Received: 26 September 2019 | Revised: 11 December 2019 | Accepted: 16 December 2019

DOI: 10.1111/mec.15342

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Speciation history of European (*Anguilla anguilla*) and American eel (*A. rostrata*), analysed using genomic data

Natacha Nikolic[1] | Shenglin Liu[2] | Magnus W. Jacobsen[2] | Bjarni Jónsson[3] | Louis Bernatchez[4] | Pierre-Alexandre Gagnaire[5] | Michael M. Hansen[2]

[1]Agence de Recherche pour la Biodiversité à la Réunion, ARBRE, Saint-Leu, Réunion

[2]Department of Bioscience, Aarhus University, Aarhus C, Denmark

[3]North West Iceland Nature Center, Sudárkrókur, Iceland

[4]IBIS (Institut de Biologie Intégrative et des Systèmes), Université Laval, Québec, QC, Canada

[5]ISEM, CNRS, IRD, EPHE, Univ. Montpellier, Montpellier, France

Correspondence
Michael M. Hansen, Department of Bioscience, Aarhus University, Ny Munkegade 114-116, DK-8000 Aarhus C, Denmark.
Email mmh@bios.au.dk

Present address
Magnus W. Jacobsen, National Institute of Aquatic Resources, Technical University of Denmark, Silkeborg, Denmark

Funding information
EU Interreg (Øresund-Kattegat-Skagerrak) funds; Independent Research Fund Denmark, Natural Sciences, Grant/Award Number: 7014-00167B

Abstract

Speciation in the ocean could differ from terrestrial environments due to fewer barriers to gene flow. Hence, sympatric speciation might be common, with American and European eel being candidates for exemplifying this. They show disjunct continental distributions on both sides of the Atlantic, but spawn in overlapping regions of the Sargasso Sea from where juveniles are advected to North American, European and North African coasts. Hybridization and introgression are known to occur, with hybrids almost exclusively observed in Iceland. Different speciation scenarios have been suggested, involving either vicariance or sympatric ecological speciation. Using RAD sequencing and whole-genome sequencing data from parental species and F1 hybrids, we analysed speciation history based on the joint allele frequency spectrum (JAFS) and pairwise sequentially Markovian coalescent (PSMC) plots. JAFS supported a model involving a split without gene flow 150,000–160,000 generations ago, followed by secondary contact 87,000–92,000 generations ago, with 64% of the genome experiencing restricted gene flow. This supports vicariance rather than sympatric speciation, likely associated with Pleistocene glaciation cycles and ocean current changes. Whole-genome PSMC analysis of F1 hybrids from Iceland suggested divergence 200,000 generations ago and indicated subsequent gene flow rather than strict isolation. Finally, simulations showed that results from both approaches (JAFS and PSMC) were congruent. Hence, there is strong evidence against sympatric speciation in North Atlantic eels. These results reiterate the need for careful consideration of cases of possible sympatric speciation, as even in seemingly barrier-free oceanic environments palaeoceanographic factors may have promoted vicariance and allopatric speciation.

KEYWORDS

allele frequency spectrum, *Anguilla* spp, PSMC, secondary contact, sympatric speciation, vicariance

Nikolic and Liu equally contributed as first author.

1 | INTRODUCTION

Inference of patterns and processes of speciation remains one of the most important research topics in evolutionary biology, now increasingly empowered by developments in next-generation sequencing and genomics (Coyne & Orr, 2004; Seehausen et al., 2014; Sousa & Hey, 2013). Speciation is usually thought of as a multifactorial process that ultimately leads to complete reproductive isolation (Abbott et al., 2013). However, it is commonly categorized along different dimensions to emphasize the role of particular factors (Smadja & Butlin, 2011), such as the spatial context of divergence ranging from sympatric to allopatric (Bolnick & Fitzpatrick, 2007), the timing and intensity of gene flow (Butlin, Galindo, & Grahame, 2008) and the relative importance of local adaptation and genetic incompatibilities in the build-up of reproductive isolation (Bank, Burger, & Hermisson, 2012; Bierne, Welch, Loire, Bonhomme, & David, 2011; Hoffmann & Rieseberg, 2008; Rundle & Nosil, 2005; Schluter, 2009). Whereas much progress has been made in speciation genomics research, the discussion and controversy regarding how to relate genomic patterns to underlying evolutionary processes and the ubiquity and general importance of different speciation scenarios is still ongoing (Burri, 2017; Cruickshank & Hahn, 2014; Feder, Egan, & Nosil, 2012; Nosil, 2008; Ravinet et al., 2017; Seehausen et al., 2014).

Much of our current knowledge of speciation is derived from terrestrial and freshwater systems. Important steps have also been taken towards understanding speciation processes in marine environments (Bernardi, 2013; Kelley, Brown, Therkildsen, & Foote, 2016; Miglietta, Faucci, & Santini, 2011; Palumbi, 1994; Rocha, Robertson, Roman, & Bowen, 2005), but the complexity of marine environments is a challenge. Marine and especially oceanic environments are seemingly characterized by few natural barriers to gene flow, and marine animals often exhibit very high fecundity and pelagic eggs and/or larval stages, which at first sight might seem to preclude geographic divergence (Palumbi, 1994). Nevertheless, spatial heterogeneity often exists with respect to spawning sites and oceanographic and environmental conditions (Palumbi, 1994; Riginos & Liggins, 2013; Selkoe et al., 2016), and just as in terrestrial environments, historical vicariance in the marine realm could have major effects on speciation (DiBattista et al., 2013; Grant & Bowen, 1998; Hou & Li, 2018; Patarnello, Volckaert, & Castilho, 2007).

Atlantic eels of the genus *Anguilla* represent a particularly noteworthy case of speciation. They are subdivided into two sister species that spend the major part of their life cycle on opposite sides of the North Atlantic Ocean (Tesch, 2003). The American eel (*Anguilla rostrata*) is distributed from Florida to Labrador, Caribbean islands and Greenland, while the distribution of European eel (*Anguilla anguilla*) extends from Morocco to Scandinavia, Mediterranean and Iceland. Despite their large continental distribution ranges, both species are considered to be genetically panmictic (Als et al., 2011; Côté et al., 2013; Gagnaire, Normandeau, Côté, Hansen, & Bernatchez, 2012; Palm, Dannewitz, Prestegaard, & Wickstrom, 2009; Pujolar, Jacobsen, Als, Frydenberg, Munch, et al., 2014), which is a direct consequence of their exceptional migratory life cycle. Hence, once in their lives, adult eels from American and European continents undertake a long reproductive migration to spawn in thermal fronts of the Subtropical Convergence Frontal Zone in the Sargasso Sea (Aarestrup et al., 2009; Kleckner & McCleave, 1987; McCleave & Kleckner, 1987; McCleave, Kleckner, & Castonguay, 1987; Munk et al., 2010; Schmidt, 1923; Tesch, 2003). The larvae are then advected by ocean currents, notably the Gulf Stream (Schmidt, 1923; Tesch, 2003). Although they use the same surface currents for dispersal, American eels recruit to North American shores (>1,500 km from the spawning area), while European eel larvae drift for at least one more year towards European and North African habitats, >5,000 km from the spawning area. The developmental and behavioural mechanisms ensuring this spatial segregation remain unknown. However, transcriptome analysis of eel larvae of both species collected in the Sargasso Sea shows a differential timing of gene regulation in the two species, suggesting a genetic control of pelagic larval duration (Bernatchez et al., 2011). Also, a genome scan involving the two species suggests that genomic regions showing the highest divergence are enriched for genes related to developmental processes and phosphorylation, again suggesting association with differences in larval phase and migration distance (Jacobsen, Pujolar, Bernatchez, et al., 2014).

Reproduction of the two species overlaps temporally and spatially within the frontal zone (McCleave et al., 1987), leaving ample opportunities for hybridization. However, to date most newly hatched larvae sampled in the Sargasso Sea consist of pure genotypes of both species that co-occur throughout much of the Sargasso Sea, with only small proportions of different backcross categories (Jacobsen et al., 2017; Pujolar, Jacobsen, Als, Frydenberg, Magnussen, et al., 2014). No F1 hybrids have been observed in the spawning region among 552 analysed larvae (Jacobsen et al., 2017). In continental Europe and North America, a few later-generation backcrosses have been identified (one out of 225 European eel and one out of 30 American eel; Pujolar, Jacobsen, Als, Frydenberg, Munch, et al., 2014), whereas in Iceland, it is well established that both F1 hybrids and backcrosses are relatively abundant (up to 10%–15%; Albert, Jónsson, & Bernatchez, 2006; Avise et al., 1990; Gagnaire, Albert, Jónsson, & Bernatchez, 2009; Pujolar, Jacobsen, Als, Frydenberg, Magnussen, et al., 2014). Evidence for asymmetric introgression from American to European eel has been revealed using nuclear markers (Albert et al., 2006; Gagnaire et al., 2009; Pujolar, Jacobsen, Als, Frydenberg, Magnussen, et al., 2014; Wielgoss, Gilabert, Meyer, & Wirth, 2014), and genome-wide SNP analysis has shown the presence of admixed individuals among American eel that have a European ancestor tracing back to only 3–6 generations ago (Jacobsen et al., 2017). The extent to which hybridization provides an effective bridge to interspecific gene flow therefore remains unclear, and so is the degree of reproductive isolation between species.

