



A Chromosome-Level Genome Assembly of Yellowtail Kingfish (*Seriola lalandi*)

Shuo Li^{1,2,3}, Kaiqiang Liu^{1,2}, Aijun Cui^{1,2}, Xiancai Hao¹, Bin Wang^{1,2}, Hong-Yan Wang^{1,2}, Yan Jiang^{1,2}, Qian Wang^{1,2}, Bo Feng¹, Yongjiang Xu^{1,2*}, Changwei Shao^{1,2*} and Xuezhou Liu^{1,2}

¹Key Laboratory of Sustainable Development of Marine Fisheries, Ministry of Agriculture and Rural Affairs, Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Qingdao, China, ²Laboratory for Marine Fisheries Science and Food Production Processes, Qingdao National Laboratory for Marine Science and Technology, Qingdao, China, ³China State Key Laboratory for Managing Biotic and Chemical Threats to the Quality and Safety of Agroproducts, Ningbo University, Ningbo, China

OPEN ACCESS

Edited by:

Roger Huerlimann,
Okinawa Institute of Science and
Technology Graduate University,
Japan

Reviewed by:

Zhenhua Ma,
Chinese Academy of Fishery Sciences
(CAFS), China
Qiong Shi,
Beijing Genomics Institute (BGI), China

*Correspondence:

Yongjiang Xu
xuyj@ysfri.ac.cn
Changwei Shao
shaocw@ysfri.ac.cn

Specialty section:

This article was submitted to
Livestock Genomics,
a section of the journal
Frontiers in Genetics

Received: 30 November 2021

Accepted: 22 December 2021

Published: 19 January 2022

Citation:

Li S, Liu K, Cui A, Hao X, Wang B,
Wang H-Y, Jiang Y, Wang Q, Feng B,
Xu Y, Shao C and Liu X (2022) A
Chromosome-Level Genome
Assembly of Yellowtail Kingfish
(*Seriola lalandi*).
Front. Genet. 12:825742.
doi: 10.3389/fgene.2021.825742

Yellowtail kingfish (*Seriola lalandi*) is a pelagic marine piscivore with a circumglobal distribution. It is particularly suitable for open ocean aquaculture owing to its large body size, fast swimming, rapid growth, and high economic value. A high-precision genome is of great significance for future genetic breeding research and large-scale aquaculture in the open ocean. PacBio, Illumina, and Hi-C data were combined to assemble chromosome-level reference genome with the size of 648.34 Mb (contig N50: 28.52 Mb). 175 contigs was anchored onto 24 chromosomes with lengths ranging from 12.28 to 34.59 Mb, and 99.79% of the whole genome sequence was covered. The BUSCOs of genome and gene were 94.20 and 95.70%, respectively. Gene families associated with adaptive behaviors, such as olfactory receptors and HSP70 gene families, expanded in the genome of *S. lalandi*. An analysis of selection pressure revealed 652 fast-evolving genes, among which *mkxb*, *popdc2*, *dlx6*, and *ifitm5* may be related to rapid growth traits. The data generated in this study provide a valuable resource for understanding the genetic basis of *S. lalandi* traits.

Keywords: *Seriola lalandi*, genome, adaptation, rapid growth, aquaculture

INTRODUCTION

To develop environmentally friendly and economically sustainable aquaculture, it is necessary to understand the genetic basis of traits that currently limit/enhance development of domestic aquaculture (Rondeau et al., 2013). Genetic resources have been developed and widely used in agriculture and animal husbandry for decades, but only recently have they been used in selected aquaculture species (Ozaki et al., 2013; Dunham et al., 2014). There is still limited information on genetic variation on commercially important traits (Peterson et al., 2020). The methods used to develop these resources offer the best possibilities for genetic improvement or culture practices

Abbreviations: N50, median size; Gb, gigabase pairs; Mb, megabase pairs; kb, kilobase pairs; PacBio, Pacific Biosciences; SMRT, single molecule, real-time; BUSCOs, Benchmarking Universal Single-Copy Orthologs; GC, guanine-cytosine; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; Mb, megabase pairs; ML, maximum likelihood; MYA, million years ago; NCBI, national center for biotechnology information; HSP70, heat shock protein 70 family.

(Sodeland et al., 2013). Third-Generation Sequencing (TGS) has improved this area of research through high quality assemblies and decreasing costs, and this has enabled development of genetic resources for a greater number of species (Huete-Pérez and Quezada, 2013; Lee et al., 2016; Lv et al., 2020).

Yellowtail kingfish (*Seriola lalandi*) is an excellent marine economic fish. It has a number of beneficial traits for open ocean aquaculture systems, including large body size, rapid growth, and high-quality flesh (Orellana et al., 2014; Sanchís-Benlloch et al., 2017). Similar in taste to tuna or mackerel, yellowtail kingfish have a large market worldwide and are a popular fish used in sushi (Purcell et al., 2015). These make them a good candidate for aquaculture. Since the 1990s, extensive research in Japan has focused on artificial breeding and breeding technology for *S. lalandi* (Sano, 1998). In China, aquaculture of *S. lalandi* began in 2001 (Jiang et al., 2001), along with biological research, including studies of embryogenesis, seedling cultivation, and effects of salinity stress on growth (Shi et al., 2019; Xu et al., 2019; Liu A et al., 2021).

