CHAPTER II

A tension hybrid zone between the North Atlantic eels of the genus *Anguilla*

Maes G.E., Pujolar J.M., Jonsson B. and Volckaert F.A.M.

SUMMARY

We reassessed the hypothesis of hybridisation between the partly sympatric European eel (*Anguilla anguilla* L.) and American eel (*A. rostrata*) by testing the joint distribution of microsatellite markers and vertebral numbers. We first characterized genetic variability and differentiation between both species in populations throughout Europe and America using 8 microsatellite polymorphic loci. Using classical population genetic techniques, we screened several populations of *Anguilla* on Iceland, where both species have been reported to co-occur, for indications of introgressive hybridisation with the American eel. Icelandic allelic richness, heterozygosity values and allele frequencies were intermediate between both species. Microsatellite loci, selected as moderately variable to avoid homoplasy, yielded a high genetic differentiation between both species (F_{ST} = 0.14, R_{ST} = 0.11; $p < 0.001$), congruent with earlier mitochondrial DNA studies. Subsequent multivariate and individual based assignment tests separated both species with a high level of confidence (> 95% assignment score). Global admixture proportions in Icelandic eel populations attributed 8% of the genomic material to American eel. Classical as well as model-based Bayesian individual assignment tests detected a total admixture within Iceland of 11-15 %, mainly composed of F_1 hybrids (6.3%) and pure *A. rostrata* (3.8%) individuals. The Total Number of Vertebrae was lower in Iceland and co-varied strongly with the admixture coefficient. Our results suggest a narrow tension zone, with asymmetric introgressive hybridisation towards the European eel. The incomplete/imperfect genetic isolation of eels spawning sympatrically affects additional locations in Europe with cryptic American like genomes. Differential oceanic migration of both species is thought to be safeguarded through reinforcement and selection against hybrids.

Keywords: Anguillids; Atlantic Ocean; Bayesian assignment; homoplasy; introgression; marine organisms; population genetics; sympatric speciation

This chapter is presented as an unpublished manuscript.

INTRODUCTION

Hybridization between individuals from genetically distinct populations or species is a central theme in biology, as it interacts with any genetic species concept and species delineation. Regions where genetically distinct individuals interbreed to form genotypes of mixed origin, named hybrid zones, have been studied in numerous ways in view of speciation, adaptation, and co-evolution (Burke & Arnold, 2001; Anderson & Thompson, 2002; Arnold, 2004). Selection acting in hybrid zones may be either endogenous or exogenous (Barton & Hewitt, 1985; Jiggins and Mallet, 2000). The first type is environment-independent and results from the interaction between alleles and genes from distinct taxa. The second type acts postzygotically and is mediated by differential fitness of parental and hybrid genotypes in various environments (Jiggins and Mallet, 2000; Burke & Arnold, 2001; Seehausen, 2004).

Two models have been proposed to explain the presence of hybrid zones (Barton & Hewitt 1985; Burke & Arnold, 2001). Under the "tension zone" model, hybrids exhibit a lower fitness than both parental species (hybrid inferiority) (Barton & Hewitt, 1985). As a consequence, natural selection acts against hybrids directly or through the process of reinforcement, a much-debated pre-zygotic isolating mechanism between species, decrease the chance of hybrid formation (Servedio, 2004). A binomial distribution is then observed consisting of pure individuals of both species plus occasional F_1 and F_2 hybrids (Jiggins & Mallet, 2000). A tension zone is maintained by a selection-dispersal equilibrium in time and space, resulting in gametic disequilibrium in offspring (Barton & Hewitt, 1985). Ecological divergence between most bimodal hybrid zones suggests that ecology contributes more to speciation than genomic incompatibility. Alternatively, hybrids zones may be due to hybrid superiority (Burke & Arnold, 2001). Heterosis and favorable epistatic gene combinations explain the higher fitness of hybrids in comparison with parental species (hybrid vigor). A unimodal genotype distribution is expected, with intermediate genotypes being predominant over parental genotypes (Jiggins & Mallet 2000). Exogenous selection is believed to play a crucial role in the establishment of fit hybrids. Hybrid matings often occur near the edges of a species range or in marginal habitats, where conditions prevail under which new variants are likely to survive and thrive (Burke and Arnold, 2001).

