

ORIGINAL ARTICLE OPEN ACCESS

Recent Polygenic Adaptation in Heavily Fished Malawi Cichlids

Alexander Hooft van Huysduynen^{1,2}  | Francisco Campuzano Jiménez¹ | Julia Camacho Garcia¹ | Gudrun De Boeck² | Bosco Rusuwa³ | Hannes Svoldal^{1,4} 

¹Evolutionary Ecology Group, Department of Biology, University of Antwerp, Antwerp, Belgium | ²ECOSPHERE, Department of Biology, University of Antwerp, Antwerp, Belgium | ³School of Natural and Applied Sciences, University of Malawi, Zomba, Malawi | ⁴Naturalis Biodiversity Center, Leiden, the Netherlands

Correspondence: Alexander Hooft van Huysduynen (alex.hooftvanhuysduynen@uantwerpen.be) | Hannes Svoldal (hannes.svoldal@uantwerpen.be)

Received: 20 November 2025 | **Revised:** 14 April 2026 | **Accepted:** 27 May 2026

ABSTRACT

Intense fishing pressure can drive rapid evolution in wild populations, yet the underlying genomic mechanisms often remain elusive. Here, we investigate the genomic consequences of five decades of intense harvesting on the cichlid fish *Copadichromis mloti* in Lake Malombe, Malawi. By comparing whole-genome data from almost 200 individuals from the heavily fished Lake Malombe with less-fished populations in Lake Malawi, we identified a widespread signal of recent, population-specific positive selection. This genomic footprint, spanning hundreds of loci, is consistent with a rapid, polygenic adaptive response. The candidate genes under selection were significantly enriched for high-level developmental processes, including neurogenesis and the formation of the heart and muscle, providing a plausible mechanistic link to the dramatic reduction in size and age at maturity observed in this population. Furthermore, our strongest candidate loci include genes with known roles in controlling somatic growth (*zdhhc13*, *myh7*, *csrp3*, *gnb1l*) and reproductive timing (*mppd2a*, *kitlga*, *kat6a*, *stra6*). We also found that candidate genes were nearly five-fold more likely to be differentially expressed in male gonad tissue, providing a strong link between selection and regulatory changes in a key reproductive organ. Our findings point towards a polygenic basis for fisheries-induced evolution in this tropical system and highlight how intense harvesting can rapidly reshape the genetic architecture of complex life-history traits, with important implications for the long-term resilience of exploited fish stocks.

1 | Introduction

Human-induced environmental changes can impose strong selective pressures that drive rapid evolution in wild populations, ultimately determining whether they adapt or face extinction (Allendorf and Hard 2009; Hendry et al. 2017). While such contemporary evolutionary responses are now widely documented across many species (Baltazar-Soares et al. 2021), the underlying molecular mechanisms remain largely unknown (Hutchings and Kuperinen 2021; Pelletier and Coltman 2018). A critical challenge is to move beyond observing phenotypic shifts and identify the specific genetic loci that facilitate adaptation to anthropogenic stress. Uncovering this genetic basis is fundamental for distinguishing heritable evolutionary change

from phenotypic plasticity and for building a mechanistic understanding of how populations respond to anthropogenic pressures.

Classic examples of rapid evolutionary responses to human activities include the spread of industrial melanism in peppered moths (Hof et al. 2016), the reduction of tusk size in elephants due to ivory poaching (Campbell-Staton et al. 2021), and decreased horn size in bighorn sheep targeted by trophy hunting (Pigeon et al. 2016). Perhaps the most widespread example is fisheries-induced evolution (FIE), which describes shifts in the life-history traits of commercially harvested fish. This phenomenon has been documented across diverse aquatic systems and taxa, with compelling field, experimental

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *Molecular Ecology* published by John Wiley & Sons Ltd.

and theoretical evidence demonstrating that intense fishing pressure selects for traits like earlier maturation at a smaller size (Conover and Munch 2002; Heino et al. 2015; Kuparinen and Festa-Bianchet 2017).

Despite strong evidence for phenotypic change, identifying the genetic basis of FIE has proven exceptionally challenging (Han et al. 2025; Hutchings and Kuparinen 2021). This knowledge gap limits our understanding of the molecular pathways driving these rapid life-history shifts and understanding of whether they represent heritable adaptation or phenotypic plasticity. For instance, a temporal genomic study of Atlantic cod, a classic example of FIE, found no evidence of strong, fisheries-driven genomic change (Pinsky et al. 2021). The authors suggested that the observed life-history shifts were likely due to highly polygenic adaptation or plastic responses, rather than selection on a few major-effect loci. This highlights the complexity of FIE, where confounding factors such as changes in population density, climate or predation may also contribute significantly to phenotypic shifts (Hutchings and Kuparinen 2021).

While most studies of fisheries-induced evolution have focused on marine systems or temperate fish species, tropical freshwater taxa offer a vital but understudied context for this research (Heino et al. 2015). The cichlid fishes of Lake Malawi provide a compelling system to investigate rapid, human-driven evolutionary change. Fisheries in Lake Malawi and its peripheral lakes are a cornerstone of the regional economy, constituting 40% of all dietary animal protein and 4% of Malawi's national GDP (M'balaka et al. 2024). However, this vital resource is under threat, as many of the lake's over 800 cichlid species (Konings 2016) have experienced dramatic population declines due to intense artisanal and commercial fishing (M'balaka et al. 2024; Turner 1995). Species in shallow habitats have faced increasing fishing pressure over the last five decades, creating a strong selective landscape. This pressure is not uniform, varying substantially across the range of many species and reaching its most extreme intensity in Lake Malombe.

Lake Malombe is a young (<200 years old), shallow and highly productive water body at the outflow of Lake Malawi (Figure 1a), making it particularly vulnerable to environmental change. In the 1980s, the introduction of purse seine fishing (locally called 'Nkatcha') led to extremely intense and indiscriminate harvesting of fish of all sizes. This practice first caused the collapse of the local tilapia ('Chambo') fishery around 1990 and has since been directed at the lake's diverse haplochromine cichlids, threatening many with extinction (Hara et al. 1999; Weyl et al. 2004; Zwieter et al. 2011). While Nkatcha fishing was restricted under fisheries regulations in Lake Malawi and is largely impractical across most of the lake's greater depths, it remained prevalent in the shallow waters of Lake Malombe (Makwinja et al. 2021; Weyl et al. 2004; Zwieter et al. 2003). The direct consequence of this method has been a stark decline in fish stocks; despite increasing fishing effort, total catch biomass plummeted from a peak of ~10,000 t per year in 1990 to less than 1000 t per year in the early 2000s and a stunted recovery since (Figure 1b) (Makwinja et al. 2021; Zwieter et al. 2003).

To investigate the genomic consequences of this intense harvest, we focused on *Copadichromis mloto*, a species of high market value that used to be abundant and economically significant in the fisheries of both Lake Malawi and Lake Malombe. Coinciding with the increase in fishing pressure in Lake Malombe, the local *Copadichromis mloto* population has shown a strong decline in relative abundance and a marked reduction in size and age at maturity (see Figure 1c) (Makwinja et al. 2021; Tweddle et al. 1995). This phenotypic shift, consistent with males reproducing in their first year compared to in their third year in Lake Malawi (Iles 1971), represents a classic signature of FIE, whereby high mortality selects for earlier maturation at a smaller body size. These observations strongly suggest that the Lake Malombe population of *C. mloto* has been under intense selection. Therefore, in this study, we aim to identify genomic regions under recent selection and highlight candidate genes plausibly involved in these critical life-history shifts. By comparing whole-genome and transcriptome data from Lake Malombe with that of less intensely fished populations from Lake Malawi, we investigate the genomic basis of this apparent rapid adaptation.

2 | Materials and Methods

2.1 | Sample Collection and Sequencing

Copadichromis mloto ($N=224$) were obtained from local fishers at landing sites across southern Malawi during expeditions in 2002, 2016 and 2017. For genomic analysis, a right pelvic fin was excised from each individual, preserved in absolute ethanol, and stored at -80°C . Whole-body voucher specimens were also collected, fixed in 10% formalin, preserved in ethanol, and are housed at the Royal Museum for Central Africa, Belgium (2002 collection) and the Museum of Zoology, University of Cambridge, UK (2016–2017 collections). Samples have previously been included in Sawasawa et al. (2024) and assigned to one of three *C. mloto* clades (A, B, C) or re-classified as *C. virginalis*. DNA was extracted from fin clips using standard protocols and short reads (100 to 150 bp paired-end) were obtained on Illumina HiSeq platforms as described in Sawasawa et al. (2024) and Blumer et al. (2025) reaching genomic coverages between 7.2 $\times$ and 36.8 $\times$ (median of 14.5 $\times$) (Table S1).