Here, we report a chromosomal-level genome assembly of *S. lalandi*. Our evolutionary and comparative genomic analysis provide insights into the adaptability of the species to the external environment. Furthermore, the genome analysis provide a valuable resource for further studies of the genetic basis of traits of *S. lalandi*.

Value of the Data

This is the first chromosomal-level genome assembly in *Seriola* genus. It could be a valuable resource to conduct a comparative analysis among the species in the genome of the *Seriola* genus and for further studies of the genetic basis of traits of *S. lalandi*.

MATERIALS AND METHODS

Sampling and Sequencing

Yellowtail kingfish specimens were collected from Dalian Futai Marine Products Farming Co., Ltd. (Dalian, China). Total genomic DNA of a male fish muscle sample was extracted using the QIAamp DNA Mini Kit (QIAGEN, Hilden, Germany) following the manufacturer's protocols. We constructed two paired-end libraries (insert sizes of 200 and 500 bp) following the manufacturer's protocol (Chromium Genome v1, PN120229). The libraries were sequenced on the BGISEQ-500 platform to obtain PE 2 × 150 bp reads. The extracted DNAs were also used to construct a 20 kb library following the PacBio protocol (Pacific Biosciences, Menlo Park, CA, United States). The libraries were then sequenced on the PacBio Sequel platform. We obtained 48.74 and 106.76 Gb of raw sequence data using the BGISEQ-500 and PacBio platforms, respectively (Supplementary Table 1).

To construct chromosome-level assemblies, the Hi-C technique was used. A Hi-C library was prepared following the strategy described by Rao et al. (Rao et al., 2014) using blood samples with an ~300 bp insert size. Using the

BGISEQ-500 platform to sequence the Hi-C library, we obtained 87.60 Gb of raw Hi-C data (Supplementary Table 1).

Four tissues (brain, pituitary, liver, and muscle) were collected for RNA sequencing. RNA from each tissue was extracted and treated with DNase I (TAKARA, Kusatsu, Japan) to remove genomic DNA. For each tissue, a paired-end RNA-sequencing library was constructed with an insert size of 300 bp and then sequenced on the Illumina HiSeq 2,500 platform to generate PE 2 × 150 bp. One muscle specimen was also used to construct an Iso-Seq library and then sequenced on the PacBio Sequel platform. In total, we obtained 307.14 and 26.89 Gb of raw sequence data using the Illumina HiSeq 2,500 and PacBio platforms, respectively (Supplementary Table 1).

Genome Assembly, Chromosome Anchoring, and Genome Annotation

Before genome assembly, we estimated the genome by a k-mer analysis using Jellyfish v2.2.6 (Marçais and Kingsford, 2011). For this, a series of k-mers (17, 19, and 21) were extracted from the BGISEQ-500 sequencing data and the frequency of each kmer was calculated. The heterozygosity rate was estimated using 17-mers using GenomeScope v2.0.0 (Supplementary Figure 1). Considering the C-value (0.7) from the Animal Genome Size Database, the estimated genome size of *S. lalandi* was 684.60 Mb.

Canu v1.8 was used for the self-correction of long reads sequenced with the PacBio Sequel platform. Then, the corrected reads were assembled using wtdbg2 v2.5 (options: -x rs -g 750 m) (Ruan and Li, 2020). Pilon v1.23 (Walker et al., 2014) was used to polish contigs with short reads by three rounds of alignment. The Hi-C short reads were aligned to the scaffolds using Juicer (Durand et al., 2016) and anchoring was performed using 3D-DNA v180419 (Dudchenko et al., 2017). We finally used Juicebox Assembly Tools v1.9.9 (Durand et al., 2016) to correct the connections. The completeness of the final assembly was assessed using BUSCO v4.0 (Simão et al., 2015).

Both homology-based and *de novo* predictions were used to annotate repetitive sequences. Transposable elements were identified using RepeatMasker v4.0.7 (<http://www.repeatmasker.org>) and RepeatProteinMask v1.36 with Repbase v17.01 (Bao et al., 2015). A *de novo* transposable element library was constructed using RepeatModeler v1.0.11 (<http://www.repeatmasker.org/RepeatModeler.html>) and was then used to predict repeats using RepeatMasker.

To annotate gene structures, we used homology-based prediction, transcriptome-based prediction, and *de novo* prediction. For homology-based annotation, the protein sequences of eight teleost species downloaded from NCBI, including *Seriola lalandi dorsalis*, *Seriola dumerili*, *Seriola quinqueradiata*, *Seriola rivoliana*, *Echeneis naucrates*, *Oryzias latipes*, *Danio rerio*, and *Takifugu rubripes*, were aligned to the genome assembly by BLAT v3.6 (Kent, 2002), and then GENEWISE v2.4.0 (Birney et al., 2004) was used to predict gene structures. For next-generation

RNA-sequencing annotation, data were aligned to the genome assembly using HISAT2 v2.1.0 (Kim et al., 2015) and the alignments were fed to StringTie v1.3.5 (Pertea et al., 2015) to assemble the transcriptome. TransDecoder v5.0.2 (<https://github.com/TransDecoder/TransDecoder/>) was used to predicate ORFs and identify candidate gene structures. For third-generation RNA-sequencing annotation, long-read RNA-seq (PacBio Iso-Seq) transcripts were obtained by removing the redundant sequences using cd-hit-est v4.8.1 (Li and Godzik, 2006). Then, the non-redundant transcripts were mapped to the genome by BLAT and assembled using PASA v2.0.2 (<https://github.com/PASAPipeline/PASAPipeline/>). For *de novo* prediction, the gene structures were analyzed on the repeat-masked genome assembly using AUGUSTUS v2.5.5 (Stanke et al., 2006), GlimmerHMM v3.0.4 (Allen et al., 2006), and GENSCAN (Burge and Karlin, 1998). Finally, genes predicted from the above methods were merged to obtain a consensus gene set using Evidence Modeler (EVM). For the functional annotation of the gene sets, the protein sequences of these genes were aligned against sequences in public protein databases, including, NR, KEGG, SwissProt, GO, InterPro, and TrEMBL, to identify homologues using Blastp v2.2.26 with an E-value cutoff of $1e-5$.