TABLE 1 The 12 models of divergence history tested, using δaδi (Gutenkunst et al., 2009), against the JAFS between the two Atlantic eel species

	Model name	Description
Basic models	SI	Strict isolation
	IM	Isolation with migration
	AM	Ancient migration
	SC	Secondary contact
	AMSC	Ancient migration followed by strict isolation and secondary contact
Derived models	IM2M	IM with heterogeneous migration rates across genome
	AM2M	AM with heterogeneous migration rates across genome
	SC2M	SC with heterogeneous migration rates across genome
	AMSC2M	AMSC with heterogeneous migration rates across genome
	SCG	SC with exponential growth
	SC2MG	SC2M with exponential growth
	AMSC2MG	AMSC2M with exponential growth

The specific mechanisms underlying speciation of European and American eel are equally unclear. Different speciation scenarios have been proposed to account for the separation of American and European eels, including a vicariance and a dispersal scenario (Avise et al., 1990; Jacobsen, Pujolar, Gilbert, et al., 2014). The vicariance scenario, which essentially represents allopatric speciation, proposes a disjunction of the spawning area during glacial periods of the Pleistocene, possibly leading to multiple episodes of allopatric divergence and secondary contact. In the alternative scenario, which can be characterized as sympatric ecological speciation (Schluter, 2009), an ancestral species initially present in a single continent progressively colonized new habitats on the other continent, promoting the emergence of disruptive selection acting on traits related to larval duration, length of migration and assortative mating. So far, most attempts to date speciation have relied on mitochondrial DNA (Avise et al., 1990; Jacobsen, Pujolar, Gilbert, et al., 2014; Minegishi et al., 2005). However, the evolutionary history of mitogenomes does not necessarily reflect that of species (Ballard & Whitlock, 2004), and this class of markers also leaves little if any opportunity for reconstructing the past history of divergence and gene flow.

Here, we investigated genome-wide variation patterns between European (*Anguilla anguilla*) and American eels (*A. rostrata*) based on data derived from whole-genome sequencing and RAD (restriction site-associated DNA) sequencing (Hohenlohe et al., 2010) using two different demographic inference methods: one based on the joint allele frequency spectrum (JAFS) (Gutenkunst, Hernandez, Williamson, & Bustamante, 2009) and another estimating demographic history based on the genome of F1 hybrids between the two species (Li & Durbin, 2011). We particularly addressed the following questions: (a) What is the speciation history of European and American eel, in particular with respect to the temporal dynamics of gene flow during species divergence? (b) Are all genomic regions equally affected by interspecific gene flow? (c) How can Atlantic eels in general inform us about speciation processes in oceanic environments?

2 | MATERIALS AND METHODS

2.1 | RAD sequencing data: SNP filtering and JAFS construction

We used a previously published SNP data set (Jacobsen, Pujolar, Bernatchez, et al., 2014; http://datadryad.org/resource/doi:10.5061/dryad.f2313) to obtain the joint allele frequency spectrum (JAFS) between the European and the American eels. The original data set contained 328,300 SNPs identified by restriction site-associated DNA (RAD) sequencing (Hohenlohe et al., 2010) using the enzyme *Eco*RI from 28 American and 30 European eels sampled as glass eels or yellow eels from six different geographic areas between 1999 and 2010. No minor allele frequency threshold was applied by the authors, and therefore, the data set contains all frequency classes including singletons (see Jacobsen, Pujolar, Bernatchez, et al. (2014) for details on the bioinformatic pipeline for SNP detection). The archived file (in genepop format) was first converted to the VCF format to apply species-specific filters with VCFTOOLS (Danecek et al., 2011). For each species, we removed loci with fewer than 25 genotypes available. We also excluded SNPs showing significant deviation from Hardy–Weinberg equilibrium within each species using a p-value significance threshold of 0.01. Finally, we randomly selected one SNP per RAD locus in order to limit the impact of linkage disequilibrium in our analyses. A total of 9,481 SNPs retained in our filtered data set were used to construct the JAFS after downsizing the sample size to 20 individuals per species. No out-group was available to polarize the SNPs. Therefore, the JAFS was folded, that is a lower triangular matrix representing the observed counts of the minor frequency alleles and making no distinction between ancestral and derived allelic states for each SNP. Singletons were masked to minimize the impact of genotyping errors during the inferences. The size of the European eel genome is approximately 1 Gbp (Henkel et al., 2012), and linkage disequilibrium is low, showing rapid decay within ca. 1,000 bp (Jacobsen,

Pujolar, Bernatchez, et al., 2014). Hence, the 9,481 SNPs are assumed to constitute a random sample of unlinked markers that are representative of genome-wide variation patterns, although they obviously do not cover the entire genome.

2.2 | Inferring divergence history using the JAFS

In order to infer the divergence history of the two Atlantic eel species, we analysed their JAFS using δaδi v1.7.0 (Gutenkunst et al., 2009). This program, given an observed JAFS, estimates the likelihood of an assumed divergence history model for two species (or two populations) and its relevant evolutionary parameters (e.g. population size, duration of isolation and migration rate). Here, we explored 12 models capturing important demographic and selective aspects related to the divergence history for the two eel species (Table 1, Figure S1).

The simplest model of strict isolation (SI, no gene flow) is one in which an ancestral population (of effective size N_a) splits into two derived populations (of effective sizes N_{EU} and N_{AM}, corresponding to European eel and American eel, respectively) which subsequently diverge in the absence of gene flow during T_{SI} generations. We also considered four basic extensions of that model, which differ in the timing of the gene flow since their initial split from the common ancestor. In these models, migration occurs with constant but possibly asymmetrical rates between species, replacing a fraction m_{12} of population 1 (European eel) by migrants from population 2 (American eel) every generation, and m_{21} in the other direction. These four gene flow models, isolation with migration (IM, continuous gene flow during T generations), ancient migration (AM, ancient gene flow during T_{AM} followed by isolation during T_{SI} generations), secondary contact (SC, recent gene flow during T_{SC} following isolation during T_{SI} generations) and ancient migration followed by secondary contact (AMSC, ancient gene flow during T_{AM} followed by isolation during T_{SI} and recent gene flow during T_{SC} generations), are represented in Figure S1. Each of the four gene flow models was then extended to account for heterogeneous effective migration rates across the genome. These semipermeability models (called -2M models, Figure S1) capture the effect of selection against immigrant alleles by considering that the genome contains a fraction P of neutral loci (with migration rate parameters m_{12} and m_{12}) and a fraction (1-P) of loci that are affected by selection (with reduced effective migration rate parameters me_{12} and me_{21}) (Tine et al., 2014). Finally, we evaluated the need to include an exponential population size growth (-G models, Figure S1) only for the best-fit models, in order to further improve model fit. In these models, the ancestral population unequally splits into two derived populations that take a fraction s for population 1 and (1-s) for population 2 of the ancestral population size. When growth occurs (i.e. simultaneously to gene flow in SC models), population 1 effective size exponentially changes from s^*N_a to N_{EU}, and from (1-s) $^*N_a$ to N_{AM} for population 2.

Each of the 12 models was fitted to the JAFS using a modified version of the δaδi program that includes two simulated annealing (SA) procedures (one hot and one cold) before quasi-Newton (BFGS) optimization (Tine et al., 2014). For each model, we implemented 20 independent runs of optimization, and only the best fit (highest likelihood) was kept after checking for convergence. We then compared the fit of different models using the Akaike information criterion (AIC) to select the best-fit model while penalizing for the number of parameters in each model.

All estimated parameters were scaled by ancestral population size (N_a), obtained from the formula: $N_a = \theta/(4^*L^*\mu)$, where θ is a scaling factor calculated by δaδi, L is the total length of sequence used to discover SNPs, and μ is mutation rate per nucleotide per generation. Since the original data set contained 328,300 SNPs from 67,583 RAD loci of 75 bp each (Jacobsen, Pujolar, Bernatchez, et al., 2014), after accounting for the extra filtering steps as mentioned above, the L in our analyses is 146,380 bp. We assumed a standard mutation rate of 10^{-8} per site per generation for parameter conversion. This value is close to the inferred mutation rate of 1.1×10^{-8} in humans (Roach et al., 2010), but could obviously be different in eels. Parameter confidence intervals were estimated only for the best-fit model with 1,000 nonparametric bootstrapped data sets using the Godambe information matrix method in δaδi v1.7. The Godambe method provides estimates of standard deviation for all parameters, which were converted into 95% confidence intervals by removing (lower bound) or adding (upper bound) two times the standard deviation to each estimated parameter value.

