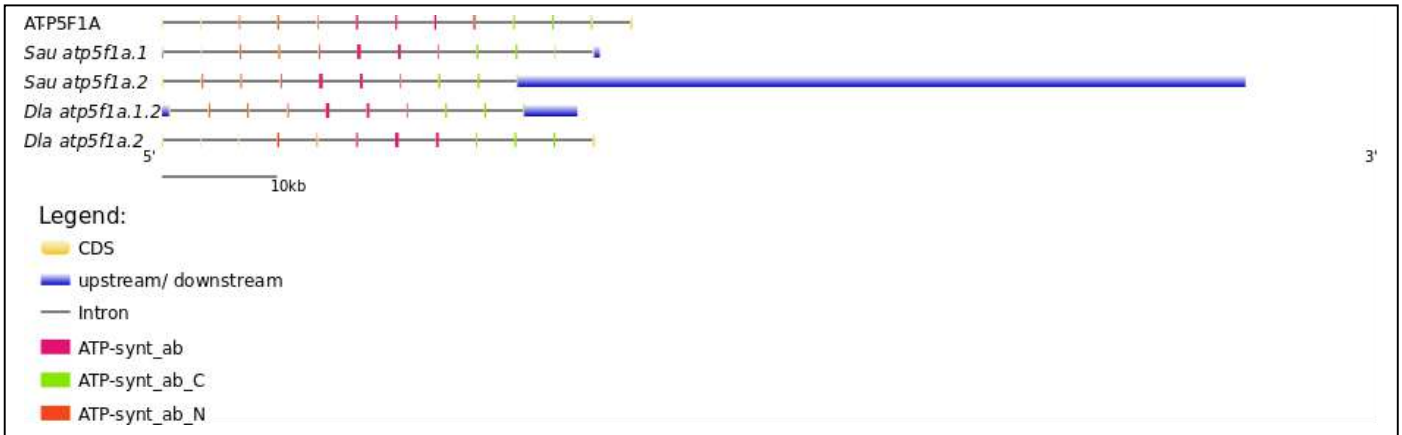


Figure S2.38. Gene structure and protein domains for *atp5f1a* gene family in gilthead seabream (Sau), European seabass (Dla) and human (capital letters).

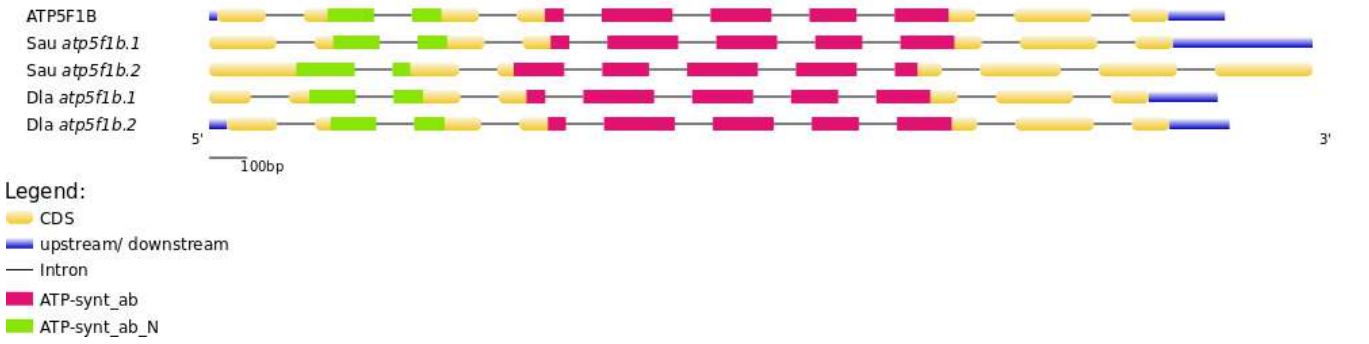


Figure S2.39. Gene structure and protein domains for *atp5f1b* gene family in gilthead seabream (Sau), European seabass (Dla) and human (capital letters).

Figure S2.40. Gene structure and protein domains for *atp5if1* gene family in gilthead seabream (Sau), European seabass (Dla) and human (capital letters).

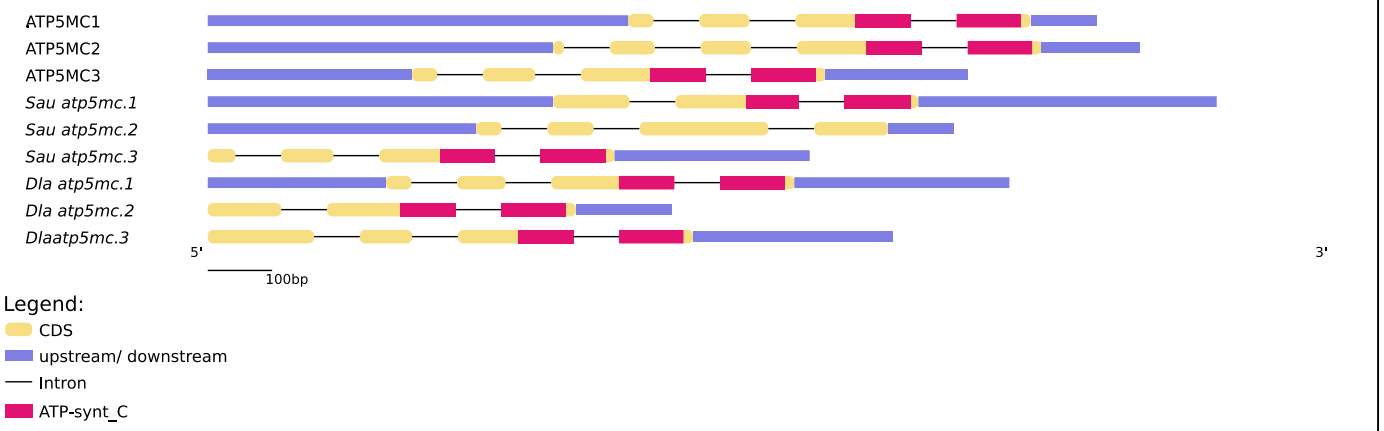


Figure S2.41. Gene structure and protein domains for *atp5mc* gene family in gilthead seabream (Sau), European seabass (Dla) and human (capital letters).

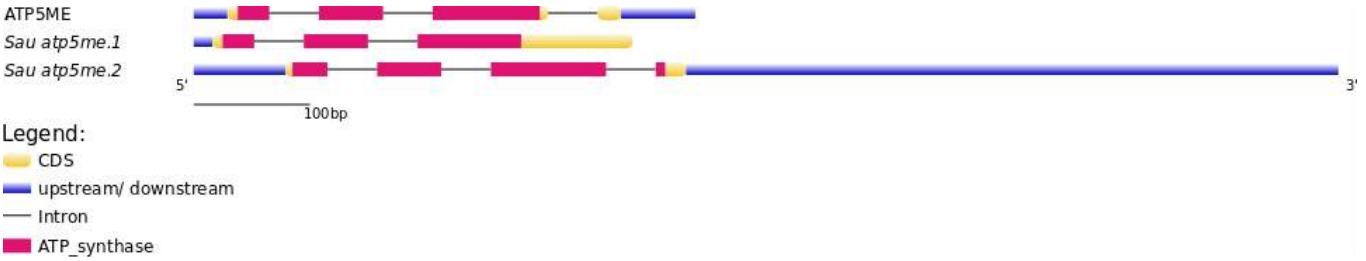


Figure S2.42. Gene structure and protein domains for *atp5me* gene family in gilthead seabream (Sau), European seabass (Dla) and human (capital letters).

