The Innovation

Report

Genomic basis of evolutionary adaptation in a warm-blooded fish

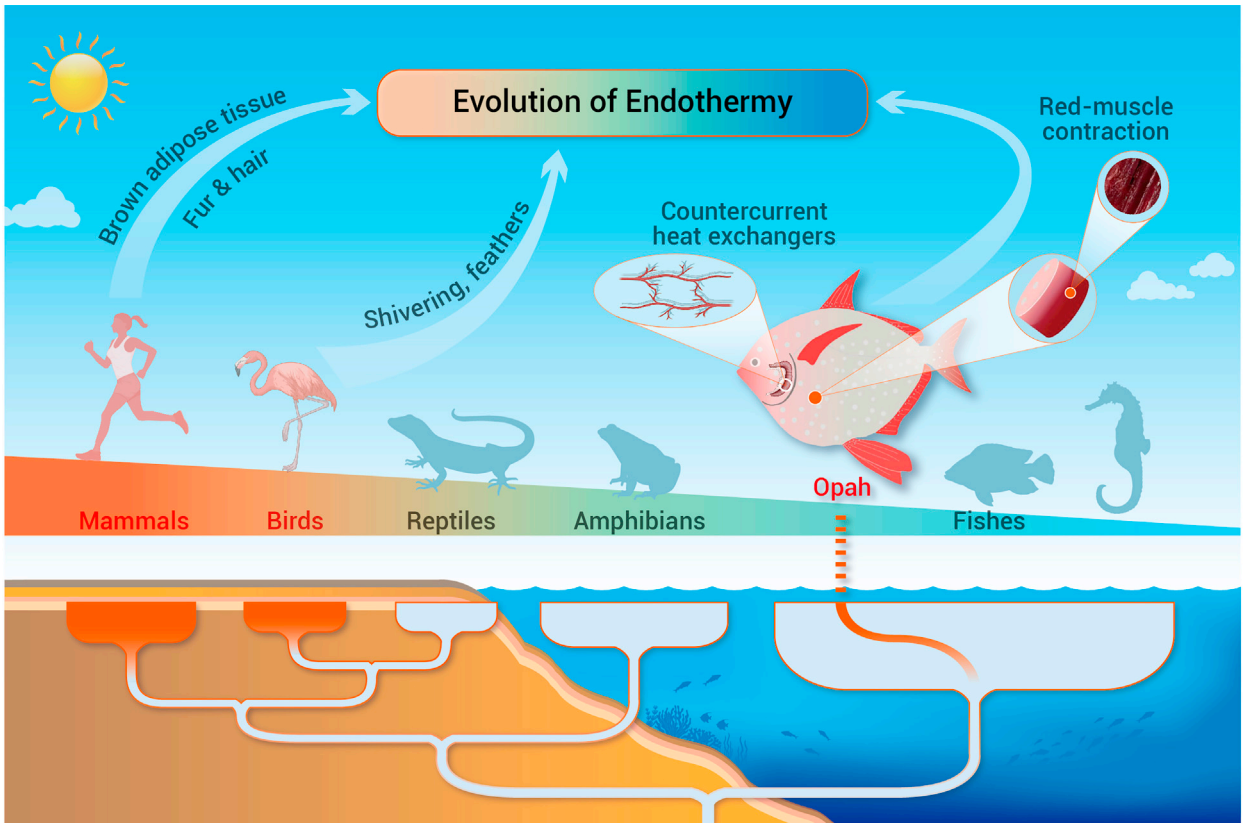
Xin Wang,[1,2,10] Meng Qu,[1,2,10] Yali Liu,[1,2,10] Ralf F. Schneider,[3] Yue Song,[4] Zelin Chen,[1,2] Hao Zhang,[1,2,5] Yanhong Zhang,[1,2] Haiyan Yu,[1] Suyu Zhang,[4] Dongxu Li,[5] Geng Qin,[1,2] Shaobo Ma,[1,2] Jia Zhong,[1,2] Jianping Yin,[1,2] Shuaishuai Liu,[1,2] Guangyi Fan,[4,6] Axel Meyer,[7,*] Dazhi Wang,[5,*] and Qiang Lin[1,2,8,9,*]

*Correspondence: axel.meyer@uni-konstanz.de (A.M.); dzwang@xmu.edu.cn (D.W.); linqiang@scsio.ac.cn (Q.L.)

Received: September 9, 2021; Accepted: November 7, 2021; Published Online: ■■ ■■, ■■; https://doi.org/10.1016/j.xinn.2021.100185

© 2021 The Author(s). This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Graphical abstract



Public summary

- Endothermy has evolved multiple times not only in mammals and birds but also in fishes (teleosts and chondrichthyans)
- A chromosome-level genome sequence of the whole-body endothermic opah was generated, explaining genetic changes in heat production and the sensory and adaptive immune system
- Convergent evolution in endothermic vertebrate lineages was investigated, and genes essential for heart function and metabolic heat production were screened
- Analyses of the unique pectoral muscle of opah revealed that numerous proteins were co-opted from dorsal swimming muscles for thermogenesis and oxidative phosphorylation

Genomic basis of evolutionary adaptation in a warm-blooded fish

Xin Wang,[1,2,10] Meng Qu,[1,2,10] Yali Liu,[1,2,10] Ralf F. Schneider,[3] Yue Song,[4] Zelin Chen,[1,2] Hao Zhang,[1,2,5] Yanhong Zhang,[1,2] Haiyan Yu,[1] Suyu Zhang,[4] Dongxu Li,[5] Geng Qin,[1,2] Shaobo Ma,[1,2] Jia Zhong,[1,2] Jianping Yin,[1,2] Shuaishuai Liu,[1,2] Guangyi Fan,[4,6] Axel Meyer,[7,*] Dazhi Wang,[5,*] and Qiang Lin[1,2,8,9,*]

[1]CAS Key Laboratory of Tropical Marine Bio-resources and Ecology, South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou 510301, China
[2]Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Guangzhou 511458, China
[3]Marine Evolutionary Ecology, GEOMAR Helmholtz Centre for Ocean Research Kiel, 24105 Kiel, Germany
[4]BGI-Qingdao, Qingdao 266555, China
[5]State Key Laboratory of Marine Environmental Science/College of the Environment and Ecology, Xiamen University, Xiamen 361102, China
[6]BGI-Shenzhen, Shenzhen 518083, China
[7]Department of Biology, University of Konstanz, Konstanz 78464, Germany
[8]University of the Chinese Academy of Sciences, Beijing 100101, China
[9]Laboratory for Marine Fisheries Science and Food Production Processes, Pilot National Laboratory for Marine Science and Technology (Qingdao), Qingdao 266237, China

[10]These authors contributed equally
*Correspondence: axel.meyer@uni-konstanz.de (A.M.); dzwang@xmu.edu.cn (D.W.); linqiang@scsio.ac.cn (Q.L.)
Received: September 9, 2021; Accepted: November 7, 2021; Published Online: ■■ ■■, ■■; https://doi.org/10.1016/j.xinn.2021.100185
© 2021 The Author(s). This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
Citation: Wang X., Qu M., Liu Y., et al., (2021). Genomic basis of evolutionary adaptation in a warm-blooded fish. The Innovation ■■(■■), 100185.

Few fishes have evolved elevated body temperatures compared with ambient temperatures, and only in opah (*Lampris* spp) is the entire body affected. To understand the molecular basis of endothermy, we analyzed the opah genome and identified 23 genes with convergent amino acid substitutions across fish, birds, and mammals, including *slc8b1*, which encodes the mitochondrial Na^+/Ca^{2+} exchanger and is essential for heart function and metabolic heat production. Among endothermic fishes, 44 convergent genes with suggestive metabolic functions were identified, such as *glrx3*, encoding a crucial protein for hemoglobin maturation. Numerous genes involved in the production and retention of metabolic heat were also found to be under positive selection. Analyses of opah's unique inner-heat-producing pectoral muscle layer (PMI), an evolutionary key innovation, revealed that many proteins were co-opted from dorsal swimming muscles for thermogenesis and oxidative phosphorylation. Thus, the opah genome provides valuable resources and opportunities to uncover the genetic basis of thermal adaptations in fish.

Keywords: endothermy; opah; whole-genome sequencing; convergent evolution; thermogenesis

INTRODUCTION

Endothermy, the ability of an organism to metabolically produce heat to achieve a high, stable body temperature, is one of the most fascinating traits that has evolved repeatedly.[1,2] This mechanism provides some independence from environmental thermal fluctuations by permitting high activity levels even with low or variable temperatures.[3] Mammals and birds were long considered the only lineages of endothermic vertebrates, and whole-body endothermy is thought to have evolved independently in these two lineages.[1] Except for muscular thermogenesis, in placental mammals, metabolic heat is mainly produced by nonshivering thermogenesis in brown adipose tissue. Heat production in birds likely occurs through muscular thermogenesis, the production of heat through the uncoupling of calcium ions in the sarcoplasmic reticulum of muscle cells, activated by sarcolipin.[1] However, endothermy has also been discovered in other vertebrate clades, with different types of endothermy having evolved independently at least five times in fishes (teleosts and chondrichthyans).[4] The thermogenic strategies of teleosts and chondrichthyans are mainly attributed to contractions of the slow-twitch aerobic red muscles during constant swimming, a configuration often called "red muscle endothermy".[1,4] For example, tuna (Scombridae) and mackerel sharks (Lamnidae) can elevate their core body temperature, whereas the brain and eye regions can become warmer than ambient temperatures in billfishes (Istiophoridae). In contrast to such regional endothermy, opahs are the only known teleosts that can warm their entire body by circulating locally warmed-up blood and are thus the only whole-body endotherm species aside from mammals and birds.[4–6]

Opah is a large mesopelagic circumglobal species with the unique ability to produce large quantities of metabolic heat in its dark red, aerobic pectoral musculature, putatively an evolutionary key innovation that induces forward thrust during continuous swimming by powering pectoral fin oscillation, and this is insulated from the cold surrounding waters by a thick layer of fatty connective tissue.[7,8] Furthermore, to reduce the cooling of warm blood in the gills, where most heat is lost, opah has evolved a crucial second key change, a counter-current retia mirabilia system in its gills.[7] The convoluted alternating arteries with cold-oxygenated and warm-deoxygenated blood form the basis of the heat exchange system. With whole-body endothermy, the opah has the capacity for enhanced physiological functions in their cold, deep, and nutrient-rich habitats, below the ocean thermocline. Its adaptation to whole-body endothermy facilitates enhanced physiological functions, such as increased muscle power, increased capacity for more sustained performance, enhanced temporal resolution and neural conductance for the eye and brain, and an increased rate of food digestion and assimilation.[7] Moreover, as endothermy is thought to represent an optimum level in the trade-off between metabolism and fitness, the case of the opah can inform us of how endothermy affects organism evolution and survivorship.[9,10]

The ability of the opah to maintain body regions at a warmer temperature than the ambient water challenges the general concept of "cold-bloodedness" in fish.[3] However, unlike mammals and birds, the genetic characteristics associated with the occurrence of endothermy in fish have rarely been investigated. In many respects, the features of the entirely endothermic opah provide an unparalleled model to explore the molecular basis of endothermy as an adaptive trait in the ocean environment. In this study, we generated a chromosome-level genome of the Smalleye Pacific Opah (*Lampris incognitus*) (Figure 1A). We then conducted comparative genomic, transcriptomic, and proteomic analyses to identify the molecular adaptations accompanying endothermy. In addition, we explored the genetic patterns that are shared during the convergent evolution of endothermy among endothermic fishes, mammals, and birds.

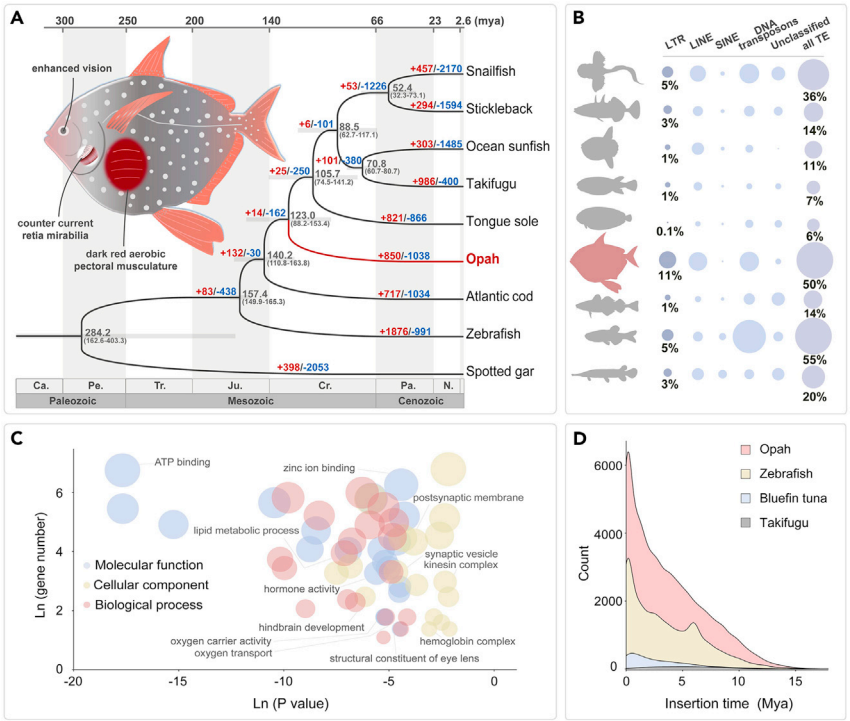


Figure 1. Evolutionary history of opah (A) Evolutionary key innovations and specialized physiological features involved in the endothermy of the Smalleye Pacific Opah, *Lampris incognitus*, and gene family expansion (red)/contraction (blue) in opah and other teleosts. The bootstrap value of all nodes is 100.
(B) Percentages of transposable elements (TEs) in the different teleost genomes studied. LTR, long terminal repeats; LINE, long interspersed element; SINE, short interspersed element.
(C) Enriched GO functional categories that were closest to the LTR retrotransposons (2 kb up and downstream).
(D) Expansion of LTR retrotransposon in opah, zebrafish, bluefin tuna, and takifugu.

RESULTS

Genome features and long terminal repeat (LTR) retrotransposons

PacBio sequencing technology was used to generate a high-quality genome of a male opah. We generated 183.23 Gb of PacBio clean data with approximately 136-fold coverage. These data were assembled into a 1,367.47 Mb genome, which approximated the size estimated by *k*-mer distribution (1,374.37 Mb, Figure S1), with a contig N50 size of 3.819 Mb and a scaffold N50 size of 21.58 Mb (Table S1). Hi-C reads were used to scaffold the contig-level assemblies, resulting in 22 chromosome-level scaffolds, which comprised the opah genome (Figures S2 and S3, Tables S2 and S3). In total, 24,658 genes were predicted (Tables S4–S6) using a combination of de novo gene prediction programs and homology-based methods, along with RNA-seq data. The evaluation of the genome for completeness based on Benchmarking Universal Single-Copy Orthologs (BUSCO; database: vertebrata_odb9) and the Core Eukaryotic Genes Mapping Approach (CEGMA) resulted in values of 94.08% and 98.91%, respectively (Tables S7 and S8). Genome annotation completeness was also evaluated using BUSCO, and the results showed that our gene set contained 95.78% complete and 2.28% fragmented ortholog genes (Table S9), showing that our gene annotation was highly complete. In total, 167 unique gene families were identified in the opah genome compared with those in zebrafish, fugu, and spotted gar (Figure S4).

A phylogenomic analysis assigned the opah to a basal position within the advanced spiny-rayed Teleostei clade (Acanthomorpha) (Figure S5), which is in agreement with recent studies.[11,12] The opah genome was found to contain a relatively high percentage of transposable elements (TEs) and other interspersed repeats (taken together, ~50%). Specifically, the genome contained the highest percentage of the LTR retrotransposon superfamily (11%) reported in any teleost to date (Figure 1B, Tables S10, and S11). Enriched genes in spatial proximity to the LTR retrotransposons (2 kb up and downstream) were investigated to discover the evolutionary role of the extremely high LTR content in the opah genome. Among these genes, many gene ontology terms related to metabolism and the sensory system were significantly overrepresented, including the lipid metabolic process (GO:0006629, P = 4.2e-04), ATP binding (GO:0005524, P = 1.9e-08), glycerol transport (GO:0015793, P = 7.3e-04), oxygen transport (GO:0015671, P = 2.7e-03), and structural constituent of the eye lens (GO:0005212, P = 5.1e-03) (Figure 1C, Table S12). The expansion of LTRs during the evolution of the opah might thus have affected gene expression and translocation,[13] likely contributing to the endothermic ability of the opah and its adaptations to the deep-sea environment. In opah, the LTR retrotransposons were inferred to have started expanding approximately 15 million years ago (Figure 1D), which might have coevolved with (and potentially facilitated) whole-body endothermy.