2.2 | Alignment and Variant Calling

Sequencing reads were mapped to the *Astatotilapia calliptera* reference genome (fAstCal1.2, accession GCA_900246225.3, https://www.ncbi.nlm.nih.gov/assembly/GCF_900246225.1) using BWA-MEM (version 0.7.17; Li 2013). *C. mloto* & *A. calliptera* belongs to the Malawi haplochromine cichlid radiation and these species share low sequence divergence of ~0.096% (Malinsky et al. 2018) comparable to within human diversity (0.075%–0.1%; 1000 Genomes Project Consortium 2015). Mean mapping rates of 98.38% across samples (range: 97.4%–98.9%) are consistent with values reported for within-radiation cichlid studies, indicating that cross-species mapping artefacts are unlikely to substantially confound our results (Malinsky et al. 2018). Variant calling was conducted with bcftools (v1.14; Danecek et al. 2021). To ensure high-quality variant calls, sites

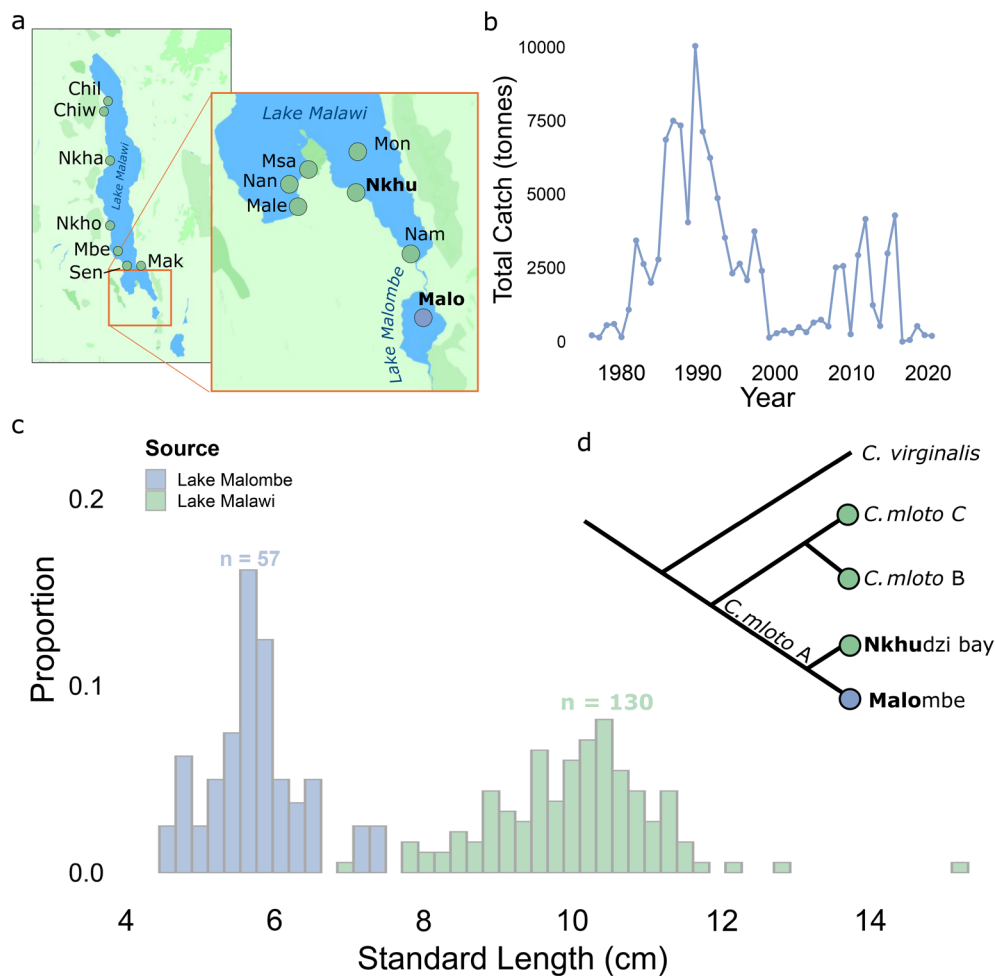


FIGURE 1 | Geographic, ecological and phylogenetic context of the study populations. (a) Map showing geographic context and sampling locations in Lake Malawi (green points) and Lake Malombe (blue point) used in this study. Denoted sampling locations are specified as follows: Chilumba (Chil), Chiweta (Chiw), Nkhata Bay (Nkha), Nkhotakota (Nkho), Mbenji Islands (Mbe), Senga Bay (Sen), Msaka (Msa), Nankhwali trawler (Nan), Malembo (Male), Monkey Bay (Chenga trawler) (Mon), Nkhudzi Bay (Nkhu), Namias-Palm Beach (Nam), Lake Malombe (Malo) (b) Temporal trend in total annual catch from Lake Malombe fisheries of haplochromine cichlids, data from: Makwinja et al. (2021). (c) Standard length distributions of *Copadichromis mloto* from field samples collected in Lake Malombe (blue bars) and Lake Malawi (green bars) across several field expeditions since 2016. (d) Scheme of phylogenetic relationships of *Copadichromis mloto* complex, including clades A, B, C as defined in Sawasawa et al. (2024).

were masked under the following conditions: if more than 10% of mapped reads had zero mapping quality, if the overall mapping quality was below 50, if mapping quality was significantly biased between strands ($p < 0.001$), or if total read depth across all samples was either above the 97.5th percentile or below the 2.5th percentile. Additionally, heterozygous sites were filtered out when a binomial test indicated a significant allele balance bias (PHRED > 20). Positions with evidence of excess heterozygosity (Inbreeding Coefficient < 0.2) or with more than 20% missing data were also excluded. Only single nucleotide polymorphisms (SNPs) were retained for further population genetic analyses. Following these filtering steps, the dataset included approximately 41.7 million biallelic SNPs spanning 22 chromosomes. The variant call format (VCF) file of whole-genome biallelic SNP dataset is available on Zenodo: (<https://doi.org/10.5281/zenodo.17657791>). For the selection statistics analysis below, the *C. mloto* callset was complemented by population SNP variation data from 45 additional species from Blumer et al. (2025) (Table S2).

2.3 | Genome Wide Selection Statistics

Tajima's D was calculated in 10kb windows for all individuals of each defined population with vcfTools (v0.1.16; Danecek et al. 2011). A normalized Fay and Wu's H calculation (Zeng et al. 2006) derived from the allele frequency spectra of each population was calculated in a custom python script available on Zenodo (<https://doi.org/10.5281/zenodo.17649119>).

2.4 | Genome Scans for Malombe-Specific Signals of Selection

To identify Malombe-specific allele frequency shifts we used the Population Branch Statistic (PBS) developed by Yi et al. (2010) based on the per SNP pairwise F_{ST} values of the Malombe, Nkhudzi bay and Namias Palm beach populations ('Malo', 'Nkhu' & 'Nam', respectively in Figure 1a). PBS gives an estimate of F_{ST} differentiation since the last common ancestor of

Malombe and sister population Nkhudzi Bay (Figure 1d) allowing us to determine SNP differentiation unique to the Malombe population. The Namiasi Palm Beach population is used as an outgroup in the PBS calculation, which despite its geographic proximity to Lake Malombe phylogenetically belongs to the *C. mloto* B sister clade (Figure 1d; Sawasawa et al. (2024)). Weir and Cockerham (1984) F_{ST} values were estimated using vcfTools (v0.1.16; Danecek et al. 2011).

To detect hard and soft selective sweeps across the genome we explored the haplotype homozygosity in the Malombe population around each SNP using the software *H*-scan (version 1.3, downloaded from: <https://messerlab.org/resources/>; Schlamp et al. 2016). We performed *H*-scan on the Malombe population and the Nkhudzi bay populations and increased the specificity for Malombe-specific homozygosity patterns by subtracting the observed *H*-statistic of the Nkhudzi bay population from the *H*-statistic estimated for the Malombe population. This modification isolates the selective sweep signals that are unique to the Malombe population (i.e., occurred after its divergence from the Nkhudzi bay population).

To integrate complementary signals of recent selection, we computed a Fisher's Combined Score (FCS) per SNP by combining the percentile ranks of the PBS and the normalized *H*-value from *H*-scan, following Choin et al. (2021). For each SNP, we computed the negative base-10 logarithm of the percentile rank for PBS and *H* separately, then summed these to obtain per-SNP FCSs. This method prioritizes SNPs with strong rank-based signals in both selection scans while reducing dependence on raw magnitudes.

2.5 | Neutral Demographic Simulations

To evaluate the statistical significance of observed selection scan results, we estimated a distribution of the FCS statistic under a neutral demographic scenario using msprime (v1.3.3; Baumdicker et al. 2022). Specifically, we simulated three-population genealogies under the standard coalescent and superimposed neutral mutations to produce VCF files that were later processed identically to empirical data. Model parameters were based on genome-wide estimates derived from our dataset and used recombination and mutation rate estimates from earlier studies on Malawi cichlids (Malinsky et al. 2018; Talbi et al. 2024). Effective population sizes were estimated from π (nucleotide diversity) using vcfTools --window-pi and the relation $N_e = \pi / (4\mu)$, with a mutation rate of $\mu = 3.5 \times 10^{-9}$. Recombination rate was fixed at a genome-wide average of 6.17×10^{-9} , based on linkage map data for *A. calliptera* (Talbi et al. 2024). The mutation rate used in simulations was 3.5×10^{-9} , as previously estimated in Malawi cichlids (Malinsky et al. 2018). Split times were derived by subtracting the average within species coalescent times from between species coalescent times. Simulations were run over 1 Mb windows and summary statistics (PBS, H, FCS) were calculated per SNP for comparison with empirical values. Although residuals were consistently small (< 0.06 FCS points) across most of the matched percentile between the simulated and empirical FCS distributions, the upper tail showed a pronounced deviation towards higher empirical values (See Results: Figure 3c

and Figure S1). This deviation started at an FCS of 5.512, corresponding approximately to the 99.99th percentile of the empirical values, and increased to an extreme residual exceeding 2.5 FCS points. Based on this trend we established the FCS value of 5.512 as an empirical selection threshold in further analyses. Scripts and configuration files for demographic simulations are available on Zenodo (<https://doi.org/10.5281/zenodo.17649119>).

2.6 | Functional Annotation and Enrichment

To assess the potential functional consequences of selection on specific variants, we annotated all variants surpassing the empirical selection threshold with functional effect categories using SnpEff (v5.2.1; Cingolani et al. 2012) based on an annotated genome of *A. calliptera* (fAstCal1.2, ensemble release 105; Dyer et al. 2025) which contains 28,001 coding-genes. Where ensembl stable IDs had not yet been assigned a gene name in this release, the name resolved in the updated fAstCal1.3 annotation (Ensembl release 112) is given in parentheses.

