Aquatic Conservation: Marine and Freshwater Ecosystems

ARTICLE OPEN ACCESS

A Universal Genomic Approach to Detect Hybridization Using MIPs: Insights From a Case Study on Pike

Annalisa Scapolatiello[1] | Chiara Samassa[1,2] | Andrea Gandolfi[2,3] | Maria Chiara Valerin[1] | Lorenzo Zane[1,2,4] | Elisa Boscari[1] | Leonardo Congiu[1,2,4]

[1]Department of Biology, University of Padova, Padova, Italy | [2]National Biodiversity Future Center, Palermo, Italy | [3]Conservation Genomics Research Unit, Research and Innovation Centre, Fondazione Edmund Mach (FEM), San Michele all'Adige, Trentino, Italy | [4]Consorzio Nazionale Interuniversitario Per le Scienze del Mare (CoNISMa), Rome, Italy

Correspondence: Elisa Boscari (elisa.boscari@unipd.it)

Received: 1 July 2025 | **Revised:** 15 December 2025 | **Accepted:** 8 January 2026

Keywords: diagnostic markers | hybridization | introns | species identification

ABSTRACT

Interspecific hybridization poses a major challenge to biodiversity conservation, as genetic introgression can threaten the persistence of native species. This phenomenon occurs when individuals from distinct species interbreed, producing hybrids with a combination of genetic traits from both parents. The implications are especially concerning when these hybrids are fertile and able to backcross with the parental species, potentially leading to the permanent integration of foreign genetic material into native populations. Human-mediated translocations further increase hybridization risk by bringing reproductively compatible species into contact. Detecting introgression in individuals that cannot be reliably distinguished morphologically is therefore crucial for conservation planning. In this study, we investigate hybridization between the southern pike (*Esox flaviae*) and northern pike (*Esox lucius*) in Italy and Central Europe. We evaluate multilocus intron polymorphisms (MIPs) as a molecular tool for species identification and hybrid detection by analysing pike samples using a combination of previously available MIP loci and newly isolated pike-specific MIPs. Our results show that MIP loci successfully reproduce patterns previously identified with microsatellites, confirming their suitability for detecting hybridization and interspecific genetic structure. We also developed a panel of diagnostic loci enabling rapid species/hybrids identification for routine applications. This is the first study demonstrating the effectiveness of MIPs for accurate species assignment and assessment of genetic diversity in pike, particularly in the context of hybridization and its conservation implications. Overall, our findings highlight the value of MIPs as complementary molecular markers for biodiversity studies, providing practical diagnostic tools for species monitoring and management in conservation programmes.

1 | Introduction

Anthropogenic-mediated interspecific hybridization is a significant concern for biodiversity conservation, especially given its potential consequences for native species. When individuals from distinct species interbreed, hybrids are produced with a mosaic of genetic traits inherited from both parental species. The issue becomes particularly alarming if these hybrids are fertile and capable of backcrossing with parental species, leading to the permanent introgression of exotic genetic traits. This process may disrupt unique coevolving genetic loci within native populations, ultimately jeopardizing their survival and destabilizing ecological balance (Abbott et al. 2013; Allendorf et al. 2001; Meilink et al. 2015). A major driver of interspecific

Annalisa Scapolatiello and Chiara Samassa contributed equally to this work.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *Aquatic Conservation: Marine and Freshwater Ecosystems* published by John Wiley & Sons Ltd.

hybridization is the intentional or unintentional transfer of species beyond their natural distribution ranges (Meraner et al. 2013 and 2014). Such translocations allow species to come into contact with reproductively compatible native species, creating conditions conducive to hybridization. This issue is notably pronounced in freshwater fish populations, where numerous species have been repeatedly stocked or translocated to new environments for recreational or aquaculture purposes (Rhymer and Simberloff 1996; Allendorf et al. 2001; Mandeville et al. 2019; Casanova et al. 2022; Bekkevold et al. 2024; Tarkan et al. 2024). Such human-induced reshuffling of species substantially increases the risk of hybridization and introgression, potentially altering the genetic diversity and ecological dynamics of native populations (Tarkan et al. 2024). Moreover, freshwater species might be particularly prone to hybridization processes as they often originate in allopatry and do not develop reproductive isolation mechanisms (Bernardi 2013; Tarkan et al. 2024).

Within freshwater ecosystems, the detection of hybridization events between different pike species (*Esox* spp.) has emerged as a pressing scientific challenge. The genus *Esox* comprises seven primary freshwater species, each with distinct distribution patterns across the Holarctic region and notable for their ecological and evolutionary significance. While *Esox americanus* (Gmelin, 1789), *Esox masquinongy* (Mitchill, 1824) and *Esox niger* (Lesuer, 1818) are confined to North America, *Esox reichertii* (Dybowski, 1869) is limited to the Amur River Basin in Northeast Asia (Kottelat and Freyhof 2007; Skog et al. 2014). In stark contrast, *Esox lucius* (Linnaeus, 1758), the top predator known as the northern pike, boasts a circumpolar distribution and was historically considered the only esocid in Europe (Gandolfi et al. 2017). With substantial commercial and recreational value, *E. lucius* is also recognized as a model organism in ecology and evolutionary biology (Koch-Schmidt et al. 2015). Recent morphometric/meristic and genetic studies, however, have contested its sole presence in Europe, revealing distinct local forms now recognized as separate species (Bianco and Delmastro 2011; Lucentini et al. 2011; Denys et al. 2014). *Esox cisalpinus* (Bianco and Delmastro 2011) and *Esox flaviae* (Lucentini et al. 2011) were described from Italian waters and *Esox aquitanicus* in Western France (Denys et al. 2014, 2018; Jeanroy and Denys 2019). *Esox flaviae* was considered a junior synonym of *E. cisalpinus* (Bianco 2014), but some authors have disputed this conclusion and consider the former to remain valid (e.g., Gandolfi et al. 2015, 2017). Notably, the taxonomic status of *E. cisalpinus* and *E. flaviae* (henceforth referred to as *E. flaviae*), the Italian pike and the southern pike, is still debated, with arguments for synonymy lacking thorough revision (Gandolfi et al. 2015). Genetic, meristic and morphological investigations revealed distinct genetic and morphological traits in southern pike populations compared to *E. lucius* (Lucentini et al. 2011). These findings suggest a monophyletic origin for the southern pike and pronounced genetic differentiation from the northern pike, which appears to be non-native in Italy. The invasion of *E. lucius* in Italian waters, attributed to human translocation and stocking campaigns, has led to hybridization and genetic introgression (Gandolfi et al. 2017). The potential consequences of this hybridization extend beyond mere taxonomic interest, as they can impact the genetic integrity of wild populations, influencing adaptive potential and ecological interactions as hypothesized by the evident decline of Southern Italian populations (Lucentini et al. 2011). Given that traditional morphological characteristics often prove insufficient in distinguishing hybrids from pure species, the integration of molecular tools has become indispensable for accurate and reliable identification. Native and exotic pike species or hybrid identification has been accomplished using a panel of 16 microsatellite markers and comparing the resulting genetic profile to reference samples from putatively pure *E. lucius* or *E. flaviae* (Gandolfi et al. 2017).

To streamline this process and in order to promote a multilocus genetic support to management and conservation practices, we propose the application of the multilocus intron polymorphism (MIP) approach, an intron-targeted amplicon sequencing recently developed by Boscari et al. (2024) which demonstrated the effectiveness of these nuclear markers in differentiating closely related species and populations, particularly among teleost fishes. Building on this approach, our study presents a new genetic panel of markers, including a subset of previously characterized MIP loci (Boscari et al. 2024) alongside newly isolated, pike-specific MIPs. By examining pike specimens from Italy (*E. flaviae*) and Central Europe (*E. lucius*), we aim to (i) replicate the findings shown by Gandolfi et al. (2017), (ii) validate MIP markers as effective alternatives to other molecular markers and (iii) develop and apply a reduced panel of diagnostic loci for rapid species detection without sequencing to expedite routine identification of pure and introgressed individuals. Moreover, to verify the effectiveness of the reduced panel of 10 MIP loci in detecting hybridization and introgression between *E. lucius* and *E. flaviae*, we implemented simulations generating hybrid genotypes with known ancestry and varying degrees of introgression. This approach ensured the accurate detection of genetic contributions from *E. lucius*, even at low levels.