Genome-wide signatures of convergent evolution in endothermic lineages

A whole-body endothermic phenotype evolved independently in mammals, birds, and opahs, and we analyzed these three lineages to detect molecular signals accompanying the convergent evolution of this phenotype (Figure 2A). After selecting five endothermic species (opah, turkey, chicken, mouse, and human) as foreground branches, 204 genes were found to have a convergent signature for positive selection, which were enriched in terms related to primary metabolic processes (GO: 0044238, P = 2.3e-05), organic substance metabolic processes (GO: 0071704, P = 8.0e-05), and carbohydrate derivative biosynthetic processes (GO: 1901137, P = 1.7e-03) (Table S13).

Genes with convergent amino acid substitutions were further screened among the selected mammals, birds, and opah (Table S14). Notably, *slc8b1* (mitochondrial sodium/potassium/calcium exchanger 6) showed signs of convergence, and it encodes the mitochondrial Na^+/Ca^{2+} exchanger (NCLX)[14] and is essential for heart function and metabolic heat production. A substitution site (M460I) was found to be located in the Na_Ca_ex domain (Figure S6A, Table S14), which is conserved among ectothermic teleosts, chondrichthyans, and reptiles. Because our orthologous gene sets comprised only two mammals and two birds, we manually expanded the scope of this comparison with 64 species, including 10 species of mammals, birds, and reptiles, five amphibians, and 29 fishes (Table S15). We found that the convergent amino acid change (M460I) was restricted to the whole-body endothermic mammals, birds, and opah and thus not found in the regional endothermic fishes, bluefin tuna, and great white shark (Figure S6A), implying potentially different mechanisms for evolving whole-body and regional endothermy. Additionally, another 22 convergent genes were detected among four of the five endothermic species (including opah; Figure 2A, Table S14).

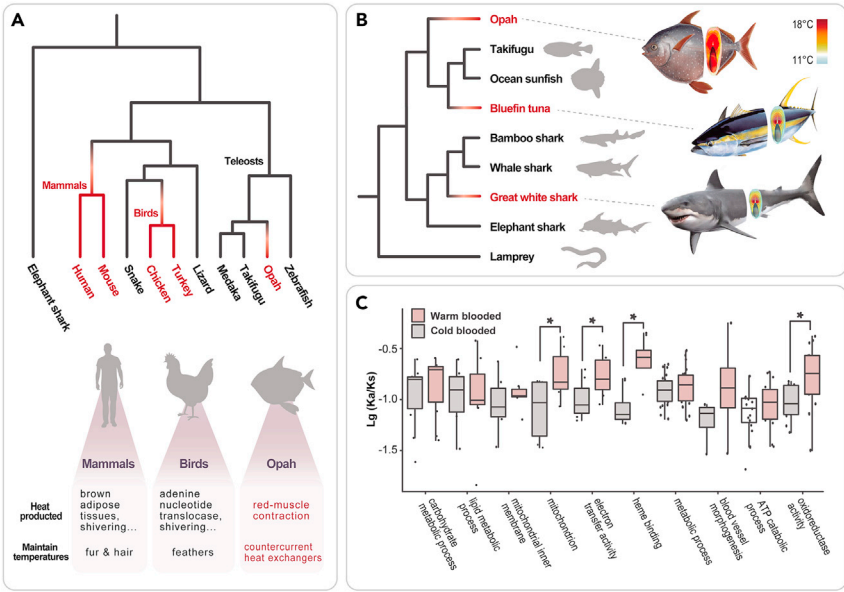


Figure 2. Genomic convergence analysis of endothermy in fishes (A) Phylogeny of independently evolved endothermic vertebrates.
(B) The phylogeny consists of red muscle endothermic fishes (opah, Pacific bluefin tuna, and great white shark) and outgroups.
(C) Comparison of the Ka/Ks ratios of genes involved in 10 GO terms related to heat production and metabolism between the endothermic foreground branch and ectothermic background branch of panel (B).

Despite these similarities, teleosts and chondrichthyans generally generate heat mainly with their swim musculature, which differs from the heat production strategy of mammals and birds.[15,16] To account for this discrepancy, we performed another convergent evolution analysis comprising only fishes, namely the opah, Pacific bluefin tuna, and the great white shark (Figure 2B). Firstly, positive selection analyses were conducted with three different datasets by using either the opah, the Pacific bluefin tuna, or the great white shark as the foreground branch, revealing the genetic landscape of 36 genes with adaptive mutations that occurred during the evolution of endothermic fishes (Table S16). Among them, the *opa1* gene is crucial for the maintenance of mtDNA,[17] and *pyroxd2* plays important roles in regulating mitochondrial functions.[18] Then, we set all three endothermic fishes as foreground branches simultaneously. A total of 112 genes were found to have convergent signatures of positive selection, which were enriched in gene ontology terms related to a citrate lyase complex (GO: 0009346, P = 4.4e-03), energy homeostasis (GO: 0097009, P = 5.5e-03), and mitochondrial DNA metabolic process (GO: 0032042, P = 5.5e-03) (Table S17). In addition, we calculated the Ka/Ks ratio for each gene and found that four GO terms related to heat production and metabolism had significantly higher Ka/Ks ratios in the three endothermic fishes (Figure 2C, Table S18).

Among the endothermic fishes, 78 convergent amino acid replacements in 44 genes were identified (Table S19). One particular convergent amino acid substitution was identified in all three endothermic fish species at the V198A site of *glutaredoxin 3* (*glrx3*) (Table S19), a gene that plays a crucial role in Fe/S protein biogenesis and is indispensable for hemoglobin maturation.[19] The convergent site V198A was found to be located in the GRX domain, which has thiol oxidoreductase activity that is responsible for the reduction in protein disulfides or glutathione-protein mixed disulfides and is involved in cellular redox regulation. Convergent amino acid substitutions might affect the activity of GRX, which could in turn influence the function of Glrx3. We manually analyzed 45 bony fishes and six chondrichthyans and found that this convergent amino acid change (V198A) only existed in the three endothermic fishes (Figure S6B). This substitution is located in the GRX domain, which is conserved among teleosts, chondrichthyans, and lamprey, indicating the ancient and conserved role of *glrx3* in binding Fe/S clusters. In addition, three convergently changed sites (R8L, R159Q, and S201A) in the citrate lyase subunit β -like protein (Clybl) were found to be shared between opah and bluefin tuna (Table S19). Clybl plays important roles in energy metabolism, as it encodes a malate/β-methylmalate synthase, which can convert glyoxylate and acetyl-CoA to malate or glyoxylate and propionyl-CoA to β-methylmalate.[20] Additionally, we performed convergent analysis for background (nonendothermic species), and no convergent genes were found in either nonendothermic vertebrates (zebrafish, takifugu, medaka, lizard, snake, and elephant shark) or nonendothermic fish (takifugu, ocean sunfish, bamboo shark, whale shark, elephant shark, and lamprey). In summary, our genome scans provide evidence of convergent selective pressures and genetic responses shared by these endothermic fish species.

Evolutionary basis of whole-body endothermy

We performed comparative genomic analyses with other fish genomes to uncover opah-specific expanded/contracted gene families and positively selected genes (PSGs). We identified a total of 850 (36 genes with $P < 0.05$) and 1,038 (38 genes with $P < 0.05$) gene families to be expanded or contracted in opah, respectively (Figure 1A, Tables S20 and S21), and identified 233 genes with signatures of positive selection (Table S22). A notably large number of expanded gene families was enriched for GO terms related to the composition of muscle fibers (e.g., actin cytoskeleton GO: 0015629, myofibril GO: 0030016) and the contractile machinery of skeletal muscle (e.g., actin filament organization GO: 0007015, skeletal muscle thin filament assembly GO: 0030240) (Figure 3A, Table S23). In addition, many PSGs, such as *med12*, *rmnd1*, and *alg2*, are considered to be involved in the regulation of muscle formation and development. The expanded gene families of the structural proteins myosin, actin, and troponin encode key components of the sarcomere structure (Figure 3B). The expanded gene families were also enriched in the term cardiac muscle cell development (GO: 0055013), and three PSGs, *utp15* (U3 small nucleolar RNA-associated protein 15), *utp14a* (U3 small nucleolar RNA-associated protein 14 homolog A), and *robo4* (roundabout homolog 4) (Figures 3C and S7), were related to the genesis, regulation, and patterning of vasculature. A mutation in *utp15* disrupts vascular patterning in zebrafish embryos,[21] *utp14a* promotes angiogenesis,[22] and *robo4* is expressed in emerging blood vessels and can neutralize signaling to maintain vessel integrity through an angiogenic factor.[23] These expanded gene families and PSGs found in opah might thus be related to the evolution of the counter-current retia mirabilia.

The identified 233 PSGs were significantly enriched in organic substance metabolic processes and nitrogen compound metabolic processes (Figure 4A, Table S24). Among them, *agrp* regulates energy homeostasis;[24] *acsf4* and *acot8* are involved in lipid metabolism (Figure 4B); and *ndufa10*, *ndufs4*, and *cox15* encode core proteins in the mitochondrial electron transport chain (Figure 4C), which is crucial for oxidative phosphorylation. Remarkably, *ndufa10* also exhibited convergent amino acid changes among the fast-running ruminant species, which require a sufficient quantity of red muscle and efficient contractions to guarantee a supply of metabolic heat.[25] Moreover, gene families involved in lipid metabolism (*mcadd*),

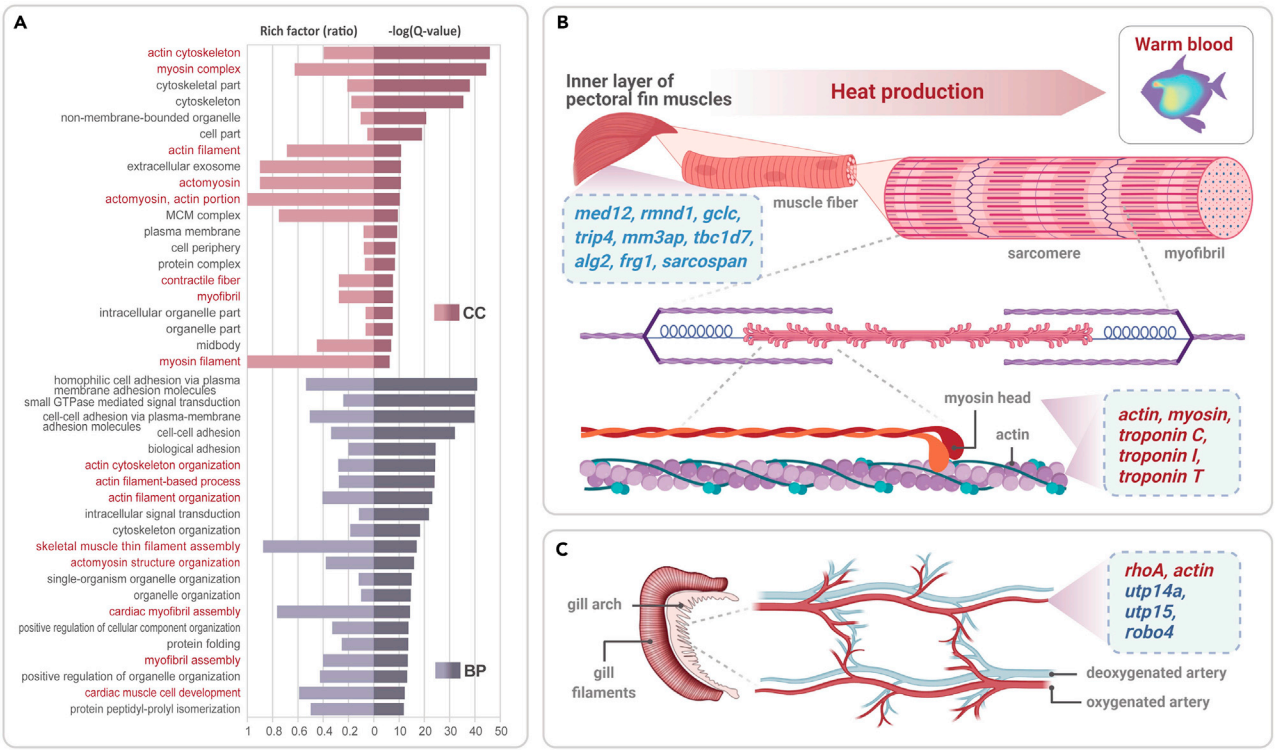


Figure 3. Genomic features related to the production and retention of metabolic heat in opah (A) The top 20 significantly enriched GO terms for expanded gene families in the opah genome are shown for cellular components (CCs) and biological processes (BPs). (B) PSGs (blue) and expanded genes (red) related to muscle development and skeletal muscle contraction. (C) PSGs (blue) and expanded genes (red) are involved in vasculature genesis, regulation, and patterning.

carbohydrate metabolism (*gapd*, *enolase*), and oxidative phosphorylation (*ndufv*, *cytc*, *cox7a2*, and *atp1a1*) were found to be expanded in the opah genome (Figure 4C). The genetic changes detected in opah indicate enhanced physiological functions, including increased aerobic performance, food digestion, and assimilation.

Transcriptomic and proteomic profiles of opah pectoral musculature

Being the major "heat motor" for the whole-body endothermy of opah, further studies were conducted to evaluate the evolutionary origin of the dark red, aerobic pectoral musculature using its transcriptomic and proteomic profiles. Seven tissues were used to identify the tissue-specific expressed genes as follows: the inner (PMI) and superficial layer (PMS) of the pectoral musculature, dorsal musculature (DM), ventral musculature (VM), liver, kidney, and gill (Table S25). Additionally, five tissues were used to identify the tissue-specific expressed proteins, namely PMI, PMS, VM, liver, and gill. Transcriptomic profiles revealed that the 100 genes with the highest expression in PMI, which included many expanded gene families such as cytochrome c oxidase (*cytc*), ATP synthase F0 subunit 6 (*atp6*), glyceraldehyde-3-phosphate dehydrogenase (*gapdh*), and troponin (*troponin*) (Figure 5, Table S26). Consistent with the transcriptomic profiles, the 100 most abundant proteins included Cytc, Atp6, Gapdh, troponin, and myosin (Figure 5, Table S27). These proteins have been reported to be involved in cytochrome heat production, ATP biosynthesis, glycolysis, and skeletal muscle contraction.[26–28] In addition, hemoglobins were also highly expressed, which might indicate a type of functional compensation, considering the high oxygen consumption (Tables S26 and S27).[29] Taken together, the PMI highly expressed proteins are likely involved in endothermy, providing insights into the genetic basis for the evolutive "whole-body endothermy" of opah.

At both the structural and molecular levels, the PMI is a unique tissue that differs from muscles in other opah regions. Previous studies have indicated that the evolutionary origin of "new organs" or key innovations such as pectoral musculature often depends on the recruitment and co-option of genes that were originally expressed in other tissues.[30] In this study, the comparative transcriptome analysis indicated that the PMI depends largely on genes recruited from the segmental swimming musculature, specifically the DM (Table S28), whereas the gene expression profile of the PMS was closest to that of VM. To determine the disparity between the PMI and other musculature, we conducted further analyses and identified 1,480, 6,645, and 4,771 differentially expressed genes in PMI versus DM, PMI versus VM, and PMI versus PMS, respectively (Figures S8 and S9, Tables S29, 30 , and S31). In addition, a comparative label-free proteomic approach identified 470 (149 downregulated, 321 upregulated) and 573 (402 downregulated, 171 upregulated) differentially expressed proteins (DEPs) in PMI versus VM and PMI versus PMS, respectively (Tables S32 and S33). A Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of these DEPs revealed that the highly expressed proteins in the PMI were mainly involved in thermogenesis and oxidative phosphorylation, indicating that there has been an adaptive change in the heat production of the evolved musculature (Figure S8).