2.3 | Whole-genome sequencing and mapping of reads

We selected two European eels, two American eels and two F1 hybrids between the two species for whole-genome resequencing (see Table S1 for details on the sequenced individuals). The hybrids were sampled in Iceland, and their identity as F1 hybrids was determined using diagnostic genetic markers (Pujolar, Jacobsen, Als, Frydenberg, Magnussen, et al., 2014). Library construction (using insert size ca. 500 bp) and Illumina sequencing (2* 100 bp paired-end reads; HiSeq 2500 platform) were outsourced to AROS Applied Biotechnology. We targeted a sequencing depth of ca. 20× for each individual.

The last five base pairs of each read were trimmed off with FASTX v0.0.13 (http://hannonlab.cshl.edu/fastx_toolkit/index.html) due to high "N" content. Reads were subsequently mapped to the European eel reference genome (Henkel et al., 2012) using the "mem" subfunction of BWA-0.7.12 with default parameters (Li & Durbin, 2009). The resulting BAM files were filtered for mapping quality using a MAPQ threshold of 20.

2.4 | PSMC analysis

We ran a separate pairwise sequentially Markovian coalescent (PSMC) (Li & Durbin, 2011) analysis on the genome sequence of each of the six individuals to reconstruct population size histories.

The PSMC plots of F1 hybrids were used to analyse the divergence history of two parental species as suggested in Li and Durbin (2011). Mutations accumulating independently in either parental lineage after divergence are manifested as heterozygotes in F1 hybrids and interpreted in a F1 PSMC analysis as a drastic population expansion. Hence, if two lineages split from a common ancestor and experienced no subsequent gene flow, the PSMC plots of their hybrids will "blow up" (i.e. exhibit seemingly vertical asymptotic growth of population size) at the time of their split (Li & Durbin, 2011). However, if there was gene flow between the two species after the split, the plots will vary according to the timing and strength of gene flow (Mazet, Rodriguez, Grusea, Boitard, & Chikhi, 2016).

For each individual, a "psmcfa" file was created from the BAM file following the instructions in the GitHub page of the PSMC program (https://github.com/lh3/psmc). Because the reference genome assembly of the European eel (Henkel et al., 2012) was rather fragmented (N50: 2,544, L50: 84,717) and PSMC requires long sequences of DNA to provide information concerning coalescent with recombination, we decided to use only scaffolds longer than 100 kb. We retained 1,664 out of 1,104,447 scaffolds/contigs of the assembly, which together accounted for 37.1% of the total genome size. For converting the BAM files into genomic sequences, we applied a depth filter (minimum of 7× and maximum of 40×) for each individual. All the parameters for running PSMC were set as recommended by the PSMC GitHub page except for the bin size and the iteration length. The bin size was set as 20 instead of the default value of 100 due to high heterozygosity level (the PSMC method being tuned by default to the human genome). The iteration length was set to 25. The estimates of the effective population sizes (N_e) and the time points in the past generated by PSMC were rescaled with a standard mutation rate of 10^{-8} per site per generation. We tested through simulations whether the fragmented assembly would compromise the inferential power of PSMC, and found no support for that potential limitation (Figure S2).

2.5 | Assessing conformance between JAFS- and PSMC-inferred divergence history

In order to validate the correspondence of results derived from the fundamentally different methods implemented in δaδi and PSMC, we made use of the F1 hybrids to assess the divergence history inferred above from the JAFS-based analyses. This was achieved by visually comparing the PSMC plots of simulated hybrids with those of the real F1 hybrids. Specifically, we simulated new hybrid genomes by incorporating the divergence history inferred by δaδi into the population size and gene flow history of the two parental populations using ms (Hudson, 2002). Then, we ran PSMC on the simulated hybrids. Our rationale is that a high resemblance between the PSMC plots of the simulated hybrids and the two real hybrids would provide stronger support to the divergence history inferred by δaδi. We used the models detected by δaδi as having the highest support (lowest AIC) along with other models considered in the analyses. Because the nonhybrids showed highly similar population size histories (see Section 3), we assumed the same population size history for the two species in the simulations. We calculated the geometric mean of the PSMC results of all the nonhybrids and defined that as the expected population size history of the nonhybrids.

We used two scripts, "psmc2history.pl" and "history2ms.pl", provided in the PSMC package to convert the expectation into an ms command and manually edited the command to make it simulate hybrids and to incorporate population splits, secondary contacts and migration according to the divergence model being applied. The time (T) and the migration rate estimates (m) from δaδi were converted to fit into the ms command using the following equations (with θ and L as defined previously):

$$m_{ms} = m_{\delta a \delta i} \cdot \frac{2 \cdot \theta_{ms} \cdot L_{\delta a \delta i}}{\theta_{\delta a \delta i} \cdot L_{ms}} \quad (1)$$

$$T_{ms} = T_{\delta a \delta i} \cdot \frac{\theta_{\delta a \delta i} \cdot L_{ms}}{2 \cdot \theta_{ms} \cdot L_{\delta a \delta i}} \quad (2)$$

For each divergence model, we simulated 30 hybrid genomes and ran PSMC on those. The geometric mean of the 30 PSMC results was calculated and was compared with the geometric mean of the PSMC results of the two real hybrids. We defined the latter as the expected population size history of the hybrids. The support for the different divergence models was then evaluated according to the resemblance between the expectation and the simulation results. As a safety measure for the validity of the simulations, we also simulated 30 nonhybrids for each species under each divergence model, ran PSMC on them and compared the results with the expectation for the nonhybrids.

3 | RESULTS

3.1 | Divergence history inferred from the JAFS

Among the 12 alternative models fitted to the JAFS (Table 1), the different versions of the secondary contact model (SC) outperformed other alternative models whatever the level of complexity. Overall, the best-fit model that received the lowest Akaike information criterion (AIC) was the SC2MG model (Figure 1, Table 2). The difference in AIC with the second best model (AMSC2MG) was 28, and this model also involved a secondary contact (Table 2). The SC2MG model represents a scenario whereby two species split from their common ancestor and then diverged during a period of strict isolation (for T_{SI} generations), before exchanging genes during a secondary contact period (for T_{SC} generations). During this secondary contact, the two species experienced exponential growth in population size and variation in effective migration rate across the genome to account for differential introgression among loci.

According to the SC2MG model, the two species split ca. 150,000–160,000 generations ago. After a long period of isolation

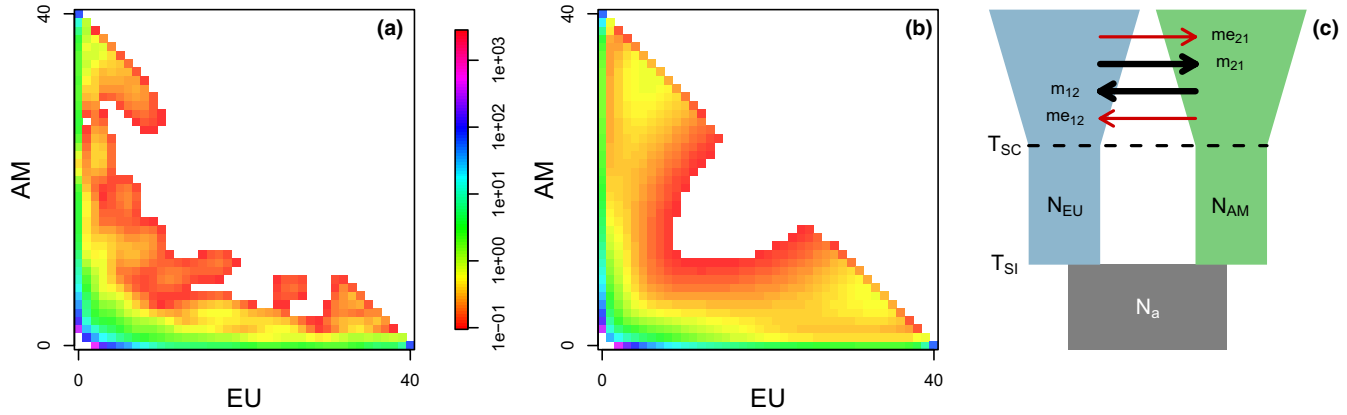

FIGURE 1 The best-fit divergence history (SC2MG) of the two Atlantic eel species according to δaδi (Gutenkunst et al., 2009) analyses. (a) The observed JAFS between the European eel (EU, *x*-axis) and the American (AM, *y*-axis) eel. The colour bar illustrates the number of SNPs in the bins of the JAFS. (b) The expected JAFS under the SC2MG model. The colours follow the same scale as (a). (c) Diagram of the divergence history represented by the SC2MG model (see Table 1 for details on parameters) [Colour figure can be viewed at wileyonlinelibrary.com]

(ca. 63,000–73,000 generations), they came into contact ca. 87,000–92,000 generations ago. The split involved European eels representing the largest fraction (0.7) of the ancestral population. The two species then underwent a demographic expansion at the onset of secondary contact. Gene flow upon secondary contact occurred in both directions, but most of the genome (ca. 64%) was inferred to have experienced very restricted effective gene flow.

