

RESEARCH

Open Access



Chromosome-level genome assembly of *Trichiurus japonicus* resolves species boundaries and evolutionary history of cutlassfishes in Northwest Pacific

Vanthu Giap^{1,2†}, Tianqin Wu^{1,2†}, Liang Lu^{1,2}, Jiantao Hu^{1,2}, Mingwei Zhou^{1,2} and Chenhong Li^{1,2*}

Abstract

Persistent taxonomic uncertainty has plagued the Trichiuridae family (hairtails), a group of ecologically and commercially significant fishes, a result of conflicting morphological criteria and historical misclassifications. While molecular approaches have advanced understanding their taxonomy, mitochondrial-based analyses and misapplied nomenclature have obscured species boundaries and population dynamics. Here, we integrate a chromosome-level genome assembly of *Trichiurus japonicus* with whole-genome resequencing of congeneric species to resolve the taxonomic ambiguity in *Trichiurus*. Combining phylogenomic, demographic, and population genetic analyses, we resolve taxonomic uncertainties by redefining species boundaries, reconstructing population dynamics, and identifying genomic divergence between populations. Our analyses suggest that *T. japonicus* and *T. nanhaiensis* are distinct species. Crucially, we identify a cryptic, genetically distinct lineage putatively derived from the Indian Ocean, misclassified as *T. lepturus* in previous studies, underscoring the impact of commercial trade on taxonomic confusion. Our analyses reveal low spatial genetic differentiation in *T. japonicus* across the sampled Northwest Pacific localities, consistent with weak population structure within our sampling frame. This study establishes a genomic framework of *Trichiurus* to resolve taxonomic conflicts in morphologically cryptic taxa and advances strategies for sustainable fisheries management under climate change.

Keywords Cutlassfish, *Trichiurus*, Population genomics, Taxonomy, Pleistocene glaciations

[†]Vanthu Giap and Tianqin Wu contributed equally to this work.

*Correspondence:

Chenhong Li
chli@shou.edu.cn

¹Shanghai Universities Key Laboratory of Marine Animal Taxonomy and Evolution, Shanghai Ocean University, Shanghai 201306, China

²Engineering Research Center of Environmental DNA and Ecological Water Health Assessment, Shanghai Ocean University, Shanghai 201306, China



© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Significance statement

This study employs genomic approaches to resolve taxonomic uncertainties in the *Trichiurus* genus, providing genome-wide evidence of *T. japonicus* and *T. nanhaiensis* as distinct species while identifying a cryptic, genetically distinct lineage putatively derived from the Indian Ocean (*T. sp.*). Our analyses indicate that *T. japonicus* maintained stable population dynamics throughout Pleistocene glaciations. These results may advance management strategies for commercially important cutlassfish species and establish a crucial genomic framework for evaluating marine organisms' vulnerability to contemporary climate-driven environmental changes.

Introduction

The Trichiuridae family (cutlassfishes or hairtails) comprises ecologically and economically significant fish [1, 2]. Their role as mid-to-upper trophic level predators makes them vital for maintaining the structure and function of marine ecosystems. Fisheries in the Northwest Pacific rely heavily on species of the genus *Trichiurus*, including *T. japonicus*, *T. lepturus*, *T. nanhaiensis*, and *T. brevis* [3, 4], although pronounced morphological conservatism among these taxa has long complicated accurate identification [5–7]. This taxonomic ambiguity directly undermines stock assessments and sustainable management, as effective conservation strategies relies on the accurate delineation of species and their population units.

Initial taxonomic work regarded cutlassfishes as a single cosmopolitan species [8, 9], but subsequent morphological revisions recognized multiple East Asian taxa [10, 11]. The advent of molecular markers, particularly mitochondrial DNA (mtDNA), revealed deep evolutionary divergences within the genus, challenging the traditional view and supporting the existence of cryptic species [8, 12, 13]. While mitochondrial DNA (mtDNA) is useful for initial screening and barcoding, its maternal, non-recombining, single-locus nature makes it susceptible to stochastic processes such as genetic drift and incomplete lineage sorting. These can produce evolutionary signals that conflict with genome-wide divergence patterns and true species relationships [14, 15]. Historical nomenclatural acts established the foundational framework for current taxonomy of *Trichiurus* [16–20], yet the relationships and validity of many nominal species remained contested [21].

Despite these advances, critical knowledge gaps persist due to methodological limitations. First, earlier investigations were often constrained by geographically restricted sampling and a primary reliance on mitochondrial markers, which has frequently resulted in cytonuclear discordance and ambiguous species boundaries [22–24]. This discordance arises when genealogies from nuclear and mitochondrial genomes conflict, a common

issue in recently diverged species or those with complex demographic histories, obscuring true evolutionary relationships [22, 25]. Second, several described taxa retain uncertain morphological diagnoses [3, 8, 26], and key lineages, including *T. brevis* [16, 26], *T. sp.2* [27–30], and an unresolved Indian Ocean lineage [21, 31], lack clear phylogenetic placement and genomic characterization. Third, while early isozyme studies suggested fine-scale population structure for *T. japonicus* [4, 32], subsequent genetic data with limited resolution supported only two broad groups in the East and South China Seas [11, 33], leaving its contemporary population connectivity and historical demography poorly understood.

In this study, we have three objectives: firstly, using genome-wide nuclear data to resolve long-standing species boundaries in the *Trichiurus* genus, which have been obscured by morphological conservatism and mitochondrial discordance in previous studies; secondly, test whether Pleistocene sea-level fluctuations in the Northwest Pacific are a key driver of allopatric speciation in *Trichiurus*; thirdly, evaluate spatial genetic structure of *T. japonicus* exhibits across its Northwest Pacific range, if it is consistent with a panmictic population hypothesis. To address these interconnected limitations in taxonomy, species delimitation, and population genomics, we established a robust phylogenomic framework for Northwest Pacific *Trichiurus*. By integrating extensive specimen collections across the region's major marine basins with a chromosome-level genome assembly and whole-genome resequencing, we apply genome-wide SNP analysis and coalescent-based modeling. The specific goal of this study is to clarify the phylogenetic status of ambiguous lineages; delineate the population structure, connectivity, and demographic history of the commercially dominant *T. japonicus*; and thereby create an empirical genomic foundation for informed regional fisheries management and conservation.

Result

Genome assembly and annotation

We present a high-quality, chromosome-level genome assembly for *T. japonicus* (Fig. 1C). This 964.99 Mbp assembly, anchored to 24 chromosomes (98.87% anchoring rate), was generated by integrating 112 GB of Oxford Nanopore long reads, 66.37 GB of Illumina short reads, and 104.11 GB of Hi-C scaffolding data, resulting in scaffold and contig N50 values of 39.19 Mbp and 6.07 Mbp respectively (Supplementary Table 1). The genome exhibits a GC content of 41.21% and achieved a high BUSCO completeness score of 96.9%.

Genome composition and repetitive elements. Repetitive elements constituted a substantial portion (40.48%) of the *T. japonicus* genome. DNA transposons were the most abundant class (10.95%), followed by retroelements

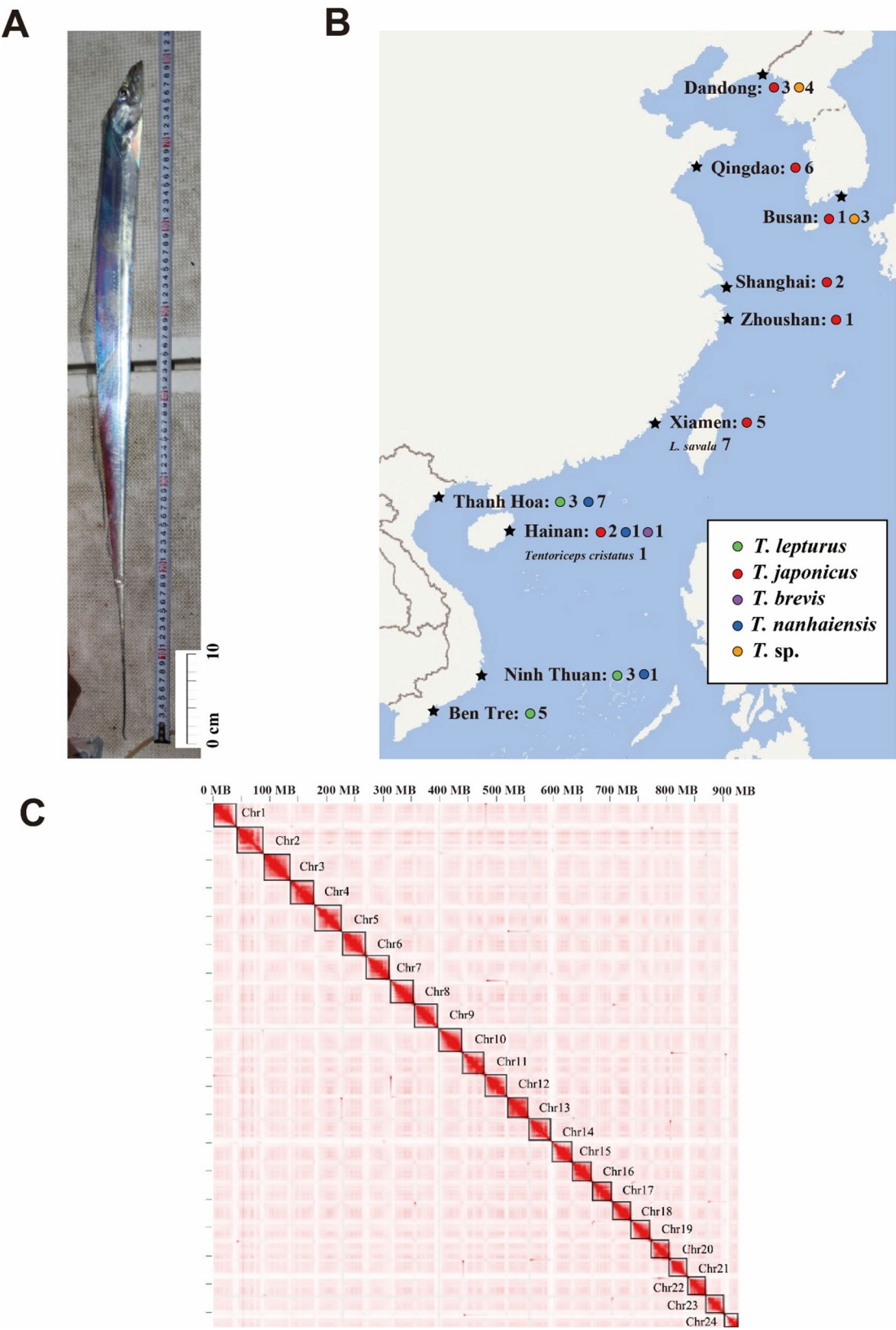


Fig. 1 **A**. A specimen of *Trichirus japonicus* used for genome sequencing, collected by hand rod nearby Gouqi Island, Zhoushan, China; **B**. Sampling locations for genome resequenced specimens, with numbers indicating the count of individuals of each species collected at each site; **C**. Chromosome contact map of *T. japonicus* (resolution in 1 kb bins)

(10.50%), which comprised primarily long terminal repeats (LTRs; 6.10%) and long interspersed nuclear elements (LINEs; 3.89%). Unclassified repeats accounted for an additional 14.08% of the genome. Repetitive elements dominated the *T. japonicus* genome, accounting for nearly half of its composition (40.48%). DNA transposons (10.95%) and retroelements (10.50%) were the two major repeat classes, with unclassified repeats comprising the remainder (14.08%). Integrating transcriptomic, protein homology, and ab initio evidence, we annotated 23,993 high-confidence protein-coding genes. This gene set provides a robust foundation for comparative genomic analyses.