Adaptive sensory and immune systems pave the way to the deep sea

In opah, approximately 5.6% (13 out of 234 genes) of the identified PSGs play important roles in the visual system (Figure 6A, Table S22): the centrosomal protein POC5 (*poc5*) is a causative gene for retinal disease, and previous studies have shown that the loss of *poc5* caused retinal degeneration in zebrafish;[31] the mutant dynamin binding protein (*dnmbp*) and mitochondrial isoleucyl-tRNA synthetase 2 (*iars2*) are associated with cataracts in mammals;[32,33] tripartite motif-containing protein 44 (*trim44*) is crucial for iris development, and the variants in *trim44* are considered the major cause of aniridia;[34] biallelic mutations in the S-phase cyclin-A-associated protein in the ER (*scaper*) can cause autosomal recessive retinitis pigmentosa with intellectual disability;[35] optic atrophy protein 1 (*opa1*) plays an important role in optic nerve development.[36] In addition, significant expansions of the β- and γ-crystallin family have been found in the opah genome, which also provide support for the enhanced visual capability of opah (Figure 6A). However, in contrast to their enhanced eyesight, opah might have a reduced

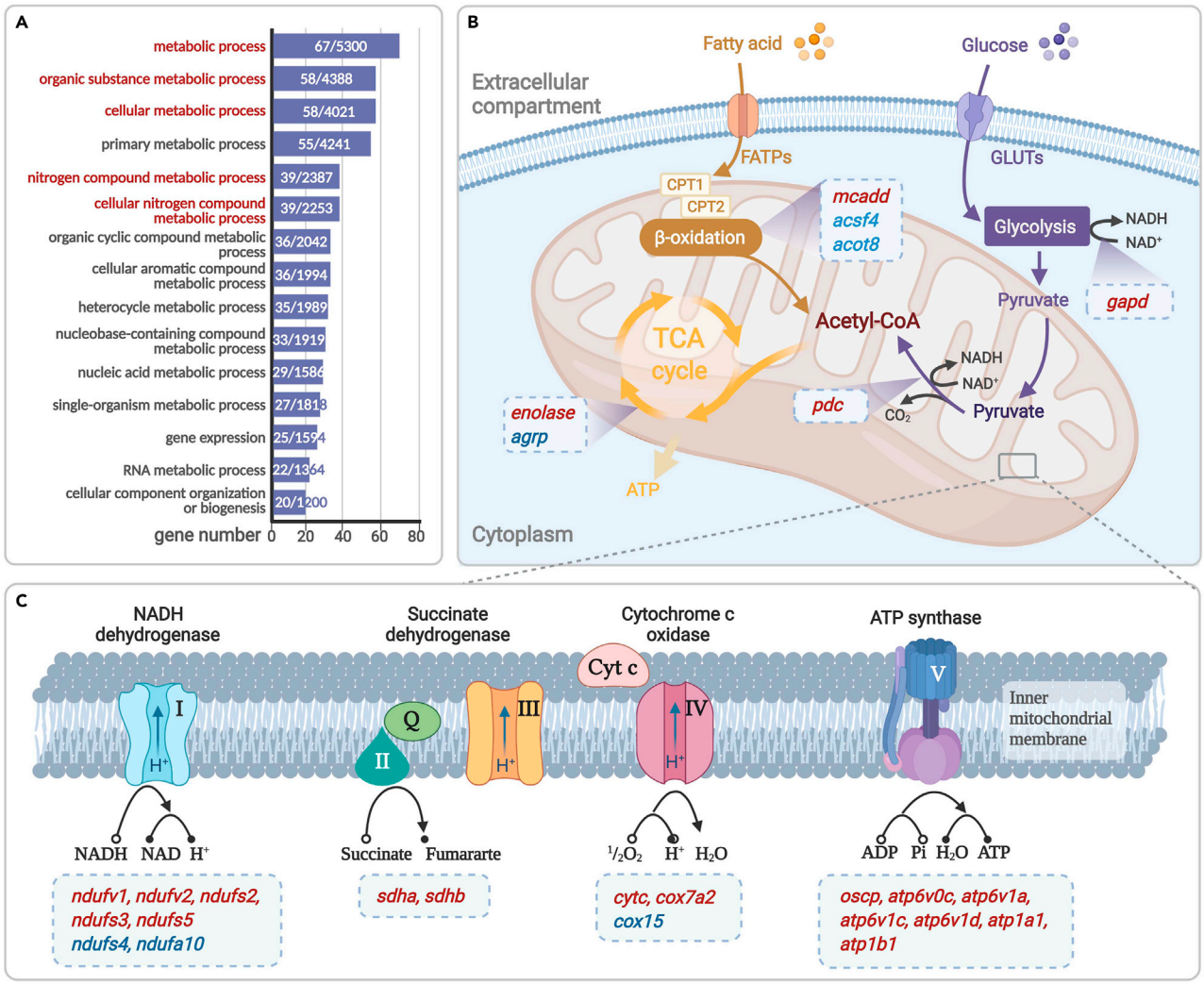


Figure 4. Genomic features related to the whole-body endothermy of opah (A) Top 15 enriched GO terms (category: biological process) among PSGs in the opah genome. (B and C) Expanded genes (red) and PSGs (blue) in opah are involved in lipid metabolism, carbohydrate metabolism, and (C) oxidative phosphorylation.

olfactory performance, as they have the lowest number of olfactory receptor genes reported to date for teleosts (only 24, compared with 26 in *Hippocampus comes* and 156 in *Danio rerio*; Figure 6B).[37,38] Additionally, in the opah genome, *MHCII* genes were significantly expanded, especially in the DA group of the *MHCII* β-chain (Figure S10). Compared with the DB and DE groups, the DA group exhibited the classical characteristics of *MHCII*, and these genes were found to be specifically and highly expressed in the immune tissues (liver, kidney, and gill) of opah (Figure S11).

DISCUSSION

The evolution of endothermy in mammals and birds has been intensively studied during the last few decades, and it has been hypothesized to confer numerous physiological and ecological advantages.[15,39] Different thermogenic strategies have been adopted by those vertebrate clades that evolved endothermy independently;[1] however, the adaptive changes that accompanied the evolution of endothermy exhibit many similarities, including a higher metabolic rate, a higher capability of sustained aerobic performance, and enhanced neural conductance.[7,15,16] Adaptive phenotypic convergence between distantly related taxa has occurred repeatedly in nature and has generally been attributed to convergent genetic changes driven by similar selective pressures.[40–42] Several lineages of elasmobranchs and teleosts independently evolved specific forms of muscular endothermy allowing them to elevate their body temperature relative to that of the ambient water.[1] Previous studies on tuna and the great white shark revealed convergent genetic changes involved in oxidative phosphorylation, the Krebs cycle, muscle contraction, and lipid and fat metabolism,[15,16] and a recent study on tuna and billfish observed convergent amino acid replacements in proteins related to heat production and the visual system.[43]

Fish endothermy is mainly dependent on red muscle activity, which produces heat as a by-product during muscle contraction[7,15] and is located close to the center of the body, in most endothermic species (e.g., tuna and great white sharks).[44] To achieve the extraordinarily high respiratory metabolic rate, endothermic fishes have evolved a number of key adaptations. In this study, we found a convergent amino acid substitution at the V198A site located in the conserved GRX domain of *glrx3* for the three endothermic fishes among the 51 fishes investigated. As *glrx3* is indispensable for hemoglobin maturation, the replaced amino acid in the convergent site of the endothermic species, relative to that in ectothermic groups, might change the ability to provide oxygen to the musculature, thus affecting muscle strength and endurance in the opah, bluefin tuna, and great white sharks.

As tuna, billfish, and the great white shark are only able to warm up specific regions of their bodies, they are not typically considered warm blooded. Opah exhibit "whole-body endothermy," through the production of vast amounts of heat, by constantly flapping their enlarged pectoral fins when continuously swimming in a labrid fashion.[7,8] In the opah genome, a substantial portion of PSGs and expanded gene families have been identified as involved in

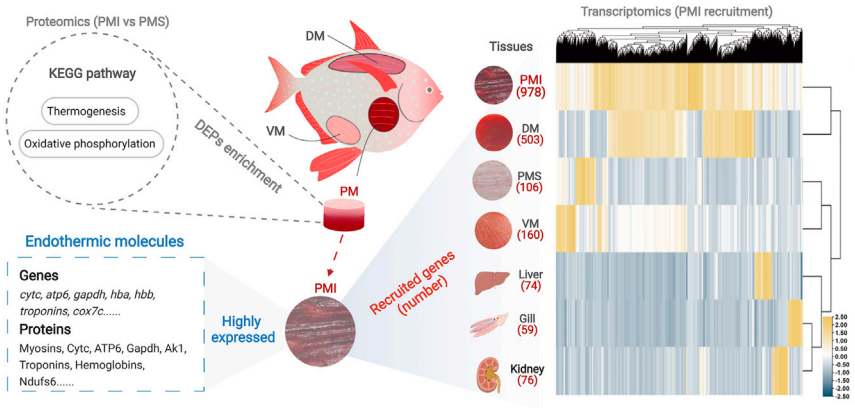


Figure 5. Transcriptomic and proteomic profiles of the novel pectoral musculature that was co-opted for heat production Proteomics: KEGG pathway enrichment analysis of the upregulated DEPs of the PMI and the PMS, including thermogenesis and oxidative phosphorylation. Endothermic molecules: the 100 most highly expressed genes and proteins in the PMI contributing to whole-body endothermy. Transcriptomics: the number and expression profiles of genes recruited from other tissues for the development of the dark red, aerobic PMI.

not only skeletal muscle contraction but also lipid metabolism, carbohydrate metabolism, and oxidative phosphorylation, which coincides with expectations that endothermic species have increased metabolic rates and aerobic capacities.[15] Furthermore, the expansion of LTR retrotransposons in opah was also suspected to regulate the transposition and expression of genes related to heat-generating processes. Genes involved in muscle contraction and energy metabolism were found to have high expression levels in the PMI, which have thick fibers and show a much deeper red color compared with muscles in other regions. Interestingly, the genes in the PMI were largely recruited from the DM rather than the PMS, suggesting that the PMI is more derived and specialized in opah. Hence, we agree with previous studies suggesting that the novel structure and function of PMI is the key to the evolution of endothermy in opah.

Adaptive changes have been hypothesized to have arisen in the sensory systems of opah alongside endothermy, including enhanced temporal resolution and neural conductance for the eye and brain.[7] As a high-performance predator, opah have large eyes and are thought to rely heavily on vision to seek and capture active prey such as fast swimming squid and fish.[45] This was also reflected by its genome, as crystallin genes were duplicated, and several PSGs related to the visual system were identified; this enhanced visual capability is likely one of the benefits from endothermy, since an improvement in vision has also been reported in billfish and tuna, which are able to warm their brains and eyes.[43,46] The reduced olfactory performance might be compensated by the enhanced vision, which has been reported to be an adaptation to facilitate habitat utilization of deep and cold waters.[38] Enhanced vision, along with an expanded thermal

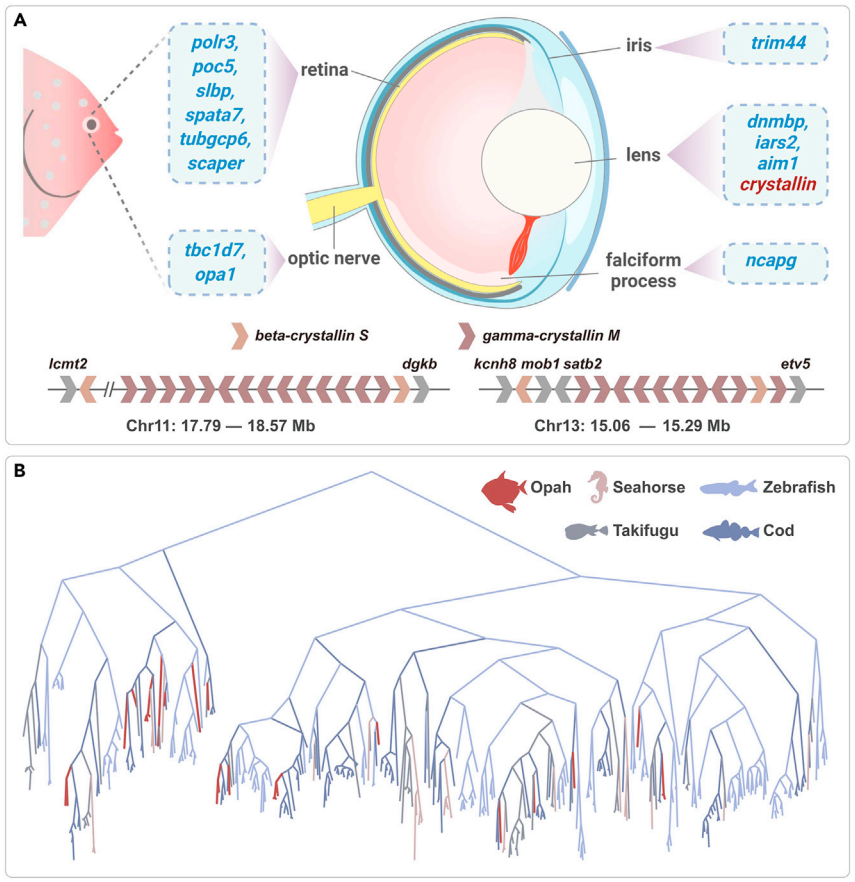


Figure 6. Genomic analysis of genes related to vision and olfaction in opah (A) Expanded genes (red) and PSGs (blue) in opah are involved in visual development.
(B) Gene phylogeny of olfactory receptors across the genomes of opah, zebrafish, fugu, and cod.

tolerance to the nutrient-poor pelagic environment, would enable increased access to high-energy prey and provide a strong selective benefit.[39] The MHCII molecules are heterodimeric surface proteins involved in the presentation of exogenous antigens during the adaptive immune response.[47] Since endothermy represents an optimum in the trade-off between metabolism and fitness of immunity,[9,10] the adaptive genetic changes in the immune system might be coordinated with the evolution of endothermy in opah. The unique endothermy of opah has put the species in a unique situation where it can exploit cold, nutrient-rich deep waters without the shortcomings of reduced metabolic rates.

Conclusions

In this study, we generated a high-quality chromosome-level genome assembly of the Smalleye Pacific Opah (*Lampris incognitus*). By comparing the opah with mammals, birds, and other regionally endothermic fishes, strong genomic signatures of convergent evolution in the heat production process and visual system were identified. We identified a large number of positively selective genes and expanded gene families in the opah genome that are involved in the processes of skeletal muscle contraction, lipid metabolism, carbohydrate metabolism, and oxidative phosphorylation, among others, indicating that they are part of the genomic basis for whole-body endothermy. Our transcriptomic and proteomic analyses showed that genes involved in muscle contraction and energy metabolism have high expression levels in the PMI, which might be the evolutionary key innovation involved in the evolution of whole-body endothermy. These adaptations might explain how opahs can have such an active lifestyle in cold and deep-sea environments.