The PSMC plots of the two European X American F1 hybrids showed a pattern that differed strikingly from the expected "blow-up" in the case of no gene flow after divergence (Figure 2). The plots diverged from those of the nonhybrids ca. 200,000 generations ago and exhibited a much faster population expansion, but not a "blow-up," suggesting that gene flow occurred after initial divergence.

3.2 | Whole-genome sequencing, demographic history and divergence time inferred using PSMC

The sequencing depth over the six genomes (two of each species and two F1 hybrids) averaged 19× per individual and mapping was of high quality (see Table S2 for mapping statistics). The observed mean nucleotide diversity (Pi) was 0.0072 for the two European eels, 0.0076 for the two American eels and 0.0096 for the two F1 hybrids.

PSMC analysis suggested very similar demographic histories for the two species of the Atlantic eels (Figure 2). They both experienced a gradual increase of population size for most of their histories, with American eels showing slightly higher population sizes. This was followed by a drastic expansion in the recent period (since ca. 14,000 generations ago). However, this recent expansion should be interpreted carefully because PSMC estimation tends to suffer higher stochasticity for recent history (Li & Durbin, 2011). It should be mentioned that the slightly higher population size for American eel contrasts with the result based on JAFS suggesting that European eel constitutes a major fraction of the ancestral population. Whereas we are unable to identify the exact cause of the differences, we note that the PSMC results are in accordance with the slightly higher nucleotide diversity in American relative to European eel, both for the whole-genome sequences and for the original RAD data set (Jacobsen, Pujolar, Bernatchez, et al., 2014).

3.3 | Assessing conformance between JAFS- and PSMC-inferred divergence history

δaδi detected the SC2MG as the best model (see above), but incorporation of heterogeneous migration rates across the genome was too complex to be implemented in ms (Hudson, 2002) without making assumptions on the unknown underlying genome architecture. We therefore chose to use the results from the SCG model to run the simulations. The SCG model yielded the lowest AIC value among the models with homogeneous migration rate, and it provided very similar estimations as SC2MG concerning the time of divergence and the time of secondary contact (see Table 2). In addition, we included results from the models SI and IM in the simulations.

The PSMC plots of the simulated hybrids obtained without gene flow (i.e. based on the SI model) "blew up" immediately after the split, showing no resemblance to the plots of the real hybrids (Figure 3). In contrast, hybrids simulated under the IM and particularly the SCG model showed a much better resemblance. Hence, the simulations demonstrated congruence between the results obtained using JAFS and PSMC. To test the validity of these simulations, we also compared PSMC plots based on simulated and real parental genomes of the two species. A nearly perfect resemblance was observed (Figure S3), supporting the validity of the simulations.

TABLE 2 Likelihoods and parameter estimates of the 12 models of divergence history tested using δaδi (Gutenkunst et al., 2009). The values were generated upon fitting the models to the JAFS between the two Atlantic eel species. The models are ranked in decreasing order of the AIC (Akaike information criterion) value. The last row specifies the standard deviations (*SD*) of the parameter estimates in the best model (SC2MG). Parameters' 95% confidence intervals can be obtained by removing (lower bound) or adding (upper bound) 2**SD* to each estimate for the SC2MG model. NP: number of parameters in the model. LL: log-likelihood. AIC: the AIC value. ΔAIC: the difference in AIC compared with the best model (SC2MG). N_a: ancient population size. s: proportion of the ancestral population going into the European eel after the initial split. N_{EU}: population size of the European eel. N_{AM}: population size of the American eel. *P*: proportion of the genome falling within the category of m_{12} and m_{21} in migration rates (valid only when the heterogeneous migration rate across the genome was assumed). m_{12}, m_{21}: migration rate from the American eel to the European eel and vice versa. me_{12}, me_{21}: reduced effective migration rates for loci influenced by selection when semipermeability was assumed. T_{AM}: duration of the ancient migration period. T_{SI}: duration of the strict isolation. T_{SC}: duration of the secondary contact. T: time from the initial split to present. Time estimates are provided in unit of generations

Model	NP	LL	AIC	ΔAIC	N_a	s	N_{EU}	N_{AM}	P	m_{12}	m_{21}	me_{12}	me_{21}	T_{AM}	T_{SI}	T_{SC}	T
SI	3	−1487	2,980	1771	37,252	–	744,515	34,495	–	–	–	–	–	–	86,946	–	86,946
AM	6	−1355	2,721	1512	12,295	–	114,267	43,339	–	4E-07	1E-06	–	–	0	188,650	–	188,650
IM	5	−1354	2,719	1509	10,077	–	113,356	42,737	–	3E-07	1E-06	–	–	–	–	–	195,434
AMSC	7	−1310	2,634	1,425	11,435	–	109,769	44,347	–	7E-07	2E-06	–	–	0	168,536	26,027	194,563
SC	6	−1310	2,632	1,423	12,457	–	110,124	44,560	–	8E-07	3E-06	-	–	–	166,830	24,840	191,671
IM2M	8	−1112	2,240	1,031	33,956	–	668,155	17,555	0.28	0	4E-04	0	1E-06	–	–	–	91,410
AM2M	9	−1108	2,234	1,025	114,031	–	314,955	47,665	0.95	7E-05	0	2E-08	2E-07	1,832,486	25,543	–	1,858,029
AMSC2M	10	−1075	2,171	961	32,234	–	644,517	18,019	0.29	0	4E-04	0	3E-06	0	86,903	4,900	91,802
SC2M	9	−1075	2,168	959	33,087	–	661,038	17,536	0.29	0	4E-04	2E-08	3E-06	–	86,356	5,095	91,452
SCG	7	−793	1601	391	27,018	0.83	1,349,002	386,328	–	2E-07	8E-07	–	–	–	63,060	98,831	161,891
AMSC2MG	11	−608	1,238	28	32,558	0.77	637,168	638,503	0.40	4E-06	5E-06	0	2E-07	83,545	0	89,210	172,755
SC2MG	10	−595	1,209	0	30,791	**0.69**	**1,202,060**	**1,063,746**	**0.36**	**3E-06**	**4E-06**	**5E-08**	**1E-07**	**–**	**69,119**	**89,234**	158,452
SD (SC2MG)						***0.02***	***8,776***	***3,818***	***0.08***	***4E-06***	***3E-06***	***9E-06***	***3E-06***	**–**	***5,850***	***2,402***	

4 | DISCUSSION

Using two independent methods and data sets for analysing the demographic history of Atlantic eels, we obtained new insights into the speciation history of these remarkable migratory species. Our analysis of the JAFS based on RADseq data found models involving secondary contact to be the most likely. Also, the PSMC analysis of F1 hybrid genomes showed that the dynamics of speciation deviated significantly from a simple split model and involved recent gene flow during the divergence history. By simulating genome-wide sequence data under the best-fit model derived from the JAFS analysis, we found that the PSMC plots of simulated F1 hybrids closely resembled those obtained from real empirical genomes. Thus, the congruence of the results obtained under the two complementary approaches and data sets further supports the secondary contact scenario.

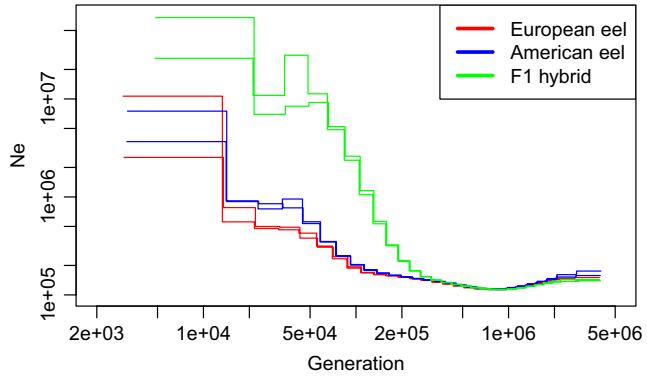

FIGURE 2 Population size history inferred by PSMC (Li & Durbin, 2011) using the six whole-genome sequenced individuals, including two European eels, two American eels and two F1 hybrids of the two species [Colour figure can be viewed at wileyonlinelibrary.com]

4.1 | The demographic divergence history of Atlantic eels

Our results showing isolation followed by secondary contact between the two species reject the hypothesis of sympatric speciation, and also the "Dispersal" scenario previously suggested by Avise et al. (1990). This comes with the small caveat that the second most supported model (AMSC2MG) includes both ancient migration and secondary contact, the former of which could indicate an initial phase of sympatric divergence followed by vicariance if this scenario is in fact correct. However, the difference in AIC between the first (SC2MG) and the second best model (AMSC2MG) was well above 10, a value considered as a threshold to reject the second most supported model (Burnham & Anderson, 2002). In contrast, the "Vicariance" scenario also proposed by Avise et al. (1990) is compatible with the results of our study. Assuming a mutation rate of 10^{-8} per site per generation, the JAFS analyses suggested that divergence has started around 160,000 (±15,000) generations ago, an estimate not far from the ca. 200,000 generations suggested by the PSMC analysis. Generation length is notoriously difficult to estimate in Atlantic eels, with proposed estimates ranging from 8.7 (Vollestad, 1992) to 15 years (https://www.iucnredlist.org/species/60344/45833138) for the European eel. Based on these extreme values and our divergence time estimates obtained from the JAFS analysis, the beginning of divergence may have occurred between 1.3 and 2.4 mya during the early Pleistocene. The time of secondary contact similarly inferred by our JAFS analysis was estimated to 87,000 – 92,000 generations ago, corresponding to between ca. 0.8 and 1.4 mya.