Compared with the published HiFi-based *T. japonicus* genome of Hsu et al. (2025) [34], our Nanopore-based assembly differs in assembly size, gene count, and BUSCO completeness, but both studies consistently recover 24 chromosomes, indicating consistency in chromosome number; furthermore. Synteny mapping between the assemblies also suggest structural concordance (Supplementary Fig. 5), despite differences in total size (964.99 Mb vs. 1,138 Mb), gene count (23,993 vs. 26,541), and BUSCO completeness (96.9% vs. 99.4%). Beyond genome description, our study further links the reference assembly to population genomics, phylogenomics, and demographic inference to address species boundaries and evolutionary history within *Trichiurus*, whereas Hsu et al. primarily presented a genomic resource. Our comparative analyses also reveal strong chromosomal synteny conservation with related perciforms (*Scomber japonicus*, *Thunnus albacares*), as evidenced by large collinear blocks spanning entire chromosomes (Supplementary Fig. 1A). Furthermore, spatial mapping uncovers significant correlations between syntenic regions, GC content gradients, gene density, and transposable element distribution (Supplementary Fig. 1B), highlighting the interplay between genome architecture and sequence composition.

Population genomics and phylogenomics

Northwest Pacific *Trichiurus* taxa revealed clear genetic divergence in PCA (Fig. 2A, left). *Trichiurus japonicus* formed a tight cluster, indicating low genomic diversity. *Trichiurus lepturus* separated distinctly from all other species. While closely grouped, *T. sp.* and *T. nanhaiensis* separate from both *T. japonicus* and *T. lepturus*. *Trichiurus nanhaiensis* and *Trichiurus sp.* formed two distinct genetic clusters, which is more apparent in the two-species PCA (Fig. 2A, right). Applying a MAF threshold of 0.05 further accentuated this genetic separation. Because MAF thresholds can affect population-structure inference, this pattern should be interpreted together with the unfiltered PCA, F_{ST} and phylogenomic analyses.

Genetic differentiation among the four *Trichiurus* species was medium to high ($F_{ST} = 0.1728$ – 0.3800 ; Fig. 2B), Divergence was greatest between *T. lepturus* and *T. sp.*, and lowest between *T. lepturus* and *T. japonicus*. In contrast, northern and southern *T. japonicus* populations showed negligible differentiation ($F_{ST} = 0.0188$), consistent with a single cohesive coastal population within our 2022–2023 multi-season sampling frame. Genome-wide D_{xy} , F_{ST} and π patterns supported divergence among *Trichiurus* species and low differentiation within the sampled *T. japonicus* populations (Supplementary Fig. 2). The apparent difference between F_{ST} and PCA/ADMIXTURE patterns likely reflects the use of different SNP-filtering schemes and the fact that PCA and ADMIXTURE emphasize clustering patterns, whereas F_{ST} measures allele-frequency differentiation. Therefore, these analyses were interpreted as complementary but method-dependent lines of evidence.

ADMIXTURE analysis revealed three major genetic clusters at $K=3$, corresponding to *T. lepturus*, *T. japonicus*, and a combined *T. nanhaiensis*–*Trichiurus sp.* group (Fig. 2C). These clusters were broadly consistent with the major species-level groupings inferred from PCA and phylogenomic analyses. At higher K values ($K=5$ to $K=7$), *T. japonicus* individuals formed distinct subclusters, but these showed no correlation with geographic sampling locations.

Intraspecific genomic variation within *T. japonicus* was low compared with interspecific differentiation. The weighted F_{ST} between northern and southern populations was only 0.01181, suggesting that *T. japonicus* individuals across the sampled range share a relatively uniform genetic background and are consistent with a single, weakly differentiated population within our sampling frame. After applying a MAF threshold of 0.05, ADMIXTURE analysis showed clearer separation between *T. nanhaiensis* and *Trichiurus sp.* This threshold-dependent pattern should be interpreted cautiously and in combination with other analyses.

Two individuals showed possible evidence of genetic admixture: *T. japonicus* h11 and *Trichiurus sp.* 661 (Dandong Fig. 1A). Because these individuals differ in sequencing depth, these signals should be interpreted cautiously and require confirmation with additional high-coverage samples. These findings may reflect localized gene flow, historical hybridization or coverage related uncertainty. Most *T. japonicus* across its range showed a relatively uniform genetic background, consistent with weak population structure in the present dataset. *Trichiurus brevis* showed as a distinct cluster, although this result may be influenced by the small sample size ($n=1$).

Overall, the ADMIXTURE results support three principal genetic patterns: distinction between *T. lepturus*, *T. japonicus*, and the *T. nanhaiensis*–*T. sp.* group; a

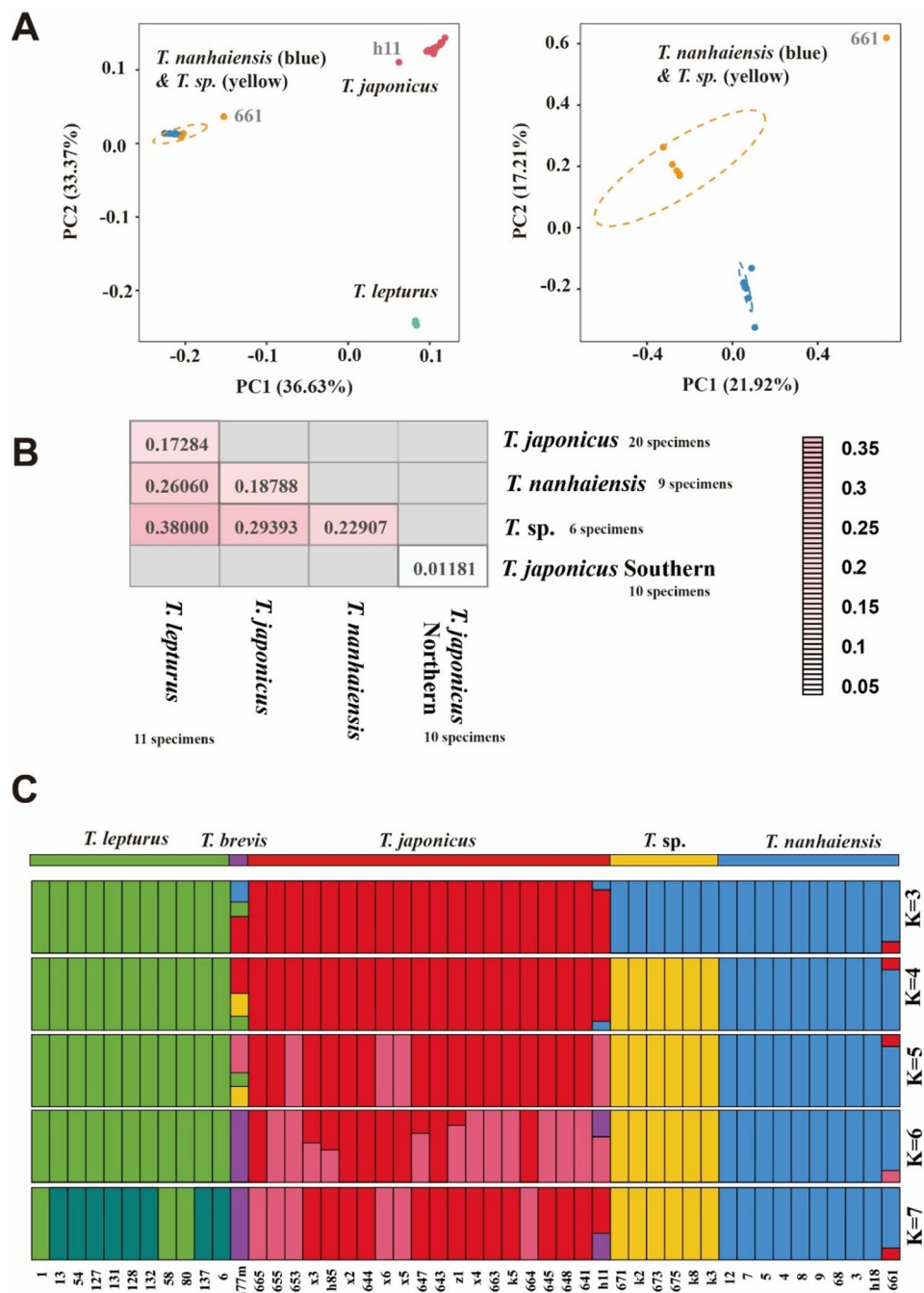


Fig. 2 **A.** Principal Component Analysis (PCA) of four *Trichiurus* taxa (left) and the PCA of *T. nanhaiensis* and *T. sp.* (right). **B.** The heatmap of weighted F_{ST} values calculated by VCFtools among the four species and between the presumed Yellow Sea and East China Sea populations of *T. japonicus*. **C.** Admixture proportions for five *Trichiurus* taxa

cohesive *T. japonicus* population along the Chinese coast; and localized admixture suggesting past or limited contemporary gene flow at species boundaries.

Phylogenomics

For clarity, we present the nuclear phylogeny first (Fig. 3A) and the mitochondrial phylogeny second

(Fig. 3B), allowing direct comparison between genome-wide and mitochondrial signals. Phylogenomic analysis resolved the evolutionary relationships within *Trichiurus* but revealed cytonuclear discordance (Fig. 3). The nuclear phylogeny (Fig. 3A) supported a basal lineage of South China Sea *T. lepturus* and resolved *T. japonicus*, *T. nanhaiensis*, and *T. sp.* as distinct genomic lineages, with *T.*

