

Reviewed Preprint

v1 • August 20, 2026

Not revised

✉ For correspondence:

whli@uchicago.edu

jphuang@sinica.edu.tw

ijtsai@sinica.edu.tw

shensf@sinica.edu.tw

† These authors contributed equally to this work.

Competing interests: No competing interests declared

Funding: See [page 39](#)

Reviewing editor: Yingguang Frank Chan, Groningen Institute for Evolutionary Life Sciences, University of Groningen, Netherlands

© 2026, Lee et al. This article is distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use and redistribution provided that the original author and source are credited.

Eco-genomic analysis uncovers precision-conservation targets for the western Pacific's southernmost salmonid

Yi-Chien Lee^{1,†}, Zong-Yu Shen^{1,2,3,†}, Wei-Ren Lin^{1,†}, Tzi-Yuan Wang^{1,†}, Mark Liu¹, Yu-Chen Yeh^{1,4}, Shih-Fan Chan¹, Mei-Yeh Jade Lu¹, Hui-Yu Wang¹, Lin-Yan Liao⁵, Wen-Hsiung Li^{1,6}✉, Jen-Pan Huang^{1,2}✉, Isheng Jason Tsai^{1,2,4}✉, Sheng-Feng Shen^{1,2}✉

¹Biodiversity Research Center, Academia Sinica, Taipei, Taiwan • ²Department of Life Science, National Taiwan Normal University, Taipei, Taiwan • ³The University Museum, The University of Tokyo, Tokyo, Japan • ⁴Genome and Systems Biology Degree Program, Academia Sinica and National Taiwan University, Taipei, Taiwan • ⁵Wuling Station, Shei-Pa National Park, Taichung, Taiwan • ⁶Department of Ecology and Evolution, University of Chicago, Chicago, United States

eLife Assessment

This **valuable** study combines a chromosome-level genome assembly with population resequencing and demographic modelling to reassess the status of the Formosan landlocked salmon, a critically endangered salmonid at the southern edge of the genus. The assembly and the evidence for extensive lineage-specific chromosomal fusions are genuinely well done, and the discovery of an overlooked, genetically distinct population in Hehuan Creek is sound and of clear management relevance. Evidence for the headline claims is **incomplete**: species rank is asserted but nowhere argued, and gene flow is excluded using a coalescent model that contains no migration parameter. Support for the population-viability conclusions is also not complete, as the demographic mechanisms invoked in the discussion are not borne out by the supplementary tables.

<https://doi.org/10.7554/eLife.111444.1.sa2>

Abstract

Understanding how isolated small populations persist and adapt in adverse environments is instrumental to evolutionary and conservation biology. We combine a chromosome-level genome assembly, population resequencing and forward-time simulations to reconstruct the history and viability of the Formosan landlocked salmon (*Oncorhynchus formosanus*), now restricted to a handful of high-mountain headwaters in Taiwan. We estimate that this lineage has diverged from Japanese masu salmon over one million years ago and has no detectable gene flow for ~50,000 years. It has accumulated extensive chromosome fusions and expansions of cold-adaptation gene families, qualifying it as a new species rather than a subspecies of Japanese masu salmon. Whole-stream sampling reveals an overlooked Hehuan-Creek population that retains high heterozygosity and has gained unique alleles. Life-table simulations show that the Hehuan population has a notably lower extinction risk and can persist or even grow under low-to-moderate typhoon frequency, whereas Qijiawan-Creek population would decline precipitously under the same or higher frequency. These findings contradict the notion that peripheral populations are likely genetically depleted and support stream-specific “precision conservation” in place of broad, untargeted translocations that could erode local adaptation potential. Thus, our genomic-ecological analysis has uncovered hidden strong resilience in a critically endangered, climate-threatened salmonid lineage.

Introduction

How do small populations at the margins of a species' range maintain genetic diversity and adaptive potential under long-term isolation and extreme environmental conditions^{1,2}? Moreover, what are the implications of these dynamics in the face of accelerating climate change^{2,3}? These questions lie at the heart of modern conservation biology. Marginal populations, though often small and fragmented, may harbor unique adaptive variation shaped by intense local selection⁴. For example, edge populations of Australian vertebrates in hot, arid regions have evolved enhanced drought or heat tolerance, potentially serving as reservoirs of climate-relevant alleles⁵. In salmonids, local adaptation is both rapid and fine-scaled: within just a few generations, native populations frequently outperform immigrants in survival and reproductive success⁶.

However, evolutionary outcomes are not dictated by selection alone. In very small populations, genetic drift can overwhelm selection, eroding adaptive variation and fixing deleterious alleles^{7,8}. Migrants may occasionally restore lost variation or even outperform residents⁹. This interplay between drift and selection complicates conservation decisions. Peripheral populations are often recognized as evolutionarily significant units (ESUs) for their unique genetic legacies¹⁰, but also face elevated extinction risks due to inbreeding, reduced heterozygosity, and loss of adaptive potential^{3,11,12}. Genetic rescue—introducing individuals from other populations to restore diversity—is increasingly promoted^{9,13}, though not without risks of outbreeding depression¹⁴ and sociopolitical resistance, especially in culturally iconic taxa.

The Formosan landlocked salmon (*Oncorhynchus formosanus*, Jordan & Oshima 1919) exemplifies this conservation paradox. Endemic to Taiwan, it is the southernmost lineage and one of the most tropical salmonids, classified within the *O. masou* species complex alongside the Japanese Masu (*O. m. masou*), Amago (*O. m. ishikawae*)¹⁵, and Biwa salmon—now recognized as the distinct species *O. biwaensis*¹⁶ rather than *O. m. rhodurus* (Fig 1a [↗](#)). Morphologically distinct and geographically isolated by over 1,000 km of ocean and mountains¹⁷, its taxonomic status remains contested. Although the IUCN lists it as a distinct and Critically Endangered species¹⁸, genetic studies have often placed it within the *O. masou* clade, based on limited molecular data^{15,19–21}. Although species-level recognition has been proposed²², this view was based on earlier datasets, lacking new genetic evidence.

This taxonomic ambiguity complicates conservation efforts. After decades of habitat degradation and overharvesting, wild populations fell to ~600 individuals by the 1980s²³. Conservation actions, including captive breeding and habitat restoration, gradually increased the population to ~1,000 individuals. However, at that time the species remained confined to a single confirmed stronghold—a 7-km stretch of Chichiawan (Qijiawan) Creek, with the Luoyewei population being hatchery-supplemented and the Hehuan population presumed extirpated since the 1960s (Fig. S1 [↗](#)). In 2017, reintroduction began in Hehuan Creek using juveniles bred from Qijiawan-derived broodstock. Over three years, more than 6,000 juveniles were released across three sectors of Hehuan Creek^{24,25}. However, post-reintroduction assessments have focused narrowly on census numbers, with limited attention to genetic monitoring. Early molecular studies identified low diversity, likely reflecting historical bottlenecks and long-term isolation, but these inferences relied on sparse sampling and low-resolution markers. The extent of genome-wide variation and its relevance to adaptation remains poorly characterized. More recently, a controversial proposal suggested hybridizing *O. formosanus* with Japanese masu salmon to boost genetic diversity and resilience²⁶. This proposal lacks comprehensive genomic validation and remains divisive, particularly given the species' iconic status in Taiwan. Much of the controversy stems from insufficient genomic data, hampering rigorous assessment of the risks and benefits of such genetic rescue interventions.

In this study, we integrate whole-genome sequencing, extensive population sampling, demographic modeling, and ecological monitoring to reevaluate the evolutionary legacy and conservation outlook of *O. formosanus*, a critically endangered salmonid persisting at the tropical edge of its range. We assess whether this long-isolated lineage retains meaningful adaptive potential despite historical bottlenecks, and how genomic and demographic patterns vary across

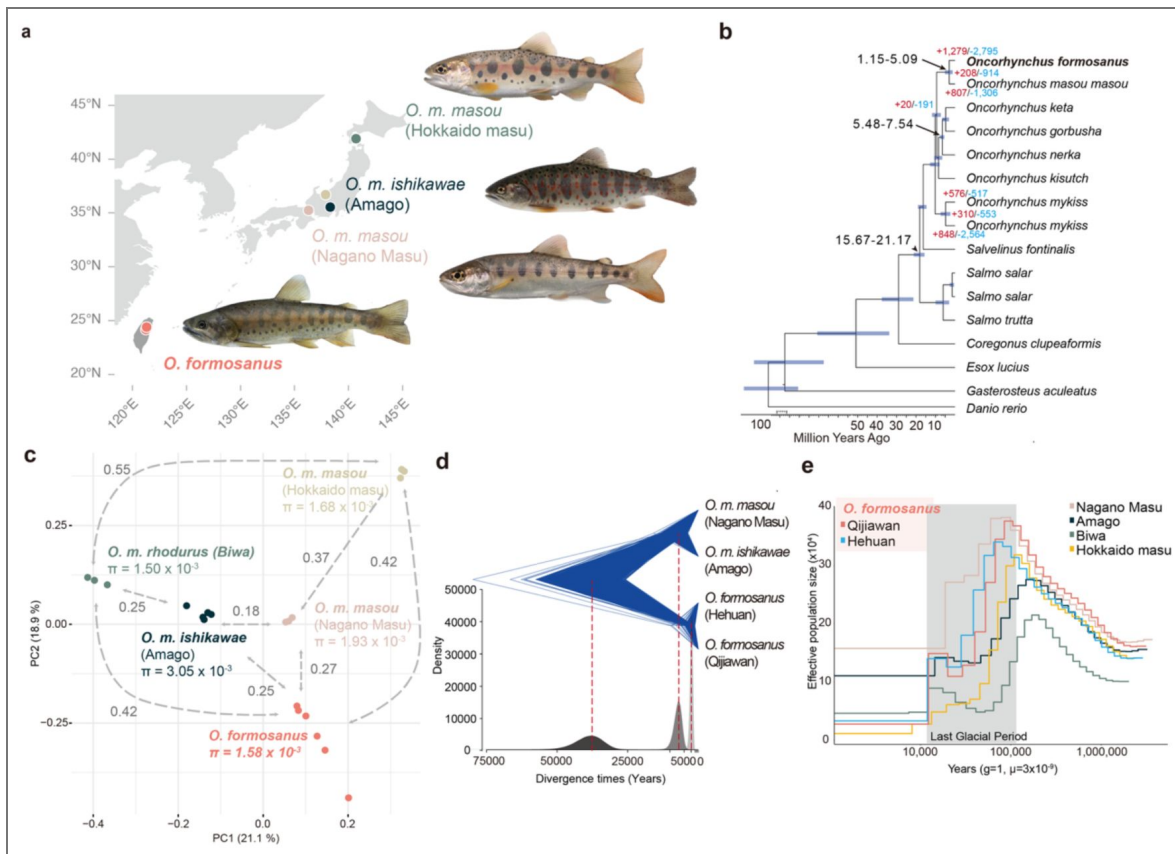


Figure 1. Phylogenomics and demographic history of *O. formosanus*.

(a) Geographic distribution and phenotypic variation of selected lineages within the *Oncorhynchus masou* species complex. (b) Species tree inferred from a concatenated alignment of 1,483 single-copy orthogroups. Blue horizontal bars representing the 95% confidence intervals. Numbers in red and blue denote the gained and lost orthogroups at each node, respectively. (c) Principal component analysis (PCA) showing genetic differentiation among masu salmon lineages. Numbers along dashed arrows represent mean pairwise homozygous SNP differences (in millions) between populations. (d) Time-calibrated diversification within Formosan (*O. formosanus*) and Japanese (Nagano Masu and Amago) salmon lineages visualized by DensiTree, indicating lineage splits occurring approximately 50,000–20,000 years ago. (e) Demographic histories inferred by PSMC, showing changes in effective population size for Japanese and Taiwanese salmon individuals.

relictual stream populations. By linking genome-wide diversity with stream-specific life-history dynamics and climate-linked disturbances, we uncover mechanisms of persistence—and vulnerability—that have broader relevance for managing small, fragmented populations in a rapidly changing world.

Results

Sequencing, assembly and annotation of the *O. formosanus* genome

We sequenced the genome of *O. formosanus* from a single female individual from the Qijiawan Creek, one of the last remaining habitats. An initial assembly of ~2.9 Gb was generated from 50.0x PacBio HiFi reads (Table S1 [\[link\]](#)), similar to genome sizes reported for other salmonids (Table S2 [\[link\]](#)). This assembly was further scaffolded with 31.4x Omni-C reads using YaHS²⁷ (Fig. S2 [\[link\]](#)). The final chromosome-level assembly anchored 97.3% of the genome onto 25 scaffolds (Table S2 [\[link\]](#)), representing another high-quality genomic resource within the *Oncorhynchus* species; the assembly represents one of the fewest salmon chromosome numbers²⁸. In addition, the mitochondrial genome was separately assembled into a single circular contig of 16.6kb²⁹. The *O. formosanus* genome exhibits 56-65% repetitive content per chromosome, predominantly comprising the DNA transposon family (Table S3 [\[link\]](#)), like other salmonids^{30,31}.

Using a combination of transcriptome sequencing from multiple tissues and proteomes from eight representative salmonids as references, we predicted 43,110 coding genes (Methods) in *O. formosanus*, which was estimated to be 95.7% complete based on BUSCO (benchmarking universal single-copy orthologs)³² assessment. Of these predictions, 84.1% could be assigned to 6,696 protein families using Pfam scan and 89.3% could be assigned Gene Ontology terms using eggNOG-mapper³³. Orthology analysis with Orthofinder³⁴ placed these predicted gene models alongside protein sequences from eleven representative salmonid and four outgroup genomes (Table S4 [\[link\]](#)) into 30,011 orthologous groups (OGs), 97.7% of which were shared broadly across salmonid species, and only 0.45% were specific to *O. formosanus*. We mapped genomic sequencing data from three to five representative individuals of Japanese Amago, Biwa, Nagano and Hokkaido Masu, as well as six individuals from different Taiwanese populations, to assess genomic diversity within and between lineages (Table S1 [\[link\]](#)). Sequencing depth ranged from 18 - 41.8x coverage, with high mapping rates (99.49% to 99.74%;) achieved for all samples against our *O. formosanus* assembly compared to the existing *O. m. masou* reference genome, indicating close genomic affinity and reliable alignment quality (Table S5 [\[link\]](#)). Further genomic diversity was assessed using double-digest restriction-site-associated DNA sequencing (ddRADseq) data generated from 21, 17, and 40 individuals collected from Qijiawan, Luoyewei, and Hehuan creek populations, respectively (Table S1 [\[link\]](#)).

Delimitation and demography of the *Oncorhynchus masou* species complex

We conducted phylogenomic analyses using 1,484 genome-wide single-copy ortholog sets to infer the evolutionary history and taxonomic position of *O. formosanus* relative to other salmonids. Maximum likelihood analyses of a concatenated supermatrix and coalescent-based species tree analyses placed *O. formosanus* within the *Oncorhynchus masou* species complex as the sister lineage to *O. m. masou* (Fig. 1a [\[link\]](#) and Fig. S3 [\[link\]](#)). Using MCMCTree³⁵ with two fossil calibrations³⁶, we estimated the most recent common ancestor of these two lineages to be 1.2 to 5.1 million years ago (Mya) during the Pliocene and Pleistocene epoch (Fig. 1b [\[link\]](#)), indicating long term evolutionary isolation. Principal component analysis (PCA) based on genome-wide SNP data clearly separated different taxa from the previously designated masu salmon species complex along two primary axes (Fig. 1c [\[link\]](#)). The Biwa (*O. m. rhodurus*) lineage was distinctly differentiated along PC1, while the divergence between *O. formosanus* and *O. m. masou* appeared prominently along PC2, consistent with their geographic extremes (Central Taiwan versus Hokkaido, Japan). Pairwise genetic differentiation *F*_{st} was highest among the geographically isolated lineages (Biwa, *O. m.*

masou and *O. formosanus*, Fig. S4 [↗](#)), whereas lower differentiation among geographically southern lineages (Amago (*O. m. ishikawae*), Nagano Masu, and *O. formosanus*; 0.18–0.27) suggests relatively recent divergence or historical gene flow among these lineages (Fig. S4 [↗](#)).

To statistically delimit the taxonomic status of the Taiwanese *O. formosanus* and two Japanese southern lineages within the *O. masou* species complex, we first reconstructed a species tree among the three taxa based on a multi-species coalescent model³⁷ and the ddRADseq data sets (Fig. 1d [↗](#)): the earliest divergence (~50,000 years ago) separated the Taiwanese lineages (Hehuan and Qijiawan) from the Japanese lineages (Nagano Masu and Amago), followed by the divergence between Nagano Masu and Amago ~ 20,000 years ago. By using a species delimitation model while allowing the species tree to vary among MCMC searches (A11 model; Methods), our genome-wide data sets suggested three species, or three evolutionarily independent lineages³⁸. Specifically, the two populations within Taiwan belong to the same evolutionary lineage, which is independent from the other two evolutionary lineages, Nagano Masu and Amago. The support for three independent lineages in our data sets received > 99% posterior probability. In addition, the genealogical divergence index (GDI) estimated from our genome-wide data sets further supported substantial differentiation between Taiwanese and Japanese salmon lineages, with notably high values (Fig. S5 [↗](#)), exceeding previously reported thresholds distinguishing biological species^{39,40}. The topology of the maximum likelihood phylogeny from both the nuclear and mitochondrial genome trees also placed the *O. formosanus* species sister to all Japanese lineages (Fig. S6 [↗](#)), suggesting an ancient split of Taiwanese salmon well before the diversification of regional populations in the Japanese archipelago. Together, these results strongly indicate that *O. formosanus* represents a distinct evolutionary significant unit (ESU) separate from the Japanese masu salmon and will be referred as *O. formosanus* as previously suggested^{18,22}.

Analyses of historical demographic trajectories under a pairwise sequentially Markovian coalescent (PSMC) model indicated that all lineages experienced demographic expansions approximately 139,900– 64,600 years ago, prior to the Last Glacial Maximum (Fig. 1e [↗](#)). Peak effective population sizes (N_e) varied considerably, with Nagano Masu (~381,000) and Qijiawan (~375,000) having the largest historical N_e . However, all populations underwent pronounced contractions during and post glaciation, particularly severe in the Taiwanese populations, resulting in recent N_e below 28,000 and 34,000 individuals for Qijiawan and Hehuan creeks, respectively. We further calculated the relative proportions of homozygous and heterozygous variants in both synonymous and loss-of-function (LOF) mutations. Across all populations, heterozygous derived LOF variants were more frequent than homozygous derived LOF variants (Fig S7), consistent with the expectations that smaller populations would accumulate higher frequencies of deleterious alleles and increased homozygosity. In particular, the Qijiawan creek population showed a significantly higher proportion of both heterozygous and homozygous derived LOF variants compared to synonymous variants, highlighting its heightened genetic vulnerability (Fig. S7 [↗](#)). Together, these findings emphasize the profound evolutionary distinctiveness and genetic vulnerability of *O. formosanus* populations.

Extensive chromosome fusions and sex chromosome rearrangements in *O. formosanus*

Genome-wide synteny analyses uncovered substantial chromosomal rearrangements differentiating masu salmon lineages, using rainbow trout (*O. mykiss*), *O. keta* and *S. fontinalis* as outgroups (Fig. 2a [↗](#) and Table S6 [↗](#)). Eight chromosomes of the masu salmon complex originated exclusively from fusion events involving 16 ancestral chromosomes identified through comparison with homologous chromosomes in *O. mykiss*, reflecting considerable genomic reorganization since their divergence (Fig. 2a [↗](#) and Table S6 [↗](#)). Notably, *O. formosanus* underwent further chromosomal reorganizations through telomere-to-telomere fusion events involving at least 14 ancestral chromosomes, in contrast to only two in *O. m. masou* (Fig. 2a [↗](#) and Table S6 [↗](#)). For example, chromosome 1 of *O. formosanus* corresponds to chromosomes 23 and 24 in *O. m. masou* and chromosomes 5, 29, 1, and 9 in *O. mykiss*, while chromosome 4 corresponds to chromosomes 4 and 5 in *O. m. masou* and chromosome 17 in *O. mykiss*, illustrating significant

ancestral chromosome fusions. These fusion sites are enriched in repeats including satellite DNA, characteristic of centromeric repositioning or neocentromere formation⁴¹. In addition, genes residing at these *O. formosanus* chromosomes displayed significantly elevated evolutionary rates compared to non-fused chromosomes (dN/dS ratios, $P=0.009$, Wilcoxon Rank Sum test, Fig. S8⁴²), suggesting that structural genomic rearrangements facilitated adaptive evolution in response to ecological pressures such as cold tolerance and metabolic specialization^{42,43}.

The sex-linked region demonstrated particularly pronounced chromosomal rearrangements. An earlier study linked the contig containing the sex-determining gene (sdY) to chromosome 7 in *O. m. masou*, although alternative arrangements were suggested for the Amago and Biwa lineages²⁸. Our genome assembly originated from a single female revealed that this *O. m. masou* contig was aligned uniquely to the central region (~50 Mb) of chromosome 13 (Fig. S9⁴⁴). This location corresponds exactly to a known ancient chromosome fusion site homologous to chromosomes 14 and 25 in rainbow trout (*O. mykiss*) (Fig. 2a⁴⁵ and Fig. S10⁴⁶) and contains abundant 281-mer satellite repeats typical of centromeric regions.

To further clarify this genomic region, we produced an additional genome assembly from a male individual (Methods). Three scaffolds were mapped uniquely to this fusion region and revealed clear coverage differences between males and females, enabling us to unambiguously distinguish XY-shared contigs from Y-specific ones (Fig. 2b⁴⁷). An inversion was identified between the XY-shared region between the two assemblies, and the sdY gene was identified specifically on a Y-specific contig, closely adjacent to the Y chromosome's centromeric region (Fig. 2c⁴⁸). Interestingly, sdY is also located ~5Mb adjacent to the centromere of *O. mykiss* non-homologous Chr29, consistent with sdY's mobility in fishes^{44,45}. Pairwise identity comparisons between homologous X and Y-specific regions from these contigs revealed elevated sequence divergence averaging 9.6%, indicating that XY divergence was restricted to the genomic segment spanning from sdY to the centromere, a pattern similarly observed in rainbow trouts⁴⁴. Our findings demonstrate significant differences in sex chromosome configurations within the *O. masou* complex, highlighting how chromosomal rearrangements have shaped genomic evolution and potentially contributed to reproductive isolation.