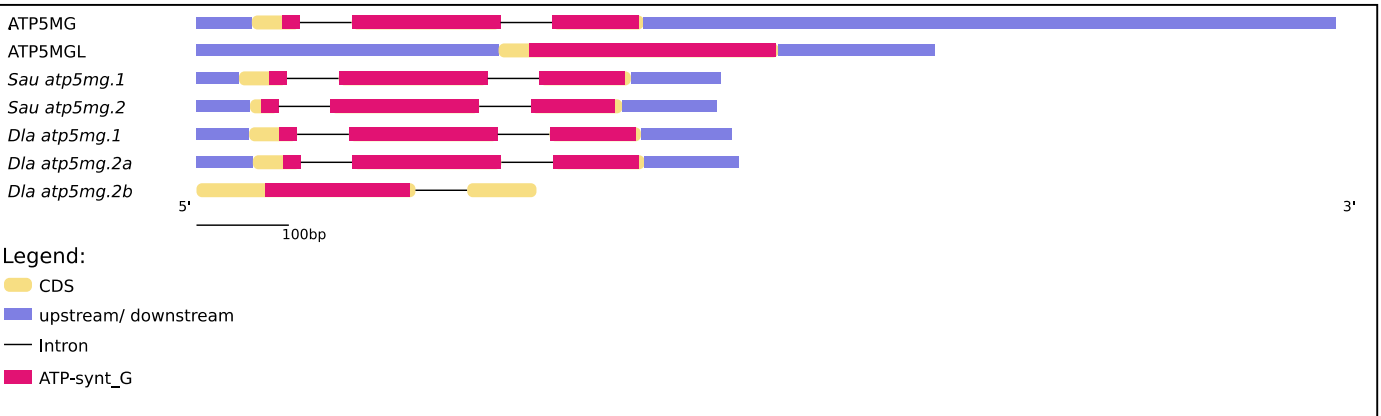


Figure S2.43. Gene structure and protein domains for *atp5mg* gene family in gilthead seabream (Sau), European seabass (Dla) and human (capital letters).

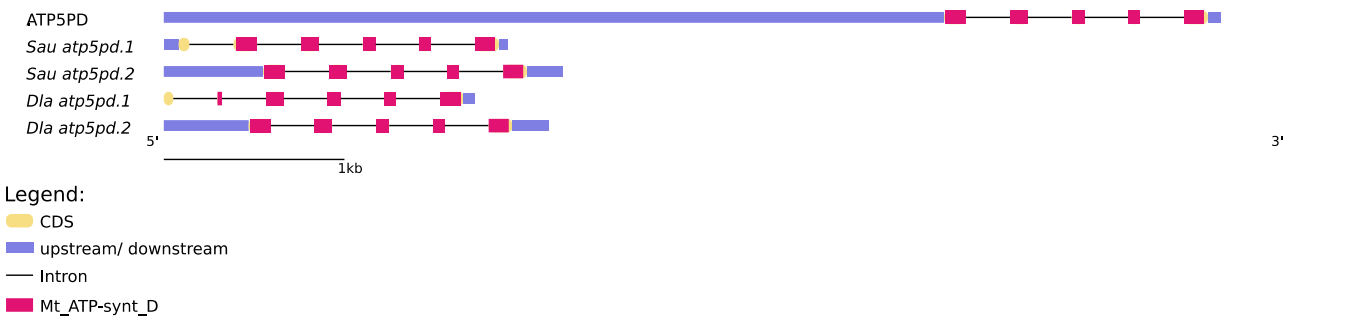


Figure S2.44. Gene structure and protein domains for *atp5pd* gene family in gilthead seabream (Sau), European seabass (Dla) and human (capital letters).

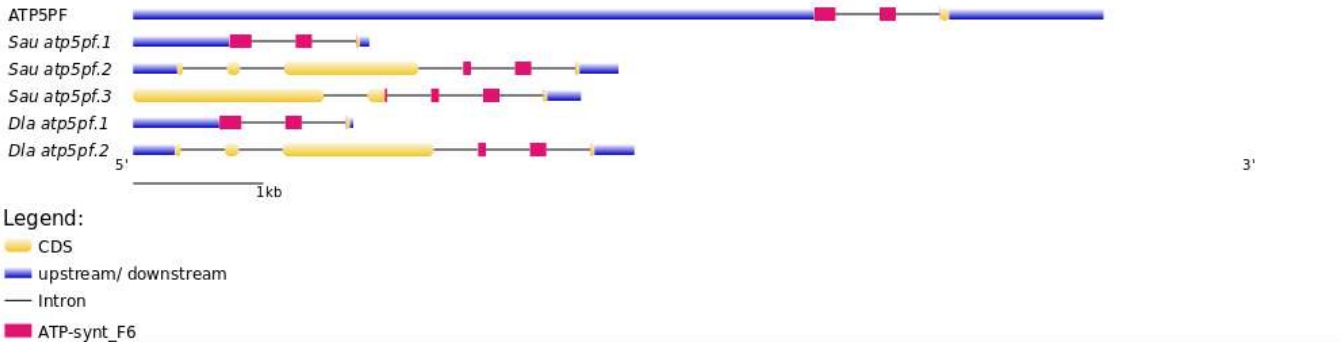


Figure S2.45. Gene structure and protein domains for *atp5pf* gene family in gilthead seabream (Sau), European seabass (Dla) and human (capital letters).

Chapter 3

Table S3.1. Number, location, and selection mode of polymorphic sites for gilthead seabream OXPHOS paralogs. Dashes used when there were not enough unique sequences existed for selection analysis.

Gene name	Complex	Purifying	Diversifying	Neutral	Synonymous	Missense	5UTR	3UTR	Splice region
ndufs1.1	CI	19	1	18	24	16		3	2
ndufs1.2	CI	1		3	3	1			
ndufs8.1	CI	2		7	5	4			
ndufs8.2	CI	-	-	-		1		3	
ndufv1.1	CI	-	-	-	1				
ndufv1.2	CI	-	-	-					
sdhb.1	CII			3	1	2		1	
sdhb.2	CII	3		1	4			1	1
uqcr11.1	CIII	1		2	3		1	8	
uqcr11.2	CIII	-	-	-	1				
uqcrc2.1	CIII	-	-	-		1	2	2	
uqcrc2.2	CIII	1		4	2	3		1	
uqcrfs1.1	CIII	-	-	-	1			3	
uqcrfs1.2	CIII	-	-	-					
uqcrh.1	CIII	-	-	-			1		
uqcrh.2	CIII	-	-	-	1			2	
cox4i.1	CIV	-	-	-				3	2
cox4i.2	CIV	1		2	2	1		4	
cox5a.1	CIV	2		3	5				
cox5a.2	CIV	-	-	-		1		3	
cox5b.1	CIV	-	-	-	1				
cox5b.2	CIV	7		1	7	1	2	16	1
cox6a.1	CIV	1		2	2	1		7	
cox6a.2	CIV	-	-	-					
cox6b2.1	CIV	3		2	4	1	4	5	
cox6b2.2	CIV	1		1	1	1		1	
cox6c.1	CIV	-	-	-			3	2	1
cox6c.2	CIV	-	-	-		1	3		1
cox7a1.1	CIV	3		2	4			5	
cox7a1.2	CIV	1		1	1	1	1		
cox7a2.1	CIV	-	-	-					
cox7a2l.1	CIV	-	-	-	1			2	
cox7a2l.2	CIV	-	-	-	1		1		
cox7a2l.3	CIV	1		3	4			2	1
atp5f1a.1	CV	4		3	4	3		11	
atp5f1a.2	CV	6		3	8	1		2	1
atp5f1b.1	CV	10	1	5	13	4		6	1
atp5f1b.2	CV	2			2		2		
atp5if1.1	CV	-	-	-		1	5	1	
atp5if1.2	CV	-	-	-	1			2	

atp5mc.1	CV			2		2		1	1
atp5mc.2	CV	-	-	-					
atp5mc.3	CV	-	-	-	1				
atp5me.1	CV	-	-	-					
atp5me.2	CV	3		2	5		4	36	
atp5mg.1	CV	-	-	-			1	2	1
atp5mg.2	CV	-	-	-					
atp5pd.1	CV	-	-	-		1			
atp5pd.2	CV	3		4	7		7	5	
atp5pf.1	CV	2		4	5	1	2	1	1
atp5pf.2	CV	24	7	20	38	14	3	5	
atp5pf.3	CV	13	9	31	19	36		2	

Table S3.2. Number, location, and selection mode of polymorphic sites for European seabass OXPHOS paralogs. Dashes used when there were not enough unique sequences existed for selection analysis.