The dependence of Atlantic eels on the Subtropical Convergence Zone in the Sargasso Sea for spawning and the Gulf Stream and other North Atlantic currents for dispersal of larvae (Kleckner & McCleave, 1987; Schmidt, 1923; Tesch, 2003) points towards

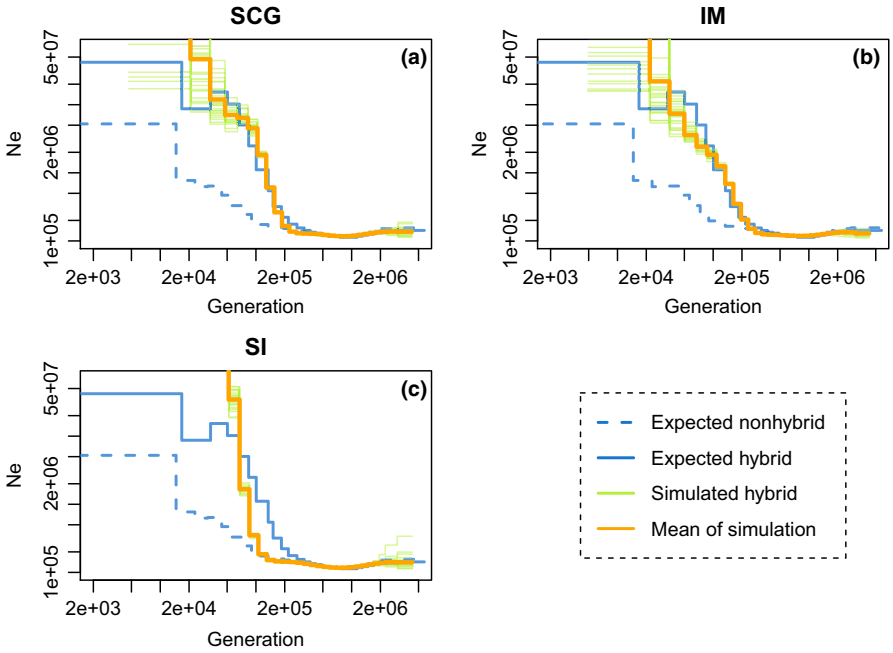

FIGURE 3 PSMC (Li & Durbin, 2011) results of the simulated hybrids using the demographic history parameters inferred by δaδi (Gutenkunst et al., 2009) under three different models, that is SCG, IM and SI [Colour figure can be viewed at wileyonlinelibrary.com]

Pleistocene ocean current changes as a probable factor causing both vicariance and secondary contact. Although divergence must have taken place several glaciation cycles back in time, data covering the Last (Weichselian) Glaciation may illustrate the dynamics. During the Last Glacial Maximum ca. 21 kya, subfossil data of European eel suggest that the continental distribution of the species was restricted to south-western Europe, and the Gulf Stream was located further south compared to present time (Kettle, Heinrich, Barrett, Benecke, & Locker, 2008). Hence, the core areas of spawning for the two species may have been further displaced compared to the present situation and population sizes of both species are expected to have been smaller, leading to a reduced overlap between the spawning areas of the two species in this region. Moreover, palaeoceanographic data have revealed intermittent massive discharges of icebergs, presumably from the Laurentide Ice Sheet in North America, into the North Atlantic (Broecker, 1994, 2003; Heinrich, 1988). These so-called Heinrich events have occurred during several glaciations spanning the last ca. 640,000 years (Hodell, Channell, Curtis, Romero, & Rohl, 2008). The resulting melting of freshwater is assumed to have significantly changed and weakened oceanic circulation, including the Gulf Stream (Broecker, 1994, 2003), hence potentially leading to vicariance in Atlantic eels.

If ocean current changes during Pleistocene glaciations indeed underlie vicariance in Atlantic eels, then this would imply that multiple episodes could have taken place involving divergence and partial reproductive isolation followed by secondary contact. In this context, it should be noted that our models that assume a single episode of isolation and secondary contact cannot detect such cyclic connectivity, if it has occurred. Hence, testing scenarios of multiple events of isolation and secondary contact against the more simple scenario involving a single episode of vicariance and secondary contact would be difficult with the data and methods at hand.

Interestingly, the PSMC plots of the parental species (Figure 2) provide little evidence for reduced N_e during glaciations and expanding populations during interglacial periods, whereas both microsatellite and mitogenome data have suggested population declines during the Weichselian Glaciation and subsequent population expansion (Jacobsen, Pujolar, Gilbert, et al., 2014; Wirth & Bernatchez, 2003). The higher mutation rate at mtDNA and microsatellites and the statistical methods used in these studies make them more suitable for exploring the more recent demographic history, whereas PSMC has limitations for capturing the most recent demographic events (Li & Durbin, 2011). Moreover, PSMC tends to smooth demographic history (Li & Durbin, 2011). Possible population expansions further back in time during interglacial and interstadial periods typically encompassing <20,000 years may therefore be difficult to detect. This is likely to underlie the different results obtained by different methods.

Previous attempts to date divergence between the two species have relied on mitochondrial sequence data (Avise, Helfman, Saunders, & Hales, 1986; Avise et al., 1990; Jacobsen, Pujolar, Gilbert, et al., 2014; Minegishi et al., 2005). The most extensive of these studies based on whole mitogenome sequencing yielded a divergence time estimate of 3.4 mya (Jacobsen, Pujolar, Gilbert, et al., 2014), which significantly predates the estimates obtained in the present study. The differences between these estimates could be explained by uncertainties of the mutation rates assumed in both approaches. However, it also raises the intriguing possibility that initial divergence has occurred earlier at the mitogenome level due to mitonuclear co-evolution, a process which could have started before the species split. Evidence for differential selection between the two species at the mitochondrial ATP6 gene has previously been presented (Gagnaire, Normandeau, & Bernatchez, 2012; Jacobsen, Pujolar, Gilbert, et al., 2014), and this furthermore involves co-adaptation with functional variation at the nuclear interactor atp5c1. ATP6 is involved in the oxidative phosphorylation pathway and thereby energy production. It is therefore a strong candidate for being under selection associated with the different lengths of migration routes from either continents to the Sargasso Sea (1,500–3,000 km in American eel and 5,000–6,000 km in European eel). Under this scenario, the mitonuclear co-evolution would have been one of the first genetic barriers to evolve before the accumulation of additional barriers genome-wide during the allopatric divergence period.

4.2 | Differential gene flow across the genome

Our results clearly suggest that gene flow between the two Atlantic eel species is still ongoing, as also supported by the detection of admixed individuals in both species dating back to several generations post F1 (Albert et al., 2006; Jacobsen et al., 2017; Pujolar, Jacobsen, Als, Frydenberg, Magnussen, et al., 2014; Pujolar, Jacobsen, Als, Frydenberg, Munch, et al., 2014). Interspecific gene flow resulting from admixture occurs at variable rates across the genome, indicating the presence of genomic regions that are resistant to introgression (Gagnaire et al., 2009). Interestingly, numerous regions of increased divergence have been found that are enriched in genes involved in phosphorylation and development, which may therefore play a role in reproductive isolation (Jacobsen, Pujolar, Bernatchez, et al., 2014). Nevertheless, disentangling the different factors underlying genome-wide variation in between-species divergence remains challenging, due to the confounding effects of linked selection and barrier loci on genome divergence (Cruickshank & Hahn, 2014; Ravinet et al., 2017). Here, accounting for heterogeneous introgression rates among loci clearly improved the fit of the secondary contact model, thus supporting the view that gene flow occurs at variable rates across the genome. Overall, our analyses based on the JAFS indicated that a rather large fraction of the genome (i.e. 64%) may be experiencing restricted gene flow. Thus, accounting for heterogeneous gene flow might be especially important in this case, to better integrate the information of contrasted gene flow histories among loci for the inference of divergence parameters. This finding moreover indicates that the two species lie at the end of the speciation continuum (Roux et al.,

2016) and only remain compatible in a relatively small fraction of their genome.