One issue remains unclear for many hybrid zones, namely whether they have a sympatric origin or is due to subsequent secondary contact (natural or anthropogenic). This knowledge is crucial to define conservation issues for endangered species (Allendorf *et al.*, 2001). This can be detected using mtDNA haplotype phylogenies and timing estimates between species

(e.g. molecular clock or mismatch distributions). In a secondary contact zone, the genetic distance between shared and non-shared haplotypes in both species should be much higher than the average distance between non-shared haplotypes within species. Additionally, vicariance enables the calculation of the most likely divergence time between species and the assessment of complete isolation or not.

Marine organisms exhibit a high dispersal potential, a broad distribution, numerous progeny and the absence of reproductive barriers. All these factors favor cross-species fertilization (Palumbi, 1994). Interspecific and intraspecific hybrid zones are more common in the ocean than initially thought (eg gastropods: Rolan-Alvarez, 1997; Cruz *et al.*, 2004; mussels: Bierne *et al.*, 2002; fish: She *et al.*, 1987; Planes & Doherty, 1997, McMillan, 1999; Nielsen *et al.*, 2003). The stability and size of hybrid zones seems not strongly dependent on the dispersal ability (O'Mullan *et al.*, 2001; Roques *et al.*, 2001), but more so on the forces shaping the fitness of hybrids (hybrid superiority vs tension zone) and the evolutionary potential of the hybrids (Moore, 1977; Barton & Hewitt, 1985; Burke & Arnold, 2001).

There are currently 15 *Anguilla* species recognized, all exhibiting a catadromous life history and highly adapted to oceanic currents to complete their life-cycle (Tsukamoto *et al.*, 2002). This study focuses on the two North Atlantic *Anguilla* species, the European eel (*A. anguilla*) and the American eel (*A. rostrata*) (Anguillidae; Teleostei) (Tesch, 2003). Although both taxa spend most of their lifetime in freshwater systems or estuaries, their early and late life-history is entirely comparable to that of marine organisms. They breed in the Sargasso Sea, between 20°N and 35°N, but show some spatio-temporal reproductive segregation (McCleave *et al.*, 1993) and larval migratory segregation (Arai *et al.*, 2000). Glass eels ascend rivers to feed and mature. After several years, in freshwater partially mature silver eels migrate back to their natal spawning grounds. Here they complete maturation, produce up to 2.10^6 eggs per female and die (Tesch, 2003).

The biological and taxonomic integrity of *Anguilla anguilla* and *A. rostrata* have long been under discussion. The only quasi-diagnostic morphometric marker known to distinguish both species is the number of vertebrae. *A anguilla* exhibits 110-119 (mean = 114.7) vertebrae, while *A. rostrata* exhibits 103-110 (mean = 107.1) vertebrae (Tesch, 2003). Tucker (1959) questioned the use of this trait based on possible selective influences on it. The genetic status of Icelandic eels has gained attention after the observation of a small fraction (0.3 % and 5.6 %, respectively) of eels with a reduced number of vertebrae ($\leq$ 110) in Northern Europe and in Iceland. They were proposed to represent either (1) pure American eel expatriates, (2) pure European eel individuals with ontogenic abnormalities or (3) F_1 hybrids.

Avise *et al.* (1990) detected hybrids in Iceland through cytonuclear disequilibria after combining meristic, mitochondrial and allozyme marker. Their data suggested ongoing gene flow between both species, including pure individuals of both species and hybrids incorporating American eel genomic material from 2 to 4%. Mank & Avise (2003) reassessed these conclusions with highly polymorphic microsatellites markers. No indications for hybridisation were detected in this latter study. A potential problem with this study is that highly polymorphic loci can be prone to homoplasy. In a recent paper, Maes *et al.* (2005, submitted) showed that eel species (including *A. anguilla* and *A. rostrata*) can be discriminated with high confidence based on four moderately variable microsatellite loci. The use of four loci is sufficient to detect various hybrid classes (Boecklen & Howard, 1997). We thus set out to use this four-marker set to reassess the hybrid status of the Icelandic eels.