Gene name	Complex	Purifying	Diversifying	Neutral	Synonymous	Missense	5UTR	3UTR	Splice region
ndufs1.1	CI	5		7	10	3			
ndufs1.2	CI	5		3	7	1		2	
ndufs8.1	CI	1		1	2			2	1
ndufs8.2	CI	2		8	8	2		2	
ndufv1.1	CI		-	-					
ndufv1.2	CI	5		2	6	1		3	
sdhb.1	CII			2	2				
sdhb.2	CII				1				
uqcr11.1	CIII				1		3	4	
uqcr11.2	CIII		-	-					
uqcrc2.1	CIII	2		2	3	1			
uqcrc2.2	CIII			2	1	1			
uqcrfs1.1	CIII		-	-					
uqcrfs1.2	CIII	2		5	5	2			
uqcrh.1	CIII		-	-					
uqcrh.2	CIII			5	3	2		4	
cox4i.1	CIV	1		1	1	1	1		
cox4i.2	CIV	1		1	2			4	
cox5a.1	CIV	1		1	2			1	
cox5a.2	CIV				1				
cox5b.1	CIV			5	3	2	1	1	
cox5b.2	CIV	1		1	2	1			
cox6a.1	CIV				1			1	
cox6a.2	CIV			1	1			6	
cox6b2.1	CIV	-	-	-				3	
cox6b2.2	CIV	1		1	1	1			
cox6b2.3	CIV					2		1	
cox7a1.1	CIV	1			1				

cox7a1.2	CIV			3	1	2		5	
cox7a2.1a	CIV					1	2		
cox7a2.1b	CIV	-	-	-					
cox7a2l.1	CIV	-	-	-					
cox7a2l.3	CIV				1			1	
atp5f1a.1	CV	3		5	7	1			
atp5f1a.2	CV	1		3	4			3	
atp5f1b.1	CV	-	-	-					
atp5f1b.2	CV	5		2	7			3	
atp5if1.1	CV	-	-	-			3	1	
atp5if1.2	CV	-	-	-				3	
atp5mc.1	CV	1		2	3			3	
atp5mc.2	CV	-	-	-					
atp5mc.3	CV				1			2	
atp5mg.1	CV				1			1	
atp5mg.2a	CV				1		1		
atp5mg.2b	CV			3	2	1			
atp5pd.1	CV	-	-	-					
atp5pd.2	CV	-	-	-				2	
atp5pf.1	CV			2	1	1			
atp5pf.2	CV	2		3	3	2	4	2	
atp5pf.3	CV			2		2			

Chapter 4

Table S4.1. Gilthead seabream OXPHOS DEGs between first feeding (FF) and flexion (FL) stages

Gene ID	logFC	logCPM	PValue	FDR	neglog10FDR	DEG	gene_name	assigned_name	OXPHOS complex
ENSSAUG00010023201	-0.60838	6.76655	0.003721	0.015471225	1.810475294	**DOWN**	NDUFA11	*ndufa11*	CI
ENSSAUG00010018631	-0.65729	6.958678	0.001683	0.007861031	2.104520501	**DOWN**	NDUFA12	*ndufa12*	CI
ENSSAUG00010020916	-0.8491	6.394953	6.05E-05	0.000435645	3.360867736	**DOWN**	NDUFA3	*ndufa3*	CI
ENSSAUG00010007465	-0.79699	6.61397	0.000155	0.000996071	3.001709531	**DOWN**	NDUFA7	*ndufa7*	CI
ENSSAUG00010023733	-0.6547	6.822671	0.001791	0.008293791	2.081246903	**DOWN**	NDUFB10	*ndufb10*	CI
ENSSAUG00010013498	-0.677	7.401605	0.001149	0.005676951	2.245884879	**DOWN**	NDUFB2	*ndufb2*	CI
ENSSAUG00010025848	-0.71878	6.686957	0.000629	0.003367447	2.472699217	**DOWN**	NDUFB3	*ndufb3*	CI
ENSSAUG00010018909	-0.60662	7.081087	0.00367	0.015275237	1.816012039	**DOWN**	NDUFB7	*ndufb7*	CI
ENSSAUG00010012757	-0.73931	6.739576	0.000434	0.002440533	2.612515383	**DOWN**	NDUFC1	*ndufc1*	CI
ENSSAUG00010000182	1.287757	7.95149	8.24E-10	1.75E-08	7.755780885	**UP**	NDUFC2	*ndufc2*	CI
ENSSAUG00010021984	-0.71608	6.959896	0.000627	0.003359146	2.47377114	**DOWN**	ndufs4	*ndufs4*	CI
ENSSAUG00010018928	-1.62518	4.752917	1.69E-12	5.56E-11	10.25479479	**DOWN**	NDUFV1	*ndufv1.1*	CI
ENSSAUG00010000613	-0.75775	6.641202	0.000319	0.001867339	2.728776785	**DOWN**	SDHB	*sdhb.1*	CII
ENSSAUG00010008991	-0.60111	7.503902	0.003826	0.015840092	1.800242307	**DOWN**	SDHC	*sdhc*	CII
ENSSAUG00010003983	-0.6081	8.083334	0.003296	0.013959407	1.855133027	**DOWN**	UQCRB	*uqcrb*	CIII
ENSSAUG00010009981	-0.79974	7.113707	0.000132	0.000865331	3.062817713	**DOWN**	UQCRH	*uqcrh.1*	CIII
ENSSAUG00010018590	-0.61284	7.932847	0.003091	0.013220967	1.878736787	**DOWN**	UQCRQ	*uqcrq*	CIII
ENSSAUG00010008794	-0.73361	8.80013	0.000382	0.002189293	2.659696197	**DOWN**	Cox4a	*cox4i.1*	CIV
ENSSAUG00010018002	1.744559	0.209029	0.00179	0.008292567	2.081311003	**UP**	Cox4b	*cox4i.2*	CIV
ENSSAUG00010022937	-1.15388	7.943167	3.36E-08	5.26E-07	6.279155231	**DOWN**	COX5a2	*cox5a.1*	CIV
ENSSAUG00010021488	-3.19077	2.26313	1.57E-19	1.16E-17	16.93542513	**DOWN**	COX5b1	*cox5b.1*	CIV

ENSSAUG00010020219	-1.19595	5.211188	6.66E-08	9.77E-07	6.010095673	**DOWN**	cox6b2	*cox6b2.1*	CIV
ENSSAUG00010014837	-0.68203	8.423963	0.000969	0.004895556	2.310198009	**DOWN**	cox6b2	*cox6b2.2*	CIV
ENSSAUG00010005990	-0.73822	4.943381	0.00086	0.004411896	2.355374712	**DOWN**	COX6c1	*cox6c.1*	CIV
ENSSAUG00010005612	-1.44712	5.252524	7.83E-11	1.98E-09	8.703058983	**DOWN**	COX6c1	*cox6c.2*	CIV
ENSSAUG00010007755	-0.68023	7.170087	0.00112	0.005551347	2.2556016	**DOWN**	COX7a2	*cox7a1.2*	CIV
ENSSAUG00010023048	1.283496	2.473303	1.37E-05	0.000117078	3.93152438	**UP**	-	*cox7a2l.2*	CIV
ENSSAUG00010006458	-0.74144	8.033359	0.000349	0.002023832	2.693825598	**DOWN**	-	*cox7b*	CIV
ENSSAUG00010024581	-0.7192	8.002034	0.000522	0.002867575	2.542485263	**DOWN**	COX7c	*cox7c*	CIV
ENSSAUG00010003151	-1.11451	5.732481	2.68E-07	3.47E-06	5.460127927	**DOWN**	atp5if1b	*atp5if1.2*	CV
ENSSAUG00010022955	-0.91205	5.995557	1.99E-05	0.000163118	3.787498189	**DOWN**	atp5mea	*atp5me.1*	CV
ENSSAUG00010012064	-0.68817	7.378575	0.000954	0.004826685	2.316351061	**DOWN**	-	*atp5mj*	CV
ENSSAUG00010001316	0.857325	6.760232	4.68E-05	0.000348307	3.458037353	**UP**	si:ch211-140m22.7	*atp5pf.2*	CV
ENSSAUG00010012045	1.018596	5.146167	4.24E-06	4.16E-05	4.380951318	**UP**	-	*atp5pf.3*	CV

Table S4.2. Gilthead seabream OXPHOS DEGs between flexion (FL) and mid metamorphosis (MM) stages

Gene ID	logFC	logCPM	PValue	FDR	neglog10FDR	DEG	gene_name	assigned_name	OXPHOS complex
ENSSAUG00010018219	0.771216	6.717771	0.000235	0.001104	2.957227	**UP**	NDUFA1	*ndufa1*	CI