This simulation framework allowed for a systematic evaluation of marker performance across a gradient of hybridization scenarios, providing robust validation of their effectiveness in identifying both first-generation hybrids (F1) and successive backcrosses.

This case study shows the utility of MIPs as alternative molecular markers for species identification and detection of hybridization and illustrates the possibility to easily assign pike individuals through MIP-based diagnostic markers in routine analysis, highlighting the potential effectiveness of MIPs' application in genetic monitoring for conservation purposes.

2 | Material and Methods

2.1 | Sampling and DNA Purification

The study included a total of 209 pike individuals representing various populations of *E. flaviae*, *E. lucius* and hybrids between the two species (Figure 1), encompassing both wild specimens from different Italian basins and hatchery samples collected in different years and used for restocking initiatives in Italy (Table 1). Of these, 188 individuals were used for MIP library preparation and sequencing. This group comprised 104 specimens that had already been analysed for mitochondrial DNA and microsatellite loci in Gandolfi et al. (2017) (from Lake Garda, Lake Caldonazzo, Lake Trasimeno, River Adda, Monticolo Lake,

10990755, 2026, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aqc.70329 by Fondazione Edmund Mach, Wiley Online Library on [10/02/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

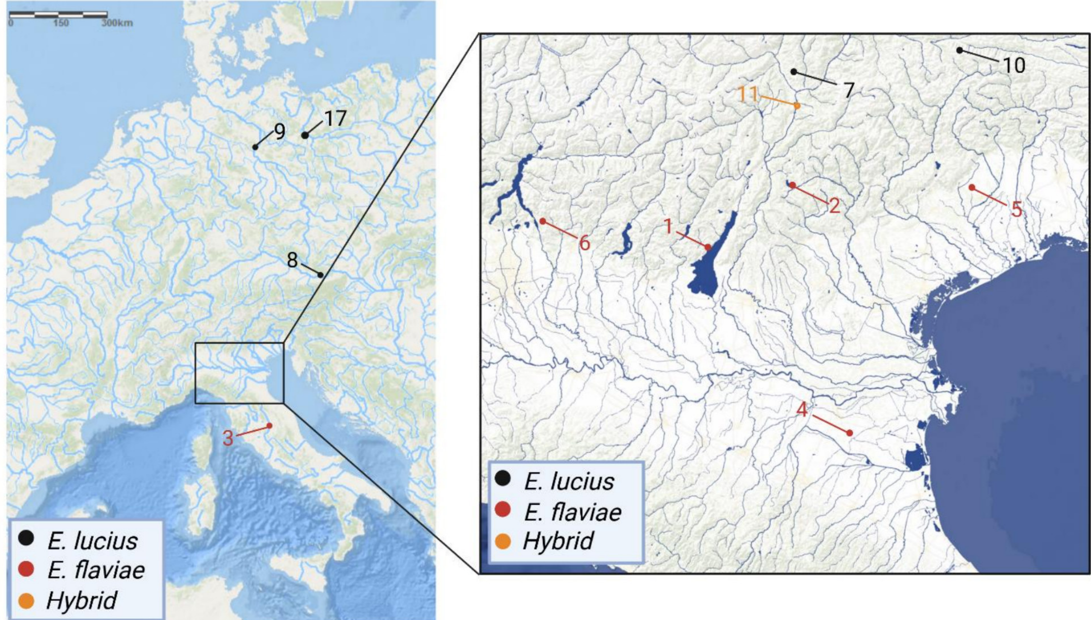

FIGURE 1 | Geographical distribution of the sampling sites included in this study (only wild populations are reported): 1, Garda Lake; 2, Caldonazzo Lake; 3, Trasimeno Lake; 4, Argenta Valleys; 5, Livenza River; 6, Adda River; 7, Monticolo Lake; 8, Danube River; 9, Elbe River; 10, Cadore Lake; 11, Caldaro/Salorno ditches; and 17, Parstein Lake. The populations with corresponding sample details are reported in Table 1.

TABLE 1 | Populations list and samples analysed. The study includes 188 individuals analysed with MIPs (104 analysed with microsatellites from Gandolfi et al. [2017] and 84 that were analysed later with the same markers but remained unpublished), plus 21 additional individuals (highlighted in grey) tested only at diagnostic MIP loci by PCR and agarose gel electrophoresis. For each population, identified by an ID number in the first column, the following information is reported: species, sampling site, origin (wild, W; hatchery, H), number of individuals included in the MIP library preparation (N), number of successfully sequenced/analysed individuals (Ns), number of individuals validated by PCR with the diagnostic loci and the reference for sample provenance.

Pop ID	Species	Sampling site	W/H	N	Ns	Nv	Reference
1	*E. flaviae*	Garda Lake	W	13	12	—	Gandolfi et al. 2017
2	*E. flaviae*	Caldonazzo Lake	W	15	15	—	Gandolfi et al. 2017
3	*E. flaviae*	Trasimeno Lake	W	4	4	—	Gandolfi et al. 2017
4	*E. flaviae*	Argenta Valleys	W	14	14	—	Gandolfi A. (unpublished)
5	*E. flaviae*	Livenza River	W	15	14	—	Gandolfi A. (unpublished)
6	*E. flaviae*	Adda River	W	14	14	—	Gandolfi et al. 2017
7	*E. lucius*	Monticolo Lake	W	14	13	—	Gandolfi et al. 2017
8	*E. lucius*	Danube River	W	16	16	—	Gandolfi et al. 2017
9	*E. lucius*	Elbe River	W	15	15	—	Gandolfi et al. 2017
10	*E. lucius*	Cadore Lake	W	14	14	—	Gandolfi A. (unpublished)
11	Hybrid	Trentino Alto Adige (Caldaro/Salorno)	W	11	11	—	Gandolfi et al. 2017
12	Hybrid	Hatchery2011	H	2	1	—	Gandolfi et al. 2017
13	*E. flaviae*	Hatchery2018	H	14	14	—	Gandolfi A. (unpublished)
14	*E. flaviae*	Hatchery2020	H	13	12	—	Gandolfi A. (unpublished)
15	*E. flaviae*	Hatchery2021	H	14	14	—	Gandolfi A. (unpublished)
16	*E. flaviae*	Hatchery2023 Trasimeno	H	—		15	Tentoni E. (newly collected)
17	*E. lucius*	Parstein Lake (Brandeburg)	W	—		6	Ludwig A. (newly collected)
		Total		**188**	**183**	**21**	

10990755, 2026, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aqc.70329 by Fondazione Edmund Mach, Wiley Online Library on [10/02/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

River Danube, River Elbe, Caldaro/Salorno ditches and the 2011 hatchery stock). The remaining 84 samples (from River Argenta, River Livenza, Lake Cadore and hatchery stocks from 2018, 2020 and 2021) were not included in that publication but had been analysed later with the same mtDNA and microsatellite markers during subsequent monitoring programmes. Species determination was therefore based on these genetic results, and within the 188 individuals selected for MIP library preparation, putatively pure reference samples from Lake Trasimeno (Italy) for *E. flaviae* and from the Elbe (Germany) and Danube (Austria) basins for *E. lucius* were included.

In addition, 21 individuals (15 putative *E. flaviae* and six putative *E. lucius*, highlighted in grey in Table 1) with no prior genetic information were tested only by PCR at the 10 diagnostic MIP loci here identified to validate the species identification procedure. All samples were derived from existing fin-clip collections of other research institutions, and no live animals were used. The study was carried out in compliance with the ARRIVE guidelines.

All genomic DNAs were purified using the Euroclone spinNAker Universal Genomic DNA mini kit, following the manufacturer's protocol.

2.2 | MIP Markers' Selection

For the marker selection, we started using the panel of 121 MIP loci developed by Boscari et al. (2024). These loci were tested through PCR and agarose gel electrophoresis in a subset of individuals from various populations to confirm amplification reliability. Forty-seven MIP loci with optimal amplification efficiency and amplicon size between 200 and 600 bp were chosen for further analyses.