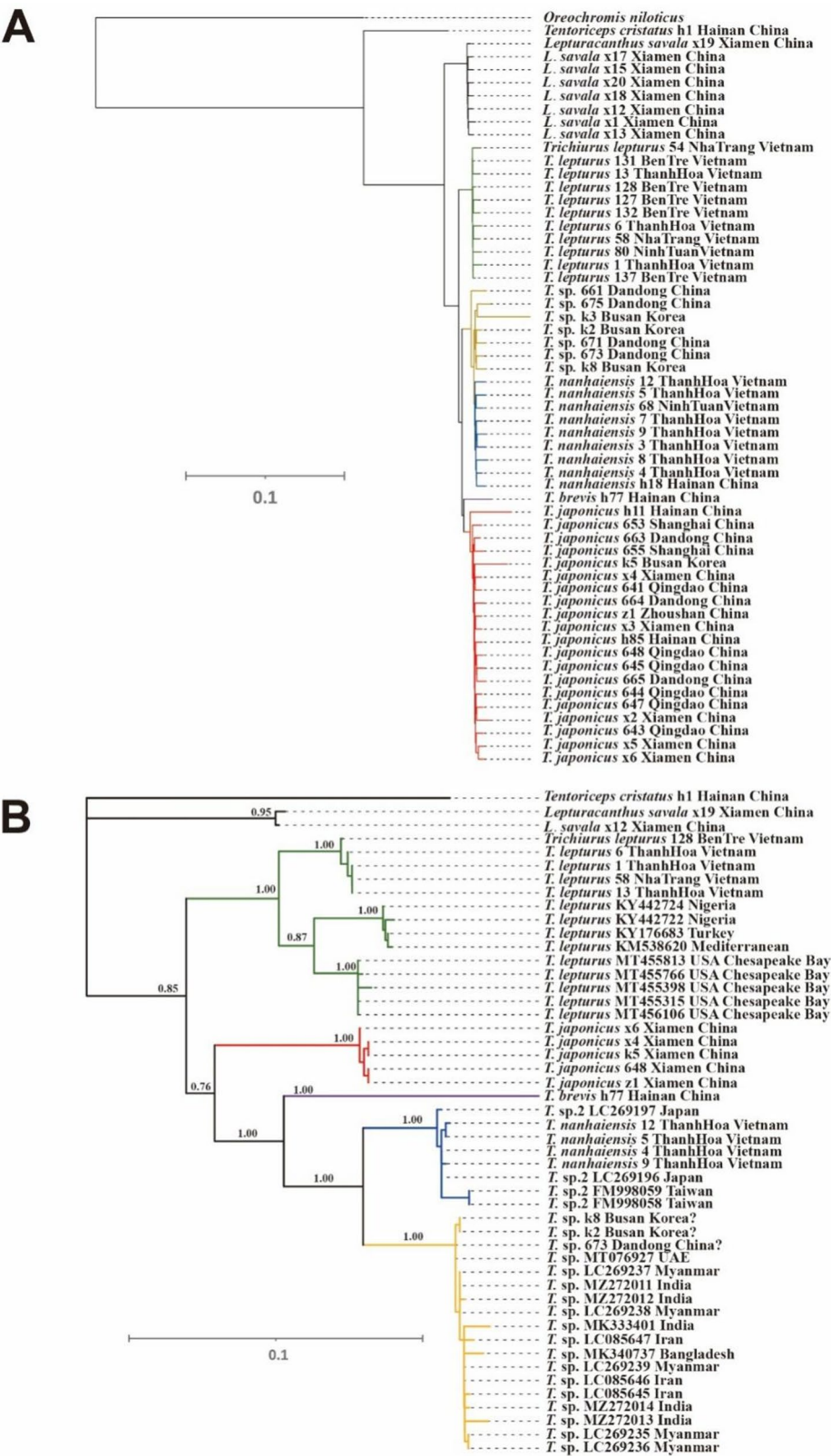
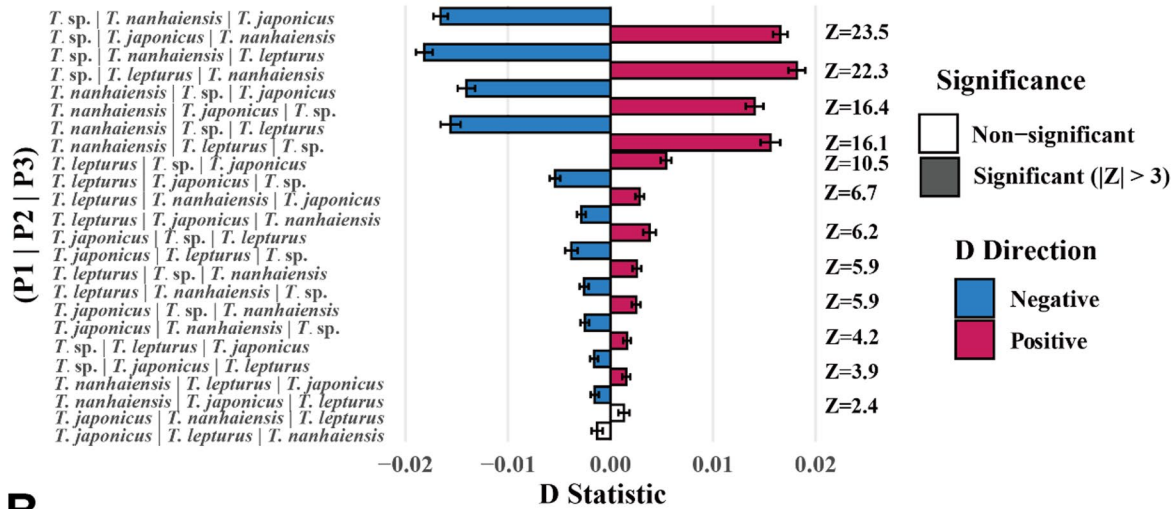


Fig. 3 Phylogenetic relationships of *Trichiurus* species. **A** The phylogenetic tree reconstructed from 4,434 nuclear single-copy orthologs using RaxML with 1,000 bootstrap. **B** Mitochondrial phylogeny inferred from COX1 sequences, combined resequencing data and sequences retrieved from NCBI (accessions shown on the taxon labels)

A f4-statistics for *Trichiurus* ((Outgroup: *Lepturacanthus savala*, P1), (P2, P3))



B

Effective population size dynamics (MSMC-IM)

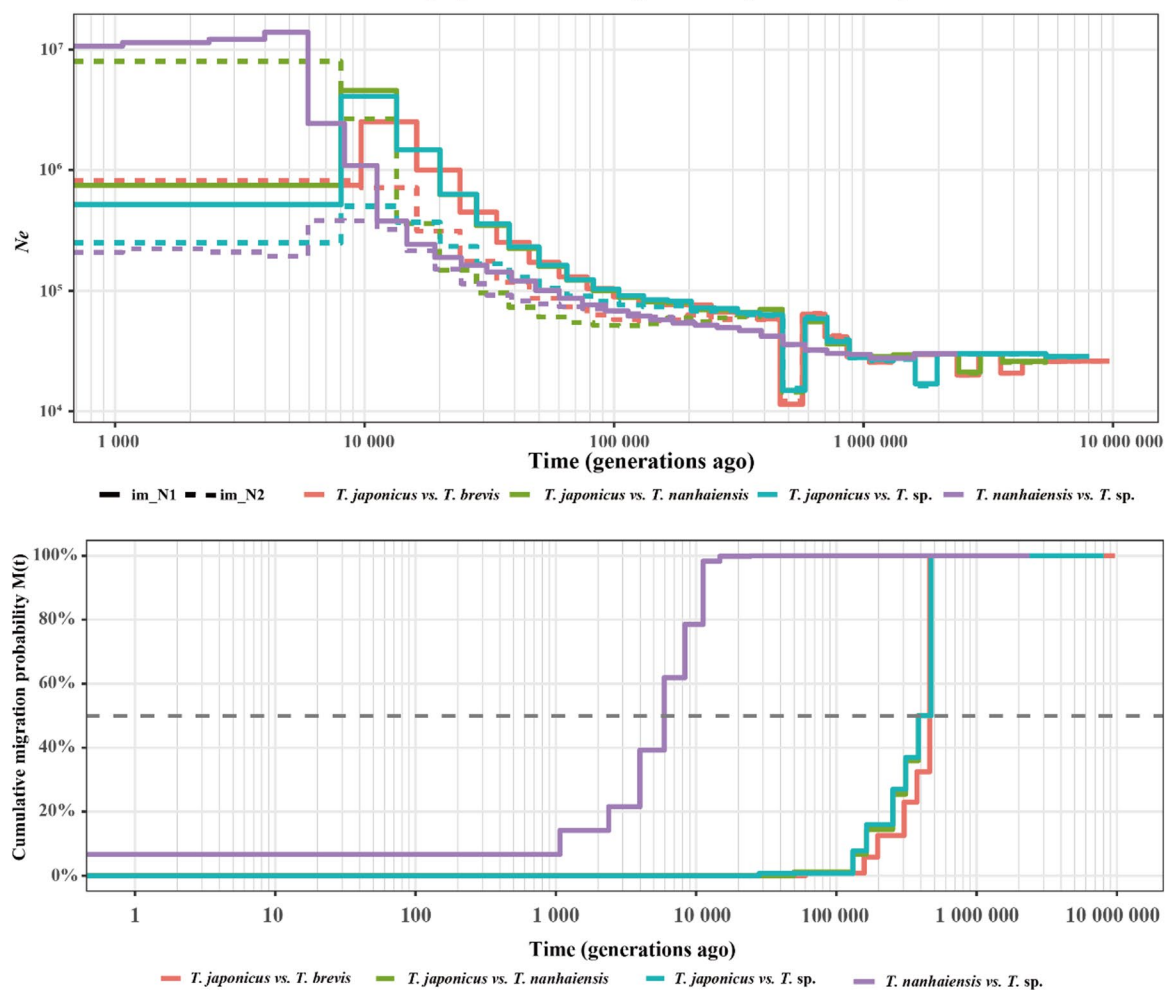


Fig. 4 (See legend on next page.)

(See figure on previous page.)

Fig. 4 A f4-statistics ((Outgroup: *Lepturacanthus savala*, P1), (P2, P3)) for various *Trichiurus* species comparisons, with bars indicating the D statistic and Z-scores (right) denoting significance. Colors represent the direction of D (blue: negative; red: positive), and significance is defined as $|Z| > 3$ (gray fill). **B** Pairwise MSMC-IM results: effective population sizes of the two populations ($im_N1 = N_1(t)$, solid; $im_N2 = N_2(t)$, dashed) inferred for each *Trichiurus* species pair (colors), and cumulative migration probability $M(t)$ (bottom). The x-axis is scaled by time (generations ago, log scale), with N_e on a log scale (top) and $M(t)$ on a linear scale (bottom)

japonicus sister to a clade containing the latter two. The mitochondrial gene tree (Fig. 3B) recovered three well-supported monophyletic clades of *T. lepturus* (West African, American, Vietnamese). Nuclear sequences for the West African and American clades were unavailable, so their nuclear differentiation remains unassessed. All three lineages are still called *T. lepturus* because Linnaeus originally applied this name to three distinct hairtail genera in 1758. Other related taxa have since been resolved, leaving only these three monophyletic lineages as *T. lepturus* in the strict sense. The two trees are largely congruent at the species level (e.g., *T. japonicus* forms a consistent clade in both analyses). The difference we refer to is primarily within *T. lepturus*, where the mitochondrial tree shows a more admixed and geographically inconsistent structure compared to the more resolved nuclear phylogeny. These discrepancies are reflected in local topology and node support rather than in major species-level relationships. Given that the mitochondrial inference is based on a single locus (COX1), this pattern may result from incomplete lineage sorting or introgression. Therefore, we interpret the mitochondrial signal cautiously as provisional evidence of lineage complexity rather than definitive mito-nuclear discordance.

The analysis also clarified the identity of market specimens. Specimens labeled *Trichiurus* sp. from Dandong (China) and Busan (South Korea) clustered with an unnamed Indian Ocean lineage, matching the GenBank sequence MK333401 identified as *T. lepturus* [35]. Furthermore, COX1 data supported the interpretation that specimens previously called *Trichiurus* sp. 2 correspond to *T. nanhaiensis*.

Gene flow and divergence history

D-statistic analyses detected significant asymmetric allele sharing between *T. nanhaiensis* and *Trichiurus* sp., whereas comparable signals involving *T. japonicus* and other congeners were weak or not significant (Fig. 4A). Divergence timing estimates revealed that most species pairs separated approximately 383,000–398,000 generations ago. In contrast, *T. nanhaiensis* and *Trichiurus* sp. diverged much more recently, at only 4,903 generations ago. Using a generation time of 2.9 years for *Trichiurus*, this recent divergence is consistent with incomplete lineage separation between the two lineages into the Holocene.

Our analysis identified a major historical population bottleneck in the genus *Trichiurus* occurring

approximately 474,010 generations ago, equivalent to ~1.12 million years ago using a generation time of 2.9 years (Fig. 4B). This bottleneck coincided with Pleistocene glacial-interglacial cycles and associated sea-level fluctuations. Subsequent major divergence events followed this bottleneck. The separation of *T. japonicus* from its relatives (~0.39–0.80 Mya) coincided with the closure of the Taiwan Strait during glacial periods, indicating geographic isolation as the driver of speciation.