Proteome specialization of *O. formosanus*

We investigated gene and protein domain expansions specific to *O. formosanus* by annotating protein family (Pfam) domains across salmonid genomes (Fig. S11a⁴⁹). Several Pfam domains exhibited significant expansions, notably the immune-related C1-set immunoglobulin domain (>60 copies), chromatin-associated Chromo domain (~60 copies), glycosyltransferase (Glyco_transf_49; ~60 copies), and glycoside hydrolase (Glyco_hydro_56; ~13 copies). Interestingly, we identified an enrichment of microtubule-associated protein MAP6 (~30 copies), which are known to confer cold stability to microtubules⁴⁶ and upregulated during acute cold stress in Japanese flounder⁴⁷. Aligning with this cold-adaptive expansion, we observed a substantial enrichment of the ice-binding lectin domain (LbR_Ice_bind; ~19 copies) in *O. formosanus*. Ice-binding proteins (IBPs) protect cells by binding to and inhibiting ice crystal growth. Our phylogenetic analyses indicated that these expanded LbR_Ice_bind proteins in *O. formosanus* formed a lineage distinct from previously characterized ice-structuring proteins (ISP types I, II, and IV⁴⁸) identified in other *Oncorhynchus* species (Table S7⁵⁰ and Fig. S12⁵¹), suggesting coordinated functional specialization to colder freshwater habitats. Orthologous group expansions via CAFE analyses further highlighted significant enrichment in 331 genes associated with exocrine and salivary gland development, chromatin assembly, DNA conformation changes, homophilic cell adhesion, innate immune responses, and calcium ion transport (Table S8⁵²), underpinning key physiological adaptations to freshwater environments. In contrast, closely related *O. m. masou* exhibited distinctly different expanded domains (e.g., Beach and Polycystin domains, Fig. S11b⁵³), reflecting divergent evolutionary trajectories post-geographic separation⁴⁹.

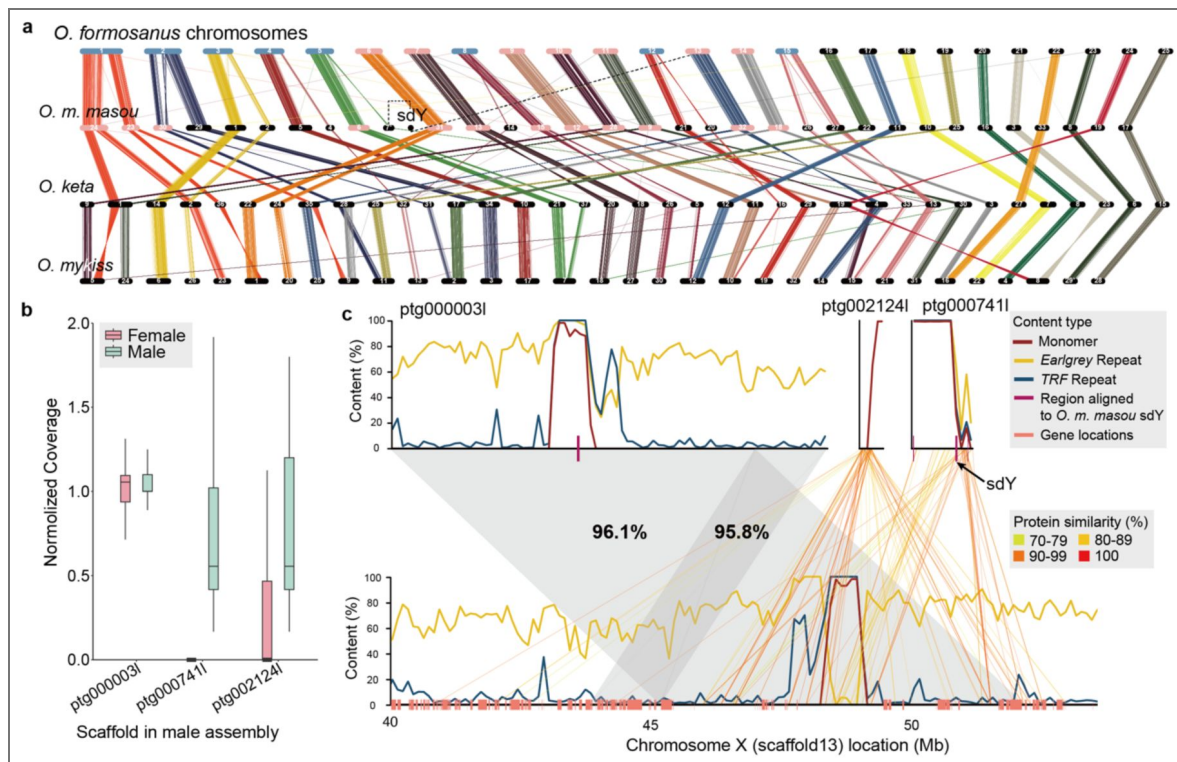


Figure 2. Chromosome synteny and sex chromosome evolution in *O. formosamensis*

(a) Genome-wide synteny comparison across *Oncorhynchus* species. Chromosomes highlighted in blue denote fusions unique to *O. formosanus*, and those in red indicate fusions in the common ancestor of masu salmon lineages. Colored lines represent pairwise single-copy orthologs. The scaffold containing the sex-determining gene (*sdY*) in *O. m. masou* follows data from Christensen et al. ²⁸. (b) Normalized coverage of female (pink) and male (blue) sequencing reads mapped onto selected scaffolds from the male *O. formosanus* genome assembly. (c) Detailed annotation of repetitive elements and sequence alignments between putative X-linked (ptg000003l) and Y-specific scaffolds (ptg000741l and ptg002124l) from the male genome, aligned to chromosome X (scaffold13) from the female assembly. Numbers above shaded grey area indicate nucleotide identity between X chromosome sequences of male and female assemblies. Lines connecting Y-specific scaffolds (ptg000741l and ptg002124l) and scaffold13 represent homologous regions, with colors indicating amino acid sequence similarity. The exact location of the *sdY* gene on scaffold ptg000741l is indicated.

Divergent population structures among *O. formosanus* populations

To resolve fine-scale genetic structure and evolutionary differentiation within *O. formosanus*, we analyzed 1,509 SNPs derived from whole-genome sequencing and ddRADseq datasets (Table S1 [↗](#) and Fig. S13 [↗](#)), applying a 1-Mb window selection approach across individuals from the three major stream systems. Principal component analysis (PCA) revealed distinct clustering patterns that mirrored geographic distributions (Fig. 3a [↗](#)). Importantly, populations from the Hehuan creek (HEH) exhibited substantially greater dispersion along both principal component axes compared to the other groups, occupying an expanded region in multivariate space. In contrast, Luoyewei creek populations and the downstream sites of Hehuan creek (HEH_01, HEH_02) formed tightly clustered, discrete genetic groups, suggestive of localized genetic homogenization and elevated genetic diversity within the Hehuan creek.

Long-term native status of the Hehuan lineage

To further investigate if the Hehuan creek system harbor potential, previously unrecognized reservoir of genetic variation within *O. formosanus*, we assessed multilocus heterozygosity (MLH) and runs of homozygosity (ROH) across individuals. Hehuan creek samples consistently exhibited the highest heterozygosity and lowest levels of inbreeding among all creeks (MLH: HEH–QIJ mean difference = 0.07, $P < 0.05$; HEH–LUO = 0.23, $P < 0.05$; ROH: HEH–QIJ = −310,968.3, $P < 0.05$; HEH–LUO = −1,108,330, $P < 0.05$) (Fig. 3b–c [↗](#)), followed by Qijiawan creek (MLH: QIJ–LUO = 0.15, $P < 0.05$) and lower inbreeding levels (ROH: QIJ–LUO = −797,361.8, $P < 0.05$), while Luoyewei creek samples showed markedly reduced heterozygosity and elevated inbreeding. The ROH patterns in Hehuan individuals—characterized by fewer and shorter homozygous segments—are indicative of demographic continuity and argue against a recent or artificially-introduced origin. Within Qijiawan creek, we also detected a clear longitudinal gradient in genetic diversity (Pearson’s r : −0.82, $P < 0.05$): downstream populations retained higher heterozygosity and lower inbreeding compared to upstream ones. This pattern aligns with theoretical expectations from metapopulation dynamics, wherein downstream connectivity supports gene flow while upstream isolation fosters drift⁵⁰. In contrast, the Luoyewei creek showed no such spatial structure, consistent with its known history of hatchery stocking and limited founder input.

Demographic patterns across stream populations

Population genetic analyses revealed distinct demographic parameters across streams. Effective population size (N_e) measurements showed significant differences among populations: Luoyewei Creek had $N_e < 1$, while Hehuan and Qijiawan populations maintained substantially higher values of 98.5 and 84.5, respectively (Fig. 4a [↗](#)). Temporal reconstruction analyses of historical N_e revealed divergent trajectories (Fig. 4b [↗](#)). The data showed parallel demographic patterns between Hehuan and Qijiawan until approximately 100 years ago, when they began to diverge coinciding with the 1930s conservation designation of Qijiawan Creek populations as a natural monument during the Japanese administration. Both populations showed marked declines beginning in the 1960s, corresponding with the construction of the Central Cross-Island Highway. Our N_e reconstruction from Fig. 4b [↗](#) showed different recovery patterns in effective population size: Qijiawan Creek has exhibited gradual recovery over the past two decades, while Hehuan Creek’s N_e increased notably following the 2017 reintroduction program^{24,25}. Critically, post-2017 genomic data from Hehuan Creek revealed higher-than-expected genetic diversity—reflected in elevated multilocus heterozygosity (MLH), reduced runs of homozygosity (ROH), and the presence of unique allelic variants not found in the Qijiawan source population (Fig. 3b–c [↗](#), Table S1 [↗](#)). Additionally, population structure analyses identified a discrete genetic cluster within Hehuan Creek, suggesting the persistence of a previously undetected native lineage, now likely admixed with reintroduced individuals.

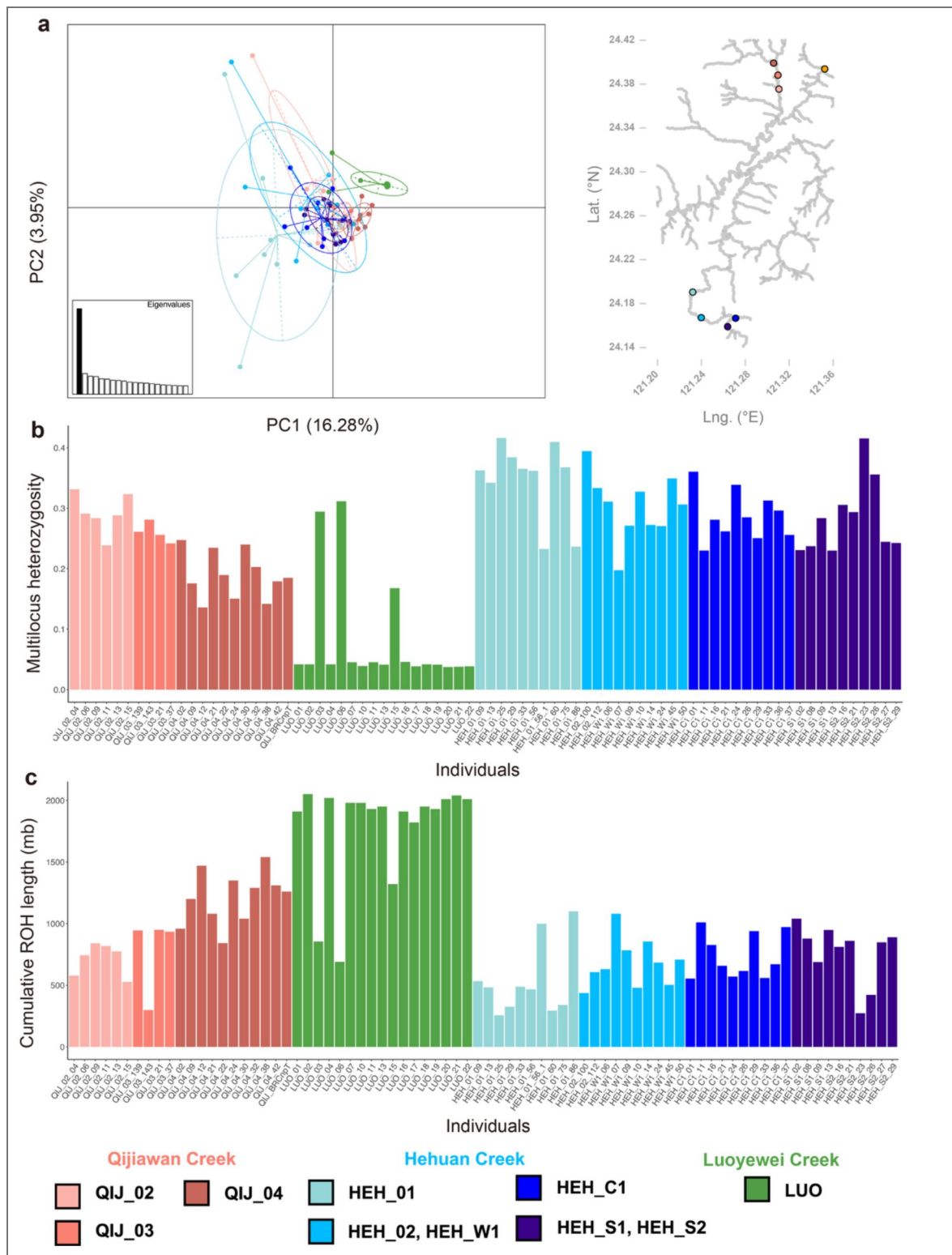


Figure 3. Population structure, genetic diversity and level of inbreeding of the Formosan salmon.

(a) PCA result of the population structure shows no significant differentiation among most populations, except for those from Luoyewei Creek and the downstream sites of Hehuan Creek (HEH_01, HEH_02, HEH_W1). Populations from Hehuan Creek occupy a broader range in PCA space (PC1 vs. PC2), indicating a higher genetic diversity. (b) Multilocus heterozygosity across individuals from the three creek systems. Hehuan Creek populations exhibit the highest genetic diversity, followed by Qijiawan Creek and lastly by Luoyewei Creek. (c) The level of inbreeding, estimated using runs of homozygosity (ROHs), is highest in Luoyewei Creek, intermediate in Qijiawan Creek, and lowest in Hehuan Creek.

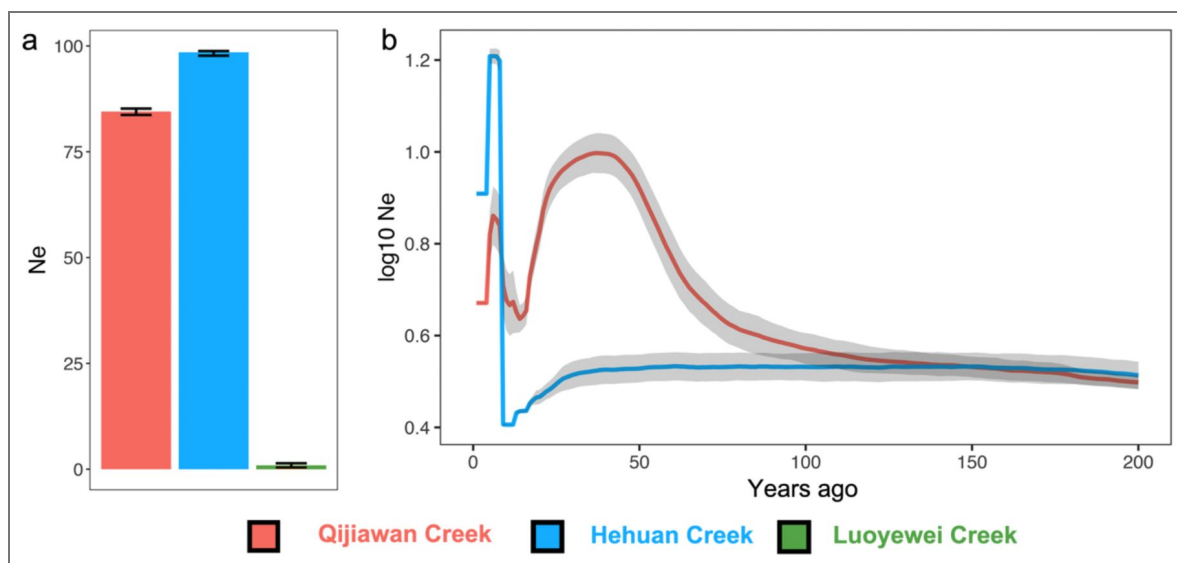


Figure 4. Current effective population size N_e and demographic history of Formosan salmon populations.

(a) Estimated current N_e for Formosan salmon populations across multiple stream sites. The Qijiawan and Hehuan populations appear to maintain effective sizes above thresholds associated with inbreeding depression, while the Luoyewei Creek population, which likely comprises individuals from hatchery releases, has a very small effective population size. (b) Inferred demographic histories of populations from Qijiawan and Hehuan Creeks. The contrasting trajectories reflect differences in conservation management and habitat development across the two regions.

Life table analyses reveal stream-specific demographic asymmetries and typhoon-induced shifts

Using improved non-lethal sampling methods and bootstrapped statistical inference, we generated age-structured life table models for *O. formosanus* populations in the two headwater stream systems (Methods). These estimates revealed significant impacts of typhoon on population dynamics and important differences between stream systems. Under non-typhoon scenarios, both streams exhibited strong intrinsic growth potential (dominant eigenvalue $\lambda > 1.0$), with QIJ at $\lambda = 2.23$ (1.53 – 2.65) and HEH at $\lambda = 2.25$ (1.84 – 2.83), indicating rapid population growth capacity. However, under typhoon disturbance scenarios, λ values dropped substantially below 1 (QIJ = 0.405 (0.01–0.55); HEH = 0.477 (0.26–0.66)), indicating population contraction (Table S9 [↗](#)). Notably, HEH maintained a relatively higher growth rate post-typhoon, demonstrating greater resistance to disturbance. Survival rate analysis showed that annual survival during the typhoon period (January → November) was markedly lower than during baseline periods (November → June) (Table S10 [↗](#)). Under typhoon influence, annual survival rates for age-0 fish in QIJ and HEH were 0.202 and 0.184, respectively, while under baseline conditions they were 0.616 and 0.615. Adult fish (ages 1–2) exhibited reduced survival rates during typhoon periods (0.046–0.131), indicating that typhoons imposed substantial survival pressure across all age classes. Recruitment analysis revealed significant declines in recruitment capacity due to typhoons. In January 2024 (non-typhoon), adjusted recruitment for QIJ and HEH was 1,492 and 2,354 individuals/1000m², respectively, but decreased to 657 and 634 individuals/1000m² in January 2025 (post-typhoon), representing declines of 56–74%. Fecundity coefficients were similarly affected (Table S11 [↗](#)), with maximum fecundity under non-typhoon scenarios reaching 19.3 and 8.9 for QIJ and HEH, respectively, while post-typhoon values declined to 1.53 and 1.21, indicating substantial reductions in reproductive capacity. Overall, these life table estimates demonstrate that typhoon disturbance causes comprehensive negative impacts on *O. formosanus* populations, including concurrent declines in survival, recruitment, and fecundity. HEH retained higher post-disturbance λ , exhibited superior resistance to disturbance compared to QIJ, a difference that has important implications for conservation strategy development under climate change.

Divergent long-term responses of Qijiawan and Hehuan populations to typhoon disturbances

Based on empirically derived demographic parameters—including age-specific survival rates, recruitment levels, and fecundity estimates corrected for sampling bias—we developed a stochastic, discrete-time simulation model to assess the long-term population dynamics of *O. formosanus* under varying typhoon disturbance regimes. The model incorporates a two-stage Leslie matrix structure representing pre-disturbance (January–November) and non-disturbance (November–June) periods, with transitions modulated by observed survival patterns (s_{1-11} and S_{11-6} ; Table S10 [↗](#)). Simulation results revealed divergent trajectories and extinction probabilities between the QIJ and HEH systems under varying typhoon disturbance frequencies (Fig. 5a–d [↗](#)). At low disturbance ($p = 0.3$) both populations stabilised after an initial spike, but at different plateaus: HEH settled at $\approx 1,100$ – $1,150$ fish $\cdot 10^3$ m⁻² whereas QIJ levelled off at $\approx 1,000$ – $1,050$ fish $\cdot 10^3$ m⁻². When p rose to 0.5, HEH maintained a gently declining yet robust abundance of ~ 550 – 650 individuals, while QIJ declined steadily from ≈ 600 to ≈ 450 individuals over 20 years. Under high disturbance ($p = 0.7$) both curves plunged, yet HEH fell from ~ 600 to ~ 50 individuals—hovering just above the quasi-extinction threshold—whereas QIJ dropped from ~ 450 to < 40 individuals, crossing the threshold in most replicates. At $p = 0.9$ both streams trended toward local extinction: QIJ fell below 10 individuals by year 8 and approached zero thereafter, while HEH slipped below the threshold by year 9 and dwindled to < 5 individuals by the simulation endpoint.