In this study, we sampled putative *A. rostrata* and *A. anguilla* eel together with Icelandic eels. The objective of this study was threefold: (1) We first reassessed the genetic variability and differentiation between the European and American eel using moderately polymorphic microsatellite markers, hence avoiding homoplasy and enabling a highly reliable discrimination of both North Atlantic eel species. (2) We then tested the hypothesis of random hybridisation or unequal introgression between the American and the European eel at several Icelandic sites using meristic (Total Number of Vertebrae) and microsatellite markers. Under a tension zone model, we should detect a bimodal genotypic distribution of pure individuals of both species and occasional F_1 or F_2 hybrids. Under a hybrid superiority model, first generation hybrids should be common and the genotypic distribution more unimodal. The detection of second generation hybrids would imply assortative mating between hybrids, while backcrosses with one of both species are much more likely. (3) Finally, we discuss the possible causes of the presence of hybrids in Iceland and the nature of the hybrid zone in the Sargasso Sea.

MATERIAL AND METHODS

Material and meristic counts - Samples of adult and juvenile eels were collected from four American (N = 210), 9 Icelandic (N = 342) and 10 European (N = 591) locations. Sample sizes ranged from 6 to 60 individuals (see Table 1 & Figure 1 for details). The eels were captured with fyke nets or electrofishing. A piece of muscle or finclip tissue was sampled and stored in 100% ethanol until processing. Only glass eels were digitally X-ray photographed and the number of vertebrae counted by eye.

Table 1: Location of *Anguilla anguilla* and *A. rostrata* samples taken across Europe and North America; LAT : Latitude; LON : Longitude; *CODE*: sampling code; *N* : number of individuals; Life stages: G : glass eel; Y : yellow eel; S : silver eel.

COUNTRY	SITE	LAT	LON	CODE	STAGE	YEAR	N	SPECIES
Iceland	Vogslækur	64°23'N	21°22'W	IC01	G	2001	60	*A.anguilla* ?
	Vogslækur	64°23'N	21°22'W	IC02	G	2003	52	*A.anguilla* ?
	Reykjavik	64°00'N	21°11'W	IC03	Y	2001	11	*A.anguilla* ?
	Floi	63°50'N	20°40'W	IC05	S	1999	6	*A.anguilla* ?
	Reykhólar A	65°26'N	22°13'W	IC06	YS	2001	45	*A.anguilla* ?
	Reykhólar B	65°26'N	22°12'W	IC07	YS	2001	16	*A.anguilla* ?
	Vatnsdalsá	65°35'N	20°20'W	IC08	YS	2000	32	*A.anguilla* ?
	Vífilsstaðavatn	64°07'N	21°52'W	IC09	YS	2002	60	*A.anguilla* ?
Ireland	Burrishoole	53°55'N	09°55'W	AA01	G	2001	60	*A.anguilla*
				AA02	S	2001	60	*A.anguilla*
Netherlands	Den Oever	53°01'N	05°13'E	AA03	G	2001	60	*A.anguilla*
				AA04	S	2001	60	*A.anguilla*
W France	Loire	47°12'N	01°44'W	AA05	G	2001	60	*A.anguilla*
				AA06	S	2001	60	*A.anguilla*
S France	Tour du	43°33'N	04°38'E	AA07	G	2001	60	*A.anguilla*
	Valat			AA08	S	2001	51	*A.anguilla*
Morocco	Sebou	34°16'N	06°34'W	AA09	G	2001	60	*A.anguilla*
				AA10	Y	2001	60	*A.anguilla*
N-America	Florida	27°12'N	80°13'W	AR01	S	1999	30	*A.rostrata*
	Maine	44°02'N	69°58'W	AR02	G	2003	60	*A.rostrata*
		43°59'N	69°50'W	AR03	G	2003	60	*A.rostrata*
		43°51'N	69°37'W	AR04	G	2003	60	*A.rostrata*

DNA purification and microsatellite amplification – Purification of genomic DNA and amplification of microsatellites loci was performed following methods described in Dannewitz *et al.* (2005). We analysed the following eight nuclear microsatellite loci: AAN 01, AAN 03, AAN 05 (Daemen *et al.*, 2001), ANG 151, ANG 075, ARO 054, ARO 063 and ARO 095 (Wirth & Bernatchez, 2001). Electrophoresis and size determination of alleles was made on a LICOR 4200 automated sequencer (Westburg, Leusden, The Netherlands) using a 6% acrylamide 7 M urea sequencing gel. A molecular ladder (supplied by the manufacturer) was run along with the PCR products, and allele lengths and genotypes were assessed with the GeneImagIR 4.03 software (Scanalytics inc, Fairfax, USA).

