

Genome of striped eel catfish (*Plotosus lineatus*) provides insights into the adaptive evolution of amphidromous fish

DEAR EDITOR,

Dramatic changes in habitat can have a range of effects on organisms (Bi & Zhang, 2021). *Plotosus lineatus* is a representative fish within Ostariophysi showing adaptation from freshwater habitats to the ocean. Here, we sequenced the genome of *P. lineatus* using long-read and Hi-C sequencing technology. Based on the high-quality chromosome-level genome, we explored the molecular basis of various *P. lineatus* features, including large genome with fewer chromosomes, novel dendritic organ (DO), and amphidromous and venomous nature. Our results showed that changes in the *P. lineatus* habitat may have resulted in genomic enlargement, owing to the wide range of transposon activity caused by genomic shock. Increased copy number of transposons may have promoted chromosome fusion and translocation. Positive selection and gene family expansion also played a role in the adaptive evolution of *P. lineatus*, particularly in immunity and gametogenesis. Tissue-specific expression analysis indicated that some developmental stages of the DO may be similar to those of the kidney, and existing genes may play an important role in the formation of novel organs. We identified candidate genes associated with toxins in *P. lineatus* and found that these genes may have undergone rapid evolution. Tandem duplication of the *stonustoxin* gene may be one reason for *P. lineatus* toxicity. Our findings provide novel insights into the evolution of the striped eel catfish genome, chromosomes, adaptations, novel organs, and virulence.

Using Nanopore and HiFi long-read technologies, we assembled the *P. lineatus* genome, which was approximately 1.33 Gb in size and consisted of 111 contigs with a contig N50 of 36.8 Mb (Supplementary Table S1). These contigs were further assembled into 24 pseudochromosomes at the chromosome scale with an assembly mount ratio of 99.74% (Supplementary Figure S1 and Table S2). Benchmarking Universal Single-Copy Orthologs (BUSCO) assessment showed that 97.69% of the genes were complete (metazoa_odb10 database), indicating high completeness of the assembled genome (Supplementary Table S3). We identified 21 969 protein-coding genes with a BUSCO assessment of 92.80% (actinopterygii_odb9 database), of which 20 378 (92.76%) were functionally annotated (Supplementary Tables S4, S5). Transposable elements (TEs)

accounted for 65.44% of the genome (Supplementary Table S6). DNA and long interspersed nuclear element (LINE) transposons were significantly increased in *P. lineatus*. The increase in LINE transposons was primarily due to the increase in the RTE superfamily (Supplementary Figure S2). The *P. lineatus* genome is the largest among published catfish genomes (<https://www.ncbi.nlm.nih.gov/assembly/?term=Siluriformes>), which is likely due to the proliferation of transposons. According to the genomic shock hypothesis proposed by McClintock (1984), stress and regulatory interference due to environmental changes in the habitat may lead to the mobilization of TEs. Within Ostariophysi, *P. lineatus* is a representative fish showing adaptation from freshwater habitats to the ocean, thus experiencing dramatic changes in habitat during its evolutionary history. This may have caused genomic shock and transposon activation in the genome of its ancestors, thereby affecting genome size.

Chromosome number varies greatly among catfish species (Arai, 2011). Our results showed that *P. lineatus* has fewer chromosomes compared to other catfish with published genomes. To explain this, we used chromosome data from different species (*Cranoglanis boudierius*, *Hemibagrus wyckioides*, *Ictalurus punctatus*, *Leiocassis longirostris*, *Silurus meridionalis*, *Tachysurus fulvidraco*, and *Electrophorus electricus*) and reconstructed 34 proto-chromosomes (Figure 1A; Supplementary Table S7) for the last common ancestor (LCA) of Siluriformes and identified many more chromosome fusion events in *P. lineatus* (Supplementary Figure S3). We selected 15 fusion and translocation sites (FTSs, Figure 1A; Supplementary Table S8), and found that TE content was significantly higher in seven FTSs (c, d, h, i, k, n, and o) than in other regions (Figure 1B; Supplementary Table S9). Only *Tc1-Mariner* and RTE were at higher concentrations (Figure 1C; Supplementary Table S10), and the Mann-Whitney test was used to compare differences in *Tc1-Mariner* and RTE content among the seven sites and other FTSs by examining the main factors causing differences in transposon content at sites. The RTE superfamily ($P=0.00373$) played a dominant role in the enrichment of transposons at FTSs. We speculate that transposons may participate in chromosome fusion and translocation in *P. lineatus*, with RTE playing an important role in these processes. Ectopic recombination is an important cause of chromosomal rearrangements; for example, non-allelic homologous recombination between different ancestral muntjac chromosomes is thought to have led to tandem