Phylogenetic Analysis and Gene Family Expansion

To determine single-copy genes of *S. lalandi* and other species (*S. dumerili*, *S. quinqueradiata*, *S. rivoliana*, *E. naucrastes*, *O. latipes*, *D. rerio*, *T. rubripes*, *Larimichthys crocea*, *Oreochromis niloticus*, and *Caranx melampygus*), the TreeFam pipeline (Li et al., 2006) was used. Before generating the alignment, the longest transcript of each gene was selected and protein sequences shorter than 50 amino acids were filtered out. Then, Blastp searches were performed for all protein sequences with an E-value cut-off of $1e-5$, and fragmented alignments were merged using SOLAR. Hcluster was used to filter segments, group genes, and determine single-copy orthologue families. The phylogenetic tree was inferred using multiple alignments from the single-copy genes using RaxML-ng v0.9.0 (Kozlov et al., 2019) under the site-heterogeneous GTR + G4 model with maximum likelihood estimation (ML).

An ultrametric tree was inferred using r8s v1.71 with fossil records from the TimeTree website (<http://www.timetree.org>) for calibration. An MCMCTREE analysis implemented in PAML v4.5 (Yang, 1997) was employed to estimate divergence times. CAFÉ v5.0 (De Bie et al., 2006) was used to assess gene family size dynamics, and families with $p < 0.05$ showed significant expansion or contraction. GO and KEGG pathway enrichment analysis were used to analyze the expanded and contracted genes.

Positive Selection Analysis

To identify positively selected genes (PSGs), we re-determined single-copy orthologues shared among five species (*E. naucrastes*, *T. rubripes*, *O. latipes*, *D. rerio*, and *S. lalandi*) and constructed a phylogenetic tree. Based on the

new phylogenetic tree and single-copy genes, we estimated the rate ratio (ω) of non-synonymous to synonymous nucleotide substitutions using CodeML (PAML package) to examine selective constraint. After obtaining high-quality alignments using prank v.100802 (Löytynoja and Goldman, 2010), Gblocks v0.91b (Castresana, 2000) was used to eliminate poorly aligned positions and divergent regions. Finally, the signature of positive selection ($d_N/d_S > 1$) was identified using the PAML branch site model. GO and KEGG pathway enrichment analysis were used to evaluate PSGs.