Data analysis of genotypes - Genetic diversity estimates such as the level of polymorphism, and observed and expected heterozygosity (H_o and H_e) were calculated in GENETIX version 4.05 (Belkhir *et al.*, 1999). Allelic Richness (R) and Gene Diversity (H_S) comparisons between species (*A. anguilla* – AA and *A. rostrata* AR) and Icelandic samples, as well as

departures from Hardy-Weinberg equilibrium (F_{IS}) and R_{ST} values (Rho ST) were calculated using the software FSTAT version 3.9.5 (Goudet, 1996). Genetic differentiation was characterized using hierarchical F-statistics (θ, Weir & Cockerham, 1984) and G_{ST}-values as implemented in the GENETIX 4.05 software package (Belkhir *et al.*, 1999). Significance was assessed with permutation tests (1000 replicates). In all cases, significance levels were corrected for multiple comparisons using a sequential Bonferroni correction (Rice, 1989). Unbiased pairwise genetic distances (Nei, 1978), a Neighbour-joining dendrogram and branch boostrap values (1000 iterations) were calculated using the software package PHYLIP (Felsenstein, 1996). Due to homoplasy or uninformative loci, misassignments can easily occur or the confidence interval may become very broad, especially when analysing admixed individuals. We used the WHICHLOCI software (Banks *et al.*, 2003) to pinpoint the most discriminating loci between species and avoid possible misclassifications due to homoplasy. The software GENECLASS (Piry *et al.*, 2000) was used to assign individuals to their respective species. We first used a classical Bayesian assignment method (Rannala & Mountain, 1997), assuming known baseline populations without admixture analysis. To assess the species integrity within each species and to detect admixture, a fully Bayesian model based individual clustering algorithm was used, as implemented in the software STRUCTURE 2.1 (Pritchard *et al.*, 2000). This method assumes no prior knowledge about population structure. The program organises individuals into a predefined number of clusters (K), which may represent putative populations or species, and returns log likelihood values for different Ks. Initial analyses were performed with K=1 to K=5, to account for possible population structure within species. The non-admixture model was used in first instance (to detect pure non-admixed genotypes within each species) in combination with the correlated allele frequency model (low genetic differentiation at several loci) to discriminate the two species with high confidence. A burn in length of 10,000 iterations followed by 100,000 additional Monte Carlo iterations were performed as recommended by Pritchard *et al.* (2000). Each assessment of K was repeated three times to check the confidence of the results. The admixture model was then applied with prior morphological and genetic species information and the correlated allele frequencies model to detect admixed genotypes (hybrids) within Icelandic and baseline samples. We considered individual with admixture coefficients (q) between 0.80 and 1 to be a pure species and scores between 0.80 and 0.20 to be hybrids or backcrosses. A recently developed specific test to accurately discriminate between hybrid classes, implemented in the software NEWHYBRIDS version 1.1 (Anderson & Thompson, 2002) was subsequently used. This algorithm can assess fine scale hybridisation between two

species, enabling the detection of six different genotype classes, namely pure species, F_1, F_2 hybrids and backcrosses between hybrids and pure species. The program computes by MCMC the Bayesian posterior likelihood that individuals fall into different hybrid categories (based on allele frequency distributions). To test the discriminative power of all loci versus the most discriminating ones detected before, we repeated this last analysis using both datasets (discarding the non informative loci in the second analysis). We then compared classical assignment (GENECLASS), model based Bayesian assignment (STRUCTURE) and NEWHYBRIDS assignment results to increase the confidence of hybrid detection. Finally, we calculated the parental contribution of both species within the Icelandic group of individuals using the ADMIX software (Bertorelle & Excoffier, 1998), yielding two estimators of admixture proportions based on a coalescent approach that explicitly takes into account gene frequencies and molecular information. We used both the gene frequency (m_R) as the coalescent (m_Y) methods, assuming a large genetic differentiation between both species. Standard deviations were calculated by bootstrapping technique for each parental and hybrid population (1000 bootstraps). The power of detection of hybrid individuals through H&W disequilibrium is low in marine species with high expected heterozygosities (H_E) (Nielsen *et al.*, 2003). Tension zones are often characterized by gametic phase disequilibrium due to dispersal of parental gene combinations into the center of the zone (Barton & Hewitt, 1985). Gametic phase disequilibrium, corresponding to the correlations of gene identities across loci within samples, was estimated to detect admixture and infer the nature of a possible hybrid zone (tension vs. hybrid vigor zone) using the program ESTIM 1.1 (Vitalis & Couvet, 2001).