Additional putatively informative loci were newly isolated from the panel of 2441 intronic loci extracted through the multi-alignment of 48 teleost genomes (Boscari et al. 2024). Among these, loci represented in the reference pike genome were prioritized, and exon-flanking regions were directly aligned with the annotated *E. lucius* genome using BLAST (Altschul et al. 1990) to select the ones in the target size range (200–600 bp). This yielded a total of 40 additional suitable loci, for which locus-specific primer pairs were designed in Geneious v8.1.8 (Kearse et al. 2012), adhering to the following criteria: (i) primer length between 19 and 24 bp, (ii) primer melting temperature around 58°C–60°C, (iii) low hairpin or dimer formation (verified with FastPCR; Kalendar et al. 2017) and (iv) proximity to the intron junction. Their testing on a reduced pike sample set led to the exclusion of a single marker, due to amplification failure (locus c134621_scn8aa), thus resulting in a final panel of 86 loci, overall.

2.3 | MIP Library Preparation, Sequencing and Bioinformatics Data Post-Processing

The 188 individuals were processed at 86 loci using the optimized method by Boscari et al. (2024) for multilocus intronic region sequencing on the Illumina platform. This protocol utilized a two-step PCR approach. The first step amplified the 86 loci subdivided in five multiplex reactions (M1–M5) for each individual. The second step incorporated individual-specific barcodes and Illumina-compatible adapters (P5–P7). Four individuals were processed in replicate. Sequencing was conducted at BMR Genomics s.r.l. in Padua on the Illumina MiSeq platform with 300-bp paired-end reads.

Raw data processing utilized the Snakemake pipeline by Scapolatiello et al. (2025) for automation. Retained allele sequences in FASTA format were aligned per locus and visualized with MEGA11 (Tamura et al. 2021). This process allowed for excluding nonspecific alleles and loci with ambiguous genotypes (loci with more than two alleles per individual, chimeras or low-quality alleles). Loci were retained if they met criteria including (i) genotype presence in at least 75% of individuals, (ii) concordant genotypes between replicates, (iii) no more than two called alleles per genotype and (iv) absence of homopolymer stretches or microsatellites in the merging region of the forward-reverse reads.

The final panel of informative loci was organized in GenePop format for the subsequent genetic analysis (Rousset 2008). The software CREATE (Coombs et al. 2008) was employed to convert GenePop files into other formats compatible with population genetic software (e.g., ARLEQUIN and STRUCTURE).

2.4 | Data Analysis

Data analysis primarily focused on verifying the potential of MIPs in tracing the genetic differentiation between the southern and the northern pike and possible signals of hybridization as well as on identifying diagnostic markers for rapid species identification. Additionally, the study also explored the potential of MIPs in studying population differentiation, in comparison to the already available data from microsatellite loci (Gandolfi et al. 2017). All analyses described below were performed independently for the microsatellite and MIP datasets to allow direct comparison of patterns derived from the two marker types.

The percentage of loci with more than one allele is reported as P_N (%). Expected heterozygosity (H_E), observed heterozygosity (H_O), inbreeding coefficient (F_{IS}) and average number of alleles (N_A) were calculated using ARLEQUIN 3.5 (Excoffier and Lischer 2010). Allelic richness (A_R) and private allelic richness (pA_R) were assessed at the species level using HP-Rare 1.1 (Kalinowski 2005), standardized to the smallest sample size. Hardy–Weinberg equilibrium tests with 1,000,000 Markov chain Monte Carlo (MCMC) steps and 100,000 dememorization steps were performed with ARLEQUIN 3.5. The significance level was set at 0.05, and *p*-values were adjusted using Benjamini and Hochberg (1995) correction for multiple testing.

To visualize dataset variability at the species, principal component analysis (PCA; Pearson 1901) was conducted in R (R Core Team 2022) using the adegenet (Jombart 2008) and ggplot2 (Wickham 2016) packages.

To detect differentiated clusters between and within species and to assess the putative presence of hybridization, various

10990755, 2026, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aqc.70329 by Fondazione Edmund Mach, Wiley Online Library on [10/02/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

datasets including different sets of individuals were analysed using STRUCTURE 2.3.4 (Pritchard et al. 2000). STRUCTURE works on multilocus genotype data, and given a priori defined *K* groups, it associates to each individual the proportion of ancestry in any group. Implementing the analysis for different values of *K*, the software suggests the solution which better fits the data. STRUCTURE uses a Bayesian iterative algorithm based on MCMC. It starts by randomly assigning the individuals to the *K* groups, estimating allele frequencies in each group and reassigning individuals based on the frequency estimates.

Clustering analysis was performed, firstly, over the entire dataset (including *E. flaviae* populations, *E. lucius* populations and hybrid individuals) to test the ability of MIPs to discriminate the two species' contributions. The analysis tested *K* values from 1 to 7, utilizing a 100,000 MCMC burn-in and 1,000,000 MCMC steps. The analysis was replicated five times for each *K* tested. Since evidence of hybridization between the two pike species was reported (Gandolfi et al. 2017), an admixture model (assuming that individuals can have partial ancestry from multiple genetic groups) with correlated allele frequencies was implemented. STRUCTURE-HARVESTER (Earl and vonHoldt 2012) was used to estimate the most likely number of *K* through the ΔK method (Evanno et al. 2005). CLUMPAK (Kopelman et al. 2015) generated graphical representations of STRUCTURE results by collapsing replicates for each *K*.

Population clustering analyses were also performed at the intraspecies level to verify the ability of MIPs to detect subtler levels of differentiation in comparison with available microsatellite data (Gandolfi et al. 2017). In this context, the *E. flaviae* dataset was also analysed, including only wild populations because hatchery stocks might unbalance the analysis, reducing the possibility to detect differences between populations likely due to a higher probability of kinship relationships among reared individuals. Clustering settings were the same as for the inference at the species level.

2.5 | Diagnostic Marker Identification and Validation

Based on the allele frequencies calculated for each locus across the entire dataset, potentially diagnostic loci were selected for further analysis. Starting from the initial 86 loci, the search for diagnostic markers focused on identifying loci showing high-frequency alleles in one species and low frequencies in the other. Through sequence analysis of the identified diagnostic loci, regions suitable for designing species-specific primers were identified. Two alternative primers per locus, specific for *E. flaviae* and *E. lucius*, were designed with the 3′-end complementary to the diagnostic mutation and the penultimate nucleotide intentionally designed to be non-complementary to its target nucleotide. This design ensures that, when species-specific primers are used on a non-target species, the two-nucleotide mismatch prevents successful amplification. The species-specific primers for each locus were paired with a third primer designed on the opposite strand in a region common to both species (Supporting Information S1, Table S1). The goal was to develop a set of loci that, through simple amplification and verification of the presence or absence of the amplicon, would enable rapid species identification without the need for sequencing.

The primer design was performed using the software GENEIOUS (v8.1.9—https://www.geneious.com) according to the following criteria: (i) the 3′ end of the diagnostic primers was positioned at the SNP of interest to ensure species specificity; (ii) a deliberate double mismatch was introduced at the penultimate nucleotide of each diagnostic primer when binding occurs in the non-target species; and (iii) the annealing temperatures were calibrated for compatibility within each primer set.

All designed primers were tested and optimized for amplification via PCR. The amplification protocol employed Taq polymerase (QIAGEN) in a 10-μL reaction volume, including 1 μL of genomic DNA. The thermal cycling conditions consisted of an initial denaturation at 95°C for 15 min, followed by 35 cycles of denaturation at 95°C for 30 s, primer-specific annealing and extension at 72°C, with a final elongation step at 60°C for 30 min. Following amplification, PCR products were subjected to gel electrophoresis. Band presence or absence was recorded and translated into genotypes by replacing each observed band with the corresponding nucleotide in an Excel table. The genotypic data for each individual across the loci were then converted into the GenePop format and further into the STRUCTURE format using the software CREATE (Coombs et al. 2008). This allowed for the assignment of individuals to genetic clusters through STRUCTURE analysis, based on their genotype set-up across the diagnostic loci.

To validate the final panel of 10 diagnostic loci detected, PCR products were initially tested on 28 individuals with known genotypes. Subsequently, the panel was tested on an additional 21 individuals of known geographic origin (15 from Trasimeno and six from Parstein Lake) but not previously genetically characterized, allowing for further validation of the diagnostic loci.