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright ©2023 Editorial Office of Zoological Research, Kunming Institute of Zoology, Chinese Academy of Sciences

Received: 03 December 2022; Accepted: 20 February 2023; Online: 20 February 2023

Foundation items: This work was supported by the National Natural Science Foundation of China (31872204) and Fundamental Research Funds for the Central Universities (SWU120049)

fusions (Yin et al., 2021). Furthermore, the copy number of transposons is positively correlated with the frequency of heterotopic recombination (Petrov et al., 2003). This process may be exhibited in *P. lineatus*, whereby transition from freshwater to seawater habitat potentially caused genomic shock, resulting in large-scale transposon activation, increased transposon copy number, and increased heterotopic recombination, and finally the fusion and translocation of *P. lineatus* chromosomes. Synteny analysis (Figure 1D) indicated the occurrence of many chromosomal rearrangements during catfish evolution, possibly explaining the diversity of chromosome numbers in these species.

A phylogenetic tree was constructed based on 3 774 single-copy orthologous genes from *P. lineatus* and nine other fish genomes (Supplementary Figure S4). To detect the potential genetic basis of *P. lineatus* traits compared to other species, we performed positive selection and gene family analyses. We

identified various positively selected genes (PSGs) that may be involved in osmotic pressure regulation. Notably, *efhc2* is crucial for the distal segmentation of the pronephros (primitive urinary organ) in zebrafish. In addition, *pdzk1* is involved in the regulation of proximal tubular Na (+)-dependent inorganic phosphate co-transport. The protein encoded by *atp1b1* is a non-catalytic component of the active enzyme that catalyzes adenosine triphosphate (ATP) hydrolysis, coupled with the exchange of Na⁺ and K⁺ ions across the plasma membrane. Interestingly, we identified several PSGs that may be involved in gametogenesis, specifically genes associated with spermatogenesis, such as *crta*, *ccdc136*, *cfap43*, and *iho1*. Furthermore, the *marf1* gene, which is involved in oogenesis, was shown to be under positive selection. We also identified several possible genes related to nocturnal adaptation in *P. lineatus*, including *ciart*, which is associated with circadian rhythm, and *crba1*, which is associated with lens development.