Discussion

This study helps resolve long-standing taxonomic ambiguity within the genus *Trichiurus* in the Northwest Pacific by employing a phylogenomic approach. Our integration of chromosome-level genome assembly with genome-wide SNP data reduces the limitations of previous studies that relied solely on morphology or mitochondrial DNA, which can be inadequate for delineating closely related marine species due to phenotypic plasticity and incomplete lineage sorting. We delineate four distinct evolutionary lineages: *T. japonicus*, *T. nanhaiensis*, *T. lepturus*, and a newly identified cryptic lineage (*T. sp.*), thereby providing a robust genomic framework for the genus. Our research presents the first chromosome-level genome assembly for *T. japonicus*, an ecologically and commercially vital Northwest Pacific hairtail. Our high-quality assembly anchors to 24 chromosomes with good completeness (BUSCO: 96.9%), providing a great genomic resource for the understudied Trichiuridae family. Our assembly (GCA_046254865.1) was submitted to NCBI in December 2024 and released on bioRxiv shortly thereafter, predating the recently reported *T. japonicus* genome (Sci. Data, 2025, 12:1860). Prior to this work, no high-quality chromosome-level reference or genome-wide SNP data existed for *Trichiurus*, precluding robust analysis of nuclear divergence, population structure, and demographic history.

Compared to the HiFi-based assembly by Hsu et al. (2025) [34] (1,138 Mb; 26,541 genes; 99.4% BUSCO), our Nanopore-based assembly is more compact (931.81 Mb; 23,993 genes; 96.9% BUSCO), reflecting differences in sequencing platform, assembly algorithms, and haplotype resolution. Whole-genome synteny analysis (Supplementary Fig. 5) revealed high structural concordance across all chromosomes, confirming the overall karyotype stability of *T. japonicus*. Importantly, both assemblies converge on 24 chromosomes, confirming karyotype stability and supporting the chromosome level organization

of our assembly. While Hsu et al. focused on genome description, our study extends to comparative genomics, gene family evolution, and adaptive signatures, filling a critical interpretive gap.

Our divergence estimates of *Trichiurus* species (383,689–398,073 generations ago, corresponding to ~1.11–1.15 million years ago with a generation time of 2.9 years) align with the timeline of Pleistocene glacial lowstands, when the emergence of the Taiwan Strait Land Bridge acted as a vicariant barrier. This geographic isolation would have separated ancestral populations between the East China Sea and the South China Sea/Indo-Pacific, initiating genetic divergence, a plausible scenario that is consistent with the timing of known glacial-interglacial cycles and the observed species divergence pattern. Our genomic data support the recognition of *T. japonicus* as a distinct species, consistent with Temminck & Schlegel (1844) [36] and rejecting the long-misapplied synonymy with *T. haumela* Forsskal, (1775) [37]. This controversy arises because the name *T. haumela* has been historically misapplied as a catch-all label for *T. japonicus* in regional collections, creating nomenclatural inconsistency where a single biological entity is referenced by multiple names [3, 38]. Thus, we clarify that *T. haumela* is not a true junior synonym but a distinct name erroneously assigned to *T. japonicus* lineages due to phenotypic similarity [27]. The persistence of this taxonomic error stems from the pronounced morphological conservatism within *Trichiurus*, which confounds traditional identification [4, 21]. This initial ambiguity was then cemented through the uncritical repetition of misidentifications in the literature [39, 40] and practical challenges in accessing verifiable type specimens for comparison [41]. Our genome-wide clustering and species delimitation analyses argue against conspecificity under the name *haumela*, and provide evidence to help resolve this long-standing confusion. Consequently, this work underscores the utility of genome-scale data in clarifying species boundaries.

Furthermore, we uncovered a significant case of market mislabeling with direct implications for trade and conservation [42, 43]. Specimens sold as locally caught were genetically identified as a novel lineage (*T. sp.*) that clusters with Indian Ocean populations, matching the GenBank sequence MK333401 [35] (Fig. 3, B). This suggests that these specimens may represent imported products, rather than a resident Yellow Sea population. This finding highlights the critical need for forensic seafood science, demonstrating how genetic tools can be leveraged for market surveillance to combat fraud, ensure accurate labeling, and ultimately support sustainable fisheries management [44, 45].

Finally, we document two concurrent speciation processes (Fig. 4A). The first involves *T. nanhaiensis* and *T. sp.*, where high gene flow across weak ecological

gradients maintains incomplete divergence, trapping them in an early stage of speciation [46]. In contrast, *T. japonicus* and *T. nanhaiensis* exhibit a classic pattern of allopatric speciation: Pleistocene isolation in separate glacial refugia established strong reproductive barriers [47], and although subsequent secondary contact permitted limited hybridization [48, 49], species integrity remains intact, a pattern observed in numerous other taxa [50–52]. This evolutionary dynamics mirrors patterns seen in terrestrial taxa (e.g., human [53] or brown bears [54]) and demonstrates how cyclical climate-driven isolation and reconnection have shaped marine biodiversity.

The main limitations of this study are as follows: first, sequencing depth was uneven among resequenced individuals, which may influence genotype accuracy and downstream PCA or ADMIXTURE patterns, particularly for individuals showing possible admixture. Second, the use of different filtering thresholds, especially MAF thresholds, can affect population-structure inference; therefore, ADMIXTURE patterns based on MAF=0.05 should be interpreted together with PCA, F_{ST} , D-statistics, and phylogenomic results. Third, the single sample of *T. brevis* ($n=1$) precludes a definitive assessment of its taxonomic status. While this individual formed a distinct cluster at high K-values in our structure analysis, this result requires validation with additional specimens to rule out potential artifacts or exceptional individual variation. Fourth, sampling gaps persist in the northern distribution range of *T. lepturus*, the southern South China Sea, and the broader Indian Ocean, which limits our ability to fully assess the global phylogeographic pattern of the *T. lepturus* complex. Fifth, our demographic inference is based on a single high-coverage individual per species and a fixed mutation rate and generation time, which introduces uncertainty in the absolute timing of demographic events; however, bootstrap validation supported the relative population size trajectory and bottleneck signal. Sixth, our sampling was conducted within a limited seasonal window, which does not capture potential temporal population dynamics of *T. japonicus*. Nevertheless, the four major evolutionary lineages (*T. japonicus*, *T. nanhaiensis*, *T. lepturus*, and the cryptic *T. sp.* lineage), are supported by genome-wide SNP data and show general congruence across multiple analytical frameworks (e.g., F_{ST} PCA, ADMIXTURE, and phylogenomic analysis) [55].

Beyond resolving taxonomy, this study demonstrates the decisive advantage of nuclear genomic data for understanding marine diversification. Genome-wide SNPs provide a biparentally inherited, multilocus signal [56], that can help disentangle contemporary gene flow from ancestral polymorphism [57], enabling accurate reconstructions of evolutionary history and revealing cryptic diversity [58–60].

Although the nuclear and mitochondrial phylogenies are largely congruent at the species level (Fig. 3), we observe localized discordance within *T. lepturus*, where the COX1-based tree exhibits a more admixed (Fig. 3B) and geographically inconsistent structure compared to the well-resolved nuclear phylogeny (Fig. 3A). Importantly, this discrepancy is restricted to shallow nodes and is not associated with conflicts in major species-level relationships. Such patterns are expected under incomplete lineage sorting (ILS) [61], particularly in lineages that have undergone recent divergence and retain large effective population sizes [14, 62, 63], as is likely for the widely distributed *T. lepturus*. In addition, the possibility of historical or ongoing introgression cannot be excluded, given the overlapping distributions and ecological continuity across the Northwest Pacific.

We emphasize that the mitochondrial inference is based on a single locus and is therefore inherently more susceptible to stochastic lineage sorting, selective sweeps, and introgression, whereas the nuclear phylogeny integrates signals from 4,434 independent loci and thus provides a more robust estimate of the species tree. The consistency of species-level relationships across both datasets further supports the reliability of the nuclear framework.

Taken together, we interpret the mitochondrial pattern not as definitive evidence of mito-nuclear discordance, but as an expected outcome of lineage complexity in a recently diversified and potentially reticulated system. Resolving this complexity will require broader population-level sampling and the application of coalescent-based approaches that explicitly model gene tree heterogeneity. Until such analyses are conducted, the mitochondrial signal should be viewed as provisional and hypothesis-generating rather than conclusive. Our phylogenomic, population structure, and demographic inference results converge to support a coherent evolutionary history of *Trichiurus* in the Northwest Pacific. The nuclear phylogeny resolved four distinct monophyletic lineages, which are further supported by clear species-level separation in PCA and F_{ST} analyses. MSMC-IM results revealed that the major interspecific divergence events occurred immediately after the genus-wide population bottleneck at ~474,010 generations ago, consistent with allopatric speciation driven by glacial vicariance. The only exception is the very recent divergence between *T. nanhaiensis* and *T. sp.*, which is accompanied by persistent gene flow, as evidenced by D-statistic and MSMC-IM migration rate results. The apparent mito-nuclear incongruence within the *T. lepturus* complex may reflect mitochondrial introgression, incomplete lineage sorting, or limited resolution of the current nuclear sampling, and should therefore be interpreted cautiously until denser geographic and genomic sampling becomes available.

Our findings have direct, actionable implications for fisheries science and management. The low genetic differentiation observed in *T. japonicus* across the sampled localities suggests that this taxon may be managed as a weakly differentiated regional fishery stock, pending broader temporal and geographic validation. Conversely, the recognition of *T. nanhaiensis* and *T. sp.* as distinct evolutionary and likely ecological units necessitates the development of species-specific assessment frameworks and trade regulations to prevent overexploitation and mislabeling, thereby aligning conservation policy with biological reality.

In summary, this integrated phylogenomic study establishes a resolved framework for *Trichiurus* in the Northwest Pacific and provides a concrete model of how Pleistocene sea-level fluctuations catalyzed marine speciation. Looking forward, we propose three priorities: (1) expanded sampling of *T. brevis* and under-represented regions to refine species boundaries and distributions; (2) the integration of high-throughput morphometric or otolith geochemical analyses with genomic data to develop rapid, diagnostic tools for fisheries monitoring; and (3) a broader biogeographic study of the *T. lepturus* complex across the Indo-Pacific to elucidate global diversification drivers. Collectively, our work underscores the power of phylogenomics to unveil hidden marine biodiversity and provides the essential scientific foundation for evidence-based management of these ecologically and economically vital fisheries.

Conclusion

This study presents a chromosome-level genome assembly for *T. japonicus*. Analysis of population resequencing data from China and Vietnam reveals low genetic differentiation across the Chinese coast, suggesting a weakly differentiated stock within the present sampling frame. In contrast, genetic divergence supports *T. japonicus*, *T. lepturus*, *T. nanhaiensis*, and a putative Indian Ocean lineage (*T. sp.*) as distinct species. Multiple analytical approaches suggest a recent divergence (~10,000 years ago) between *T. sp.* and *T. nanhaiensis*, accompanied by substantial ongoing gene flow. These findings help clarify species boundaries within *Trichiurus* and provide a scientific basis for improved fisheries management.

Materials and methods

Specimen collection and sampling

A single specimen of *T. japonicus* was captured by fishing rod-at Gouqi Island, China (30°N, 122°E) on 23 October 2023 (Fig. 1A). Following capture, tissues from ten distinct organs were dissected and preserved for downstream genomic analyses [64]. The liver was designated for long-read Nanopore sequencing, and skeletal muscle was reserved for Hi-C library construction [65].