RESULTS

Meristic characteristics – The Total Number of Vertebrae (TNV) within the American samples (N = 3) ranged from 104 to 110, with a mean of 107.2 (± 2.08). The number of vertebrae within the European eel samples (N = 5) ranged from 110 to 119, with a mean of 114.9 (± 1.25). The meristic counts of Icelandic samples (N = 2) ranged from 106 to 120, with a mean of 113.6 (± 2.17). Out of 107 X-ray photographed Icelandic glass eels, 8 exhibited intermediate TNV counts of 110 or less (7.4 %) compared to 2 out of 296 (0.6 %) in European glass eels (Figure 1).

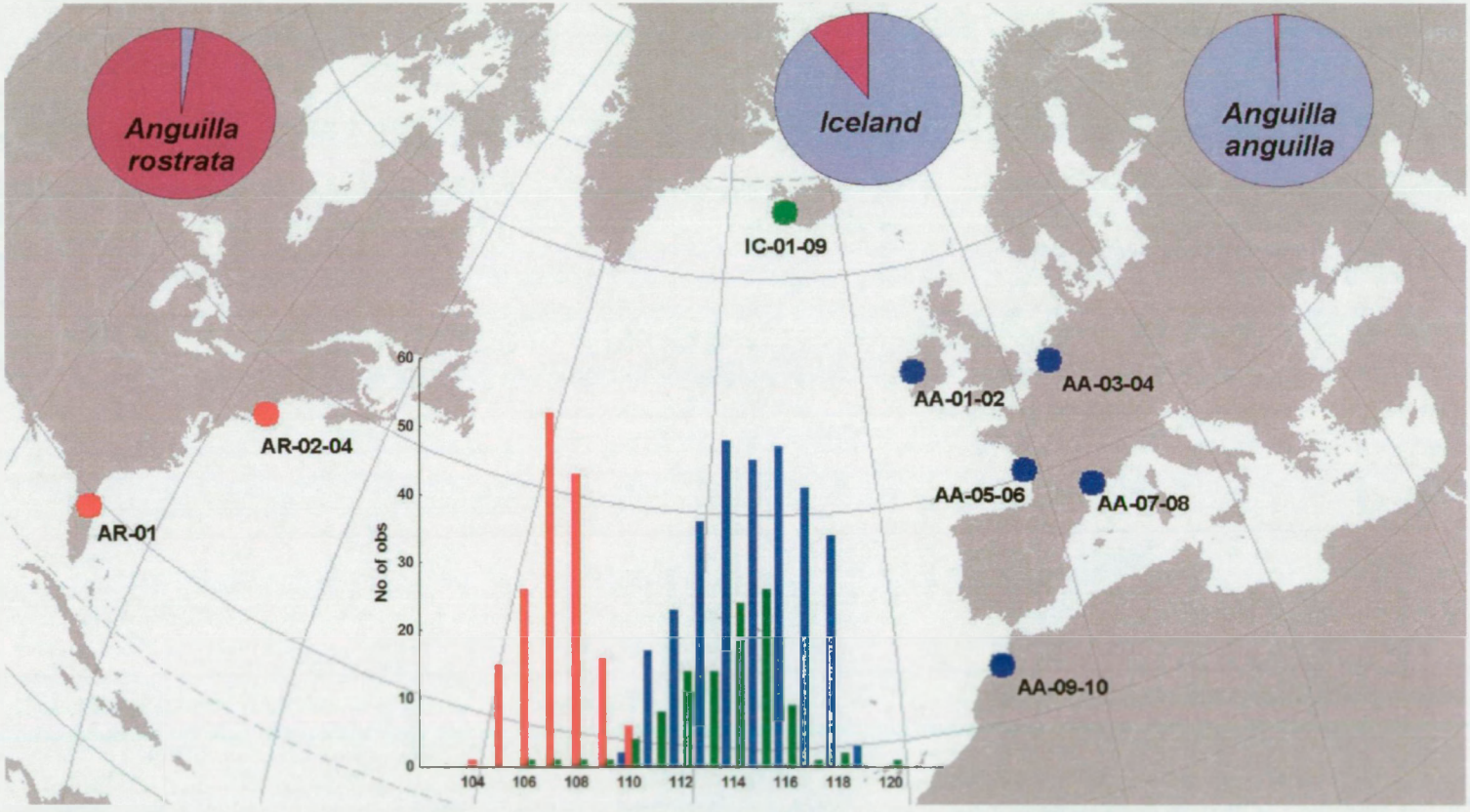

Figure 1: Sampling locations of European eel (*Anguilla anguilla* L., AA-01-10), American eel (*A. rostrata*, AR-01-04) and Icelandic eel (IC-01-09) (for sampling codes, see Table 1). The pie diagram represents the admixture proportion within each species and in Iceland based on combined results of the three assignment methods (Table 3). The histogram represents the Total Number of Vertebrae (TNV) distribution in European eel, American eel and Icelandic individuals.

Genetic variability and departures from Hardy-Weinberg proportions – Genetic variability estimates were similar for European eel (H_E = 0.73-0.77; R = 5.54; H_S = 0.77) and American eel (H_E = 0.78-0.80; R = 5.78; H_S= 0.80). Icelandic samples (IC) mostly exhibited intermediate diversity values: 0.72-0.79 (H_E), 0.79 (H_S) and 5.60 (R) (Table2). All three groups differed significantly in gene diversity and allelic richness amongst each other ($p < 0.05$). Allelic distribution differed highly across loci (Figure 2).