To determine whether the observed proportions of functional annotations (e.g., synonymous, non-synonymous, etc.) differed from expectation, we computed an empirical distribution of annotation effects in minor allele frequency (MAF)-matched set of SNPs using a custom Python script (script available on Zenodo: <https://doi.org/10.5281/zenodo.17649119>). For each SNP surpassing the selection threshold, control SNPs of similar MAF were randomly sampled from the full SNP dataset, over 1000 iterations. Significant overrepresentation in the proportions of functional annotation categories was assessed using binomial tests; *p*-values were corrected for multiple testing using Benjamini-Hochberg false discovery rate (FDR) considering results significant at $FDR < 0.05$.

Gene Ontology enrichment analysis was performed on genes located within candidate regions, defined as genomic regions extending ± 25 kb around each locus under selection, reflecting the typical range of linkage disequilibrium in Malawi cichlids. Genes located within these regions were considered candidate genes. We conducted a Gene Ontology (GO) enrichment analysis using the topGO package (v2.38.1; Alexa and Rahnenfuhrer 2023) in R/Bioconductor (www.bioconductor.org). As a background, we used all annotated protein-coding genes from the *A. calliptera* reference genome (fAstCal1.2, ensemble release 105, Dyer et al. 2025) and mapped to GO term assignments of *Dania rerio* (org.Dr.eg.db, v3.19.1). In total, 11,522 genes were successfully mapped to GO terms within the Biological Process (BP) ontology. The minimum category size was set to five genes (nodeSize = 5). Enrichment significance was evaluated using Fisher's exact test implemented in the *weight* algorithm of topGO ('weightCount' class). This algorithm improves interpretability by accounting for hierarchical relationships among GO terms (Alexa et al. 2006). Reported *p*-values (Table S3) are not explicitly adjusted for multiple testing, but the authors note that for the topology aware algorithm used, *p*-values can be regarded as effectively corrected, as the methods account for GO-term dependencies (Alexa and Rahnenfuhrer 2023).

2.7 | Differential Gene Expression Analysis

Tissue samples were collected from five tissues (brain, gill, gonad, liver, muscle) from six *C. mloto* males in breeding coloration from Msaka (Lake Malawi) and six males in breeding coloration from Lake Malombe. Differential expression analyses were restricted to males in breeding coloration to minimize expression variation associated with sex or reproductive state that would confound differential expression analyses across populations. Samples were stored in RNAlater at -80°C , then homogenized in Trizol with zirconia/silica beads. RNA was extracted using the Direct-Zol RNA Purification Kit (Zymo Research), and quality assessed with Qubit and TapeStation (Agilent). RNA libraries (75bp paired-end) were sequenced at the Wellcome Sanger Institute on an Illumina HiSeq 4000. RNA-seq was performed on a subset of 12 individuals also included in the whole genome sequencing dataset (Table S1); RNA and DNA library preparation were conducted independently, ensuring no shared technical biases between analyses.

The differential gene expression analysis comparing patterns of expression between the two populations was conducted using the RASflow pipeline (Zhang and Jonassen 2020). RNA-seq reads were aligned to the *Astatotilapia calliptera* reference genome (fAstCal1.2) using HISAT2, and gene-level counts were obtained using featureCounts (Liao et al. 2014). Differential expression analysis was conducted with PyDESeq2 (Muzellec et al. 2023), the Python implementation of DESeq2, applying $\log_2$ fold change >1 and FDR-adjusted p -value <0.01 as significance thresholds. To test whether differentially expressed genes were significantly enriched among genomic candidate genes from selection scans, we quantified pairwise overlaps and compared them to expectations under a random model using a hypergeometric test. For each DE gene list and selection analysis candidate gene set we compared the observed number of overlapping genes (O) to the expected overlap under random association ($E = A \times B/N$) where A and B are the sizes of the two gene sets being compared and N is the total number of genes tested. Enrichment is expressed as an observed/expected (O/E) ratio and p -values were corrected for multiple testing using the Benjamini-Hochberg FDR procedure.

3 | Results

3.1 | Recent Positive Selection in the Lake Malombe Population

To investigate the evolutionary forces shaping the heavily fished Lake Malombe population, we first analysed genome-wide allele frequency spectra for all *C. mloto* populations. We found that all populations exhibited negative mean values for Tajima's D , indicating a general excess of rare variants. While consistent with either population expansion or positive selection, the Lake Malombe population showed the most negative Tajima's D value, suggesting it has experienced the strongest demographic or selective event (Figure 2a).

To disentangle these two possibilities, we calculated Fay and Wu's H (H_{FW}), a statistic that detects the excess of high-frequency

derived alleles characteristic of recent selective sweeps. In this analysis, the Lake Malombe population was a clear outlier; it was the only *C. mloto* population to show a strongly negative H_{FW} value (Figure 2b). To place this result in a broader phylogenetic context, we reanalyzed SNP data from Blumer et al. (2025), computing Tajima's D and H_{FW} for all cichlid populations with at least five sequenced individuals (Figure 2c). While Tajima's D and H_{FW} values varied considerably across species, the Lake Malombe population remained a notable outlier: it showed the most negative Tajima's D value among all species and populations and a significantly negative H_{FW} , despite belonging to the species with overall most positive H_{FW} values.

In summary, the combination of a negative Tajima's D with a negative H_{FW} specific to the Malombe population of *C. mloto* provides compelling evidence that recent positive selection, in the form of pervasive selective sweeps, has occurred in the population since its divergence from the Lake Malawi populations.

3.2 | Elevated Signatures of Selection Across the Malombe Genome

To further confirm that the Lake Malombe *C. mloto* population has been subject to recent, lineage-specific selection, we performed genome-wide selection scans using the Population Branch Statistic (PBS) and H -scan. Quantile-quantile plots showed that the Lake Malombe population had globally elevated values for both selection statistics compared to identical analyses for its sister-population from Lake Malawi's Nkhudzi Bay (Figure 2d,e). This indicates that, on average, allele frequency differentiation and haplotype lengths are greater in the Malombe population. This genome-wide pattern provides additional support for the conclusion that positive selection has been more pervasive or intense in the Lake Malombe lineage following its split from Lake Malawi populations. More critically, the extreme upper tail of both distributions was highly divergent between the sister populations, with higher values in the Malombe population. This extreme tail of the distribution corresponds to the specific genomic regions expected to have experienced the strongest selection.

3.3 | Genomic Regions Under Malombe-Specific Selection

To pinpoint specific genomic regions under recent selection in the Lake Malombe population, we combined population differentiation (PBS) and Malombe-specific haplotype homozygosity (H -scan) statistics into a Fisher's Combined Score (FCS), allowing us to detect loci with both Malombe-specific allele frequency shifts and signatures of selective sweeps (see Materials and Methods for details).

A genome-wide scan of FCS values revealed numerous outlier regions across the genome (Figure 3a). Among these signals, ten major peaks stood out as the most prominent candidates for recent, strong selection. Each peak corresponds to a genomic region extending approximately 25kb on either side of clustered outlier SNPs. These ten regions collectively encompass 24 annotated genes, namely; *slc22a18*, *nr2e3*,