Additional samples from the gill, heart, kidney, spleen, brain, gas bladder, skin, and blood were collected for subsequent transcriptomic profiling [64]. For population-level morphological and molecular analyses, an additional 439 ethanol-preserved specimens from the genus *Trichiurus* were examined [3]. These specimens originated from coastal waters of Vietnam ($n=151$), China ($n=223$), and South Korea ($n=65$).

DNA extraction and genome sequencing

Genomic DNA was isolated from tissue samples using a standard cetyltrimethylammonium bromide (CTAB) protocol [66], followed by purification with a commercial QIAGEN kit [67]. Long-read data were generated using Oxford Nanopore sequencing technology to aid in genome assembly [68]. To support assembly polishing and provide deep coverage, short-read data were also generated from the same DNA sample using the Illumina HiSeq platform [69].

Mitochondrial DNA amplification and species identification

To confirm species identity and assess phylogenetic relationships, the cytochrome c oxidase subunit I (COXI) gene segment of the mitochondrial genome was amplified and Sanger sequenced from 81 representative specimens, using universal primers (Forward: 5'-CTACAACCCACCGCTTACTC-3'; Reverse: 5'-GTTGTAATAAGTTAATGGCGC-3'). This analysis confirmed the presence of *T. lepturus*, *T. nanhaiensis*, *T. japonicus* (from Zhoushan, Xiamen, Qingdao, and Hainan, China), and an undescribed *T. sp.* within the collection from Vietnam, China, and Korea. For a broader phylogenetic context, all publicly available COXI sequences for the genus *Trichiurus* were retrieved from the NCBI database. This comparative dataset included sequences annotated as *T. lepturus* (e.g., KM538620, MT455315), putative "*T. lepturus*" or *T. sp.* (e.g., LC085645, MZ272011), and *Trichiurus* sp2 (e.g., FM998058, LC269196).

Whole-genome resequencing and sample composition

To investigate genomic variation across the genus, shallow whole-genome resequencing was performed on 56 individuals. Genomic DNA libraries were prepared according to an inline-indexing protocol following ultrasonic shearing to a fragment size range of 250–1000 base pairs [70]. All libraries were sequenced as 150 bp paired-end reads on an Illumina HiSeq2500 platform, achieving an average genomic coverage of approximately 8× per individual [71]. Per-sample read numbers and mapping statistics are provided in Supplementary Tables 3 to allow independent verification of coverage estimates. The resequenced panel included one individual of *Tenitoriceps cristatus* from Hainan, China, and eight individuals of *Lepturacanthus savala* from Xiamen, China

as outgroups. The remaining 47 individuals comprised multiple *Trichiurus* species: *T. lepturus* ($n=11$ from Vietnam), *T. sp.* ($n=6$ from South Korea and China), *T. nanhaiensis* ($n=9$ from Vietnam and China), *T. brevis* ($n=1$ from China), and *T. japonicus* ($n=20$ from China and South Korea).

Genome assembly and scaffolding

The genome size was estimated using GenomeScope 2.0 by analyzing 21-mer frequency profiles generated with Jellyfish (v2.3.0) [71]. A de novo assembly was constructed from Nanopore long-read data using NextDenovo with non-default parameters: read_cutoff=1k, seed_cutoff=25k. The same individual used for Nanopore long-read sequencing also provided the DNA for Illumina short-read polishing and Hi-C library construction. The initial assembly was subsequently polished in a two-stage process; the first polishing step was performed using Illumina short-read data with NextPolish (v1.4.1) with non-default parameters: sgs_options = -max_depth 100, to correct sequencing errors.

Hi-C-based chromosome construction

To order and orient contigs into chromosome-scale scaffolds, Hi-C scaffolding was performed [72]. The Hi-C sequencing data were processed and aligned to the polished assembly using HiC-Pro (v3.1.0) [73] and Bowtie2 (v2.4.5) [74]. Chromosomal-level scaffolding was then executed using Juicer (v1.6.2) and 3D-DNA (v180922) [74]. The resulting scaffolds were visually inspected and manually curated in Juicebox (v1.11.08) [75].

Assembly quality

The completeness of the assembled genome was assessed using BUSCO (v5.4.3) [76] against the actinopterygii_odb10 database [77].

Repeat annotation and structural gene prediction

Repetitive elements within the genome were identified and characterized using a combination of RepeatModeler v2.0.3 [78] for de novo repeat library construction and RepeatMasker v4.1.2-pl [79] for genome-wide annotation.

Genome structural annotation was performed using MAKER3 v3.01.03 [80]. A cleaned protein database was first prepared from five closely related teleost species: *Danio rerio*, *Perca flavescens*, *Scomber japonicus*, *Thunnus albacares*, and *Thunnus maccoyii*. Transcriptomic evidence was generated by aligning RNA-Seq reads and assembling them into a deduplicated transcriptome using Trinity and CD-HIT-EST [81]. The MAKER pipeline was executed iteratively in three rounds: an initial annotation run; a second round incorporating ab initio gene predictions from Augustus v3.4.0 and SNAP, which were

trained using the outputs of the first round; and a final third round integrating all available transcriptomic, protein homology, and ab initio evidence [82]. Post-processing of the resulting GFF file included filtering, extraction of coding sequences and protein sequences, and functional annotation. Evidence-based gene prioritization was performed using InterProScan v5.56-89.0 [83].

Comparative genomics

For evolutionary context, a comparative genomic analysis focused on synteny was performed [84]. Genome synteny between *Scomber japonicus* and *Thunnus albacare* was investigated using JCVI (v1.1.14) [85], and the results were visualized using pyCircos [86].

Resequencing data processing and variant calling

Raw sequencing reads were demultiplexed using unique dual-index identifiers, which comprised 8-bp i7/i5 adapters paired with 6-bp inline barcodes [87]. Demultiplexed reads were then processed for adapter removal and quality trimming [88] using TrimGalore (v0.6.7) [89] and Cutadapt (v1.2.1) [90], with a quality threshold of Phred score < 20 [91]. Variant calling was performed with GATK (v4.2.0.0) to identify single nucleotide polymorphisms (SNPs) [92]. The resulting variant calls were filtered with the following thresholds: QUAL < 30, QD < 2, FS > 60, and SOR > 4. We describe these variants as a filtered SNP dataset rather than as definitively high-quality SNPs, because low and uneven sequencing depth can affect heterozygote recovery and genotype accuracy [93].

Population structure analysis

The filtered SNP dataset was used to investigate population structure [94]. Principal component analysis was conducted using PLINK (v1.9) [95]. Ancestry proportions and population clustering were inferred with ADMIXTURE (v1.3.0) [96], testing a range of hypothetical ancestral populations from K = 1 to K = 7 [97].

For downstream analyses, the filtered SNP dataset was further processed with analysis-specific thresholds: For PCA analysis: 1,875,043 SNPs were retained with a maximum missing data rate of 20% per site (--geno 0.2) and no minor allele frequency (MAF) filtering, to maximize interspecific variation information; For ADMIXTURE analysis: 120,445 filtered SNPs were retained after filtering with a MAF threshold of 0.05, depth and missingness filtering with minimum depth (minDP) = 2, maximum depth (maxDP) = 80, and a maximum missing data rate of 30% per site (--geno 0.3); For D-statistic analysis: 793,598 SNPs were retained after applying MAF > 0.01, removing sites deviating from Hardy-Weinberg equilibrium at $P < 1 \times 10^{-6}$, filtering sites with missing genotype rate ≥ 0.8 , and pruning linkage disequilibrium with --indep-pairwise 50 10 0.2. For F_{ST} , Dxy, and π analyses, the same

793,598-SNP dataset was used for 50-kb window-based divergence and diversity estimates. For MSMC-IM, 834,650 SNPs were retained following the standard filtering protocol of the MSMC-IM pipeline (Supplementary Table 4).

Demographic history inference

The demographic histories of *T. japonicus*, *T. brevis*, *T. nanhaiensis*, and *T. sp.* were reconstructed using MSMC2 and MSMC-IM to model changes in effective population size and migration over time [98]. For this analysis, the individual from each of the four species with the deepest sequencing coverage (> 20×) was selected [99]. Data phasing was performed with WhatsHap (v2.8) prior to analysis [100]. Two primary comparison groups were analyzed jointly: *T. japonicus* versus the other three species, and *T. sp.* versus *T. nanhaiensis*. Analyses followed the standard MSMC2 and MSMC-IM protocols, using an average mutation rate (μ) of 5.97×10^{-9} per base per generation, and a standard generation time of 2.9 years for *Trichiurus* obtained from the authoritative FishBase database. Bootstrap validation was performed using MSMC-Tools, with 50 bootstrap datasets generated from per-chromosome data, and MSMC2 and MSMC-IM analyses were re-run for each bootstrap replicate to assess the robustness of the demographic inference.

Phylogenetic analysis of orthologous genes

To reconstruct a species phylogeny, single-copy orthologous genes were identified and analyzed [101]. A modified version of the Assexon pipeline [102], adapted for whole-genome resequencing data, was employed. Reads were aligned to a conserved set of 4,434 single-copy orthologs derived from the *Oreochromis niloticus* reference genome [103]. Consensus sequences were assembled using UBLAST and SGA, and exon boundaries were precisely refined with Exonerate [102]. A maximum likelihood phylogeny was reconstructed from the concatenated alignment using RAxML (v8.2.12) [104] under a GTR+GAMMA model of sequence evolution, with branch support assessed from 1,000 bootstrap replicates [105].

Admixture and population divergence analysis

For tests of admixture and window-based divergence, the SNP dataset was further filtered using PLINK [105] by retaining variants with minor allele frequency (MAF) > 0.01 [105], removing sites deviating from Hardy-Weinberg equilibrium at $P < 1 \times 10^{-6}$ [106], and excluding sites with missing genotype rate (≥ 0.8) [107], resulting in 793,598 filtered SNPs. Linkage disequilibrium was pruned with the parameter --indep-pairwise 50 10 0.2. The D-statistic (ABBA-BABA test) was then calculated using the qpDstat program from the Admixtools

package [108], with *Lepturacanthus savala* designated as the outgroup. Each test involved a trio of ingroup populations selected from *Trichiurus lepturus*, *T. japonicus*, *T. nanhaiensis*, and *T. sp.* Finally, to quantify genome-wide patterns of divergence and diversity, pairwise genetic differentiation (D_{xy} , F_{ST}) and nucleotide diversity (π) were calculated from this 793, 598 – SMP dataset in 50-kb non-overlapping windows across all chromosomes using pixy [108]. These results were summarized and visualized using the CMplot R package [109].

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-026-13009-x>.

Supplementary Material 1. Supplementary Table 1 Summary of genome assembly of *Trichiurus japonicus*. Supplementary Table 2 Summary of hairtails specimens. Supplementary Table 3 Summary of mappings. Supplementary Table 4 Summary of SNP filtering parameters and datasets used for each analysis.

Supplementary Material 2. Supplementary Fig. 1 A. Synteny plot showing the genomic collinearity between the *Trichiurus japonicus*, *Scomber japonicus* and *Thunnus albacares*. B. Genomic features of the *T. japonicus*. From top clockwise are Chr1–Chr24 and from the inner to the outer layer: genome-wide synteny, GC content heatmap, gene density distribution, LTR transposon distribution, and TE transposon distribution, from 0–1. Supplementary Fig. 2 The π values of the four species, along with the D_{xy} and F_{ST} values between every pair of species or between each species and the presumed populations of *T. japonicus*. Supplementary Fig. 3 Analysis of four species using MSMC2 and MSMC-IM: Each column represents a pair of species included in the analysis, and each row sequentially lists population size, migration rates, coalescence rate, and relative cross-coalescence rate. Supplementary Fig. 4 Bootstrap result of MSMC-IM. Supplementary Fig. 5 The synteny of this *T. japonicus* genome and another genome assembled by Hsu et al.