ENSSAUG00010027521	0.644622	7.907949	0.001846	0.006786	2.168374	**UP**	ndufa10	*ndufa10*	CI
ENSSAUG00010023201	0.652024	6.32463	0.001976	0.007202	2.142543	**UP**	NDUFA11	*ndufa11*	CI
ENSSAUG00010018631	1.099147	6.775862	1.81E-07	1.69E-06	5.772288	**UP**	NDUFA12	*ndufa12*	CI
ENSSAUG00010017787	0.905105	7.49443	1.40E-05	8.95E-05	4.048269	**UP**	NDUFA13	*ndufa13*	CI
ENSSAUG00010002150	0.813611	6.381413	0.000115	0.000582	3.235436	**UP**	NDUFA2	*ndufa2*	CI
ENSSAUG00010020916	0.959625	5.997943	6.58E-06	4.49E-05	4.347421	**UP**	NDUFA3	*ndufa3*	CI
ENSSAUG00010007020	0.882536	6.63665	2.70E-05	0.00016	3.795327	**UP**	NDUFA6	*ndufa6*	CI
ENSSAUG00010007465	1.137219	6.36987	8.19E-08	8.26E-07	6.083017	**UP**	NDUFA7	*ndufa7*	CI
ENSSAUG00010008168	0.85443	7.624379	3.99E-05	0.000226	3.645654	**UP**	NDUFA9	*ndufa9*	CI
ENSSAUG00010021118	0.618377	7.555135	0.002877	0.009981	2.000823	**UP**	ndufab1b	*ndufab1*	CI
ENSSAUG00010023733	1.125042	6.659024	9.81E-08	9.75E-07	6.011127	**UP**	NDUFB10	*ndufb10*	CI
ENSSAUG00010012559	0.686491	6.381765	0.001116	0.004378	2.358699	**UP**	NDUFA11	*ndufb11*	CI
ENSSAUG00010025848	0.87938	6.320167	3.17E-05	0.000184	3.734938	**UP**	NDUFB3	*ndufb3*	CI
ENSSAUG00010027131	0.88502	6.776047	2.48E-05	0.000148	3.828414	**UP**	NDUFB5	*ndufb5*	CI
ENSSAUG00010018717	0.814084	6.024703	0.000127	0.000635	3.197111	**UP**	NDUFB6	*ndufb6*	CI
ENSSAUG00010018909	0.956861	6.833796	5.17E-06	3.63E-05	4.440299	**UP**	NDUFB7	*ndufb7*	CI
ENSSAUG00010023063	0.668573	6.149064	0.001572	0.005891	2.229834	**UP**	ndufb8	*ndufb8*	CI
ENSSAUG00010000482	1.133814	7.394714	6.15E-08	6.36E-07	6.196223	**UP**	NDUFB9	*ndufb9*	CI
ENSSAUG00010012757	0.980987	6.426523	3.43E-06	2.50E-05	4.602244	**UP**	NDUFC1	*ndufc1*	CI
ENSSAUG00010021399	0.942584	3.241936	0.000172	0.000834	3.079074	**UP**	NDUFS1	*ndufs1.2*	CI
ENSSAUG00010004194	0.849018	7.342747	4.64E-05	0.000259	3.586628	**UP**	ndufs3	*ndufs3*	CI
ENSSAUG00010021984	1.100184	6.741542	1.78E-07	1.67E-06	5.777538	**UP**	ndufs4	*ndufs4*	CI
ENSSAUG00010020202	0.917909	6.855913	1.21E-05	7.80E-05	4.107957	**UP**	NDUFS5	*ndufs5*	CI
ENSSAUG00010008813	0.765986	5.952428	0.000315	0.001425	2.846034	**UP**	NDUFS6	*ndufs6*	CI
ENSSAUG00010009712	1.226513	0.884088	0.001789	0.006604	2.180217	**UP**	NDUFS8	*ndufs8.2*	CI
ENSSAUG00010012337	0.661651	7.031377	0.001516	0.005698	2.244293	**UP**	ndufv2	*ndufv2*	CI
ENSSAUG00010001379	0.740143	5.913822	0.000503	0.002155	2.666553	**UP**	NDUFV3	*ndufv3*	CI
ENSSAUG00010000613	0.70026	6.136861	0.000936	0.003732	2.428073	**UP**	SDHB	*sdhb.1*	CII

ENSSAUG00010006809	0.662778	7.955723	0.001365	0.005194	2.284497	**UP**	Cyc1	*cyc1*	CIII
ENSSAUG00010021526	0.694521	7.112121	0.000865	0.003476	2.458954	**UP**	UQCR10	*uqcr10*	CIII
ENSSAUG00010018172	0.671483	6.78801	0.001333	0.005088	2.293466	**UP**	UQCR11-B	*uqcr11.1*	CIII
ENSSAUG00010003983	0.665923	7.650416	0.001325	0.005061	2.295766	**UP**	UQCRB	*uqcrb*	CIII
ENSSAUG00010010721	0.702831	8.617001	0.000662	0.002753	2.560168	**UP**	uqcrc1	*uqcrc1*	CIII
ENSSAUG00010022860	0.757147	7.393519	0.000274	0.001261	2.899274	**UP**	uqcrc2a	*uqcrc2.2*	CIII
ENSSAUG00010000160	0.653075	7.698811	0.001632	0.006085	2.215767	**UP**	UQCRFS1	*uqcrfs1.1*	CIII
ENSSAUG00010025510	1.152369	7.351788	3.82E-08	4.14E-07	6.383079	**UP**	UQCRFS1	*uqcrfs1.2*	CIII
ENSSAUG00010009981	0.611918	6.529101	0.00356	0.012038	1.919428	**UP**	UQCRH	*uqcrh.1*	CIII
ENSSAUG00010018590	0.950094	7.677698	5.05E-06	3.55E-05	4.450039	**UP**	UQCRQ	*uqcrq*	CIII
ENSSAUG00010008794	0.869992	8.418707	2.66E-05	0.000158	3.801778	**UP**	Cox4a	*cox4i.1*	CIV
ENSSAUG00010018002	4.466957	3.734163	7.38E-50	1.59E-47	46.79971	**UP**	Cox4b	*cox4i.2*	CIV
ENSSAUG00010022937	1.53874	7.750779	2.95E-13	6.83E-12	11.16557	**UP**	COX5a2	*cox5a.1*	CIV
ENSSAUG00010021488	6.698974	5.151455	9.58E-96	7.18E-93	92.144	**UP**	COX5b1	*cox5b.1*	CIV
ENSSAUG00010019745	1.281683	8.712407	7.98E-10	1.14E-08	7.94381	**UP**	COX6a1	*cox6a1*	CIV
ENSSAUG00010024406	10.81198	2.166252	3.03E-38	4.20E-36	35.37663	**UP**	COX6a2	*cox6a2*	CIV
ENSSAUG00010020219	1.00298	4.609326	8.14E-06	5.46E-05	4.262502	**UP**	cox6b2	*cox6b2.1*	CIV
ENSSAUG00010014837	0.886065	8.084874	1.95E-05	0.00012	3.919706	**UP**	cox6b2	*cox6b2.2*	CIV
ENSSAUG00010005990	1.160463	4.749528	2.27E-07	2.09E-06	5.678988	**UP**	COX6c1	*cox6c.1*	CIV
ENSSAUG00010005612	1.224155	4.621897	6.04E-08	6.25E-07	6.203904	**UP**	COX6c1	*cox6c.2*	CIV
ENSSAUG00010002730	6.441541	3.678219	2.28E-68	8.09E-66	65.0918	**UP**	COX7a1	*cox7a1.1*	CIV
ENSSAUG00010007755	0.746639	6.742953	0.000367	0.001631	2.787562	**UP**	COX7a2	*cox7a1.2*	CIV
ENSSAUG00010026799	0.993402	6.352207	2.66E-06	1.99E-05	4.701727	**UP**	cox7a2l	*cox7a2l.3*	CIV
ENSSAUG00010006458	1.211318	7.876359	6.83E-09	8.38E-08	7.076527	**UP**	-	*cox7b*	CIV
ENSSAUG00010024581	0.914044	7.658047	1.13E-05	7.33E-05	4.13499	**UP**	COX7c	*cox7c*	CIV
ENSSAUG00010021868	6.975317	4.623382	1.48E-89	9.54E-87	86.02036	**UP**	COX8b	*cox8a*	CIV
ENSSAUG00010001552	0.645129	9.579008	0.001713	0.00635	2.197239	**UP**	atp5fa1	*atp5f1a.1*	CV
ENSSAUG00010008214	0.980939	10.55919	2.03E-06	1.55E-05	4.808315	**UP**	atp5fa1	*atp5f1a.2*	CV

