bioRxiv preprint doi: https://doi.org/10.1101/2021.12.04.471201; this version posted December 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Population genetics reveals cryptic lineages and ongoing hybridization in a declining migratory fish species complex

Quentin Rougemont[1,2*], Charles Perrier[3], Anne-Laure Besnard[1], Isabelle Lebel[4,5], Yann Abdallah[4,6], Eric Feunteun[7], Elodie Réveillac[7,8*], Emilien Lasne[7,1*], Anthony Acou[8,9], David José Nachon[10], Fernando Cobo[10],Guillaume Evanno[1] Jean-Luc Baglinière[1], Sophie Launey[1]

[1]ESE, Ecology and Ecosystems Health, Agrocampus Ouest, INRAE, Rennes, France.

[2] CEFE, Univ Montpellier, CNRS, EPHE, IRD, Montpellier, France.

[3]UMR CBGP, INRAE, CIRAD, IRD, Institut Agro, Univ Montpellier, Montpellier, France.

[4]Migrateurs Rhône Méditerranée, Arles, France.

[5] Fédération Nationale de la Pêche en France et de la protection du milieu aquatique (FNPF), Paris, France.

[6]SCIMABIO, Thonon-les-Bains, France

[7]UMS OFB-CNRS-MNHN PatriNat, Station marine du Museum National d'Histoire Naturelle, Dinard, France.

[8]LIENSs, Univ La Rochelle CNRS, La Rochelle, France.

[9]Management of Diadromous Fish in their Environment OFB-INRAE-Institut Agro-UPPA, Rennes, France

[10]Department of Zoology, Genetics and Physical Anthropology, Universidade de Santiago de Compostela, Santiago de Compostela

Corresponding author: quentinrougemont@orange.fr

Abstract

Deciphering the effects of historical and recent demographic processes responsible for the spatial patterns of genetic diversity and structure is a key objective in evolutionary and conservation biology. Using genetic analyses, we investigated the demographic history, the contemporary genetic diversity and structure, and the occurrence of hybridization and introgression, of two species of anadromous fish with contrasted life history strategies and which have undergone recent demographic declines, the allis shad (*Alosa alosa*) and the twaite shad (*Alosa fallax*). We genotyped 706 individuals from 20 rivers and 5 sites at sea in Southern Europe at microsatellite markers. Genetic structure between populations was lower for the nearly semelparous species allis shad that disperse stronger distance compared to the iteroparous species, twaite shad. Individuals caught at sea were assigned at the river level for twaite shad and at the region level for allis shad. Using an approximate Bayesian computation framework, we inferred that the most likely long term historical divergence scenario between both species implicated historical separation followed by secondary contact accompanied by strong population size decline. Accordingly, we found evidence of contemporary hybridization and introgression between both species. Besides, our results support the existence of cryptic species in the Mediterranean sea. Overall, our results shed light on the interplay between historical and recent demographic processes and life history strategies in shaping population genetic diversity and structure of closely related species. The recent demographic decline of these species' populations and their hybridization should be carefully considered while implementing conservation programs.

bioRxiv preprint doi: https://doi.org/10.1101/2021.12.04.471201; this version posted December 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Introduction

Reconstructing how historical and recent processes shape present genetic diversity is an important step in evolutionary biology. In particular, long-term processes such as vicariance events, long term contraction during the last ice age or postglacial recolonization are well known to shape current patterns of genetic diversity (Hewitt 1996). In addition, several species and/or populations are currently declining at fast pace due to human activity or climate change (Ceballos et al. 2020, Ryan et al. 2018). The interplay of these processes may leave counter-intuitive signatures on contemporary patterns of genetic diversity, it is not clear whether they can be decorrelated and their respective roles quantified. In addition, populations within species or closely related species may differ in terms of life history traits, which can results in contrasting levels of population genetic structure and diversity.

In particular, the movement of species or dispersal processes are key factors affecting genetic structure through space and time (Cayuela et al. 2018). Many migratory species have undergone sharp decline worldwide (Wilcove and Wikelski 2008; Limburg and Waldman 2009). These declines in population size can have evolutionary consequences for these species. Indeed, small populations can display higher rates of genetic drift, accumulation of deleterious alleles and loss of genetic variability, which can ultimately threaten their genetic integrity and adaptive potential (Frankham 2005). Therefore, quantifying a species' evolutionary potential is vital for the conservation of wild and managed populations. This task can benefit from reconstruction of demographic history from genetic data (Rougemont et al. 2020), monitoring of population genetic parameters including genetic diversity, effective population size (*Ne*), and quantifying hybridization between closely related groups or species (Barton and Hewitt 1985) as well as their positive (Abbott et al. 2013) or negative consequences on hybrid fitness (Mikkelsen and Irwin 2021).

Diadromous fish are keystone species that move between freshwater and the marine environment. Many of these species, including salmonids, eels, sturgeons or shads, have undergone steep decline in their population size due to human activities such as dams building, over-harvesting, habitat degradation, and pollution (Parrish et al. 1998; Waters et al. 2000; Limburg and Waldman, 2009). Some of them (salmonids and shads) display a tendency to return to their natal river for breeding (i.e "homing" behavior), as demonstrated by otolith analysis (Tomás et al. 2005; Walther and Thorrold 2008; Perrier et al. 2011; Martin et al. 2015; Randon et al. 2017). This homing behavior fosters local adaptation, which is beneficial in stable environmental conditions (Keefer and Caudill 2014) but can also isolate some populations and results in smaller local effective population sizes. Consequently, in the context of environmental changes, this strategy may lead to less stable populations as compared to those displaying higher levels of gene flow and effective population size (Frankham 1997, 2005), although this question gave rise to recent debates (e.g. Kyriazis et al. 2020; Teixeira & Huber et al. 2021; Garcia-Dorado et al. 2021).

The allis shad, *Alosa alosa* (Linnaeus, 1758), and twaite shad, *Alosa fallax* (Lacépède, 1803), are two closely related anadromous clupeidae species that have undergone steep declines in their abundance and distribution range (Baglinière et al. 2003a). Since the middle of the 20th century, both species decline has mostly been attributed to freshwater habitat degradation, inducing loss of spawning grounds and to over-harvesting (Aprahamian et al. 2003; Limburg and Waldman 2009). Both species are now mostly restricted to large rivers in Portugal, Spain, France and the United Kingdom (mainly twaite shad). Presently, this decline appears stronger for the allis shad. For example, the last assessment of IUCN status for France show the allis shad is now considered as critically endangered (Baglinière et al. 2020a) while twaite shad is considered as vulnerable only (Baglinière et al. 2020b). It is expected that this decrease in abundance has had important repercussions on the genetic diversity of both species. Especially, a decrease in census population size will reduce effective population size and decreased selection efficacy. This decrease of abundance might have affected both species differently since they differ in their life history strategy despite their close phylogenetic relationship. First, *A. alosa* is nearly-semelparous while *A. fallax* is iteroparous (Mennesson-Boisneau et al. 2000). Second, *A. alosa* displays a stronger dispersal behavior than *A. fallax* (Taverny and Elie 2001; Jolly et al. 2012; Martin et al. 2015; Nachón et al. 2020). Following (Hasselman et

bioRxiv preprint doi: https://doi.org/10.1101/2021.12.04.471201; this version posted December 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

al. 2013) we predict that iteroparity sets a constraint on *A. fallax* dispersal and migratory distance due to stronger selection for homing every year for breeding (Jolly et al. 2012). On the contrary, longer migratory distance may facilitate gene flow in the semelparous *A. alosa* (Martin et al. 2015; Randon et al. 2017). Because of these difference in dispersal and parity mode, we expect stronger gene flow and lower population genetic structure in *A. alosa* compared to *A. fallax*. The genetic distinctiveness of *A. alosa* and *A. fallax* has been confirmed (Faria et al. 2004; Alexandrino et al. 2006; Jolly et al. 2012), but the details of their demographic history of prior isolation and subsequent contact has not been fully resolved (Faria et al. 2012; Taillebois et al. 2020). This issue could benefit from demographic modeling based on genetic data. Such an approach would notably help to decipher whether these species evolved separately before a secondary contact, or whether they could have evolved in sympatry, which has important implications for understanding speciation processes, as well as for the conservation of the species. Especially, this might influence the rate at which hybridization could result in a generalized meltdown of both species into one. In fact, hybrids between both species have been documented in some geographic areas like Portugal, United Kingdom or France (Alexandrino et al. 2006; Coscia et al. 2010; Jolly et al. 2011; Taillebois et al. 2020) suggesting that barriers to gene flow between species are permeable. In the context of river fragmentation, barriers (e.g. dams) generate "forced" common spawning grounds between *A. alosa* and *A. fallax* (Alexandrino et al. 2006), which would otherwise be spatially segregated along the river network (Mennesson-Boisneau et al. 2000). It is therefore expected that both species hybridize frequently in rivers with high levels of fragmentation (see also Taillebois et al. 2020). In addition, the taxonomic status of *A. fallax* within the Mediterranean Sea is debated (Chiesa et al. 2014; Baglinière et al. 2020c), despite genetic and morphological support for the existence of two lineages (Le Corre et al. 2005; Bianco, 2005) and our sampling design may help resolve this taxonomic conundrum.

In this study we used 13 microsatellite markers (Rougemont et al. 2015) to *i)* determine the extent of genetic diversity and structure within the genus *Alosa* along the Atlantic and Mediterranean coasts, *ii)* determine the historical process of divergence between both species using approximate Bayesian computations, *iii)* contrast the patterns of long term historical gene flow between species and within major genetic groups versus contemporary dispersal at sea in each species, i*v)* test the power of this marker set to assign individuals captured at sea. Doing so, we tested the hypotheses that *i)* both species were formerly isolated and came into secondary contact, *ii)* both species differ in their genetic structure, with the more dispersive, semelparous species showing lower genetic structure than the iteroparous species, *iii)* hybrids might be frequently found, potentially as a result of frequent "forced" common spawning grounds and *iv)* individuals caught at sea can be confidently assigned to rivers or geographic region, depending on the species.