2.6 | Simulation of Virtual Hybrids

To evaluate the diagnostic accuracy of the selected sub-panel of MIP loci in detecting genetic admixture, we conducted simulations of hybrid genotypes with different proportions of parental species contribution. These simulations were based on real genotypic data from *E. lucius* and *E. flaviae* to create F1 and subsequent backcrosses for up to three generations. The goal was to replicate realistic hybridization scenarios and test the robustness of the loci in identifying *E. lucius* ancestry at decreasing proportions.

The parental genotypes were derived from the purest populations in our dataset, specifically, the Garda and Trasimeno populations for *E. flaviae* and the Elbe and Danube populations for *E. lucius*. These datasets were used as inputs for computational simulations of genetic recombination and successive progeny generations. A Python-based computational approach (Supporting Information S2) was employed to perform the simulations in a local computational environment. Two tab-delimited datasets containing genetic information of real parental genotypes

10990755, 2026, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aqc.70329 by Fondazione Edmund Mach, Wiley Online Library on [10/02/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

(10_A.txt and 10_B.txt) were processed using the *pandas* library (McKinney 2010). Each dataset represented the genotypes of individual fish. At each locus, custom functions were implemented to randomly select either the first or the second allele, simulating the generation of gametes. Gametes from different parents were combined to generate the virtual progeny genotype. The simulation workflow included two steps. First is the 'F1' in which gametes from *E. lucius* and *E. flaviae* were combined using a custom function to create 100 unique F1 individuals. Random selection without repetition ensured the uniqueness of each genotype, and the parental origins were tracked throughout the process. Second is the 'Backcrossing', where, starting from the F1 individuals, three consecutive backcross cycles were simulated, each involving crossing F1 or later-generation hybrids with gametes from *E. flaviae*. This process generated hybrid classes (100 individuals per generation) ranging from first-generation backcrosses (BC1) to third-generation backcrosses (BC3), progressively reducing the genetic contribution from *E. lucius*. Finally, data were analysed by STRUCTURE as the real data. This allowed us to assess the loci's ability to detect *E. lucius* ancestry across the hybrid classes, even when its genetic contribution was substantially reduced.

3 | Results

3.1 | Sequencing Data Processing and Selection of Loci

The sequencing yielded a total of 47,691,492 raw reads, averaging 248,393 reads per sample across 192 samples (188 individuals, of which four were processed twice). Following paired-end merging, trimming and filtering, a total of 16,435,823 reads were acquired. Four individuals did not retain any reads: Individual No. 279 from the Garda population (*E. flaviae*), Individual No. 20003 from the Hatchery2020 stock (*E. flaviae*), Individual No. 21070 from the Livenza population (*E. flaviae*) and Individual No. 132_B from the Monticolo population (*E. lucius*). Read counts ranged from a minimum of 18,605 reads observed in Sample No. 111 from the Hatchery2011 stock (Hybrid) to a maximum of 257,861 reads observed in Sample No. 270 from the Garda population (*E. flaviae*). Sample No. 111 from the Hatchery2011 stock was excluded from the analysis due to missing data on over 40% of loci due to a low number of reads. On average, a total of 162,731 reads per locus were obtained, with an average of 147 reads per individual per allelic variant considering the 187 individuals (183 + 4 replicates) retained for the analysis.

Across the four technical replicates, the allele proportion identity for each pair of samples was 91% on average (ranging from 89% to 93%). This indicates a robust reproducibility of both the amplification and sequencing procedures. The majority of discrepancies between replicates was observed in loci exhibiting amplification issues, such as extensive missing data, low coverage or supernumerary alleles.

Among the initial 86 loci, 25 (29%) were excluded due to missing data exceeding 25% across individuals. The remaining 61 loci demonstrated consistent genotypes across replicates. However, 15 of these loci (17.4%) were subsequently eliminated due to either lack of variability (resulting in monomorphic loci) or the presence of additional alleles in at least one individual. The occurrence of the latter case was primarily attributed to two factors: the presence of repeated sequence stretches, which could pose challenges during sequencing and post-processing (such as paired-end merging), or the existence of duplicated loci. The final dataset was composed of 46 MIP loci. Given that loci showing alleles with low or imbalanced sequencing depth in heterozygous individuals may lack complete reliability, analyses were also performed in parallel by using a subset of 33 loci considered the most consistent.

3.2 | Descriptive Statistics

H_O, H_E, F_{IS}, N_A per locus, A_R and pA_R are reported in Table S2 (Supporting Information S3) for MIPs (33 loci) and for microsatellite (16 loci) data at the species level. Both MIP and microsatellite data show lower H_E and H_O in southern pike, indicating a lower genetic variability compared to northern pike, although the N_A per locus is comparable. Also, F_{IS} is higher in southern pike, indicating a higher level of inbreeding compared to northern pike. Both marker types show similar A_R and pA_R in the two species; pA_R is lower compared to A_R, indicating that most alleles are shared between the two species and only a few allelic variants are private. It must be noted that microsatellites show a clearly higher N_A, A_R and pA_R compared to MIPs, indicating a higher level of variability captured by these markers. Interestingly, the relative trend between the two taxa is reversed across marker types: *E. flaviae* shows lower A_R and pA_R values than *E. lucius* when analysed with MIPs, but higher values when analysed with microsatellites.

None of the analysed MIP loci showed significant deviation from the Hardy–Weinberg equilibrium after correction for multiple testing.

3.3 | Species Identification

Clustering analysis using STRUCTURE was conducted on the entire dataset to evaluate the discriminatory power of the markers in distinguishing between the two species compared to the microsatellites (Gandolfi et al. 2017). Results obtained from MIP markers and microsatellites (Figure 2a) exhibit an almost perfect alignment. Both types of markers consistently identify two distinct groups corresponding to the species *E. flaviae* and *E. lucius*. Within southern pike populations, Garda, Trasimeno and the three hatcheries demonstrate the highest levels of ancestry in *E. flaviae*, whereas Caldonazzo, Argenta, Livenza and Adda display some individuals with varying degrees of admixture with *E. lucius*. Both datasets reveal a high degree of uniformity among northern pike populations, except for three admixed individuals in the Cadore population, identified by both analyses as hybrids, thus confirming the concordance between MIP and microsatellite data. As expected, the hybrid populations of Trentino Alto Adige and Hatchery2011 exhibit a high level of admixture, with all individuals displaying mixed ancestry in the two clusters, and some overlap in results between the two analyses. The overall agreement between the data obtained from the two marker systems demonstrates a similar efficacy in resolving species-level distinctions for both MIPs and microsatellites.

In the PCA, without a priori information of species identity, MIP data showed a clear distinction between the two species along the

10990755, 2026, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aqc.70329 by Fondazione Edmund Mach, Wiley Online Library on [10/02/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