ENSSAUG00010007970	0.830788	11.54342	5.43E-05	0.000297	3.527221	**UP**	atp5f1b	*atp5f1b.1*	CV
ENSSAUG00010022681	0.76334	9.174677	0.000214	0.001017	2.99278	**UP**	ATP5C1	*atp5f1c*	CV
ENSSAUG00010005650	1.05242	8.46095	4.06E-07	3.57E-06	5.446873	**UP**	ATP5D	*atp5f1d*	CV
ENSSAUG00010011548	0.803622	8.136408	0.000105	0.000538	3.269095	**UP**	ATP5E	*atp5f1e*	CV
ENSSAUG00010020382	1.666153	6.04345	1.64E-14	4.46E-13	12.35099	**UP**	atp5if1a	*atp5if1.1*	CV
ENSSAUG00010003151	2.123771	6.020761	7.07E-22	3.68E-20	19.43433	**UP**	atp5if1b	*atp5if1.2*	CV
ENSSAUG00010024766	0.764596	9.594603	0.000207	0.000984	3.007098	**UP**	ATP5G1	*atp5mc.1*	CV
ENSSAUG00010003390	0.86244	10.05665	2.88E-05	0.000169	3.771379	**UP**	ATP5G3	*atp5mc.3*	CV
ENSSAUG00010011394	1.319305	7.364447	3.54E-10	5.29E-09	8.276236	**UP**	ATP5L	*atp5mg.1*	CV
ENSSAUG00010012064	0.815853	6.990301	9.56E-05	0.000496	3.304914	**UP**	-	*atp5mj*	CV
ENSSAUG00010005272	0.666573	7.723071	0.001303	0.004984	2.302439	**UP**	atp5md	*atp5mk*	CV
ENSSAUG00010001248	0.634952	8.327361	0.002107	0.007619	2.118115	**UP**	atp5pb	*atp5pb*	CV
ENSSAUG00010003235	0.90566	8.382742	1.24E-05	8.02E-05	4.096036	**UP**	atp5pd	*atp5pd.1*	CV
ENSSAUG00010027145	1.799595	6.611352	5.30E-17	1.84E-15	14.73635	**UP**	atp5pd	*atp5pd.2*	CV
ENSSAUG00010012045	-0.66802	4.805868	0.002729	0.009534	2.020742	**DOWN**	-	*atp5pf.3*	CV

Table S4.3 Gilthead seabream OXPHOS DEGs between first feeding (FF) and mid metamorphosis (MM) stages

Gene ID	logFC	logCPM	PValue	FDR	neglog10FDR	DEG	gene_name	assigned_name	OXPHOS complex
ENSSAUG00010000482	0.659065	7.569029	0.001511	0.004871	2.312417	**UP**	NDUFB9	ndufb9	CI
ENSSAUG00010000182	1.016243	7.292003	1.21E-06	7.49E-06	5.125621	**UP**	NDUFC2	ndufc2	CI
ENSSAUG00010021399	0.711764	3.338327	0.004572	0.013148	1.88113	**UP**	NDUFS1	ndufs1.2	CI
ENSSAUG00010018928	-1.10834	4.433457	1.68E-06	1.01E-05	4.997089	**DOWN**	NDUFV1	ndufv1.1	CI
ENSSAUG00010000160	0.865791	7.627671	3.21E-05	0.00015	3.823841	**UP**	UQCRFS1	uqcrfs1.1	CI
ENSSAUG00010018002	6.218516	3.702216	4.39E-63	6.06E-61	60.21757	**UP**	Cox4b	cox4i.2	CIV

ENSSAUG00010021488	3.521233	5.269738	7.54E-46	6.27E-44	43.20265	**UP**	COX5b1	cox5b.1	CIV
ENSSAUG00010019745	0.775891	8.886962	0.000172	0.000691	3.160642	**UP**	COX6a1	cox6a1	CIV
ENSSAUG00010024406	6.280244	2.212241	5.55E-37	3.51E-35	34.4548	**UP**	COX6a2	cox6a2	CIV
ENSSAUG00010002730	5.146639	3.717386	1.55E-54	1.67E-52	51.77787	**UP**	COX7a1	cox7a1.1	CIV
ENSSAUG00010021868	6.166504	4.643526	9.22E-79	2.04E-76	75.68968	**UP**	COX8b	cox8a	CIV
ENSSAUG00010008214	0.67821	10.67607	0.000967	0.003264	2.486215	**UP**	atp5fa1	atp5f1a.2	CV
ENSSAUG00010022681	0.729034	9.195353	0.000407	0.001499	2.824085	**UP**	ATP5C1	atp5f1c	CV
ENSSAUG00010005650	0.797383	8.556931	0.000116	0.00048	3.318376	**UP**	ATP5D	atp5f1d	CV
ENSSAUG00010020382	1.32508	6.141802	7.25E-10	7.42E-09	8.129576	**UP**	atp5if1a	atp5if1.1	CV
ENSSAUG00010003151	1.020945	6.310654	1.52E-06	9.17E-06	5.037462	**UP**	atp5if1b	atp5if1.2	CV
ENSSAUG00010024766	0.6881	9.631272	0.00083	0.002837	2.547087	**UP**	ATP5G1	atp5mc.1	CV
ENSSAUG00010011394	1.014957	7.466486	1.20E-06	7.42E-06	5.129429	**UP**	ATP5L	atp5mg.1	CV
ENSSAUG00010003235	0.686751	8.470709	0.000888	0.003017	2.520477	**UP**	atp5pd	atp5pd.1	CV
ENSSAUG00010027145	1.676791	6.648144	5.15E-15	9.73E-14	13.01191	**UP**	atp5pd	atp5pd.2	CV

Table S4.4 European seabass OXPHOS DEGs between first feeding (FF) and flexion (FL) stages

Gene ID	logFC	logCPM	PValue	FDR	neglog10FDR	DEG	gene_name	assigned_name	OXPHOS complex
ENSDLAG00005001413	-0.70445	8.256581	0.000678	0.002999	2.522988561	**DOWN**	ndufab1b	ndufab1	CI
ENSDLAG00005002523	-1.48555	6.126992	1.03E-11	2.35E-10	9.629284241	**DOWN**	ndufs6	ndufs6	CI
ENSDLAG00005002836	-1.04983	6.558052	8.13E-07	7.49E-06	5.125725335	**DOWN**	ndufa3	ndufa3	CI
ENSDLAG00005004491	-0.68613	3.300529	0.009363	0.028908	1.538987047	**DOWN**	NDUFV1	ndufv1.1	CI
ENSDLAG00005006810	-2.02051	5.753171	3.58E-19	2.24E-17	16.65050077	**DOWN**	-	ndufc1	CI
ENSDLAG00005008059	-0.85669	8.455135	3.65E-05	0.000225	3.64816344	**DOWN**	ndufa10	ndufa10	CI
ENSDLAG00005008294	-1.24766	6.486466	5.57E-09	7.98E-08	7.09784577	**DOWN**	ndufb6	ndufb6	CI