Acknowledgements

We extend special thanks to Prof. Hanlin Wu for encouraging the research team to address taxonomic complexities in *Trichiurus*. We also acknowledge Dr. Shaoxiong Ding, Dr. Hui Wang, Mr. Jiajie Chen, and Ms. Ni Xin for their assistance with specimen collection.

Authors' contributions

V.G., T.W., and C.L. conceived the research plan. V.G. collected the samples. T.W. and V.G. performed experiments and collected data. T.W. analyzed the data and made the figures. L.L., J.H., and M.Z. contributed to data analysis and participated in experimental work. V.G., T.W., and C.L. wrote the draft. All authors edited and approved the final version of the manuscript.

Funding

This work was supported by the Capacity building project of local colleges and universities of the Shanghai Science and Technology Commission (No. 23010502500).

Data availability

The chromosome-level genome assembly of *Trichiurus japonicus* has been deposited in the NCBI genome database under accession GCA_046254865.1 (BioSample: SAMN44500129). The Chromium-linked Illumina reads, RNA-Seq reads, and Nanopore reads used for the genome assembly have been deposited under the BioProject accession number PRJNA1179659. The resequencing data and other supporting data have been deposited in the Dryad repository: https://datadryad.org/share/GpQ6KcJfWqSEUvww62AUhzWxb-axl_ZrA9TvkG45Us. The scripts and commands used for analysis are available at GitHub: <https://github.com/wutianqin123/TrichiurusGenome>.

Declarations

Ethics approval and consent to participate

Not applicable. The fish samples used in this study were obtained from commercial catches or scientific surveys, and no live animal experiments were conducted.

Consent for publication

All authors have read and approved the final manuscript and consent to its publication.

Competing interests

The authors declare no competing interests.

Received: 30 December 2025 / Accepted: 29 May 2026

Published online: 04 June 2026

References

- Imagawa T, Horie R, Matsumoto M, Tachibanagi K, Tanaka R. Seasonal changes in the body size and nutritional components of *Trichiurus* sp. 2 and *Trichiurus japonicus* from the Miyazaki coast in Japan. *J Food Meas Charact*. 2025;1–16. <https://doi.org/10.1007/s11694-025-03617-8>.
- Jeong YD, Kim S, Jung IH, Cho G. Optimal harvesting strategy for hairtail, *Trichiurus lepturus*, in Korea Sea using discrete-time age-structured model. *Appl Math Comput*. 2021;392:125743. <https://doi.org/10.1016/j.amc.2020.125743>.
- Yi M-R, Hsu K-C, Gu S, He X-B, Luo Z-S, Lin H-D, Yan Y-R. Complete mitogenomes of four *Trichiurus* species: A taxonomic review of the *T. lepturus* species complex. *ZooKeys*. 2022;10841. <https://doi.org/10.3897/zookeys.10841.71576>.
- Lin H-C, Tsai C-J, Wang H-Y. Variation in global distribution, population structures, and demographic history for four *Trichiurus* cutlassfishes. *PeerJ*. 2021;9:e12639. <https://doi.org/10.7717/peerj.12639>.
- Teletchea F. Molecular identification methods of fish species: reassessment and possible applications. *Rev Fish Biol Fish*. 2009;19(3):265–93. <https://doi.org/10.1007/s11160-009-9107-4>.
- Healey AJ, McKeown NJ, Taylor AL, Provan J, Sauer W, Gouws G, Shaw PW. Cryptic species and parallel genetic structuring in Lethrinid fish: Implications for conservation and management in the southwest Indian Ocean. *Ecol Evol*. 2018;8(4):2182–95. <https://doi.org/10.1002/ece3.3775>.
- Neves JM, Nolen ZJ, Fabr   NN, Mott T, Pereira RJ. Genomic methods reveal independent demographic histories despite strong morphological conservatism in fish species. *Heredity*. 2021;127(3):323–33. <https://doi.org/10.1038/s41437-021-00455-4>.
- Tzeng C, Chen C, Chiu T. Analysis of morphometry and mitochondrial DNA sequences from two *Trichiurus* species in waters of the western North Pacific: taxonomic assessment and population structure. *J Fish Biol*. 2007;70:165–76. <https://doi.org/10.1111/j.1095-8649.2007.01368.x>.
- Linnaeus C, Gmelin JF, editor. *Systema Naturae per regna tria naturae, secundum classes, ordines, genera, species; cum characteribus, differentiis, synonymis, locis. Editio decima tertia, aucta, reformata. Vol. 1. Lugduni [Lyon]: apud J.B. Delamolliere; 1789.* <https://doi.org/10.5962/bhl.title.36932>.
- Hsu K-C, Shih N-T, Ni I-H, Shao K-T. Genetic variation in *Trichiurus lepturus* (Perciformes: Trichiuridae) in waters off taiwan: several species or cohort contribution. *Raffles Bull Zool*. 2007;14:215–20. <https://doi.org/10.1590/s1519-69842011000400024>.
- He L, Zhang A, Weese D, Li S, Li J, Zhang J. Demographic response of cutlassfish (*Trichiurus japonicus* and *T. nanhaiensis*) to fluctuating palaeoclimate and regional oceanographic conditions in the China seas. *Sci Rep*. 2014;4(1):6380. <https://doi.org/10.1038/srep06380>.
- Chakraborty A, Iwatsuki Y. Genetic variation at the mitochondrial 16S rRNA gene among *Trichiurus lepturus* (Teleostei: Trichiuridae) from various localities: preliminary evidence of a new species from West coast of Africa. *Hydrobiologia*. 2006;563(1):501–13. <https://doi.org/10.1007/s10750-006-0105-4>.
- Wang H-Y, Dong CA, Lin H-C. DNA barcoding of fisheries catch to reveal composition and distribution of cutlassfishes along the Taiwan coast. *Fish Res*. 2017;187:103–9. <https://doi.org/10.1016/j.fishres.2016.11.015>.
- DeRaad DA, McCullough JM, DeCicco LH, Hime PM, Joseph L, Andersen MJ, Moyle RG. Mitonuclear discordance results from incomplete lineage sorting, with no detectable evidence for gene flow, in a rapid radiation of Todiramphus kingfishers. *Mol Ecol*. 2023;32(17):4844–62.