Table 2 : Genetic diversity estimates of *Anguilla anguilla*, *A. rostrata* and Icelandic samples, including expected (H_e) and observed heterozygosity (H_o), Gene diversity (H_s) and allelic richness (R). Standard deviation is showed between brackets.

LOCATION/SPECIES	SAMPLE	H_e.	H_o.	H_s	R
Icelandic samples	IC01	0.7685 (0.2118)	0.7297 (0.2016)	0.776	5.574
	IC02	0.7502 (0.2042)	0.6320 (0.2125)	0.760	5.391
	IC03	0.7013 (0.2742)	0.7260 (0.3248)	0.747	5.411
	IC04	0.7474 (0.2369)	0.6657 (0.2292)	0.757	5.390
	IC05	0.7272 (0.1242)	0.6810 (0.1952)	0.811	5.199
	IC06	0.7591 (0.1779)	0.6653 (0.2179)	0.776	5.324
	IC07	0.7703 (0.1235)	0.7946 (0.1759)	0.835	5.778
	IC08	0.7410 (0.2386)	0.6789 (0.2512)	0.756	5.321
	IC09	0.7681 (0.2041)	0.7018 (0.2180)	0.780	5.536
European samples *(Anguilla anguilla)*	AA01	0.7241 (0.2813)	0.7139 (0.2645)	0.731	5.190
	AA02	0.7411 (0.2509)	0.6268 (0.2110)	0.750	5.351
	AA03	0.7448 (0.2371)	0.6851 (0.2233)	0.752	5.271
	AA04	0.7449 (0.2170)	0.7354 (0.2067)	0.752	5.333
	AA05	0.7562 (0.2224)	0.7267 (0.2032)	0.763	5.423
	AA06	0.7513 (0.2294)	0.7430 (0.2112)	0.759	5.427
	AA07	0.7429 (0.2453)	0.6892 (0.2462)	0.750	5.307
	AA08	0.7192 (0.2819)	0.6746 (0.2629)	0.728	5.216
	AA09	0.7564 (0.1968)	0.7198 (0.1884)	0.764	5.275
	AA10	0.7441 (0.2311)	0.7024 (0.2217)	0.751	5.276
American samples *(Anguilla rostrata)*	AR01	0.7580 (0.2077)	0.7405 (0.1738)	0.772	5.519
	AR02	0.7796 (0.1744)	0.6713 (0.1794)	0.788	5.530
	AR03	0.7680 (0.2007)	0.6692 (0.1704)	0.776	5.585
	AR04	0.7786 (0.1658)	0.7138 (0.1439)	0.786	5.580