ENSDLAG00005008847	-1.30567	6.526013	1.08E-09	1.76E-08	7.75396646	**DOWN**	ndufa5	ndufa5	CI
ENSDLAG00005011028	-1.26714	7.133062	2.17E-09	3.35E-08	7.47446084	**DOWN**	ndufb10	ndufb10	CI
ENSDLAG00005013355	-1.16739	8.363301	2.27E-08	2.87E-07	6.542476368	**DOWN**	ndufv1	ndufv1.2	CI
ENSDLAG00005014329	-1.90579	6.858027	1.95E-18	1.13E-16	15.94594856	**DOWN**	ndufs4	ndufs4	CI
ENSDLAG00005015202	-1.40378	6.432511	7.99E-11	1.59E-09	8.798478693	**DOWN**	ndufs5	ndufs5	CI
ENSDLAG00005015610	-1.45502	6.826454	9.89E-12	2.26E-10	9.646257323	**DOWN**	ndufa6	ndufa6	CI
ENSDLAG00005015965	-1.7492	6.637423	8.49E-16	3.59E-14	13.44444702	**DOWN**	ndufa1	ndufa1	CI
ENSDLAG00005018458	-2.23475	5.991291	4.93E-23	4.50E-21	20.34663221	**DOWN**	NDUB4	ndufb4	CI
ENSDLAG00005019348	-1.30979	6.542041	9.48E-10	1.57E-08	7.805060546	**DOWN**	ndufb11	ndufb11	CI
ENSDLAG00005019817	-1.12811	6.791651	1.06E-07	1.18E-06	5.927707281	**DOWN**	ndufb2	ndufb2	CI
ENSDLAG00005019820	-0.97414	8.203083	2.90E-06	2.34E-05	4.630804753	**DOWN**	ndufs3	ndufs3	CI
ENSDLAG00005020963	-1.6498	7.072826	1.28E-14	4.48E-13	12.34849083	**DOWN**	ndufb7	ndufb7	CI
ENSDLAG00005021448	-1.44661	7.732764	7.47E-12	1.75E-10	9.757829178	**DOWN**	ndufa11	ndufa11	CI
ENSDLAG00005022483	-1.19363	8.198219	1.15E-08	1.56E-07	6.806493613	**DOWN**	ndufa8	ndufa8	CI
ENSDLAG00005022603	-1.54095	7.474925	3.93E-13	1.12E-11	10.95124919	**DOWN**	ndufb9	ndufb9	CI
ENSDLAG00005022657	-1.02766	8.073506	8.42E-07	7.71E-06	5.112695273	**DOWN**	ndufs7	ndufs7	CI
ENSDLAG00005022803	-0.77658	5.896394	0.000321	0.001545	2.811189383	**DOWN**	ndufb8	ndufb8	CI
ENSDLAG00005023654	-1.72051	5.899412	8.43E-15	3.01E-13	12.52158278	**DOWN**	NDUFB1	ndufb1	CI
ENSDLAG00005024193	-0.87087	7.405356	3.19E-05	0.0002	3.699065199	**DOWN**	ndufc2	ndufc2	CI
ENSDLAG00005024263	-1.02768	5.392867	3.45E-06	2.74E-05	4.562698834	**DOWN**	ndufb3	ndufb3	CI
ENSDLAG00005024309	-0.83068	8.745731	6.04E-05	0.000351	3.454096642	**DOWN**	ndufs2	ndufs2	CI
ENSDLAG00005024691	-1.42753	7.517111	1.51E-11	3.36E-10	9.474244881	**DOWN**	ndufa12	ndufa12	CI
ENSDLAG00005024977	-1.85277	6.79397	1.57E-17	8.14E-16	15.08944468	**DOWN**	NDUFA13	ndufa13	CI
ENSDLAG00005025335	-2.05257	5.877343	7.42E-20	4.99E-18	17.30160245	**DOWN**	ndufa7	ndufa7	CI
ENSDLAG00005025864	-0.85785	8.450899	3.56E-05	0.00022	3.656833182	**DOWN**	ndufs1	ndufs1.1	CI
ENSDLAG00005025947	-1.15224	8.04395	3.64E-08	4.46E-07	6.350476635	**DOWN**	ndufv2	ndufv2	CI
ENSDLAG00005026097	-0.93144	7.34769	8.96E-06	6.45E-05	4.190260138	**DOWN**	ndufb5	ndufb5	CI
ENSDLAG00005006491	-0.97705	7.129747	3.46E-06	2.74E-05	4.561917322	**DOWN**	sdhc	sdhc	CII
ENSDLAG00005008624	-0.66892	6.552356	0.001578	0.006253	2.203915991	**DOWN**	SDHB	sdhb.1	CII

ENSDLAG00005016007	-1.93826	7.352673	2.10E-19	1.36E-17	16.86630644	**DOWN**	SDHB	sdhb.2	CII
ENSDLAG00005001473	-2.10543	6.734384	9.98E-22	7.93E-20	19.10055577	**DOWN**	uqcrq	uqcrq	CIII
ENSDLAG00005003877	-1.9005	6.82499	2.63E-18	1.50E-16	15.82281476	**DOWN**	uqcrh	uqcrh.1	CIII
ENSDLAG00005007897	-1.38456	3.348448	1.99E-07	2.10E-06	5.677404691	**DOWN**	-	uqcr11.2	CIII
ENSDLAG00005009892	-0.8792	6.376813	3.69E-05	0.000227	3.643565667	**DOWN**	-	uqcrh.2	CIII
ENSDLAG00005010064	-1.03509	9.129857	6.10E-07	5.76E-06	5.239435087	**DOWN**	cyc1	cyc1	CIII
ENSDLAG00005014897	-1.15164	7.730468	4.02E-08	4.87E-07	6.312863002	**DOWN**	uqcrc2a	uqcrc2.2	CIII
ENSDLAG00005019495	-1.16702	7.313231	3.08E-08	3.82E-07	6.417925835	**DOWN**	uqcrb	uqcrb	CIII
ENSDLAG00005021271	-1.38742	7.951304	4.38E-11	9.05E-10	9.043471941	**DOWN**	UQCRFS1	uqcrfs1.2	CIII
ENSDLAG00005021401	-1.73604	7.813145	3.42E-16	1.52E-14	13.81864111	**DOWN**	uqcr10	uqcr10	CIII
ENSDLAG00005024816	-0.85254	6.66764	5.64E-05	0.00033	3.481215826	**DOWN**	-	uqcr11.1	CIII
ENSDLAG00005001169	-1.18299	8.673532	1.41E-08	1.87E-07	6.729056908	**DOWN**	cox6b2	cox6b2.1	CIV
ENSDLAG00005001205	5.409093	3.312495	5.57E-41	1.88E-38	37.72592172	**UP**	CX6B1	cox6b2.3	CIV
ENSDLAG00005007600	-1.01325	2.082775	0.002237	0.008468	2.072237343	**DOWN**	-	cox7a2l.1	CIV
ENSDLAG00005008900	-2.10076	4.410784	1.38E-17	7.19E-16	15.14310892	**DOWN**	cox6b2	cox6b2.2	CIV
ENSDLAG00005009814	-2.95887	5.18657	1.47E-33	2.98E-31	30.5263622	**DOWN**	cox5b2	cox5b.1	CIV
ENSDLAG00005009958	2.338964	2.065602	2.57E-10	4.73E-09	8.324800984	**UP**	-	cox8a	CIV
ENSDLAG00005011849	-1.41471	8.634956	1.53E-11	3.41E-10	9.467820633	**DOWN**	cox6a1	cox6a.1	CIV
ENSDLAG00005014376	-0.68445	7.23364	0.001075	0.004496	2.347195175	**DOWN**	cox5aa	cox5a.2	CIV
ENSDLAG00005014851	-1.18221	7.360492	2.02E-08	2.59E-07	6.586772729	**DOWN**	COX5B	cox5b.2	CIV
ENSDLAG00005016960	-1.06324	8.970568	3.09E-07	3.13E-06	5.504470677	**DOWN**	cox4i1	cox4i.1	CIV
ENSDLAG00005022694	-1.10311	8.059628	1.30E-07	1.42E-06	5.848129481	**DOWN**	-	cox7a1.2	CIV
ENSDLAG00005023073	-1.3114	8.59349	3.68E-10	6.60E-09	8.180484188	**DOWN**	COX5A	cox5a.1	CIV
ENSDLAG00005023340	-1.08791	7.031331	2.61E-07	2.68E-06	5.572346433	**DOWN**	COX7R	cox7a2l.3	CIV
ENSDLAG00005023940	-0.63061	8.390518	0.002307	0.008707	2.060147381	**DOWN**	-	cox7b	CIV
ENSDLAG00005025866	-1.2104	7.9336	7.79E-09	1.08E-07	6.965481308	**DOWN**	cox7c	cox7c	CIV
ENSDLAG00005000620	-2.20388	6.860376	1.47E-23	1.42E-21	20.84850123	**DOWN**	atp5if1b	atp5if1.2	CV
ENSDLAG00005003781	-0.988	9.060952	1.90E-06	1.60E-05	4.79706307	**DOWN**	atp5pd	atp5pd.1	CV
ENSDLAG00005004647	-0.64911	4.617687	0.004651	0.015897	1.798685648	**DOWN**	atp5if1a	atp5if1.1	CV