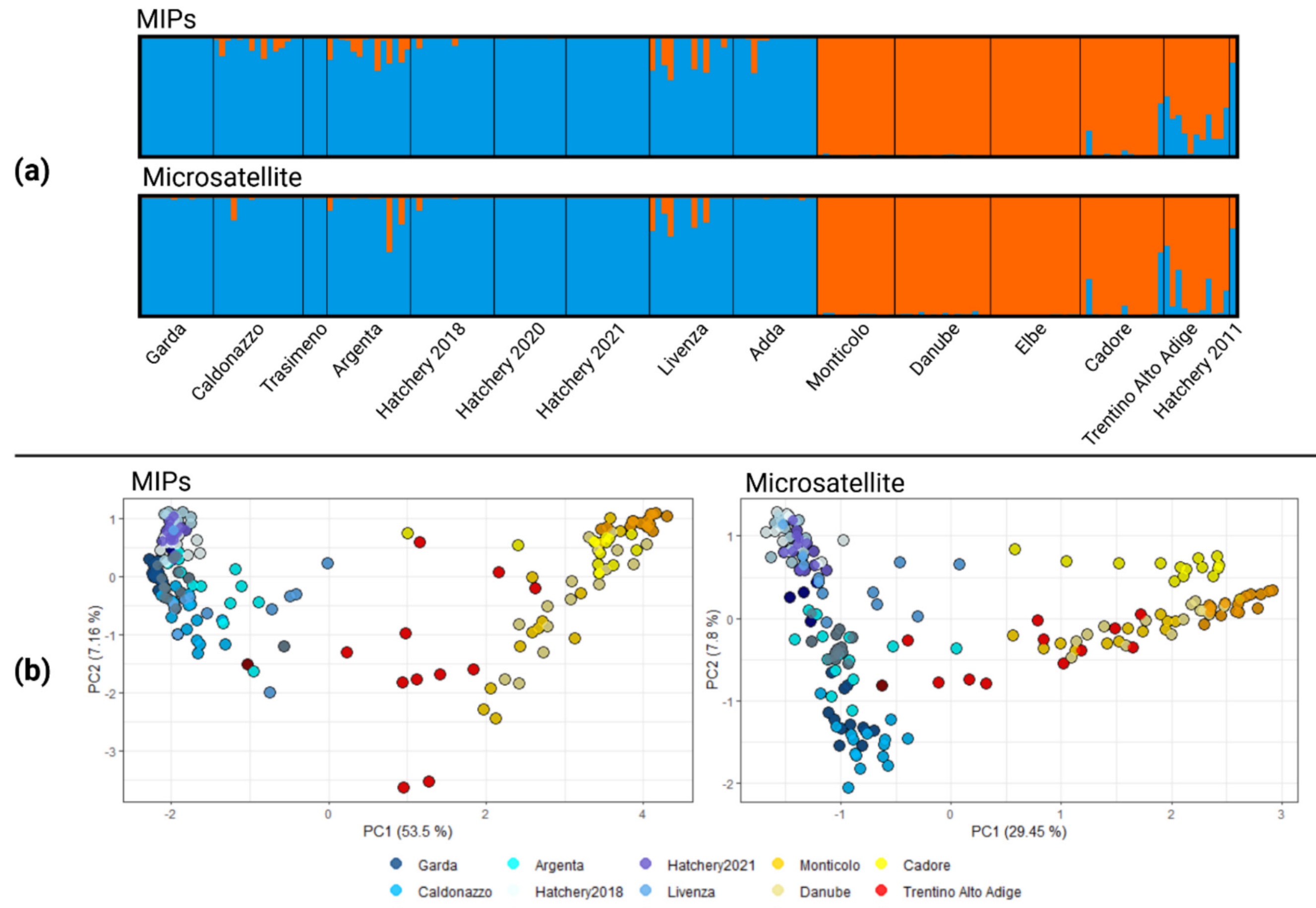

FIGURE 2 | The figure compares the genetic structure (a) and multivariate patterns (b) inferred from MIPs and microsatellite datasets, illustrating the consistency between the two marker systems in distinguishing *E. flaviae, E. lucius* and their hybrids. **(a)** Results of the STRUCTURE analysis ($K=2$) on southern pike populations based on 33 MIP loci (upper bar plot) and on 16 microsatellite loci (lower bar plot). *Each vertical bar represents one individual, coloured according to its estimated membership proportion in the two clusters. Vertical separators indicate population boundaries.* (b) Principal component analysis (PCA) based on 33 MIP loci (left) and on 16 microsatellite loci (right). PC1 vs. PC2 are shown; percentages of explained variance are reported on the axes. Colours indicate taxon: Southern pike (*E. flaviae*) populations are represented in blue tonalities, northern pike (*E. lucius*) populations are represented in yellow-orange tonalities and hybrids are in red.

first principal component (PC1, explaining 53.5% of the variance in the data), while individuals of the hybrid populations show intermediate position along the axis (Figure 2b, left). Microsatellite PCA (Figure 2b, right), similarly, is able to capture species distinction along the PC1 axis (explaining 29.45% of the variability). The hybrid population (Trentino Alto Adige) is coherently found in the intermediate position, along Axis 1, between the two species.

Regarding the ability of MIPs to distinguish between populations, the 33 MIP loci used are probably not sufficient. This consideration is better supported by the STRUCTURE bar plots (Supporting Information S4, Figures S1 and S2) rather than by the PCA in Figure 2b (left).

3.4 | Diagnostic Marker Validation

To evaluate the capacity of loci to detect hybridization and introgression, we conducted a simulation involving crosses between individuals from the two pure species *E. flaviae* and *E. lucius*. The initial cross produced F1, which was then backcrossed with one of the parental species (*E. flaviae*). This process was repeated across three generations of backcrossing. The results show that F1 hybrids were correctly assigned with nearly equal genetic contributions (50%) from both parental species. As expected, successive backcross generations displayed a progressive decrease in the genetic contribution from *E. lucius*, coupled with an increasing contribution from *E. flaviae*, the species used in the backcrosses (Figure 3). These findings highlight the robustness of our loci in detecting first-generation hybridization events and their effectiveness in identifying introgression signals even after the first backcross generation. Notably, the 10 selected loci for rapid identification through agarose gel demonstrated high amplification specificity, effectively distinguishing between species and degrees of hybridization (Figure 4). Remarkably, all newly analysed individuals, included to validate the assignment method, were correctly assigned to the respective species, either *E. flaviae* or *E. lucius*, in agreement with their geographic origin. This result underscores the reliability of this reduced panel of loci for accurate species identification and emphasizes its potential for detecting introgression events in individuals of unknown origin.

10990755, 2026, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aqc.70329 by Fondazione Edmund Mach, Wiley Online Library on [10/02/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

E. flaviae × E. lucius
First generation hybrids (F1) × E. flaviae
First generation backcrosse (BC1) × E. flaviae
Second generation backcrosse (BC2) × E. flaviae
Third generation backcrosse (BC3)
E. flaviae
E. lucius
33 loci
10 loci
Pure species
F1 hybrids
BC1
BC2
BC3

FIGURE 3 | Upper panel: schematic overview of the simulations performed to assess the sensitivity of the 33- and 10-loci panel in detecting hybridization and introgression signals. Specifically, first-generation hybrids were generated by crossing pure individuals of *E. flaviae* and *E. lucius*. F1 hybrids were then backcrossed with *E. flaviae* to produce first-generation backcrosses (BC1), which were further crossed with the same parental species to generate second- (BC2) and third-generation (BC3) backcrosses. Lower panel: two structure bar plots (generated using the dataset with 33 and 10 loci, respectively) showing the membership probabilities of assignment of individuals with different degrees of introgression.

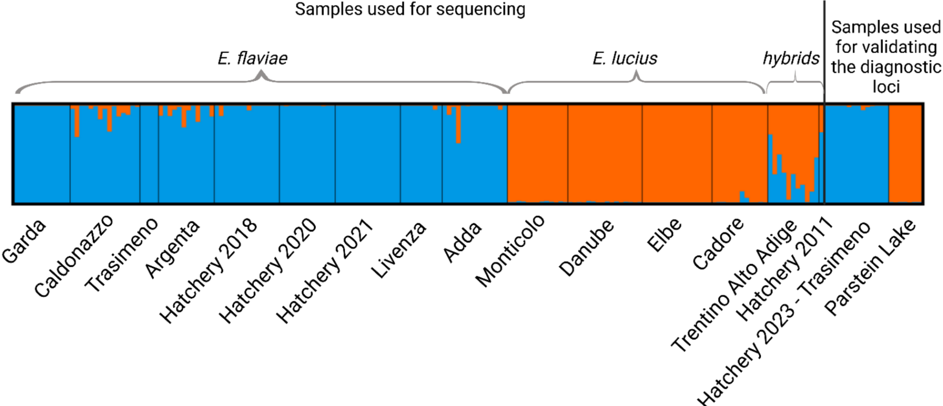

FIGURE 4 | Results of STRUCTURE analysis ($K=2$) on pike populations based on 10 MIP loci are represented by bar plots. The new individuals from Trasimeno and from Parstein Lake (with unknown genetic traits) are attributed to the *E. flaviae* and *E. lucius* species, respectively.