RESULTS AND DISCUSSION

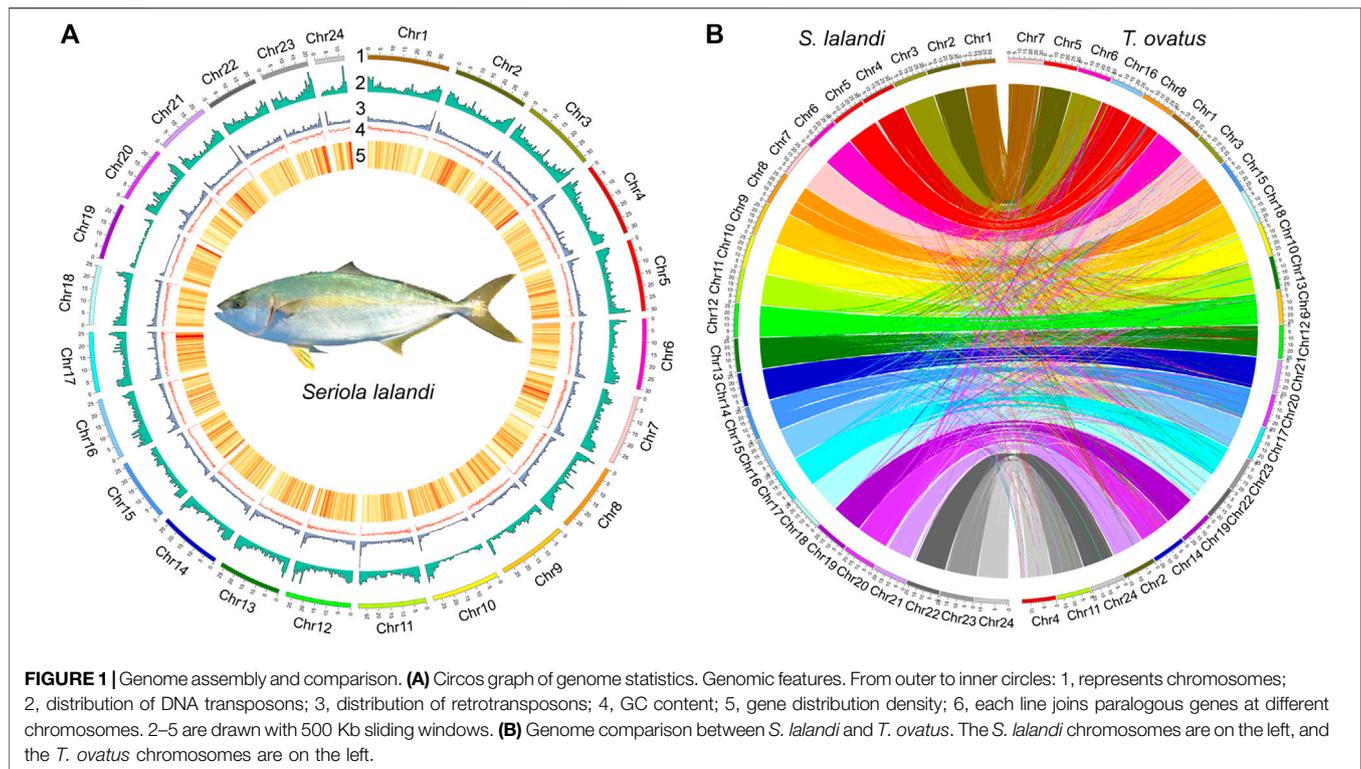
Genome Features

To generate a high-quality reference genome, we combined PacBio, Illumina, and Hi-C data (**Supplementary Table 1**). PacBio CLRs with coverage of $165 \times$ were used for genome assembly. The draft assembly was 648.34 Mb, with 277 contigs, a contig N50 of 28.52 Mb, and a GC content of 40.79% (**Supplementary Table 2**). Using ~ 52 Gb ($\sim 87 \times$) of valid Hi-C data, we anchored 175 contigs onto 24 chromosomes (**Figure 1A**, **Supplementary Figure 2**) (Shi et al., 2017). The lengths of the 24 chromosomes ranged from 12.28 to 34.59 Mb, and 99.79% of the whole genome sequence was covered (**Supplementary Table 3**). To evaluate the completeness of the assembly, the BUSCO database (actinopterygii_odb10) and RNA-seq data were used. The genome contained 94.20% complete BUSCOs and the average mapping rate of transcriptome data was 96.30% (**Supplementary Table 4**). The published *Trachinotus ovatus* chromosome-level genome was used to validate the accuracy of the assembly of the chromosomes (Zhang et al., 2019); 567.01 MB synteny blocks (each synteny block > 500 bp) were consistent with the assembled chromosomes (**Figure 1B**).

Repetitive elements comprised 22.46% of the *S. lalandi* genome, similar to the estimate in the *T. ovatus* genome (20.25%, 655 Mb) (Zhang et al., 2019). The most abundant transposable elements (TEs) were DNA transposons (11.51%), followed by long terminal repeats (LTRs, 4.93%) and long interspersed elements (LINEs, 3.85%) (**Supplementary Figure 4**). We integrated *de novo*, homology-based and transcriptome-based methods to predict a protein-coding gene set comprising 22,674 genes (**Supplementary Table 5**), and which 20,568 (90.71%) matched entries in a public database (**Supplementary Table 6**). We identified 95.70% complete BUSCOs from 22,674 protein-coding genes.

Phylogenetic Relationships and Genomic Comparison

We constructed a phylogenetic tree of *S. lalandi* and 10 teleost fish (*S. dumerili*, *S. quinqueradiata*, *S. rivoliana*, *E. naucrastes*, *O. latipes*, *D. rerio*, *T. rubripes*, *L. crocea*, *O. niloticus*, and *C. melampygus*) based on 5,067 single-copy genes (**Figure 2A**, **Supplementary Table 7**). According to the phylogeny and the



fossil record of teleosts, we dated the divergence of *Seriola* from the other teleost species to approximately 72.6 million years ago (Figure 2A).