Materials and Methods

Study area and sampling

A total of 706 individuals from both species (367 *A. alosa* and 339 *A. fallax*) were sampled in 20 rivers distributed along the French Atlantic coast (N=514), Spanish coast (N= 56 *A. fallax*), Mediterranean coast (N=92 *A. fallax*) and Corsica (N= 65 *A. fallax*). A subset of these (40 *A. alosa* and 65 *A. fallax*) were captured at sea along the Atlantic coast and were also used for population assignment (**Fig. 1, Table1** and **Table S1**). Individuals collected in rivers were captured between 2009 and 2012 either by nets or trapping, a fin clip was stored in 95% ethanol while individuals captured at sea were collected by professional fishermen and identified phenotypically. Scales were also collected from carcasses after breeding migration and conserved in paper envelopes. Collections of scales (stored at INRAE Rennes, France) from cohorts 1997 to 1999 and 2001 (Aulne and Charente rivers) were also used (**Table1**).

Molecular methods

Genomic DNA was extracted using a Chelex protocol (modified from Estoup et al. 1996). We genotyped each individual at 13 microsatellite loci specifically developed for allis shad and twaite shad (Rougemont et

bioRxiv preprint doi: https://doi.org/10.1101/2021.12.04.471201; this version posted December 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

al. 2015). Occurrence of null alleles and scoring errors due to large allelic dropout or stuttering were checked using Micro-Checker v2.2.3 (Oosterhout et al. 2004). Linkage disequilibrium (LD) was inspected using GENEPOP 4.0 (Rousset 2008).

Hybrid identification

Prior to any further analysis we were interested in accurately distinguished purebred individuals from hybrids. We performed this identification before other analyses since the presence of hybrids could bias population genetics estimates. We used a similar framework to the one of Vähä and Primmer (2006) using STRUCTURE 2.3.3 (Pritchard et al. 2000) and NEWHYBRIDS 1.1 (Anderson and Thompson 2002). For the purpose of hybrid identification we assumed K=2 where each cluster is composed of a single species contributing to the total gene pool of the sample. STRUCTURE was used without prior information about the species classification using an admixture model with (*i*) correlated allele frequency (Falush et al. 2003) and (*ii*) 500.000 burn-in steps followed by 1.000.000 MCMC iterations, replicated 10 times. Individual admixture proportions (q-values) and their 90% confidence intervals were averaged over the ten replicates and used to assign individuals to their respective cluster. NEWHYBRIDS assumes that the sample is drawn from a mixture of pure individuals and hybrids (F1, F2 or backcrosses) so that the q-value inferred with this method is a discrete variable (Anderson and Thompson, 2002). NEWHYBRIDS was used to calculate the posterior probability for an individual to belong to one of the following classes: purebred *A. alosa*, purebred *A. fallax*, hybrid F1, hybrid F2 or backcrosses. The q-values were summed over all hybrids categories (F1, F2 and backcrosses) because preliminary tests showed that performance were greater under these conditions. Hence, hybrids, regardless of their hybrid class, were distinguished from purebreds. Uniform priors were used for allele frequency and admixture estimations with a burn in of 1.000.000 steps followed by 1.000.000 iterations. Tests using Jeffrey priors, instead of uniform, yield similar results. Results are based on the average of 10 runs performed with various random seed number.

Performance of Admixture Analysis

To test the correctness of assignment of STRUCTURE and NEWHYBRIDS, the software HYBRIDLAB (Nielsen et al. 2006) was used. Individuals showing q-value >0.9 with both methods were chosen randomly and 3500 simulated genotypes of each parental species were generated and divided into 10 datasets of 350 individuals. The sample size was chosen to be close to our real dataset. Then 10 other datasets were created containing both parental and hybrid individuals (F1, F2, backcrosses). Ten individuals of each hybrid category were incorporated into the datasets of 350 individuals to represent the actual sample size of our empirical dataset and to incorporate a small proportion of hybrids. The simulated dataset was analyzed with STRUCTURE and NEWHYBRIDS to calculate (1) the hybrid proportion, which is the number of individuals classified as hybrid divided by the total number of individuals in the sample, (2) the efficiency, (3) accuracy and (4) overall performance of these methods following the definition of Vähä and Primmer (2006). We used threshold q-values of 0.90 for hybrid classification.

Genetic diversity and HWE within populations

Fish captured at sea for which the river of origin was unknown were excluded from these analyses as well as hybrids leaving a total of 586 fish (315 *A. alosa* and 271 *A. fallax)*. For each population (sample in each river) of each species, diversity indices (observed heterozygosity (*Ho*) and expected Heterozygosity (*He*) under Hardy Weinberg Equilibrium (HWE)) were calculated with GENETIX 4.0.5 (Belkhir et al. 1996). Significance of heterozygosity comparisons was assessed with a Wilcoxon sign rank test. Deviations from HWE and linkage disequilibrium were tested for each locus within each population and globally in GENEPOP 4.0 (Rousset 2008). FSTAT version 2.9.3 (Goudet 1995) was used to calculate inbreeding coefficient (*Fis*), allele number, and allelic richness (*Ar*), with *Ar* differences between species tested using 4.000 permutations.

bioRxiv preprint doi: https://doi.org/10.1101/2021.12.04.471201; this version posted December 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Population genetic structure

Genetic structure was assessed excluding hybrids (based on their genotype) and fish captured at sea. The extent of genetic structure was quantified using the pairwise F_{ST} estimator θ_{ST} of (Weir and Cockerham 1984) between sample pairs with Fstat, with significance assessed using 10,000 permutations. We tested for a signal of isolation by distance (IBD) in our data in both species separately using a mantel test on matrices of linearized F_{ST} using $F_{ST}/(1 - F_{ST})$ against the waterway distances measured in ArcGIS. The chord distance (Cavalli-Sforza and Edwards 1967) was used to quantify genetic differentiation between sample pairs and to construct neighbor-joining phylograms (Saitou and Nei 1987) with MSA 4.05 (Dieringer and Schlötterer 2003). Trees were computed using the software Phylip 3.6 (Felsenstein 1995) with 10,000 permutations and visualized using Tree view 1.6 (Page, 1996). Individual genetic clustering was further investigated without *a priori* definition of population boundaries using STRUCTURE. The admixture and correlated allele frequency model was used to detect a number of genetics cluster (K) varying from 1 to 14 separately for *A. alosa* and for *A. fallax*. Fifteen replicates were run for each K with a burn-in period of 400 000 MCMC followed by 400 000 MCMC iterations. Optimal number of cluster was evaluated using the likelihood distribution (Pritchard et al. 2000) and the ΔK method (Evanno et al., 2005). Results were combined with Structure harvester (Earl and vonHoldt 2012) and graphs were ploted using Distruct 1.1 (Rosenberg 2004). We finally used a multivariate method, the Discriminant Analysis of Principal Component Adegenet Package (Jombart 2008) implemented in R (R Development Core Team 2012), to inspect structure between populations for each species.

Genetic stock assignment

Genetic assignment of fish captured at sea was performed using three different methods. First STRUCTURE was run with fish captured at sea included in the dataset. The same settings as described above were used. Second, DAPC assignments were performed using fish from the sea as supplementary individuals projected using the previously defined parameters. Third, we used the Bayesian method implemented in GeneClass 2 (Piry et al. 2004) specifically designed for assignment tests. The likelihood that any fish captured at sea effectively came from one of the populations was tested with the (Paetkau et al. 2004) algorithm using 100 000 simulated individuals. Fish that displayed a probability <0.01 were assumed to come from an un-sampled reference population and were excluded from the assignment tests.

Demographic history: approximate Bayesian computations

The demographic history of divergence between *A. alosa* and *A. fallax* was investigated using an ABC framework. We excluded all putative hybrids to avoid favoring models of ongoing gene flow while our focus was on long-term patterns of gene flow. A total of four scenarios of divergence were compared (**Fig S1**). The model of strict isolation (SI) assumes that an ancestral population of size N_{ANC} splits instantaneously at T_{split} into two daughter populations of constant and independent size N_{pop1} and $N_{pop2.}$ The split is not followed by any gene flow and the population can either undergo an instantaneous bottleneck or an expansion at T_{split}. In contrast, the three other models assume various rates of gene flow following the instantaneous split. This migration occurs at a rate $M = 4\ N_0.m.$ with $M_{1\leftarrow 2}$ being the number of migrants from population 2 to population 1 and $M_{2\leftarrow 1}$being the reciprocal. In the model of ancestral migration (AM) the first generations of divergence are followed by gene flow until T_{am} (**Fig S1**) at which point there will be no further gene flow (looking forward in time). In the model of isolation with migration (IM) genes flow occurs continuously from T_{split} to the present and a constant rate each generation. Under the secondary contact (SC) model, T_{split} is followed by a period of strict isolation and by a period of secondary contact T_{sc} generations ago that is still ongoing. We used large and uninformative priors for model choice and parameter estimation. Simulation of microsatellites strictly followed the procedure of Illera et al. (2014) later modified by Rougemont et al. (2016). For each simulation, summary statistics were then computed and used in the ABC model choice implementing a neural network, as well as for parameter estimation (Csilléry et al. 2010). For each pairwise comparison, the robustness of the model choice was measured

bioRxiv preprint doi: https://doi.org/10.1101/2021.12.04.471201; this version posted December 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

using 4,000 pseudo observed datasets (PODS) sampled randomly from each model and the same ABC procedure as for the empirical dataset was performed but considering a given model M instead of the empirical data. Then the robustness between two given models M_1 and M_2 was computed as:

$$P(P_{M1} = P \mid M_1) / [\, P(P_{M1} = P \mid M_1) + P(P_{M1} = P \mid M_2)].$$

Where $P(P_{M1} = P \mid M_1)$ represents the probability of correctly supporting *M1* given the observed posterior probability *P* and $P(P_{M1} = P \mid M2)$ is the probability of erroneously supporting M_1 given that the true model is M2 (Fagundes et al., 2007). Parameter estimation was performed for the best model by running another round of simulation for a total of 1,4 million simulations. The whole pipeline used for ABC can be found at: https://github.com/QuentinRougemont/MicrosatDemogInference.

Results

A total of 706 individuals were genotyped at 13 loci and 180 different alleles were found with 153 in *A. alosa* and 138 in *A. fallax*. No evidence of null alleles was detected. In *A. alosa* one test of linkage disequilibrium (LD) out of 936 comparisons was significant. One LD test out of 858 comparisons was significant in *A. fallax*.