Extinction risk analysis using a quasiextinction threshold of ≤ 10 individuals per 10^3 m² corroborated these trends (< 2 mature female per 10^3 m²). Across 1,000 replicates, probabilities of QIJ climbed from 2.2 % (95 % CI 1.4–3.3 %) at $p = 0.3$ to 29.8 % (27.0–32.7 %) at $p = 0.5$, 85.7 % (83.4–87.8 %) at $p = 0.7$, and 100 % (99.6–100 %) at $p = 0.9$. The corresponding risks for HEH were 0.5 %,

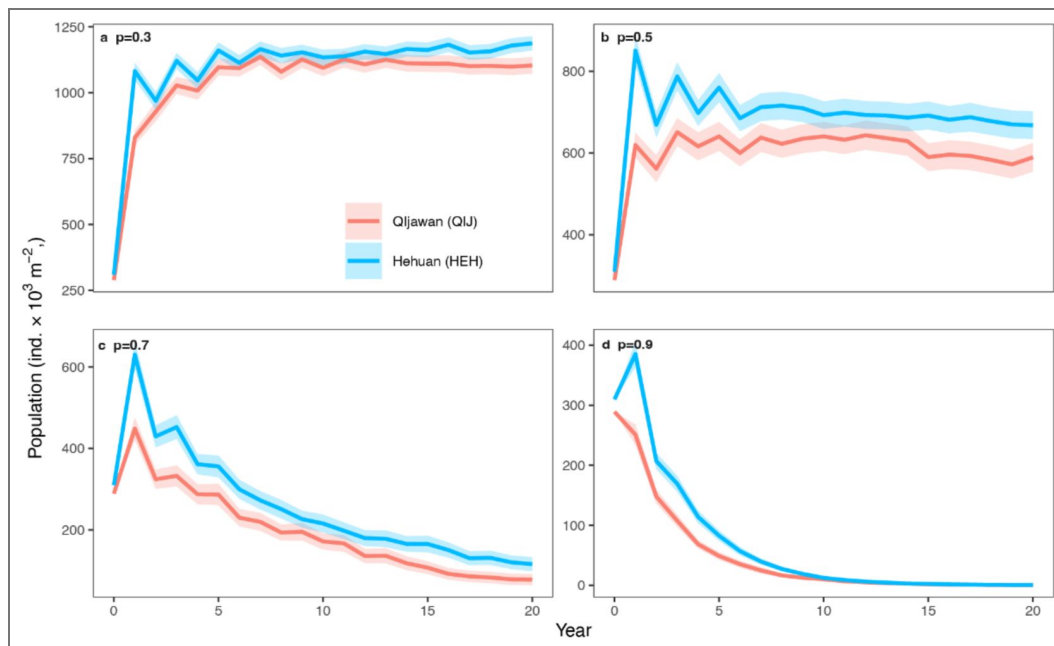


Figure 5. Population trajectories and extinction risk under varying typhoon frequencies.

(a-d) Simulated population dynamics over 20 years under different annual typhoon occurrence probabilities ($p = 0.3, 0.5, 0.7, 0.9$) over a 20-year simulation period. Panels a–d represent different disturbance scenarios, with each line indicating mean abundance (individuals aged 1–4, per 1,000 m^2) for Qijiawan (QIJ; red) and Hehuan (HEH; blue) stream populations. Shaded bands represent 95% confidence intervals across 1,000 stochastic replicates per scenario. Simulations incorporated stage-specific survival rates, Leslie matrix projections, and probabilistic typhoon impact modeled using zero-inflated Gamma distributions derived from historical records.

15.1 %, 77.4 %, and 99.9 %. Fisher's exact tests yielded odds ratios (QIJ / HEH) of 4.48 (1.69–11.86) at $p = 0.3$, 2.39 (1.92–2.97) at $p = 0.5$, and 1.75 (1.39–2.20) at $p = 0.7$, confirming that QIJ remains 1.7 to 4.5 times more vulnerable until both populations approach certain extinction (Fig. S14). LogisticGLM contrasts showed a monotonic decline in OR with increasing disturbance, reflecting convergence of relative risk as absolute probabilities approach unity. These results indicate that while both systems are vulnerable to high-frequency typhoon disturbance, QIJ exhibits substantially greater extinction risk across all disturbance scenarios.

Discussion

Adaptive capacity and conservation implications in marginal populations

Our findings reveal how small, geographically marginal populations can maintain adaptive potential under prolonged environmental constraint. *Oncorhynchus formosanus*, now confined to a few high-mountain streams in Taiwan, represents the southernmost lineage of the genus and is not a recent derivative of Japanese masu salmon. Phylogenomic analyses of 1,484 single-copy nuclear genes place its divergence from *O. m. masou* over one million years ago during the Pliocene epoch. In contrast, ddRAD-based coalescent modeling suggests that the last gene flow between Taiwanese populations and their southern Japanese relatives (Nagano Masu and Amago) occurred around 50,000 years ago. These distinct timelines reflect different evolutionary layers—deep ancestral divergence versus more recent lineage sorting—and arise from the differing temporal sensitivity of genomic markers. This ancient divergence is supported by genealogical divergence index (GDI) values exceeding established species-level thresholds^{39,40,51}. We also observed substantial chromosomal fusions accompanied by elevated dN/dS ratios, along with notable expansions in gene families associated with cold adaptation, such as *MAP6* and *LbR_Ice_bind*. These adaptive signatures have persisted despite restricted gene flow and historical bottlenecks, likely facilitated by Taiwan's complex montane river systems that both isolate populations and maintain localized connectivity. Such genomic plasticity, enabled by historical whole-genome duplications (WGD), is also documented in Atlantic salmon, where duplicated genomic regions provide raw material for rapid adaptation to local environments³¹. Similar patterns of local adaptation, where native populations outperform non-native ones, have been widely reported across salmonids⁶. Our extensive multi-stream sampling revealed genetically distinct *O. formosanus* populations with diversity levels that rival, and in some cases exceed, those of Japanese counterparts. However, effective population size (N_e) estimates show stark contrasts: Luoyewei populations exhibit extremely low N_e (<1), consistent with their hatchery origin, while Hehuan and Qijiawan creeks maintain higher values (98.5 and 84.5, respectively). These estimates exceed the threshold ($N_e < 50$) needed to minimize short-term inbreeding depression^{52,53}, but still fall within the range ($50 < N_e < 500$) that warrants targeted conservation intervention.

Historical conservation impacts and demographic trajectories

The contrasting trajectories of the Hehuan (HEH) and Qijiawan (QIJ) populations illustrate how the timing, enforcement, and ecological targeting of conservation measures shape long-term demographic outcomes. Until approximately a century ago, both lineages apparently followed similar patterns. Their histories began to diverge in the 1930s when *O. formosanus* was designated a natural monument under the Japanese colonial rule^{54,55}. Restrictions on harvest and agricultural activity were implemented across the upper Dajia River basin, and these measures appear to have been more effectively enforced in Qijiawan, contributing to early signs of demographic recovery. In contrast, similar bans in Hehuan Creek may have had limited impact due to the remoteness and inaccessibility of the area at the time.

Despite these early interventions, both populations experienced sharp declines in the 1960s, coinciding with the construction of the Central Cross-Island Highway and widespread habitat fragmentation⁵⁶. These patterns underscore that *O. formosanus* retains recovery potential when

habitats receive sustained and effective protection⁵⁷. However, the trajectory differences between HEH and QIJ in recent decades suggest that early legal protections alone are insufficient without continued ecological management tailored to each stream's conditions.

Our demographic simulations (Fig. 5 [↗](#)) support this view. Under increasing typhoon disturbance probabilities, QIJ populations declined steadily under all disturbance levels, whereas HEH populations maintained higher abundance and slower decline, remaining stable or slightly increasing only when typhoon frequency was low-to-moderate ($p \leq 0.5$). These trends align with survival estimates across age classes, and with HEH's more favorable age structure and fecundity profile. The observed divergence highlights the critical role of ongoing, stream-specific ecological support—such as habitat quality and spawning condition maintenance—in sustaining viable populations.

While genetic rescue might appear attractive for isolated or declining populations, our results caution against indiscriminate translocation in deeply diverged lineages. Introducing genetic material from non-local sources can lead to hybrid incompatibility, disruption of co-adapted gene complexes, and genomic swamping^{58,59}. The case of the Chinese giant salamanders, where unregulated stocking erased native lineages through uncontrolled admixture⁶⁰, exemplifies these risks. Thus, any intervention must be informed by both genomic distinctiveness and local demographic context to avoid undermining evolutionary legacy under the guise of recovery.

Rediscovery of native lineages and stream-specific recovery dynamics

Our findings reveal that population recovery in *O. formosanus* varies significantly across its range, with genetic composition differing markedly among streams. Most striking is the case of Hehuan Creek, previously assumed devoid of native salmon since the 1960s due to overharvesting⁵⁶. Contrary to this assumption, our whole-genome sequencing uncovered a genetically distinct lineage consistent with a surviving native population. This discovery coincided with a reintroduction program initiated in 2017²⁵, which contributed to Hehuan's population increase. However, genomic signals—elevated heterozygosity, reduced inbreeding coefficients, and unique allelic variants—indicate that this growth cannot be attributed to stocking alone. Instead, our data suggest a synergistic scenario where reintroduced individuals may have encountered and interbred with remnant natives under favorable ecological conditions. This interpretation is supported by the recent study²⁴, which demonstrated that Hehuan Creek provides superior environmental conditions for salmon reproduction compared to the long-protected Qijiawan Creek. Similarly, Qijiawan populations have shown demographic recovery over the past two decades, aligning with intensified conservation measures including habitat restoration and regulated fishing. In stark contrast, Luoyewei Creek populations, despite similar census-based recovery, display extremely low genetic diversity and no evidence of increasing effective population size. This discrepancy highlights a key limitation of conventional conservation aquaculture: numerical recovery alone is insufficient to ensure evolutionary viability. As demonstrated in other systems^{61,62}, hatchery programs relying solely on long-term broodstock often fail to maintain genetic diversity, leading to elevated inbreeding and loss of rare alleles. In comparison, approaches based on wild-origin offspring collection consistently preserve broader genetic variation and more accurately represent natural spawner diversity. The Luoyewei case mirrors such outcomes, where demographic recovery masks underlying genetic erosion due to historical reliance on hatchery stock with limited founders. Without revisiting propagation protocols to incorporate alternative strategies—such as wild-origin embryo collection, sibship-based pedigree reconstruction, or genomic monitoring—such populations may continue to erode in adaptive potential despite apparent demographic recovery. These findings underscore the importance of embedding genetic resilience as an explicit conservation goal, particularly in isolated headwater systems where natural recolonization is unlikely and demographic indicators alone may be misleading.

Integrating genomics with ecological context for effective conservation

The differing responses observed among the three stream systems suggest that traditional, uniform conservation approaches may have important limitations when applied across ecologically diverse contexts. Our demographic simulations reinforce this point: under simulated typhoon disturbances, HEH shows markedly greater persistence, but still declines under high-frequency events ($p \geq 0.7$), whereas QIJ collapses under every scenario studied (Fig. 5c–d). These divergent outcomes likely stem from differences in genetic diversity, local adaptation, and life-history traits shaped by stream-specific ecological conditions—not simply by broad variables like elevation or streamflow. Qijiawan’s vulnerability under high disturbance regimes stems not only from reduced genetic variability, but also from demographic constraints: lower juvenile survival, decreased fecundity, and narrower terminal age class representation. These factors collectively reduce its ability to recover following extreme events, as clearly observed in the modeled outcomes. As such, ecological context must inform genetic interpretations—demographic plasticity and reproductive capacity are critical buffers that genomic data alone cannot predict. Such findings are part of a broader pattern repeatedly observed in salmonids: local adaptation is both common and consequential. Numerous studies have shown that even fine-scale environmental variation can drive significant divergence in traits critical for survival and reproduction. For example, Fraser River Sockeye Salmon populations facing more challenging migration routes exhibit enhanced aerobic scope and cardiac performance⁶³, demonstrating physiological traits adapted to local demands. Similarly, thermal physiology in Pacific salmonids varies widely among populations, reinforcing the need for context-dependent conservation approaches rather than generalized prescriptions⁶⁴.

Hehuan Creek provides a compelling case: despite historical overharvesting, it now supports a genetically distinct and resilient population. Genomic data reveal elevated heterozygosity, low inbreeding, and unique allelic variants—signals not fully explained by recent reintroduction efforts²⁵. Instead, these patterns point toward a hybrid population formed through interaction between remnant natives and reintroduced fish, thriving in an ecologically favorable habitat²⁴. This underscores the importance of integrating ecological niche data with genomic signals to avoid underestimating the recovery potential of marginal populations. Yet, local adaptation also complicates well-intentioned interventions. A recent reciprocal transplant experiment in brown trout populations by Labonne et al.⁶⁵ found little evidence of genetic rescue and no consistent fitness advantages for foreign or hybrid individuals, suggesting that gene flow alone does not guarantee improved outcomes. These results caution against assuming that translocations or supplementation will necessarily benefit recipient populations, especially when local selection regimes are strong or not fully understood. Taken together, these findings argue for conservation strategies that are highly tailored, genetically informed, and ecologically grounded. Preserving intraspecific diversity requires more than demographic stabilization; it demands recognition of local evolutionary trajectories and the complex interplay between genes, environment, and behavior.

Materials and Methods

Sample Collection

For whole genome sequencing, a female adult (OIJ-BRCnpT, Table S1) and a female subadult (CHI-04-2409-01, Table S1) of Taiwanese salmon, *O. formosanus*, were collected from the Qijiawan Creek (24°23′58.5″N 121°18′19.0″E), Heping District, Taichung City, Taiwan. The brain, eye, gill and odor tissues were excised from fish body and preserved in liquid nitrogen or RNAlater™ Stabilization Solution (Cat: AM7021, Invitrogen, Thermo Fisher Scientific Inc.) at –80°C freezer. The muscle tissues were preserved in 100% ethanol or liquid nitrogen at –80°C freezer.

To investigate nucleotide diversity and population genetics of Taiwanese salmon, each fresh fin tissue of 44 Taiwanese, three Japanese amago and two Japanese Nagano Masu individuals was collected and sequenced via ddRAD-seq (Table S1 and Fig. S1). Each captured salmon was anaesthetized with 2-phenoxyethanol and immediately released to the original habitats after fin clipping. Six of 44 Taiwanese individuals (QIJ_02_04-, QIJ_02_15, LUO_07, HEH_01_13, HEH_W1_24, HEH_02_112, QIJ-BRCnpT) and five Japanese samples were further selected for Illumina PE-150 sequencing. All experiments in this study were performed in accordance with guidelines of the animal ethics committee and were approved by the Institutional Animal Care and Use Committee, Academia Sinica (IACUC 24-09-2284) and collection permissions of Ministry of Agriculture (MOA) and National Park Service (NPS), Ministry of the Interior.

DNA and RNA sequencing

For genome sequencing, 30–50 mg of muscle tissue was lysed with proteinase K and Biofluid & Solid Tissue Buffer from Quick-DNA™ HMW MagBead Kit (Catalog No. D6060, Zymo Research Corp., USA), following the manufacturer's instructions. For ddRADseq, 2–10 mg of fin tissue was lysed with proteinase K and BL2 Buffer from MagPurix Tissue DNA Extraction Kit (Catalog No. ZP02004, Zinexts Life Science Corp., Taiwan), following the manufacturer's instructions. The eluted genomic DNA were stored at –20°C before sequencing.

For PacBio long-read sequencing, the genomic DNA was enriched for fragments larger than 10 kb using the Short Reads Eliminator Kit (PacBio). DNA purity was checked using SimpliNano™ - Biochrom Spectrophotometers (Biochrom, MA, USA). DNA concentration was determined by Qubit 4.0 fluorometer (Thermo Scientific) and the fragment size was monitored by the Qsep 100TM system (Taiwan). The SMRTbell library was prepared according to the Preparing whole genome and metagenome libraries using SMRTbell® Prep Kit 3.0 procedure (PacBio). In brief, 1 µg of high molecular weight genomic DNA was digested with DNase to remove single-strand overhangs, following end repaired and added dA- tail using KAPA End Repair and A-tailing reagent (KAP HyperPlus Kit, KAPA Biosystems, Roche). The Overhang Adapter was introduced onto double-stranded DNA fragments by ligation steps. After purification with AMPure PB beads to remove the adapter dimer, the SMRTbell library was incubated with sequel II Binding Kit 3.2 for the primer annealing and polymerase binding. At last, sequencing was performed in the HiFi sequencing mode on a PacBio Sequel IIe instrument to generate the subread.bam raw data for 30 hours movie time. It was then sequenced through the services of the BIOTOOLS Co., Ltd. and the High Throughput Genomics Core Facility at Academia Sinica, Taiwan.

For Illumina short-read sequencing, genomic DNA was sheared using a Covaris 2 machine, and then quantified using a Qubit 2.0 Fluorometer. A short-read sequencing library was also constructed from 200ng of the sheared gDNA using the TruSeq Nano DNA High Throughput Library Prep Kit (Cat. No. 20015965). The PE-150 reads were then sequenced using an Illumina NovaSeq X Plus through the Genomics BioScience & Technology (New Taipei City, Taiwan). In addition, the high-resolution chromosome conformation capture (Hi-C) libraries were prepared and sequenced using Illumina NextSeq 2000 at the High Throughput Genomics Core Facility at Academia Sinica, Taiwan.

For Illumina RNA sequencing, five to 30 mg of brain, eye, gill and odor tissues were excised and homogenized with 600 µl Trizol solution for RNA sequencing. RNA was extracted from the homogenized Trizol solution using RNeasy® Plus Mini Kit (Cat: 74134, Qiagen), and quantified using the Agilent 2100 Bioanalyzer. A short-read sequencing library was constructed from ~1.5µg of the quantified RNA using the TruSeq Stranded mRNA Library Prep Kit (Cat. No. 20020595). It was then sequenced on the Illumina Novaseq X Plus through the services of GENOMICS BIOSCIENCE & TECHNOLOGY CO., LTD, Taipei, Taiwan.

Male-specific marker detection

For sex determination, polymerase chain reaction (PCR) using male-specific primer pair OtY2m (5'-TCAATCTGTGACGTCACTCA-3' and rev: 5'-GGCTTACCGTCCCAAGTAT-3')⁶⁶ were conducted on 50–100 ng of genomic DNA in 25-µl reactions from each individual of the *O. masou* complex with the

following composition: 12.5 μ l of GoTaq G2 Green Master Mix (M7822, Promega), 2 μ l of each OtY2m (5 μ M, forward and reverse) primer, 6.5 μ l of dd water and 2 μ l of genomic DNA. The amplification program was 94 °C for 2 min; followed by 35 cycles of 94 °C for 1 min, 52 °C for 1 min, and 72 °C for 1 min; followed by 72 °C for 2 min. The final products were analyzed using 2% agarose gel for 30-min electrophoresis and fragment visualization.

Genome assembly and gene annotation

Hifiasm (ver 0.19.9)⁶⁷ was used to assemble the raw PacBio Hifi reads and haplotigs were removed using HaploMerger2 (ver. 20180603)⁶⁸. The assembly were further scaffolded using HiC reads with YaHS²⁷ and subsequently curated in Juicebox⁶⁹. Repetitive DNA sequences were identified and masked using EarlGrey⁷⁰. The proteomes of 11 Salmons were obtained from NCBI (listed in Table S2³²). Transcriptome reads of eye, brain, gill and odor were mapped to the genome assemblies using STAR (ver. 2.7.7a)⁷¹. The gene models were predicted using BRAKER3 (ver 3.0.8)⁷² with proteomes and RNA-seq mappings as evidence hints. Protein families and GO terms were annotated using Pfam scan (<https://ftp.ebi.ac.uk/pub/databases/Pfam/Tools/>³³) and eggnoG mapper³³.

Phylogenetic tree and estimation of divergence time

Protein datasets from eight representative salmon species were download from NCBI. Orthogroups (OGs) were identified using OrthoFinder⁷³. A total of 1,483 single copy orthologs were used to construct a phylogenetic tree using IQtree⁷⁴ with the model JTT+F+I+R7. Gene family expansions and losses were inferred using CAFÉ⁷⁵. The divergence time of each tree node was inferred using MCMCtree of the PAML package⁷⁶. The divergence time estimates were based on a known fossil record that gave an estimated divergence date of 6 MYA between *O. nerka* + *O. gorbuscha* and *O. keta* and 15-20 MYA as the divergence time between *Salmo-Oncorhynchus* genera³⁶.

Whole-genome mapping and genotype calling

Whole-genome sequencing was conducted on two Amago, three Nagano Masu, and six additional *O. formosamensis* individuals (Illumina platform, System). Raw paired-end reads were subjected to quality trimming using fastp (ver. 0.24.0)⁷⁷ with the following parameters “-q=30, -l=140, -5 -M=30 -cut_front_window_size=1, -r -M=30 -cut_right_window_size=4”. Subsequently, the processed reads were mapped to the reference genome assembly using BWA v.2.2.1⁷⁸. Reads that were duplicated in the BAM files would be tagged with ‘DT’ using Picard MarkDuplicates of GATK v4.4.0.0⁷⁹.

Variant calling of each salmon individual was performed using bcftools v1.19⁸⁰ mpileup and call commands for minimum mapping quality of 30. Both monomorphic and polymorphic sites with read coverage were preserved in the initial output and subsequently merged to create one VCF file containing all individuals and all sites. The identification of biallelic single nucleotide polymorphisms (SNPs) was achieved through the use of ‘-m 2 -M2 -v snps’ options, followed by filters based on the criteria of ‘AC $\geq$ 2 & F_missing $\leq$ 0.1’. To account for the variation in individual sequencing depth between studies, loci with sequencing depth less than 2 in more than 30% of the samples were excluded. Tentative false positive heterozygous SNPs exhibiting depth exceeding 175% of the individual genomic median were masked as missing genotype. The final dataset comprised 31.6 million SNPs after the above filtration process.

Population genomic analysis

Principal component analysis was performed on the SNPs from the 25 largest scaffolds through PLINK v2.00a6LM⁸¹. Population structure was analyzed using ADMIXTURE v1.3.0⁸² with fivefold cross-validation of K values from 2 to 8. According to the topology of the species tree, two *Oncorhynchus kisutch* individuals (SAMN36289694, SAMN36289695)⁸³ were incorporated to the phylogeny analysis as outgroups to the *O. masou* species complex. The mapping rates of *O. kisutch* samples to the *O. formosamensis* genome assembled in this study are higher than 98%. The best-fit nucleotide evolution model, TVMe, was determined by IQTREE v2.2.6^{74,84} based on the loci of

filtered SNPs. A maximum-likelihood phylogenetic tree was then inferred on a 1000 ultrafast bootstrap approximation. The effects of SNP variants on protein-coding gene sequences were further annotated and classified as sites of loss of function (LoF) (gain or loss of stop codon and loss of start codon), missense, and synonymous (synonymous variant, start retained, stop retained variant) variants using SnpEff (ver. 5.2)⁸⁵.