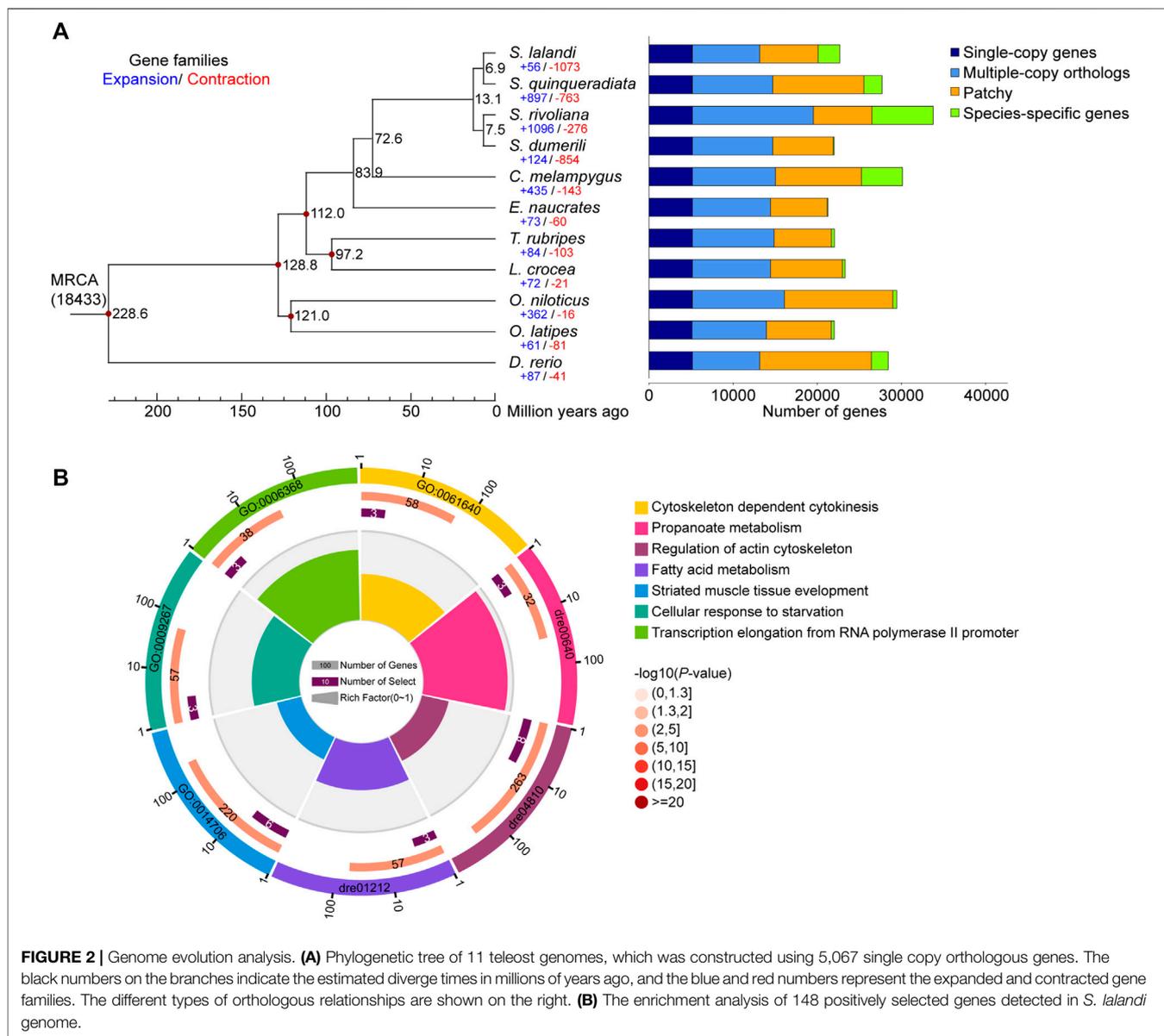
We detected 56 significantly expanded and 1,073 significantly contracted gene families ($p < 0.05$) in *S. lalandi* (Figure 2A). Compared with teleost fish except of *Seriola*, the HSP70 family with 19 HSP70 genes was expanded (Supplementary Table 8). We found five *hspa12* genes, including *hspa12a*, *hspa12b*, *hspa12l-1*, *hspa12l-2*, and *hspa12l-3*, which was more than observed in *E. naucrates* (2), *O. latipes* (2), *D. rerio* (2), *T. rubripes* (2), *C. melampygyus* (3), and *L. crocea* (3) (Liu et al., 2019). In *S. lalandi*, there were three *hspa12l* gene copies. HSP70 is a well-known stress protein (Clark and Peck, 2009), and the expansion of the HSP70 family in *S. lalandi* may contribute to its adaptation to changes in the aquatic environment.

Yellowtail kingfish is a migratory marine fish with high olfactory sensitivity (Martínez-Montañó et al., 2016). We identified 147 olfactory receptor (OR)-like genes from the *S. lalandi* genome, including subfamily "Delta" (68), "Eta" (49), "Zeta" (12), "Epsilon" (9), "Beta" (6), "Thet" (2), and "Kappa" (1) (Supplementary Table 9). The expanded subfamilies "Delta" and "Epsilon" are important for the perception of water-soluble odorants (Cong et al., 2019). Most teleosts possess one or two "Beta" OR genes, which are important for detecting both water-soluble and airborne odorants (Liu H et al., 2021). However, subfamily "Beta" of olfactory receptor was expanded in *S. lalandi*. These expansions may contribute to the olfactory detection ability of the species, which could be useful for feeding and migration (Bett and Hinch, 2016).

Fast-Evolving Genes in Yellowtail Kingfish

PSGs are often associated with adaptive evolution and may contribute to new or improved functions. To understand the selective pressure operating on *S. lalandi*, we compared the orthologues of five teleost species (*E. naucrates*, *T. rubripes*, *O. latipes*, *D. rerio*, and *S. lalandi*) and identified 652 fast-evolving genes, including 148 PSGs ($d_N/d_S > 1$) and 504 genes that contain positively selected sites in *S. lalandi* (Supplementary Table 10). Consistent with the large body size and fast swimming ability, an enrichment analysis revealed that the PSGs were involved in striated muscle tissue development (GO:0014706), regulation of actin cytoskeleton (dre04810), and fatty acid metabolism (dre01212) (Figure 2B).

Muscle tissue development is associated with the growth rate, which is a major economic trait in animal production. Several genes involved in muscle tissue development (*klf2a*, *klhl41b*, *cdk9*, *ndrg4*, *mkxb*, and *popdc2*) showed rapid evolution in *S. lalandi* and likely contribute to the rapid growth of the species (Supplementary Table 10). Fast growing muscles also require increased bone support. Two genes, *dlx6* and *ifitm5*, were involved in skeletal system development and promote bone formation to support the large body (Supplementary Table 10). Based on the strong muscle and skeletal systems, muscle contraction-related genes (*arhgef12b*, *ramp2*, *tnnt2a*, *tnn1a*, *cald1a*, and *tnnt2a*) with positively selected sites may provide support for fast swimming (Supplementary Table 10). Furthermore, fatty acid metabolism-related fast-evolving genes (*hsd17b12b*, *acadm*, *mecr*, *lipg*, and *hao1*) also contributed to energy consumption and growth (Supplementary Table 10).