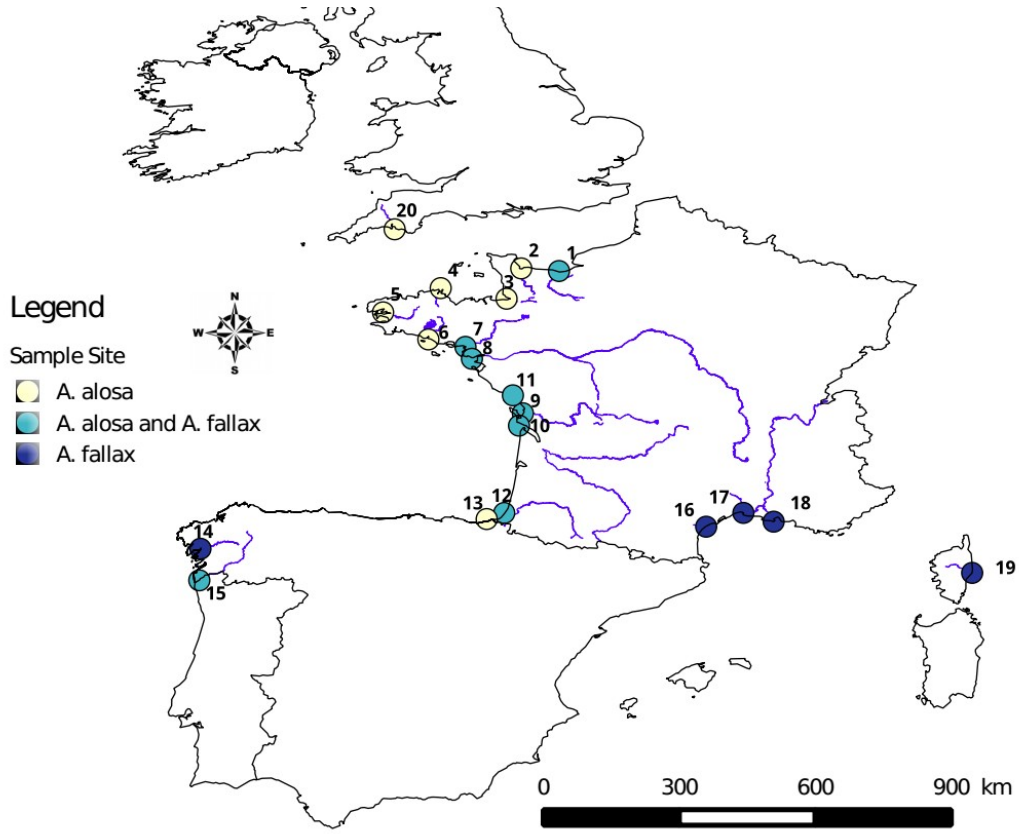

Figure 1: *Map showing the location of sampling sites across France, United-Kingdom and Spain. Sample sites numbers match those given in table 1.*

Hybrid identification

Results of admixture analysis in NEWHYBRIDS and STRUCTURE using simulated genotypes from HYBRIDLAB varied depending on the presence or absence of hybrids (**Table S2)**. Both STRUCTURE and NEWHYBRIDS produced some false hybrids in the purebreds' dataset and this overestimation was greater in STRUCTURE than in NEWHYBRIDS (**Table S2**). In brief, NEWHYBRIDS displayed slightly higher efficiency and accuracy at a q-value threshold of 0.90 in the presence of hybrids with both software displaying accuracy and efficiency above 0.90.

bioRxiv preprint doi: https://doi.org/10.1101/2021.12.04.471201; this version posted December 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Regarding our empirical data, both species were separated in two fully distinct clusters with average q-values greater than 0.99 for *A. alosa* and *A. fallax* in both STRUCTURE and NEWHYBRIDS. STRUCTURE and NEWHYBRIDS reclassified 13 individuals captured at sea, morphologically identified as *A. alosa* but genotypically classified as *A. fallax* and 1 individual wrongly identified as *A. fallax* but genotypically classified as *A. alosa*. A total of 35 individuals displayed a q-value smaller than 0.9 in either of the methods and were not classified as purebreds. STRUCTURE found 31 putative hybrids for a threshold q-value smaller than 0.9. NEWHYBRIDS identified 12 hybrids at the q-value threshold of 0.9 while 22 individuals remained unassigned. All hybrids identified by NEWHYBRIDS had a q-value greater than 0.8 and were all identified as hybrids in STRUCTURE. Among the 35 putative hybrids, nine individuals were further identified as purebred (**Table S3**). Finally, 11 individuals were not identified as either purebred or hybrids as they displayed inconsistent q-value according to each method and were left unassigned (**Table S3**). From a geographic standpoint the majority (86%) of the 35 individuals came from three major areas: namely the Charente River and Pertuis Charentais (37.1%), the Loire River (25.7%) and South Brittany/Scorff River (22.9%).

Genetic diversity

One *Fis* value was significant for the marker Alo29 in the Loire River in *A. alosa* while other populations and other markers did not show any significant deviations from HWE and no *Fis* was significant. The total number of alleles varied from 4 to 20 in *A. alosa* and from 4 to 17 in *A. fallax*. The two species shared 62% of alleles. Mean adjusted allelic richness (*Ar*) per population ranged from 3.74 to 5.94 in *A. alosa* and from 3.44 to 5.70 in *A. fallax* (**Table 1** for details by river and **Table S4, Table S5** for details for each river and each marker). Mean observed heterozygosity was 0.596 (range: 0.49–0.62) in *A. alosa* and 0.523 (range: 0.44–0.59) in *A. fallax* and differed significantly between species (P<0.005, 5000 permutations). In contrast, *Ar* did not differ between species (P=0.061, 5000 permutations). For *A. alosa*, *Ar* and *He* were the greatest in the Dordogne, Minho and Loire River while they were the lowest in the Vire, Aulne and Trieux rivers. For *A. fallax*, *Ar* and *He* were the highest in the Rhône (Mediterranean) and the lowest in the Ulla, Orne and Dordogne (Atlantic).

Population genetic structure

Global inter-species F_{ST} was 0.240 (95%CI: 0.194 – 0.286) and significant (P<0.00087. 15 000 permutations). Global F_{ST} for *A. alosa* was 0.046 (95%CI: 0.034 0.058) (P<0.0001. 10 000 permutations). Global F_{ST} for *A. fallax* was higher with a value of 0.219 (95%CI: 0.175– 0.263) (P<0.0001. 10 000 permutation). In *A. alosa*, 89% of pairwise F_{ST} comparisons were significant after Bonferonni corrections (**Table 2**) compared to 82% in *A. fallax*. Levels of differentiation between sampling localities were smaller in *A. alosa* than in *A. fallax*. Populations sampled on the same river basin (i.e Rhône, or Tavignano) were not significantly structured (**Table 2)** but were significantly differentiated from all other rivers. Similarly, Mediterranean *A. fallax* populations (Aude, Rhône, Vidourle) were not differentiated among them but were strongly differentiated from Atlantic populations. The Tavignano (Corsica) was significantly differentiated from all other rivers (F_{ST} ranged between 0.233 and 0.358). *A. fallax* populations from the Minho, Ulla and Orne were also distinct from all other populations (F_{ST} ranged between 0.08 and 0.289), but again comparisons with the Tavignano were the highest (**Table 2, Table S6** for interspecific comparison).

Structure results were as follows: In *A. alosa* the highest likelihood was for K=6 (**Fig. S2**) while ΔK showed a peak for K= 3 and another for K=6 (**Fig. S2**). For K=3 the clustering separated the population into three geographic regions (Brittany, Normandy and Atlantic, **Fig. 2A**). However, the Nivelle individuals (Southern France) clustered with individuals from Normandy in Northern France (Vire and Orne river in Fig. 2A). For K=6 (**Fig. 2A**), the Nivelle individuals were clearly separated from Normandy and formed a single cluster (contribution = 0.851). Normandy individuals formed a separated cluster, as well as Brittany individuals. Populations from the Atlantic were separated into three admixed clusters.

bioRxiv preprint doi: https://doi.org/10.1101/2021.12.04.471201; this version posted December 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

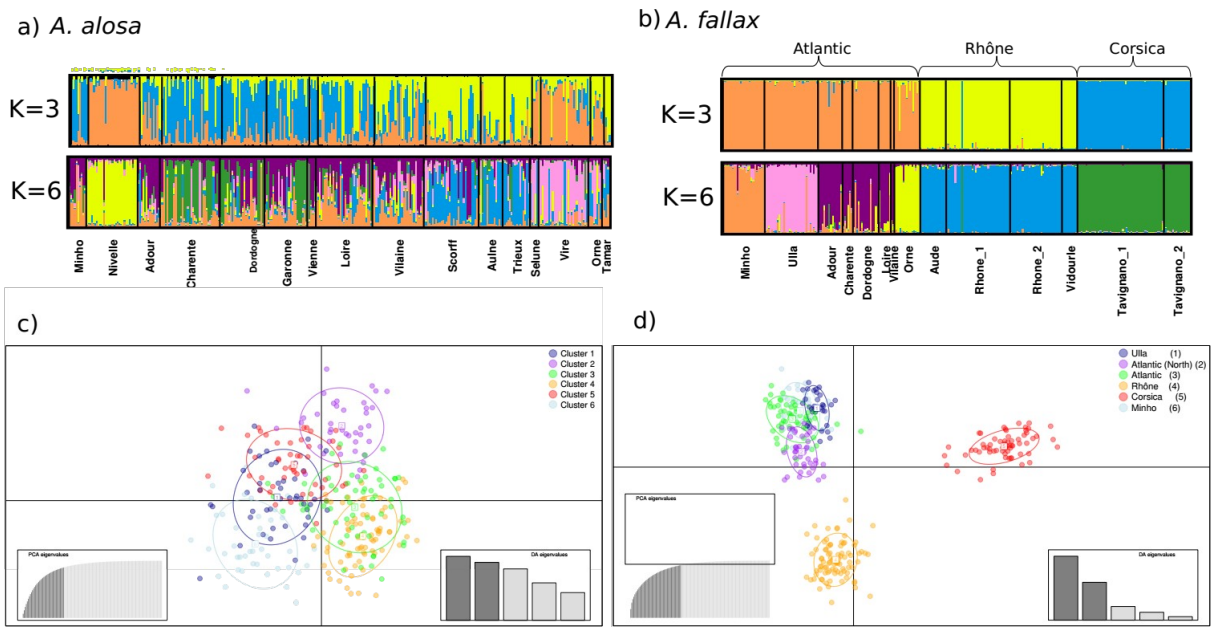

Figure 2: *A) Bayesian individual clustering performed with STRUCTURE for each species separately. In A. alosa K=6 clusters are identified. B) In A. fallax 3 and 6 clusters are identified. Individuals are represented by a bar with each colour representing its membership proportion to each cluster. C) plot of the Discriminant Analysis of Principal Components (DAPC) performed in A. alosa. Clusters are numbered from one to six given that mixed membership probability was often attributed to different rivers. Membership assignment of individuals averaged by river to each cluster can be found in* ***Table S8****). D) plot of the DAPC performed in A. fallax. Membership assignment of individuals averaged by river to each cluster can be found in* ***Table S9****)*