4.3 | Speciation in oceanic environments

The existence of sympatric speciation and its relative importance compared to allo- and parapatric speciation has been discussed for decades (Coyne & Orr, 2004; Via, 2001). It has been pointed out that a strict classification is counter-productive as most empirical cases represent a continuum between the extreme modes of speciation. Nonetheless, the seemingly continuous nature of marine and particularly oceanic environments should present ample opportunities for identifying cases of "pure" sympatric speciation not involving preliminary phases of allopatry, and it has even been suggested that sympatric speciation could be more important than allopatric speciation in pelagic species (Norris, 2000). This could particularly be the case for organisms with a highly dispersive planktonic life stage and huge population sizes, where some cases of cryptic speciation and sympatry of sister species have been ascribed to sympatric speciation [e.g. in pelagic nudibranches *Glaucus* spp. (Churchill, Alejandrino, Valdes, & Foighil, 2013) and planktonic foraminifera (Seears, Darling, & Wade, 2012)], albeit also noting that allopatric speciation cannot be ruled out entirely. In marine fishes, sympatric speciation has been suggested in cases such as angelfishes (*Holacanthus* spp.; Tariel, Longo, & Bernardi, 2016) and reef fishes of the genus *Hexagrammos* (Crow, Munehara, & Bernardi, 2010). In contrast, studies of divergent lineages within the European sea bass *Dicentrarchus labrax* (Tine et al., 2014) and European anchovy *Engraulis encrasicolus* (Le Moan, Gagnaire, & Bonhomme, 2016) employing genomic data and analysis of JAFS have found strong support for scenarios involving isolation followed by secondary contact implying allopatric divergence. Indeed, many cases of divergence and incipient speciation in marine fishes have been found to be directly associated with oceanographic or benthic barriers (Catarino et al., 2015; Johannesson & Andre, 2006; Nielsen, Hansen, Ruzzante, Meldrup, & Grønkjær, 2003; Patarnello et al., 2007).

Given the pronounced overlap of spawning region of the two Atlantic eel species and the seeming continuity of the ca. 2,000 km long frontal system in which spawning takes place (Kleckner & McCleave, 1987; Miller et al., 2019; Munk et al., 2010; Schmidt, 1923), this would seem an ideal candidate for sympatric speciation to have occurred, as also implied in one of the scenarios originally suggested by Avise et al. (1990). Our study nevertheless refutes the possibility of strict sympatric speciation and rather points towards allopatric speciation followed by secondary contact. These results thus reiterate the need for careful consideration of cases of possible sympatric speciation, as even in seemingly barrier-free oceanic environments, palaeoceanographic factors may have promoted vicariance and allopatric speciation. Fortunately, developments in speciation genomics and associated statistical and bioinformatics methodology keep accelerating (Ravinet et al., 2017; Seehausen et al., 2014). Among others, this increases the possibilities for a more rigorous assessment of the allopatric divergence hypothesis, thus addressing one of the key requirements for demonstrating sympatric speciation (Coyne & Orr, 2004). Hence, whereas the jury is still out, we expect that future studies will shed much more light on the general significance of sympatric speciation in marine environments.

ACKNOWLEDGEMENTS

We thank Annie Brandstrup for technical assistance, Michael Glad for the maintenance of computers and Shawn Narum and three anonymous reviewers for insightful and constructive comments on the manuscript. This work was financially supported by the Independent Research Fund Denmark | Natural Sciences (grant no. 7014-00167B to MMH) and EU Interreg (Øresund-Kattegat-Skagerrak) funds (MarGenII).

AUTHORS' CONTRIBUTIONS

N.N., S.L., M.W.J., L.B., P.A.G. and M.M.H. conceived the study. N.N. and P.A.G. conducted the analysis of JAFS based on RAD. data, whereas S.L. conducted all analyses of whole-genome sequence data along with simulations. BJ provided hybrid eels from Iceland. N.N., S.L., P.A.G. and M.M.H. wrote the paper with contributions from M.W.J., L.B. and B.J. All authors read and approved the final version of the paper.

DATA AVAILABILITY STATEMENT

The genome sequences of the two American, two European and two F1 hybrid eels are deposited in the NCBI Sequence Read Archive under project number PRJNA554219. The Genepop and VCF files used for the JAFS analysis are available through DRYAD (https://doi.org/10.5061/dryad.s1rn8pk40) (Nikolic et al., 2019). These latter data were generated from RAD data in Jacobsen, Pujolar, Bernatchez, et al. (2014), and sequence reads have been deposited in the NCBI Sequence Read Archive under project number PRJNA195555 (European eel) and PRJNA230782 (American eel). Finally, the commands for the simulations of pure and hybrid genomes using ms (Hudson, 2002) are also available through DRYAD (Nikolic et al., 2019).

ORCID

Michael M. Hansen https://orcid.org/0000-0001-5372-4828

REFERENCES

Aarestrup, K., Okland, F., Hansen, M. M., Righton, D., Gargan, P., Castonguay, M., … McKinley, R. S. (2009). Oceanic spawning migration of the European eel (*Anguilla anguilla*). *Science*, *325*, 1660–1660. https://doi.org/10.1126/science.1178120

Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, N., … Zinner, D. (2013). Hybridization and speciation. *Journal of Evolutionary Biology*, *26*, 229–246. https://doi.org/10.1111/j.1420-9101.2012.02599.x

Albert, V., Jónsson, B., & Bernatchez, L. (2006). Natural hybrids in Atlantic eels (*Anguilla anguilla*, *A. rostrata*): Evidence for successful reproduction and fluctuating abundance in space and time. *Molecular Ecology*, *15*, 1903–1916. https://doi.org/10.1111/j.1365-294X.2006.02917.x

Als, T. D., Hansen, M. M., Maes, G. E., Castonguay, M., Riemann, L., Aarestrup, K., … Bernatchez, L. (2011). All roads lead to home: Panmixia of European eel in the Sargasso Sea. *Molecular Ecology*, *20*, 1333–1346. https://doi.org/10.1111/j.1365-294X.2011.05011.x

Avise, J. C., Helfman, G. S., Saunders, N. C., & Hales, L. S. (1986). Mitochondrial DNA differentiation in North Atlantic eels: Population genetic consequences of an unusual life-history pattern. *Proceedings of the National Academy of Sciences of the United States of America*, *83*, 4350–4354. https://doi.org/10.1073/pnas.83.12.4350

Avise, J. C., Nelson, W. S., Arnold, J., Koehn, R. K., Williams, G. C., & Thorsteinsson, V. (1990). The evolutionary genetic status of Icelandic eels. *Evolution*, *44*, 1254–1262. https://doi.org/10.1111/j.1558-5646.1990.tb05229.x

Ballard, J. W. O., & Whitlock, M. C. (2004). The incomplete natural history of mitochondria. *Molecular Ecology*, *13*, 729–744. https://doi.org/10.1046/j.1365-294X.2003.02063.x

Bank, C., Burger, R., & Hermisson, J. (2012). The limits to parapatric speciation: Dobzhansky-Muller incompatibilities in a continent-island model. *Genetics*, *191*, 845–U345. https://doi.org/10.1534/genetics.111.137513

Bernardi, G. (2013). Speciation in fishes. *Molecular Ecology*, *22*, 5487–5502. https://doi.org/10.1111/mec.12494

Bernatchez, L., St-Cyr, J., Normandeau, E., Maes, G. E., Als, T. D., Kalujnaia, S., … Hansen, M. M. (2011). Differential timing of gene expression regulation between leptocephali of the two *Anguilla* eel species in the Sargasso Sea. *Ecology and Evolution*, *1*, 459–467.

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. https://doi.org/10.1111/j.1365-294X.2011.05080.x

Bolnick, D. I., & Fitzpatrick, B. M. (2007). Sympatric speciation: Models and empirical evidence. *Annual Review of Ecology Evolution and Systematics*, *38*, 459–487. https://doi.org/10.1146/annurev.ecolsys.38.091206.095804

Broecker, W. S. (1994). Massive iceberg discharges as triggers for global climate change. *Nature*, *372*, 421–424. https://doi.org/10.1038/372421a0

Broecker, W. S. (2003). Does the trigger for abrupt climate change reside in the ocean or in the atmosphere? *Science*, *300*, 1519–1522. https://doi.org/10.1126/science.1083797

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach*. London, UK: Springer-Verlag.