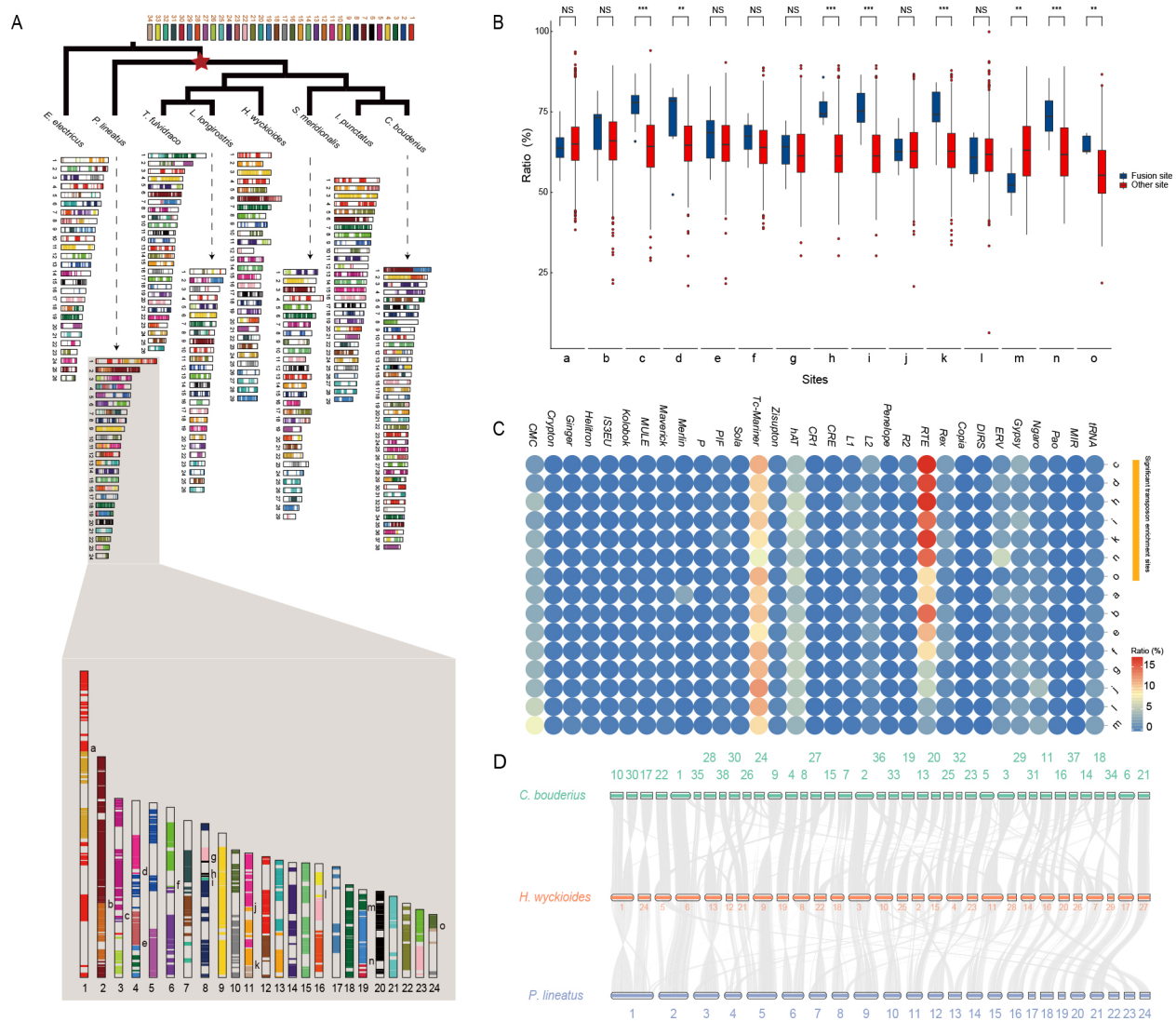


Figure 1 Chromosomal evolution of *P. lineatus*

A: Ancestral chromosome construction of Siluriformes. Conserved blocks are color-coded according to their chromosomal source in Siluriformes ancestor. Designated FTSs (distance between adjacent different colored blocks of less than 1 Mb) are identified by letters in enlarged figure. B: Comparison of TE content between selected fusion site and other sites on corresponding chromosome (100 kb window, significant differences: NS: Not significant; **: $P < 0.01$; ***: $P < 0.001$). Letters on horizontal axis indicate different fusion sites, and different colors indicate selected fusion site and other sites on corresponding chromosome. C: Heatmap of TE superfamily content at selected fusion site. D: Pairwise whole-genome alignments across *C. boudierus*, *H. wyckioides*, and *P. lineatus* genomes. Each horizontal bar represents one chromosome, and chromosome IDs of the three species are labeled.

The above PSGs and related references are presented in Supplementary Table S11. In addition, 345 and 2 942 gene families were expanded and contracted in *P. lineatus*, respectively, including 63 significantly expanded gene families (Supplementary Table S12). Immune-related genes were considerably enriched based on Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) analysis (Supplementary Tables S13, S14). Many GO terms were related to virus defense processes (Supplementary Figure S5A), suggesting that *P. lineatus* may have significantly enhanced viral defense capability, consistent with the viruses encountered in seawater and freshwater habitats. Furthermore, circadian rhythm-related pathways and terms were also enriched, which may be related to their nocturnal behavior. The *sox* gene family was also expanded (43 members) in *P. lineatus* compared with that in *C. boudierus* and *H. wyckioides* (26 and 25 members, respectively), which was mainly due to *sox30* (Supplementary Figure S6 and Table S15). The *sox30* gene is involved in spermatogenesis in fish (Wei et al., 2022), suggesting that *P. lineatus* may also have adapted to the marine environment during germ cell processes.