In *A. fallax*, the highest likelihood plateaued between K=6 and K=8. ΔK showed one strong peak for K=3 (**Fig. S2**). At K=3, individuals from the Atlantic, the North Mediterranean coast and from Corsica were separated. No sign of admixture was observed in *A. fallax* populations as opposed to the pattern observed in *A. alosa* (**Fig. 2B**). Indeed, all but the Orne River had >0.95 membership probability to form a separate cluster (the Orne still had a 0.86 membership probability to be a distinct cluster). K=6 revealed six geographic regions with low levels of admixture and allows a finer clustering of populations along the Atlantic Coast (**Fig. 2B**). The Orne (Normandy) and the two Spanish rivers (Ulla and Minho) formed three distinct clusters. The remaining populations from the French Atlantic Coast formed a single cluster on the Rhône River. One individual was clearly assigned as a migrant from the Tavignano (individual q-value = 0.981 CI: 0.894-1.00). Although some individuals displayed admixed ancestry, only one individual from the Ulla River was assigned with probability greater than 0.8 to the Minho suggesting relatively low dispersal among clusters as opposed to the pattern observed in *A. alosa*. The DAPC approach revealed a similar level of population structure in *A. alosa* (**Fig. 2C**) as well as in *A. fallax* (**Fig2D**). To reveal potential hierarchical structure and fine scale genetic structure we replicated the analysis in *A. fallax* including only populations from the Atlantic coast. This revealed the existence of 4 clusters (**Fig. S3**).

The neighbor-joining tree revealed a clear separation between *A. alosa* and *A. fallax* (**Fig. 3B**). In *A. alosa* 5 clusters can be delineated, with the Adour being apart and the Vilaine grouping with the Minho. In *A. fallax* three main clusters can be distinguished corresponding to a geographic clustering. Separation of *A. fallax* from the Tavignano was as strong as the separation between the two cryptic lineages from the Mediterranean Sea and the Atlantic coast. Finally, tests for IBD were significant in both species (Mantel tests: P<0.001. r = 0.621 in *A. alosa* and P<0.0001. r = 0.554 in *A. fallax.* **Fig. 3A).** These results were also supported by significant linear models (**Fig 3A).**

Assignment tests

According to GeneClass three individuals of *A. fallax* captured at sea displayed a probability <0.01 to originate from one of the river sampled and were excluded for further assignment with the three methods. The forty *A. alosa* had a probability >0.01 to originate from one of the rivers. Contrasting results were obtained in each species; *A. fallax* were assigned with higher probability than *A. alosa* (only fish with score

bioRxiv preprint doi: https://doi.org/10.1101/2021.12.04.471201; this version posted December 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

>0.9 were kept for assignment.) with a total of 34 to 45 and of 18 to 34 fish respectively assigned depending on the method (**Table 3**). All methods assigned the majority of *A. alosa* to the Atlantic cluster. The DAPC approach classified 6 *A. alosa* from South Brittany in Normandy and 4 in the Minho, while these fish were either assigned to the Atlantic cluster or left unassigned by the other methods. Similarly, one *A. alosa* from the Pertuis Charentais was assigned to the Minho by the DAPC while the two other methods classified this fish as belonging to the Atlantic cluster. In *A. fallax* the majority of fish from the North Sea, the South Brittany and the Southern part of the Bay of Biscay were assigned to the Orne cluster. Most fish from the Pertuis Charentais were assigned to the Atlantic cluster. One fish was assigned to the Ulla by the DAPC approach while it was left unassigned by the two other methods. Similarly, two fish from Southern Brittany were assigned to the Minho while they were left unassigned by STRUCTURE and one of them was assigned to the Atlantic cluster by GeneClass, the other being unassigned as well.

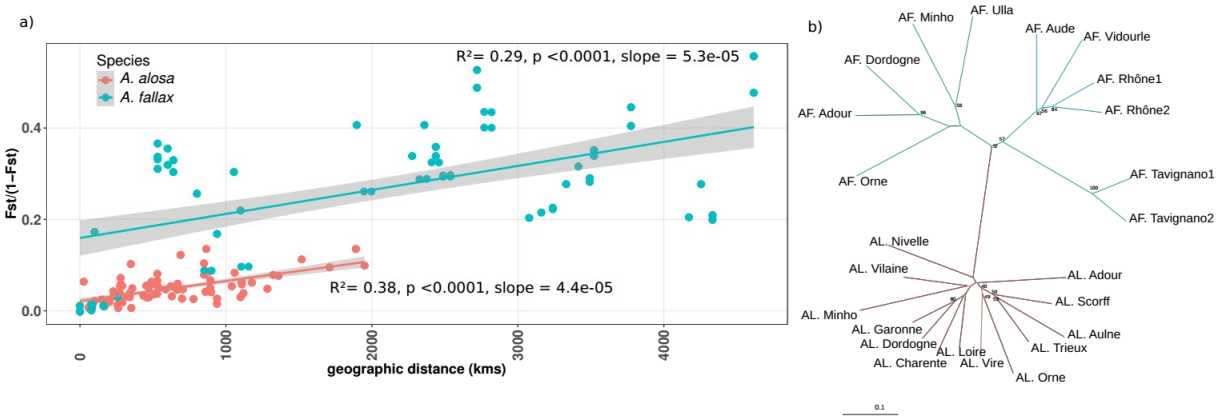

Figure 3: A) *Relationship between genetic distance as measured by $F_{ST}/(1-F_{ST})$ and distance in kilometers between populations of each species. All correlation were significant based on linear models and Mantel test.* ***B)*** *Neighbor-joining tree based on Cavalli-Svorza and Edwards (1967) genetic distances. Only bootstrap values >50% are indicated. Sample abbreviations match those given in Table 1.*

Demographic history

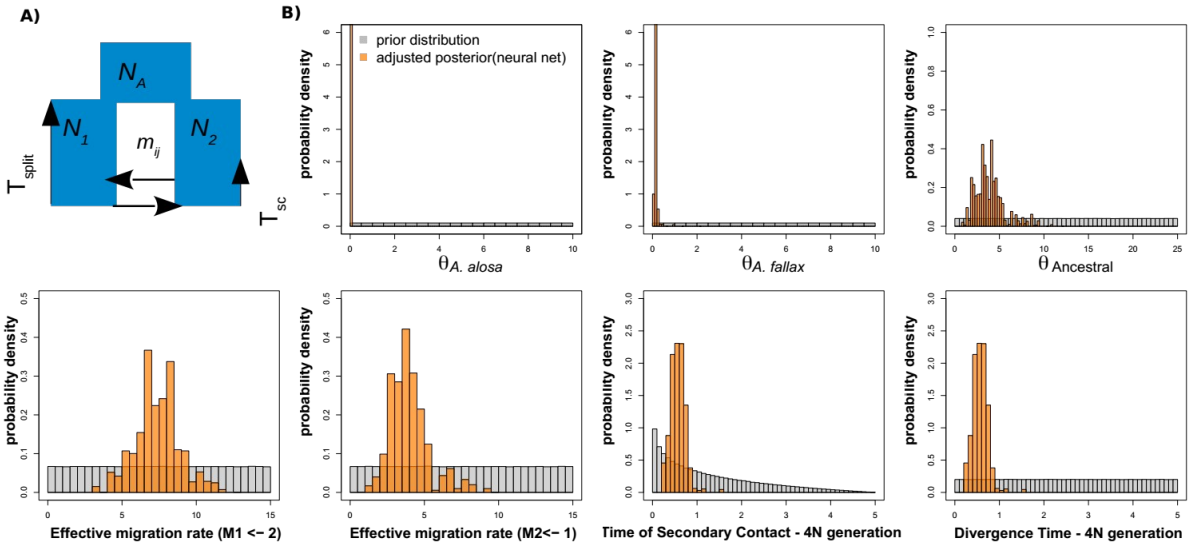

Figure 4: Best model of divergence and distribution of prior and posterior parameters*. A) Simplified representation of the best inferred scenario of divergence between Alosa alosa and A. fallax. B) Prior (grey) and Posterior (orange) parameter distribution inferred by Approximate Bayesian Computation. θ A. alosa = Scaled Effective population size of A. Alosa. θ A. alosa = Scaled Effective population size of A. fallax. $\theta_{ancestral}$= Effective population size of the ancestral population. All populations size are scaled by 4Nref*μ. T_{split} = split time (here T_{split} = tau* $4N_{ref}$). T_{sc} = Time of Secondary Contact. M12 = 4Nref*m_{12} and M21 = 4Nref*m_{21} correspond to the scaled migration rate where m_{ij} represents the fraction of subpopulation i which is made up of migrants from subpopulation j each generation. All values are provided in coalescent units and scaled by the effective reference population size.*

bioRxiv preprint doi: https://doi.org/10.1101/2021.12.04.471201; this version posted December 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Reconstruction of demographic history was performed between the two species. ABC model choice clearly rejected SI in favor of models with ongoing gene flow ($P_{(SC)}$ = 0.980 $P_{(IM)}$ = 0.851) or ancient gene flow ($P_{(AM)}$ = 0.720). In all cases the robustness was 1 for all comparisons (**Fig S4 A-F**). Comparison of model with ancient gene flow against ongoing gene flow then leads to a rejection of the former ($P_{(SC)}$ = 0.669. $P_{(IM)}$ = 0.599) with a robustness of 1 (**Fig S5A, B**). Finally, comparing IM and SC leads to a higher posterior probability of the latter with P(SC) = 0.738 and a robustness of 1 (**Fig 4a**, **Fig S5C,D**).