REFERENCES

1. Legendre, L., and Davesne, D. (2019). The evolution of mechanisms involved in vertebrate endothermy. Phil. Trans. R. Soc. B ***375***, 20190136.
2. Seebacher, F. (2020). Is endothermy an evolutionary by-product? Trends Ecol. Evol. ***35***, 503–511.
3. Watanabe, Y.Y., Goldman, K.J., Caselle, J.E., et al. (2015). Comparative analyses of animal-tracking data reveal ecological significance of endothermy in fishes. Proc. Natl. Acad. Sci. U S A ***112***, 6104–6109.
4. Block, B.A., Finnerty, J.R., Stewart, A.F., and Kidd, J. (1993). Evolution of endothermy in fish: mapping physiological traits on a molecular phylogeny. Science ***260***, 210–214.
5. Carey, F. (1982). A brain heater in the swordfish. Science ***216***, 1327–1329.
6. Wolf, N.G., Swift, P.R., and Carey, F.G. (1988). Swimming muscle helps warm the brain of lamnid sharks. J. Comp. Physiol. B ***157***, 709–716.
7. Wegner, N.C., Snodgrass, O.E., Dewar, H., and Hyde, J.R. (2015). Whole-body endothermy in a mesopelagic fish, the opah, *Lampris guttatus*. Science ***348***, 786–789.
8. Franck, J.P.C., Slight-Simcoe, E., and Wegner, N.C. (2019). Endothermy in the smalleye opah (*Lampris incognitus*): a potential role for the uncoupling protein sarcolipin. Comp. Biochem. Physiol. A. Mol. Integr. Physiol. ***233***, 48–52.
9. Bergman, A., and Casadevall, A. (2010). Mammalian endothermy optimally restricts fungi and metabolic costs. mBio ***1***, e00212–00210.
10. Evans, S.S., Repasky, E.A., and Fisher, D.T. (2015). Fever and the thermal regulation of immunity: the immune system feels the heat. Nat. Rev. Immunol. ***15***, 335–349.
11. Ravi, V., and Venkatesh, B. (2018). The divergent genomes of teleosts. Annu. Rev. Anim. Biosci. ***6***, 47–68.
12. Hughes, L.C., Ortí, G., Huang, Y., et al. (2018). Comprehensive phylogeny of ray-finned fishes (Actinopterygii) based on transcriptomic and genomic data. Proc. Natl. Acad. Sci. U S A ***115***, 6249–6254.
13. Chalopin, D., Naville, M., Plard, F., et al. (2015). Comparative analysis of transposable elements highlights mobilome diversity and evolution in vertebrates. Genome Biol. Evol. ***7***, 567–580.
14. Luongo, T.S., Lambert, J.P., Gross, P., et al. (2017). The mitochondrial Na^{+}/Ca^{2+} exchanger is essential for Ca^{2+} homeostasis and viability. Nature ***545***, 93–97.
15. Ciezarek, A.G., Dunning, L.T., Jones, C.S., et al. (2016). Substitutions in the glycogenin-1 gene are associated with the evolution of endothermy in sharks and tunas. Genome Biol. Evol. ***8***, 3011–3021.
16. Marra, N.J., Richards, V.P., Early, A., et al. (2017). Comparative transcriptomics of elasmobranchs and teleosts highlight important processes in adaptive immunity and regional endothermy. BMC Genomics ***18***, 87.
17. Elachouri1, G., Vidoni, S., Zanna, C., et al. (2011). OPA1 links human mitochondrial genome maintenance to mtDNA replication and distribution. Genome Res. ***21***, 12–20.
18. Wang, T., Xie, X.Y., Liu, H.L., et al. (2019). Pyridine nucleotide-disulphide oxidoreductase domain 2 (PYROXD2): role in mitochondrial function. Mitochondrion ***47***, 114–124.
19. Haunhorst, P., Hanschmann, E.M., Bräutigam, L., et al. (2013). Crucial function of vertebrate glutaredoxin 3 (PICOT) in iron homeostasis and hemoglobin maturation. Mol. Biol. Cell. ***24***, 1895–1903.
20. Strittmatter, L., Li, Y., Nakatsuka, N.J., et al. (2014). CLYBL is a polymorphic human enzyme with malate synthase and β-methylmalate synthase activity. Hum. Mol. Genet. ***23***, 2313–2323.
21. Mouillesseaux, K., and Chen, J.N. (2011). Mutation in *utp15* disrupts vascular patterning in a p53-dependent manner in zebrafish embryos. PLoS ONE ***6***, e25013.
22. Ren, P.W., Sun, X.Y., Zhang, C.F., et al. (2019). Human UTP14a promotes angiogenesis through upregulating PDGFA expression in colorectal cancer. Biochem. Biophys. Res. Commun. ***512***, 871–876.
23. Jones, C.A., London, N.R., Chen, H.Y., et al. (2008). Robo4 stabilizes the vascular network by inhibiting pathologic angiogenesis and endothelial hyperpermeability. Nat. Med. ***14***, 448–453.
24. Deng, J.L., Yuan, F.X., Guo, Y.J., et al. (2016). Deletion of ATF4 in AgRP neurons promotes fat loss mainly via increasing energy expenditure. Diabetes ***66***, 640–650.
25. Chen, L., Qiu, Q., Jiang, Y., et al. (2019). Large-scale ruminant genome sequencing provides insights into their evolution and distinct traits. Science ***364***, eaav6202.
26. Spudich, J.A. (2001). The myosin swinging cross-bridge model. Nat. Rev. Mol. Cell Bio. ***2***, 387–392.
27. Jesina, P., Tesarová, M., Fornůsková, D., et al. (2004). Diminished synthesis of subunit a (ATP6) and altered function of ATP synthase and cytochrome c oxidase due to the mtDNA 2 bp microdeletion of TA at positions 9205 and 9206. Biochem. J. ***383***, 561–571.
28. Shestov, A.A., Liu, X.J., Ser, Z., et al. (2014). Quantitative determinants of aerobic glycolysis identify flux through the enzyme GAPDH as a limiting step. eLIFE ***3***, e03342.
29. Saunders, P.U., Garvican-Lewis, L.A., Schmidt, W.F., and Gore, C.J. (2013). Relationship between changes in haemoglobin mass and maximal oxygen uptake after hypoxic exposure. Br. J. Sports Med. ***47***, 26–30.
30. Wang, Y., Zhang, C.Z., Wang, N.N., et al. (2019). Genetic basis of ruminant headgear and rapid antler regeneration. Science ***364***, eaav6335.
31. Hubshman, M.W., Broekman, S., Wijk, E.V., et al. (2018). Whole-exome sequencing reveals POC5 as a novel gene associated with autosomal recessive retinitis pigmentosa. Hum. Mol. Genet. ***27***, 614–624.
32. Ansar, M., Chung, H.L., Taylor, R.L., et al. (2018). Bi-allelic loss-of-function variants in DNMBP cause infantile cataracts. Am. J. Hum. Genet. ***103***, 568–578.
33. Vona, B., Maroofian, R., Bellacchio, E., et al. (2018). Expanding the clinical phenotype of IARS2-related mitochondrial disease. BMC Med. Genet. ***19***, 196.
34. Zhang, X.B., Qin, G., Chen, G.L., et al. (2015). Variants in TRIM44 cause aniridia by impairing PAX6 expression. Hum. Mutat. ***36***, 1164–1167.
35. Tatour, Y., Sanchez-Navarro, I., Chervinsky, E., et al. (2017). Mutations in SCAPER cause autosomal recessive retinitis pigmentosa with intellectual disability. J. Med. Genet. ***54***, 698–704.
36. Alexander, C., Votruba, M., Pesch, U.E., et al. (2000). OPA1, encoding a dynamin-related GTPase is mutated in autosomal dominant optic atrophy linked to chromosome 3q28. Nat. Genet. ***26***, 211–215.
37. Lin, Q., Fan, S.H., Zhang, Y.H., et al. (2016). The seahorse genome and the evolution of its specialized morphology. Nature ***540***, 395–399.
38. Jiang, H.F., Du, K., Gan, X.N., et al. (2019). Massive loss of olfactory receptors but not trace amine-associated receptors in the worlds deepest-living fish (*Pseudoliparis swirei*). Genes ***10***, 910.
39. Madigan, J., Carlisle, A.B., Gardner, L.D., et al. (2015). Assessing niche width of endothermic fish from genes to ecosystem. Proc. Natl. Acad. Sci. U S A ***112***, 8330–8355.
40. Hu, Y.B., Wu, Q., Ma, S., et al. (2017). Comparative genomics reveals convergent evolution between the bamboo-eating giant and red pandas. Proc. Natl. Acad. Sci. U S A ***114***, 1081–1086.
41. Yu, Z.P., Seim, I., Yin, M.X., et al. (2021). Comparative analyses of aging-related genes in long-lived mammals provide insights into natural longevity. Innovation ***2***, 100108. https://doi.org/10.1016/j.xinn.2021.100108.
42. Zhang, Y.H., Ravi, V., Qin, G., et al. (2020). Comparative genomics reveal shared genomic changes in syngnathid fishes and signatures of genetic convergence with placental mammals. Natl. Sci. Rev. ***7***, 964–977.
43. Wu, B.S., Feng, C.G., Zhu, C.L., et al. (2021). The genomes of two billfishes provide insights into the evolution of endothermy in teleosts. Mol. Biol. Evol. ***38***, 2413–2427.
44. Bernal, D., Dickson, K.A., Shadwick, R.E., and Graham, J.B. (2001). Review: analysis of the evolutionary convergence for high performance swimming in lamnid sharks and tunas. Comp. Biochem. Physiol. A. Mol. Integr. Physiol. ***129***, 695–726.
45. Runcie, R.M., Dewar, H., Hawn, D.R., et al. (2009). Evidence for cranial endothermy in the opah (*Lampris guttatus*). J. Exp. Biol. ***212***, 461–470.
46. Fritsches, K.A., Brill, R.W., and Warrant, E.J. (2005). Warm eyes provide superior vision in swordfishes. Curr. Biol. ***15***, 55–58.

47. Cresswell, P. (1994). Assembly, transport, and function of MHC class II molecules. Annu. Rev. Immunol. ***12***, 259–293.

ACKNOWLEDGMENTS

We are grateful to Profs. Byrappa Venkatesh, Fuwen Wei, Yanhua Qu, and Dr. Yan Hao for their insightful comments and suggestions that improved the manuscript. This research was supported by the Key Research Program of Frontier Sciences of CAS (ZDBS-LY-DQC004), the Special Foundation for National Science and Technology Basic Research Program of China (2018FY100100), the National Natural Science Foundation of China (41825013), the Key Special Project for Introduced Talents Team of Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou) (GML2019ZD0407), the Key Special Project for Introduced Talents Team of Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou) (GML2019ZD0401), the Guangdong Special Support Program of Leading Scientific and Technological Innovation (2017 T × 04 N442), and the Grants by the German Science Foundation (DFG).

AUTHOR CONTRIBUTIONS

Qiang Lin, Dazhi Wang, and Axel Meyer conceived and designed the study. Xin Wang, Shaobo Ma, and Shuaishuai Liu collected the samples. Xin Wang, Meng Qu, Yali Liu, Yue Song, Ralf Schneider, Zelin Chen, Haiyan Yu, Geng Qin, Jianping Yin, Guangyi Fan, Axel Meyer, and Jia Zhong performed genome analyses. Zelin Chen, Yanhong Zhang, and Haiyan Yu performed TE analysis. Suyu Zhang, Meng Qu, and Xin Wang performed convergent analyses. Yali Liu and Haiyan Yu performed transcriptomic analysis. Hao Zhang and Dongxu Li performed proteomic analysis. Xin Wang, Meng Qu, Yali Liu, Dazhi Wang, and Qiang Lin wrote the manuscript with input from all other authors. All authors reviewed and contributed to the final manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interest.

LEAD CONTACT WEBSITE

Prof. Qiang LIN http://people.ucas.ac.cn/~0018009?language=en.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.xinn.2021.100185.

The Innovation, Volume ■■

Supplemental Information

Genomic basis of evolutionary adaptation in a warm-blooded fish

Xin Wang, Meng Qu, Yali Liu, Ralf F. Schneider, Yue Song, Zelin Chen, Hao Zhang, Yanhong Zhang, Haiyan Yu, Suyu Zhang, Dongxu Li, Geng Qin, Shaobo Ma, Jia Zhong, Jianping Yin, Shuaishuai Liu, Guangyi Fan, Axel Meyer, Dazhi Wang, and Qiang Lin

Supplemental Information

Genomic basis of evolutionary adaptation in a warm-blooded fish

Xin Wang, Meng Qu, Yali Liu, Ralf F. Schneider, Yue Song, Zelin Chen, Hao Zhang, Yanhong Zhang, Haiyan Yu, Suyu Zhang, Dongxu Li, Geng Qin, Shaobo Ma, Jia Zhong, Jianping Yin, Shuaishuai Liu, Guangyi Fan, Axel Meyer, Dazhi Wang, Qiang Lin

This Supplemental file includes:

Materials and Methods

Figures S1-S11

Tables S1-S11, S25

Captions for Tables S12-S24, S26-S33 (.xlsx)

Supplemental Materials and Methods

Sample collection and identification

A mature male opah with a total length of 110 cm and with a body weight of approximately 65 kg was collected from the East China Sea by pelagic fishing, and seven tissues, the PMI, PMS, DM, VM, liver, kidney, and gill, were dissected. The opah was identified as a Smalleye Pacific Opah, *Lampris incognitus*, on the basis of morphological observations and genetic markers reported by Underkoffler et al in 2018[1].

Sequencing and assembly of *Lampris incognitus*

The opah genome was sequenced using DNA isolated from muscle samples. The Illumina short reads from the 250 bp and 500 bp libraries were used for *k*-mer frequency analyses of the genome. The genomic DNA was fragmented to 20 kb and sequenced with the Pacific Biosystems RSII platform (Pacific Biosciences, USA). After filtering out low-quality and duplicate reads, we generated 20.98 million PacBio clean reads (totalling 183.23 Gb, an approximately 136-fold coverage of the estimated 1.37 Gb genome) with N50 read lengths of 15 kb and max read lengths of 41 kb. The Pacbio clean reads were corrected using the error correction module of Canu v.1.5 to select for longer subreads. The output subreads of Canu were further corrected using Illumina

paired-end reads. Based on these two rounds of error-corrected subreads, we generated a draft assembly using SMARTdenovo. We constructed Hi-C fragment libraries from 300-700 bp insert size and sequencing through Illumina HiSeq 2500 platform, and 140.01 Gb Hi-C data was generated, covering approximately 102-fold of the estimated 1.37 Gb genome. HiC-Pro (version 2.8.1) was used for duplicate removal and quality controls, and the remaining reads were valid interaction pairs for further assembly. Before chromosomes assembly, we first performed a preassembly for error correction of scaffolds which required the splitting of scaffolds into segments of 50 kb on average. The Hi-C data were mapped to these segments using BWA (version 0.7.10-r789) software. The uniquely mapped data were retained to perform assembly by using LACHESIS software. Any two segments which showed inconsistent connection with information from the raw scaffold were checked manually. These corrected scaffolds were then assembled with LACHESIS. In addition, CEGMA (v2.5)[2] and BUSCO (v2, database: vertebrata_odb9)[3] were employed to evaluate the integrity of the assembled sequence.

Genome annotation

Repeat elements in the opah genome were identified using RepeatMasker[4] with the Repbase transposable element library[5]. Protein-coding genes were annotated using a combination of *de novo* gene prediction programs and homology-based methods, and along with RNA-Seq data, were used to predict gene models in the opah genome. *De novo* prediction was performed using Genscan[6], Augustus[7], GlimmerHMM[8], GeneID[9], and SNAP[10]. For the homology-based analysis, we aligned the protein sequences of *Danio rerio* (GCF000002035.6), *Oryzias latipes* (GCF002234675.1), and *Hippocampus comes* (GCA001891065.1) to the opah genome, conducted a precise spliced alignment, and predicted gene structures using GeMoMa[11]. Transcriptome-based predictions were conducted using TransDecoder (http://transdecoder.github.io), GeneMarkS-T[12], and PASA[13]. Finally, we combined the results from these three methods with EvidenceModeler[14], and all gene models were further annotated using TrEMBL[15], NR[16], GO[17], KOG[18], and KEGG[19] databases. GO enrichment analysis was performed using the OmicShare tools, a free online platform for data analysis (www.omicshare.com/tools).

Ortholog identification and phylogenetic analysis

Complete single copy genes were predicted using the Augustus software (version 3.3), combined with HMM models based on Benchmarking Universal Single-Copy Orthologs

(BUSCO) of Actinopterygii, from the genomes of opah and additional 16 ray-finned fishes, including zebrafish (*Danio rerio*, GCF000002035.6), stickleback (*Gasterosteus aculeatus*, GCA016920845.1), fugu (*Takifugu rubripes*, GCF901000725.2), ocean sunfish (*Mola mola*, GCA001698575.1), Pacific bluefin tuna (*Thunnus orientalis*, http://nrifs.fra.affrc.go.jp/ResearchCenter/5_BB/genomes/Tuna_DNAmicroarray/index.html), tongue sole (*Cynoglossus semilaevis*, GCF000523025.1), Atlantic cod (*Gadus morhua*, GCF902167405.1), snailfish (*Liparis tanakae*, GCA006348945.1), Japanese medaka (*Oryzias latipes*, GCF002234675.1), lined seahorse (*Hippocampus erectus*, PRJNA613175), Asian bonytongue (*Scleropages formosus*, GCF900964775.1), giant fin mudskipper (*Periophthalmus magnuspinnatus*, GCF009829125.1), Chinese large-mouth catfish (*Silurus meridionalis*, GCA014805685.1), Indo-Pacific tarpon (*Megalops cyprinoides*, GCF013368585.1), northern pike (*Esox Lucius*, GCF011004845.1), and spotted gar (*Lepisosteus oculatus*, GCF000242695.1) as outgroup. For each gene family, protein sequences of orthologous genes were extracted from different species and multiple alignments were performed using MAFFT. The trimmed multiple alignment of each gene family was generated using Gblocks with the allowed gap positions set to 'With Half'. The ModelFinder software was used to identify the best suitable model of the trimmed alignment, and the maximum likelihood tree was generated using IQtree with 100 bootstrap replicates. The final species tree was inferred using the ASTRAL-II software with individual IQtree-estimated protein trees of all identified single-copy gene families.