4 | Discussion

The results of this study demonstrate the effectiveness of MIPs as a powerful genetic tool for distinguishing northern pike (*E. lucius*) from southern pike (*E. flaviae*) and for accurately detecting admixture evidence between these species. Our findings show that the MIP panel provides reliable species discrimination. In some cases, such as the population from Trentino, which was identified as non-pure based on microsatellite analysis, the MIPs not only confirm this result but also appear, based on both STRUCTURE analysis and PCA, to clearly detect introgression even better than microsatellites. The slight differential resolution observed between MIPs and microsatellites may stem from their inherent molecular characteristics. Microsatellites

10990755, 2026, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aqc.70329 by Fondazione Edmund Mach, Wiley Online Library on [10/02/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

are highly variable and excel in capturing fine-scale genetic differences within populations (Hess et al. 2011; Zimmerman et al. 2020). However, this high variability may mask the overall proportion of genetic variation that can be attributed to species differentiation. Furthermore, the high mutation rate of microsatellites increases the risk of homoplasy, potentially leading to an underestimation of genetic distances between species (Zimmerman et al. 2020). In contrast, as clearly shown by the lower average N_A per locus, MIPs are less polymorphic and evolve at a slower (more conservative) rate than microsatellites (Boscari et al. 2024). This characteristic enhances their ability to capture species-level genetic differentiation over extended evolutionary timescales and therefore their robustness for interspecific diagnostic applications, whereas microsatellites remain more informative for fine-scale population structure within species. In addition, MIPs have the added advantage of providing sequence information, which can be leveraged in various contexts, including the development of single-locus targeted strategies such as in our case for the rapid and reliable identification of the pike species.

Overall, MIPs and microsatellites should be regarded as complementary rather than alternative tools: MIPs offer reproducible and transferable interspecific diagnostic markers, whereas microsatellites remain powerful for detailed intra-specific and population-level analyses.

Our study also demonstrates that MIPs effectively identify diagnostic polymorphisms suitable for specie-specific SNP-targeted PCR, which can ease robust species assignment in pike. The selected panel of 10 loci unambiguously assigned previously uncharacterized individuals to the nominal species. This panel offers a practical resource for species identification and routine monitoring, especially within conservation frameworks where detecting hybrids and introgression events is crucial to maintaining genetic integrity.

The possibility of avoiding sequencing steps—typically more time-consuming and expensive—while instead analysing individuals through simple species-specific PCR assays represents a major advantage of the diagnostic MIP panel. Similar strategies have been successfully applied in other taxa (e.g., *Pelophylax* species, Hauswaldt et al. 2012; Vucić et al. 2018; Meilink et al. 2024; sturgeon species, Boscari et al. 2013, 2017). However, many previous studies rely on one or a few loci, and species identification or hybrid detection based on single-locus information can be unreliable, particularly when different levels of introgression are involved.

The approach with MIPs overcomes this limitation by enabling the characterization of multiple diagnostic loci that are easily amplifiable across species and that display high-frequency private alleles in each taxon. This approach allows for the construction of robust, multilocus diagnostic panels that provide reliable species assignment and hybrid detection. A second advantage of this strategy lies in its applicability to single individuals. In conservation and management contexts, genetic identification is often required for isolated specimens of uncertain origin—such as individuals captured during monitoring or restocking events—and the ability to obtain results through rapid, low-cost PCR assays is extremely valuable. Conversely, the main limitation of this method is that when a large number of diagnostic loci are used, simultaneous processing of many individuals may become time-consuming due to the number of PCRs required. In such cases, the adoption of automated SNP genotyping platforms (e.g., SNP-chips; LaFramboise 2009, Pujolar et al. 2022) could represent a natural extension for large-scale applications. In our case, hybridization poses a significant threat to the genetic integrity of native pike populations, particularly in regions like Italy where the introduction of *E. lucius* has led to the admixture of the endemic *E. flaviae*. The ability to accurately detect and monitor hybridization events is essential for the development of effective management strategies aimed at preserving the unique genetic identity of *E. flaviae*. Overall, our study demonstrates that MIPs offer a robust method for achieving this goal, providing a valuable complement to existing molecular tools such as microsatellites, and the diagnostic MIP-based panel here developed offers a practical, transferable and scalable approach for the identification of *E. flaviae*, *E. lucius* and their hybrids, making it a valuable tool for conservation and management programmes where rapid and reliable genetic diagnosis is essential.

Unfortunately, we were not able to analyse additional hybrids with known genealogy to increase the sample size for testing the method's effectiveness. However, the ability of the proposed marker panel to detect different degrees of hybridization is strongly supported by the analyses of simulated hybrids with different degrees of contribution from the two species. The overlapping results obtained when analysing simulated genotypes using all 33 loci or the subset of the 10 most informative ones show that this restricted panel is suitable for the detection of *E. lucius* contribution. Moreover, this approach allows for a straightforward PCR followed by agarose electrophoresis, significantly reducing the time and costs of the purity analyses performed to inform restocking activities compared to microsatellites.

One advantage of MIPs compared to microsatellites lies in their high comparability across laboratories. While microsatellites can be standardized across laboratories, achieving full comparability typically relies on explicit calibration and quality control procedures (e.g., allele-binning harmonization and size-calling standardization). In the absence of such procedures, lab-specific scoring differences may occur, as documented in several methodological assessments (Zimmerman et al. 2020 and references therein). In contrast, MIP-based analyses are inherently more transferable and directly comparable with shared reference datasets. Furthermore, the reduced MIP panel developed in this study, due to its straightforward application based solely on PCR and agarose gel electrophoresis, can be easily implemented even in laboratories with limited resources. This approach significantly reduces both costs and processing times, making it an accessible and efficient tool for a wide range of research and monitoring applications.

5 | Conclusions

This study highlights the value of MIPs as a powerful tool for addressing conservation challenges posed by interspecific hybridization, specifically between the northern pike (*E. lucius*) and the southern pike (*E. flaviae*). MIPs proved to be effective in

10990755, 2026, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aqc.70329 by Fondazione Edmund Mach, Wiley Online Library on [10/02/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

distinguishing these two closely related species, offering good species-level resolution, similar to traditional microsatellite markers. The diagnostic potential of MIPs was underscored by their capacity to clearly identify hybrid individuals, allowing for reliable detection of genetic introgression. This tool's utility is particularly relevant for conservation efforts in areas like Italy, where the introduction of *E. lucius* has led to significant hybridization threats to native *E. flaviae* populations.

The findings suggest that MIP-based markers can represent a practical solution for routine species identification and monitoring in conservation programmes, where tracking genetic integrity is crucial. By enabling precise detection of hybrids, this MIP-based approach equips conservationists with an essential tool to manage and mitigate the impacts of hybridization on biodiversity. Overall, this study underscores the broader implications of MIPs for genetic monitoring and highlights their potential to enhance conservation strategies aimed at preserving native species in the context of human-mediated translocations and habitat changes.

Author Contributions

Annalisa Scapolatiello: formal analysis, visualization. **Chiara Samassa:** investigation, formal analysis. **Andrea Gandolfi:** conceptualization, validation, resources. **Maria Chiara Valerin:** investigation. **Lorenzo Zane:** conceptualization. **Elisa Boscari:** conceptualization, methodology, supervision, writing – original draft, validation, visualization, project administration.**Leonardo Congiu:** conceptualization, methodology, validation, writing-original draft.

Acknowledgements

The authors extend a warm thank you to Arne Ludwig (IZW, Berlin) and Enea Tentoni for kindly providing part of the samples and to Edoardo Viale for the support in the laboratory work. Open access publishing facilitated by Universita degli Studi di Padova, as part of the Wiley-CRUI-CARE agreement.

Artificial intelligence (ChatGPT) was used only to assist with the final linguistic revision of the manuscript. All scientific content was generated by the authors, who critically reviewed and take full responsibility for the final version.

Funding

A.G., L.Z. and L.C. were partially funded under the National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.4—Call for tender No. 3138 of 16 December 2021, rectified by Decree No. 3175 of 18 December 2021 of the Italian Ministry of University and Research funded by the European Union—NextGenerationEU; Award Number: project code CN_00000033, Concession Decree No. 1034 of 17 June 2022 adopted by the Italian Ministry of University and Research, CUP C93C22002810006, Project title 'National Biodiversity Future Center—NBFC'. Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission, nor can they be held responsible for them.

Ethics Statement

The authors have nothing to report.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

All Illumina raw data generated for this study have been uploaded to NCBI as bioproject PRJNA1356836.

References

Abbott, R., D. Albach, S. Ansell, et al. 2013. "Hybridization and Speciation." *Journal of Evolutionary Biology* 26, no. 2: 229–246. https://doi.org/10.1111/j.1420-9101.2012.02599.x.