The posterior distribution of parameter estimates under SC were well differentiated from the prior yielding confidence for interpretations of the mean values and credible intervals (**Fig. 4b**). These indicated a striking pattern of population size reduction as compared to the estimated ancestral effective population size (N_{ANC}). The median estimated effective population size for *A. alosa* was N_1 = 209 [IC= 47 – 646]. for *A. fallax* N_2 = 2573 [940 – 5649] and for the ancestral population N_{ANC} = 617,000 [39,074 – 1,228,000] (**Fig. 4b**). Credible intervals around the split time and of the time of secondary contact overlapped. Median estimate of T_{split} was 617,000 generations but with large credible intervals [100,000 – 916,000]. Given the difference in life history strategy we did not converted these estimates into years. The time of secondary contact would be of 157,000 generations (**Fig. 4b**) [24,000 – 620,000].

Discussion

Our results shed light on evolutionary processes affecting the rate of speciation and provide key information for the management of these declining species. We found that genetic structure among populations was lower for the nearly semelparous species, allis shad, compared to the iteroparous species, twaite shad. Our results also suggest that for individuals captured at sea, individuals can be assigned at the region level for allis shad while at the river level in twaite shad. We inferred that the most likely long term historical divergence scenario between both species implicated historical separation followed by a secondary contact accompanied by contemporary hybridization and strong population size declines. Our results also corroborate the hypothesis of a divergent lineage (*A. agone*) present along the Mediterranean coast and suggest a possible undocumented new cryptic lineage present in the Island of Corsica, which were both formerly grouped with twaite shad. All of these results have strong conservation issues showing the importance of combining catchment and region-based management.

Level of species differentiation and hybridization

Our results confirm previous estimates of introgression and hybridization between both species (Alexandrino et al. 2006; Coscia et al. 2010; Jolly et al. 2011; Faria et al. 2012; Taillebois et al. 2020). Normally, spatio-temporal segregation mechanisms exist and maintain or minimize the contact between the two species during reproduction. In particular, twaite shad is generally spawning in lower areas of the watershed than allis shad that normally spawns in the upper parts of rivers (Aprahamian et al. 2003; Baglinière et al. 2003b), minimizing hybridization opportunities between species. Yet, human river fragmentation by dams and various obstacles is increasingly restricting shad migration to downstream areas. This favors sympatry and leads to increased opportunities for hybridization. The removal of this premating reproductive barrier thus promotes hybridization and introgression (Alexandrino et al. 2006; Taillebois et al., 2020). Accordingly, we find strong support for ongoing hybridization between both species in populations distributed in the same watershed and reaching similar spawning sites (e.g. Loire and Charente River). However, the two species seem to remain genetically well distinct despite hybridization. This may indicate either that hybridization is very recent and/or the existence of pre-zygotic (e.g. behavioral) and post-zygotic genetic barriers that are likely involved in the maintenance of the two species reproductive isolation (Ravinet et al. 2017, Barth et al. 2020). Yet, genetic barriers are often semi-permeable (Wu, 2001) resulting in heterogeneous differentiation across the genome (Ravinet et al. 2017; Rougemont et al. 2017). The porosity of the species barrier is well illustrated by the numerous backcrosses found recently by Taillebois et al. (2020) in a set of partially overlapping samples. Whole genome re-sequencing of these hybridizing populations would be needed to i) quantify the mosaic of local ancestry

bioRxiv preprint doi: https://doi.org/10.1101/2021.12.04.471201; this version posted December 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

across the genome (Duranton et al. 2018) ii) identify the regions involved in adaptive introgression from one species to the other (Hedrick 2013) and iii) determine the deleterious effect of introgression and extent of selection against introgression (Kim et al. 2018). The hybridization of these two species is hence an interesting context to study speciation and also highlights the consequences of river fragmentation on the genetic integrity of these declining species (Harris et al. 2018; Rougemont et al., 2020).

Influence of life history strategy on the population genetic structure

As expected from theory and from previous research, the more dispersive species (and semelparous), allis shad, displays, comparatively to twaite shad, lower levels of genetic structure and higher levels of genetic diversity. The twaite shad having lower contemporary dispersal, displays on the opposite a higher genetic structure and lower genetic diversity (Jolly et al. 2012; Martin et al. 2015; Taillebois et al. 2020). These variations in dispersal syndrome have been well documented and associated to variation in life history traits across many species (Stevens et al. 2013; Cayuela et al. 2016, 2018). Here, this suggests a substantial trade-off associated to these life history traits. twaite shad being iteroparous, we hypothesized that they maximize their survival probability by dispersing less, and reducing the duration of their marine phase due to an earlier sexual maturity than allis shad (Bagliniere et al. 2020). Other hypotheses have been made to explain the lower genetic structure of the semelparous American shad (*A. sapidissima*) as compared to its iteroparous form and relate to environmental differences along a latitudinal gradient (discussed in Hasselman et al. 2013). Here, however, both species reproduce in the same rivers along the Atlantic coast, so the proposed mechanisms are less likely to be comparable and to apply in our case. A meta-analysis in terrestrial and semi-terrestrial animals revealed that higher dispersal was associated with higher fecundity and survival (Stevens et al. 2014). In less stable environment, higher dispersal, higher fecundity and shorter lifespan can be favored (Cayuela et al. 2016). This leads us to hypothesize that allis shad populations may be more resilient under climate change, but this hypothesis remains to be investigated.

Demography and speciation

The differentiation levels among twaite shad lineages are similar to those observed between species. This led us to identify two putative cryptic species in addition to the already documented twaite shad located along the Atlantic coast. The first putative lineage occurs in the continental coast of the Mediterranean and the second one in Corsica. These results confirm the observations of Le Corre et al. (2005) showing genetic differences between the Atlantic and Mediterranean twaite shads using protein markers. They also suggest that the Mediterranean twaite shad may be considered as an independent species, identified as *Alosa agone* (Scopoli 1786) and considered as such in the recent atlas of freshwater fish from France (Bagliniere et al. 2020d), based on the genetic studies of Le Corre et al. (2005) and the morphological observations of (Bianco 2005). Nevertheless, it was also mentioned that this species is still very poorly documented and its naming is still debated within the scientific community (Chiesa et al. 2014; Baglinière et al. 2020c). Ideally, long-term crosses between lineages should be combined along with fitness measurement of the resulting hybrids to better evaluate the taxonomic status of these putative species. The identification of a putative lineage in Corsica is a new finding. A parallel can be drawn with *Salmo trutta* that is also represented with many phylogeographic lineages (Bernatchez 2001), and locally differentiated populations in the Tyrrhenian area, including Corsica, that are attached to the Adriatic lineage rather than the Mediterranean lineage (Berrebi et al. 2019). Unfortunately, microsatellite data are poorly suited to disentangle recent high drift (resulting in increased differentiation) from long divergence time. Again, a whole genome approach would be necessary to better reconstruct the divergence history of all the species and sub-populations or cryptic lineages using measures of divergence such as Dxy. Presently, the systematics of the genus *Alosa* remains confused because of the large number of subspecies of twaite shad (Chiesa et al. 2014; Coscia et al. 2013; Bagliniere et al. 2020c.).

In addition, we have attempted to reconstruct the demographic history of the two species. We focused broadly on the between-species divergence patterns, yet many caveats need to be pointed out. First, we only

bioRxiv preprint doi: https://doi.org/10.1101/2021.12.04.471201; this version posted December 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

had a limited number of markers and it is therefore not possible to highlight the semi-permeable nature of the genome and to discriminate local regions of the genomes that are freely exchanged from those that are possibly impermeable to gene flow and that may be involved in reproductive isolation (Duranton et al. 2018). Next, this limited number of markers may explain the large credible intervals associated to some parameters such as ancestral population size or migration rate from twaite shad into allis shad. Moreover, the mutation rate is not known and the generation time remains difficult to estimate. With all these caveats in mind, we were able to find higher support for a model of secondary contact between the two species than for alternative models, indicating that isolation was necessary to initiate divergence between the two species. Most theories suggest that speciation is difficult to establish in the presence of continuous gene flow (Barton and Bengtsson 1986; Bierne et al. 2011). Only a few studies convincingly support speciation with gene flow and have explicitly attempted to discriminate among competing models (Martin et al. 2013; Malinsky et al. 2015; Tusso et al. 2021). On the contrary, an increasing number of studies have reported evidence for divergence initiated in allopatry followed by gene flow (e.g. Roux et al. 2013, 2014; Rougemont et al. 2017; Rougemont and Bernatchez 2018; Leroy et al. 2019; Cayuela et al. 2020). Interestingly, Tine et al. (2014) and Duranton et al. (2018) have reported the existence of two cryptic species of sea bass (*Dicentrarchus labrax*), the Atlantic and Mediterranean Seabass with evidence of secondary contact (Tine et al. 2014), multiple islands of reproductive isolation (Duranton et al. 2018) and cryptic genetic and demographic connectivity (Robinet et al. 2020). Similarly, Riquet et al. (2019) reported the existence of divergent lineages of seahorses (*Hippocampus guttulatus*) on the Mediterranean Sea and the Atlantic coast as well as the existence of partially reproductively isolated cryptic lineages maintained in sympatry within the Mediterranean Sea. Thus, including additional species distributed on both the Atlantic coast and Mediterranean Sea would bring fruitful insights about the mechanisms generating diversity and maintaining divergent lineages in that area.

Finally, our results have important conservation implications. First, our estimates of long-term effective population size indicated that the *Ne* of twaite shad was higher than that of allis shad. This estimate is a long-term average based on the number of coalescent events that have occurred until all lineages coalesce. It reflects the historical population size and is in line with previous phylogeographic knowledge (Alexandrino et al. 2006; Faria et al. 2012). Despite all the incertitude associated to ancestral population size estimates, our results indicate that both twaite shad and allis shad have undergone strong reduction following their divergence. Therefore, the ongoing decline of the two species (Aprahamian et al. 2003; Baglinière et al., 2003b; Rougier et al., 2012) is of further concern if the species already have a reduced adaptive potential. These long-term trends further suggests a reduced evolutionary potential (Leroy et al. 2021). Finally, we were able to identify and assign individuals captured at sea of unknown origin to putative river or regional groups with a modest power. These results suggest that our markers set could be improved (e.g. using additional markers or by moving to SNPs array) to identify individuals dispersing at sea and draw inferences about dispersal distances of the two species, an important problematic when assessing the provenance of individuals in mixed-stocks fisheries (Beacham et al. 2019; Nachón et al. 2020).