Plotosidae catfish have an extra-branchial salt-secreting DO organ, which is unique among catfish and all marine bony fish (Kolbadinezhad et al., 2018). We sequenced the transcriptomes of 13 tissues from *P. lineatus* to study DO formation. Consequently, we obtained 305 genes specifically expressed in DO tissues (Supplementary Table S16), which were thus subjected to GO and KEGG enrichment analyses (Supplementary Tables S17, S18). Results showed that the genes were enriched in two KEGG pathways, i.e., “Proximal tubule bicarbonate reclamation” and “Mineral absorption”, and in GO terms related to embryonic and kidney development (Supplementary Figure S5B). These findings are consistent with previous studies showing that the DO is osmoregulatory (Kolbadinezhad et al., 2018) and suggest similarities between DO and kidney development, as specifically expressed DO genes were enriched in kidney development. Tissue-specific expression analysis also indicated that existing genes (DO specific-expressed genes related to renal tubule development) may play an important role in novel organ development. This is essentially similar to our previous findings in lungfish, where the *sftpb* gene, important in respiratory system evolution, already existed in the common ancestors of bony fish and thus is not novel (Wang et al., 2021). We also identified transposon activation in 91 of the 305 purposely expressed genes, including 33 genes specifically in the DO (Supplementary Table S19). Of these 33 genes, *six1b*, *six4*, and *chs2* were enriched in “renal tubule morphogenesis/development”, and the corresponding activated transposons were long terminal repeat type (LTR, unknown superfamily), *Ngaro*, and *RTE*, respectively. These results suggest that transposons may be involved in DO tissue-specific expression of genes as well as DO formation.

Plotosus lineatus has received increasing attention due to its possible high toxicity. However, among catfish with published genomes, the virulence of most species is ambiguous, with only *Ameiurus melas* found to be non-toxic (Wright, 2009). We comprehensively annotated venom-like genes in *A. melas* and *P. lineatus* using a multistep approach and identified 86 and 98 venom-like genes in their genomes, respectively (Supplementary Table S20). We evaluated the molecular evolutionary characteristics of venom-like genes in

A. melas and *P. lineatus* by calculating the number of synonymous (K_s) and nonsynonymous (K_a) nucleotide substitutions per site for each pair of venom-like and other genes in the two species. Results showed that the K_a/K_s substitution ratios of venom-like gene pairs were significantly higher than those of other gene pairs (Supplementary Figure S7A–C and Table S21), suggesting that rapid evolution may be one way to form venom-like genes. Comparing venom-like genes in the two species, *stonustoxin* genes were found to be more abundant in *P. lineatus*. *Stonustoxin* is an extremely potent cytolytic toxin, mainly found in stonefish (Ellisdon et al., 2015). Here, we identified 12 *stonustoxin* genes in the *P. lineatus* genome and nine in *A. melas*. Importantly, we found tandem duplications in nine *stonustoxin* genes in the *P. lineatus* genome (Supplementary Figure S7D). In contrast, *A. melas* had only five *stonustoxin* genes at the corresponding locations (two very close clusters). We also observed the distribution of *stonustoxin* genes in corresponding regions of *T. fulvidraco*, a species with toxicity characteristics, and identified 10 *stonustoxin* genes, although no continuous tandem repeats. Tandem repeat genes can affect expression and evolution; for example, duplicated genes often accumulate more transcripts to enhance expression (Picart-Piccolo et al., 2020). However, among the toxin candidate genes in *P. lineatus*, some were immune-related genes. For example, there were 14 lectin-related genes in *P. lineatus* and 10 lectin-related genes in *A. melas*, as well as six *aep1* genes in *P. lineatus* and three *aep1* genes in *A. melas*. These results suggest that gene number expansion may be an important pathway for toxin formation in fish. Notably, there appears to be a link between immune-related genes and toxin genes (Ellisdon et al., 2015).