RNA sequencing and analysis

Raw reads from seven different tissues, in a FASTQ format, were first processed using in-house Perl scripts. Then, clean reads were mapped to the opah genome using the TopHat2 program. Gene expression levels were estimated with fragments per kilobase of exon per million fragments (FPKM values) using the Cufflinks program. A Pearson correlation across all samples was conducted to evaluate the correlation across all samples.

Transcriptomes from the seven tissues were used to identify tissue-specific expressed genes in PMI. We calculated tissue specificity indices as follows: first, we filtered out weakly expressed genes (the maximum FPKM in all tissues was less than 1), then we used the remaining opah genes to calculate τ tissue-specific expression indices according to previous studies[20,21]. Genes specifically expressed in the dark red aerobic pectoral musculature were defined as those that have a τ index exceeding 0.8 and have the highest, or second highest, expression in the dark red aerobic

pectoral musculature. The DEGs were checked using DESeq2 and were identified based on corrected p-values (Q-value) and false discovery rates with a log2 (fold change) > 1 and a corrected P value < 0.05.

Experimental procedure for the proteomic analysis

Three pseudo-replicates of each fish tissue (PMI, PMS, VM, LI, GI) were grinded while submerged in liquid nitrogen and then suspended in 10% w/v trichloroacetic acid/acetone for protein extraction, following the published protocol[22]. Protein quantification was performed using a 2D Quant kit (GE Healthcare, San Francisco, CA). After adjusting the pH to 8.5 with 1 M ammonium bicarbonate, 200 μg of protein from each sample was transferred into a Microcon-10 kDa Centrifugal Filter Unit with Ultracel-10 membrane (Merck Millipore, USA), reduced with 50 mM dithiothreitol, carboxyamidomethylated with 50 mM iodoacetamide, and then washed five times with 8 M urea to elute impurities. Proteins were then digested twice with Trypsin Gold (Promega, Madison, WI, USA), desalted with Sep-Pak C18 solid phase column (Waters, Milford, MA, USA), and then evaporated and reconstituted in 0.1% formic acid. Peptide fractionations were performed through a peptide bridged ethyl hybrid (BEH) C18 Column (2.1×150 mm, 1.7 μm, Waters, USA) using an ACQUITY UPLC H-Class system (Waters, Milford, MA, USA). Each peptide sample was separated into 10 fractions, desalted, vacuum-dried, and re-dissolved in 0.1% formic acid.

Each fraction of the peptide mixtures was analysed with an Orbitrap Fusion Lumos Tribrid mass spectrometer (Thermo Scientific, USA) interfaced to a Dionex UltiMate3000 RSLCnano HPLC system (Thermo Scientific, USA). The mass spectrometer was operated in the data dependent acquisition mode. The parameter settings for MS1 were: orbitrap resolution at 120,000, scan range at 350–1800 m/z, auto gain control (AGC) target at 4.0e5, maximum injection time of 50 ms, radio frequency lens at 40%, centroid data type, charge state at 2–7, dynamic exclusion for 30 s using a mass tolerance of 10 ppm, and internal calibration using m/z 445.12. Parameter profiles for dd MS^2 were: orbitrap resolution at 15,000, isolation window at 1.6 m/z, HCD activation type with 30% collision energy, maximum injection time of 40 ms, AGC target at 5.0e4, and with a centroid data type.

Spectral matching was performed using the MaxQuant software (Version 1.5.2.8) against a protein sequence database constructed from the genome and metatranscriptome[23]. To reduce the probability of false identification, the false discovery rate estimated by a target-decoy approach

was set to less than 1% for both peptide and protein levels. Both unique and razor peptides were used to obtain intensity-based absolute quantification (iBAQ) values calculated by MaxQuant software. DEPs were chosen based on a fold change > 1.5 and a P value < 0.05. Proteins were functionally annotated against the NCBI non-redundant protein (NCBInr), GO, and KEGG databases. GO and KEGG enrichment analyses for DEPs were conducted by GOseq R package and KOBAS software, respectively.

Gene family expansion/contraction

Protein datasets of opah and another eight ray-finned fishes were involved in the gene family expansion/contraction analysis, including zebrafish, stickleback, fugu, ocean sunfish, tongue sole, Atlantic cod, snailfish, and spotted gar. The divergence time was estimated using MCMCtree version 4.5[24] with topology of the species tree and three calibration time points, including zebrafish-stickleback: ~149.85–165.2 Mya[25], snailfish-stickleback: 32–73 Mya[26], and ocean sunfish-fugu: 60.8–80.8 Mya[27]. The expansion and contraction of gene families in the opah genome were calculated by the CAFÉ program (v 3.1) based on the birth-and-death model[28]. The parameters '-p 0.01, -r 10,000, -s' were set to search the birth and death parameters (λ) of genes based on a Monte Carlo resampling procedure, and birth and death parameters in gene families with the p-values ≤ 0.01 were reported. The gene families with no homology in the SWISS-PROT database were filtered out to reduce the potential false positive expansions or contractions caused by gene prediction.

Identification of PSGs

One-to-one orthologous genes were obtained from the same species selected for the gene family expansion/contraction analysis. Positive selection analyses were conducted with the branch-site model using PAML[29]. We compared model A (allows sites to be under positive selection; fix_omega = 0) with the null model A1 (sites may evolve neutrally or under purifying selection; fix_omega = 1 and omega = 1), with likelihood ratio tests (LRT) performed by the Codeml program in PAML. Significance (P < 0.05) (FDR adjusted P-value < 0.05) of the compared LRTs was evaluated by $\chi 2$ tests from PAML.

Convergent evolution analysis

To find convergent positively selected genes underlying the emergence of endothermy, positive selection analyses were conducted with three different datasets by either using the opah, Pacific bluefin tuna, or the great white shark as foreground branches. Then the positively selected genes were intersected for further analysis.

The Ka/Ks ratio of all orthologous genes of the foreground branches (opah, Pacific bluefin tuna, and great white sharks) were calculated by setting all three endothermic fishes as a foreground simultaneously with branch model under two assumptions: a one-ratio model that assumes the same Ka/Ks ratio for two clades (foreground and background branches) and a two ratio model in which the two clades were assigned to different Ka/Ks ratios.

We investigated genome-wide signatures of convergent evolution related to endothermy in the lineages of five endothermic species (opah, turkey, chicken, mouse, and human) and three endothermy fish species (opah, Pacific bluefin tuna, and great white sharks), respectively. PSGs were first identified by setting all endothermic species as foreground branches using a branch-site model. Then the most recent common ancestor (MRCA) sequence of each inner-node in species tree was reconstructed by using Codeml program in PAML software. Comparing the amino acid residue of each endotherm species with that of its most recent common ancestor amino acid residue, the observed number of convergence and parallel substitution sites were respectively counted according to Storz's criteria[30]. To filter out the noise caused by random mutation, expected numbers of these two types of substitution under JTT-f_{genes} amino acid substitution models[31] of each PSG were calculated between the endotherm species, and MRCA and Poisson test were conducted to verify statistically significant differences ($P < 0.05$) between observed and excepted values. Finally, a gene with non-random convergent and/or parallel substitution amino acid sites that also was identified as a PSG was considered to have undergone convergent evolution[32,33].

Supplemental References

1. Underkoffler, K.E., Luers, M.A., Hyde, J.R., and Craig, M.T. (2006). A taxonomic review of *Lampris guttatus* (Brünnich 1788) Lampridiformes; Lampridae) with descriptions of three new species. Zootaxa ***4413***, 551-565.
2. Parra, G., Bradnam, K., and Korf, I. (2007). CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. Bioinformatics ***23***, 1061-1067.
3. Simão, F.A., Waterhouse, R.M., Panagiotis, I., et al. (2015). BUSCO: assessing genome

assembly and annotation completeness with single-copy orthologs. Bioinformatics ***31***, 3210-3212.

4. Smit, A.F.A., Hubley, R., and Green, P. (2013). RepeatMasker Open-4.0; http://www.repeatmasker.org. Institute for Systems Biology.
5. Jurka, J., Kapitonov, V.V., and Pavlicek, A. (2005) Repbase Update, a database of eukaryotic repetitive elements. Cytogenet. Genome Res. ***110***, 462-467.
6. Burge, C., and Karlin, S. (1997). Prediction of complete gene structures in human genomic DNA. J. Mol. Biol. ***268***, 78-94.
7. Stanke, M., and Waack, S. (2003). Gene prediction with a hidden Markov model and a new intron submodel. Bioinformatics ***2***, 215-225.
8. Majoros, W.H., Pertea, M., and Salzberg, S.L. (2004). TigrScan and GlimmerHMM: two open source ab initio eukaryotic gene-finders. Bioinformatics ***20***, 2878-2879.
9. Alioto, T., Blanco, E., Parra, G., and Guigó, R. (2018). Using geneid to Identify Genes. Curr. Protoc. Bioinformatics ***64***, e56.
10. Korf, I. (2004). Gene finding in novel genomes. BMC Bioinformatics ***5***, 59.
11. Keilwagen, J., Wenk, M., Erickson, J.L., et al. (2016). Using intron position conservation for homology-based gene prediction. Nucleic Acids Res. ***44***, e89.
12. Tang, S., Alexandre, L., and Borodovsky, M. (2015). Identification of protein coding regions in RNA transcripts. Nucleic Acids Res. ***43***, e78.
13. Campbell, M.A., Haas, B.J., Hamilton, J.P., et al. (2006). Comprehensive analysis of alternative splicing in rice and comparative analyses with Arabidopsis. BMC Genomics. ***7***, 327.
14. Haas, B.J., Salzberg, S.L., and Zhu, W., et al. (2008). Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. Genome Biol. ***9***, R7.
15. Boeckmann, B., Bairoch, A., Apweiler, R., et al. (2003). The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. Nucleic Acids Res. ***31***, 365-370.
16. Marchler-Bauer, A., Lu, S.N., Anderson, J.B., et al. (2011). CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Res. ***39***, 225-229.

17. Conesa, A., Götz, S., García-Gómez, J.M., et al. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics ***21***, 3674-3676.
18. Tatusov, R.L., Natale, D.A., Garkavtsev, I.V., et al. (2001). The COG database: new developments in phylogenetic classification of proteins from complete genomes. Nucleic Acids Res. ***29***, 22-28.
19. Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. ***28***, 27-30.
20. Kryuchkova-Mostacci, N., and Robinson-Rechavi, M. (2017). A benchmark of gene expression tissue-specificity metrics. Brief Bioinform. ***18***, 205-214.
21. Wang, Y., Zhang, C.Z., Wang, N.N., et al. (2019). Genetic basis of ruminant headgear and rapid antler regeneration. Science ***364***, eaav6335.
22. Wang, M.H., Wang, Y.Y., Zhang, L., et al. (2013). Quantitative proteomic analysis reveals the mode-of-action for chronic mercury hepatotoxicity to marine medaka (*Oryzias melastigma*). Aquat. Toxicol. ***130***, 123-131.
23. Tyanova, S., Temu, T., and Cox, J. (2016). The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. Nat. Protoc. ***11***, 2301-2319.
24. Yang, Z.H. (2007). PAML 4: Phylogenetic Analysis by Maximum Likelihood. Mol. Biol. Evol. ***24***, 1586-1591.
25. Lin, Q., Fan, S.H., Zhang, Y.H., et al. (2016). The seahorse genome and the evolution of its specialized morphology. Nature ***540***, 395-399.
26. Pan, H.L., Yu, H., Ravi, V., et al. (2016). The genome of the largest bony fish, ocean sunfish (*Mola mola*), provides insights into its fast growth rate. Gigascience ***5***, 36.
27. Wang, K., Shen, Y.J., Yang, Y.Z., et al. (2019). Morphology and genome of a snailfish from the Mariana Trench provide insights into deep-sea adaptation. Nat. Ecol. Evol. ***3***, 823-833.
28. Bie, T.D., Cristianini, N., Demuth, J.P., and Hahn, M.W. (2006). CAFE: a computational tool for the study of gene family evolution. Bioinformatics ***22***, 1269-1271.
29. Mitchell, C.R., Harris, M.B., Cordaro, A.R., and Starnes, J.W. (2002). Effect of body temperature during exercise on skeletal muscle cytochrome c oxidase content. J. Appl. Physiol. ***93***, 526-530.
30. Storz, J.F. (2016). Causes of molecular convergence and parallelism in protein evolution.

Nat. Rev. Genet. ***17***, 239-250.

31. Jones, D.T., Taylor, W.R., and Thornton, J.M. (1992). The rapid generation of mutation data matrices from protein sequences. Bioinformatics ***8***, 275-282.
32. Hu, Y.B., Wu, Q., Ma, S., et al. (2017). Comparative genomics reveals convergent evolution between the bamboo-eating giant and red pandas. Proc. Natl. Acad. Sci. USA ***114***, 1081-1086.
33. Hao, Y., Xiong, Y., Cheng, Y.L., et al. (2019). Comparative transcriptomics of 3 high-altitude passerine birds and their low-altitude relatives. Proc. Natl. Acad. Sci. USA. ***116***, 11851-11856.

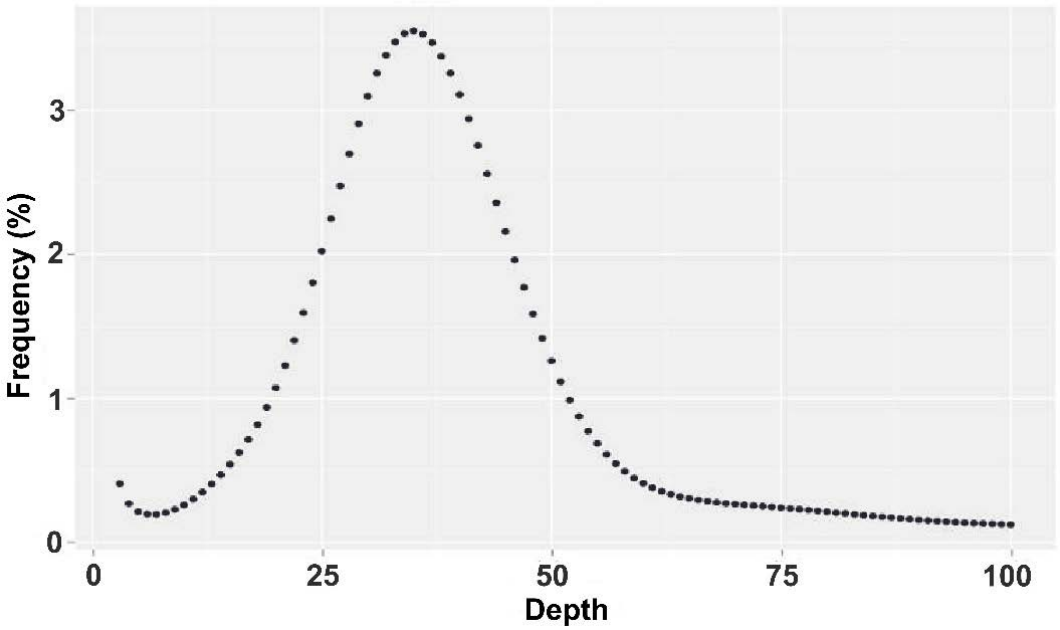


Figure S1. *K-mer* distribution of the *Lampris incognitus* genome. The *K-mer* spectrum was constructed based on a 19-*mer*. The figure shows the *K-mer* spectrum of raw reads. The x-axis represents the *K-mer* depth while the y-axis represents the proportion of 19-mer with different coverage, and the K_{num} is 48,102,985,862 based on 19-mer and K_{depth} is 35. Therefore, the estimated genome size of *L. incognitus* is around 1,374.37 Mb.