ENSDLAG00005006335	-1.78254	3.035939	4.69E-10	8.22E-09	8.084881134	**DOWN**	-	atp5mg.2b	CV
ENSDLAG00005006347	-1.80516	7.141081	4.37E-17	2.14E-15	14.66887518	**DOWN**	atp5l	atp5mg.2a	CV
ENSDLAG00005007485	-1.71922	7.42999	8.13E-16	3.45E-14	13.4621238	**DOWN**	atp5mf	atp5mf	CV
ENSDLAG00005010974	-1.2756	8.022423	1.21E-09	1.95E-08	7.709802314	**DOWN**	-	atp5me	CV
ENSDLAG00005013665	-0.98863	8.764214	1.92E-06	1.62E-05	4.791652255	**DOWN**	atp5f1d	atp5f1d	CV
ENSDLAG00005015257	-1.95292	8.103764	6.57E-20	4.44E-18	17.35229757	**DOWN**	atp5pf	atp5pf.1	CV
ENSDLAG00005016988	-1.05843	8.149599	3.92E-07	3.89E-06	5.410014104	**DOWN**	atp5f1e	atp5f1e	CV
ENSDLAG00005017184	-0.60661	9.004443	0.003267	0.011751	1.929909348	**DOWN**	atp5po	atp5po	CV
ENSDLAG00005021874	-0.92945	9.37781	7.07E-06	5.22E-05	4.282531644	**DOWN**	atp5pb	atp5pb	CV
ENSDLAG00005024092	-1.55396	7.412036	2.64E-13	7.73E-12	11.11178677	**DOWN**	atp5md	atp5mk	CV
ENSDLAG00005024589	-0.89256	10.14188	1.53E-05	0.000105	3.980090439	**DOWN**	atp5mc1	atp5mc.3	CV
ENSDLAG00005025077	-0.70499	10.10942	0.000615	0.002749	2.560901333	**DOWN**	atp5f1c	atp5f1c	CV

Table S4.5 European seabass OXPHOS DEGs between flexion (FL) and mid metamorphosis (MM) stages

Gene ID	logFC	logCPM	PValue	FDR	neglog10FDR	DEG	gene_name	assigned_name	OXPHOS complex
ENSDLAG00005002523	1.305503	5.687814	1.98E-09	5.86E-08	7.232033	**UP**	ndufs6	ndufs6	CI
ENSDLAG00005006810	1.999907	5.428119	7.30E-19	6.84E-17	16.16514	**UP**	-	ndufc1	CI
ENSDLAG00005008059	0.650139	8.016166	0.001683	0.009642	2.015851	**UP**	ndufa10	ndufa10	CI
ENSDLAG00005008294	0.786943	5.869568	0.000228	0.001835	2.736388	**UP**	ndufb6	ndufb6	CI
ENSDLAG00005008847	1.413724	6.294578	4.07E-11	1.60E-09	8.7961	**UP**	ndufa5	ndufa5	CI
ENSDLAG00005009594	0.719364	3.121275	0.00488	0.023062	1.637111	**UP**	NDUFS1	ndufs1.2	CI
ENSDLAG00005011028	1.031958	6.661889	1.01E-06	1.67E-05	4.77673	**UP**	ndufb10	ndufb10	CI
ENSDLAG00005013355	1.035511	7.964034	6.66E-07	1.15E-05	4.937461	**UP**	ndufv1	ndufv1.2	CI
ENSDLAG00005014329	1.634557	6.339475	4.00E-14	2.46E-12	11.60828	**UP**	ndufs4	ndufs4	CI
ENSDLAG00005015202	1.490327	6.186711	5.11E-12	2.29E-10	9.640976	**UP**	ndufs5	ndufs5	CI
ENSDLAG00005015610	1.386018	6.467141	8.21E-11	3.05E-09	8.515792	**UP**	ndufa6	ndufa6	CI
ENSDLAG00005015965	2.12706	6.627692	2.60E-22	3.16E-20	19.50054	**UP**	ndufa1	ndufa1	CI
ENSDLAG00005018458	1.656808	5.225019	1.31E-13	7.39E-12	11.13111	**UP**	NDUB4	ndufb4	CI
ENSDLAG00005019348	1.144852	6.117511	8.31E-08	1.76E-06	5.754157	**UP**	ndufb11	ndufb11	CI
ENSDLAG00005019817	2.071203	7.189475	7.84E-22	9.21E-20	19.03581	**UP**	ndufb2	ndufb2	CI
ENSDLAG00005019820	0.669663	7.699378	0.001247	0.007557	2.121664	**UP**	ndufs3	ndufs3	CI
ENSDLAG00005020963	1.784456	6.866862	8.79E-17	6.93E-15	14.15937	**UP**	ndufb7	ndufb7	CI
ENSDLAG00005022483	0.805961	7.630648	0.000106	0.00097	3.013023	**UP**	ndufa8	ndufa8	CI
ENSDLAG00005022603	1.397929	7.060694	3.97E-11	1.57E-09	8.805373	**UP**	ndufb9	ndufb9	CI
ENSDLAG00005022803	0.641011	5.503287	0.002998	0.015556	1.808104	**UP**	ndufb8	ndufb8	CI
ENSDLAG00005023654	1.404996	5.355437	1.98E-10	6.98E-09	8.155926	**UP**	NDUFB1	ndufb1	CI
ENSDLAG00005024193	0.710288	6.994357	0.000677	0.00457	2.340116	**UP**	ndufc2	ndufc2	CI
ENSDLAG00005024263	2.277893	6.027145	2.35E-24	3.50E-22	21.45646	**UP**	ndufb3	ndufb3	CI
ENSDLAG00005024691	1.126572	6.995004	8.68E-08	1.83E-06	5.736803	**UP**	ndufa12	ndufa12	CI
ENSDLAG00005024977	2.327198	6.868624	3.55E-26	6.03E-24	23.21948	**UP**	NDUFA13	ndufa13	CI
ENSDLAG00005025335	1.793068	5.363768	1.13E-15	8.22E-14	13.08526	**UP**	ndufa7	ndufa7	CI