The demographic history was estimated using PSMC⁸⁶. An additional three Illumina-sequenced individuals from Christensen *et al.*²⁸ were mapped to genome assembly using bwa-mem (ver. 2.2.1)⁸⁷ and the aligned mapped reads were converted to consensus sequence in FASTQ file using bcftools pipeline. The minimum (d) and maximum (D) coverage for vcftools were set to 2 and 1.75X genome reads coverage. All of the parameters used for the PSMC program were at default with the exception of $-\mu$ 3.0e-09⁸⁸ and $-g$ 1⁸⁹.

ddRADseq libraries preparation and data processing

Double-digest restriction-site-associated DNA sequencing (ddRADseq) was used to investigate the evolutionary history of Japanese and Taiwanese masou salmon populations, as well as the genetic diversity and demographic history of the Formosan salmon in Taiwan. Library preparation followed established protocols^{90,91}. Briefly, genomic DNA from each sample was digested using the restriction enzymes EcoRI and MseI (New England Biolabs, Massachusetts, USA), with ~200 ng of DNA per individual. Adaptors containing unique barcode sequences were then ligated to the digested DNA fragments using T4 ligase (New England Biolabs). The barcoded DNA samples were individually purified using 1.8× AMPure XP magnetic beads (Beckman Coulter). PCR with eight cycles was performed on the purified fragments to incorporate Illumina sequencing adaptors. The resulting PCR products were pooled based on DNA concentration and underwent an additional round of magnetic bead-based size selection and purification. DNA fragments ranging from 300 to 500 base pairs (bp) were selected using 0.5× and 0.7× AMPure XP magnetic beads. In total, 83 samples were collected, including five Japanese samples (including two individuals of *O. m. ishikawae* (Amago) and three individuals of *O. m. masou* (Nagano masu) and 78 Taiwanese samples. The Taiwanese samples were distributed as follows: 21 from the Qijiawan creek, 17 from the Luoyewei creek, and 40 from the Hehuan creek. These samples were prepared into two sequencing libraries. Single-end sequencing with a read length of 150 bp was performed on an Illumina NovaSeq X platform at Tri-I Biotech, Inc. (New Taipei City, Taiwan).

The sequencing data were processed using the STACKS v2.62 pipeline^{92,93}. First, the process_radtags program in STACKS was used to demultiplex the sequencing data with default settings. Reads were retained if they contained an intact EcoRI RAD site, had a unique barcode, and did not include Illumina adapter sequences. The trimmed reads were then mapped to the *O. formosanus* reference genome using the reference-mapping process in STACKS. Specifically, the reference genome was indexed using the bwa index command⁹⁴, and each sample was aligned using the bwa mem mode. The mapped results were then compressed while preserving the headers into individual BAM files using Samtools⁸⁰.

Following the read alignment, two datasets were generated for downstream analysis. The first dataset included both Japanese and Taiwanese samples and was used to examine the evolutionary history of Japanese and Taiwanese masou salmon. The second dataset contained only Taiwanese samples and was used to estimate the genetic diversity and demographic history of the Formosan salmon. Loci were assembled using the ref_map.pl wrapper script in STACKS, and the populations module was used to export the results. To ensure data quality, loci were retained only if they were present in at least 80% of individuals per population, shared across all populations, and had a minor allele count of at least three.

Population tree estimation and species delimitation

We used the Bayesian Phylogenetics and Phylogeography program (BPP, version 4.7.0)^{37,95} to estimate a species tree at the population/taxon level, including divergence times, and to delimit evolutionarily independent entities under the multi-species coalescent model^{38,96}. Given that our sampling design did not include outgroup samples, we first used the BPP program to estimate the

species tree topology while simultaneously delimiting evolutionarily independent entities using the A11 model. Next, we fixed the best estimated species tree topology and used it to estimate genetic diversity (theta) for each terminal taxon, as well as the ancestral node, and coalescent times (tau) for each internal node, applying the A00 model.

A time-calibrated diversification history was then generated, using a generation time of 1 year and a spontaneous genomic mutation rate of 1.06×10^{-8} bp/generation⁹⁷, via the bppr package (<https://github.com/dosreislab/bppr>) and the results from the A00 model of BPP analyses. Furthermore, the estimated theta and tau values were subsequently used to calculate the genealogical divergence index (GDI)⁹⁸ for different divergence events on the species tree. The GDI values have been shown to capture nuances that represent various stages of the speciation continuum, providing a more evolutionarily informed framework for species delimitation^{98,99}. A practical application of the GDI was recently carried out in a large-scale study on darter fishes in the U.S. for conservation purposes⁴⁰.

To facilitate the tractability of computational power while ensuring the repeatability and reproducibility of our results from BPP analyses, we generated three independent datasets for each BPP analysis, with each dataset containing 3000 randomly selected RAD loci from across the genome. We consistently included the three Amago and two Nagano Masu samples. For the Formosan salmon populations, we randomly included three individuals from the Qijiawan population and three individuals from the Hehuan population, for a total of six individuals. To generate control files for the BPP analyses, we used an online tool provided by the authors of the BPP program (<https://brannala.github.io/bpps/>). This online tool also helped with specifying appropriate prior settings for the theta and tau parameters based on the empirical sequence data. The BPP analyses were run for 5×10^5 generations, with samples taken every 5 generations, and a burn-in period of 1×10^4 generations was applied. The convergence of the analyses was assessed by ensuring that the effective sample sizes (ESSs) for all estimated parameters were greater than 200.

Population genetic structure of the Formosan salmon

To investigate the genetic relationships among Formosan salmon populations, we conducted a principal component analysis (PCA). First, we imported VCF files into R using the read.vcfR function from the vcfR package¹⁰⁰. To mitigate the effects of linkage among SNPs, we filtered SNPs using the distance_thin function in the SNPfiltR package¹⁰¹, setting a minimum distance of 1 Mb (min.distance = 1,000,000). Subsequently, the filtered SNP data were converted into genind objects using the vcfR2genind function, with populations defined based on eight collection sites across three streams. To handle missing data, we applied the scaleGen function from the adegenet package¹⁰², replacing missing values with the mean (NA.method = "mean"). Finally, we performed PCA on the transformed genind dataset (center = TRUE, scale = TRUE) using the dudi.pca function from the ade4 package¹⁰³.

Estimating the genetic diversity and level of inbreeding in the Formosan salmon

To calculate the genetic diversity of all the individuals across the three stream populations in Taiwan, we assessed the multi-locus heterozygosity (MLH)¹⁰⁴ using the MLH function within the inbreedR package¹⁰⁵ in R. The estimated MLH values were then visually represented using bar charts, with each bar representing the heterozygosity of each individual.

To estimate runs of homozygosity (ROHs) for each individual, we used PLINK v1.9¹⁰⁶ to calculate the number and length of ROHs (>1000 Kb) across the genome. ROHs were estimated using two different window sizes, 5 and 10 SNPs (--homozyg-window-snp 5; --homozyg-window-snp 10). The parameters for ROH estimation were adapted from previous studies^{107,108} as follows: allowing one heterozygous SNP per window (--homozyg-window-het 1), permitting one missing call per window

(--homozyg-window-missing 1), and setting the percentage of overlapping windows required to classify a SNP as part of a homozygous segment to 1% (--homozyg-window-threshold 0.01). Finally, segments were considered homozygous if they exceeded 1000 Kb in length (--homozyg-kb 1000).

Estimating current effective population sizes and reconstructing the demographic history

To estimate the current effective population size of the three geographical taxa, we applied NeEstimator V2¹⁰⁹ using the *glLDNe* function¹¹⁰ in dartR package in R^{111,112}. The genlight objects for downstream analyses were imported from the VCF files using the *gl.read.vcf* function in dartR, with populations defined by the three streams. To account for the linkage among SNPs from the same chromosome, we specified the argument *pairing* in the *glLDNe*, which limited the use of paired loci only when they were from different chromosomes when estimating the effective population size. We estimated the effective population size both with and without empirical corrections by using the *naïve* argument in the *glLDNe*. The effective population size for each stream populations was estimated under the assumption of random mating within each population, while considering only alleles with a frequency higher than 0. Specifically, we do not observe genetic structure within populations, and the populations appear to be relicts of hypothesized recent local extirpation events; therefore, assuming a panmictic population, or random mating system, is biologically appropriate.

To reconstruct the recent demographic history of Formosan salmon from the Qijiawan and Hehuan Streams, we applied a linkage disequilibrium-based method implemented in GONE¹¹³. The population from the Luoyewei Stream was excluded due to its high level of inbreeding, as excessively long ROHs would lead to erroneous estimates in the software. The sequence data were treated as unknown phase, and the estimation window was set to 100 generations, divided into 100 time bins, with each bin representing one generation. The maximum recombination rate between analyzed SNP pairs was set to 0.05 ($hc = 0.05$), and a maximum of 50,000 SNPs per chromosome was analyzed. For the GONE analyses, we used a generation time of 1 year and the average recombination rate of salmon (1.09 cM/Mb)¹¹⁴. Additionally, to assess run-to-run variability in GONE results, we conducted 100 independent runs using the average recombination rate of salmon.

Non-lethal field sampling and age estimation

Given the endangered status of Formosan landlocked salmon (*Oncorhynchus formosanus*) and strict legal restrictions on lethal sampling, we used low- and non-invasive monitoring techniques from 2021 to 2024. Between 2021 and 2023, individuals were sampled using hook-and-line fishing; however, this method exhibited size selectivity that could bias estimates of population structure. To address this limitation, in 2024 we switched to snorkel-based underwater visual surveys with scale-referenced imaging to non-invasively estimate body length distributions. During the late breeding season (November–December 2024), individuals were collected for morphometric and age data using non-lethal methods: night snorkeling capture in Chichiawan Creek (QIJ) and electrofishing in Hehuan Creek (HEH). Sampling periods, methods, and summaries of metrics are detailed in Table S12¹¹⁵ and Table S13¹¹⁶.

For masu salmon, scales have been shown to provide age estimates that are not only consistent, but in many cases more accurate than those derived from otoliths^{115,116}. Due to legal prohibitions on lethal sampling for otolith extraction (Taichung City Agricultural Bureau 2021), we used scales as the primary material for age inference. Previous validation experiments report scale-based age accuracy between 90–96% using standard sampling protocols¹¹⁶. We collected scales from the preferred body region (posterior to the dorsal fin and anterior to the anal fin, above the lateral line, in the 2nd–3rd scale rows^{115,117}). After rinsing and air drying, five good quality scales per individual were selected. These were cleaned and pressed using a roller device to produce acetate impressions. The rings were examined by reflected light microscopy at 30–40x. Annual rings were

identified based on abrupt transitions from widely spaced summer rings to densely packed winter rings. False rings and spawning checks were excluded from age counts according to [115,116](#). A single trained reader performed all scale readings to ensure internal consistency.

Observed lengths from underwater surveys were corrected for light refraction by dividing raw values by the refractive index of water (1.33). To further correct for detection bias across size classes, paired snorkel and electrofishing surveys were conducted at HEH in 2024. Electrofishing, assumed to be less biased across size classes, provided a reference distribution. Length-specific detection ratios were calculated by comparing the frequency distributions of the two methods in overlapping size bins (e.g., 37.6–75.2 mm, 75.2–112.8 mm). Ratios < 1 indicated underestimation by snorkeling, and snorkel-based length counts were adjusted accordingly. Corresponding detection distributions and correction curves are shown in [Figures S15](#) and [S16](#).

Snorkel observations provided the number of fish within specific length intervals ([Fig. S17](#)), but did not provide individual-level measurements. To reconstruct individual body lengths, we implemented a probabilistic sampling approach. For each observation of N individuals within a length interval $[L_{\min}, L_{\max}]$, N measurements were randomly sampled with replacement from the corresponding donor pool of scale-measured individuals. Sampling was constrained by stream and sampling period (e.g., only samples from HEH_{11–12} were used to reconstruct HEH November–December data). When sample sizes in a donor pool were insufficient, sampling was extended incrementally to similar intervals. This resulted in a reconstructed distribution of individual lengths for each snorkel dataset.

Estimation of demographic parameters

Because age-specific capture probability (q_{age}) cannot be measured directly, we adopted a pragmatic two-gear strategy in which backpack electrofishing served as the least age-biased reference method. Independent field studies on salmonids demonstrate that electrofishing efficiency rises with body size but plateaus at a high, nearly constant level once fish pass the early–juvenile stage [118,119](#), rendering capture probability effectively age-invariant for fully recruited age classes. We therefore used electrofishing counts to approximate an equal-catchability age distribution and, for the same stream reaches, calculated length-specific detection ratios between electrofishing and snorkel surveys. These ratios were applied as correction factors to snorkel counts, compensating for the under-representation of the smallest size classes. After correction, residual differences in q_{age} among fully recruited ages were negligible, satisfying the Chapman–Robson assumption of equal catchability.

To assign ages to individual fish, we implemented a two-step procedure that propagates uncertainty in the length–age relationship. First, 1,000 bootstrap resamples were drawn from the empirical length–age data set to obtain a distribution of the mean (μ_i) and standard deviation (σ_i) for each age class ([Fig. S18](#)). For Qijiawan (QIJ), which lacked prior age data, we used maximum likelihood: for each fish of length l we evaluated $\text{dnorm}(l; \mu_i, \sigma_i)$ over all ages and assigned the age with the highest likelihood. For Hehuan (HEH), where electrofishing supplied reliable priors, we adopted a Bayesian approach in which posterior probabilities were proportional to $\text{dnorm}(l; \mu_i, \sigma_i) P(\text{Age} = i)$. Repeating this across the 1,000 resamples yielded age-class proportions and 95 % confidence intervals for every site–month combination ([Table S9](#)).

Three demographic quantities were then estimated: (i) annual recruitment of age-0+ fish, (ii) seasonal survival rates (January → November and November → June), and (iii) age-specific fecundity coefficients. All site-level quantities were first calculated on an area-specific basis (ind. 1000m⁻²) and then aggregated to the reach or stream scale as area-weighted means. Seasonal, age-specific survival was estimated by tracking the same cohort between successive surveys (matching age i at the first survey to age $i + 1$ at the next), rather than by cross-sectional comparisons among ages. We analysed two intervals: January → November (10 months; typhoon season) and November → June (7 months; baseline). Site-level estimates were aggregated to the stream scale using area-weighted means, then standardized to monthly and annual scales to make intervals comparable. Because survival can vary with age, all estimates were computed and reported by age class. Estimates were flagged as unreliable when extremely low or high (below 0.05 or above 0.95)

or when the bootstrap 95% CI was very wide (width > 0.60). Flagged cells were replaced with values sampled from the empirical distribution of age-2 survival in the paired stream during the same season, preserving seasonal and age structure while avoiding unstable ratios. Replicate-level results were summarized as stream means with 95% confidence intervals.

To contrast population structure before and after the 2024 typhoon sequence, we reconstructed age vectors for January 2024 (baseline) and January 2025 (post-typhoon) by temporal back-tracking. The baseline series was anchored to the last comprehensive survey in June 2024, preceding the typhoon season. A monthly baseline survival rate (S_{base}) was extracted from mark-recapture data collected between November 2023 and June 2024, and the June abundance vector was propagated backwards five months,

$$N_{2024-01} = N_{2024-06} \times S_{\text{base}}^5$$

The post-typhoon series used the November–December 2024 survey as its anchor; a winter survival rate (S_{winter}) for the November–January interval was estimated and applied for one-month back-projection,

$$N_{2025-01} = N_{2024-12} \times S_{\text{winter}}$$

Counts of fish aged ≥ 3 y—already near the species' maximum longevity—were left unadjusted at both anchor dates.

Recruitment (R) at each anchor month was obtained from a Chapman–Robson catch-curve analysis applied to the reconstructed age distributions. Instead of using the method to estimate mortality, we interpreted the log-linear catch curve intercept (back-extrapolated to age 0) as an index of initial cohort abundance, i.e. relative recruitment strength¹²⁰. This approach relies on the standard catch-curve identity in which the intercept reflects $q \cdot N_0$ under the assumptions of equal catchability and approximately constant mortality across fully recruited ages. Catchability was equalized by using electrofishing as a reference gear and applying length-specific detection corrections to snorkel counts (see above). Because disturbance years can violate strict steady-state assumptions, we explicitly treat R as a relative index rather than an absolute number, using it only for between-stream and between-scenario comparisons. Replicate-level estimates were aggregated to stream-level means and 95% confidence intervals after area-weighting sites.

We then linked recruitment indices to mature-female density through a scaling factor B . Specifically,

$$B = \frac{R_{\text{index}}}{\sum_i (N_i \times p_{m,i})}$$

where N_i is the abundance of age class i , and $p_{m,i}$ is the empirically derived proportion of mature females at age i . Thus, B represents the number of recruits per mature-female equivalent. Age-specific fecundity coefficients were then defined as $F_i = p_m \cdot B$, representing the expected contribution of age- i females to the recruitment index. For QIJ in 2024, a correction factor of 0.719 was applied to account for spatial sampling bias associated with an upstream, high-productivity reach²³. Fecundity parameters were estimated separately for January 2024 (pre-typhoon) and January 2025 (post-typhoon back-track). Importantly, these are index-scale fecundities rather than absolute egg counts, consistent with age-structured modelling that assigns age-specific reproductive weights to drive population projections¹²¹. This framework provides the demographic inputs for the stochastic simulations described below.

Stochastic simulation of population dynamics

To assess the long-term effects of typhoon disturbance on salmonid population viability, we constructed a discrete-time, age-structured simulation model using Leslie matrices. In our simulations, we defined a 'typhoon year' as any year in which at least one typhoon event caused observable impacts on stream habitat and salmon populations during the July–October period, based on historical records from the Central Weather Bureau. We treated typhoon occurrence as a binary variable (present/absent) rather than incorporating intensity gradients, as our

demographic data did not allow for parsing intensity-specific effects. The model incorporated two distinct demographic scenarios representing typhoon and non-typhoon years, with parameters derived from empirical field estimates. The age structure included four classes (0+–3+), with age-specific fecundity coefficients and survival rates parameterized from field data. Fecundity vectors were scenario-dependent: "noTyphoon" scenario used parameters derived from January mature females contributing to July recruitment, while "afterTyphoon" scenario used parameters from post-typhoon population reconstruction. Annual survival rates were calculated from two distinct seasonal periods: baseline survival rates were derived from November → June transitions (7-month period representing non-typhoon conditions), while typhoon-impacted survival rates were estimated from January → November transitions (10-month period encompassing the typhoon season). These period-specific survival estimates were then standardized to annual rates for incorporation into the Leslie matrix framework. Carrying capacity was set at 1,500 individuals per 1,000 m², with proportional scaling applied when simulated populations exceeded this threshold. The simulation proceeded as follows: (1) determination of typhoon occurrence for the focal year; (2) selection of appropriate survival and fecundity parameter sets; (3) construction and application of the corresponding Leslie matrix; (4) density regulation if population exceeded carrying capacity; and (5) transition to the next annual cycle. Initial population structure was derived from November 2024 field observations for both Qijiawan Creek (QIJ) and Hehuan Creek (HEH). Simulations were replicated 1,000 times for each typhoon probability scenario over 20-year periods, generating age-structure trajectories and total population size distributions for comparative analysis.

For every replicate we extracted the dominant eigenvalue (λ) of the corresponding Leslie matrix and summarised its distribution by the mean and 95 % CI to compare 'no-Typhoon' and 'after-Typhoon' scenarios. Quasi-extinction was defined as a total population density ≤ 10 individual per 1,000 m². Extinction counts across 1,000 replicates were compared between streams by Fisher's exact test, yielding odds ratios (OR) with exact 95 % CI. To test for a stream $\times$ typhoon-probability interaction, we fitted a logistic GLM to the binary extinction outcome, omitting $p = 0.9$ to avoid complete separation.

Research permits and ethical compliance

All fieldwork conducted between 2021 and 2024 was approved by relevant regulatory agencies, including the Taiwan Forestry and Nature Conservation Agency (permits: 1132401290, 1130218297, 111701591), Shei-Pa National Park (1110002834, 1130003015), and Taroko National Park (1110003283, 1131005811). All procedures complied with the Wildlife Conservation Act and National Park regulations. Capture methods included non-lethal snorkel surveys, hand netting, fly fishing, and limited electrofishing. All individuals were anesthetized with 2-phenoxyethanol (0.1–0.3 mL L⁻¹), handled within 3 minutes, and released after full recovery. Only two individuals were sacrificed for genomic sequencing under the approval of Taiwan Forestry and Nature Conservation Agency, Shei-Pa National Park and Academia Sinica IACUC (24-09-2284). All research followed the principles of replacement, reduction and refinement (3Rs).

Supplementary Figures

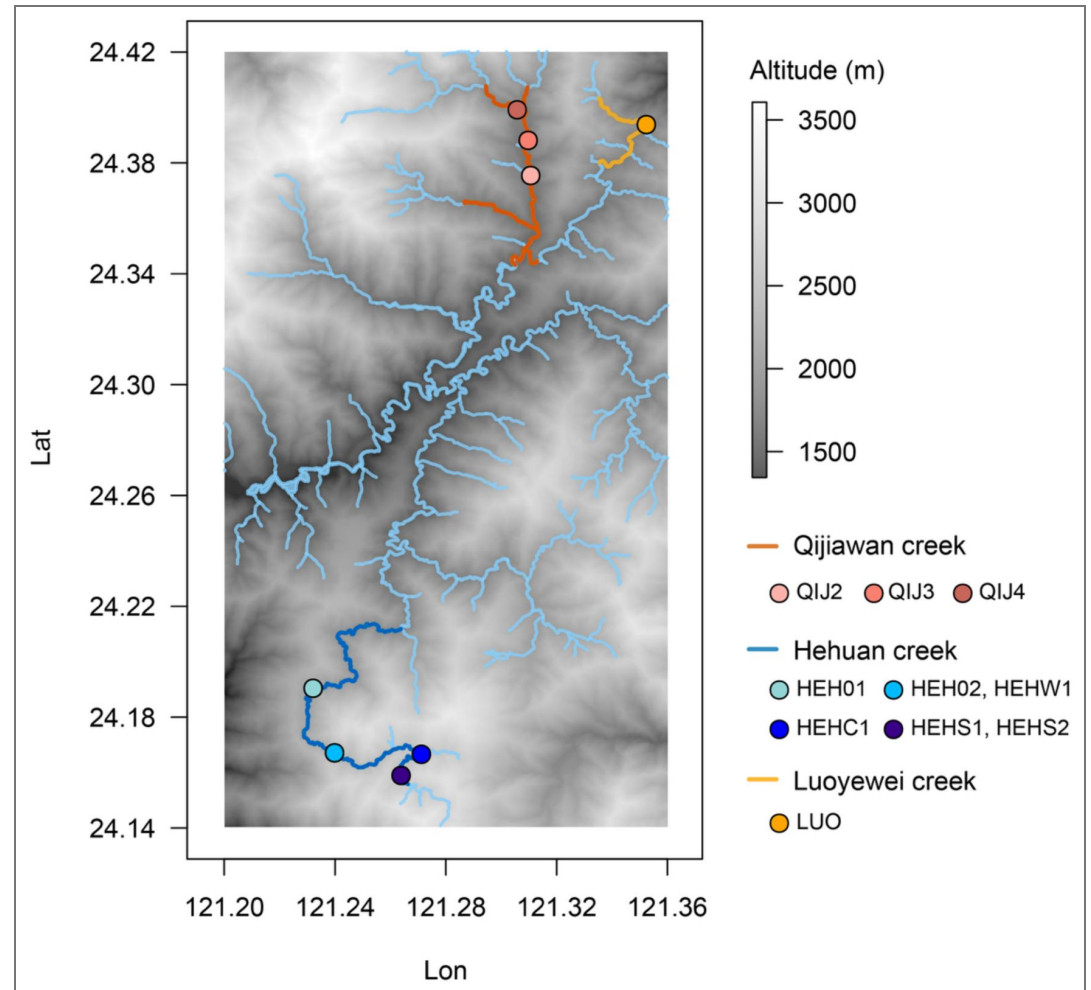


Fig. S1. Geographic distribution and sampling locations of *Oncorhynchus formosanus* populations in Taiwan's Central Mountain Range. The map shows the three creek systems harboring extant *O. formosanus* populations at elevations between 1,500-3,500 m. Qijiawan Creek (orange line) in the northern portion contains the historically protected population with sampling sites QIJ2, QIJ3, and QIJ4. Hehuan Creek (blue lines) in the southern portion includes both upstream sites (HEH01, HEHS1, HEHS2; dark blue circles) and downstream sites (HEH02, HEHW1; light blue circles) where reintroduction occurred in 2017. Luoyewei Creek (yellow line) in the eastern portion represents the hatchery-supplemented population (LUO). The grayscale gradient indicates elevation, with darker shades representing lower elevations. All populations are located within the upper tributaries of the Dajia River basin in central Taiwan.