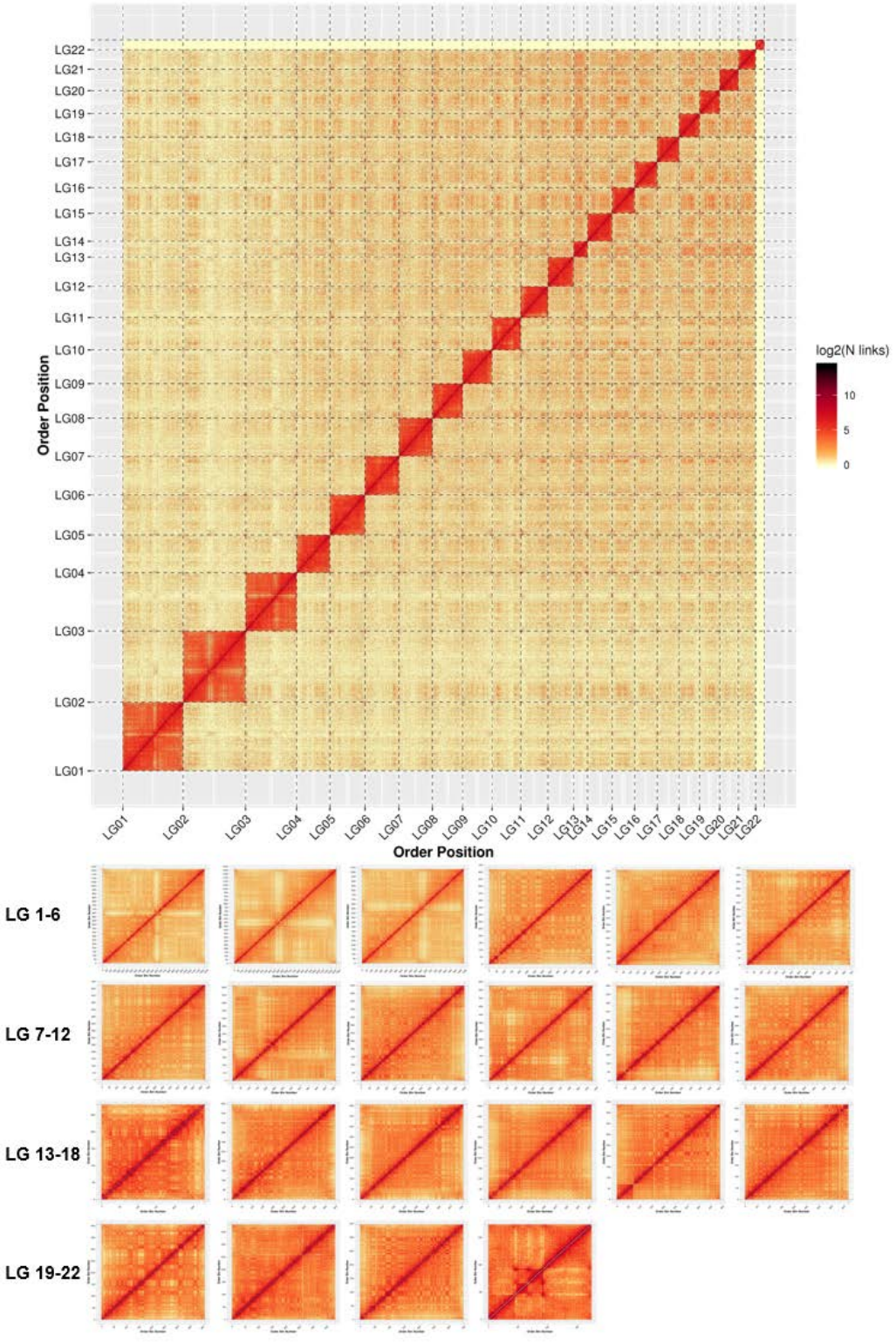


Figure S2. A Hi-C contact map of of *Lampris incognitus* genomes assembly. Chromosome-level genome assemblies comprising 22 chromosome-level scaffolds.

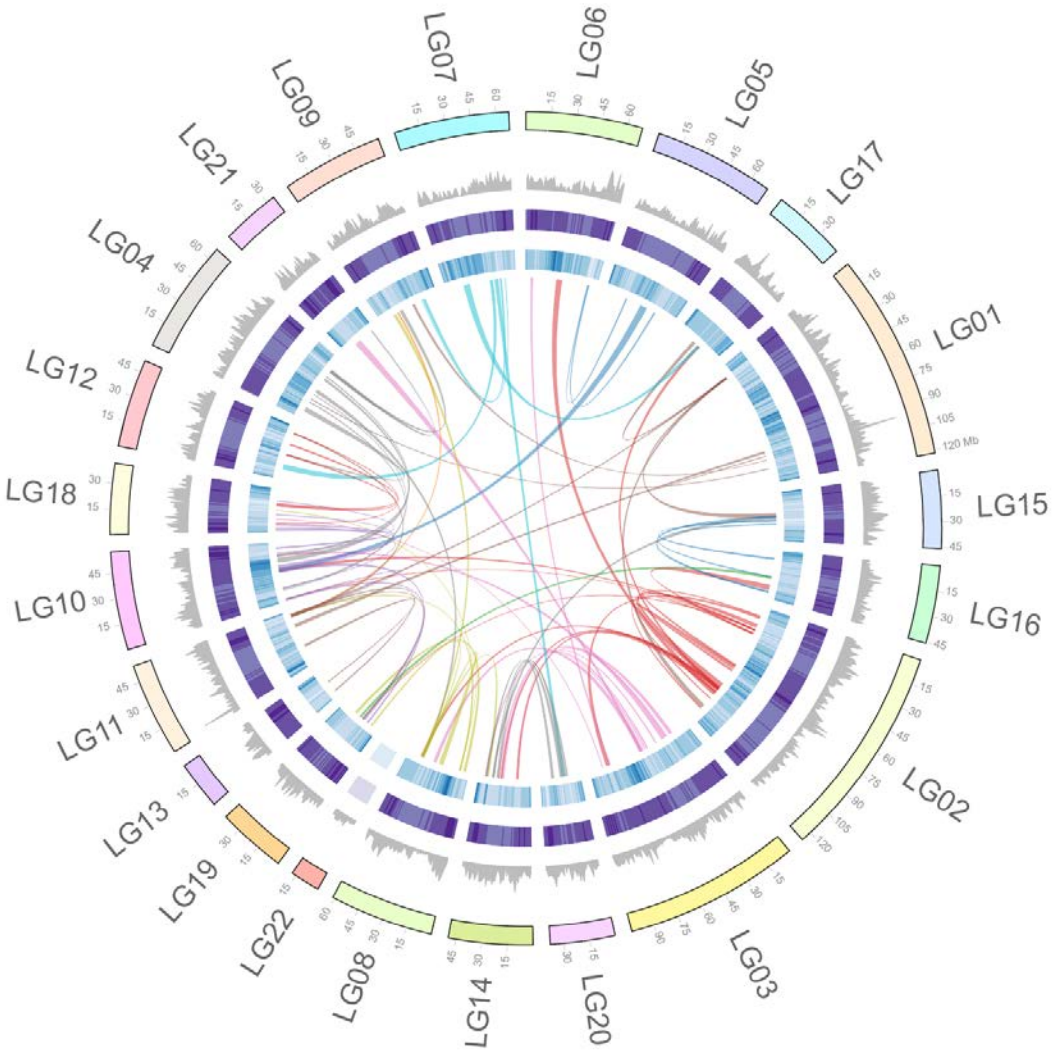


Figure S3. Genome feature of *Lampris incognitus*. The outside circle showing the 22 pseudochromosomes in opah, and the tick marks on the axis for LG01 are common for all such axes (Mb). The grey circle represents the gene density; The purple circle represents GC content. The blue circle represents density of repeated elements. The centre shows syntenic links between genes located on 22 pseudochromosomes.

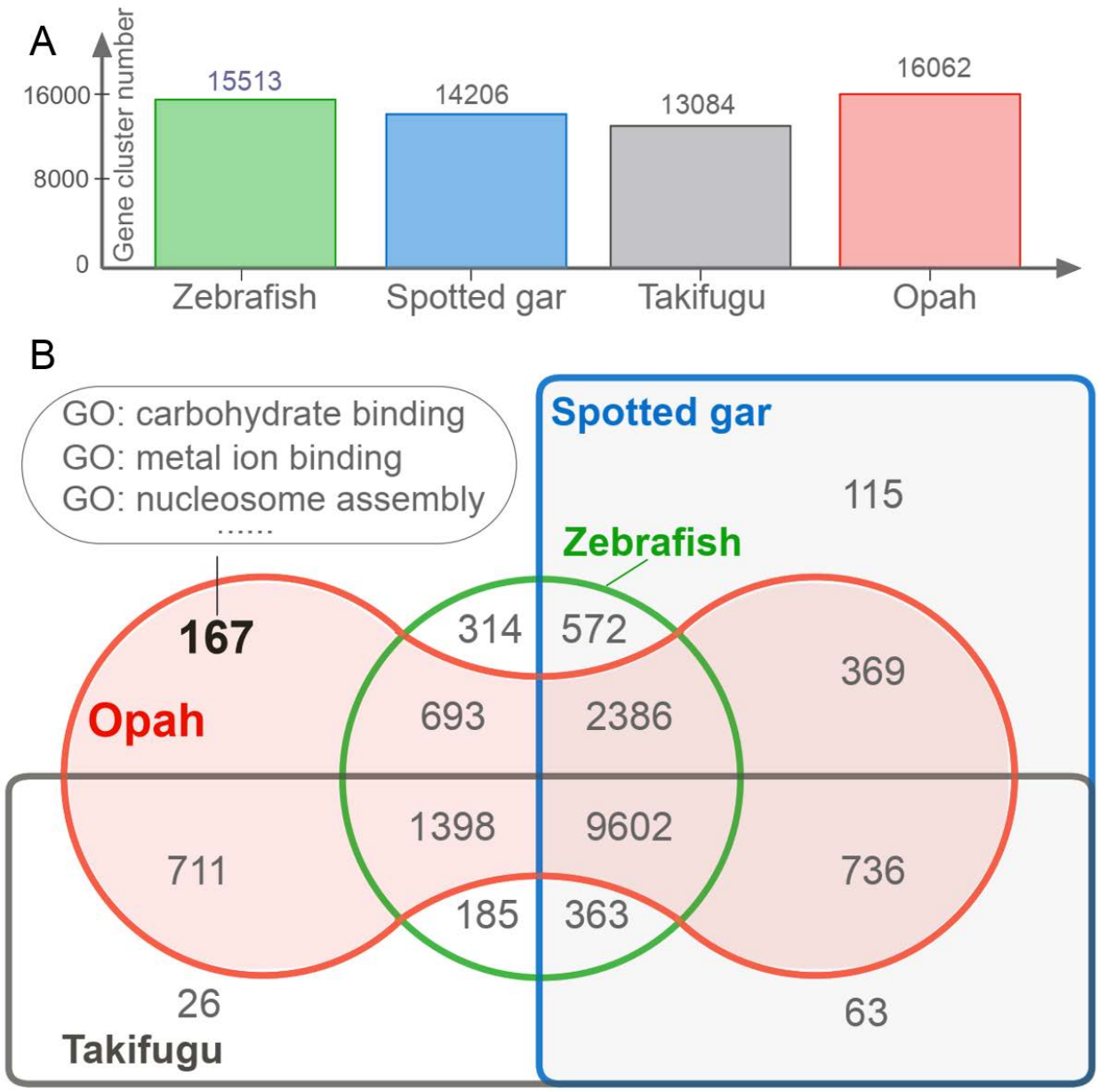


Figure S4. Venn diagram of the distribution of shared orthologous clusters among the 4 teleost species. (A) Orthologous cluster numbers detected in the genomes of opah, zebrafish, takifugu, and spotted gar. (B) Venn diagram of shared orthologous gene families in opah, zebrafish, takifugu, and spotted gar.

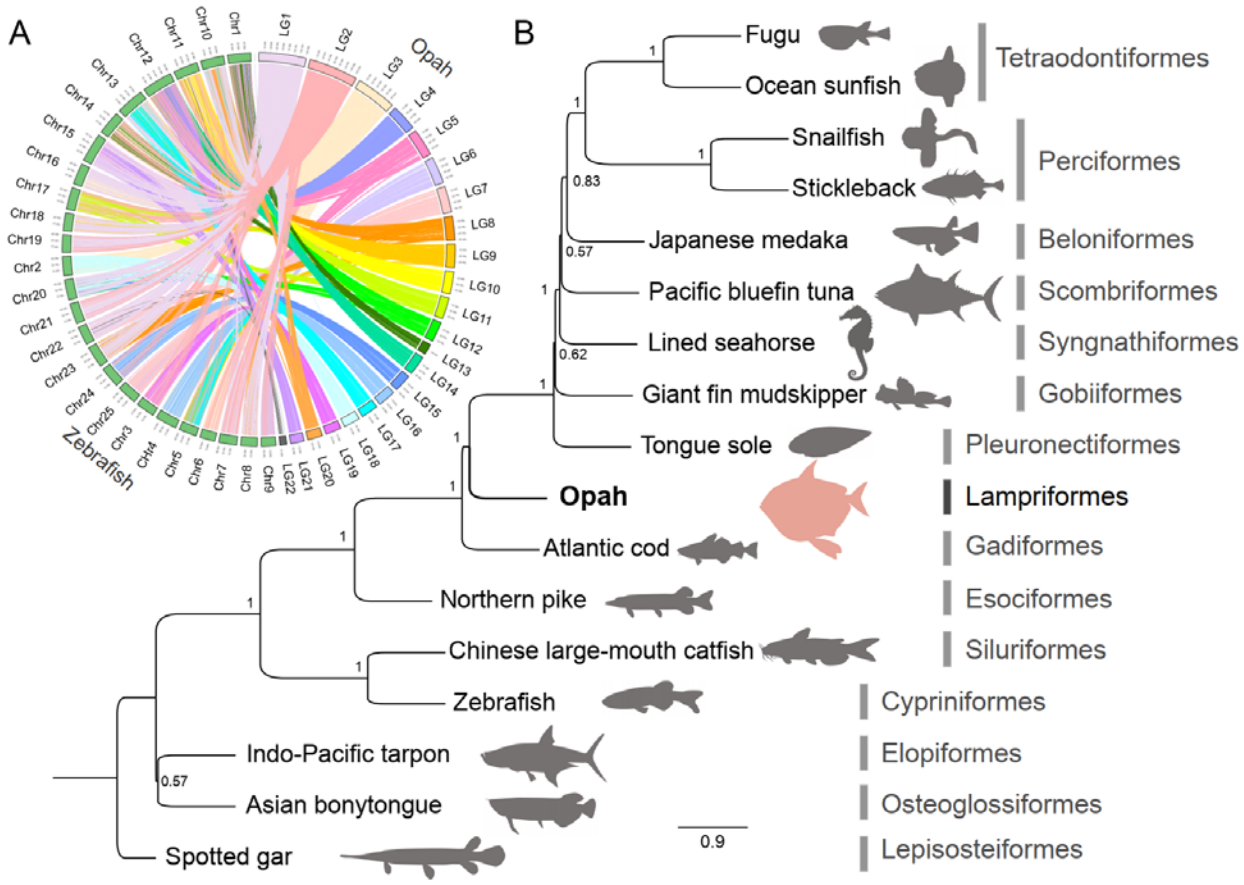


Figure S5. Syntenic and phylogenetic analysis of opah. (A) Syntenic regions between opah pseudochromosome and zebrafish chromosomes. (B) Phylogeny of opah and other teleost fishes, the spotted gar was used as the outgroup species.