Burri, R. (2017). Interpreting differentiation landscapes in the light of long-term linked selection. *Evolution Letters*, *1*, 118–131. https://doi.org/10.1002/evl3.14

Butlin, R. K., Galindo, J., & Grahame, J. W. (2008). Sympatric, parapatric or allopatric: The most important way to classify speciation? *Philosophical Transactions of the Royal Society B-Biological Sciences*, *363*, 2997–3007. https://doi.org/10.1098/rstb.2008.0076

Catarino, D., Knutsen, H., Veríssimo, A., Olsen, E. M., Jorde, P. E., Menezes, G., … Stefanni, S. (2015). The Pillars of Hercules as a bathymetric barrier to gene flow promoting isolation in a global deep-sea shark (*Centroscymnus coelolepis*). *Molecular Ecology*, *24*, 6061–6079.

Churchill, C. K. C., Alejandrino, A., Valdes, A., & Foighil, D. O. (2013). Parallel changes in genital morphology delineate cryptic diversification of planktonic nudibranchs. *Proceedings of the Royal Society B-Biological Sciences*, *280*, 20131224.

Côté, C. L., Gagnaire, P.-A., Bourret, V., Verreault, G., Castonguay, M., & Bernatchez, L. (2013). Population genetics of the American eel (*Anguilla rostrata*): FST = 0 and North Atlantic Oscillation effects on demographic fluctuations of a panmictic species. *Molecular Ecology*, *22*, 1763–1776.

Coyne, J. A., & Orr, H. A. (2004). *Speciation*. Sunderland, MA: Sinnauer Associates.

Crow, K. D., Munehara, H., & Bernardi, G. (2010). Sympatric speciation in a genus of marine reef fishes. *Molecular Ecology*, *19*, 2089–2105. https://doi.org/10.1111/j.1365-294X.2010.04611.x

Cruickshank, T. E., & Hahn, M. W. (2014). Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology*, *23*, 3133–3157. https://doi.org/10.1111/mec.12796

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., … Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*, 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

DiBattista, J. D., Berumen, M. L., Gaither, M. R., Rocha, L. A., Eble, J. A., Choat, J. H., … Bowen, B. W. (2013). After continents divide: Comparative phylogeography of reef fishes from the Red Sea and Indian Ocean. *Journal of Biogeography*, *40*, 1170–1181. https://doi.org/10.1111/jbi.12068

Feder, J. L., Egan, S. P., & Nosil, P. (2012). The genomics of speciation-with-gene-flow. *Trends in Genetics*, *28*, 342–350. https://doi.org/10.1016/j.tig.2012.03.009

Gagnaire, P. A., Albert, V., Jónsson, B., & Bernatchez, L. (2009). Natural selection influences AFLP intraspecific genetic variability and introgression patterns in Atlantic eels. *Molecular Ecology*, *18*, 1678–1691. https://doi.org/10.1111/j.1365-294X.2009.04142.x

Gagnaire, P. A., Normandeau, E., & Bernatchez, L. (2012). Comparative genomics reveals adaptive protein evolution and a possible cytonuclear incompatibility between European and American eels. *Molecular Biology and Evolution*, *29*, 2909–2919. https://doi.org/10.1093/molbev/mss076

Gagnaire, P. A., Normandeau, E., Côté, C., Hansen, M. M., & Bernatchez, L. (2012). The genetic consequences of spatially varying selection in the panmictic American eel (*Anguilla rostrata*). *Genetics*, *190*, 725–733.

Grant, W. S., & Bowen, B. W. (1998). Shallow population histories in deep evolutionary lineages of marine fishes: Insights from sardines and anchovies and lessons for conservation. *Journal of Heredity*, *89*, 415–426. https://doi.org/10.1093/jhered/89.5.415

Gutenkunst, R. N., Hernandez, R. D., Williamson, S. H., & Bustamante, C. D. (2009). Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *Plos Genetics*, *5*, e1000695. https://doi.org/10.1371/journal.pgen.1000695

Heinrich, H. (1988). Origin and consequences of cyclic ice rafting in the Northeast Atlantic-Ocean during the past 130,000 years. *Quaternary Research*, *29*, 142–152. https://doi.org/10.1016/0033-5894(88)90057-9

Henkel, C. V., Burgerhout, E., de Wijze, D. L., Dirks, R. P., Minegishi, Y., Jansen, H. J., … van den Thillart, G. E. E. J. M. (2012). Primitive duplicate Hox clusters in the European eel's genome. *PLoS ONE*, *7*, e32231. https://doi.org/10.1371/journal.pone.0032231

Hodell, D. A., Channell, J. E. T., Curtis, J. H., Romero, O. E., & Rohl, U. (2008). Onset of "Hudson Strait" Heinrich events in the eastern North Atlantic at the end of the middle Pleistocene transition (similar to 640 ka)? *Paleoceanography*, *23*, PA4218.

Hoffmann, A. A., & Rieseberg, L. H. (2008). Revisiting the impact of inversions in evolution: From population genetic markers to drivers of adaptive shifts and speciation? *Annual Review of Ecology Evolution and Systematics*, *39*, 21–42. https://doi.org/10.1146/annurev.ecolsys.39.110707.173532

Hohenlohe, P. A., Bassham, S., Etter, P. D., Stiffler, N., Johnson, E. A., & Cresko, W. A. (2010). Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *Plos Genetics*, *6*, 23. https://doi.org/10.1371/journal.pgen.1000862

Hou, Z. G., & Li, S. Q. (2018). Tethyan changes shaped aquatic diversification. *Biological Reviews*, *93*, 874–896. https://doi.org/10.1111/brv.12376

Hudson, R. R. (2002). Generating samples under a Wright-Fisher neutral model of genetic variation. *Bioinformatics*, *18*, 337–338. https://doi.org/10.1093/bioinformatics/18.2.337

Jacobsen, M. W., Pujolar, J. M., Bernatchez, L., Munch, K., Jian, J., Niu, Y., & Hansen, M. M. (2014). Genomic footprints of speciation in Atlantic eels (*Anguilla anguilla* and *A. rostrata*). *Molecular Ecology*, *23*, 4785–4798.

Jacobsen, M. W., Pujolar, J. M., Gilbert, M. T. P., Moreno-Mayar, J. V., Bernatchez, L., Als, T. D., … Hansen, M. M. (2014). Speciation and demographic history of Atlantic eels (*Anguilla anguilla* and *A. rostrata*) revealed by mitogenome sequencing. *Heredity*, *112*, 432–442. https://doi.org/10.1038/hdy.2014.44

Jacobsen, M. W., Smedegaard, L., Sørensen, S. R., Pujolar, J. M., Munk, P., Jónsson, B., … Hansen, M. M. (2017). Assessing pre- and post-zygotic barriers between North Atlantic eels (*Anguilla anguilla* and *A. rostrata*). *Heredity*, *118*, 266–275. https://doi.org/10.1038/hdy.2016.96

Johannesson, K., & Andre, C. (2006). Life on the margin: Genetic isolation and diversity loss in a peripheral marine ecosystem, the Baltic Sea. *Molecular Ecology*, *15*, 2013–2029.

Kelley, J. L., Brown, A. P., Therkildsen, N. O., & Foote, A. D. (2016). The life aquatic: Advances in marine vertebrate genomics. *Nature Reviews Genetics*, *17*, 523–534. https://doi.org/10.1038/nrg.2016.66

Kettle, A. J., Heinrich, D., Barrett, J. H., Benecke, N., & Locker, A. (2008). Past distributions of the European freshwater eel from archaeological and palaeontological evidence. *Quaternary Science Reviews*, *27*, 1309–1334. https://doi.org/10.1016/j.quascirev.2008.03.005

Kleckner, R. C., & McCleave, J. D. (1987). The northern limit of spawning by Atlantic eels (*Anguilla spp.*) in the Sargasso Sea in relation to thermal fronts and surface water masses. *Journal of Marine Research*, *46*, 647–667.

Le Moan, A., Gagnaire, P. A., & Bonhomme, F. (2016). Parallel genetic divergence among coastal-marine ecotype pairs of European anchovy explained by differential introgression after secondary contact. *Molecular Ecology*, *25*, 3187–3202. https://doi.org/10.1111/mec.13627

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., & Durbin, R. (2011). Inference of human population history from individual whole-genome sequences. *Nature*, *475*, 493–U484. https://doi.org/10.1038/nature10231

Mazet, O., Rodriguez, W., Grusea, S., Boitard, S., & Chikhi, L. (2016). On the importance of being structured: Instantaneous coalescence rates and human evolution—lessons for ancestral population size inference? *Heredity*, *116*, 362–371. https://doi.org/10.1038/hdy.2015.104

McCleave, J. D., & Kleckner, R. C. (1987). Distribution of the leptocephali of the catadromous *Anguilla* species in the western Sargasso Sea in relation to water circulation and migration. *Bulletin of Marine Science*, *41*, 789–806.

McCleave, J. D., Kleckner, R. C., & Castonguay, M. (1987). Reproductive sympatry of American and European eels and implications for migration and taxonomy. *American Fisheries Society Symposium*, *1*, 286–297.