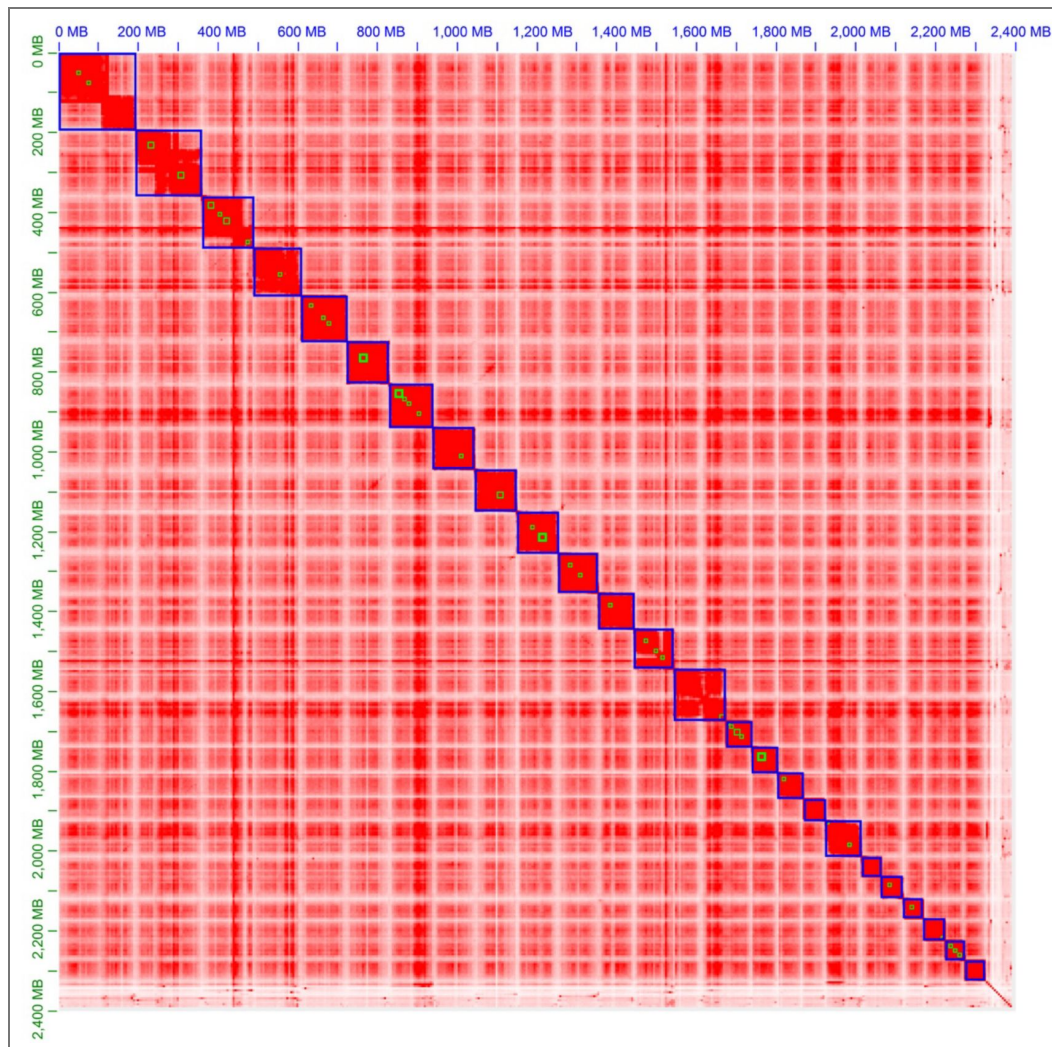


Fig. S2. Genome-wide Hi-C contact map of the *O. formosanus* genome assembly.

The Hi-C interaction heatmap illustrates the genome-wide three-dimensional organization and chromosomal-level scaffolding of the *O. formosanus* genome. Each blue box represents an individual chromosome-scale scaffold identified in this study. Stronger interaction signals are indicated in red, representing frequent physical interactions within chromosomes, while weaker or infrequent interactions appear in lighter colors.

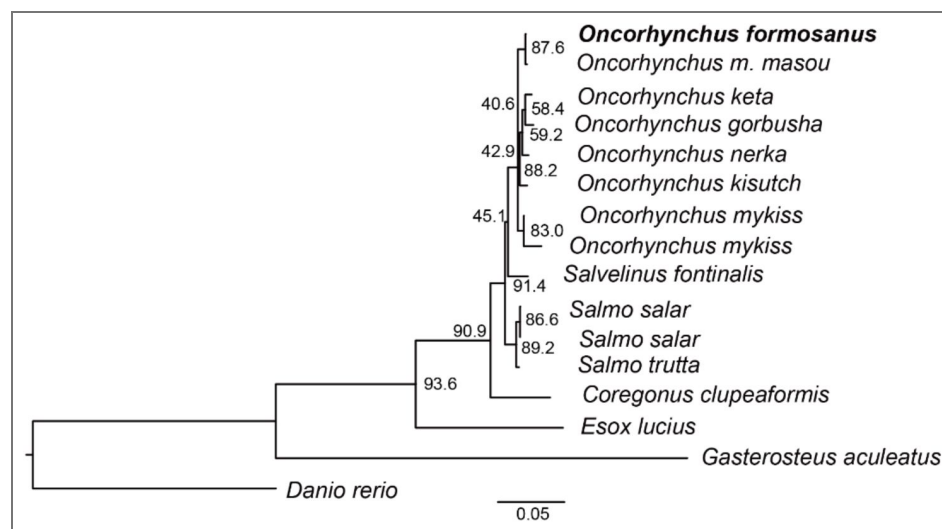


Fig. S3. Coalescent-based species tree inferred by ASTRAL for *Oncorhynchus* species.

Species phylogeny constructed using the coalescent-based approach implemented in ASTRAL from 1,483 single-copy orthologs. Branch support values represent local posterior probabilities, reflecting concordance among gene trees. This analysis provides the same topology and corroborates relationships among *Oncorhynchus* species compared to the species tree from concatenated alignments of Fig. 1b.

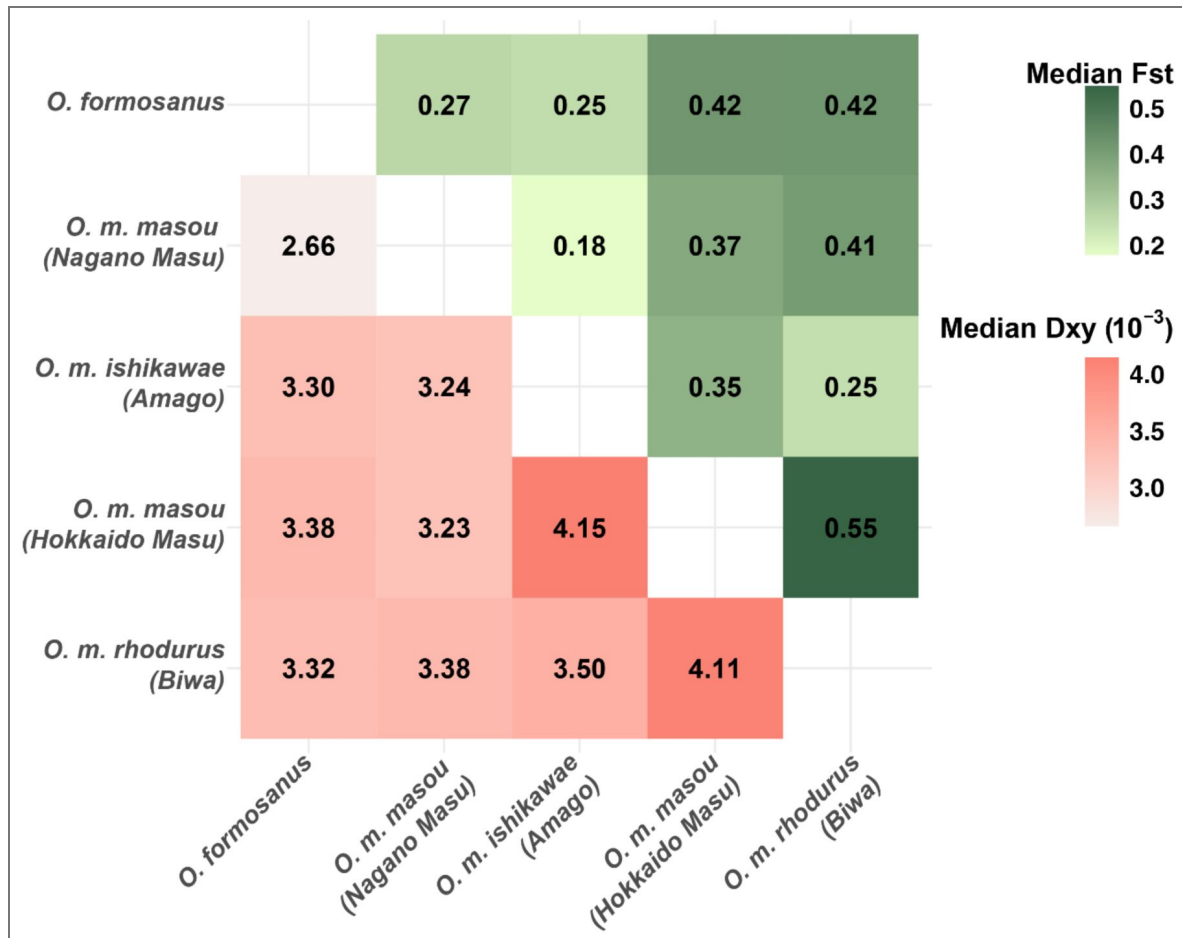


Figure S4. Pairwise genetic differentiation among masu salmon lineages.

Heatmap illustrating pairwise genetic differentiation (Fst, green shades) and absolute nucleotide divergence (Dxy, red shades) among five *Oncorhynchus masou* lineages.

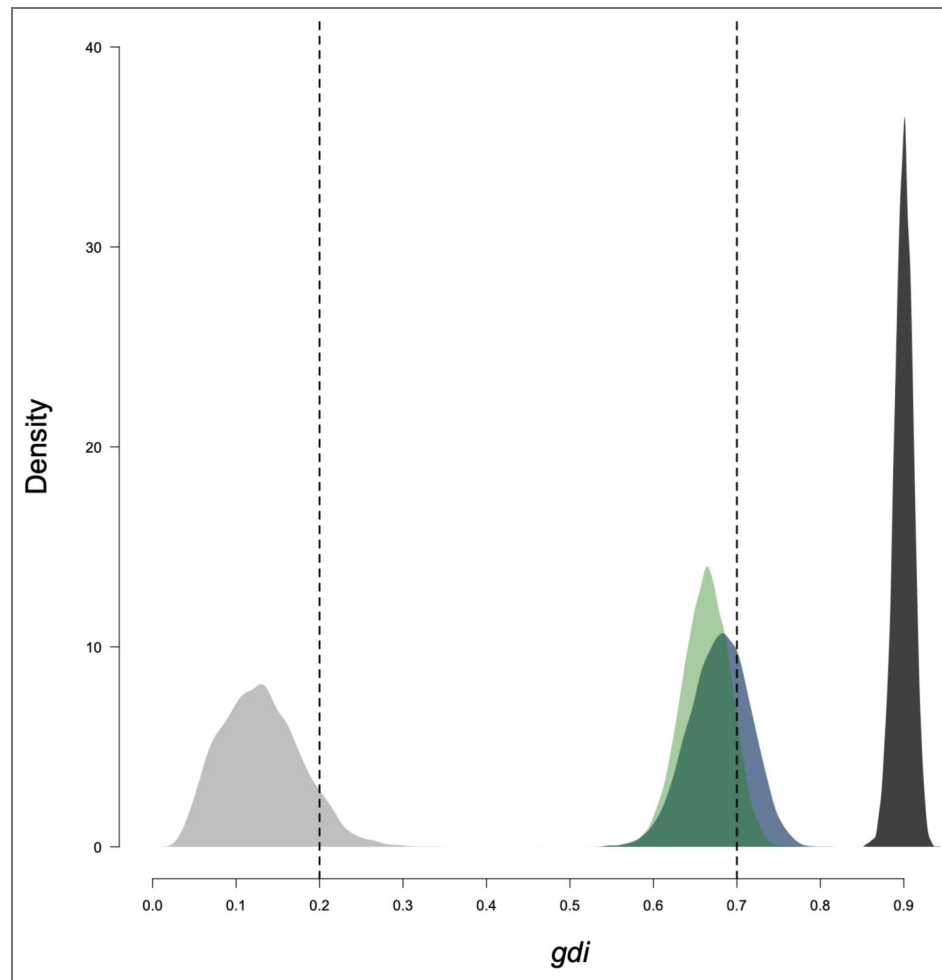


Fig. S5. Genealogical Divergence Index (GDI) analysis of Taiwanese and Japanese salmon lineages.

The genome-wide estimates of genealogical divergence index (GDI) reveals substantial genetic differentiation between Taiwanese (*O. formosanus*) and Japanese (*O. masou*) salmon lineages (the value was around 1.0; dark grey). The estimated GDI values between the two Japanese salmon lineages were highlighted in green and blue colors, where the values were around 0.7. High GDI values (>0.7) indicate strong genealogical divergence, implying long-term isolation and lineage differentiation. In contrast, the estimated GDI between the two Taiwanese population (highlighted in light grey) was smaller than 0.2, which suggests neglectable genetic divergence between the two populations.

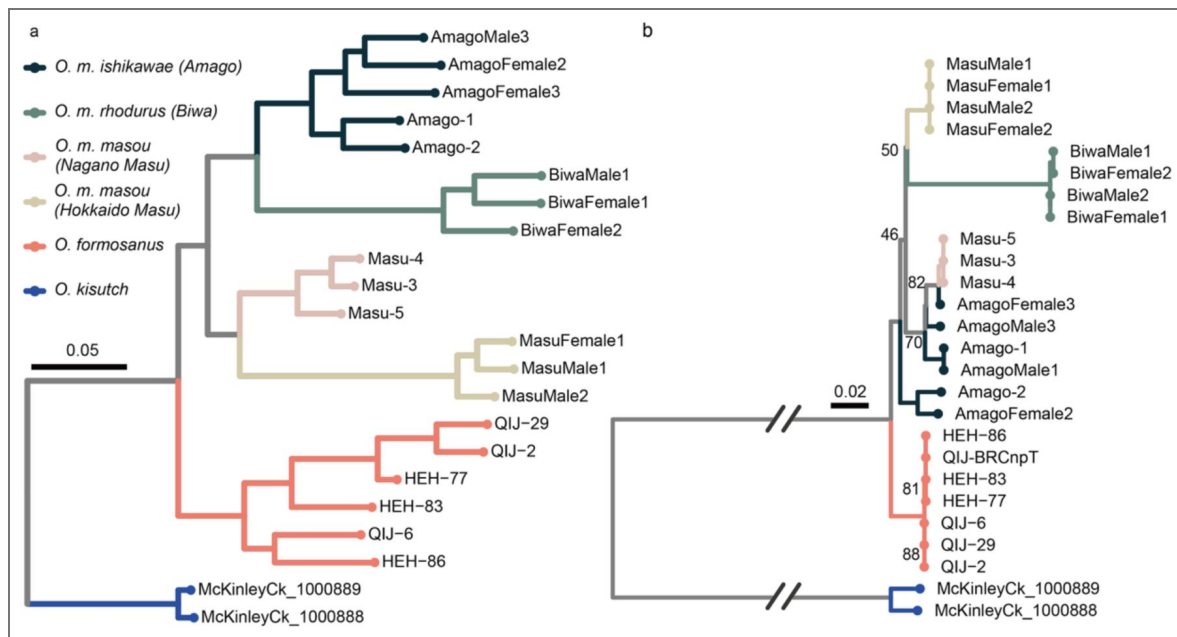


Fig. S6. Population phylogeny of Taiwanese and Japanese salmon lineages inferred from nuclear genome-wide and mitochondrial SNPs.

(a) Maximum likelihood phylogeny based on genome-wide SNP data from representative individuals of Taiwanese (*O. formosanus*: QIJ, HEH) and Japanese (Amago, Biwa, Nagano Masu and Hokkaido Masu) salmon populations, rooted with *O. kisutch* (coho salmon) as the outgroup. Branch lengths represent nucleotide substitutions per site. (b) Corresponding mitochondrial genome phylogeny inferred from mitogenome SNP data, also rooted with *O. kisutch*. Bootstrap support values (percent) for major branches are indicated. Together, these phylogenies illustrate consistent genetic differentiation patterns between nuclear and mitochondrial genomes across salmon lineages.

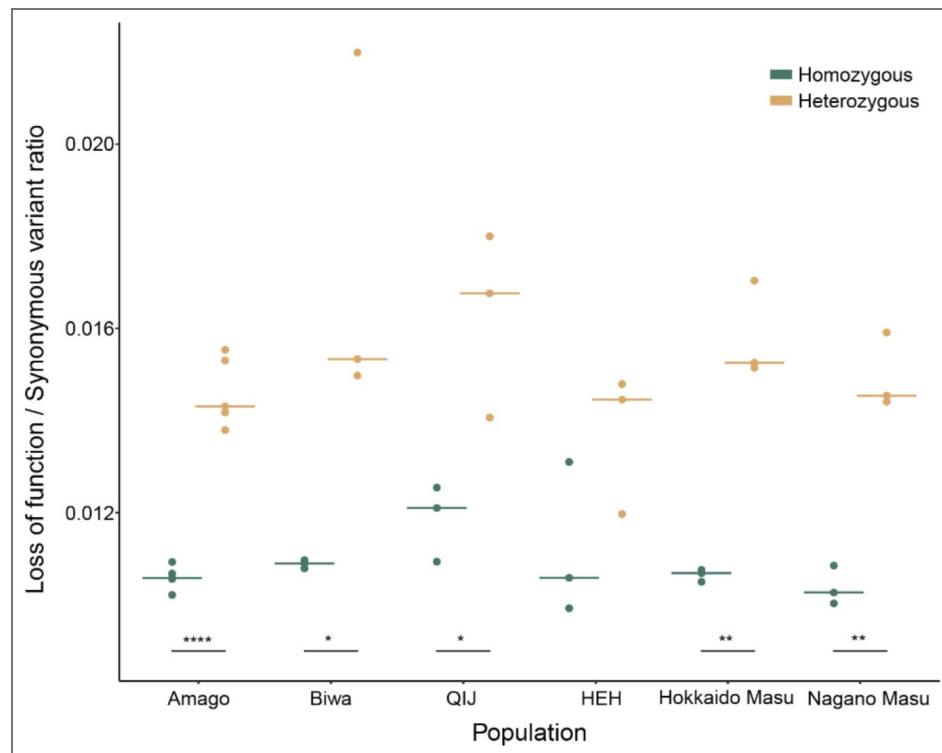


Fig. S7. Relative proportions of loss-of-function to synonymous variants in homozygous and heterozygous states across salmon populations.

Scatter plot showing the ratio of loss-of-function (LOF) variants to synonymous variants in both homozygous (green) and heterozygous (orange) states among salmon populations: Amago, Biwa, QIJ (Qijiawan), HEH (Hehuan), Hokkaido Masu, and Nagano Masu. Horizontal bars indicate mean values for each population. Heterozygous LOF variants consistently exhibit higher ratios compared to homozygous LOF variants across all populations, aligning with expectations of increased accumulation of deleterious alleles and elevated homozygosity in populations with smaller effective population sizes. Asterisks denote statistical significance of the difference between homozygous and heterozygous states (* $P < 0.05$, ** $P < 0.01$, **** $P < 0.0001$).

Figure S8. Comparison of evolutionary rates (dN/dS ratios) across different chromosome categories in *O. formosanus*.

Boxplot showing the distribution of nonsynonymous-to-synonymous substitution rate ratios (dN/dS) for genes located on chromosomes classified as non-fused, fused in the last common ancestor of the *O. masou* species complex, or specifically fused in *O. formosanus*. "ns" denotes no significant difference, ** denote $P < 0.01$ in Wilcoxon rank sum test.