Conclusion and future directions

Our results indicate that *Alosa* species constitute an interesting model to study speciation and hybridization between closely related species. Understanding the fitness effect of such ongoing hybridization will be off utmost importance to help manage the species. For instance, if the species are increasingly constrained to breed on overlapping spawning sites the rate of hybridization is expected to increase and this may lead to increased fitness of the hybrid during the first few generations of hybridization but may also increase selection against introgression if recessive deleterious alleles are expressed. This could negatively affect populations size, by reducing the fitness of backcrosses and hence decrease the species evolutionary potential. However, the application of the EU Water Framework Directive on good ecological status could reverse the hybridization trend of these species, if the ecological connectivity of rivers is restored and the

bioRxiv preprint doi: https://doi.org/10.1101/2021.12.04.471201; this version posted December 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

natural, separated breeding grounds of both species accessible again. Overall, our results revealed deep divergence followed by secondary contact and a prevailing role of gene flow between the two species as well as among new cryptic lineages or possible species of twaite shad. These results highlighted the consequences of contrasted life history strategies at the genetic level and suggest that the two species should be managed jointly given their porous reproductive boundaries. We propose that whole genome sequences will help address several questions raised through the discussion, especially the inference of historical divergence and demography of these species, and the inference of putative post-zygotic barriers across the genome.

Acknowledgments

We thank the many professional fisherman involved in gather samples. We thank the institutions involved in collecting samples, namely people at INRAE, OFB and FDPPMA 14, 22, 50, and Migado association. This study was funded by the European Regional Development Fund (Transnational program Interreg IV. Atlantic Aquatic Resource Conservation Project).

References

Abbott R, Albach D, Ansell S, Arntzen JW, Baird SJE, Bierne N, *et al.* (2013). Hybridization and speciation. *Journal of Evolutionary Biology* **26**: 229–246.

Alexandrino P, Faria R, Linhares D, Castro F, Corre ML, Sabatié R, *et al.* (2006). Interspecific differentiation and intraspecific substructure in two closely related clupeids with extensive hybridization, Alosa alosa and Alosa fallax. *Journal of Fish Biology* **69**: 242–259.

Anderson EC, Thompson EA (2002). A Model-Based Method for Identifying Species Hybrids Using Multilocus Genetic Data. *Genetics* **160**: 1217–1229.

Aprahamian MW, Baglinière J-LJ-L, Sabatié M-R, Alexandrino P, Theil R, Aprhamian CD (2003). Biology, Status and Conservation of the Anadromous Atlantic Twaite Shad Alosa fallax fallax. In: *Biodiversity, status and conservation of world's shads*, American Fisheries Society Symposium. American Fisheries Society.

Baglinière J, Sabatié M, Rochard E, Alexandrino P, Aprahamian M (2003a). The allis shad Alosa alosa: Biology, ecology, range, and status of populations. *undefined*.

Baglinière J-L, Sabatié MR, Rochard E, Alexandrino P, Aprahamian MW (2003b). The allis shad Alosa alosa: Biology, ecology, range, and status of populations. *American Fisheries Society Symposium* **2003**: 85–102.

Baglinière J.L., Launey S., Beaulaton L., 2020a. La grande Alose Alosa alosa, Linnaeus, 1758. In " Les poissons d'eau douce de France. 2nde édition, Biotope Editions, Mèze ; Muséum national d'Histoire naturelle, Keith P., Poulet N., Denys G., Changeux T., Feunteun E. et Persat H. J. (Coords.), Muséum national d'histoire naturelle, Paris, Collection inventaire et biodiversité, pp. 294-297.

Baglinière J.L., Launey S., Denys G., Beaulaton L., 2020b. L'Alose feinte atlantique Alosa fallax, Lacépède 1803. In " Les poissons d'eau douce de France. 2nde édition, Biotope Editions, Mèze ; Muséum national d'Histoire naturelle, Keith P., Poulet N., Denys G., Changeux T., Feunteun E. et Persat H. J. (Coords.), Muséum national d'histoire naturelle, Paris, Collection inventaire et biodiversité, pp. 298-300.

Baglinière J.L., Launey S., Denys G., Beaulaton L., 2020c. Ordre des Clupéiformes. Famille des Clupeidae. In " Les poissons d'eau douce de France. 2nde édition, Biotope Editions, Mèze ; Muséum national d'Histoire naturelle, Keith P., Poulet N., Denys G., Changeux T., Feunteun E. et Persat H. J. (Coords.), Muséum national d'histoire naturelle, Paris, Collection inventaire et biodiversité, pp. 291-292.

Barth JMI, Gubili C., Matschiner M., Toressen O., Watanabe S., Egger B., Han YS, Feunteun E., Sommaruga R., Jehle R., Schabetsberger R. (2020) Stable species boundaries despite 10 million years of hybridization in tropical eels. Nature Communications. 11(1) DOI 10.1038/s41467-020-15099-x

Barton N, Bengtsson BO (1986). The barrier to genetic exchange between hybridising populations. *Heredity* **57**: 357–376.

Barton NH, Hewitt GM (1985). Analysis of Hybrid Zones. *Annu Rev Ecol Syst* **16**: 113–148.

bioRxiv preprint doi: https://doi.org/10.1101/2021.12.04.471201; this version posted December 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Beacham TD, Wallace C, Jonsen K, McIntosh B, Candy JR, Willis D, *et al.* (2019). Variation in migration pattern, broodstock origin, and family productivity of coho salmon hatchery populations in British Columbia, Canada, derived from parentage-based tagging. *Ecology and Evolution* **9**: 9891–9906.

Belkhir K, Borsam P, Chikhi L, Raufaste N, Bonhomme F, Belkir K, *et al.* (1996). GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. *undefined*.

Bernatchez L (2001). The Evolutionary History of Brown Trout (salmo Trutta L.) Inferred from Phylogeographic, Nested Clade, and Mismatch Analyses of Mitochondrial Dna Variation. *Evolution* **55**: 351–379.

Berrebi P, Caputo Barucchi V, Splendiani A, Muracciole S, Sabatini A, Palmas F, *et al.* (2019). Brown trout (Salmo trutta L.) high genetic diversity around the Tyrrhenian Sea as revealed by nuclear and mitochondrial markers. *Hydrobiologia* **826**: 209–231.

Bianco PG (2005). The Status of the Twaite Shad, *Alosa agone*, in Italy and the Western Balkans. *Marine Ecology* **23**: 51–64.

Bierne N, Welch J, Loire E, Bonhomme F, David P (2011). The coupling hypothesis: why genome scans may fail to map local adaptation genes. *Molecular Ecology* **20**: 2044–2072.

Cavalli-Sforza LL, Edwards AW (1967). Phylogenetic analysis. Models and estimation procedures. *Am J Hum Genet* **19**: 233–257.

Cayuela H, Boualit L, Arsovski D, Bonnaire E, Pichenot J, Bellec A, *et al.* (2016). Does habitat unpredictability promote the evolution of a colonizer syndrome in amphibian metapopulations? *Ecology* **97**: 2658–2670.

Cayuela H, Rougemont Q, Laporte M, Mérot C, Normandeau E, Dorant Y, *et al.* (2020). Shared ancestral polymorphisms and chromosomal rearrangements as potential drivers of local adaptation in a marine fish. *Molecular Ecology* **29**: 2379–2398.

Cayuela H, Rougemont Q, Prunier JG, Moore J-S, Clobert J, Besnard A, *et al.* (2018). Demographic and genetic approaches to study dispersal in wild animal populations: A methodological review. *Mol Ecol* **27**: 3976–4010.

Ceballos G, Ehrlich PR, Raven PH (2020). Vertebrates on the brink as indicators of biological annihilation and the sixth mass extinction. *PNAS* **117**: 13596–13602.

Chiesa, S., Piccinini, A., Lucentini, L. et al. Genetic data on endangered twaite shad (Clupeidae) assessed in landlocked and anadromous populations: one or more species?. Rev Fish Biol Fisheries 24, 659–670 (2014). https://doi.org/10.1007/s11160-014-9352-z

Coscia I, Rountree V, King JJ, Roche WK, Mariani S (2010). A highly permeable species boundary between two anadromous fishes. *Journal of Fish Biology* **77**: 1137–1149.

Csilléry K, Blum MGB, Gaggiotti OE, François O (2010). Approximate Bayesian Computation (ABC) in practice. *Trends Ecol Evol* **25**: 410–418.

Dieringer D, Schlötterer C (2003). microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets.

Duranton M, Allal F, Fraïsse C, Bierne N, Bonhomme F, Gagnaire P-A (2018). The origin and remolding of genomic islands of differentiation in the European sea bass. *Nat Commun* **9**: 2518.

Earl DA, vonHoldt BM (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genet Resour* **4**: 359–361.

Estoup, A. C. R. L. (1996). Rapid one-tube DNA extraction for reliable PCR detection of fish polymorphic markers and transgenes. Molecular Marine Biology and Biotechnology, 5, 295-298.

Evanno G, Regnaut S, Goudet J (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* **14**: 2611–2620.

Fagundes NJR, Ray N, Beaumont M, Neuenschwander S, Salzano FM, Bonatto SL, *et al.* (2007). Statistical evaluation of alternative models of human evolution. *Proceedings of the National Academy of Sciences* **104**: 17614–17619.

bioRxiv preprint doi: https://doi.org/10.1101/2021.12.04.471201; this version posted December 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Falush D, Stephens M, Pritchard JK (2003). Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies. *Genetics* **164**: 1567–1587.

Faria R, Wallner B, Weiss S, Alexandrino P (2004). Isolation and characterization of eight dinucleotide microsatellite loci from two closely related clupeid species (Alosa alosa and A. fallax). *Molecular Ecology Notes* **4**: 586–588.

Faria R, Weiss S, Alexandrino P (2012). Comparative phylogeography and demographic history of European shads (Alosa alosa and A. fallax) inferred from mitochondrial DNA. *BMC Evolutionary Biology* **12**: 194.