15. Ballard JWO, Whitlock MC. The incomplete natural history of mitochondria. *Mol Ecol*. 2004;13(4):729–44. <https://doi.org/10.1046/j.1365-294x.2003.02063.x>.
16. Burhanuddin AI, Iwatsuki Y, Yoshino T, Kimura S. Small and valid species of *Trichiurus brevis* Wang and You, 1992 and *T. russelli* Dutt and Thankam, 1966, defined as the *T. russelli* complex (Perciformes: Trichiuridae). *Ichthyol Res*. 2002;49(3):211–23. <https://doi.org/10.1007/s10228-003-0168-4>.
17. Wheeler A. The Linnaean fish collection in the Linnean Society of London. *Zool J Linn Soc*. 1985;84(1):1–76.
18. Wang K. Studies on the genetic variation and systematic of the hairtail fishes from the South China Sea. *Mar Sci Acad Sinica* (Beijing). 1992;2:69–72.
19. Wang K. Hairtail fishes from Chinese coastal waters (Trichiuridae). *Oceanol Limnol Sin*. 1995;26:215–22.
20. Tokimura M. Comments on taxonomy and distributions of the species of *Trichiurus* in the East China and Yellow Seas and adjacent waters. *Seikaiku Suisan Kenkyusho News*. 1995;80:12–4.
21. Tzeng C-H, Chiu T-S. DNA barcode-based identification of commercially caught cutlassfishes (Family: Trichiuridae) with a phylogenetic assessment. *Fish Res*. 2012;127:176–81. <https://doi.org/10.1016/j.fishres.2012.01.022>.
22. Toews DP, Brelsford A. The biogeography of mitochondrial and nuclear discordance in animals. *Mol Ecol*. 2012;21(16):3907–30. <https://doi.org/10.1111/1/j.1365-294x.2012.05664.x>.
23. Pedraza-Marrón CR, Silva R, Deeds J, Van Belleghem SM, Mastretta-Yanes A, Domínguez-Domínguez O, Rivero-Vega RA, Lutackas L, Murie D, Parkyn D. Genomics overrules mitochondrial DNA, siding with morphology on a controversial case of species delimitation. *Proceedings of the Royal Society B* 2019;286(1900):20182924. <https://doi.org/10.1098/rspb.2018.2924>.
24. Marshall TL, Chambers EA, Matz MV, Hillis DM. How mitochondrial discordance and geographic variation have confounded species boundaries in a widely studied snake. *Mol Phylogenet Evol*. 2021;162:107194. <https://doi.org/10.1016/j.ympev.2021.107194>.
25. Després L. One, two or more species? Mitonuclear discordance and species delimitation. : Wiley Online Libr. 2019. <https://doi.org/10.1111/mec.15211>.
26. Song N, Zhao X, Cai C, Gao T. Profile of the genomic characteristics and comparative studies of five Trichiuridae species by genome survey sequencing. *Front Mar Sci*. 2022;9:962307. <https://doi.org/10.3389/fmars.2022.962307>.
27. Hsu K-C, Shih N-T, Ni H-H, Shao K-T. Genetic diversity and population structure of largehead hairtail, *Trichiurus japonicus*, based on mtDNA control region. *Zool Stud*. 2009;48(6):835–49. <https://doi.org/10.3109/19401736.2013.809427>.
28. Lee SJ, Kim J-K. Identification of *Trichiurus* (Pisces: Trichiuridae) eggs and larvae from Korea, with a taxonomic note. *Fisheries Aquat Sci*. 2014;17(1):137–43. <https://doi.org/10.5657/fas.2014.0137>.
29. Chakraborty A, Aranishi F, Iwatsuki Y. Genetic differences among three species of the genus *Trichiurus* (Perciformes: Trichiuridae) based on mitochondrial DNA analysis. *Ichthyol Res*. 2006;53(1):93–6. <https://doi.org/10.1007/s10228-005-0313-3>.
30. Wu R, Guo L, Liu J. Genetic variation of *Trichiurus lepturus* and phylogenetic relationship among its closed species. *Acta Zootaxonomica Sin*. 2011;36:648–55. <https://doi.org/10.1007/s10750-006-0105-4>.
31. Okamoto N, Koike K, Shibata J-Y, Tomiyama T. Species composition of hairtails (Trichiuridae) in Myanmar. *Reg Stud Mar Sci*. 2018;17:73–7. <https://doi.org/10.1016/j.rsma.2017.11.007>.
32. Lee S, Tsoi S, Chao W. Biochemical systematics of *Trichiurus lepturus* and *T. japonicus* (Perciformes, Trichiuridae) from Taiwan Strait. *J Zoological Syst Evolutionary Res*. 1993;31(3):227–32. <https://doi.org/10.1111/j.1439-0469.1993.tb00191.x>.
33. Tzeng Chih-Hsiang TC, Chiu Tsan-Yu CT, Chen Chih-Shin CC, Wang Hui-Yu WH, Chiu Tai-Sheng CT. The current population structure of the demersal hairtail (*Trichiurus japonicus*) in the western North Pacific was shaped by the Taiwan Strait, as revealed by mitochondrial DNA. 2016. <https://doi.org/10.3109/19401736.2013.809427>.
34. Hsu P-C, Lin C-Y, Hsieh P-H, Chuang W-H, Lu M-Y, Chen CA, Chen S-H. Chromosome-Level Genome Assembly and Annotation of the Japanese Cutlassfish (*Trichiurus japonicus*): A High-Quality Genomic Resource Featuring Nuclear and Mitochondrial Completeness for Future Studies. *Sci Data*. 2025;12(1):1860.
35. Mukundan LP, Sukumaran S, Sebastian W, Gopalakrishnan A. Characterization of the whole mitogenome of largehead hairtail *Trichiurus lepturus* (Trichiuridae): Insights into special characteristics. *Biochem Genet*. 2020;58(3):430–51. <https://doi.org/10.1007/s10528-020-09956-z>.
36. Temminck CJ, Schlegel H. Pisces. In: Siebold P. de, editor. *Fauna Japonica, sive descriptio animalium, quae in itinere per Japoniam ... suscepto annis 1823-1830 collegit, notis, observationibus et adumbrationibus illustravit Ph. Fr. de Siebold. Lugduni Batavorum [Leiden] (A. Arnz et soc.). Parts 5-6: 73-112. 1844. https://doi.org/10.5962/bhl.title.124951.*
37. Forsskål P. Descriptiones animalium, avium, amphibiorum, piscium, insectorum, vermium; quae in itinere orientali observavit Petrus Forsskål. Post mortem auctoris edidit Carsten Niebuhr. Hauniae [Copenhagen]: ex officina Möller; 1775. <https://doi.org/10.5962/bhl.title.2154>.
38. Chakraborty A, Aranishi F, Iwatsuki Y. Genetic differentiation of *Trichiurus japonicus* and *T. lepturus* (Perciformes: Trichiuridae) based on mitochondrial DNA analysis. *ZOOLOGICAL STUDIES-TAIPEI*. 2006;45(3):419. <https://doi.org/10.1007/s10228-005-0313-3>.
39. Talaga S, le Goff G, Arana-Guardia R, Baak-Baak CM, García-Rejón JE, García-Suárez O, Rodríguez-Valencia VM, Tolsá-García MJ, Suzán G, Roiz D. The mosquitoes (Diptera: Culicidae) of the Mexican Yucatan Peninsula: a comprehensive review on the use of taxonomic names. *J Med Entomol*. 2023;61(2):274–308. <https://doi.org/10.1093/jme/tjad168>.
40. Nekola JC, Horsák M. The impact of empirically unverified taxonomic concepts on ecological assemblage patterns across multiple spatial scales. *Ecography*. 2022;2022(5):e06063. <https://doi.org/10.1111/ecog.06063>.
41. SANTOS CMD, AMORIM DS, FACHIN KLASSAB, NIHEI SS DA, DE CARVALHO CJB, FALASCHI RL, MELLO-PATIU CA, COURI MS, OLIVEIRA SS. On typeless species and the perils of fast taxonomy. *Syst Entomol*. 2016;41(3):511–5. <https://doi.org/10.1111/syen.12180>.
42. Stawitz CC, Siple MC, Munsch SH, Lee Q, Derby SR. Financial and ecological implications of global seafood mislabeling. *Conserv Lett*. 2017;10(6):681–9. <https://doi.org/10.1111/cons.12328>.
43. Cawthorn DM, Baillie C, Mariani S. Generic names and mislabeling conceal high species diversity in global fisheries markets. *Conserv Lett*. 2018;11(5):e12573. <https://doi.org/10.1111/cons.12573>.
44. Lewis SG, Boyle M. The expanding role of traceability in seafood: tools and key initiatives. *J Food Sci*. 2017;82(51):A13–A. <https://doi.org/10.1111/1750-3841.13743>.
45. Kotsanopoulos KV, Exadactylos A, Gkafas GA, Martsikalis PV, Parlapani FF, Bozaris IS, Arvanitoyannis IS. The use of molecular markers in the verification of fish and seafood authenticity and the detection of adulteration. *Compr Rev Food Sci Food Saf*. 2021;20(2):1584–654. <https://doi.org/10.1111/1541-4337.12719>.
46. He Z, Li X, Yang M, Wang X, Zhong C, Duke NC, Wu C-I, Shi S. Speciation with gene flow via cycles of isolation and migration: insights from multiple mangrove taxa. *Natl Sci Rev*. 2018;6(2):275–88. <https://doi.org/10.1101/335166>.
47. April J, Hanner RH, Dion-Côté AM, Bernatchez L. Glacial cycles as an allopatric speciation pump in north-eastern American freshwater fishes. *Mol Ecol*. 2013;22(2):409–22. <https://doi.org/10.1111/mec.12116>.
48. Nugent CM, Kess T, Langille BL, Beck SV, Duffy S, Messmer A, Smith N, Lehnert SJ, Wringe BF, Kent M. Post-glacial recolonization and multiple scales of secondary contact contribute to contemporary Atlantic salmon (*Salmo salar*) genomic variation in North America. *J Biogeogr*. 2024;51(9):1767–82. <https://doi.org/10.1111/jbi.14852>.
49. Schield DR, Adams RH, Card DC, Perry BW, Pasquesi GM, Jezkova T, Portik DM, Andrew AL, Spencer CL, Sanchez EE. Insight into the roles of selection in speciation from genomic patterns of divergence and introgression in secondary contact in venomous rattlesnakes. *Ecol Evol*. 2017;7(11):3951–66. <https://doi.org/10.1002/ece3.2996>.
50. Qu Y, Chen C, Xiong Y, She H, Zhang YE, Cheng Y, DuBay S, Li D, Ericson PGP, Hao Y, et al. Rapid phenotypic evolution with shallow genomic differentiation during early stages of high elevation adaptation in Eurasian Tree Sparrows. *Natl Sci Rev*. 2019;7(1):113–27. <https://doi.org/10.1093/nsr/nwz138>.
51. Nadachowska-Brzyska K, Burri R, Olason PI, Kawakami T, Smeds L, Ellegren H. Demographic divergence history of pied flycatcher and collared flycatcher inferred from whole-genome re-sequencing data. *PLoS Genet*. 2013;9(11):e1003942. <https://doi.org/10.1371/journal.pgen.1003942>.
52. Eliášová K, Lucas Lledó JI, Grau JH, Loudová M, Bannikova AA, Zolotareva KI, Beneš V, Hulva P, Černá Bolíková B. Contrasting levels of hybridization across the two contact zones between two hedgehog species revealed by genome-wide SNP data. *Heredity*. 2022;129(5):305–15. <https://doi.org/10.1038/s41437-022-00567-5>.
53. Villanea FA, Schraiber JG. Multiple episodes of interbreeding between Neanderthal and modern humans. *Nat Ecol Evol*. 2019;3(1):39–44. <https://doi.org/10.1038/s41559-018-0735-8>.
54. Lan T, Leppälä K, Tomlin C, Talbot SL, Sage GK, Farley SD, Shideler RT, Bachmann L, Wiig Ø, Albert VA et al. Insights into bear evolution from a Pleistocene polar bear genome. *Proceedings of the National Academy of Sciences* 2022;119(24):e2200016119. <https://doi.org/10.1073/pnas.2200016119>.