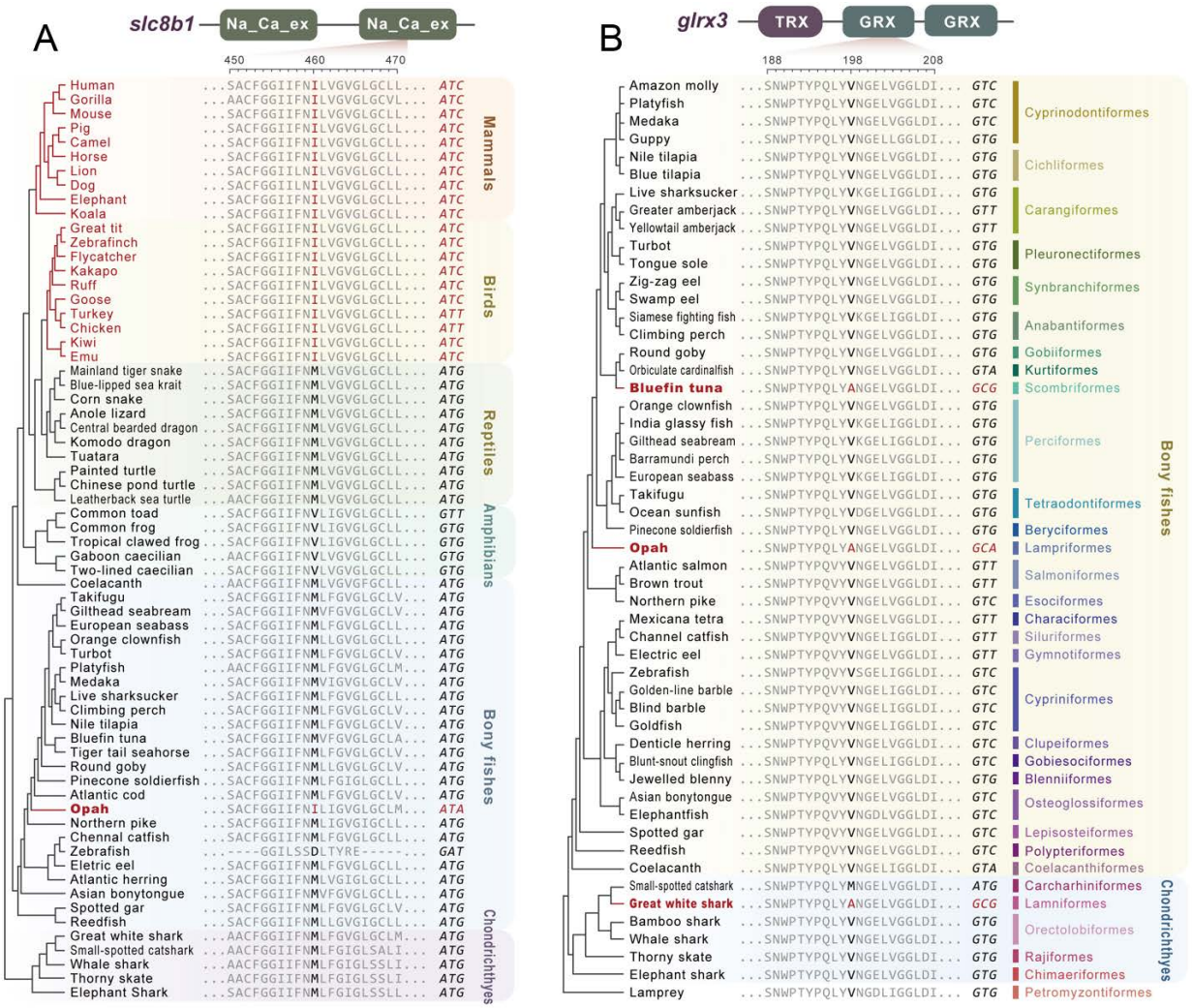


Figure S6. Convergent molecular evolution in Slc8b1 and Glrx3 proteins in endothermic species. (A) Comparison of the convergent amino acid substitution of Slc8b1 protein among 64 vertebrate species (including the species used in the genome-wide screen). The convergent site M460I is located in the Na_Ca_ex domain and the amino acid M only occurs in endothermic mammals, birds, and opah (red). (B) Comparison of the convergent amino acid substitution of Glrx3 protein among lamprey and 51 fish species (including the species used in the genome-wide screen). The convergent site V198A is located in the GRX domain and the amino acid A only occurs in three endothermic fishes (red).

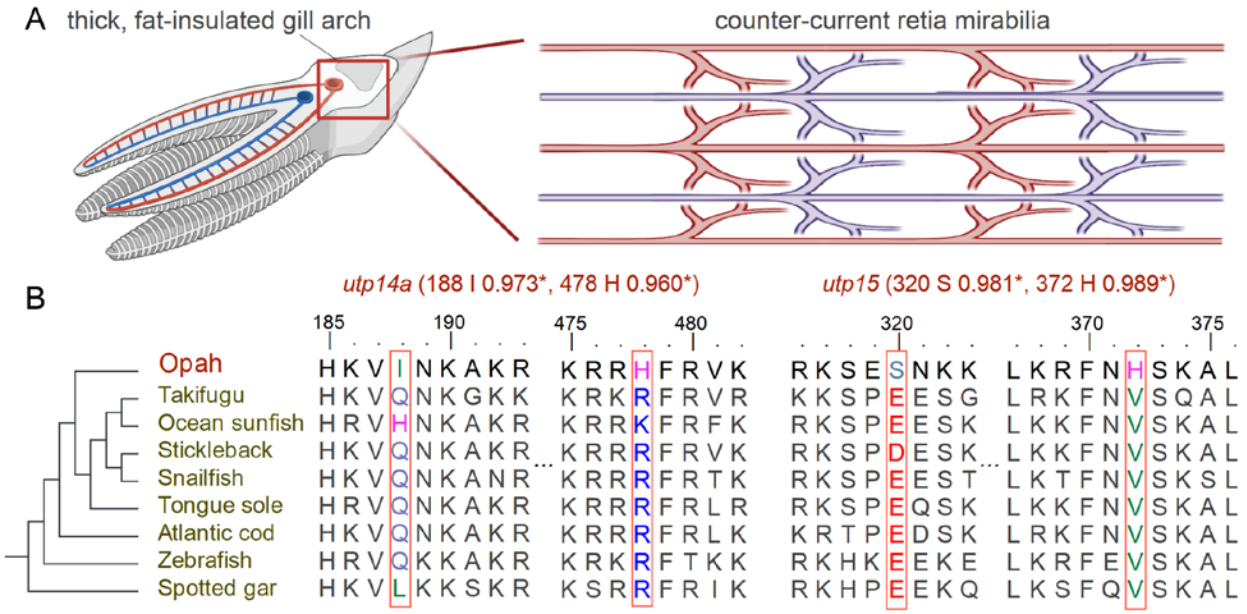


Figure S7. Comparison of the amino acid substitution of positive selected genes (*utp14a* and *utp15*) involved in the development of counter-current heat exchanger. (A) Schematic representation of the counter-current retia mirabilia inside the gill arch of opah. (B) Alignment of Utp14a and Utp15 amino acid sequences of opah and other 8 teleost species, and the positively selected sites are boxed in red.

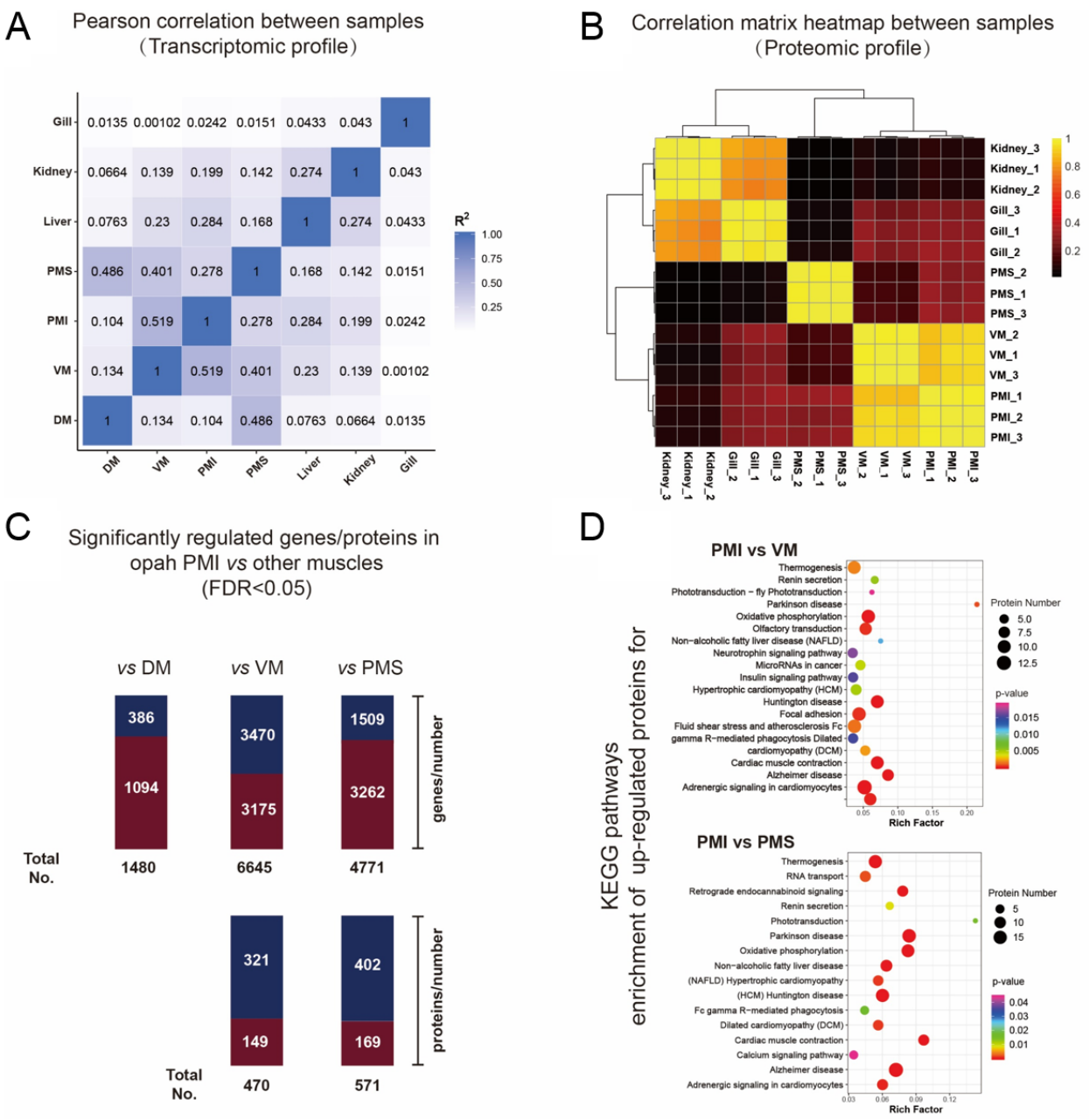


Figure S8. Transcriptomic and proteomic analysis of different opah tissues. (A) Transcriptomic correlation analysis of sequenced tissues. Pearson analysis was conducted to evaluate the correlation across all samples based on gene expression levels (FPKM values). (B) Correlation matrix heatmap between several tissues based on proteomic expression level. (C) Number of significantly regulated genes (or proteins) in opah PMI compared to other muscles. Red: up-regulated, blue: down-regulated. (D) Enriched KEGG pathways of up-regulated proteins in PMI compared to VM and PMS.

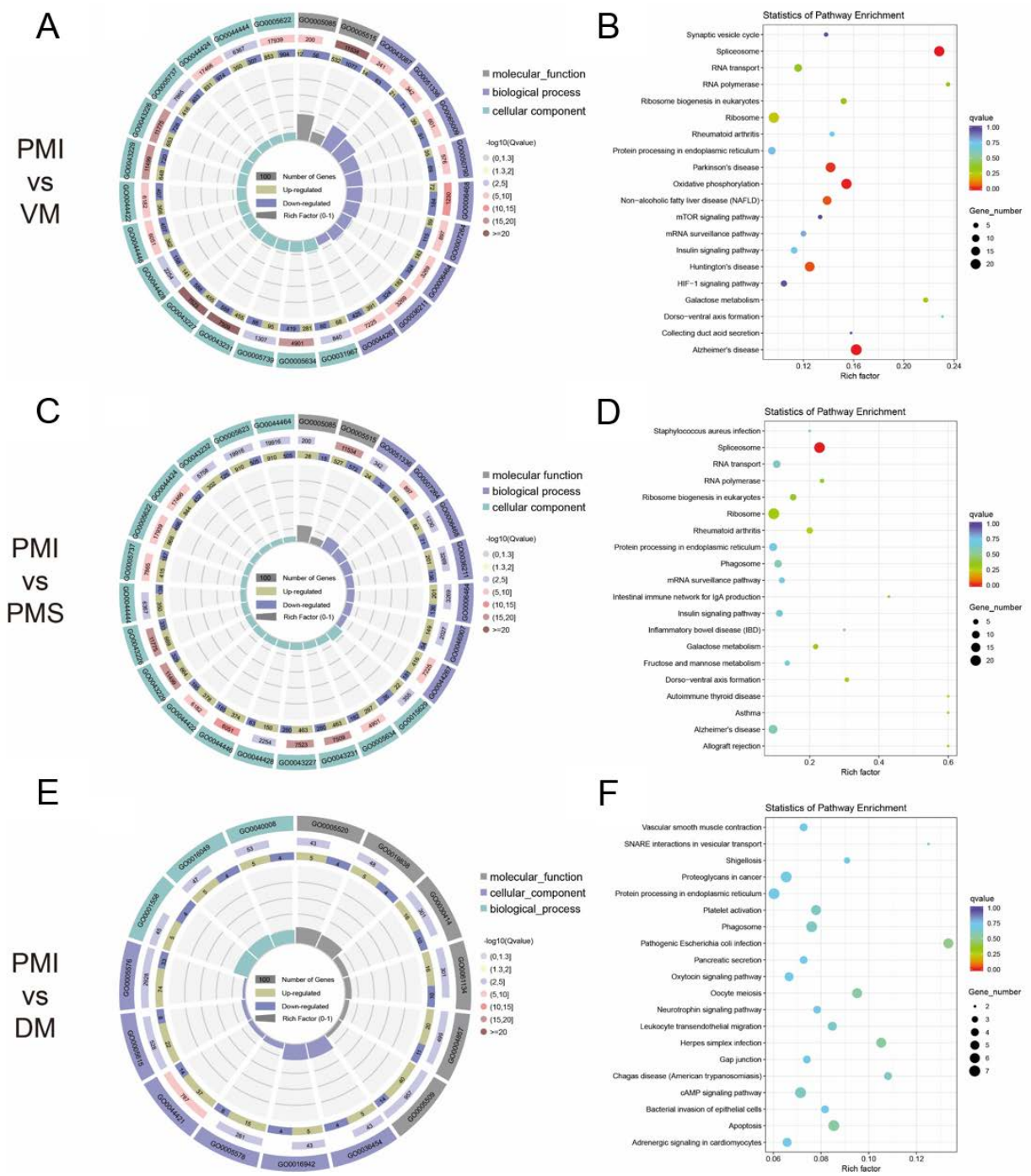


Figure S9. GO and KEGG enrichment analysis of DEGs between dark red aerobic pectoral musculature and other muscles. (A,C,E) Gene Ontology (GO) enrichment analysis of the DEGs was implemented by the GOseq R packages. KOBAS software was used to test the statistical enrichment of DEGs in KEGG pathways. (B,D,F) KEGG analysis of these genes revealed that the most significant enriched pathways of the up-regulated genes of the dark red aerobic pectoral musculature showed the higher expression of genes involved in oxidative phosphorylation, galactose metabolism and insulin signaling pathway.