Miglietta, M. P., Faucci, A., & Santini, F. (2011). Speciation in the sea: Overview of the symposium and discussion of future directions. *Integrative and Comparative Biology*, *51*, 449–455. https://doi.org/10.1093/icb/icr024

Miller, M. J., Westerberg, H., Sparholt, H., Wysujack, K., Sørensen, S. R., Marohn, L., … Hanel, R. (2019). Spawning by the European eel across 2000 km of the Sargasso Sea. *Biology Letters*, *15*, 20180835. https://doi.org/10.1098/rsbl.2018.0835

Minegishi, Y., Aoyama, J., Inoue, J. G., Miya, M., Nishida, M., & Tsukamoto, K. (2005). Molecular phylogeny and evolution of the freshwater eels genus *Anguilla* based on the whole mitochondrial genome sequences. *Molecular Phylogenetics and Evolution*, *34*, 134–146. https://doi.org/10.1016/j.ympev.2004.09.003

Munk, P., Hansen, M. M., Maes, G. E., Nielsen, T. G., Castonguay, M., Riemann, L., … Bachler, M. (2010). Oceanic fronts in the Sargasso Sea control the early life and drift of Atlantic eels. *Proceedings of the Royal Society Series B: Biological Sciences*, *277*, 3593–3599.

Nielsen, E. E., Hansen, M. M., Ruzzante, D. E., Meldrup, D., & Grønkjær, P. (2003). Evidence of a hybrid-zone in Atlantic cod (*Gadus morhua*) in the Baltic and the Danish Belt Sea revealed by individual admixture analysis. *Molecular Ecology*, *12*, 1497–1508. https://doi.org/10.1046/j.1365-294X.2003.01819.x

Nikolic, N., Liu, S., Jacobsen, M. W., Jónsson, B., Bernatchez, L., Gagnaire, P., & Hansen, M. M. (2019). Dataset for: Speciation history of European (*Anguilla anguilla*) and American eel (*A. rostrata*), analyzed using genomic data. *DRYAD*, https://doi.org/10.5061/dryad.s5061rn5068pk5040

Norris, R. D. (2000). Pelagic species diversity, biogeography, and evolution. *Paleobiology*, *26*, 236–258. https://doi.org/10.1017/S0094837300026956

Nosil, P. (2008). Speciation with gene flow could be common. *Molecular Ecology*, *17*, 2103–2106. https://doi.org/10.1111/j.1365-294X.2008.03715.x

Palm, S., Dannewitz, J., Prestegaard, T., & Wickstrom, H. (2009). Panmixia in European eel revisited: No genetic difference between maturing adults from southern and northern Europe. *Heredity*, *103*, 82–89. https://doi.org/10.1038/hdy.2009.51

Palumbi, S. R. (1994). Genetic divergence, reproductive isolation, and marine speciation. *Annual Review of Ecology and Systematics*, *25*, 547–572. https://doi.org/10.1146/annurev.es.25.110194.002555

Patarnello, T., Volckaert, F. A. M. J., & Castilho, R. (2007). Pillars of Hercules: Is the Atlantic-Mediterranean transition a phylogeographical break? *Molecular Ecology*, *16*, 4426–4444. https://doi.org/10.1111/j.1365-294X.2007.03477.x

Pujolar, J. M., Jacobsen, M. W., Als, T. D., Frydenberg, J., Magnussen, E., Jónsson, B., … Hansen, M. M. (2014). Assessing patterns of hybridization between North Atlantic eels using diagnostic single-nucleotide polymorphisms. *Heredity*, *112*, 627–637. https://doi.org/10.1038/hdy.2013.145

Pujolar, J. M., Jacobsen, M. W., Als, T. D., Frydenberg, J., Munch, K., Jónsson, B., … Hansen, M. M. (2014). Genome-wide single-generation signatures of local selection in the panmictic European eel. *Molecular Ecology*, *23*, 2514–2528. https://doi.org/10.1111/mec.12753

Ravinet, M., Faria, R., Butlin, R. K., Galindo, J., Bierne, N., Rafajlović, M., … Westram, A. M. (2017). Interpreting the genomic landscape of speciation: A road map for finding barriers to gene flow. *Journal of Evolutionary Biology*, *30*, 1450–1477. https://doi.org/10.1111/jeb.13047

Riginos, C., & Liggins, L. (2013). Seascape genetics: Populations, individuals, and genes marooned and adrift. *Geography Compass*, *7*, 197–216. https://doi.org/10.1111/gec3.12032

Roach, J. C., Glusman, G., Smit, A. F. A., Huff, C. D., Hubley, R., Shannon, P. T., … Galas, D. J. (2010). Analysis of genetic inheritance in a family quartet by whole-genome sequencing. *Science*, *328*, 636–639. https://doi.org/10.1126/science.1186802

Rocha, L. A., Robertson, D. R., Roman, J., & Bowen, B. W. (2005). Ecological speciation in tropical reef fishes. *Proceedings of the Royal Society B-Biological Sciences*, *272*, 573–579. https://doi.org/10.1098/2004.3005

Roux, C., Fraïsse, C., Romiguier, J., Anciaux, Y., Galtier, N., & Bierne, N. (2016). Shedding light on the grey zone of speciation along a continuum of genomic divergence. *Plos Biology*, *14*, e2000234. https://doi.org/10.1371/journal.pbio.2000234

Rundle, H. D., & Nosil, P. (2005). Ecological speciation. *Ecology Letters*, *8*, 336–352. https://doi.org/10.1111/j.1461-0248.2004.00715.x

Schluter, D. (2009). Evidence for ecological speciation and its alternative. *Science*, *323*, 737–741. https://doi.org/10.1126/science.1160006

Schmidt, J. (1923). The breeding places of the eel. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *211*, 179–208. https://doi.org/10.1098/rstb.1923.0004

Seears, H. A., Darling, K. F., & Wade, C. M. (2012). Ecological partitioning and diversity in tropical planktonic foraminifera. *BMC Evolutionary Biology*, *12*, 54. https://doi.org/10.1186/1471-2148-12-54

Seehausen, O., Butlin, R. K., Keller, I., Wagner, C. E., Boughman, J. W., Hohenlohe, P. A., … Widmer, A. (2014). Genomics and the origin of species. *Nature Reviews Genetics*, *15*, 176–192. https://doi.org/10.1038/nrg3644

Selkoe, K. A., D'Aloia, C. C., Crandall, E. D., Iacchei, M., Liggins, L., Puritz, J. B., … Toonen, R. J. (2016). A decade of seascape genetics: Contributions to basic and applied marine connectivity. *Marine Ecology Progress Series*, *554*, 1–19. https://doi.org/10.3354/meps11792

Smadja, C. M., & Butlin, R. K. (2011). A framework for comparing processes of speciation in the presence of gene flow. *Molecular Ecology*, *20*, 5123–5140. https://doi.org/10.1111/j.1365-294X.2011.05350.x

Sousa, V., & Hey, J. (2013). Understanding the origin of species with genome-scale data: Modelling gene flow. *Nature Reviews Genetics*, *14*, 404–414. https://doi.org/10.1038/nrg3446

Tariel, J., Longo, G. C., & Bernardi, G. (2016). Tempo and mode of speciation in Holacanthus angelfishes based on RADseq markers. *Molecular Phylogenetics and Evolution*, *98*, 84–88. https://doi.org/10.1016/j.ympev.2016.01.010

Tesch, F. (2003). *The Eel*. Oxford: Blackwell Science Ltd.

Tine, M., Kuhl, H., Gagnaire, P.-A., Louro, B., Desmarais, E., Martins, R. S. T., … Reinhardt, R. (2014). European sea bass genome and its variation provide insights into adaptation to euryhalinity and speciation. *Nature Communications*, *5*, 5770. https://doi.org/10.1038/ncomms6770

Via, S. (2001). Sympatric speciation in animals: The ugly duckling grows up. *Trends in Ecology & Evolution*, *16*, 381–390. https://doi.org/10.1016/S0169-5347(01)02188-7

Vollestad, L. A. (1992). Geographic variation in age and length at metamorphosis of maturing European eel – Environmental effects and phenotypic plasticity. *Journal of Animal Ecology*, *61*, 41–48. https://doi.org/10.2307/5507

Wielgoss, S., Gilabert, A., Meyer, A., & Wirth, T. (2014). Introgressive hybridization and latitudinal admixture clines in North Atlantic eels. *BMC Evolutionary Biology*, *14*, 61. https://doi.org/10.1186/1471-2148-14-61

Wirth, T., & Bernatchez, L. (2003). Decline of North Atlantic eels: A fatal synergy? *Proceedings of the Royal Society of London Series B-Biological Sciences*, *270*, 681–688.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Nikolic N, Liu S, Jacobsen MW, et al. Speciation history of European (*Anguilla anguilla*) and American eel (*A. rostrata*), analysed using genomic data. *Mol Ecol*. 2020;29:565–577. https://doi.org/10.1111/mec.15342