Moreover, some genes with positively selected sites were associated with growth (*ficn*, *tgfb3*, *igf1ra*, *gpc1b*, *rnf11*, and *tgfbap1*) (Supplementary Table 10).

We identified other fast-evolving genes, such as the pheromone response gene *ora2*, ear development gene *ddt*, and sensory perception gene *pppf1*, with potential roles in the perception of the external environment (Supplementary Table 10). Fast-evolving genes involved in nervous system development and the regulation of neurotransmitter secretion (*rab33a*, *fstl5*, and *cplx2*) provided a tissue basis for sensitive sensory systems (Supplementary Table 10). The fast-evolving genes *ldha* and *rnf152* may contribute to adaptation to adverse environmental conditions, such as hypoxia and starvation (Supplementary Table 10).

CONCLUSIONS

We sequenced and assembled the genome of *S. lalandi* using Illumina shotgun, PacBio SMRT, and Hi-C data, providing the first chromosome-level genome assembly for the genus *Seriola*. Basing on multiple annotation strategies, we obtained 22,674 protein-coding genes with minimal redundancy. Further genomic analysis revealed gene families associated with expansions of HSP70 and olfactory receptor gene families, and the rapid evolution of muscle and skeletal system development genes, providing insight into the genetic basis underlying the physiological characteristics of *S. lalandi* and its adaptability to the external environment. We believe these

new resources will promote genetic research and accelerate the genetic breeding process for *S. lalandi*.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: GenBank, JAIQDC010000001.1-JAIQDC010000031.1; NCBI BioProject, PRJNA754209.

ETHICS STATEMENT

The animal study was reviewed and approved by Experimental Animal Care, Ethics and Safety Inspection Form Yellow Sea Fisheries Research Institute, CAFS. Written informed consent was obtained from the owners for the participation of their animals in this study.

AUTHOR CONTRIBUTIONS

LX, CS, and XY conceived and designed the experiments. BW, AC, and YJ collected, identified, and photographed the

specimens. SL, KL, and XH analyzed the genome and transcriptome data. WH, WQ, and BF performed gene analysis. SL drafted the manuscript. CS, YX, and ZL provided advice on manuscript writing. All authors reviewed the manuscript.

FUNDING

This work was supported the National Key R&D Program of China (2018YFD0900301, 2019YFD0900901, 2018YFD0901204); the Marine S&T Fund of Shandong Province for Pilot National Laboratory for Marine Science and Technology (Qingdao) (2018SDKJ0303-1); the Central Public-interest Scientific Institution Basal Research Fund, CAFS (No.2020TD19, 2020TD47, 2021GH05); the Taishan Scholar Project Fund of Shandong of China; the China Agriculture Research System of MOF and MARA (CARS-47).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2021.825742/full#supplementary-material>