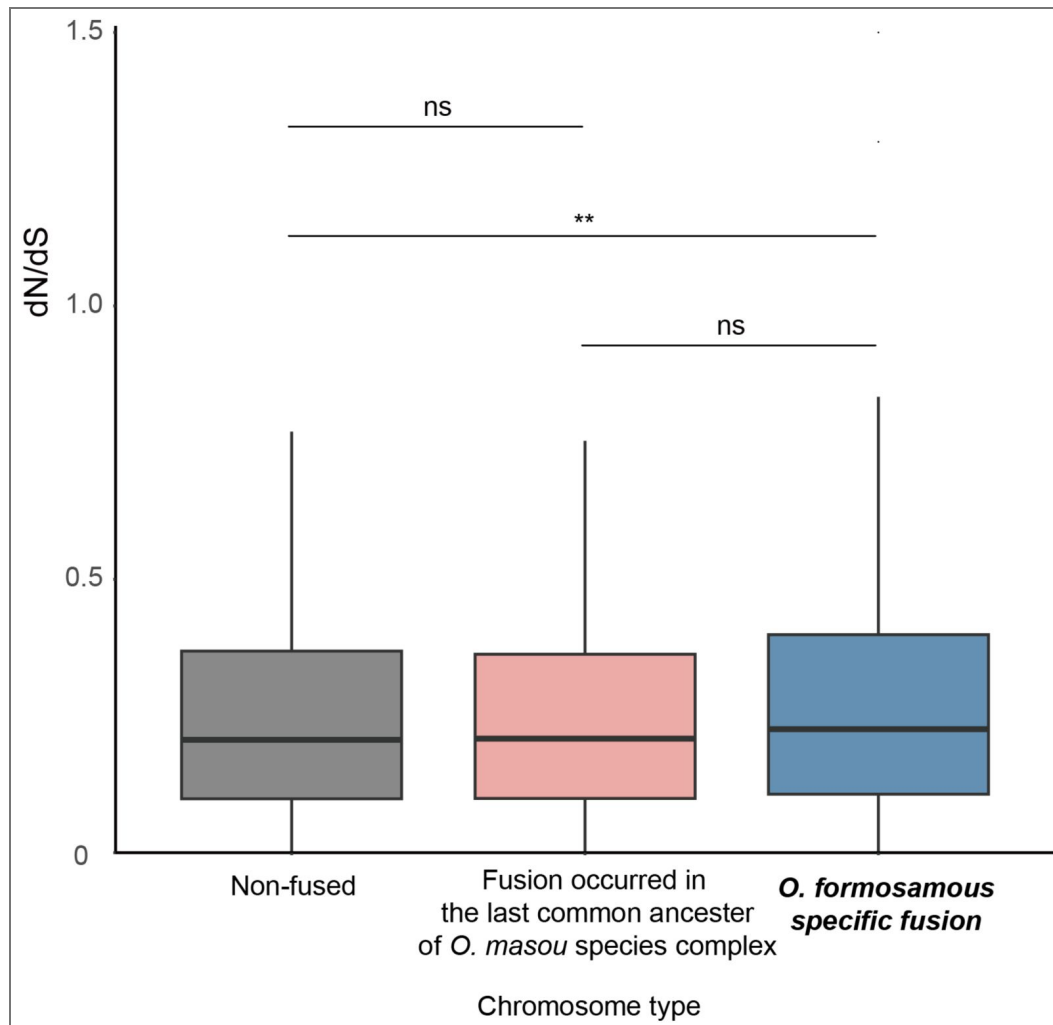
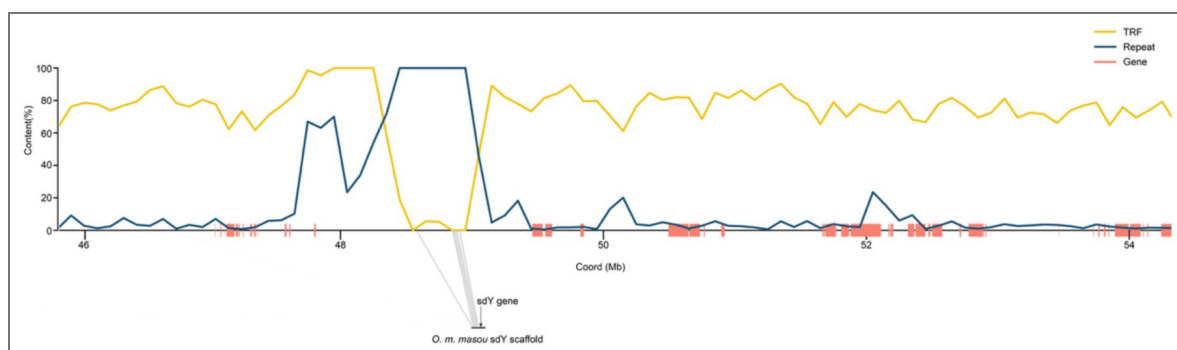


Fig. S9. Alignment of the female-derived *O. m. masou* sdY-containing contig to chromosome 13.

Dot plot visualization illustrating the unique alignment of the female-derived *O. m. masou* contig containing the sex-determining gene (*sdY*) to the central genomic region (~50 Mb) of chromosome 13. Alignment identity is indicated by a color scale, with higher identity in red. This confirms the precise chromosomal location of the sex-linked region at an ancestral chromosome fusion site.



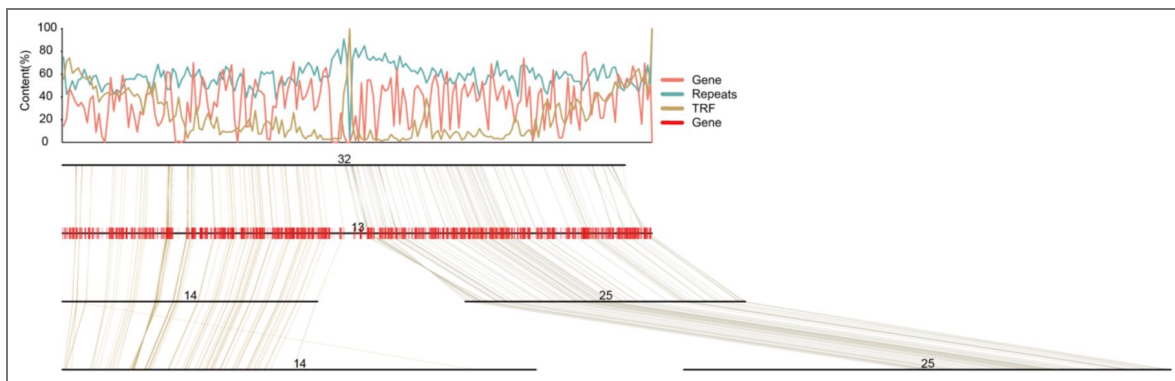
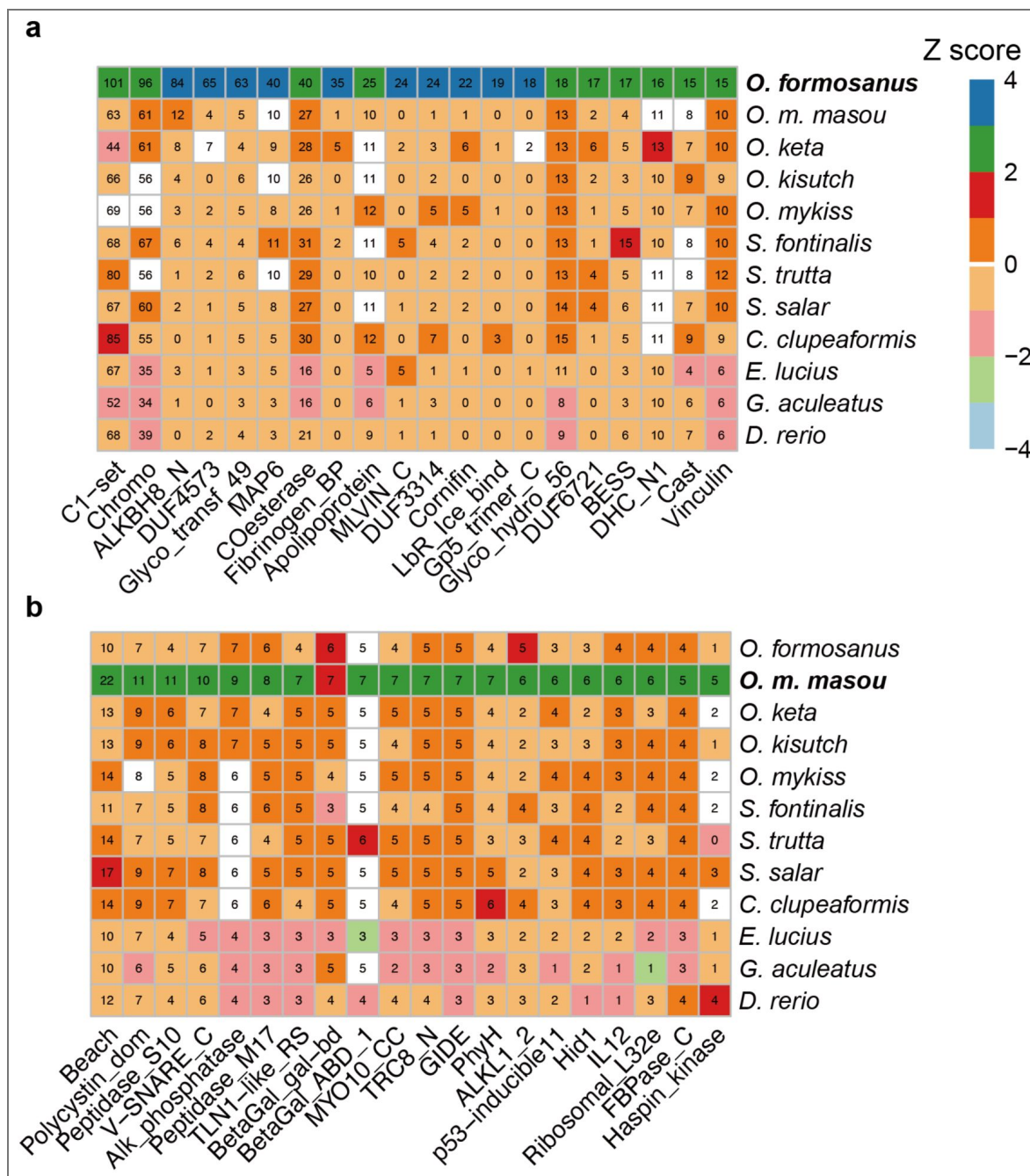
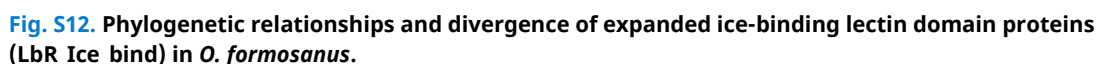


Fig. S10. Comparative alignment showing ancestral chromosome fusion between chromosome 13 of *O. m. masou* and chromosomes 14 and 25 of *O. mykiss*.

Visualization illustrating the genomic correspondence and alignment between chromosome 13 of *O. m. masou* and chromosomes 14 and 25 in rainbow trout (*O. mykiss*). The ancestral chromosome fusion site includes enriched regions of repetitive sequences, notably 281-mer satellite repeats, indicative of centromeric or neocentromeric activity and consistent with chromosome fusion events. Sequence similarity is depicted by color intensity.





Lee, Shen, Lin et al., 2026 *eLife* **15**:RP111444. <https://doi.org/10.7554/eLife.111444.1>

Fig. S13. ddRAD-seq mapping results across the genome of *O. formosanus*.

SNPs were identified across the newly assembled *O. formosanus* genome using ddRAD-seq data with the window size of 1Mb. The mapping results show even genome-wide coverage, indicating the ddRAD-seq data effectively mapped variation across most regions of the reference genome.

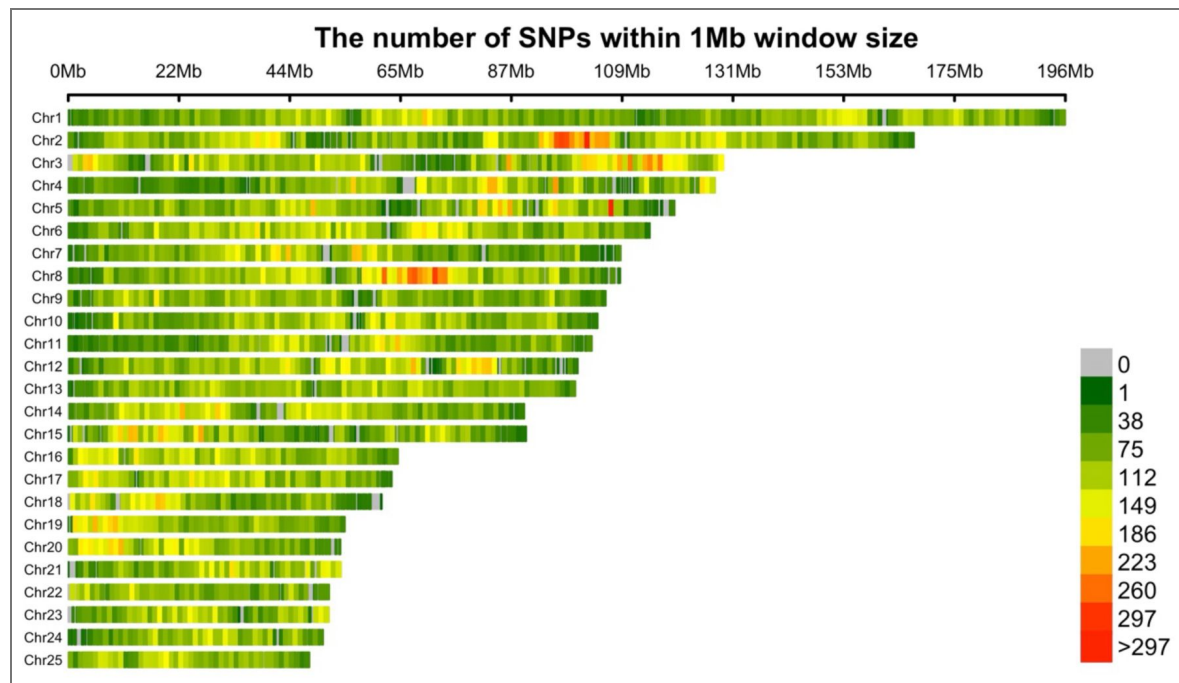


Fig. S14. Quasi-extinction probabilities and odds ratios under increasing typhoon disturbance.

(a) Quasi-extinction probability of Qijiawan (QIJ, light gray) and Hehuan (HEH, dark gray) across four typhoon frequencies. Bars represent means of 1,000 replicates; whiskers indicate exact 95% binomial confidence intervals. (b) Extinction odds ratios (QIJ / HEH) across typhoon scenarios. Gray points and error bars show Fisher's exact test results with exact 95% confidence intervals; asterisks indicate significance levels ($p < 0.05$ *, $p < 0.01$ **, $p < 0.001$ ***, ns = not significant). The black line represents logistic GLM estimates of OR across scenarios, showing a monotonic decline with increasing typhoon frequency. The OR at $p = 0.9$ is omitted due to complete extinction in both populations.

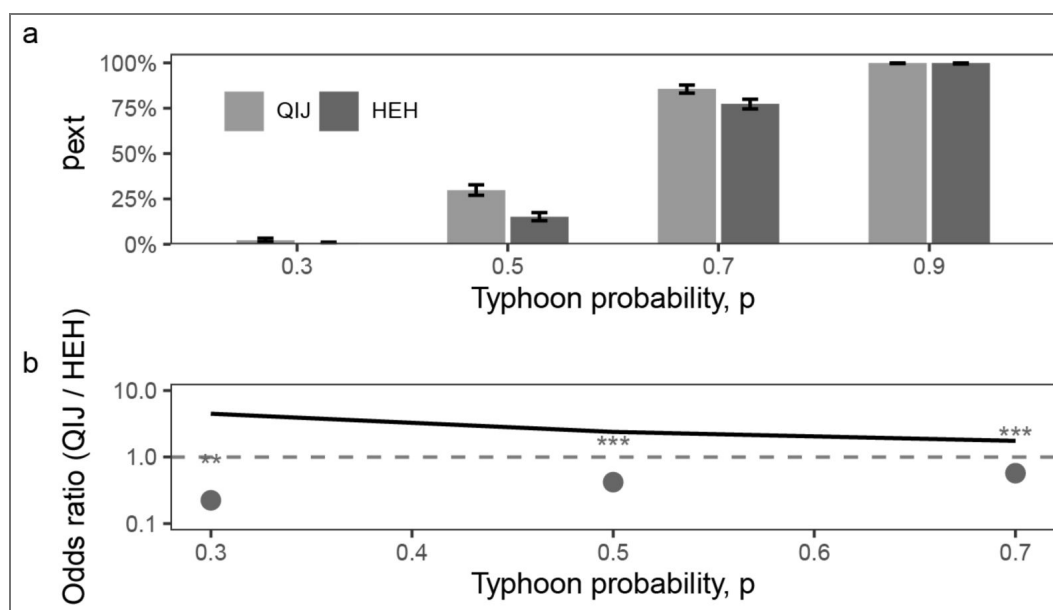


Fig. S15. Electrofishing vs. snorkel sampling frequency and detection ratio by length bin.

Stacked bars show the proportional frequency of fish in each corrected length bin (mm) as detected by electrofishing (purple) and snorkel surveys (orange) during the same sampling period in 2024 at Hehuan Creek. The green line with markers indicates the detection ratio (electrofishing frequency ÷ snorkel frequency) for each length bin, plotted against the secondary y-axis. Bins with a detection ratio >1 suggest underestimation by snorkel sampling, while ratios <1 suggest potential overestimation.

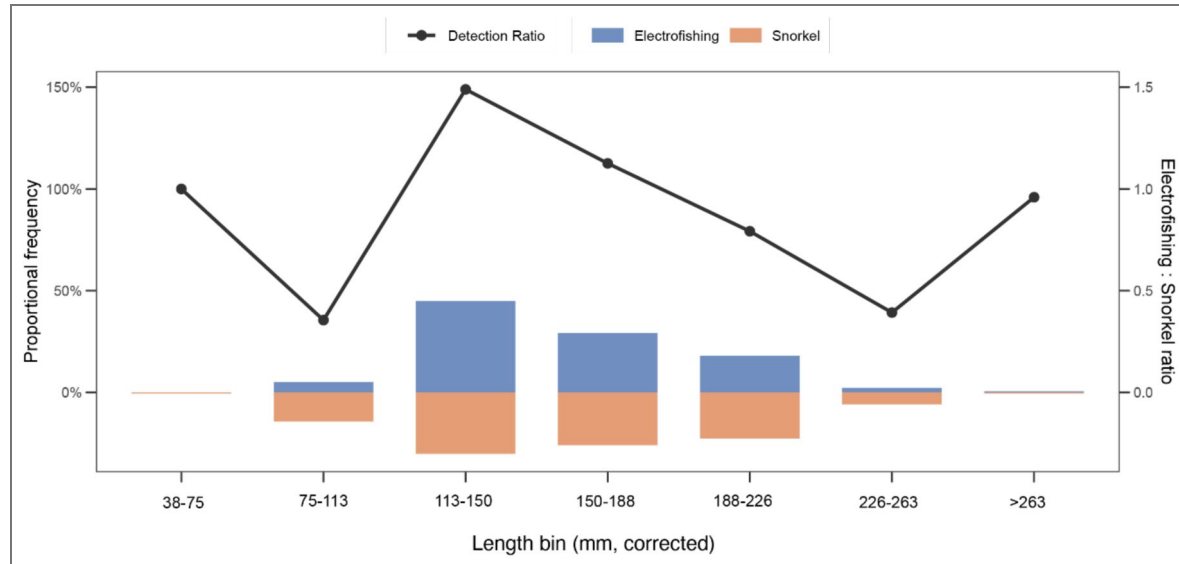
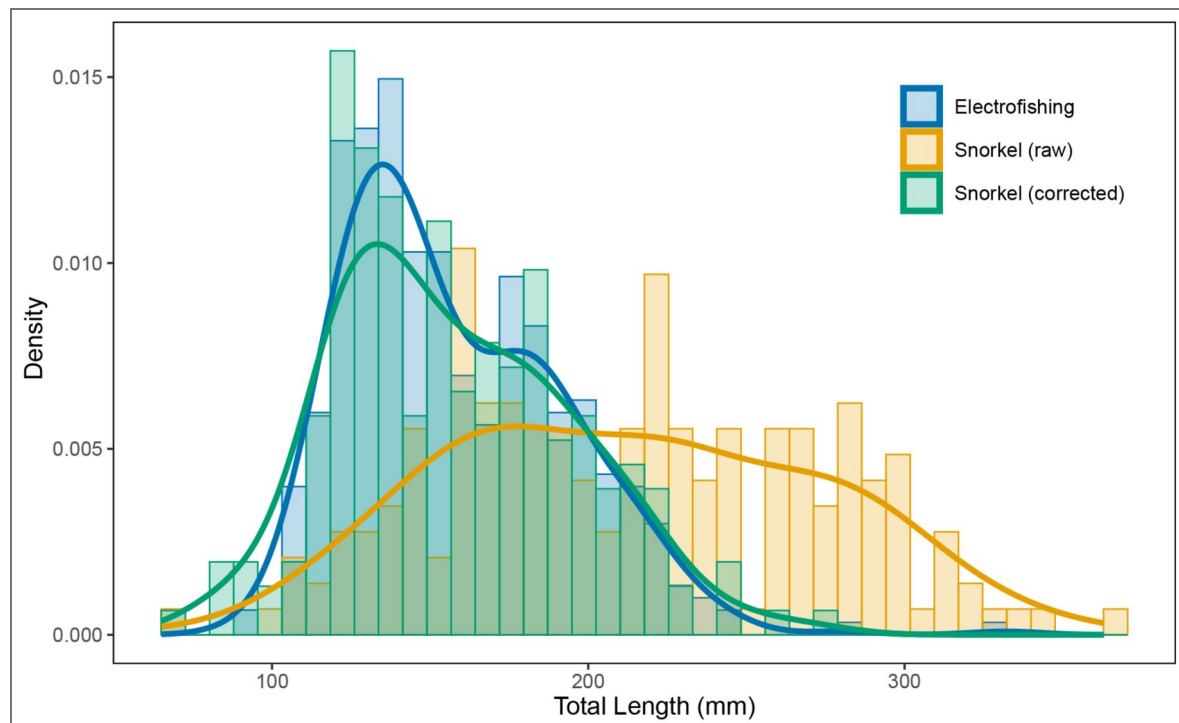


Fig. S16. Total length distributions of electrofishing, raw snorkel, and corrected snorkel samples in Hehuan Creek.

Histograms show fish total length distributions (mm) obtained via electrofishing (blue), unadjusted snorkel observations (orange), and corrected snorkel estimates (green) during the 2024 sampling period in Hehuan Creek. Corrections for snorkel-derived lengths were applied using refractive index adjustment and detection ratio-based resampling to account for observational biases across length classes.