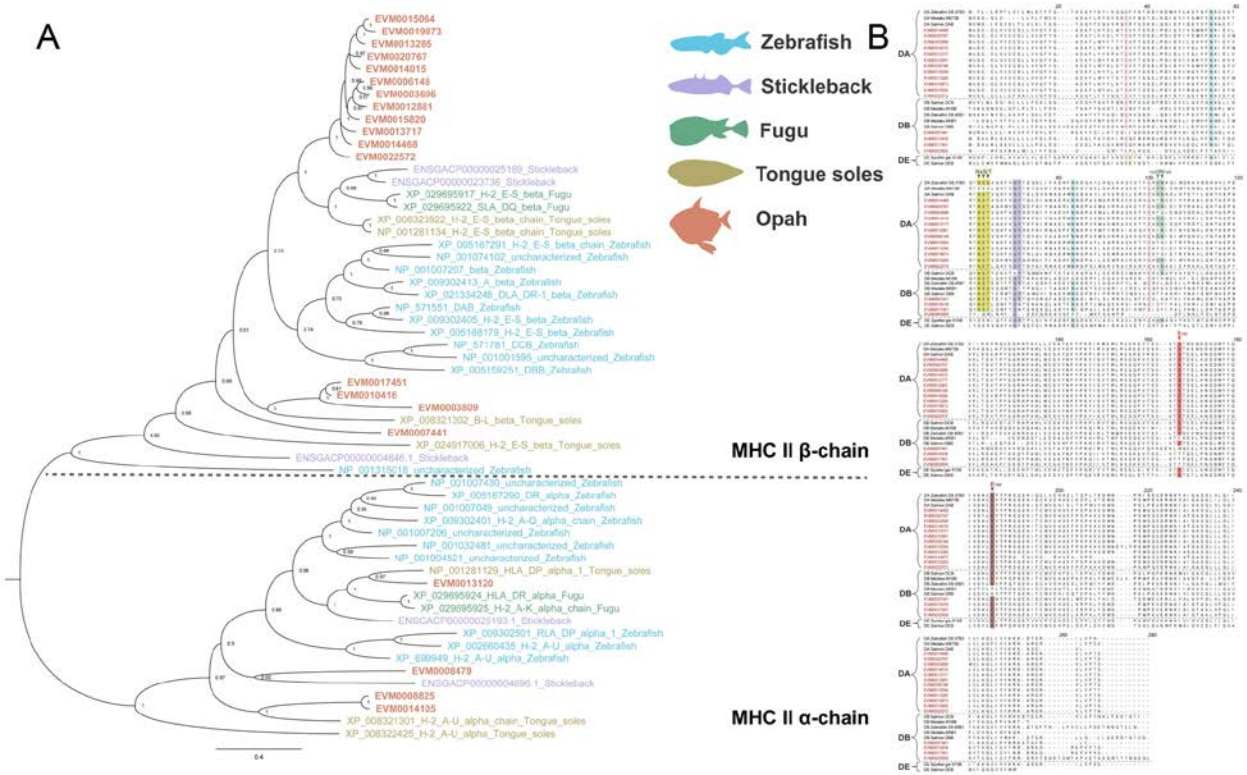


Figure S10. Phylogenetic analysis of the expanded MHCII genes. (A) Phylogenetic tree of MHCII genes of opah (orange), zebrafish (blue), stickleback (purple), fugu (green), and tongue sole (yellow). (B) Alignment of MHCII amino acid sequences.

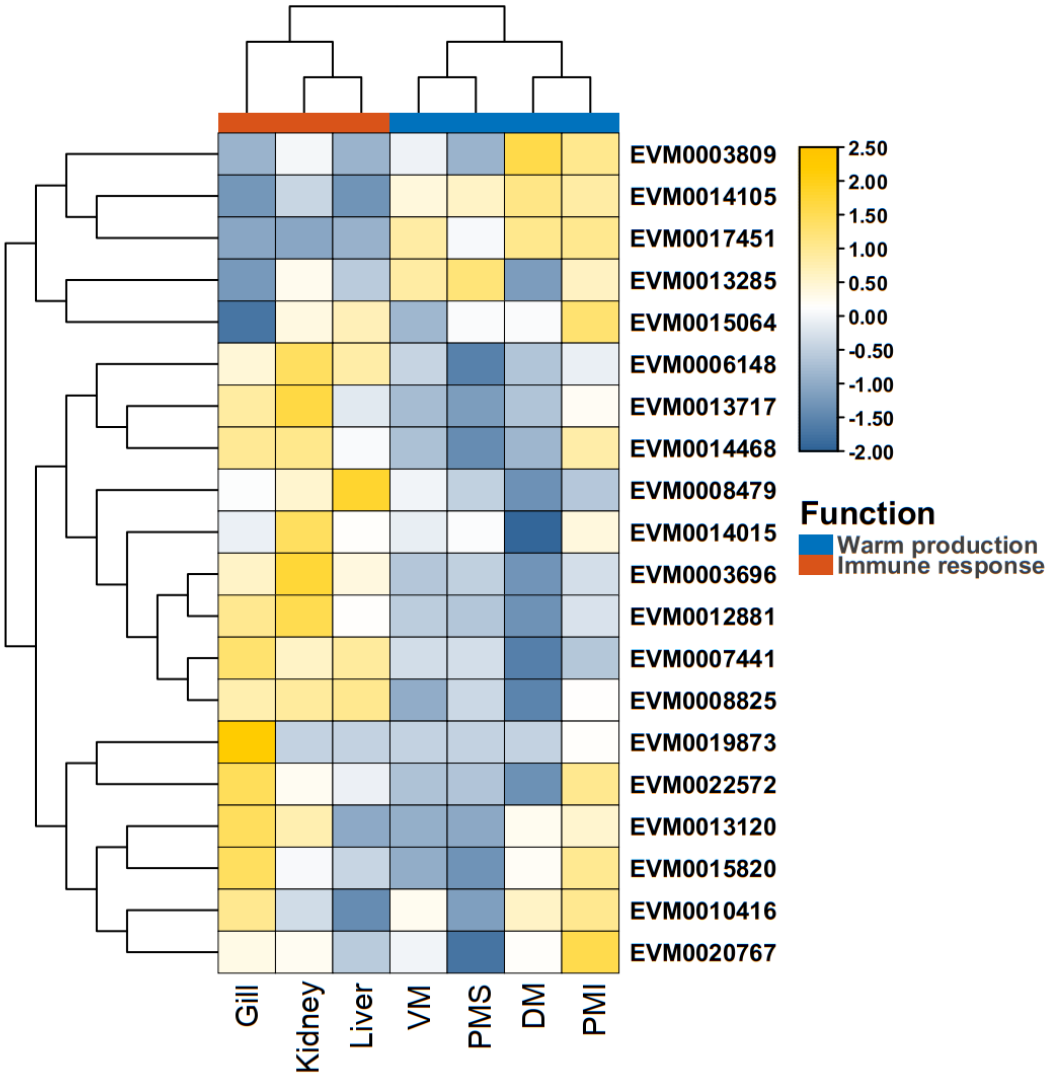


Figure S11. The heatmap of the expression of expanded MHC Ⅱ genes.

Table S1. Genome statistics of the *Lampris incognitus* genome.

N50 Reads length	15.00 Kb
Contig number	1584
N50 Contig length	3.71 Mb
Maximum contig length	21.58 Mb
Scaffold number	809
N50 scaffold length	61.11 Mb
Maximum scaffold length	126.23 Mb
Estimated genome size	1.37 Gb
Assembled genome size	1.37 Gb
GC content	45.20%
Repeat content	49.83%
Genome coverage	136 ×

Table S2. Quality control parameters for the Hi-C sequencing reads.

Parameters	**Details**
Number of Reads	467,444,773
Data Size (bp)	140,008,358,626
Q20 of fq1	97.45%
Q20 of fq2	95.61%
Q30 of fq1	93.42%
Q30 of fq2	89.47%
GC of fq1	46.04%
GC of fq2	46.22%

Table S3. Statistics of pseudochromosome length after Hi-C scaffolding.

Group	Cluster Num	Cluster Len (bp)	Order Num	Order Len (bp)
LG01	132	128,503,786	64	121,812,641
LG02	140	131,911,057	87	126,224,445
LG03	123	109,382,113	65	103,392,204
LG04	90	71,955,455	37	67,103,001
LG05	45	72,309,922	29	71,062,774
LG06	38	69,529,287	28	68,666,869
LG07	56	69,184,032	40	67,661,657
LG08	60	63,696,297	36	61,102,148
LG09	43	60,867,799	29	59,844,798
LG10	53	59,729,760	29	57,605,961
LG11	72	58,636,374	26	55,181,348
LG12	59	55,016,936	29	51,983,464
LG13	53	30,358,726	27	28,181,107
LG14	50	51,475,209	28	49,621,984
LG15	71	48,721,393	36	45,661,502
LG16	61	47,609,483	38	45,800,424
LG17	61	46,995,622	29	43,881,051
LG18	73	45,382,762	32	41,761,948
LG19	79	44,694,566	36	40,628,000
LG20	44	40,353,164	23	38,080,480
LG21	63	38,055,883	29	34,747,212
LG22	31	18,650,159	20	17,444,319
Total (Ratio %)	1497(94.51)	1363019785(99.67)	797(53.24)	1297449337(95.19)

Table S4. Statistics of predicted gene models in the *L. incognitus* genome.

Method	Software	Species	Gene number
Ab initio	Genscan	-	29,632
	Augustus	-	40,529
	GlimmerHMM	-	113,914
	GeneID	-	27,994
	SNAP	-	61,439
Homology-based	GeMoMa	*Danio rerio*	23,091
		Oryzias latipes	20,030
		Hippocampus comes	21,300
RNAseq	TransDecoder	-	98,419
	GeneMarkS-T	-	56,976
	PASA	-	27,110
Integration	EVM	-	24,658

Table S5. Statistics of the functional annotation.

Annotation database	Annotated number	Percentage (%)
GO annotation	12,137	49
KEGG annotation	14,538	59
KOG annotation	16,040	65
TrEMBL annotation	23,328	95
nr annotation	23,530	95
All annotated	23,628	96

Table S6. Statistics of the gene annotation.

Gene annotation	
Gene number	24,658
Gene length (bp)	468,925,661
Average gene length (bp)	19,017
Exon length (bp)	59,378,389
Average exon length (bp)	2,408.08
Exon number	230,852
Average exon number	9.36
Intron length (bp)	409,547,272
Average intron length (bp)	16,609.1
Intron number	20,6194
Average intron number	8

Table S7. Completeness of *L. incognitus* genome assembly assessed using BUSCO.

Complete BUSCOs	Complete and single-copy BUSCOs	Complete and duplicated BUSCOs	Fragmented BUSCOs	Missing BUSCOs	Total Lineage BUSCOs
2,433 (94.08%)	2,396 (92.65%)	37 (1.43%)	35 (1.35%)	118 (4.56%)	2,586

Note: BUSCO, Benchmarking Universal Single Copy Orthologs.

Table S8. Completeness of *L. incognitus* genome assembly assessed using CEGMA.

Number of 458 CEG present in assembly	%of 458 CEGs present in assemblies	Number of 248 highly conserved CEGs present	% of 248 highly conserved CEGs present
453	98.91%	243	97.98%

Note: CEGMA, Core Eukaryotic Genes Mapping Approach.

Table S9. Completeness of *L. incognitus* genome annotation assessed using BUSCO.

Complete BUSCOs	Complete and single-copy BUSCOs	Complete and duplicated BUSCOs	Fragmented BUSCOs	Missing BUSCOs	Total Lineage BUSCOs
2,477 (95.78%)	2,412 (93.27%)	65 (2.51%)	59 (2.28%)	50 (1.93%)	2,586

Note: BUSCO, Benchmarking Universal Single Copy Orthologs.

Table S10. Statistics of repeat and TEs ratios in *L. incognitus*.

Type	Number	Length (bp)	Rate (%)
ClassI	1,248,088	468,406,778	34.25
ClassI/DIRS	25,712	9,585,498	0.7
ClassI/LARD	316,316	74,610,824	5.46
ClassI/LINE	544,517	173,916,088	12.72
ClassI/LTR/Copia	18,599	9,728,995	0.71
ClassI/LTR/Gypsy	148,672	125,768,342	9.2
ClassI/LTR/Unknown	21,544	13,565,327	0.99
ClassI/PLE	159,738	52,729,802	3.86
ClassI/SINE	5,226	1,444,033	0.11
ClassI/TRIM	6,893	6,905,200	0.5
ClassI/Unknown	871	153,980	0.01
ClassII	626,035	192,736,991	14.09
ClassII/Crypton	762	184,648	0.01
ClassII/Helitron	8,660	2,474,294	0.18
ClassII/MITE	1,642	610,034	0.04
ClassII/Maverick	2,103	316,940	0.02
ClassII/TIR	589,633	183,944,265	13.45
ClassII/Unknown	23,235	5,206,810	0.38
Potential Host Gene	6,887	1,489,875	0.11
SSR	3,621	2,299,004	0.17
Unknown	93,426	16,536,115	1.21
Total	1,978,057	681,467,061	49.83

Table S11. Statistics of TE content in 9 teleost species.

Species	TE content (%)	LINE content (%)	LTR content (%)	SINE (%)	DNA transposons (%)	Unclassified (%)
Snailfish	36.38	9.58	4.61	0.69	14.26	6.98
Stickleback	13.91	3.66	2.66	0.55	4.20	2.84
Ocean sunfish	11.49	5.59	1.12	0.53	2.20	0.08
Takifugu	6.72	2.73	1.09	0.27	1.30	1.33
Tougue sole	5.85	1.04	0.08	0.22	2.45	2.06
Opah	49.83	12.72	10.90	0.11	14.09	1.21
Atlantic cod	14.26	1.40	1.23	0.43	4.69	6.51
Zebrafish	54.94	3.62	5.00	1.81	41.04	3.47
Spotted gar	19.77	5.41	2.55	2.73	3.54	5.54

Table S12. GO enrichment of LTRs (Gypsy) nearest coding genes in *L. incognitus* genome.
(Excel table)

Table S13. GO enrichment of 204 genes with convergent signature for positive selection in endothermic mammals, bires, and opah.
(Excel table)

Table S14. The convergent gene list in *Mus musculus*, *Homo sapiens*, *Gallus gallus*, *Meleagris gallopavo*, and *Lampris incognitus* genome.
(Excel table)

Table S15. Sequences of *slc8b1* and *glrx3* downloaded from public databases
(Excel table)

Table S16. The intersected positively selected genes analysed by using the opah, Pacific bluefin tuna, and great white sharks as foreground branches respectively.
(Excel table)

Table S17. GO enrichment of 112 genes with convergent signature for positive selection in three endothermic fishes.
(Excel table)

Table S18. Ka/Ks ratio of all orthologous genes of lamprey and 8 fishes used in convergent analysis.
(Excel table)

Table S19. The convergent gene list in *L. incognitus*, *Thunnus orientalis*, and *Carcharodon carcharias* genome.
(Excel table)

Table S20. The list of expanded gene families in *L. incognitus* genome.
(Excel table)

Table S21. The list of contracted gene families in *L. incognitus* genome.
(Excel table)

Table S22. The list of positively selected genes in *L. incognitus* genome.

(Excel table)

Table S23. GO enrichment of expanded genes in opah genome.

(Excel table)

Table S24. GO enrichment of positively selected genes in opah genome.

(Excel table)

Table S25. Transcriptome sequencing tissues and datasets of *L. incognitus*.

Sample ID	Tissue	Clean reads	Clean bases (Gb)	Q30 (%)
DM	dorsal musculature	51,682,984	7.75	93.03
VM	ventral musculature	61,351,546	9.2	92.36
PMS	superficial layer of the pectoral musculature	51,978,982	7.8	92.89
PMI	dark red aerobic pectoral musculature	55,717,750	8.36	92.95
LI	liver	48,311,722	7.25	93.51
KI	Kidney	54,761,218	8.21	92.98
GI	gill	61,733,880	9.26	93.14

Table S26. Top 100 highly expressed genes in the dark red aerobic pectoral musculature (PMI).

(Excel table)

Table S27. Top 100 highly expressed proteins in the dark red aerobic pectoral musculature (PMI).

(Excel table)

Table S28. Dark red aerobic pectoral musculature (PMI) -specifically expressed genes.

(Excel table)

Table S29. The differentially expressed genes (DEGs) of the dark red aerobic pectoral musculature (PMI) and other musculatures.

(Excel table)

Table S30. GO enrichment of the differentially expressed genes (DEGs) of the dark red aerobic pectoral musculature (PMI) and other musculatures.

(Excel table)

Table S31. Top 20 KEGG pathways of the up-regulated genes of the dark red aerobic pectoral musculature (PMI) and other musculatures.

(Excel table)

Table S32. The differentially expressed proteins (DEPs) of the dark red aerobic pectoral musculature (PMI) and the superficial layer of pectoral musculature (PMS).

(Excel table)

Table S33. The differentially expressed proteins (DEPs) of the dark red aerobic pectoral musculature (PMI) and the ventral musculature (VM).

(Excel table)