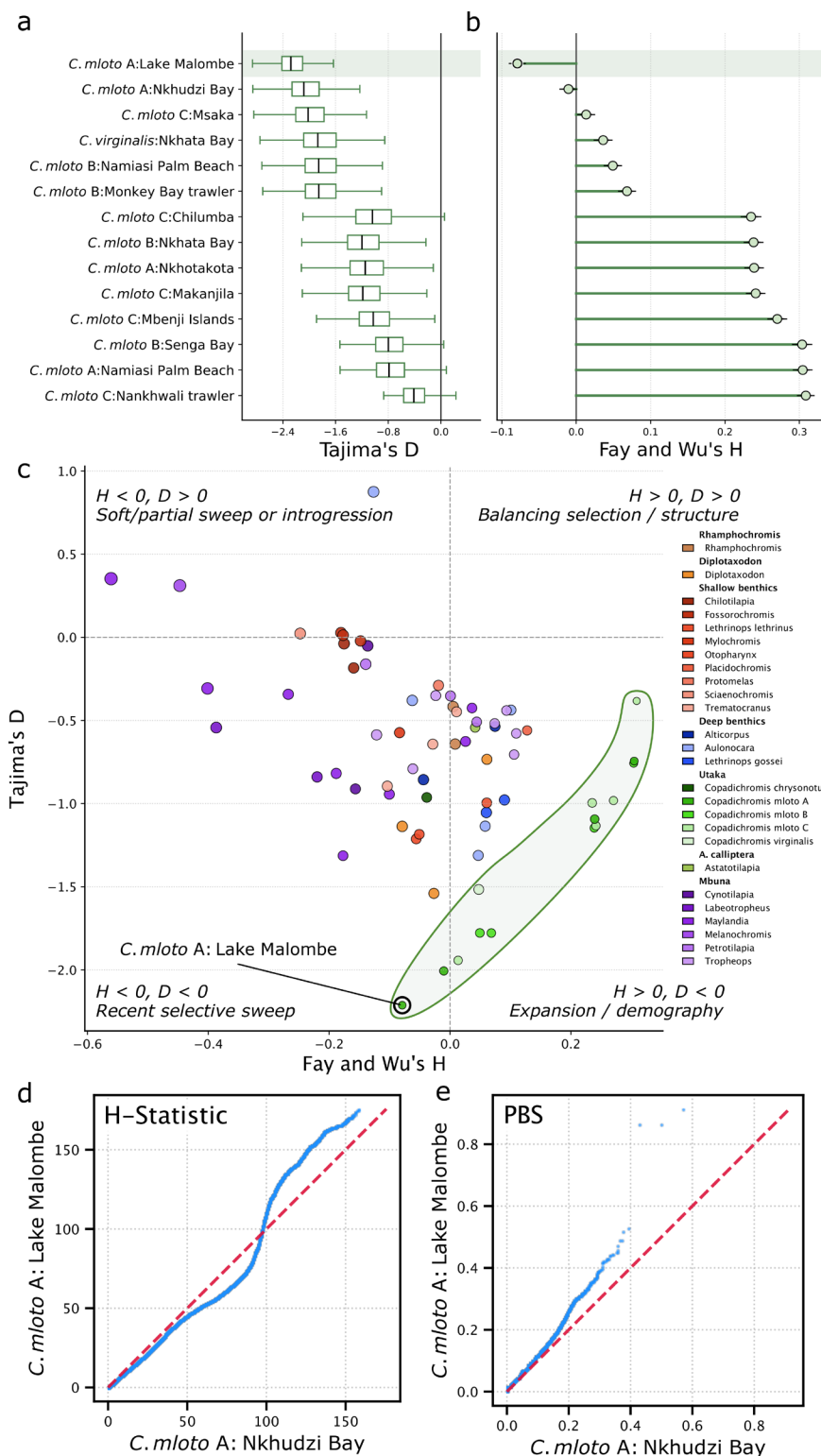


FIGURE 2 | Genome-wide summary statistics and selection signals in *Copadichromis mloto* (a, b) Distributions of Tajima's D (a) and Fay and Wu's H (b) across populations of the *C. mloto* / *C. virginalis* complex. Clades A, B and C are denoted as defined in Sawasawa et al. (2024). (c) Genome-wide Tajima's D and Fay and Wu's H values in a broader phylogenetic context of Malawi cichlids. Genera are grouped by clade as defined in Blumer et al. (2025). Populations of the *C. mloto* complex are highlighted in green, with the Lake Malombe population annotated. (d, e) Quantile–quantile (Q–Q) plots comparing the Lake Malombe and Nkhudzi Bay *C. mloto* populations for the distributions of the H -statistic (d) and Population Branch Statistic (PBS; e).

ENSACLG00000001162 (*stra6*), ENSACLG00000001233 (*islr2*),
 mpped2a, zdhhc13, csrp3, ENSACLG00000016851, ank2,
 ENSACLG00000009480, ENSACLG00000013470, si:dkey-12e7.4,

ENSACLG00000021888 (*cnksr1*), ENSACLG00000021917
 (myh7), ENSACLG00000022020, ENSACLG00000022038
 (nlrc5), kitlga, kat6a, ENSACLG00000003998, tbx1, gnb1l,

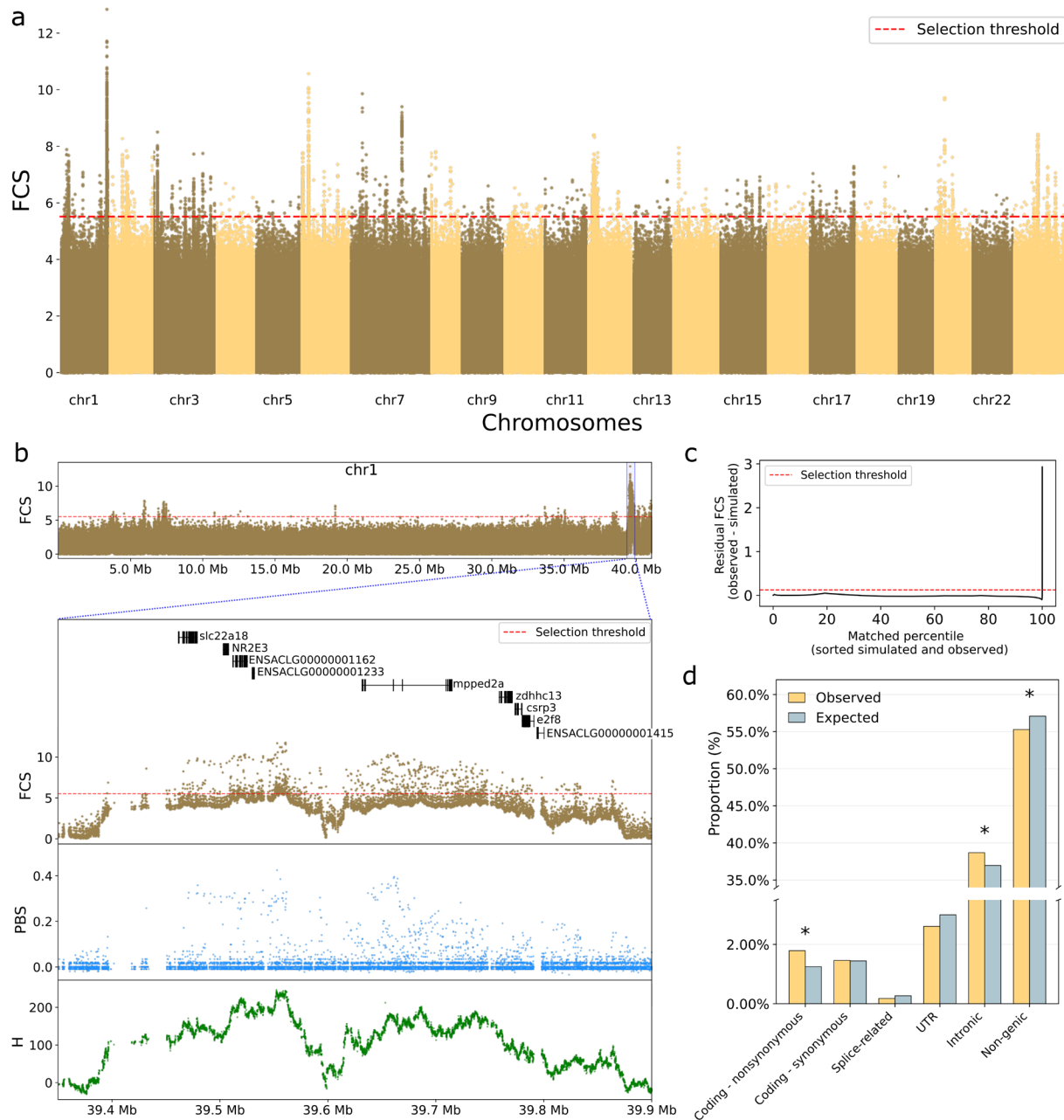


FIGURE 3 | Characterization of selection signals in the Lake Malombe population. (a) Manhattan plot showing positive values of Fisher's Combined Score (FCS) per SNP across the genome. The dashed red line indicates the empirical selection threshold derived from neutral demographic simulations (FCS = 5.51). (b) Detailed view of the strongest selection peak on chromosome 1, showing positions of gene exons and symbols, as well as the SNP score for FCS, Population Branch Statistic (PBS) and *H* statistic (*H*). (c) Residual FCS value (observed—simulated) plotted against the ranked/matched simulated and empirical FCS value. The dashed red line along y axis indicates the threshold surpassed in the 99.99 percentile of empirical observations. (d) Bar plot showing proportional distributions of functional categories among SNPs surpassing the empirical selection threshold, and expectation from minor allele frequency (MAF) matched resampling. Asterisks (*) denote categories with significant overrepresentation (FDR < 0.05, binomial test).

ENSACLG00000024047, *magi1*, *ENSACLG00000008501* (*si:ch211-108d22.2*). Detailed plots showing the Fisher's Combined Score (FCS), *H*-score and PBS for the highest peak on Chromosome 1 are presented in Figure 3b (see Figures S2–S10 for other major peaks). Intriguingly, functional annotations for genes among these ten major peaks reveal their involvement in key developmental processes, including somatic growth, muscle development and the timing of reproductive maturation (see Discussion).

3.4 | A Widespread Signal Consistent With Polygenic Adaptation

While the most prominent peaks highlight major targets of selection, a key question is what proportion of the genome shows evidence of adaptation. To address this and identify a comprehensive set of candidate SNPs, we established a significance threshold that separates signals of positive selection from neutral background variation. We achieved this by comparing our

TABLE 1 | Intersection between genes identified as candidates for selection in genome scans ('All Candidates') and the subset of genes containing SNPs surpassing the empirical selection threshold in intronic regions ('Intronic') and non-synonymous sites ('ProtChanges') with differentially expressed (DE) genes detected in each tissue. Columns indicate the number of genes per set (Selection candidates and DE genes), the number of overlapping genes (Overlap) and the number expected by chance based on random resampling (Expected).

Selection candidate gene subset	DE tissue	# Selection candidate genes	# DE genes	Overlap	Expected (CI: low-High)	Fold	Adjusted <i>p</i>
All candidates	Gonads	1155	197	18	3.8 (1–8)	4.74***	3.3e-05
All candidates	Muscle	1155	4544	215	48.5 (37–61)	4.43***	3.3e-05
All candidates	Liver	1155	3620	174	39.6 (29–51)	4.39***	3.3e-05
All candidates	Brain	1155	1428	72	20.7 (13–29)	3.48***	3.3e-05
All candidates	Gills	1155	3821	209	59.7 (47–73)	3.50***	3.3e-05
Intronic	Gonads	301	197	5	1.0 (0–3)	5.14**	0.0028
Intronic	Brain	301	1428	27	2.2 (1–10)	5.10***	3.3e-05
Intronic	Liver	301	3620	54	10.8 (5–17)	4.99***	3.3e-05
Intronic	Muscle	301	4544	65	13.2 (7–20)	4.88***	3.3e-05
Intronic	Gills	301	3821	63	15.1 (9–22)	4.17***	3.3e-05
ProtChanges	Liver	37	3620	9	2.3 (0–5)	3.99***	8.6e-05
ProtChanges	Gills	37	3821	12	3.2 (0–7)	3.77***	3.3e-05
ProtChanges	Muscle	37	4544	8	2.7 (0–6)	2.94**	0.0022
ProtChanges	Brain	37	1428	3	1.1 (0–3)	2.74	0.1024
ProtChanges	Gonads	37	197	0	0.2 (0–1)	0	1.0

Note: 'CI: Low' and 'CI:High' denote the 95% confidence interval bounds for the expected overlap. 'Fold' represents the ratio of observed to expected overlap, with asterisks indicating empirical significance levels (*p* values from 50,000 randomizations; ***FDR < 0.001, **FDR < 0.01). Full summary of overlapping genes in Table S5.

empirical FCS distribution to a null distribution generated from neutral demographic simulations (see Material and Methods), which allowed us to define a stringent 'empirical selection threshold' at FCS > 5.51 corresponding to the top 99.99% of observations (Figure 3c; Figure S1).