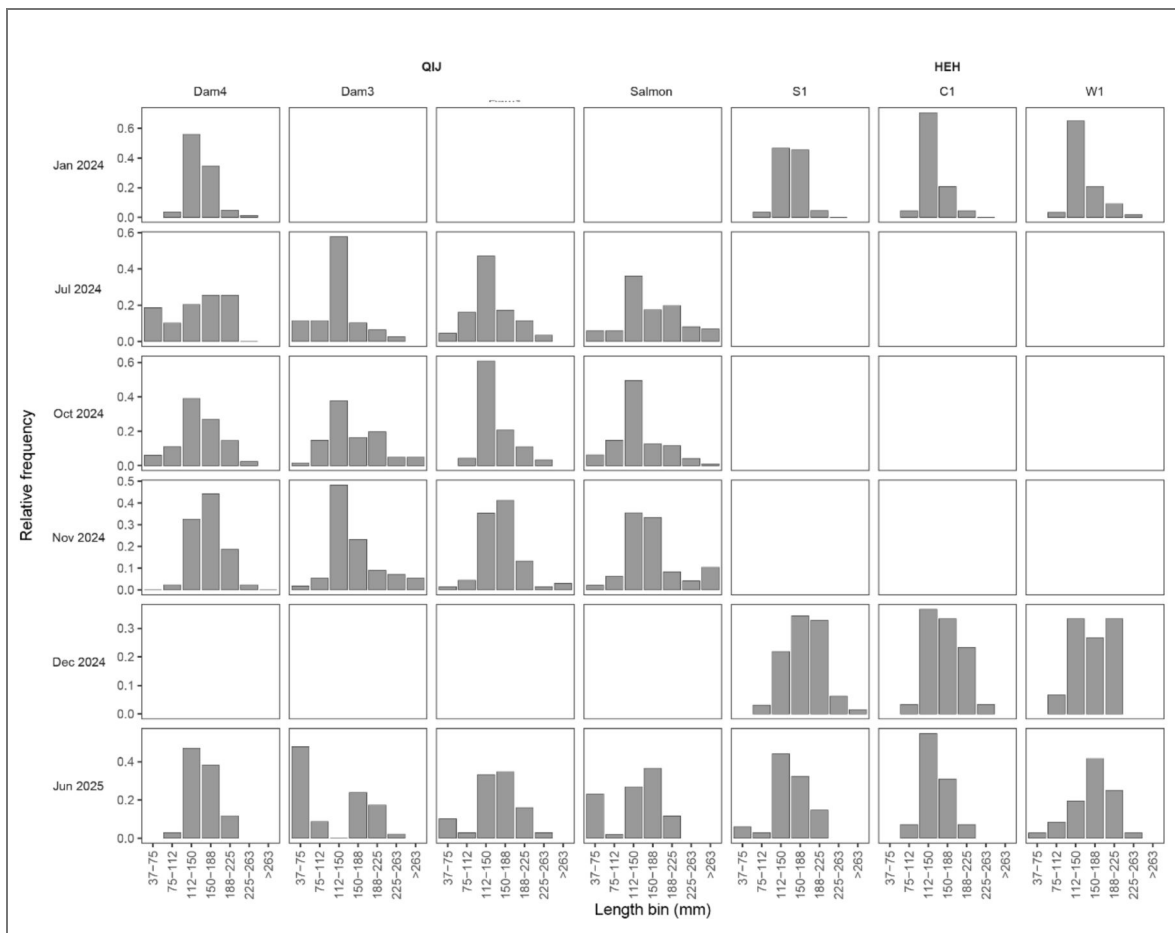


Fig. S17. Monthly body-length distributions of Formosan land-locked salmon in two creeks (QIJ and HEH).

Relative-frequency histograms (grey bars) show the proportional length composition (mm) of snorkel-derived counts—after detection-ratio and refraction corrections—at seven study sites (Dam4, Dam3, Dam2, Salmon, S1, C1 and W1). Columns are ordered left-to-right by site within each tributary; the tributary names (QIJ, HEH) appear once above their respective site block. Rows correspond to individual monthly surveys (Jan, Jul, Oct, Nov, Dec 2024 and Jun 2025). Bars within each panel sum to 1, enabling direct comparison of length-class structure between site-months regardless of sample size. Blank panels indicate months or sites with no observations.

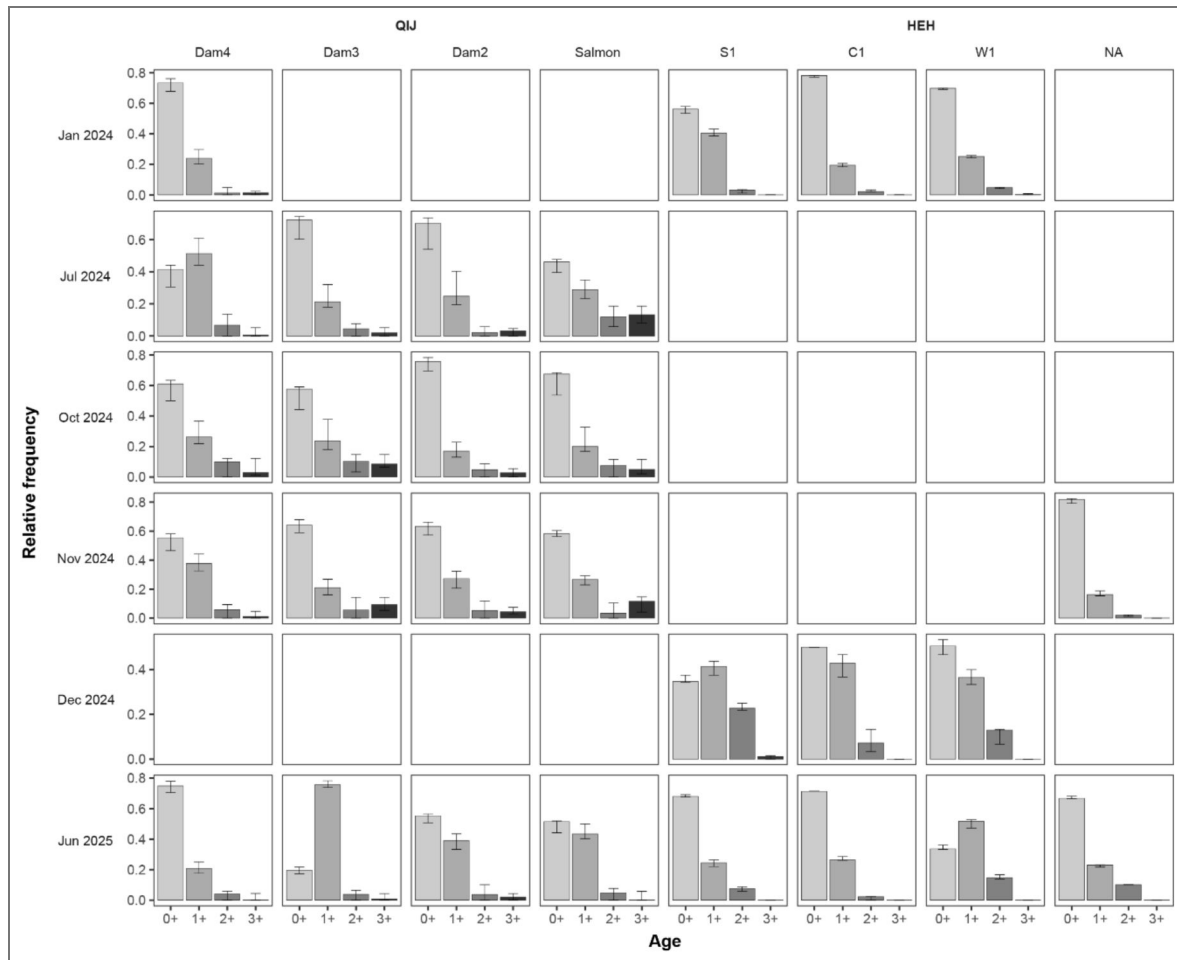


Fig. S18. Monthly age-class composition of Formosan land-locked salmon in the QIJ and HEH tributaries of Hehuan Creek.

Stacked bars (shaded grey from light to dark for age-classes 0, 1, 2 and 3, respectively) show the posterior mean relative frequency of each age-class at seven study sites (Dam4, Dam3, Dam2, Salmon, S1, C1 and W1). Columns are ordered left-to-right by site within each tributary block (stream labels appear once above the block), while rows correspond to individual monthly surveys (January, July, October, November and December 2024, and June 2025). T-shaped lines denote the 95 % Bayesian credible interval around each bar. Within every panel the four bars sum to 1, allowing direct comparison of age structure across site-month combinations irrespective of sample size. Blank panels indicate months or sites with no observations.

Data availability

The raw sequencing reads have been deposited in the NCBI Sequence Read Archive (SRA) under BioProject PRJNA1277495; sample-level accessions are listed in Supplementary Table S1. Upon publication, the chromosome-level assembly and its annotations will be released at GenBank under accession [to be provided at acceptance]. Custom scripts for read processing, variant calling/filtration, demographic inference (PSMC/GONE), species-delimitation (BPP), and the stochastic two-season Leslie simulations are available at <https://anonymous.4open.science/r/salmongenomics-CDD7>. The commit hash used to generate the results is documented in the repository README.

Acknowledgements

We thank Dr. Chyng-Shyan Tzeng (National Chung Hsing University, Taiwan), Dr. Shih-Pin Huang (Biodiversity Research Museum, Academia Sinica, Taiwan) and Yu-Ching Liu (Biodiversity Research Center, Academia Sinica, Taiwan) for their assistance.

Additional information

Funding

JPH is funded by National Science and Technology Council, Taiwan (Grant NSTC 114-2621-B-001-006-MY3). IJT is funded by National Science and Technology Council, Taiwan (Grant NSTC 113-2628-B-001-002 and 114-2628-B-001-014) and Academia Sinica (Grant AS-IA-113-L04). SFS is funded by Academia Sinica (Grant AS-SS-110-05 and AS-SS-114-03).

Funding

Funder	Grant reference number	Author
Academia Sinica (AS)	AS-SS-110-05	Sheng-Feng Shen
Academia Sinica (AS)	AS-SS-114-03	Sheng-Feng Shen
Academia Sinica (AS)	AS-IA-113-L04	Isheng Jason Tsai
National Science and Technology Council (NSTC)	114-2621-B-001-006-MY3	Jen-Pan Huang
National Science and Technology Council (NSTC)	113-2628-B-001-002	Isheng Jason Tsai
National Science and Technology Council (NSTC)	114-2628-B-001-014	Isheng Jason Tsai

Author ORCID iDs

Shih-Fan Chan:  <https://orcid.org/0000-0003-2955-9433>

Isheng Jason Tsai:  <https://orcid.org/0000-0002-2123-5058>

Additional files

[Supplementary Tables](#)

References

- 1 Kawecki TJ (2008) Adaptation to Marginal Habitats. *Annual Review of Ecology, Evolution, and Systematics* **39**:321-342 <https://doi.org/10.1146/annurev.ecolsys.38.091206.095622>
- 2 Lesica P, Allendorf FW (1995) When Are Peripheral Populations Valuable for Conservation?. *Conservation Biology* **9**:753-760 <https://doi.org/10.1046/j.1523-1739.1995.09040753.x>

- 3 Eckert CG, Samis KE, Loughheed SC (2008) Genetic variation across species' geographical ranges: The central-marginal hypothesis and beyond. *Molecular Ecology* **17**:1170-1188 <https://doi.org/10.1111/j.1365-294X.2007.03659.x> | PubMed
- 4 Angert AL, Bontrager MG, Ågren J (2020) What Do We Really Know About Adaptation at Range Edges?. *Annual Review of Ecology, Evolution, and Systematics* **51**:341-361 <https://doi.org/10.1146/annurev-ecolsys-012120-091002>
- 5 SgrÒ CM, et al. (2010) A comprehensive assessment of geographic variation in heat tolerance and hardening capacity in populations of *Drosophila melanogaster* from eastern Australia. *Journal of Evolutionary Biology* **23**:2484-2493 <https://doi.org/10.1111/j.1420-9101.2010.02110.x>
- 6 Fraser DJ, Weir LK, Bernatchez L, Hansen MM, Taylor EB (2011) Extent and scale of local adaptation in salmonid fishes: review and meta-analysis. *Heredity* **106**:404-420 <https://doi.org/10.1038/hdy.2010.167> | PubMed
- 7 Lande R (1994) Risk of population extinction from fixation of new deleterious mutations. *Evolution* **8**:1460-1469 <https://doi.org/10.1111/j.1558-5646.1994.tb02188.x>
- 8 Whitlock MC (2000) Fixation of new alleles and the extinction of small populations: Drift load, beneficial alleles, and sexual selection. *Evolution* **54**:1855-1861 <https://doi.org/10.1111/j.0014-3820.2000.tb01232.x> | PubMed
- 9 Tallmon DA, Luikart G, Waples RS (2004) The alluring simplicity and complex reality of genetic rescue. *Trends in Ecology & Evolution* **19**:489-496 <https://doi.org/10.1016/j.tree.2004.07.003> | PubMed
- 10 Moritz C (1994) Defining "evolutionarily significant units" for conservation. *Trends in Ecology & Evolution* **9**:373-375 [https://doi.org/10.1016/0169-5347\(94\)90057-4](https://doi.org/10.1016/0169-5347(94)90057-4) | PubMed
- 11 Spielman D, Brook BW, Frankham R (2004) Most species are not driven to extinction before genetic factors impact them. *Proceedings of the National Academy of Sciences* **101**:15261-15264 <https://doi.org/10.1073/pnas.0403809101> | PubMed
- 12 Frankham R (2015) Genetic rescue of small inbred populations: Meta-analysis reveals large and consistent benefits of gene flow. *Molecular Ecology* **24**:2610-2618 <https://doi.org/10.1111/mec.13139> | PubMed
- 13 Hedrick PW, Fredrickson R (2010) Genetic rescue guidelines with examples from Mexican wolves and Florida panthers. *Conservation Genetics* **11**:615-626 <https://doi.org/10.1007/s10592-009-9999-5>
- 14 Edmands S (2007) Between a rock and a hard place: Evaluating the relative risks of inbreeding and outbreeding for conservation and management. *Molecular Ecology* **16**:463-475 <https://doi.org/10.1111/j.1365-294X.2006.03148.x> | PubMed
- 15 Yamamoto S, et al. (2020) Phylogeography of a salmonid fish, masu salmon *Oncorhynchus masou* subspecies-complex, with disjunct distributions across the temperate northern Pacific. *Freshwater Biology* **65**:698-715 <https://doi.org/10.1111/fw.b.13460>
- 16 Fujioka Y, Kuwahara M, Tabata R, Fuke Y, Nakabo T (2025) The Biwa salmon, a new species of *Oncorhynchus* (Salmonidae) endemic to Lake Biwa, Japan. *Ichthyological Research* <https://doi.org/10.1007/s10228-025-01032-z>
- 17 Behnke RJ, Koh TP, Needham PR (1962) Status of the Landlocked Salmonid Fishes of Formosa with a Review of *Oncorhynchus masou* (Brevoort). *Copeia* **1962**:400-407 <https://doi.org/10.2307/1440908>
- 18 Kottelat M (1996) *Oncorhynchus formosanus*. *The IUCN Red List of Threatened Species* **1996** <https://doi.org/10.2305/IUCN.UK.1996.RLTS.T15323A4513507.en>
- 19 Tzeng CS, Lin YS, Lin SM, Wang TY, Wang FY (2006) The phylogeography and population demographics of selected freshwater fishes in Taiwan. *Zoological Studies* **45**:285-297
- 20 Gwo JC, Hsu TH, Lin KH, Chou YC (2008) Genetic relationship among four subspecies of cherry salmon (*Oncorhynchus masou*) inferred using AFLP. *Mol Phylogenet Evol* **48**:776-781 <https://doi.org/10.1016/j.ympev.2007.12.023> | PubMed

- 21 Hsu TH, et al. (2010) Use of microsatellite DNA and amplified fragment length polymorphism for Cherry salmon (*Oncorhynchus masou*) complex identification. *Aquaculture Research* **41**:e316-e325 <https://doi.org/10.1111/j.1365-2109.2010.02533.x>
- 22 Nakabo T (2009) Zoogeography of Taiwanese Fishes. *Korean Journal of Ichthyology* **21**:311-321
- 23 Lin YS, Tsao SS, Chang KH (1990) Population and distribution of the Formosan landlocked salmon (*Oncorhynchus masou formosanus*) in Chichiawan stream. *Bulletin of the Institute of Zoology, Academia Sinica* **29**:73-85
- 24 Chan S.-F, et al. (2025) Assessment across life stages reveals superior habitat suitability in reintroduced historical habitats for an endangered salmon species. *Journal of Applied Ecology* <https://doi.org/10.1111/1365-2664.70067>
- 25 Liao LY, Chen YX, Chang CY (2024) Population size, distribution, and next year's proposed work for the Formosan landlocked salmon (*Oncorhynchus masou formosanus*) in Taiwan, 2023. Taiwan: Shei-Pa National Park Administration & Taroko National Park Administration.
26. Burrige CP (2019) Politics and pride: Maintaining genetic novelty may be detrimental for the conservation of Formosa landlocked salmon *Oncorhynchus formosanus*. *Aquatic Conservation: Marine and Freshwater Ecosystems* **29**:840-847 <https://doi.org/10.1002/aqc.3097>
- 27 Zhou C, McCarthy SA, Durbin R, Alkan C (2023) YaHS: yet another Hi-C scaffolding tool. *Bioinformatics* **39** <https://doi.org/10.1093/bioinformatics/btac808> | PubMed
- 28 Christensen KA, et al. (2025) Masu salmon species complex relationships and sex chromosomes revealed from analyses of the masu salmon (*Oncorhynchus masou masou*) genome assembly. *G3: Genes, Genomes, Genetics* **15** <https://doi.org/10.1093/g3journal/jkae278> | PubMed
- 29 Ho CW, Chen JJW, Lee TH, Lin HJ (2016) Complete mitochondrial genome of *Oncorhynchus masou formosanus* (Jordan & Oshima, 1919) (Pisces, Salmonidae). *Mitochondrial DNA B Resour* **1**:295-296 <https://doi.org/10.1080/23802359.2016.1166084> | PubMed
- 30 Zhu W, et al. (2022) A chromosome-level genome of *Brachymystax tsinlingensis* provides resources and insights into salmonids evolution. *G3 Genes | Genomes | Genetics* **12** <https://doi.org/10.1093/g3journal/jkac162> | PubMed
- 31 Lien S, et al. (2016) The Atlantic salmon genome provides insights into rediploidization. *Nature* **533**:200-205 <https://doi.org/10.1038/nature17164> | PubMed
- 32 Simao FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM (2015) BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**:3210-3212 <https://doi.org/10.1093/bioinformatics/btv351> | PubMed
- 33 Cantalapiedra CP, Hernandez-Plaza A, Letunic I, Bork P, Huerta-Cepas J (2021) eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. *Mol Biol Evol* **38**:5825-5829 <https://doi.org/10.1093/molbev/msab293> | PubMed
- 34 Emms DM, Kelly S (2019) OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biology* **20** <https://doi.org/10.1186/s13059-019-1832-y> | PubMed
- 35 Yang Z (2007) PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Molecular Biology and Evolution* **24**:1586-1591 <https://doi.org/10.1093/molbev/msm088> | PubMed
- 36 Smith GR (1992) Introgression in Fishes: Significance for Paleontology, Cladistics, and Evolutionary Rates. *Systematic Biology* **41**:41-57 <https://doi.org/10.1093/sysbio/41.1.41>
- 37 Flouri T, Jiao X, Rannala B, Yang Z, Yoder AD (2018) Species Tree Inference with BPP Using Genomic Sequences and the Multispecies Coalescent. *Molecular Biology and Evolution* **35**:2585-2593 <https://doi.org/10.1093/molbev/msy147> | PubMed
- 38 Sukumaran J, Knowles LL (2017) Multispecies coalescent delimits structure, not species. *Proc Natl Acad Sci U S A* **114**:1607-1612 <https://doi.org/10.1073/pnas.1607921114> | PubMed
- 39 Kornai D, Jiao X, Ji J, Flouri T, Yang Z (2024) Hierarchical heuristic species delimitation under the multispecies coalescent model with migration. *Systematic Biology* **73**:1015-1037 <https://doi.org/10.1093/sysbio/syae050> | PubMed