REFERENCES

- Allen, J. E., Majoros, W. H., Pertea, M., and Salzberg, S. L. (2006). JIGSAW, GeneZilla, and GlimmerHMM: Puzzling Out the Features of Human Genes in the ENCODE Regions. *Genome Biol.* 7 Suppl 1, S9–S13. doi:10.1186/gb-2006-7-s1-s9
- Bao, W., Kojima, K. K., and Kohany, O. (2015). Repbase Update, a Database of Repetitive Elements in Eukaryotic Genomes. *Mob DNA* 6, 11–16. doi:10.1186/s13100-015-0041-9
- Bett, N. N., and Hinch, S. G. (2016). Olfactory Navigation during Spawning Migrations: a Review and Introduction of the Hierarchical Navigation Hypothesis. *Biol. Rev.* 91, 728–759. doi:10.1111/brv.12191
- Birney, E., Clamp, M., and Durbin, R. (2004). GeneWise and Genomewise. *Genome Res.* 14, 988–995. doi:10.1101/gr.1865504
- Burge, C. B., and Karlin, S. (1998). Finding the Genes in Genomic DNA. *Curr. Opin. Struct. Biol.* 8, 346–354. doi:10.1016/s0959-440x(98)80069-9
- Castresana, J. (2000). Selection of Conserved Blocks from Multiple Alignments for Their Use in Phylogenetic Analysis. *Mol. Biol. Evol.* 17, 540–552. doi:10.1093/oxfordjournals.molbev.a026334
- Clark, M. S., and Peck, L. S. (2009). Triggers of the HSP70 Stress Response: Environmental Responses and Laboratory Manipulation in an Antarctic marine Invertebrate (*Nacella concinna*). *Cell Stress and Chaperones* 14, 649–660. doi:10.1007/s12192-009-0117-x
- Cong, X., Zheng, Q., Ren, W., Chéron, J.-B., Fiorucci, S., Wen, T., et al. (2019). Zebrafish Olfactory Receptors ORAs Differentially Detect Bile Acids and Bile Salts. *J. Biol. Chem.* 294, 6762–6771. doi:10.1074/jbc.ra118.006483
- De Bie, T., 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. doi:10.1093/bioinformatics/btl097
- Dudchenko, O., Batra, S. S., Omer, A. D., Nyquist, S. K., Hoeger, M., Durand, N. C., et al. (2017). *De Novo* assembly of the *Aedes aegypti* Genome Using Hi-C Yields Chromosome-Length Scaffolds. *Science* 356, 92–95. doi:10.1126/science.aal3327
- Dunham, R. A., Taylor, J. F., Rise, M. L., and Liu, Z. (2014). Development of Strategies for Integrated Breeding, Genetics and Applied Genomics for Genetic
- Improvement of Aquatic Organisms. *Aquaculture* 420–421, S121–S123. doi:10.1016/j.aquaculture.2013.10.020
- Durand, N. C., Shamim, M. S., Machol, I., Rao, S. S. P., Huntley, M. H., Lander, E. S., et al. (2016). Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. *Cel Syst.* 3, 95–98. doi:10.1016/j.cels.2016.07.002
- Huete-Pérez, J. A., and Quezada, F. (2013). Genomic Approaches in marine Biodiversity and Aquaculture. *Biol. Res.* 46, 353–361. doi:10.4067/S0716-97602013000400007
- Jiang, D., Lin, L., and Chen, Y. (2001). Indoor wintering and growth of *Seriola aureorvtata* Temminck et Schegel. *J. Dalian Fish. Univ.* 3, 69–73. doi:10.3969/j.issn.1000-9957.2001.03.012
- Kent, W. J. (2002). BLAT-the BLAST-like Alignment Tool. *Genome Res.* 12, 656–664. doi:10.1101/gr.229202
- Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a Fast Spliced Aligner with Low Memory Requirements. *Nat. Methods* 12, 357–360. doi:10.1038/nmeth.3317
- Kozlov, A. M., Darriba, D., Flouri, T., Morel, B., and Stamatakis, A. (2019). RAXML-NG: a Fast, Scalable and User-Friendly Tool for Maximum Likelihood Phylogenetic Inference. *Bioinformatics* 35, 4453–4455. doi:10.1093/bioinformatics/btz305
- Lee, H., Gurtowski, J., Yoo, S., Nattestad, M., Marcus, S., Goodwin, S., et al. (2016). Third-generation Sequencing and the Future of Genomics. *BioRxiv*, 048603. doi:10.1101/048603
- Li, H., Coghlan, A., Ruan, J., Coin, L. J., Hériché, J.-K., Osmotherly, L., et al. (2006). TreeFam: a Curated Database of Phylogenetic Trees of Animal Gene Families. *Nucleic Acids Res.* 34, D572–D580. doi:10.1093/nar/gkj118
- Li, W., and Godzik, A. (2006). Cd-hit: a Fast Program for Clustering and Comparing Large Sets of Protein or Nucleotide Sequences. *Bioinformatics* 22, 1658–1659. doi:10.1093/bioinformatics/btl158
- Liu, A., Pirozzi, I., Codabaccus, B. M., Stephens, F., Francis, D. S., Sammut, J., et al. (2021). Effects of Dietary Choline on Liver Lipid Composition, Liver Histology and Plasma Biochemistry of Juvenile Yellowtail Kingfish (*Seriola lalandi*). *Br. J. Nutr.* 125, 1344–1358. doi:10.1017/S0007114520003669
- Liu, H., Chen, C., Lv, M., Liu, N., Hu, Y., Zhang, H., et al. (2021). A Chromosome-Level Assembly of Blunt Snout Bream *Megalobrama amblycephala* Genome