Loci AAN 05, AAN 03, AAN 01, ARO 054 and ARO 063 showed strong differences between species and various species-specific alleles. Locus ANG 075 showed a strong deviation from H&W expectations (F_{IS} > 30%), probably because of the presence of null alleles. As this locus does not provide any discrimination power to separate the two species (see lower), it was removed from the dataset. In the remaining dataset, among 161 tests (23 samples × 7 loci) for HWE, 17 (10 %) showed significant deviations from expected genotype

frequencies after sequential Bonferroni correction (α = 0.05, k = 23). All deviations represented heterozygote deficiencies at the loci ARO 054 and ARO 095. The risk of encountering heterozygote deficiencies as a result of large-allele dropouts has been mentioned to increase when multiplexing primers, especially for highly variable loci (O'Connell & Wright, 1997).

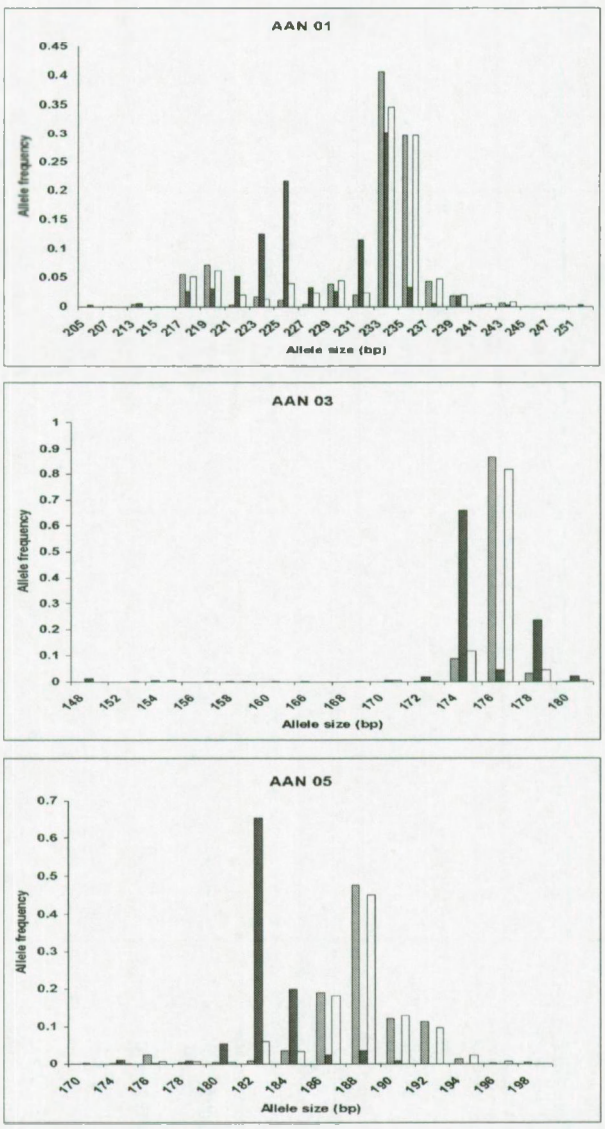

Figure 2: Allelic frequency distribution in *Anguilla anguilla* (light grey bars), *A. rostrata* (dark grey bars) and Icelandic samples (white bars) at the three most discriminative microsatellite loci: AAN 01 (upper panel); AAN 03 (middle panel); AAN 05 (lower panel).

To test for this, we re-amplified the two deviating loci separately in two different samples, but found identical genotypes, ruling out this reason for deviation. We kept to our original dataset because both loci did not influence genetic differentiation values in a previous study on a larger scale (Dannewitz *et al.*, 2005), and because populations from Iceland possibly consisted of admixed individuals. Once corrections for null alleles are applied, genotypic information is lost (Dannewitz *et al.*, 2005). Although not significant, the Icelandic populations exhibited higher F_{IS} values than both species separately. Raw allele frequencies provide further evidence of affinities between both species. Icelandic populations show a decrease of the most common AA allele and an increased frequency of the most common AR allele at loci AAN 01, AAN 03 and AAN 05 (Figure 2). Overall, several Icelandic samples take an intermediate position at the most divergent loci (data not shown).