A total of 4185 SNPs, located within 473 genomic regions, surpassed this threshold (Figure 3a). The widespread genomic footprint of outlier regions is consistent with a polygenic adaptation model, where selection acts on many loci across the genome. Further supporting the functional relevance of the outlier SNPs, we found that they were significantly enriched for non-synonymous (amino acid changing) mutations but not enriched for synonymous mutations, indicating that selection has targeted functional variants (Figure 3d). Intronic variants were also significantly enriched among candidate selection SNPs, consistent with selection acting on cis-regulatory variation and suggesting that gene expression changes contribute to the adaptive response alongside protein-level divergence (Figure 3d).

3.5 | Candidate Genes Are Enriched for Differential Expression Across Tissues

To test whether candidate regions were also associated with regulatory changes, we performed a differential expression

analysis between Lake Malombe and Lake Malawi individuals across five tissues (see Table S4 for differential expression results for each tissue). Although differential expression and non-synonymous sequence changes represent distinct molecular outcomes, overlap between them can reveal genes where regulatory and protein-level changes converge and provides stronger evidence for compounded involvement in adaptive divergence. Candidate genes were significantly more likely to be differentially expressed than non-candidate genes in all tissues tested (Table 1, Figure 4). Notably, this enrichment was strongest in gonad, muscle and liver tissue, where candidate genes were between 4.4 and 4.7-fold more likely to show differential expression than expected by chance (Table 1, Table S5). This strong link between signatures of selection and gene expression changes, particularly in the gonads, provides functional support for the role of these loci in the adaptive divergence of life-history traits in the Lake Malombe population.

Genes harbouring candidate SNPs in intronic regions showed even stronger enrichment, between 4.2 and 5.1-fold across all five tissues (Table 1, Figure 4). Genes with non-synonymous candidate SNPs were also significantly enriched for differential expression in liver, gills and muscle (2.9 to 4.0-fold; Table 1). The large number of differentially expressed genes together with the observed significant enrichment of selection signals in intronic variants suggests an important contribution of regulatory changes driving this rapid adaptive response.

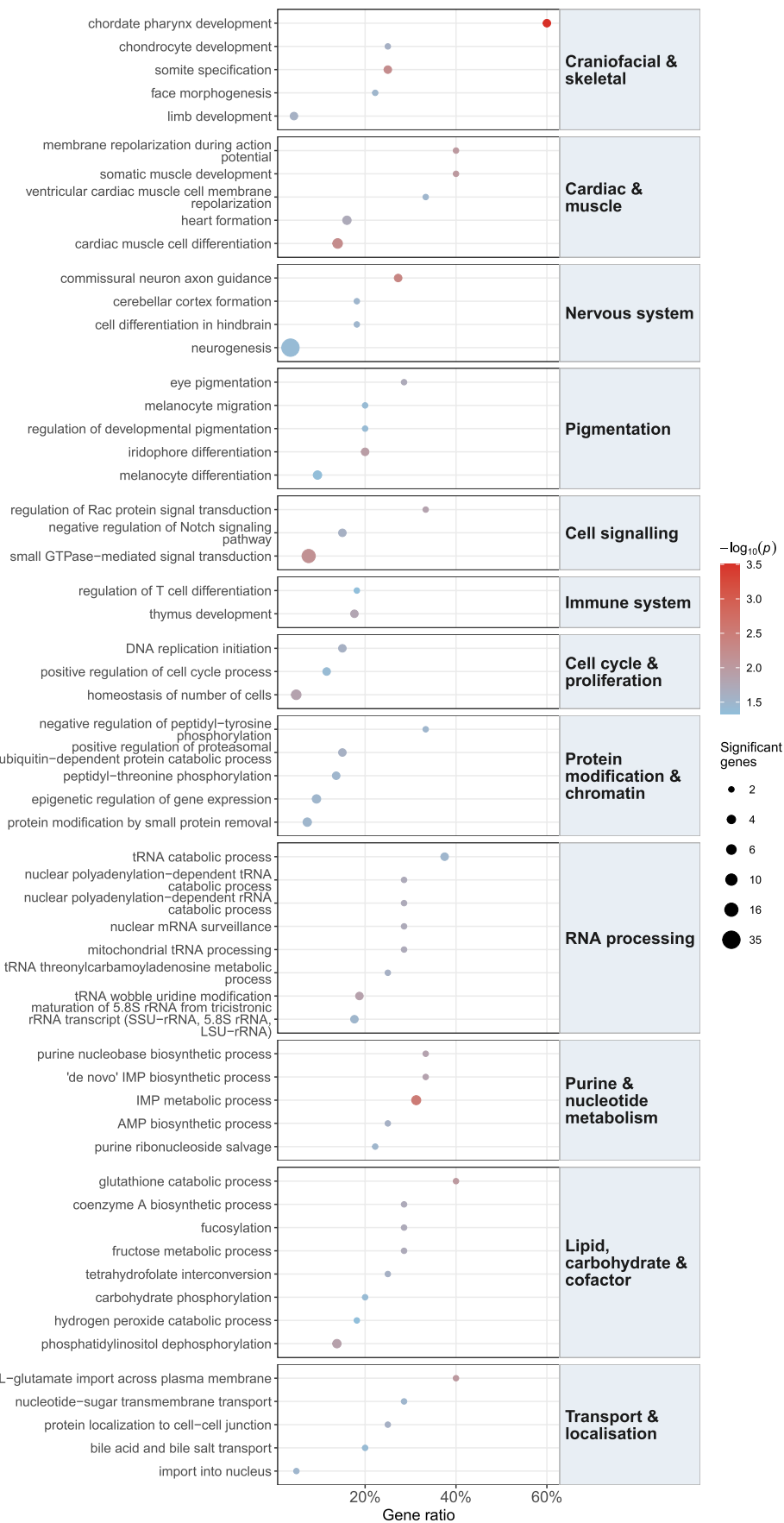


FIGURE 5 | Legend on next page.

FIGURE 5 | Gene Ontology enrichment of 1155 candidate genes. Dot plot showing significantly enriched Biological Process GO terms among genes ± 25 kb of a SNP surpassing the empirical selection threshold. Each dot represents one GO term, grouped into manually curated functional clusters. Dot size reflects the number of significant genes annotated to that term; dot colour indicates statistical significance ($-\log_{10}$ Fisher's exact test p -value). The x-axis shows the gene ratio (significant genes annotated to the term/total genes annotated to the term). Only terms with Fisher's exact test $p < 0.05$ are shown.

partially heritable (e.g., see Therkildsen et al. (2019) for empirical evidence), an evolutionary response is expected under the extreme mortality imposed by intense fishery. Therefore, we conclude that while the Malombe population has been shaped by multiple factors, the widespread signatures of selection on growth and developmental pathways are compelling candidates for the footprint of recent fisheries-induced evolution. Future studies could aim to quantify the heritability of these life-history traits, leveraging the fact that—in contrast to most marine and temperate FIE targets—cichlid fishes are generally easy to breed in captivity and have relatively short generation times.

The widespread distribution of outlier loci across the genome strongly indicates a polygenic architecture of adaptation in the Lake Malombe population. This aligns with findings from experimental FIE systems, where parallel phenotypic responses to size selection were primarily driven by small allele frequency shifts across many genes (Therkildsen et al. 2019). The selective pressure in Lake Malombe, however, is distinct. The indiscriminate 'Nkatcha' fishery imposes high mortality without regard to size, a regime predicted by life-history theory to favour a fundamental shift in energy allocation away from long-term growth and towards rapid, early-life reproduction—the classic r/K trade-off (Heino et al. 2015; Michod 1979).

Given the recent and intense nature of this pressure, the observed polygenic response likely reflects selection on standing genetic variation, resulting in numerous incomplete soft sweeps across the genome. Our methodological approach is particularly well-suited to detect such signals; by contrasting the H -statistic against the sister population, we specifically targeted signatures of recent, Malombe-specific sweeps, combining this with population differentiation in our FCS score.

Our results further suggest that this adaptation is achieved through a dual mechanism at the molecular level. The strong enrichment of our outlier set for non-synonymous mutations indicates that direct protein-coding changes are a key component of the adaptive response. Simultaneously, the significant enrichment of candidate genes among those that are differentially expressed points to the critical role of regulatory evolution. This regulatory signal was particularly striking in gonad tissue, where candidate genes were almost five-fold more likely to be differentially expressed than expected by random association. This provides a direct link between the selected loci and the organs central to reproductive timing and reinforces the hypothesis that fishing has selected for an accelerated life history. Exploring expression patterns at different life history stages and in female ovaries and other tissue in future studies may allow for a comprehensive understanding of the regulatory control and biological process involved in early maturation.