- 40 Ghezelayagh A, et al. (2025) Comparative species delimitation of a biological conservation icon. *Current Biology* **35**:398-406.e394 <https://doi.org/10.1016/j.cub.2024.11.053> | PubMed
- 41 Schubert I (2018) What is behind “centromere repositioning”? *Chromosoma* **127**:229-234 <https://doi.org/10.1007/s00412-018-0672-y> | PubMed
- 42 Bertolotti AC, et al. (2020) The structural variation landscape in 492 Atlantic salmon genomes. *Nature Communications* **11** <https://doi.org/10.1038/s41467-020-18972-x> | PubMed
- 43 Wellband K, et al. (2018) Chromosomal fusion and life history-associated genomic variation contribute to within-river local adaptation of Atlantic salmon. *Molecular Ecology* **28**:1439-1459 <https://doi.org/10.1111/mec.14965> | PubMed
- 44 Pearse DE, et al. (2019) Sex-dependent dominance maintains migration supergene in rainbow trout. *Nature Ecology & Evolution* **3**:1731-1742 <https://doi.org/10.1038/s41559-019-1044-6> | PubMed
- 45 Lubieniecki KP, et al. (2015) Genomic instability of the sex-determining locus in Atlantic salmon (*Salmo salar*). *G3* **5**:2513-2522 <https://doi.org/10.1534/g3.115.020115> | PubMed
- 46 Delphin C, et al. (2012) MAP6-F Is a Temperature Sensor That Directly Binds to and Protects Microtubules from Cold-induced Depolymerization. *Journal of Biological Chemistry* **287**:35127-35138 <https://doi.org/10.1074/jbc.M112.398339> | PubMed
- 47 He J, et al. (2023) Transcriptomic Networks Reveal the Tissue-Specific Cold Shock Responses in Japanese Flounder (*Paralichthys olivaceus*). *Biology* **12** <https://doi.org/10.3390/biology12060784> | PubMed
- 48 Bar Dolev M, Braslavsky I, Davies PL (2016) Ice-Binding Proteins and Their Function. *Annu Rev Biochem* **85**:515-542 <https://doi.org/10.1146/annurev-biochem-060815-014546> | PubMed
- 49 Bar Dolev M, Braslavsky I, Davies PL (2016) Ice-Binding Proteins and Their Function. *Annual Review of Biochemistry* **85**:515-542 <https://doi.org/10.1146/annurev-biochem-060815-014546> | PubMed
- 50 Thomaz AT, Christie MR, Knowles LL (2016) The architecture of river networks can drive the evolutionary dynamics of aquatic populations. *Evolution* **70**:731-739 <https://doi.org/10.1111/evo.12883> | PubMed
- 51 Leaché AD, Zhu T, Rannala B, Yang Z (2018) The spectre of too many species. *Systematic Biology* **68**:168-181 <https://doi.org/10.1093/sysbio/syy051>
- 52 Franklin IR (1980) Evolutionary change in small populations. In: Soulé ME, Wilcox BA (Eds). *Conservation Biology: An Evolutionary-Ecological Perspective* Sinauer Associates. pp. 135-150
- 53 Jamieson IG, Allendorf FW (2012) How does the 50/500 rule apply to MVPs?. *Trends in Ecology & Evolution* **27**:578-584 <https://doi.org/10.1016/j.tree.2012.07.001> | PubMed
- 54 Yogi K, Nakamura H (1938) A land-locked salmon from Formosa (a natural monument in Taiwan). Home Affairs Bureau of the Office of the Governor-General of Taiwan.
- 55 Taiwan GG. o (1941) Governor-General of Taiwan, Taihoku, Taiwan.
- 56 Zheng MN (1987) Final report on the ecological survey of Formosan landlocked salmon and alpine fish, Natural Resources in Taroko National Park. Construction and Planning Agency, Ministry of the Interior.
- 57 Baiotto T, Guzman LM (2025) Maintaining Local Adaptation Is Key for Evolutionary Rescue and Long-Term Persistence of Populations Experiencing Habitat Loss and a Changing Environment. *Evolutionary Applications* **18**:e70081 <https://doi.org/10.1111/eva.70081> | PubMed
- 58 Rhymer JM, Simberloff D (1996) Extinction by Hybridization and Introgression. *Annual Review of Ecology and Systematics* **27**:83-109 <https://doi.org/10.1146/annurev.ecolsys.27.1.83>
- 59 Todesco M, et al. (2016) Hybridization and extinction. *Evolutionary Applications* **9**:892-908 <https://doi.org/10.1111/eva.12367> | PubMed
- 60 Yan F, et al. (2018) The Chinese giant salamander exemplifies the hidden extinction of cryptic species. *Current Biology* **28**:R590-R592 <https://doi.org/10.1016/j.cub.2018.04.004> | PubMed

- 61 Johnson P, et al. (2025) Conservation aquaculture of wild-origin offspring preserves genetic diversity in an endangered population of white sturgeon. *Conservation Genetics* **26**:335-346 <https://doi.org/10.1007/s10592-024-01670-2>
- 62 Kalinowski ST, Van Doornik DM, Kozfkay CC, Waples RS (2012) Genetic diversity in the Snake River sockeye salmon captive broodstock program as estimated from broodstock records. *Conservation Genetics* **13**:1183-1193 <https://doi.org/10.1007/s10592-012-0363-9>
- 63 Eliason EJ, et al. (2011) Differences in Thermal Tolerance Among Sockeye Salmon Populations. *Science* **332**:109-112 <https://doi.org/10.1126/science.1199158> | PubMed
- 64 Zillig KW, FitzGerald AM, Lusardi RA, Cocherell DE, Fangue NA (2023) Intraspecific variation among Chinook Salmon populations indicates physiological adaptation to local environmental conditions. *Conservation Physiology* **11** <https://doi.org/10.1093/conphys/coad044> | PubMed
- 65 Labonne J, et al. (2021) Using Reciprocal Transplants to Assess Local Adaptation, Genetic Rescue, and Sexual Selection in Newly Established Populations. *Genes* **12** <https://doi.org/10.3390/genes12010005> | PubMed
- 66 Hsu T.-H, Gwo J.-C (2010) A PCR-based method for sex identification of critically endangered Formosa landlocked salmon. *Fisheries Science* **76**:613-618 <https://doi.org/10.1007/s12562-010-0255-1>
- 67 Cheng H, Concepcion GT, Feng X, Zhang H, Li H (2021) Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nature Methods* **18**:170-175 <https://doi.org/10.1038/s41592-020-01056-5> | PubMed
- 68 Huang S, Kang M, Xu A (2017) HaploMerger2: rebuilding both haploid sub-assemblies from high-heterozygosity diploid genome assembly. *Bioinformatics* **33**:2577-2579 <https://doi.org/10.1093/bioinformatics/btx220> | PubMed
- 69 Durand NC, et al. (2016) Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. *Cell Syst* **3**:99-101 <https://doi.org/10.1016/j.cels.2015.07.012> | PubMed
- 70 Baril T, Galbraith J, Hayward A (2024) Earl Grey: A Fully Automated User-Friendly Transposable Element Annotation and Analysis Pipeline. *Mol Biol Evol* **41** <https://doi.org/10.1093/molbev/msae068> | PubMed
- 71 Dobin A, Gingeras TR (2015) Mapping RNA-seq Reads with STAR. *Curr Protoc Bioinformatics* **51**:11 14 11-11 14 19 <https://doi.org/10.1002/0471250953.bi1114s51> | PubMed
- 72 Gabriel L, et al. (2024) BRAKER3: Fully automated genome annotation using RNA-seq and protein evidence with GeneMark-ETP, AUGUSTUS, and TSEBRA. *Genome Res* **34**:769-777 <https://doi.org/10.1101/gr.278090.123> | PubMed
- 73 Emms DM, Kelly S (2015) OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol* **16**:157 <https://doi.org/10.1186/s13059-015-0721-2> | PubMed
- 74 Minh BQ, et al. (2020) IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution* **37**:1530-1534 <https://doi.org/10.1093/molbev/msaa015> | PubMed
- 75 Mendes FK, Vanderpool D, Fulton B, Hahn MW, Robinson P (2020) CAFE 5 models variation in evolutionary rates among gene families. *Bioinformatics* **36**:5516-5518 <https://doi.org/10.1093/bioinformatics/btaa1022> | PubMed
- 76 dos Reis M, Yang Z (2011) Approximate Likelihood Calculation on a Phylogeny for Bayesian Estimation of Divergence Times. *Molecular Biology and Evolution* **28**:2161-2172 <https://doi.org/10.1093/molbev/msr045> | PubMed
- 77 Chen S (2023) Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. *iMeta* **2**:e107 <https://doi.org/10.1002/imt2.107> | PubMed
- 78 Vasimuddin M, Misra S, Li H, Aluru S (2019) Efficient Architecture-Aware Acceleration of BWA-MEM for Multicore Systems. In: IEEE International Parallel and Distributed Processing Symposium (IPDPS). pp. 314-324

- 79 McKenna A, et al. (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* **20**:1297-1303 <https://doi.org/10.1101/gr.107524.110> | PubMed
- 80 Danecek P, et al. (2021) Twelve years of SAMtools and BCFtools. *Gigascience* **10** <https://doi.org/10.1093/gigascience/giab008> | PubMed
- 81 Chang CC, et al. (2015) Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* **4** <https://doi.org/10.1186/s13742-015-0047-8> | PubMed
- 82 Alexander DH, Novembre J, Lange K (2009) Fast model-based estimation of ancestry in unrelated individuals. *Genome Res* **19**:1655-1664 <https://doi.org/10.1101/gr.094052.109> | PubMed
- 83 Christensen KA, et al. (2024) Revealing the evolutionary history and contemporary population structure of Pacific salmon in the Fraser River through genome resequencing. *G3* **14** <https://doi.org/10.1093/g3journal/jkae169> | PubMed
- 84 Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS (2017) UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution* **35**:518-522 <https://doi.org/10.1093/molbev/msx281> | PubMed
- 85 Cingolani P, et al. (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly* **6**:80-92 <https://doi.org/10.4161/fly.19695> | PubMed
- 86 Li H, Durbin R (2011) Inference of human population history from individual whole-genome sequences. *Nature* **475**:493-496 <https://doi.org/10.1038/nature10231> | PubMed
- 87 Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**:1754-1760 <https://doi.org/10.1093/bioinformatics/btp324> | PubMed
- 88 Subramanian S, Kumar M (2023) Genomic footprints of bottleneck in landlocked salmon population. *Sci Rep* **13**:6706 <https://doi.org/10.1038/s41598-023-34076-0> | PubMed
- 89 Sahashi G, Morita K (2024) Partial Migration in Salmonids: Focusing on Asian Endemic Masu Salmon (*Oncorhynchus masou*) and White-Spotted Charr (*Salvelinus leucomaenis*). In: Lobon-Cervia Javier, Budy Phaedra, Gresswell Robert (Eds). *Advances in the Ecology of Stream-Dwelling Salmonids* Springer International Publishing. pp. 255-278
- 90 De Vivo M, et al. (2023) Genomic tools for comparative conservation genetics among three recently diverged stag beetles. *Insect Conserv Diver* **16**:853-869 <https://doi.org/10.1111/icad.12678>
- 91 Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One* **7**:e37135 <https://doi.org/10.1371/journal.pone.0037135> | PubMed
- 92 Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool set for population genomics. *Mol Ecol* **22**:3124-3140 <https://doi.org/10.1111/mec.12354> | PubMed
- 93 Rivera-Colon AG, Catchen J (2022) Population Genomics Analysis with RAD, Reprised: Stacks 2. *Methods Mol Biol* **2498**:99-149 https://doi.org/10.1007/978-1-0716-2313-8_7 | PubMed
- 94 Li H (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* <https://doi.org/10.48550/arXiv.1303.3997>
- 95 Yang ZH (2015) The BPP program for species tree estimation and species delimitation. *Curr Zool* **61**:854-865 <https://doi.org/10.1093/czoolo/61.5.854>
- 96 Degnan JH, Rosenberg NA (2009) Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends Ecol Evol* **24**:332-340 <https://doi.org/10.1016/j.tree.2009.01.009> | PubMed
- 97 King TL, et al. (2007) Biodiversity and Population Structure. In: Verspoor E, Stradmeyer L, Nielsen J (Eds). *The Atlantic Salmon: Genetics, Conservation and Management* Wiley-Blackwell. pp. 117-166
- 98 Jackson ND, Carstens BC, Morales AE, O'Meara BC (2017) Species Delimitation with Gene Flow. *Syst Biol* **66**:799-812 <https://doi.org/10.1093/sysbio/syw117> | PubMed

- 99 Huang JP (2021) The Genealogical Divergence Index Across a Speciation Continuum in Hercules Beetles. *Insect Syst Diver* **5** <https://doi.org/10.1093/isd/ixab021>
- 100 Knaus BJ, Grunwald NJ (2017) VCFR: a package to manipulate and visualize variant call format data in R. *Mol Ecol Resour* **17**:44-53 <https://doi.org/10.1111/1755-0998.12549> | PubMed
- 101 DeRaad DA (2022) SNPfiltR: an R package for interactive and reproducible SNP filtering. *Mol Ecol Resour* **22**:2443-2453 <https://doi.org/10.1111/1755-0998.13618> | PubMed
- 102 Jombart T (2008) Adegnet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**:1403-1405 <https://doi.org/10.1093/bioinformatics/btn129> | PubMed
- 103 Dray S, Dufour AB (2007) The ade4 package: Implementing the duality diagram for ecologists. *J Stat Softw* **22**:1-20 <https://doi.org/10.18637/jss.v022.i04>
- 104 Coltman DW, Pilkington JG, Smith JA, Pemberton JM (1999) Parasite-Mediated Selection against Inbred Soay Sheep in a Free-Living Island Populaton. *Evolution* **53**:1259-1267 <https://doi.org/10.1111/j.1558-5646.1999.tb04538.x>
- 105 Stoffel MA, et al. (2016) inbreedR: an R package for the analysis of inbreeding based on genetic markers. *Methods in Ecology and Evolution* **7**:1331-1339 <https://doi.org/10.1111/2041-210x.12588>
- 106 Purcell S, et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**:559-575 <https://doi.org/10.1086/519795> | PubMed
- 107 Brüniche-Olsen A, Kellner KF, Anderson CJ, DeWoody JA (2018) Runs of homozygosity have utility in mammalian conservation and evolutionary studies. *Conservation Genetics* **19**:1295-1307 <https://doi.org/10.1007/s10592-018-1099-y>
- 108 Li SH, et al. (2024) Low but highly geographically structured genomic diversity of East Asian Eurasian otters and its conservation implications. *Evol Appl* **17**:e13630 <https://doi.org/10.1111/eva.13630> | PubMed
- 109 Do C, et al. (2014) NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Mol Ecol Resour* **14**:209-214 <https://doi.org/10.1111/1755-0998.12157> | PubMed
- 110 Waples RK, Larson WA, Waples RS (2016) Estimating contemporary effective population size in non-model species using linkage disequilibrium across thousands of loci. *Heredity* **117**:233-240 <https://doi.org/10.1038/hdy.2016.60> | PubMed
- 111 Gruber B, Unmack PJ, Berry OF, Georges A (2018) dartr: An r package to facilitate analysis of SNP data generated from reduced representation genome sequencing. *Mol Ecol Resour* **18**:691-699 <https://doi.org/10.1111/1755-0998.12745> | PubMed
- 112 Mijangos JL, Gruber B, Berry O, Pacioni C, Georges A (2022) dartR v2: an accessible genetic analysis platform for conservation, ecology and agriculture. *Methods in Ecology and Evolution* **13**:2150-2158 <https://doi.org/10.1111/2041-210x.13918>
- 113 Santiago E, et al. (2020) Recent demographic history inferred by high-resolution analysis of linkage disequilibrium. *Mol Biol Evol* **37**:3642-3653 <https://doi.org/10.1093/molbev/msaa169> | PubMed
- 114 Smith SR, et al. (2022) A chromosome-anchored genome assembly for Lake Trout (*Salvelinus namaycush*). *Mol Ecol Resour* **22**:679-694 <https://doi.org/10.1111/1755-0998.13483> | PubMed
- 115 Ito S.-o, Ishida Y (1998) Species identification and age determination of Pacific Salmon (*Oncorhynchus* spp.) by scale patterns [original in Japanese]. *Bulletin of the National Research Institute of Far Seas Fisheries* **35**:131-154
- 116 Kaneto Y, Katayama S, Iida M (2017) Age determination of chum salmon using scales and otoliths [original in Japanese]. *Nippon Suisan Gakkaishi* **83**:758-763 <https://doi.org/10.2331/suisan.16-00071>
- 117 Shearer WM (1992) Atlantic Salmon Scale Reading Guidelines. Copenhagen, Denmark: International Council for the Exploration of the Sea.
- 118 Borgström R, Skaala Ø (1993) Size-dependent catchability of Brown trout and Atlantic salmon parr by electrofishing in a low conductivity stream. *Nordic Journal of Freshwater Research* **68**:14-21

- 119 Hagen J, Decker S, Korman J, Bison RG (2010) Effectiveness of night snorkeling for estimating steelhead parr abundance in a large river basin. *North American Journal of Fisheries Management* **30**:1303-1314 <https://doi.org/10.1577/M09-160.1>
- 120 Maceina MJ (1997) Simple application of using residuals from catch-curve regressions to assess year-class strength in fish. *Fisheries Research* **32**:115-121 [https://doi.org/10.1016/S0165-7836\(97\)00051-9](https://doi.org/10.1016/S0165-7836(97)00051-9)
- 121 Waples RS (2022) TheWeight: A simple and flexible algorithm for simulating non-ideal, age-structured populations. *Methods in Ecology and Evolution* **13**:2030-2041 <https://doi.org/10.1111/2041-210X.13926>

Peer reviews

Reviewer #1 (Public review):

Summary:

This is an interesting paper on an important topic, the taxonomic and conservation status of some unusual salmonid populations in Taiwan.

Strengths:

The first part of the manuscript is quite strong: the authors sequence and build a reference genome and conduct a phylogenomic analysis. They examine chromosome structure and rearrangements, test for loss-of-function mutations, and do a proteomic analysis. As a stand-alone, this could serve as its own manuscript, perhaps for a more specialized journal.

Weaknesses:

I find this manuscript rather disjointed. The first part of the manuscript is related to phylogenomics of the taxon in question, compared to other nearby species from Japan. The authors go on to describe chromosomal rearrangements, sex-chromosome location, and proteomics. All of these fit within a paper about taxon-level issues. I do find the proteomic analysis perhaps unnecessary. I'm not sure we learn much of substance through this analysis, which is highly speculative.

PSMC analysis seems highly questionable for taxa with such strong genetic structure. If historical N_e and past changes in structure are confounded, what does this analysis provide? I recommend deletion of the analysis included in Figure 1e.

The second portion of the manuscript deals with population structure of three *O. formosanus* populations, based on RADSeq data. This part reads as a separate manuscript, in my opinion. I think the authors are trying to squeeze too much into one manuscript.

For the second part on population genomics, not enough detail is provided to evaluate the methods, results, and interpretations. For example, not enough detail is provided about each of the three Taiwan populations, the stocking history, and the demographic data collection. The only information available is a brief paragraph in the introduction. Was the Luoyewei (L) population stocked from a brook derived from this population or from Qijiawan (Q)? Why do three L fish have such different levels of MLH? Are these stocked from somewhere else? Are the rest of the fish from one pool, and maybe one family (this would also explain the extremely low contemporary N_e)? Only 17 fish were examined from L, and apparently from one site in the stream; more detail is needed. Are L, Q, and H currently isolated? What is the stocking history? The authors conclude that the Hehuan (H) population has more genetic variation and is likely the result of an unknown native population that bred with stocked fish (which arise from Q). This story does align with the genetic results, but again, more detail is needed. Are there alternative explanations? A more careful treatment would be helpful.

The demographic modeling is not convincing. Not enough detail is provided, and the lack of individual identification of fish makes it so the modeling is very general. It is hard to place too much stock in these vital rate estimates. The methods were fishing, snorkeling, and some electrofishing. Scales were used for ageing, and catch curve analysis was employed. Overall, this is an underdeveloped portion of the paper that is important, but not convincing as written.

<https://doi.org/10.7554/eLife.111444.1.sa1>

Reviewer #2 (Public review):

Summary:

Lee et al. is a comprehensive conservation genomics study that combines a chromosome-level genome assembly (sex-specific, too), population resequencing, coalescent species delimitation, and simulations to reassess the evolutionary status and conservation outlook of the Formosan landlocked salmon, *Oncorhynchus formosanus*. The authors showed a distinctive genome structure, replete with chromosome fusions and an unusual placement of the sex-determining gene *sdY*. Across sampling sites, they observed variable levels of genetic diversity, but in a way that was surprising given previous census numbers and conservation history. In particular, the authors report a previously unrecognised native population in Hehuan Creek, and conclude that Hehuan is more resilient to typhoon disturbance than the long-protected Qijiawan population - motivating stream-specific rather than range-wide conservation.

Overall, this is a well-written paper that combines a number of elements that are timely and relevant. It uses state-of-the-art techniques to reach its conclusions and is generally performed to a high standard. It describes a critically endangered species that poses its unique conservation challenges. There are a number of things to like, as well as some substantial shortcomings in this paper.

Strengths:

The genomic resource is excellent. The assembly is well validated (97.3% anchored to 25 scaffolds, 95.7% BUSCO), and the authors generated a separate male assembly specifically to resolve the sex-determining region, allowing XY-shared and Y-specific contigs to be distinguished on coverage rather than inference. This is truly well done, and at a high standard. The synteny evidence for telomere-to-telomere fusions involving at least 14 ancestral chromosomes, against two in *O. m. masou*, is convincing.

The Hehuan Creek result is the paper's most valuable contribution. Elevated heterozygosity, short and infrequent runs of homozygosity, and private alleles absent from the Qijiawan broodstock are difficult to reconcile with a purely reintroduced origin. The contrast with Luoyewei is a clean and useful cautionary case for hatchery supplementation.

Weaknesses:

(1) The species-rank claim is featured in the abstract, but it is made with any level of rigour in the paper. "New species" appears once, in the abstract (l. 32). The Results conclude only that *O. formosanus* is a distinct evolutionarily significant unit (ll. 188-191), which itself can be well-justified, but it's far from a taxonomic rank (see author's own ref 10). No species concept is explicitly named anywhere, and the taxon is referred to across the manuscript as a subspecies (l. 68), a "new species" (l. 32), and an ESU (l. 189) in turn.

(2) Gene flow is asserted, not tested, and two divergence estimates disagree twentyfold. The abstract reports "no detectable gene flow for ~50,000 years." That figure is a divergence time from BPP under the A00 model, which contains no migration parameter; a model that cannot

fit gene flow cannot report its absence. Separately, Figure 1b shows a split at 1.15-5.09 Mya (Figure 1b), while the ddRAD coalescent places the same split at ~50 kya (Figure 1d). The explanation offered (ll. 417-421, "differing temporal sensitivity of genomic markers") is not a mechanism.

(3) The placement of sdY is unresolved, and the paper's own figures conflict with its text. Figure S10 and Table S6 both make *O. formosanus* chr13 homologous to *O. m. masou* chr32, whereas reference 28 - on which the authors rely - places the sdY contig on *O. m. masou* chr7, whose *O. formosanus* homologue is chr5 (Table S6). These cannot both be correct, and Figure S10's caption compounds the confusion by attributing chr13 to masou and omitting the chr32 track entirely.

(4) The population-viability model's stated mechanisms are contradicted by the authors' own supplementary tables. The Discussion attributes Qijiawan's vulnerability to "lower juvenile survival, decreased fecundity, and narrower terminal age class representation" (ll. 522-525). Table S10 gives Qijiawan higher age-0 survival (0.202/0.616 vs 0.184/0.615); Table S11 gives it higher fecundity at every reproductive age (7.68/19.27/7.86 vs 6.10/8.95/4.14); Table S9 gives it a broader terminal age class (5.0% vs 0.8% age-3 in November). The only parameter favouring Hehuan is age-1 survival - 0.087 (95% CI 0.000-0.180) versus 0.131 (0.093-0.187) under typhoon, and 0.054 (0.000-0.167) versus 0.087 (0.047-0.149) at baseline. Both Qijiawan intervals include zero and overlap Hehuan's, yet a reported extinction odds ratio of 4.48 rests on this difference.

(5) The two streams were not measured equivalently, and every asymmetry favours the conclusion.

<https://doi.org/10.7554/eLife.111444.1.sa0>