Allendorf, F. W., R. F. Leary, P. Spruell, and J. K. Wenburg. 2001. "The Problems With Hybrids: Setting Conservation Guidelines." *Trends in Ecology & Evolution* 16, no. 11: 613–622. https://doi.org/10.1016/S0169-5347(01)02290-X.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. "Basic Local Alignment Search Tool." *Journal of Molecular Biology* 215, no. 3: 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

Bekkevold, D., F. Besnier, T. Frank-Gopolos, E. E. Nielsen, and K. A. Glover. 2024. "Introgression Affects *Salmo trutta* Juvenile Life-History Traits Generations After Stocking With Non-Native Strains." *Evolutionary Applications* 17: e13725. https://doi.org/10.1111/eva.13725.

Benjamini, Y., and Y. Hochberg. 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society, Series B: Statistical Methodology* 57, no. 1: 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

Bernardi, G. 2013. "Speciation in Fishes." *Molecular Ecology* 22, no. 22: 5487–5502. https://doi.org/10.1111/mec.12494.

Bianco, P. G. 2014. "An Update on the Status of Native and Exotic Freshwater Fishes of Italy." *Journal of Applied Ichthyology* 30, 62–77.

Bianco, P. I. E. R., and G. Delmastro. 2011. "Recenti novità tassonomiche riguardanti i pesci d'acqua dolce autoctoni in Italia e descrizione di una nuova specie di luccio." In *Researches on Wildlife Conservation*, 1–14. IGF Publication USA.

Boscari, E., A. Barmintseva, J. M. Pujolar, P. Doukakis, N. Mugue, and L. Congiu. 2013. "Species and Hybrid Identification of Sturgeon Caviar: A New Molecular Approach to Detect Illegal Trade." *Molecular Ecology Resources* 14, no. 3: 489–498. https://doi.org/10.1111/1755-0998.12203.

Boscari, E., S. D. Palle, N. Vitulo, et al. 2024. "MIPs: Multi-Locus Intron Polymorphisms in Species Identification and Population Genomics." *Scientific Reports* 14: 17870. https://doi.org/10.1038/s41598-024-68065-8.

Boscari, E., N. Vitulo, A. Ludwig, et al. 2017. "Fast Genetic Identification of the Beluga Sturgeon and Its Sought-After Caviar to Stem Illegal Trade." *Food Control* 75: 145–152. https://doi.org/10.1016/j.foodcont.2016.11.039.

Casanova, A., S. Heras, A. Abras, et al. 2022. "Genomic Hatchery Introgression in Brown Trout (*Salmo trutta* L.): Development of a Diagnostic SNP Panel for Monitoring the Impacted Mediterranean Rivers." *Genes* 13, no. 2: 255. https://doi.org/10.3390/genes13020255.

Coombs, J. A., B. H. Letcher, and K. H. Nislow. 2008. "CREATE: A Software to Create Input Files From Diploid Genotypic Data for 52 Genetic Software Programs." *Molecular Ecology Resources* 8, no. 3: 578–580. https://doi.org/10.1111/j.1471-8286.2007.02036.x.

Denys, G. P., A. Dettai, H. Persat, M. Hautecœur, and P. Keith. 2014. "Morphological and Molecular Evidence of Three Species of Pikes *Esox* spp. (Actinopterygii, Esocidae) in France, Including the Description of a New Species." *Comptes Rendus Biologies* 337, no. 9: 521–534. https://doi.org/10.1016/j.crvi.2014.07.002.

Denys, G. P. J., T. Lauga, G. B. Delmastro, and A. Dettai. 2018. "S7 Characterization of Western European Pikes *Esox* spp. (Actinopterygii, Esociformes)." *Cybium* 42: 221–228.

10990755, 2026, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aqc.70329 by Fondazione Edmund Mach, Wiley Online Library on [10/02/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Earl, D. A., and B. M. vonHoldt. 2012. "Structure Harvester: A Website and Program for Visualizing Structure Output and Implementing the Evanno Method." *Conservation Genetics Resources* 4: 359–361. https://doi.org/10.1007/s12686-011-9548-7.

Evanno, G., S. Regnaut, and J. Goudet. 2005. "Detecting the Number of Clusters of Individuals Using the Software STRUCTURE: A Simulation Study." *Molecular Ecology* 14, no. 8: 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x.

Excoffier, L., and H. E. L. Lischer. 2010. "Arlequin Suite ver 3.5: A New Series of Programs to Perform Population Genetics Analyses Under Linux and Windows." *Molecular Ecology Resources* 10, no. 3: 564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x.

Gandolfi, A., C. Ferrari, B. Crestanello, M. Girardi, L. Lucentini, and A. Meraner. 2017. "Population Genetics of Pike, Genus *Esox* (Actinopterygii, Esocidae), in Northern Italy: Evidence for Mosaic Distribution of Native, Exotic and Introgressed Populations." *Hydrobiologia* 794: 73–92. https://doi.org/10.1007/s10750-016-3083-1.

Gandolfi, A., D. Fontaneto, M. Natali, and L. Lucentini. 2015. "Mitochondrial Genome of *Esox flaviae* (Southern Pike): Announcement and Comparison With Other Esocidae." *Mitochondrial DNA Part A DNA Mapping, Sequencing, and Analysis* 27, no. 4: 3037–3038. https://doi.org/10.3109/19401736.2015.1063123.

Hauswaldt, J. S., M. Hoer, M. Ogielska, et al. 2012. "A Simplified Molecular Method for Distinguishing Among Species and Ploidy Levels in European Water Frogs (*Pelophylax*)." *Molecular Ecology Resources* 12, no. 5: 797–805. https://doi.org/10.1111/j.1755-0998.2012.03160.x.

Hess, J. E., A. P. Matala, and A. R. Narum. 2011. "Comparison of SNPs and Microsatellites for Fine-Scale Application of Genetic Stock Identification of Chinook Salmon in the Columbia River Basin." *Molecular Ecology Resources* 11, no. 1: 137–149. https://doi.org/10.1111/j.1755-0998.2010.02958.x.

Jeanroy, C., and G. P. J. Denys. 2019. "Morphological Traits Allow Distinguishing Their Hybrids From the Northern Pike, *Esox lucius*, and the Aquitanian Pike, *Esox aquitanicus* (Actinopterygii, Esociformes)." *Cybium* 43, no. 3: 227–232. https://doi.org/10.26028/CYBIUM/2019-433-003.

Jombart, T. 2008. "adegenet: A R Package for the Multivariate Analysis of Genetic Markers." *Bioinformatics* 24, no. 11: 1403–1405. https://doi.org/10.1093/bioinformatics/btn129.

Kalendar, R., B. Khassenov, Y. Ramankulov, O. Samuilova, and K. I. Ivanov. 2017. "FastPCR: An In Silico Tool for Fast Primer and Probe Design and Advanced Sequence Analysis." *Genomics* 109, no. 3–4: 312–319. https://doi.org/10.1016/j.ygeno.2017.05.005.

Kalinowski, S. T. 2005. "HP-RARE 1.0: A Computer Program for Performing Rarefaction on Measures of Allelic Richness." *Molecular Ecology Notes* 5, no. 1: 187–189. https://doi.org/10.1111/j.1471-8286.2004.00845.x.

Kearse, M., R. Moir, A. Wilson, et al. 2012. "Geneious Basic: An Integrated and Extendable Desktop Software Platform for the Organization and Analysis of Sequence Data." *Bioinformatics* 28, no. 12: 1647–1649. https://doi.org/10.1093/bioinformatics/bts199.

Koch-Schmidt, P., P. Tibblin, O. Engstedt, P. Larsson, and A. Forsman. 2015. "Pike *Esox lucius* as an Emerging Model Organism for Studies in Ecology and Evolutionary Biology: A Review." *Journal of Fish Biology* 87, no. 3: 472–479. https://doi.org/10.1111/jfb.12712.

Kopelman, N. M., J. Mayzel, M. Jakobsson, N. A. Rosenberg, and I. Mayrose. 2015. "CLUMPAK: A Program for Identifying Clustering Modes and Packaging Population Structure Inferences Across K." *Molecular Ecology Resources* 15, no. 5: 1179–1191. https://doi.org/10.1111/1755-0998.12387.

Kottelat, M., and J. Freyhof. 2007. *Handbook of European Freshwater Fishes*. Vol. 13. Publications Kottelat.