A widespread polygenic signal does not preclude the existence of individual loci with larger effects. For complex traits, the underlying genetic architecture is often characterized by a distribution of effect sizes, with many genes of small effect and a tail of a few genes with more substantial impacts. This 'mixed model' of adaptation is powerfully illustrated by the above-mentioned experimental evolution in Atlantic silversides (Therkildsen et al. 2019). While a parallel polygenic response was observed across all replicate lines, some lines also featured strong, idiosyncratic selective sweeps on large chromosomal blocks, demonstrating that both modes of evolution can occur simultaneously in response to the same pressure. Similarly, in wild Eastern Baltic cod, a combined association and selection scan approach successfully identified several key loci under strong directional selection contributing to changes in growth, even against a complex genetic background (Han et al. 2025).

Our findings in *C. mloto* are consistent with this mixed architecture. Against the backdrop of the widespread polygenic signal we detected, our analysis also identified prominent FCS peaks that likely represent the loci with the largest effects on FIE-related traits in this population. Among the ten most prominent FCS peaks, several genes emerge as compelling candidates with established roles in pathways directly relevant to FIE. In particular, *mppd2a*, *kitlga*, *kat6a* and *ENSACLG00000001162* (*stra6*), have functions linked to the timing of reproductive development. In humans, the region containing *mppd2a* is associated with puberty timing as the gene is understood to influence the development of gonadotropin-releasing hormone (GnRH) neurons, which are central to the neuroendocrine control of maturation via the HPG axis (Laik et al. 2018; Lund et al. 2020). Even more directly, *kitlga* is a well-characterized regulator of oocyte growth and meiotic arrest in zebrafish (Yao and Ge 2013, 2015). *Kat6a*, a histone acetyltransferase, regulates gonadal development and primordial follicles activation in mammals (Zhang et al. 2025), providing an epigenetic route through which selection could accelerate the onset of maturity. Likewise, *stra6*, which encodes a receptor mediating retinol uptake and retinoic acid signalling, could influence gametogenesis and developmental rate through altered retinoid metabolism, a pathway known to affect both oocyte development and puberty onset in teleosts (Medina et al. 2019; Pradhan and Olsson 2015). Notably, independent support for the importance of neuroendocrine regulation of maturation comes from the set of genes carrying non-synonymous candidate SNPs. Among these, *npffr1*, encoding the receptor for gonadotropin-inhibitory hormone (GnIH), harbours a candidate C $\rightarrow$ T missense variant on chromosome 12 (position chr12:5,310,593). GnIH acts as a conserved inhibitor of gonadotropin release across vertebrates, suppressing the HPG axis by directly targeting GnRH neurons via *npffr1* (Dufour et al. 2020; Son et al. 2019). A role that has been demonstrated in cichlid fish specifically (Di Yorio et al. 2016). A functional change in *npffr1* could therefore shift the threshold

for reproductive activation in the Malombe population. The convergence of protein-level selection on *npffr1* with the regulatory signal at the *mppd2a* locus, both targeting the same neuroendocrine axis strengthens the hypothesis that accelerated HPG axis activation is a primary molecular target of fisheries-induced selection in this population. Together, these genes suggest that selection in the Lake Malombe population may have acted not only on specific reproductive regulators but also on higher-order epigenetic and signalling processes that coordinate the timing of maturation.

Complementing this, other top candidates are directly involved in somatic growth. The gene *zdhhc13*, for example, is a palmitoyltransferase whose disruption in mice leads to severe postnatal growth retardation (Song et al. 2014). Furthermore, ENSACLG00000021917 (*myh7*), a slow-muscle myosin heavy chain, is essential for muscle development, and its expression level is correlated with growth rate in other teleost fish, including closely related tilapia cichlids (Liu et al. 2020). Additionally, *csrp3* is a key regulator of myogenesis and muscle fibre integrity supporting the link between growth rate and muscle development. Finally, *gnb1l* has been associated with increased body weight and feed efficiency in chickens (Ren et al. 2020), suggesting a conserved role in regulating somatic growth and metabolic performance across vertebrates. Selection on these genes could therefore directly underpin the reduction in adult body size.

Beyond the classic life-history shifts in growth and maturation, intense harvesting is expected to select for other traits that reduce capture probability, such as changes in fecundity or the evolution of avoidance behaviours (Heino et al. 2015). While we did not measure these traits directly, our genomic results provide intriguing clues that such selection may be occurring. The strong enrichment of candidate genes involved in high-level nervous system development—including terms such as neurogenesis and commissural neuron axon guidance—suggests that selection is actively shaping the neural architecture of the Lake Malombe population. Such changes could plausibly underpin the evolution of adaptive behaviours like increased wariness or altered shoaling. This behavioural component may also be linked to our top candidate genes; for instance, selection on the slow-muscle myosin gene ENSACLG00000021917 (*myh7*) could not only influence growth but also modulate the power and endurance of the flight response, a critical component of net avoidance. Together, this suite of candidate genes connects the genomic signal to a plausible biological response and provides promising targets for functional follow ups to establish the mechanistic links between the intense fishing pressure in Lake Malombe and the observed evolution of these key physiological traits.

The potential for such traits to evolve in this system is high. For example, recent work has documented significant variation in physiological and behavioural traits, including swimming performance and exploration, both within Lake Malombe *C. mloto* and compared to closely related species (Sawasawa et al. 2025). This suggests that heritable variation is present in these traits, making an evolutionary response to selection within *C. mloto* highly plausible. Future studies, leveraging the experimental tractability of the system, could directly test for divergent behavioural syndromes and other traits like fecundity between the

heavily- and weakly-fished populations and experimentally map their genetic basis.

In conclusion, our study revealed a widespread, polygenic signal of recent selection in an exceptionally heavily fished population of *C. mloto*. The concentration of this genomic footprint on the developmental pathways that govern growth and maturation makes a strong argument that this represents an adaptive response to the intense harvesting pressure. The suite of candidate genes we identified provides promising targets for functional studies, while the biology of this system presents a unique opportunity to experimentally dissect the genomic basis of affected traits and their evolutionary potential, a task intractable for most marine fisheries targets.

From a management perspective, our findings present a complex picture. The predominance of soft, incomplete sweeps on standing variation suggests that much of the original genetic diversity has been maintained. However, the highly polygenic nature of the response means that reversing these widespread genetic shifts may be slow and difficult, potentially creating a 'Darwinian debt' that could compromise the stock's long-term resilience (Pandolfi 2009). Therefore, monitoring the evolutionary trajectory of the population under future management interventions will be crucial for understanding and mitigating the long-term consequences of human-induced evolution.

Author Contributions

A.H.H. and H.S. conceived the study. H.S. and B.R. collected samples. A.H.H. performed all analyses with contributions from F.C.J., J.C.G., and H.S. G.B. and H.S. provided supervision. A.H.H. and H.S. wrote the manuscript. All Authors approved the final manuscript.

Funding

This work was supported by Fonds Wetenschappelijk Onderzoek, 11E0623N, G047521N. Cambridge-Africa ALBORADA Research Fund.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Raw sequencing data are publicly available in the NCBI Sequence Read Archive (SRA); accession numbers are listed in Table S1. All scripts and analysis pipelines used in this study are archived on Zenodo (<https://doi.org/10.5281/zenodo.17649119>) to ensure reproducibility. Variant call format (VCF) files are available on Zenodo: (<https://doi.org/10.5281/zenodo.17657791>). Benefit Sharing: A research collaboration was developed with scientists from Malawi through collaboration with the University of Malawi, and the Department of Fisheries. This research was conducted in full compliance with the Nagoya Protocol on Access and Benefit Sharing. Research permits and material transfer agreements were obtained from the Government of Malawi. The results of research have been shared with the local communities and the broader scientific community and the research addresses a priority concern, in this case the conservation of studied organisms. Benefits arising from this research have been shared through collaboration with Malawian research institutions, training and capacity building and co-authorship with in-country scientists. More broadly, our group is committed to international scientific partnerships, as well as institutional capacity building.