DATA AVAILABILITY

The striped eel catfish chromosome-level genome and genome sequencing data used in the present study, including Nanopore, PacBio, MGISEQ2000, Hi-C, and transcriptome data, are available from the Sequence Read Archive (SRA) with the BioProjectID PRJNA720277. The above data were also archived in the Science Data Bank database (DOI: 10.57760/sciencedb.06933) and Genome Sequence Archive (GSA) database (CRA009258).

SUPPLEMENTARY DATA

Supplementary data to this article can be found online.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

F.S. and Z.G.P. conceived the study and wrote the manuscript. F.S. and P.L. conducted field sample collection and prepared samples for sequencing. F.S. conducted bioinformatics analysis. Z.C.W. manually checked the genes under study. Y.X. participated in the construction of the ancestral genome of catfish. All authors read and approved the final version of the manuscript.

ACKNOWLEDGMENTS

We thank Lu-Yun Ni, Min-Zhi Zeng, Yan-Qiu Zhang, Huai-Yi Fang, and Xue-Lan Chen for their help with sample collection. We are grateful to Drs. Cheng-Long Zhu, Xu-Peng Bi, and Long Zhou for help with the reconstruction of ancestral karyotype. We also thank Drs. Lei Chen, Ru Zhang, Ze-Shan Lin, Bao Wang, Zhong-Kai Wang, Yan-Dong Ren, Yuan Yin, and Luo-Hao Xu for helpful discussions.

Feng Shao¹, Ze-Chen Wu¹, Yuan Xu¹, Ping Li²,
Zuo-Gang Peng^{1,3,*}

¹ Key Laboratory of Freshwater Fish Reproduction and Development (Ministry of Education), Southwest University School of Life Sciences, Chongqing 400715, China

² College of Fisheries, Southwest University, Rongchang, Chongqing 402460, China

³ Academy of Plateau Science and Sustainability, Qinghai Normal University, Xining, Qinghai 810008, China

*Corresponding author, E-mail: pzg@swu.edu.cn

REFERENCES

- Arai R. 2011. Fish Karyotypes: A Check List. Tokyo: Springer.
- Bi XP, Zhang GJ. 2021. Ancestral developmental potentials in early bony fish contributed to vertebrate water-to-land transition. *Zoological Research*, **42**(2): 135.
- Ellisdon AM, Reboul CF, Panjekar S, et al. 2015. Stonefish toxin defines an ancient branch of the perforin-like superfamily. *Proceedings of the National Academy of Sciences of the United States of America*, **112**(50): 15360–15365.
- Kolbadezhad SM, Coimbra J, Wilson JM. 2018. Osmoregulation in the Plotosidae catfish: role of the salt secreting dendritic organ. *Frontiers in Physiology*, **9**: 761.
- Mcclintock B. 1984. The significance of responses of the genome to challenge. *Science*, **226**(4676): 792–801.
- Petrov DA, Aminetzach YT, Davis JC, et al. 2003. Size matters: non-LTR retrotransposable elements and ectopic recombination in drosophila. *Molecular Biology and Evolution*, **20**(6): 880–892.
- Picart-Piccolo A, Grob S, Picault N, et al. 2020. Large tandem duplications affect gene expression, 3D organization, and plant-pathogen response. *Genome Research*, **30**(11): 1583–1592.
- Wang K, Wang J, Zhu CL, et al. 2021. African lungfish genome sheds light on the vertebrate water-to-land transition. *Cell*, **184**(5): 1362–1376.e18.
- Wei L, Tang YH, Zeng XH, et al. 2022. The transcription factor Sox30 is involved in Nile tilapia spermatogenesis. *Journal of Genetics and Genomics*, **49**(7): 666–676.
- Wright JJ. 2009. Diversity, phylogenetic distribution, and origins of venomous catfishes. *BMC Evolutionary Biology*, **9**: 282.
- Yin Y, Fan HZ, Zhou BT, et al. 2021. Molecular mechanisms and topological consequences of drastic chromosomal rearrangements of muntjac deer. *Nature Communications*, **12**(1): 6858.