LaFramboise, T. 2009. "Single Nucleotide Polymorphism Array: A Decade of Biological, Computational and Technological Advances." *Nucleic Acids Research* 37, no. 13: 4181–4193. https://doi.org/10.1093/nar/gkp552.

Lucentini, L., M. E. Puletti, C. Ricciolini, et al. 2011. "Molecular and Phenotypic Evidence of a New Species of Genus *Esox* (Esocidae, Esociformes, Actinopterygii): The Southern Pike, *Esox flaviae*." *PLoS ONE* 6: e25218.

Mandeville, E. G., A. W. Walters, B. J. Nordberg, K. H. Higgins, J. C. Burckhardt, and C. E. Wagner. 2019. "Variable Hybridization Outcomes in Trout Are Predicted by Historical Fish Stocking and Environmental Context." *Molecular Ecology* 28, no. 16: 3738–3755. https://doi.org/10.1111/mec.15175.

McKinney, W. 2010. "Data Structures for Statistical Computing in Python." *SciPy Proceedings* 445.1: 51–56. https://doi.org/10.25080/Majora-92bf1922-00a.

Meilink, W. R. M., J. W. Arntzen, J. J. C. W. van Delft, and B. Wielstra. 2015. "Genetic Pollution of a Native Threatened Crested Newt Species Through Hybridization With an Invasive Congener in the Netherlands." *Biological Conservation* 184: 145–153. https://doi.org/10.1016/j.biocon.2015.01.022.

Meilink, W. R. M., F. D. van Beek, I. van de Breevaart, et al. 2024. "Difficulties in Species Identification in Water Frogs (Genus *Pelophylax*) Using Morphological and Molecular Markers in the Netherlands." *Alytes* 41, no. 1–4: 49–55. https://doi.org/10.5281/zenodo.16903428.

Meraner, A., L. Cornetti, and A. Gandolfi. 2014. "Defining Conservation Units in a Stocking-Induced Genetic Melting Pot: Unravelling Native and Multiple Exotic Genetic Imprints of Recent and Historical Secondary Contact in Adriatic Grayling." *Ecology and Evolution* 4, no. 8: 1313–1327.

Meraner, A., A. Venturi, G. F. Ficetola, S. Rossi, A. Candiotto, and A. Gandolfi. 2013. "Massive Invasion of Exotic *Barbus barbus* and Introgressive Hybridization With Endemic *Barbus plebejus* in Northern Italy: Where, How and Why?" *Molecular Ecology* 22: 5295–5312. https://doi.org/10.1111/mec.12470.

Pearson, K. 1901. "On Lines and Planes of Closest Fit to Systems of Points in Space." *Philosophical Magazine* 2, no. 11: 559–572. https://doi.org/10.1080/14786440109462720.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. "Inference of Population Structure Using Multilocus Genotype Data." *Genetics* 155, no. 2: 945–959. https://doi.org/10.1093/genetics/155.2.945.

Pujolar, J. M., M. T. Limborg, M. Ehrlich, and C. Jaspers. 2022. "High Throughput SNP Chip as Cost Effective New Monitoring Tool for Assessing Invasion Dynamics in the Comb Jelly *Mnemiopsis leidyi*." *Frontiers in Marine Science* 9: 1019001. https://doi.org/10.3389/fmars.2022.1019001.

R Core Team. 2022. "R: A Language and Environment for Statistical Computing." R Foundation for Statistical Computing Vienna, Austria. https://www.R-project.org/.

Rhymer, J. M., and D. Simberloff. 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.

Rousset, F. 2008. "genepop'007: A Complete Re-Implementation of the GenePop Software for Windows and Linux." *Molecular Ecology Resources* 8, no. 1: 103–106. https://doi.org/10.1111/j.1471-8286.2007.01931.x.

Scapolatiello, A., E. Boscari, L. Schiavon, N. Vitulo, and L. Congiu. 2025. "Intronomics-MIP: A Snakemake Pipeline for Analyzing Multilocus Intron Polymorphisms in Species Identification and Population Genomics." *BMC Research Notes* 18: 203. https://doi.org/10.1186/s13104-025-07264-6.

10990755, 2026, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aqc.70329 by Fondazione Edmund Mach, Wiley Online Library on [10/02/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Skog, A., L. A. Vøllestad, N. C. Stenseth, A. Kasumyan, and K. S. Jakobsen. 2014. "Circumpolar Phylogeography of the Northern Pike (*Esox lucius*) and Its Relationship to the Amur Pike (*E. reichertii*)." *Frontiers in Zoology* 11: 1–8.

Tamura, K., G. Stecher, and S. Kumar. 2021. "MEGA11: Molecular Evolutionary Genetics Analysis Version 11." *Molecular Biology and Evolution* 38, no. 7: 3022–3027. https://doi.org/10.1093/molbev/msab120.

Tarkan, A. S., I. Kurtul, D. Błońska, J. R. Britton, and P. J. Haubrock. 2024. "Resolving the Issues of Translocated Species in Freshwater Invasions." *NeoBiota* 93: 177–186. https://doi.org/10.3897/neobiota.93.122837.

Vucić, M., D. Jelić, G. I. Klobučar, B. Prkljačić, and M. Jelić. 2018. "Molecular Identification of Species and Hybrids of Water Frogs (Genus *Pelophylax*) From Lake Skadar, Southeast Adriatic Drainages (Amphibia: Ranidae)." *Salamandra* 54, no. 2: 147–157.

Wickham, H. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://doi.org/10.1007/978-3-319-24277-4.

Zimmerman, S. J., C. L. Aldridge, and S. J. Oyler-McCance. 2020. "An Empirical Comparison of Population Genetic Analyses Using Microsatellite and SNP Data for a Species of Conservation Concern." *BMC Genomics* 21: 382. https://doi.org/10.1186/s12864-020-06783-9.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Supporting Information S1:** This Supporting Information section includes the set of diagnostic MIP loci designed for rapid species identification and screening of hybridization between *E. flaviae* and *E. lucius*. The primer combinations and diagnostic SNPs enable low-cost PCR-based genotyping for conservation and monitoring programmes. **Table S1:** Panel of diagnostic loci for species identification. Each locus was targeted by three primers: two diagnostic primers each specific to a single species, based on identified polymorphisms, and a conserved primer that binds to a region common to both species. The nucleotides listed in the 'SNP' column are those that, in the case of a positive amplification with the corresponding primer pair, will form the genotype to be included in the GenePop file used for subsequent analyses. Sometimes they match the selective bases of the primers, while other times, they are their reverse complements, depending on whether the diagnostic primer is forward or reverse. **Supporting Information S2:** Python-based computational approach to conduct simulations of hybrid genotypes with varying levels of parental species contribution. **Supporting Information S3:** This Supporting Information section includes the genetic diversity estimates obtained for *Esox flaviae* and *E. lucius* using both marker systems. The tables summarize the variability patterns revealed by the 33 MIP loci and the 16 microsatellite loci, allowing a direct comparison of their genetic informativeness. **Table S2:** Genetic variability statistics evaluated on 33 MIP loci (a) and 16 microsatellite loci (b) in the two species *E. flaviae* and *E. lucius*. Sample size (No.), proportion of polymorphic loci (with more than one haplotype—P_N), observed heterozygosity (H_O), expected heterozygosity (H_E), inbreeding coefficient (F_{IS}), number of alleles (N_A) per locus, allelic richness (A_R) and private allelic richness (pA_R) are reported. A_R and pA_R are calculated considering a sample size of 58 individuals. The 12 hybrid individuals (Table 1) were not included in this analysis. **Supporting Information S4:** Structure-based analyses at the population level. **Figure S1:** Results of structure analysis ($K=4$) on southern pike wild populations based on 33 intronic loci (a) and on 16 microsatellite loci (b) are represented by bar plots. Vertical lines separate populations. Each vertical bar corresponds to one individual and shows different colours proportionally to the probability of assignment of the sample to each cluster. **Figure S2:** Results of structure analysis ($K=2$) on northern pike populations based on 33 intronic loci (a) and on 16 microsatellite loci (b) are represented by bar plots.

10990755, 2026, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aqc.70329 by Fondazione Edmund Mach, Wiley Online Library on [10/02/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License