References

- 1000 Genomes Project Consortium. 2015. "A global reference for human genetic variation." *Nature* 526, no. 7571: 68–74. <https://doi.org/10.1038/nature15393>.
- Alexa, A., and J. Rahnenführer. 2023. "TopGO: Enrichment Analysis for Gene Ontology. (R Package Version 2.56.0)." <https://doi.org/10.18129/B9.bioc.topGO>.
- Alexa, A., J. Rahnenführer, and T. Lengauer. 2006. "Improved Scoring of Functional Groups From Gene Expression Data by Decorrelating GO Graph Structure." *Bioinformatics* 22, no. 13: 1600–1607. <https://doi.org/10.1093/bioinformatics/btl140>.
- Allendorf, F. W., and J. J. Hard. 2009. "Human-Induced Evolution Caused by Unnatural Selection Through Harvest of Wild Animals." *Proceedings of the National Academy of Sciences of the United States of America* 106, no. Supp. 1: 9987–9994. <https://doi.org/10.1073/pnas.0901069106>.
- Ashburner, M., C. A. Ball, J. A. Blake, et al. 2000. "Gene Ontology: Tool for the Unification of Biology." *Nature Genetics* 25, no. 1: 25–29. <https://doi.org/10.1038/75556>.
- Baltazar-Soares, M., K. I. Brans, and C. Eizaguirre. 2021. "Human-Induced Evolution: Signatures, Processes and Mechanisms Underneath Anthropogenic Footprints on Natural Systems." *Evolutionary Applications* 14, no. 10: 2335–2341. <https://doi.org/10.1111/eva.13305>.
- Baumdicker, F., G. Bisschop, D. Goldstein, et al. 2022. "Efficient Ancestry and Mutation Simulation With Msprime 1.0." *Genetics* 220, no. 3: iyab229. <https://doi.org/10.1093/genetics/iyab229>.
- Blumer, L. M., V. Burskaia, I. Artiushin, et al. 2025. "Introgression Dynamics of Sex-Linked Chromosomal Inversions Shape the Malawi Cichlid Radiation." *Science* 388, no. 6752: eadr9961. <https://doi.org/10.1126/science.adr9961>.
- Campbell-Staton, S. C., B. J. Arnold, D. Gonçalves, et al. 2021. "Ivory Poaching and the Rapid Evolution of Tusklessness in African Elephants." *Science* 374, no. 6566: 483–487. <https://doi.org/10.1126/science.abe7389>.
- Choin, J., J. Mendoza-Revilla, L. R. Arauna, et al. 2021. "Genomic Insights Into Population History and Biological Adaptation in Oceania." *Nature* 592, no. 7855: 583–589. <https://doi.org/10.1038/s41586-021-03236-5>.
- Cingolani, P., A. Platts, L. L. Wang, 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, no. 2: 80–92. <https://doi.org/10.4161/fly.19695>.
- Conover, D. O., and S. B. Munch. 2002. "Sustaining Fisheries Yields Over Evolutionary Time Scales." *Science* 297, no. 5578: 94–96. <https://doi.org/10.1126/science.1074085>.
- Danecek, P., A. Auton, G. Abecasis, et al. 2011. "The Variant Call Format and VCFtools." *Bioinformatics* 27, no. 15: 2156–2158. <https://doi.org/10.1093/bioinformatics/btr330>.
- Danecek, P., J. K. Bonfield, J. Liddle, et al. 2021. "Twelve Years of SAMtools and BCFtools." *GigaScience* 10, no. 2: giab008. <https://doi.org/10.1093/gigascience/giab008>.
- Di Yorio, M. P., D. I. Pérez Sirkin, T. H. Delgadín, et al. 2016. "Gonadotrophin-Inhibitory Hormone in the Cichlid Fish *Cichlasoma dimerus*: Structure, Brain Distribution and Differential Effects on the Secretion of Gonadotrophins and Growth Hormone." *Journal of Neuroendocrinology* 28, no. 5: 390–401. <https://doi.org/10.1111/jne.12377>.
- Dufour, S., B. Quérat, H. Tostivint, C. Pasqualini, H. Vaudry, and K. Rousseau. 2020. "Origin and Evolution of the Neuroendocrine Control of Reproduction in Vertebrates, With Special Focus on Genome and Gene Duplications." *Physiological Reviews* 100: 869–943. <https://doi.org/10.1152/physrev.00009.2019>.
- Dyer, S. C., O. Austine-Orimoloye, A. G. Azov, et al. 2025. "Ensembl 2025." *Nucleic Acids Research* 53, no. D1: D948–D957. <https://doi.org/10.1093/nar/gkae1071>.
- Gene Ontology Consortium. 2026. *Gene Ontology Data Archive (2026-01-23)*. Zenodo.
- Gu, Z., and D. Hübschmann. 2023. "simplifyEnrichment: A Bioconductor Package for Clustering and Visualizing Functional Enrichment Results." *Genomics, Proteomics & Bioinformatics* 21, no. 1: 190–202. <https://doi.org/10.1016/j.gpb.2022.04.008>.
- Han, K. Y., R. S. Brennan, C. T. Monk, et al. 2025. "Genomic Evidence for Fisheries-Induced Evolution in Eastern Baltic Cod." *Science Advances* 11, no. 26: eadr9889. <https://doi.org/10.1126/sciadv.adr9889>.
- Hara, M. M., S. J. Donda, and F. Njaya. 1999. *An Evaluation of the Lake Malombe Co-Management Program*, 9. International Conference on Fisheries Co-Management.
- Heino, M., B. Díaz Pauli, and U. Dieckmann. 2015. "Fisheries-Induced Evolution." *Annual Review of Ecology, Evolution, and Systematics* 46, no. 1: 461–480. <https://doi.org/10.1146/annurev-ecolsys-112414-054339>.
- Hendry, A. P., K. M. Gotanda, and E. I. Svensson. 2017. "Human Influences on Evolution, and the Ecological and Societal Consequences." *Philosophical Transactions of the Royal Society, B: Biological Sciences* 372, no. 1712: 20160028. <https://doi.org/10.1098/rstb.2016.0028>.
- Hof, A. E., P. Campagne, D. J. Rigden, et al. 2016. "The Industrial Melanism Mutation in British Peppered Moths Is a Transposable Element." *Nature* 534, no. 7605: 102–105. <https://doi.org/10.1038/nature17951>.
- Hutchings, J. A., and A. Kuparinen. 2021. "Throwing Down a Genomic Gauntlet on Fisheries-Induced Evolution." *Proceedings of the National Academy of Sciences of the United States of America* 118, no. 20: 30–32. <https://doi.org/10.1073/pnas.2105319118>.
- Iles, T. D. 1971. "Ecological Aspects of Growth in African Cichlid Fishes." *Journal du Conseil* 33, no. 3: 363–385. <https://doi.org/10.1093/icesjms/33.3.363>.
- Konings, A. 2016. *Malawi Cichlids in Their Natural Habitat*. Cichlid Press.
- Kuparinen, A., and M. Festa-Bianchet. 2017. "Harvest-Induced Evolution: Insights From Aquatic and Terrestrial Systems." *Philosophical Transactions of the Royal Society, B: Biological Sciences* 372, no. 1712: 20160036. <https://doi.org/10.1098/rstb.2016.0036>.
- Laisk, T., V. Kukuškina, D. Palmer, et al. 2018. "Large-Scale Meta-Analysis Highlights the Hypothalamic–Pituitary–Gonadal Axis in the Genetic Regulation of Menstrual Cycle Length." *Human Molecular Genetics* 27, no. 24: 4323–4332. <https://doi.org/10.1093/hmg/ddy317>.
- Li, H. 2013. "Aligning Sequence Reads, Clone Sequences and Assembly Contigs With BWA-MEM." <https://arxiv.org/abs/1303.3997>.
- Liao, Y., G. K. Smyth, and W. Shi. 2014. "FeatureCounts: An Efficient General Purpose Program for Assigning Sequence Reads to Genomic Features." *Bioinformatics* 30, no. 7: 923–930. <https://doi.org/10.1093/bioinformatics/btt656>.
- Liu, X., S. Zeng, S. Liu, et al. 2020. "Identifying the Related Genes of Muscle Growth and Exploring the Functions by Compensatory Growth in Mandarin Fish (*Siniperca chuatsi*)." *Frontiers in Physiology* 11: 553563. <https://doi.org/10.3389/fphys.2020.553563>.
- Lund, C., V. Yellapragada, S. Vuoristo, et al. 2020. "Characterization of the Human GnRH Neuron Developmental Transcriptome Using a GNRH1-TdTomato Reporter Line in Human Pluripotent Stem Cells." *Disease Models & Mechanisms* 13, no. 3: dmm040105. <https://doi.org/10.1242/dmm.040105>.