55. Khan MMH, Rafii MY, Ramlee SI, Jusoh M, Al Mamun M, Halidu J. DNA fingerprinting, fixation-index (Fst), and admixture mapping of selected Bambara groundnut (*Vigna subterranea* [L.] Verdc.) accessions using ISSR markers system. *Sci Rep*. 2021;11(1):14527. <https://doi.org/10.1038/s41598-021-93867-5>.
56. Brumfield RT, Beerli P, Nickerson DA, Edwards SV. The utility of single nucleotide polymorphisms in inferences of population history. *Trends Ecol Evol*. 2003;18(5):249–56. [https://doi.org/10.1016/s0169-5347\(03\)00018-1](https://doi.org/10.1016/s0169-5347(03)00018-1).
57. WILLING EM, Bentzen P, Van Oosterhout C, Hoffmann M, Cable J, Breden F, Weigel D, Dreyer C. Genome-wide single nucleotide polymorphisms reveal population history and adaptive divergence in wild guppies. *Mol Ecol*. 2010;19(5):968–84. <https://doi.org/10.3410/f.2478973.2118073>.
58. Wang B, Mojica JP, Perera N, Lee C-R, Lovell JT, Sharma A, Adam C, Lipzen A, Barry K, Rokhsar DS. Ancient polymorphisms contribute to genome-wide variation by long-term balancing selection and divergent sorting in *Boechera stricta*. *Genome Biol*. 2019;20(1):126. <https://doi.org/10.1186/s13059-019-1729-9>.
59. Chan KO, Hutter CR, Wood PL Jr, Su Y-C, Brown RM. Gene flow increases phylogenetic structure and inflates cryptic species estimations: a case study on widespread Philippine puddle frogs (*Occidozyga laevis*). *Syst Biol*. 2022;71(1):40–57. <https://doi.org/10.1093/sysbio/syab034>.
60. Chan KO, Alexander AM, Grismer LL, Su YC, Grismer JL, Quah ES, Brown RM. Species delimitation with gene flow: a methodological comparison and population genomics approach to elucidate cryptic species boundaries in Malaysian Torrent Frogs. *Mol Ecol*. 2017;26(20):5435–50. <https://doi.org/10.1111/mec.14296>.
61. Hendrich L, Pons J, Ribera I, Balke M. Mitochondrial cox1 sequence data reliably uncover patterns of insect diversity but suffer from high lineage-idiosyncratic error rates. *PLoS ONE*. 2010;5(12):e14448.
62. Klimov PB, Skoracki M, Bochkov AV. Cox 1 barcoding versus multilocus species delimitation: validation of two mite species with contrasting effective population sizes. *Parasites vectors*. 2019;12(1):8.
63. Wang W, Wang Y, Lei F, Liu Y, Wang H, Chen J. Incomplete lineage sorting and introgression in the diversification of Chinese spot-billed ducks and mallards. *Curr Zool*. 2019;65(5):589–97.
64. Yu Y, Zhao C, Su Z, Wang C, Fuscoe JC, Tong W, Shi L. Comprehensive RNA-Seq transcriptomic profiling across 11 organs, 4 ages, and 2 sexes of Fischer 344 rats. *Sci data*. 2014;1(1):1–8. <https://doi.org/10.1038/sdata.2014.13>.
65. Willemart C, Strazisar M, De Pooter T, Stroobants T, Demuyser T, Jorens P, Hoste E, Menschaert G, Vanden Bergh T. Nanopore sequencing enables tissue-of-origin and pathogen detection in plasma cell-free DNA from critically ill patients. *Cell Death Discovery*. 2025;11(1):484. <https://doi.org/10.1038/s41420-025-02828-8>.
66. Allen GC, Flores-Vergara M, Krasynanski S, Kumar S, Thompson W. A modified protocol for rapid DNA isolation from plant tissues using cetyltrimethylammonium bromide. *Nat Protoc*. 2006;1(5):2320–5. <https://doi.org/10.1038/nprot.2006.384>.
67. Zhou Y, Zhang Y, He W, Wang J, Peng F, Huang L, Zhao S, Deng W. Rapid regeneration and reuse of silica columns from PCR purification and gel extraction kits. *Sci Rep*. 2018;8(1):12870. <https://doi.org/10.1038/s41598-018-30316-w>.
68. Jain M, Koren S, Miga KH, Quick J, Rand AC, Sasani TA, Tyson JR, Beggs AD, Dilthey AT, Fiddes IT. Nanopore sequencing and assembly of a human genome with ultra-long reads. *Nat Biotechnol*. 2018;36(4):338–45. <https://doi.org/10.3410/f.732576831.793554178>.
69. Wick RR, Judd LM, Holt KE. Assembling the perfect bacterial genome using Oxford Nanopore and Illumina sequencing. *PLoS Comput Biol*. 2023;19(3):e1010905. <https://doi.org/10.1371/journal.pcbi.1010905>.
70. Head SR, Komori HK, LaMere SA, Whisenant T, Van Nieuwerburgh F, Salomon DR, Ordoukhanian P. Library construction for next-generation sequencing: overviews and challenges. *Biotechniques*. 2014;56(2):61–77. <https://doi.org/10.2144/000114133>.
71. de Muinck EJ, Trosvik P, Gilfillan GD, Hov JR, Sundaram AY. A novel ultra high-throughput 16S rRNA gene amplicon sequencing library preparation method for the Illumina HiSeq platform. *Microbiome*. 2017;5(1):68. <https://doi.org/10.1101/124057>.
72. Wang Y-B, Zhang X. Chromosome scaffolding of diploid genomes using ALLHiC. *Bio Protoc*. 2022;12(18):1–8. <https://doi.org/10.21769/bioprotoc.4503>.
73. Servant N, Varoquaux N, Lajoie BR, Viara E, Chen C-J, Vert J-P, Heard E, Dekker J, Barillot E. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol*. 2015;16(1):259. <https://doi.org/10.1186/s13059-015-0831-x>.
74. Wang Y, Wang M, Li J, Zhang J, Zhang L. A chromosome-level genome assembly of a deep-sea symbiotic Aplousobranchia mollusc *Chaetoderma* sp. *Sci Data*. 2024;11(1):133. <https://doi.org/10.1038/s41597-024-02940-x>.
75. Mansfeld BN, Yocca A, Ou S, Harkess A, Burchard E, Gutierrez B, van Nocker S, Gottschalk C. A haplotype resolved chromosome-scale assembly of North American wild apple *Malus fusca* and comparative genomics of the fire blight Mfu10 locus. *Plant J*. 2023;116(4):989–1002. <https://doi.org/10.1111/TPJ.16433>.
76. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*. 2015;31(19):3210–2. <https://doi.org/10.1093/bioinformatics/btv351>.
77. Cheng F, Dennis AB, Osuoha JI, Canitz J, Kirschbaum F, Tiedemann R. A new genome assembly of an African weakly electric fish (*Campylomormyrus compressirostris*, Mormyridae) indicates rapid gene family evolution in Osteoglossomorpha. *BMC Genomics*. 2023;24(1):129. <https://doi.org/10.1186/s12864-023-09196-6>.
78. Peterson JK, MacDonald ML, Ellis VA. First whole-genome sequence of *Triatoma sanguisuga* (Le Conte, 1855), vector of Chagas disease. *G3 Genes Genomes Genetics*. 2025;15(3). <https://doi.org/10.1093/g3journal/jkae308>.
79. Hanrahan BJ, Alreja K, Reis AL, Chang JK, Dissanayake DS, Edwards RJ, Bertozzi T, Hammond JM, O'Meally D, Deveson IW. A genome assembly and annotation for the Australian alpine skink *Bassiana duperreyi* using long-read technologies. *G3: Genes Genomes Genet*. 2025;15(6):jka046.
80. Chen T, Yang Y, Wang X, Qin Z, Xie Z, Fan D, Ren C, Sun H, Luo P, Jiang X. Chromosome-level genome assembly and annotation of the tropical sea cucumber *Stichopus monotuberculatus*. *Sci Data*. 2024;11(1):1245.
81. Chen L-Y, Morales-Briones DF, Passow CN, Yang Y. Performance of gene expression analyses using de novo assembled transcripts in polyploid species. *Bioinformatics*. 2019;35(21):4314–20.
82. Meunier L, Baurain D, Cornet L. AMAW: automated gene annotation for non-model eukaryotic genomes. *bioRxiv* 2021;2021.2012.2007.471566.
83. Mifsud JC, Lytras S, Oliver MR, Toon K, Costa VA, Holmes EC, Grove J. Mapping glycoprotein structure reveals Flaviviridae evolutionary history. *Nature*. 2024;633(8030):695–703.
84. Liu D, Hunt M, Tsai JJ. Inferring synteny between genome assemblies: a systematic evaluation. *BMC Bioinformatics*. 2018;19(1):26. <https://doi.org/10.1011/149989>.
85. Chen M, Yang D, Yang S, Yang X, Chen Z, Yang T, Yang Y, Yang Y. Chromosome-level genome assembly of *Hippophae gyanensis*. *Sci Data*. 2024;11(1):126. <https://doi.org/10.1038/s41597-024-02909-w>.
86. Green AG, Vargas R Jr, Marin MG, Freschi L, Xie J, Farhat MR. Analysis of genome-wide mutational dependence in naturally evolving *Mycobacterium tuberculosis* populations. *Mol Biol Evol*. 2023;40(6):msad131. <https://doi.org/10.1093/molbev/msad131>.
87. MacConaill LE, Burns RT, Nag A, Coleman HA, Slevin MK, Giorda K, Light M, Lai K, Jarosz M, McNeill MS. Unique, dual-indexed sequencing adapters with UMIs effectively eliminate index cross-talk and significantly improve sensitivity of massively parallel sequencing. *BMC Genomics*. 2018;19(1):30. <https://doi.org/10.1186/s12864-017-4428-5>.
88. Lindgreen S. AdapterRemoval: easy cleaning of next-generation sequencing reads. *BMC Res Notes*. 2012;5(1):337. <https://doi.org/10.1186/1756-0500-5-337>.
89. Krueger F, James F, Ewels P, Afyounian E, Schuster-Boeckler B. FelixKrueger/TrimGalore: v0. 6.7-doi via zenodo. *zenodo* 2021. <https://doi.org/10.5281/zenodo.5127899>.
90. Chen Y, Feng L, Lin H, Liu J, Hu Q. Chromosome-level genome assembly of *Helwingia omeiensis*: the first genome in the family Helwingiaceae. *Sci Data*. 2024;11(1):719. <https://doi.org/10.1038/s41597-024-03568-7>.
91. Modolo L, Lerat E. UrQ: an efficient software for the Unsupervised Quality trimming of NGS data. *BMC Bioinformatics*. 2015;16(1):137. <https://doi.org/10.1186/s12859-015-0546-8>.
92. Bonfiglio F, Lasorsa VA, Cantalupo S, D'Alterio G, Aievola V, Boccia A, Ardito M, Furini S, Renieri A, Morini M. Inherited rare variants in homologous recombination and neurodevelopmental genes are associated with increased risk of neuroblastoma. *EBioMedicine*. 2023;87:104395. <https://doi.org/10.1016/j.ebiom.2022.104395>.
93. Jones DA, Rybak K, Hossain M, Bertazzoni S, Williams A, Tan K-C, Phan HT, Hane JK. Repeat-induced point mutations driving *Parastagonospora nodorum* genomic diversity are balanced by selection against non-synonymous mutations. *Commun Biology*. 2024;7(1):1614. <https://doi.org/10.1038/s42003-024-07327-7>.

94. Linck E, Battey C. Minor allele frequency thresholds strongly affect population structure inference with genomic data sets. *Mol Ecol Resour.* 2019;19(3):639–47. <https://doi.org/10.1111/1755-0998.12995>.
95. Somenzi E, Partel E, Barbato M, Chero Osorio AM, Colli L, Franceschi N, Mantovani R, Pilla F, Komjanc M, Achilli A. Genetic legacy and adaptive signatures: investigating the history, diversity, and selection signatures in Rendena cattle resilient to eighteenth century rinderpest epidemics. *Genet Selection Evol.* 2024;56(1):32. <https://doi.org/10.1186/s12711-024-00900-y>.
96. Yang D, Griffen A, Isshiki M, Greally JM, Rastogi D, Raj SM. Identifying the Genetic Ancestry of the Pediatric Obesity-Related Asthma Variant (rs6494395) at Rps27l. *Pediatr Pulmonol.* 2025;60(7):e71209. <https://doi.org/10.1002/ppul.71209>.
97. Mbole-Kariuki MN, Sonstegard T, Orth A, Thumbi SM, Bronsvoort BC, Kiara H, Toye P, Conradie I, Jennings A, Coetzer K. Genome-wide analysis reveals the ancient and recent admixture history of East African Shorthorn Zebu from Western Kenya. *Heredity.* 2014;113(4):297–305. <https://doi.org/10.1038/hdy.2014.31>.
98. Schiffels S, Wang K. MSMC and MSMC2: the multiple sequentially Markovian coalescent. In: *Statistical population genomics. Humana.* 2020;147–65. https://doi.org/10.1007/978-1-0716-0199-0_7.
99. Sims D, Sudbery I, Illott NE, Heger A, Ponting CP. Sequencing depth and coverage: key considerations in genomic analyses. *Nat Rev Genet.* 2014;15(2):121–32. <https://doi.org/10.1038/nrg3642>.
100. McGinty SP, Kaya G, Sim SB, Makunin A, Corpuz RL, Quail MA, Abuelanin M, Lawniczak MK, Geib SM, Korlach J. CIFI: accurate long-read chromosome conformation capture with low-input requirements. *Nat Commun.* 2025. <https://doi.org/10.1038/s41467-025-66918-y>.
101. Yin X, Huang F, Liu X, Guo J, Cui N, Liang C, Lian Y, Deng J, Wu H, Yin H. Phylogenetic analysis based on single-copy orthologous proteins in highly variable chloroplast genomes of *Corydalis*. *Sci Rep.* 2022;12(1):14241. <https://doi.org/10.1038/s41598-022-17721-y>.
102. Yuan H, Atta C, Tornabene L, Li C. Assexon: assembling exon using gene capture data. *Evolutionary Bioinf.* 2019;15:1176934319874792. <https://doi.org/10.1177/1176934319874792>.
103. Conte MA, Joshi R, Moore EC, Nandamuri SP, Gammerdinger WJ, Roberts RB, Carleton KL, Lien S, Kocher TD. Chromosome-scale assemblies reveal the structural evolution of African cichlid genomes. *Gigascience.* 2019;8(4):giz030. <https://doi.org/10.1101/383992>.
104. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics.* 2014;30(9):1312–3. <https://doi.org/10.1093/bioinformatics/btu033>.
105. Shapiro B, Rambaut A, Drummond AJ. Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. *Mol Biol Evol.* 2006;23(1):7–9. <https://doi.org/10.1093/molbev/msj021>.
106. Rohlf S, Weir B. Distributions of Hardy–Weinberg equilibrium test statistics. *Genetics.* 2008;180(3):1609–16. <https://doi.org/10.1534/genetics.108.088005>.
107. Sun YV, Kardia SL. Imputing missing genotypic data of single-nucleotide polymorphisms using neural networks. *Eur J Hum Genet.* 2008;16(4):487–95. <https://doi.org/10.1038/sj.ejhg.5201988>.
108. Miao B, Wang Z, Li Y. Genomic analysis reveals hypoxia adaptation in the Tibetan mastiff by introgression of the gray wolf from the Tibetan Plateau. *Mol Biol Evol.* 2017;34(3):734–43. <https://doi.org/10.1093/molbev/msw274>.
109. Aiyesa LV, Kaufmann D, Zumbach B, Link W, Scholten S, Beissinger T. Individual plant genetics reveal the control of local adaptation in European maize landraces. *BMC Biol.* 2025;23(1):138. <https://doi.org/10.1186/s12915-025-02241-8>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.