ENSDLAG00005006491	0.664461	6.621099	0.001553	0.009003	2.045589	**UP**	sdhc	sdhc	CII
ENSDLAG00005001473	2.449269	6.710147	2.35E-28	4.58E-26	25.33916	**UP**	uqcrq	uqcrq	CIII
ENSDLAG00005003877	1.438636	6.165096	2.55E-11	1.03E-09	8.985321	**UP**	uqcrh	uqcrh.1	CIII
ENSDLAG00005009892	0.617132	5.903605	0.003748	0.018652	1.729268	**UP**	-	uqcrh.2	CIII
ENSDLAG00005010064	0.623337	8.557156	0.002519	0.013484	1.870176	**UP**	cyc1	cyc1	CIII
ENSDLAG00005014897	0.765566	7.166484	0.000245	0.001947	2.710695	**UP**	uqcrc2a	uqcrc2.2	CIII
ENSDLAG00005019495	1.267753	7.074362	1.86E-09	5.52E-08	7.257712	**UP**	uqcrb	uqcrb	CIII
ENSDLAG00005021271	0.873952	7.289913	2.85E-05	0.000314	3.502522	**UP**	UQCRFS1	uqcrfs1.2	CIII
ENSDLAG00005021401	1.100961	7.042437	1.64E-07	3.25E-06	5.487892	**UP**	uqcr10	uqcr10	CIII
ENSDLAG00005024816	0.824832	6.340637	9.56E-05	0.000884	3.05347	**UP**	-	uqcr11.1	CIII
ENSDLAG00005001205	1.76026	5.078701	7.70E-15	5.13E-13	12.28994	**UP**	CX6B1	cox6b2.3	CIV
ENSDLAG00005001713	2.340445	3.066815	6.53E-17	5.21E-15	14.28302	**UP**	cox7a2a	cox7a2.1a	CIV
ENSDLAG00005001717	-2.5571	-1.24886	0.009611	0.039753	1.400625	**DOWN**	-	cox7a2.1b	CIV
ENSDLAG00005007600	2.541384	2.933389	1.57E-18	1.46E-16	15.83708	**UP**	-	cox7a2l.1	CIV
ENSDLAG00005009814	3.722372	5.571764	6.98E-51	4.27E-48	47.36913	**UP**	cox5b2	cox5b.1	CIV
ENSDLAG00005009958	3.305936	4.889453	8.42E-40	3.26E-37	36.48709	**UP**	-	cox8a	CIV
ENSDLAG00005011849	1.293799	8.238723	6.18E-10	1.98E-08	7.702424	**UP**	cox6a1	cox6a.1	CIV
ENSDLAG00005012733	4.615523	2.874778	4.35E-37	1.53E-34	33.81607	**UP**	cox7a1	cox7a1.1	CIV
ENSDLAG00005014851	1.560707	7.32396	1.81E-13	1.01E-11	10.99761	**UP**	COX5B	cox5b.2	CIV
ENSDLAG00005016960	0.712433	8.43357	0.000567	0.003931	2.405445	**UP**	cox4i1	cox4i.1	CIV
ENSDLAG00005020602	3.34326	1.18869	3.74E-13	1.97E-11	10.70449	**UP**	cox4i2	cox4i.2	CIV
ENSDLAG00005022694	0.710734	7.494601	0.00063	0.004317	2.364795	**UP**	-	cox7a1.2	CIV
ENSDLAG00005023073	1.031325	8.090279	7.23E-07	1.24E-05	4.906026	**UP**	COX5A	cox5a.1	CIV
ENSDLAG00005023340	1.024938	6.679605	1.18E-06	1.93E-05	4.714597	**UP**	COX7R	cox7a2l.3	CIV
ENSDLAG00005025866	1.263506	7.66156	1.71E-09	5.12E-08	7.2911	**UP**	cox7c	cox7c	CIV
ENSDLAG00005000620	2.003383	6.388881	5.50E-20	5.70E-18	17.24433	**UP**	atp5if1b	atp5if1.2	CV
ENSDLAG00005003781	0.75454	8.600669	0.000261	0.00205	2.688221	**UP**	atp5pd	atp5pd.1	CV
ENSDLAG00005004321	0.631083	9.720717	0.00215	0.011862	1.925836	**UP**	-	atp5mc.1	CV
ENSDLAG00005004647	2.09098	5.339869	3.41E-20	3.56E-18	17.44797	**UP**	atp5if1a	atp5if1.1	CV

ENSDLAG00005006335	1.877528	2.804059	2.44E-11	9.94E-10	9.002827	**UP**	-	atp5mg.2b	CV
ENSDLAG00005006347	2.086176	7.054863	4.80E-22	5.75E-20	19.2406	**UP**	atp5l	atp5mg.2a	CV
ENSDLAG00005007485	2.178434	7.48516	6.19E-24	8.80E-22	21.05543	**UP**	atp5mf	atp5mf	CV
ENSDLAG00005010974	0.958746	7.496384	4.36E-06	6.07E-05	4.216478	**UP**	-	atp5me	CV
ENSDLAG00005013665	0.739748	8.294264	0.00035	0.002624	2.581108	**UP**	atp5f1d	atp5f1d	CV
ENSDLAG00005015257	1.918095	7.766905	2.64E-19	2.57E-17	16.59065	**UP**	atp5pf	atp5pf.1	CV
ENSDLAG00005016988	1.443251	8.111028	6.43E-12	2.83E-10	9.548729	**UP**	atp5f1e	atp5f1e	CV
ENSDLAG00005019335	1.218705	5.950208	1.50E-08	3.71E-07	6.430965	**UP**	atp5l	atp5mg.1	CV
ENSDLAG00005021874	1.136585	9.207519	4.42E-08	9.96E-07	6.001714	**UP**	atp5pb	atp5pb	CV
ENSDLAG00005024092	1.875057	7.34873	1.89E-18	1.74E-16	15.75987	**UP**	atp5md	atp5mk	CV
ENSDLAG00005024589	0.918996	9.849702	8.57E-06	0.00011	3.958601	**UP**	atp5mc1	atp5mc.3	CV

Table S4.6 European seabass OXPHOS DEGs between first feeding (FF) and mid metamorphosis (MM) stages

Gene ID	logFC	logCPM	PValue	FDR	neglog10FDR	DEG	gene_name	assigned_name	OXPHOS complex
ENSDLAG00005002836	-0.99366	6.292763	2.96E-06	1.79E-05	4.746712	**DOWN**	ndufa3	ndufa3	CI
ENSDLAG00005018458	-0.62893	6.149405	0.003039	0.009588	2.01825	**DOWN**	NDUB4	ndufb4	CI
ENSDLAG00005019817	0.892106	7.478111	1.93E-05	0.000101	3.996015	**UP**	ndufb2	ndufb2	CI
ENSDLAG00005021448	-1.08597	7.554912	2.12E-07	1.55E-06	5.808687	**DOWN**	ndufa11	ndufa11	CI
ENSDLAG00005022657	-0.67889	7.91397	0.001058	0.003762	2.424626	**DOWN**	ndufs7	ndufs7	CI
ENSDLAG00005024263	1.199223	6.253495	2.08E-08	1.78E-07	6.749777	**UP**	ndufb3	ndufb3	CI
ENSDLAG00005025947	-0.75228	7.896465	0.000289	0.001179	2.928508	**DOWN**	ndufv2	ndufv2	CI
ENSDLAG00005016007	-1.59726	7.146347	6.08E-14	1.05E-12	11.97778	**DOWN**	SDHB	sdhb.2	CII
ENSDLAG00005007897	-0.82602	3.24294	0.001234	0.004315	2.36503	**DOWN**	-	uqcr11.2	CIII
ENSDLAG00005021401	-0.68606	7.847987	0.000942	0.003385	2.470442	**DOWN**	uqcr10	uqcr10	CIII
ENSDLAG00005001169	-1.05954	8.428544	3.50E-07	2.47E-06	5.606549	**DOWN**	cox6b2	cox6b2.1	CIV
ENSDLAG00005001205	7.118177	4.696071	2.92E-80	1.05E-77	76.97697	**UP**	CX6B1	cox6b2.3	CIV
ENSDLAG00005001713	1.717545	3.172191	2.26E-10	2.53E-09	8.596459	**UP**	cox7a2a	cox7a2.1a	CIV
ENSDLAG00005007600	1.477009	3.1287	4.62E-08	3.73E-07	6.427825	**UP**	-	cox7a2l.1	CIV
ENSDLAG00005008900	-1.66973	4.219499	2.89E-12	4.06E-11	10.39133	**DOWN**	cox6b2	cox6b2.2	CIV
ENSDLAG00005009814	0.712465	6.12852	0.000809	0.002961	2.528537	**UP**	cox5b2	cox5b.1	CIV
ENSDLAG00005009958	5.593796	4.759639	1.02E-69	3.04E-67	66.51674	**UP**	-	cox8a	CIV
ENSDLAG00005012733	4.040594	2.89203	3.24E-30	2.17E-28	27.66409	**UP**	cox7a1	cox7a1.1	CIV
ENSDLAG00005020602	3.642027	1.178278	2.87E-13	4.64E-12	11.33355	**UP**	cox4i2	cox4i.2	CIV
ENSDLAG00005004647	1.390874	5.477727	3.17E-10	3.49E-09	8.457749	**UP**	atp5if1a	atp5if1.1	CV
ENSDLAG00005015632	0.754024	2.051071	0.016586	0.042482	1.371795	**UP**	-	atp5pf.3	CV
ENSDLAG00005019335	0.697829	6.10246	0.001041	0.003705	2.43116	**UP**	atp5l	atp5mg.1	CV
ENSDLAG00005021473	0.648233	6.003445	0.002362	0.007673	2.115057	**UP**	atp5pd	atp5pd.2	CV
ENSDLAG00005023876	-0.77976	8.593294	0.000163	0.000703	3.153165	**DOWN**	atp5mc3b	atp5mc.2	CV
ENSDLAG00005025077	-0.60516	9.864306	0.003243	0.010162	1.993	**DOWN**	atp5f1c	atp5f1c	CV