Felsenstein J. (1995) PHYLIP (Phylogeny Inference Package) version 3.6. Seattle. WA: Department of Genome Sciences. University of Washington.

Frankham R (1997). Do island populations have less genetic variation than mainland populations? *Heredity* **78**: 311–327.

Frankham R (2005). Genetics and extinction. *Biological Conservation* **126**: 131–140.

García-Dorado A, Caballero A (2021). Neutral genetic diversity as a useful tool for conservation biology. *Conserv Genet* **22**: 541–545.

Goudet J (1995). FSTAT (Version 1.2): A Computer Program to Calculate F-Statistics. *J Hered* **86**: 485–486.

Harris RB, Sackman A, Jensen JD (2018). On the unfounded enthusiasm for soft selective sweeps II: Examining recent evidence from humans, flies, and viruses. *PLOS Genetics* **14**: e1007859.

Hasselman DJ, Ricard D, Bentzen P (2013). Genetic diversity and differentiation in a wide ranging anadromous fish, American shad (Alosa sapidissima), is correlated with latitude. *Mol Ecol* **22**: 1558–1573.

Hedrick PW (2013). Adaptive introgression in animals: examples and comparison to new mutation and standing variation as sources of adaptive variation. *Molecular Ecology* **22**: 4606–4618.

Hewitt GM (1996). Some genetic consequences of ice ages, and their role in divergence and speciation. *Biol J Linn Soc* **58**: 247–276.

Illera JC, Palmero AM, Laiolo P, Rodríguez F, Moreno ÁC, Navascués M (2014). Genetic, Morphological, and Acoustic Evidence Reveals Lack of Diversification in the Colonization Process in an Island Bird. *Evolution* **68**: 2259–2274.

Jolly MT, Aprahamian MW, Hawkins SJ, Henderson PA, Hillman R, O'Maoiléidigh N, *et al.* (2012). Population genetic structure of protected allis shad (Alosa alosa) and twaite shad (Alosa fallax). *Mar Biol* **159**: 675–687.

Jolly MT, Maitland PS, Genner MJ (2011). Genetic monitoring of two decades of hybridization between allis shad (Alosa alosa) and twaite shad (Alosa fallax). *Conserv Genet* **12**: 1087–1100.

Jombart T (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**: 1403–1405.

Keefer ML, Caudill CC (2014). Homing and straying by anadromous salmonids: a review of mechanisms and rates. *Rev Fish Biol Fisheries* **24**: 333–368.

Kim BY, Huber CD, Lohmueller KE (2018). Deleterious variation shapes the genomic landscape of introgression. *PLOS Genetics* **14**: e1007741.

Le Corre ML, Alexandrino P, Sabatie MR, Aprahamian MW, Baglinière JL (2005). Genetic characterisation of the Rhodanian twaite shad, Alosa fallax rhodanensis. *Fisheries Management and Ecology* **12**: 275–282.

Leroy T, Rougemont Q, Dupouey J-L, Bodénès C, Lalanne C, Belser C, *et al.* (2019). Massive postglacial gene flow between European white oaks uncovered genes underlying species barriers. *New Phytologist* **0**.

Leroy T, Rousselle M, Tilak M-K, Caizergues AE, Scornavacca C, Recuerda M, *et al.* (2021). Island songbirds as windows into evolution in small populations. *Current Biology*.

Limburg KE, Waldman JR (2009). Dramatic Declines in North Atlantic Diadromous Fishes. *BioScience* **59**: 955–965.

Malinsky M, Challis RJ, Tyers AM, Schiffels S, Terai Y, Ngatunga BP, *et al.* (2015). Genomic islands of speciation

bioRxiv preprint doi: https://doi.org/10.1101/2021.12.04.471201; this version posted December 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

separate cichlid ecomorphs in an East African crater lake. *Science* **350**: 1493–1498.

Martin SH, Dasmahapatra KK, Nadeau NJ, Salazar C, Walters JR, Simpson F, *et al.* (2013). Genome-wide evidence for speciation with gene flow in Heliconius butterflies. *Genome Res* **23**: 1817–1828.

Martin J, Bareille G, Berail S, Pecheyran C, Daverat F, Bru N, *et al.* (2013). Spatial and temporal variations in otolith chemistry and relationships with water chemistry: a useful tool to distinguish Atlantic salmon Salmo salar parr from different natal streams. *Journal of Fish Biology* **82**: 1556–1581.

Martin J, Rougemont Q, Drouineau H, Launey S, Jatteau P, Bareille G, *et al.* (2015). Dispersal capacities of anadromous Allis shad population inferred from a coupled genetic and otolith approach. *Can J Fish Aquat Sci* **72**: 991–1003.

Mennesson-Boisneau. C.. Aprahamian. M. W.. Sabatie´ . M. R. & Cassous-Leins. J. J. (2000a). Caractéristiques des adultes. In Les aloses Alosa alosa et Alosa fallax spp. (Baglinièr e. J. L. & Elie. P.. eds). pp. 33–53. Paris: INRA-Cemagref.

Mikkelsen EK, Irwin, (2021). D Ongoing production of low-fitness hybrids limits range overlap between divergent cryptic species. *Molecular Ecology* **n/a**.

Nachón DJ, Bareille G, Drouineau H, Tabouret H, Taverny C, Boisneau C, *et al.* (2020). 1980s population-specific compositions of two related anadromous shad species during the oceanic phase determined by microchemistry of archived otoliths. *Can J Fish Aquat Sci* **77**: 164–176.

Nielsen EE, Bach LA, Kotlicki P (2006). hybridlab (version 1.0): a program for generating simulated hybrids from population samples. *Molecular Ecology Notes* **6**: 971–973.

Oosterhout CV, Hutchinson WF, Wills DPM, Shipley P (2004). micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4**: 535–538.

Paetkau D, Slade R, Burden M, Estoup A (2004). Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Molecular Ecology* **13**: 55–65.

Page RDM (1996). Tree View: An application to display phylogenetic trees on personal computers. *Bioinformatics* **12**: 357–358.

Parrish DL, Behnke RJ, Gephard SR, McCormick SD, Reeves GH (1998). Why aren't there more Atlantic salmon (Salmo salar)? **55**: 7.

Perrier C, Guyomard R, Bagliniere J-L, Evanno G (2011). Determinants of hierarchical genetic structure in Atlantic salmon populations: environmental factors vs. anthropogenic influences. *Molecular Ecology* **20**: 4231–4245.

Piry S, Alapetite A, Cornuet J-M, Paetkau D, Baudouin L, Estoup A (2004). GENECLASS2: A Software for Genetic Assignment and First-Generation Migrant Detection. *Journal of Heredity* **95**: 536–539.

Pritchard JK, Stephens M, Donnelly P (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics* **155**: 945–959.

Randon, M, Daverat, F, Bareille, G, Jatteau, Ph, Martin, J, Pecheyran, C, Drouineau, H (2017). Quantifying exchanges of Allis shads between river catchments by combining otolith microchemistry and abundance indices in a Bayesian model. ICES Journal of Marine Science.

Ravinet M, Faria R, Butlin RK, Galindo J, Bierne N, Rafajlović M, *et al.* (2017). Interpreting the genomic landscape of speciation: a road map for finding barriers to gene flow. *Journal of Evolutionary Biology* **30**: 1450–1477.

Riquet F, Liautard-Haag C, Woodall L, Bouza C, Louisy P, Hamer B, *et al.* (2019). Parallel pattern of differentiation at a genomic island shared between clinal and mosaic hybrid zones in a complex of cryptic seahorse lineages. *Evolution* **73**: 817–835.

Robinet T, Roussel V, Cheze K, Gagnaire P-A (2020). Spatial gradients of introgressed ancestry reveal cryptic connectivity patterns in a high gene flow marine fish. *Molecular Ecology* **29**: 3857–3871.

bioRxiv preprint doi: https://doi.org/10.1101/2021.12.04.471201; this version posted December 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Rosenberg NA (2004). distruct: a program for the graphical display of population structure. *Molecular Ecology Notes* **4**: 137–138.

Rougemont Q, Bernatchez L (2018). The demographic history of Atlantic salmon (Salmo salar) across its distribution range reconstructed from approximate Bayesian computations*. *Evolution* **72**: 1261–1277.

Rougemont Q, Besnard A-L, Baglinière J-L, Launey S (2015). Characterization of thirteen new microsatellite markers for allis shad (Alosa alosa) and twaite shad (Alosa fallax). *Conservation Genet Resour* **7**: 259–261.

Rougemont Q, Gagnaire P-A, Perrier C, Genthon C, Besnard A-L, Launey S, *et al.* (2017). Inferring the demographic history underlying parallel genomic divergence among pairs of parasitic and nonparasitic lamprey ecotypes. *Molecular Ecology* **26**: 142–162.

Rougemont Q, Moore J-S, Leroy T, Normandeau E, Rondeau EB, Withler RE, *et al.* (2020). Demographic history shaped geographical patterns of deleterious mutation load in a broadly distributed Pacific Salmon. *PLOS Genetics* **16**: e1008348.

Rougemont Q, Roux C, Neuenschwander S, Goudet J, Launey S, Evanno G (2016). Reconstructing the demographic history of divergence between European river and brook lampreys using approximate Bayesian computations. *PeerJ* **4**.

Rougier T, Lambert P, Drouineau H, Girardin M, Castelnaud G, Carry L, *et al.* (2012). Collapse of allis shad, Alosa alosa, in the Gironde system (southwest France): environmental change, fishing mortality, or Allee effect? *ICES Journal of Marine Science* **69**: 1802–1811.

Rousset F (2008). genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources* **8**: 103–106.

Roux C, Fraïsse C, Castric V, Vekemans X, Pogson GH, Bierne N (2014). Can we continue to neglect genomic variation in introgression rates when inferring the history of speciation? A case study in a Mytilus hybrid zone. *Journal of Evolutionary Biology* **27**: 1662–1675.

Roux C, Tsagkogeorga G, Bierne N, Galtier N (2013). Crossing the Species Barrier: Genomic Hotspots of Introgression between Two Highly Divergent Ciona intestinalis Species. *Molecular Biology and Evolution* **30**: 1574–1587.