- Makwinja, R., S. Mengistou, E. Kaunda, and T. Alamirew. 2021. "Lake Malombe Fish Stock Fluctuation: Ecosystem and Fisherfolks." *Egyptian Journal of Aquatic Research* 47, no. 3: 321–327. <https://doi.org/10.1016/j.ejar.2021.07.001>.
- Malinsky, M., H. Svardal, A. M. Tyers, et al. 2018. "Whole-Genome Sequences of Malawi Cichlids Reveal Multiple Radiations Interconnected by Gene Flow." *Nature Ecology & Evolution* 2, no. 12: 1940–1955. <https://doi.org/10.1038/s41559-018-0717-x>.
- M'balaka, M. S., E. Kaunda, G. Z. Kanyerere, D. Jamu, and A. Msukwa. 2024. "Changes in Fish Species Diversity, Size Structure and Distribution in the Trawlable Demersal Zones of Lake Malawi, Malawi." *Environment, Development and Sustainability* 27: 1–21. <https://doi.org/10.1007/s10668-024-04649-8>.
- Medina, P., A. Gómez, S. Zanuy, and M. Blázquez. 2019. "Involvement of the Retinoic Acid Signaling Pathway in Sex Differentiation and Pubertal Development in the European Sea Bass *Dicentrarchus labrax*." *Heliyon* 5, no. 2: e01201. <https://doi.org/10.1016/j.heliyon.2019.e01201>.
- Michod, R. E. 1979. "Evolution of Life Histories in Response to Age-Specific Mortality Factors." *American Naturalist* 113, no. 4: 531–550.
- Muzellec, B., M. Teleńczuk, V. Cabeli, and M. Andreux. 2023. "PyDESeq2: A Python Package for Bulk RNA-Seq Differential Expression Analysis." *Bioinformatics* 39, no. 9: btad547. <https://doi.org/10.1093/bioinformatics/btad547>.
- Pandolfi, J. M. 2009. "Evolutionary Impacts of Fishing: Overfishing's 'Darwinian Debt'." *F1000 Biology Reports* 1: 1–43. <https://doi.org/10.3410/b1-43>.
- Pelletier, F., and D. W. Coltman. 2018. "Will Human Influences on Evolutionary Dynamics in the Wild Pervade the Anthropocene?" *BMC Biology* 16, no. 1: 7. <https://doi.org/10.1186/s12915-017-0476-1>.
- Pigeon, G., M. Festa-Bianchet, D. W. Coltman, and F. Pelletier. 2016. "Intense Selective Hunting Leads to Artificial Evolution in Horn Size." *Evolutionary Applications* 9, no. 4: 521–530. <https://doi.org/10.1111/eva.12358>.
- Pinsky, M. L., A. M. Eikeset, C. Helmerson, et al. 2021. "Genomic Stability Through Time Despite Decades of Exploitation in Cod on Both Sides of the Atlantic." *Proceedings of the National Academy of Sciences of the United States of America* 118, no. 15: e2025453118. <https://doi.org/10.1073/pnas.2025453118>.
- Pradhan, A., and P.-E. Olsson. 2015. "Inhibition of Retinoic Acid Synthesis Disrupts Spermatogenesis and Fecundity in Zebrafish." *General and Comparative Endocrinology* 217–218: 81–91. <https://doi.org/10.1016/j.ygcen.2015.02.002>.
- Ren, T., Y. Yang, W. Lin, et al. 2020. "A 31-Bp Indel in the 5' UTR Region of GNB1L Is Significantly Associated With Chicken Body Weight and Carcass Traits." *BMC Genetics* 21, no. 1: 91. <https://doi.org/10.1186/s12863-020-00900-z>.
- Sawasawa, W., A. Hooft van Huysduynen, S. Gresham, et al. 2024. "Population Genomic Analysis Reveals Cryptic Population Structure in the Commercially Important Lake Malawi Cichlid *Copadichromis mloto* (Iles, 1960)." *Hydrobiologia* 852: 1–17. <https://doi.org/10.1007/s10750-024-05743-8>.
- Sawasawa, W., S. Kibor, B. Shi, et al. 2025. *Physiological and Behavioural Responses in Closely Related Lake Malawi Cichlids: A Comparative Study of Copadichromis Mloto and Copadichromis chrysonotus*. SSRN. <https://doi.org/10.2139/ssrn.5286638>.
- Schlamp, F., J. van der Made, R. Stambler, L. Chesebrough, A. R. Boyko, and P. W. Messer. 2016. "Evaluating the Performance of Selection Scans to Detect Selective Sweeps in Domestic Dogs." *Molecular Ecology* 25, no. 1: 342–356. <https://doi.org/10.1111/mec.13485>.
- Son, Y. L., T. Ubuka, and K. Tsutsui. 2019. "Molecular Mechanisms of Gonadotropin-Inhibitory Hormone (GnIH) Actions in Target Cells and Regulation of GnIH Expression." *Frontiers in Endocrinology* 10: 110. <https://doi.org/10.3389/fendo.2019.00110>.
- Song, I.-W., W.-R. Li, L.-Y. Chen, et al. 2014. "Palmitoyl Acyltransferase, Zdhhc13, Facilitates Bone Mass Acquisition by Regulating Postnatal Epiphyseal Development and Endochondral Ossification: A Mouse Model." *PLoS One* 9, no. 3: e92194. <https://doi.org/10.1371/journal.pone.0092194>.
- Talbi, M., G. F. Turner, and M. Malinsky. 2024. "Rapid Evolution of Recombination Landscapes During the Divergence of Cichlid Ecotypes in Lake Masoko." *Evolution; International Journal of Organic Evolution* 79: 364–379. <https://doi.org/10.1093/evolut/qpae169>.
- Therkildsen, N. O., A. P. Wilder, D. O. Conover, S. B. Munch, H. Baumann, and S. R. Palumbi. 2019. "Contrasting Genomic Shifts Underlie Parallel Phenotypic Evolution in Response to Fishing." *Science* 365: 487–490. <https://doi.org/10.1126/science.aaw7271>.
- Turner, G. F. 1995. "Management, Conservation and Species Changes of Exploited Fish Stocks in Lake Malawi." In *The Impact of Species Changes in African Lakes*, edited by T. J. Pitcher and P. J. B. Hart, 365–395. Springer Netherlands. https://doi.org/10.1007/978-94-011-0563-7_17.
- Tweddle, D., G. F. Turner, and M. B. D. Seisay. 1995. "Changes in Species Composition and Abundance as a Consequence of Fishing in Lake Malombe, Malawi." In *The Impact of Species Changes in African Lakes*, edited by T. J. Pitcher and P. J. B. Hart, 413–424. Springer Netherlands. https://doi.org/10.1007/978-94-011-0563-7_19.
- Weir, B. S., and C. C. Cockerham. 1984. "Estimating F-Statistics for the Analysis of Population Structure." *Evolution* 6, no. 36: 1358–1370. <https://doi.org/10.2307/2408641>.
- Weyl, O. L. F., K. R. Mwakiyongo, and D. S. Mandere. 2004. "An Assessment of the Nkacha Net Fishery of Lake Malombe, Malawi." *African Journal of Aquatic Science* 29, no. 1: 47–55. <https://doi.org/10.2989/16085910409503791>.
- Yao, K., and W. Ge. 2013. "Spatial Distribution and Receptor Specificity of Zebrafish Kit System-Evidence for a Kit-Mediated Bi-Directional Communication System in the Preovulatory Ovarian Follicle." *PLoS One* 8, no. 2: e56192. <https://doi.org/10.1371/journal.pone.0056192>.
- Yao, K., and W. Ge. 2015. "Differential Regulation of Kit Ligand A (Kitlga) Expression in the Zebrafish Ovarian Follicle Cells-Evidence for the Existence of a Cyclic Adenosine 3', 5' Monophosphate-Mediated Binary Regulatory System During Folliculogenesis." *Molecular and Cellular Endocrinology* 402: 21–31. <https://doi.org/10.1016/j.mce.2014.12.005>.
- Yi, X., Y. Liang, E. Huerta-Sanchez, et al. 2010. "Sequencing of 50 Human Exomes Reveals Adaptation to High Altitude." *Science* 329, no. 5987: 75–78. <https://doi.org/10.1126/science.1190371>.
- Zeng, K., Y.-X. Fu, S. Shi, and C.-I. Wu. 2006. "Statistical Tests for Detecting Positive Selection by Utilizing High-Frequency Variants." *Genetics* 174, no. 3: 1431–1439. <https://doi.org/10.1534/genetics.106.061432>.
- Zhang, X., and I. Jonassen. 2020. "RASflow: An RNA-Seq Analysis Workflow With Snakemake." *BMC Bioinformatics* 21, no. 1: 110. <https://doi.org/10.1186/s12859-020-3433-x>.
- Zhang, Y., X. Nie, L. Yang, et al. 2025. "The Role of KAT6A in Regulating Primordial Follicle Activation in Mouse Ovary." *American Journal of Physiology-Cell Physiology* 329, no. 1: C38–C49. <https://doi.org/10.1152/ajpcell.00055.2025>.
- Zwieten, P. A. M., M. Banda, and J. Kolding. 2011. "Selecting Indicators to Assess the Fisheries of Lake Malawi and Lake Malombe: Knowledge Base and Evaluative Capacity." *Journal of Great Lakes Research* 37: 26–44. <https://doi.org/10.1016/j.jglr.2010.11.001>.
- Zwieten, P. A. M. v., F. Njaya, and O. Weyl. 2003. "Effort Development and the Collapse of the Fisheries of Lake Malombe: Does Environmental Variability Matter?" *FAO Fisheries Technical Paper* 2: 132–164.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Figure S1:** Quantile–quantile plot of empirical Fisher's Combined Score (FCS) values against FCS values from neutral demographic simulations. Deviation above the 1:1 line (solid red) indicates an excess of high-scoring loci relative to neutral expectation. The dashed red line marks the empirically derived selection threshold (FCS = 5.51), above which SNPs are classified as selection outliers. **Figure S2:** Detailed plot of Fisher's combined score (yellow), Population Branch statistic (blue) and *H*-score (green) and gene positions of a major FCS peak on chr2: 11606365–11,656,365. **Figure S3:** Detailed plot of Fisher's combined score (yellow), Population Branch statistic (blue) and *H*-score (green) and gene positions of a major FCS peak on chr3: 3326518–3,376,518. **Figure S4:** Detailed plot of Fisher's combined score (yellow), Population Branch statistic (blue) and *H*-score (green) and gene positions of a major FCS peak on chr6: 6472995–6,582,963. **Figure S5:** Detailed plot of Fisher's combined score (yellow), Population Branch statistic (blue) and *H*-score (green) and gene positions of a major FCS peak on chr7:10285777–10,338,291. **Figure S6:** Detailed plot of Fisher's combined score (yellow), Population Branch statistic (blue) and *H*-score (green) and gene positions of a major FCS peak on chr7:44175985–44,244,688. **Figure S7:** Detailed plot of Fisher's combined score (yellow), Population Branch statistic (blue) and *H*-score (green) and gene positions of a major FCS peak on chr12:5505879–5,557,210. **Figure S8:** Detailed plot of Fisher's combined score (yellow), Population Branch statistic (blue) and *H*-score (green) and gene positions of a major FCS peak on chr12:6036511–6,086,511. **Figure S9:** Detailed plot of Fisher's combined score (yellow), Population Branch statistic (blue) and *H*-score (green) and gene positions of a major FCS peak on chr20:8335345–8,385,356. **Figure S10:** Detailed plot of Fisher's combined score (yellow), Population Branch statistic (blue) and *H*-score (green) and gene positions of a major FCS peak on chr23:20791821–20843857. **Table S1:** Metadata of sequenced specimens of *Copadichromis mloto*, used in this study. Including sampling location, sequencing depth, assigned population, *C. mloto* clade & NCBI SRA accession numbers. Individuals also sampled for differential gene expression analysis (RNA-seq) are indicated. **Table S2:** Metadata of sequences from Blumer et al. 2025 used in this study. Including Population assignment, Malawi cichlid clade & NCBI SRA accession numbers. **Table S3:** Output of Gene Ontology test (topGO) for all candidate genes within ± 25 kb of outlier SNPs and (ii) the subset of those candidate genes that are differentially expressed between Lake Malombe and Lake Malawi. **Table S4:** Differential expression results log₂ fold change (log₂FC) and $-\log_{10}$ (adjusted *p*-value) in five tissues (Liver, Brain, Muscle, Gonads, Gills). **Table S5:** Summary table of candidate genes, genes with non-synonymous and/or intronic variants above defined empirical selection threshold, and differentially expressed genes in five tissues (Liver, Brain, Muscle, Gonads, Gills).