- Reveals an Expansion of Olfactory Receptor Genes in Freshwater Fish. *Mol. Biol. Evol.* 38, 4238–4251. doi:10.1093/molbev/msab152
- Liu, K., Hao, X., Wang, Q., Hou, J., Lai, X., Dong, Z., et al. (2019). Genome-wide Identification and Characterization of Heat Shock Protein Family 70 Provides Insight into its Divergent Functions on Immune Response and Development of *Paralichthys Olivaceus*. *PeerJ* 7, e7781. doi:10.7717/peerj.7781
- Löytynoja, A., and Goldman, N. (2010). webPRANK: a Phylogeny-Aware Multiple Sequence Aligner with Interactive Alignment Browser. *BMC Bioinform* 11, 1–7. doi:10.1186/1471-2105-11-579
- Lv, M., Zhang, Y., Liu, K., Li, C., and Wang, J. (2020). A Chromosome-Level Genome Assembly of the Anglerfish *Lophius Litulon*. *Front. Genet.* 11. doi:10.3389/fgene.2020.581161
- Martínez-Montaño, E., González-Álvarez, K., Lazo, J. P., Audelo-Naranjo, J. M., and Vélez-Medel, A. (2016). Morphological Development and Allometric Growth of Yellowtail Kingfish *Seriola lalandi* V. Larvae under Culture Conditions. *Aquac. Res.* 47, 1277–1287. doi:10.1111/are.12587
- Marçais, G., and Kingsford, C. (2011). A Fast, Lock-free Approach for Efficient Parallel Counting of Occurrences of K-Mers. *Bioinformatics* 27, 764–770. doi:10.1093/bioinformatics/btr011
- Orellana, J., Waller, U., and Wecker, B. (2014). Culture of Yellowtail Kingfish (*Seriola lalandi*) in a marine Recirculating Aquaculture System (RAS) with Artificial Seawater. *Aquacultural Eng.* 58, 20–28. doi:10.1016/j.aquaeng.2013.09.004
- Ozaki, A., Yoshida, K., Fuji, K., Kubota, S., Kai, W., Aoki, J.-y., et al. (2013). Quantitative Trait Loci (QTL) Associated with Resistance to a Monogenean Parasite (*Benedenia Seriolae*) in Yellowtail (*Seriola quinqueradiata*) through Genome Wide Analysis. *PLOS ONE* 8, e64987. doi:10.1371/journal.pone.0064987
- Perteau, M., Perteau, G. M., Antonescu, C. M., Chang, T.-C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie Enables Improved Reconstruction of a Transcriptome from RNA-Seq Reads. *Nat. Biotechnol.* 33, 290–295. doi:10.1038/nbt.3122
- Peterson, B. C., Burr, G. S., Pietrak, M. R., and Proestou, D. A. (2020). Genetic Improvement of North American Atlantic Salmon and the Eastern Oyster *Crassostrea virginica* at the U.S. Department of Agriculture-Agricultural Research Service National Cold Water Marine Aquaculture Center. *North. Am. J. Aquac.* 82, 321–330. doi:10.1002/naaq.10144
- Purcell, C. M., Chabot, C. L., Craig, M. T., Martinez-Takeshita, N., Allen, L. G., and Hyde, J. R. (2015). Developing a Genetic Baseline for the Yellowtail Amberjack Species Complex, *Seriola lalandi* Sensu Lato, to Assess and Preserve Variation in Wild Populations of These Globally Important Aquaculture Species. *Conserv Genet.* 16, 1475–1488. doi:10.1007/s10592-015-0755-8
- Rao, S. S. P., Huntley, M. H., Durand, N. C., Stamenova, E. K., Bochkov, I. D., Robinson, J. T., et al. (2014). A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell* 159, 1665–1680. doi:10.1016/j.cell.2014.11.021
- Rondeau, E. B., Messmer, A. M., Sanderson, D. S., Jantzen, S. G., von Schalburg, K. R., Minkley, D. R., et al. (2013). Genomics of Sablefish (*Anoplopoma fimbria*): Expressed Genes, Mitochondrial Phylogeny, Linkage Map and Identification of a Putative Sex Gene. *BMC Genomics* 14, 1–9. doi:10.1186/1471-2164-14-452
- Ruan, J., and Li, H. (2020). Fast and Accurate Long-Read Assembly With Wtdbg2. *Nat. Methods* 17, 155–158. doi:10.1038/s41592-019-0669-3
- Sanchis-Benlloch, P. J., Nocillado, J., Ladisa, C., Aizen, J., Miller, A., Shpilman, M., et al. (2017). *In-vitro* and *In-Vivo* Biological Activity of Recombinant Yellowtail Kingfish (*Seriola lalandi*) Follicle Stimulating Hormone. *Gen. Comp. Endocrinol.* 241, 41–49. doi:10.1016/j.ygcen.2016.03.001
- Sano, T. (1998). Control of Fish Disease, and the Use of Drugs and Vaccines in Japan. *J. Appl. Ichthyol.* 14, 131–137. doi:10.1111/j.1439-0426.1998.tb00630.x
- Shi, B., Liu, X., Liu, Y., Zhang, Y., Gao, Q., Xu, Y., et al. (2019). Effects of Gradual Salinity Change on Osmotic Regulation of Juvenile Yellowtail Kingfish (*Seriola Aureovittata*). *Coast Eng.* 38, 63–70.
- Shi, B., Liu, Y., Liu, X., Xu, Y., Li, R., Song, X., et al. (2017). Study on the Karyotype of Yellowtail Kingfish (*Seriola Aureovittata*). *PROGRESS FISHERY SCIENCES* 38, 136–141. doi:10.11758/yykxjz.20160816004
- Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. (2015). BUSCO: Assessing Genome Assembly and Annotation Completeness with Single-Copy Orthologs. *Bioinformatics* 31, 3210–3212. doi:10.1093/bioinformatics/btv351
- Sodeland, M., Gaarder, M., Moen, T., Thomassen, M., Kjøglum, S., Kent, M., et al. (2013). Genome-wide Association Testing Reveals Quantitative Trait Loci for Fillet Texture and Fat Content in Atlantic salmon. *Aquaculture* 408–409, 169–174. doi:10.1016/j.aquaculture.2013.05.029
- Stanke, M., Keller, O., Gunduz, I., Hayes, A., Waack, S., and Morgenstern, B. (2006). AUGUSTUS: Ab Initio Prediction of Alternative Transcripts. *Nucleic Acids Res.* 34, W435–W439. doi:10.1093/nar/gkl200
- Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., et al. (2014). Pilon: An Integrated Tool for Comprehensive Microbial Variant Detection and Genome Assembly Improvement. *PLOS ONE* 9, e112963. doi:10.1371/journal.pone.0112963
- Xu, Y., Zhang, Z., Liu, X., Wang, B., Shi, B., Liu, Y., et al. (2019). Morphometric Characteristics of the Embryonic and Postembryonic Development of Yellowtail Kingfish, *Seriola Aureovittata*. *J. Fish. Sci. China* 26, 172. doi:10.3724/sp.j.1118.2019.18094
- Yang, Z. (1997). PAML: a Program Package for Phylogenetic Analysis by Maximum Likelihood. *Bioinformatics* 13, 555–556. doi:10.1093/bioinformatics/13.5.555
- Zhang, D.-C., Guo, L., Guo, H.-Y., Zhu, K.-C., Li, S.-Q., Zhang, Y., et al. (2019). Chromosome-level Genome Assembly of golden Pompano (*Trachinotus Ovatus*) in the Family Carangidae. *Sci. Data* 6, 216. doi:10.1038/s41597-019-0238-8

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Li, Liu, Cui, Hao, Wang, Wang, Jiang, Wang, Feng, Xu, Shao and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.