Ryan SF, Deines JM, Scriber JM, Pfrender ME, Jones SE, Emrich SJ, *et al.* (2018). Climate-mediated hybrid zone movement revealed with genomics, museum collection, and simulation modeling. *PNAS* **115**: E2284–E2291.

Saitou N, Nei M (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**: 406–425.

Semmens DJ, Diffendorfer JE, López-Hoffman L, Shapiro CD (2011). Accounting for the ecosystem services of migratory species: Quantifying migration support and spatial subsidies. *Ecological Economics* **70**: 2236–2242.

Stevens VM, Whitmee S, Galliard J-FL, Clobert J, Böhning-Gaese K, Bonte D, *et al.* (2014). A comparative analysis of dispersal syndromes in terrestrial and semi-terrestrial animals. *Ecology Letters* **17**: 1039–1052.

Taillebois L, Sabatino S, Manicki A, Daverat F, Nachón DJ, Lepais O (2020). Variable outcomes of hybridization between declining Alosa alosa and Alosa fallax. *Evolutionary Applications* **13**: 636–651.

Tavernie P & Elie E. 2001. Répartition spatio-temporelle de la grande alose (Alosa alosa, Linné, 1766) et de l'alose feinte (Alosa fallax, Lacépéde, 1803) dans le golfe de gascogne. 803-821 Knowl. Manag. Aquat. Ecosyst. 16.

Teixeira JC, Huber CD (2021). The inflated significance of neutral genetic diversity in conservation genetics. *PNAS* **118**.

Tine M, Kuhl H, Gagnaire P-A, Louro B, Desmarais E, Martins RST, *et al.* (2014). European sea bass genome and its variation provide insights into adaptation to euryhalinity and speciation. *Nature Communications* **5**: 5770.

bioRxiv preprint doi: https://doi.org/10.1101/2021.12.04.471201; this version posted December 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Tomás J, Augagneur S, Rochard E (2005). Discrimination of the natal origin of young-of-the-year Allis shad (Alosa alosa) in the Garonne–Dordogne basin (south-west France) using otolith chemistry. *Ecology of Freshwater Fish* **14**: 185–190.

Tusso S, Nieuwenhuis BPS, Weissensteiner B, Immler S, Wolf JBW (2021). Experimental evolution of adaptive divergence under varying degrees of gene flow. *Nat Ecol Evol* **5**: 338–349.

Vähä J-P, Primmer CR (2006). Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Molecular Ecology* **15**: 63–72.

Walther BD, Thorrold SR (2008). Continental-scale variation in otolith geochemistry of juvenile American shad (Alosa sapidissima). *Can J Fish Aquat Sci* **65**: 2623–2635.

Waters JM, Epifanio JM, Gunter T, Brown BL (2000). Homing behaviour facilitates subtle genetic differentiation among river populations of Alosa sapidissima: microsatellites and mtDNA. *Journal of Fish Biology* **56**: 622–636.

Weir BS, Cockerham CC (1984). Estimating F-Statistics for the Analysis of Population Structure. *Evolution* **38**: 1358–1370.

Wilcove DS, Wikelski M (2008). Going, Going, Gone: Is Animal Migration Disappearing. *PLOS Biology* **6**: e188.

Wu C-I (2001). The genic view of the process of speciation. *Journal of Evolutionary Biology* **14**: 851–865.

Data accessibility: data will be deposited on dryad and code for ABC simulations are available on github.

Table 1: Sampling information for each species. N = sample Size. Genetic diversity indices are as follows; A: number of allele. Ar: Allelic richness adjusted for sample size (8 in *A. alosa*. 15 in *A. fallax*). Ho: Observed Heterozygosity. He: Unbiased expected heterozygosity. Fis: Coefficient of inbreeding. Details of shad populations captured at sea are available in the supplementary material (Table S1)

Species	Region	Population	Site ID	N	Collection date	Lattitude (N)	Longitude (W)	A	Ar	Ho	Hs	Fis
A. alosa	**Normandy**	Vire	1	29	2010-2011	49°19'37	1°06'46	5.08	3.78	0.54	0.54	0.010
		Orne	2	8	2011	49°16'51	0°14'40	4.00	4.00	0.60	0.56	0.069
		Sélune	3	5	2010	48°38'37	1°23'39	-	-	-	-	-
	Brittany	Trieux	4	16	2009	48°47'37	3°05'45	4.62	3.93	0.57	0.57	0.002
		Aulne	5	14	2001	48°14'53	4°11'32	4.46	3.80	0.55	0.49	0.120
		Scorff	6	34	2011	47°45'36	3°21'35	6.15	4.10	0.60	0.57	0.041
		Vilaine	7	30	2011	47°30'00	2°28'22	5.77	4.28	0.62	0.62	-0.001
	Atlantic	Loire	8	36	2011-2012	47°17'22	2°06'39	7.46	4.74	0.65	0.61	0.062
		Dordogne	9	26	2011	45°27'45	0°52'48	6.69	4.73	0.67	0.62	0.065
		Garonne	10	25	2011	45°27'45	0°52'48	6.46	4.61	0.64	0.62	0.028
		Vienne	10	6	1999	47°12'42	0°04'36	-	-	-	-	-
		Charente	11	41	1998-2011	45°57'44	1°01'06	7.31	4.62	0.64	0.57	0.108
	Adour	Nivelle	12	30	2012	43°23'11	1°39'54	5.08	3.95	0.62	0.61	0.011
		Adour	13	13	2011	43°30'45	1°29'54	4.85	4.25	0.62	0.61	0.015
	Galicia	Minho	15	10	2012	41°53'27	8°53'31	4.92	4.62	0.66	0.66	-0.005
	UK	Tamar	20	4	2011	50°25'	4°12'	-	-	-	-	-
A. fallax	**Galicia**	Minho	15	24	2012	41°53'27	8°53'31	4.39	3.55	0.53	0.50	0.043
		Ulla	14	32	2011	42°25'21	8°39'27	4.46	3.03	0.43	0.44	-0.035
	Adour	Adour	13	16	2011	43°30'45	1°29'54	4.69	3.98	0.58	0.45	0.228
	Atlantic	Charente	11	6	2011	45°57'44	1°01'06	-	-	-	-	-
		Dordogne	9	15	2011	45°27'45	0°52'48	4.46	3.40	0.50	0.44	0.107
		Loire	8	7	2011-2012	47°17'22	2°06'39	-	-	-	-	-
		Vilaine	7	2	2011	47°30'00	2°28'22	-	-	-	-	-
	Normandy	Orne	2	15	2011	49°16'51	0°14'40	4.08	3.38	0.48	0.46	0.053
	Mediterranean	Rhône1	18a	30	2010	43°22'	4°50'	7.00	4.42	0.60	0.56	0.069
		Rhône2	19b	37	2010	43°22'	4°50'	7.00	4.56	0.59	0.59	0.002
		Vidourle	16	10	2010	42°35'15	3°02'	4.77	4.58	0.60	0.57	0.052
		Aude	17	15	2011	43°12'52	3°14	4.77	4.03	0.62	0.57	0.083
	Corsica	Tavignano1	19a	50	2010-2011	42°06'19	9°32'42	4.62	3.44	0.56	0.53	0.043
		Tavignano2	19b	15	2010-2011	42°06'19	9°32'42	3.92	3.46	0.53	0.54	-0.032

Table 2: pairwise F_{ST} estimates for each species (non significant values are in bold italics).

A fallax

	Minho	Ulla	Adour	Dordogne	Orne	Aude	Rhône1	Rhône2	Vidourle	Tavignano
Ulla	0.147									
Adour	0.081	0.204								
Dordogne	0.088	0.233	0.029							
Orne	0.207	0.289	0.180	0.144						
Aude	0.224	0.253	0.169	0.217	0.170					
Rhône S	0.229	0.253	0.182	0.220	0.173	***0.010***				
Rhône V	0.227	0.264	0.184	0.225	0.166	***0.011***	***0***			
Vidourl	0.245	0.289	0.177	0.240	0.217	***0.011***	***0.001***	***0.016***		
Tavignano1	0.286	0.328	0.253	0.288	0.323	0.233	0.237	0.252	0.242	
Tavignano2	0.303	0.345	0.260	0.308	0.358	0.248	0.249	0.268	0.262	***0.011***

A alosa

	Minho	Nivelle	Adour	Charente	Dordogne	Garonne	Loire	Vilaine	Scorff	Aulne	Trieux	Vire
Nivelle	0.119											
Adour	0.094	0.060										
Charent	0.058	0.093	0.032									
Dordogne	0.059	0.055	***0.013***	0.008								
Garonne	0.055	0.067	***0.013***	***0.006***	***-0.001***							
Loire	0.046	0.052	***0.032***	0.019	***0.006***	***0.009***						
Vilaine	0.073	0.056	0.029	0.032	0.018	0.022	***0.005***					
Scorff	0.071	0.075	0.035	0.051	0.029	0.037	0.024	0.021				
Aulne	0.101	0.109	0.056	0.049	0.062	0.052	0.047	0.052	***0.023***			
Trieux	0.087	0.079	0.040	0.051	0.048	0.045	0.036	0.028	***0.006***	***0.010***		
Vire	0.119	0.077	***0.051***	0.073	0.063	0.048	0.045	0.044	0.045	0.060	0.041	
Orne	0.090	0.037	***0.029***	***0.039***	***0.026***	***0.015***	***0.025***	0.026	0.031	***0.048***	***0.038***	***0.018***

Table 3: Results of assignment test of *A. alosa* and *A. fallax* captured at sea

	Sampling area	Number	Assignement Population	Structure	DAPC	Geneclass
				%	%	%
A. Alosa	**South Brittany**	36	Atlantic	28	33	36
			Brittany	11	8	11
			Normandy	0	22	0
			Minho	0	11	0
			Nivelle	3	14	3
	Pertuis Charentais	4	Atlantic	75	75	50
			Brittany	0	0	25
			Normandy	0	0	25
			Minho	0	25	0
A. fallax	**South Brittany**	16	Atlantic	27	38	20
			Orne	53	56	53
			Minho	0	13	0
	Pertuis Charentais	21	Atlantic	66	62	62
			Orne	5	19	5
			Ulla	0	5	0
	Gascogne Golfe	5	Orne	100	100	80
	North sea	6	Atlantic	0	0	20
			Orne	75	75	75
			Atlantic	0	25	25
	North Brittany	20	Orne	75	90	65