Multi-locus analysis of genetic differentiation – Overall F_{ST} between all samples was 0.05; global G_{ST} 0.06, while overall R_{ST} amounted to a similar value of 0.04. When considering all samples within species together (removing possibly introgressed Icelandic samples), overall genetic differentiation between species was high (F_{ST} = 0.14, G_{ST} = 0.075 and R_{ST} = 0.11), with a maximum pairwise F_{ST} value of 16.3 % between samples AA1 and AR1. There is no diagnostic difference between species at any single locus, but there are three main differentiating loci: AAN 03 (highest F_{ST}), AAN 05 (highest R_{ST}) and AAN 01. Loci ARO 054, ANG 151 and ARO 063 have the same order of magnitude of differentiation (F_{ST} around 2 %, data not shown). All values were highly significant ($p < 0.001$), except for ANG 075, a locus we discarded in the further individual analyses due to the presence of null alleles and the lack of a differentiation signal. Within species, differentiation was very low ($F_{ST} < 0.01$) and only showed some significant values in the European eel, without any evidence of geographical clustering (data not shown). A phenogram constructed using unbiased Nei (1978) distance, separates *A. rostrata* and *A. anguilla* with high confidence (98%) and places several Icelandic populations at an intermediate position between both species or within the *A. anguilla* cluster (Figure 3).

Power of assignment – In total, 1000 new populations (N = 100) were simulated from each species based on the allele frequencies data to target the most discriminating loci for species identification. One locus (AAN 05) was sufficient to discriminate both species with high accuracy (98.5 %). In order to increase the detection power for F_1 and further hybrid classes, two datasets were constructed. A first dataset included the best classifying loci, namely AAN

05, AAN 03 (94.5 %), AAN 01 (88%) and ARO 063 (76.5 %). A second dataset included all loci (excluding ANG 075). Estimates of correlation of gene identities (eta_S = 0.000832) at pairs of loci (gametic phase disequilibrium) revealed a positive significant outcome (95 % Confidence interval: [0.000281, 0.001920]) in the pooled Icelandic sample, but not in the pure species (eta_S < 0).

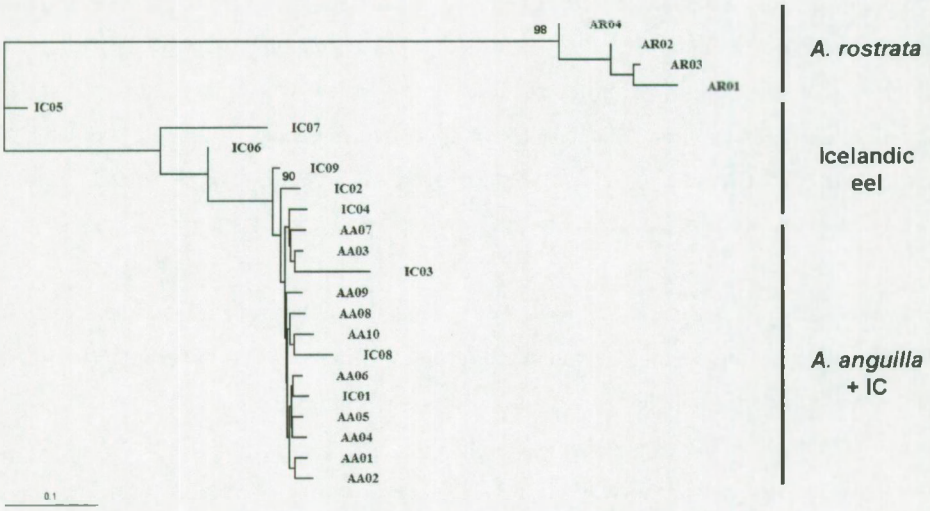

Figure 3: Neighbour-joining tree of all *Anguilla* samples based on pairwise unbiased Nei (1978) distances. Confidence values after 1000 bootstraps.

Detection, extent and level of introgression in Anguilla species - We first used classical assignment to test the discriminative power of the applied method. When both presumably pure species were analysed, more than 99 % of all individuals were correctly assigned to their morphological species class (Table 3). We detected 0.5 % American material (3 pure AR, 3 partially admixed individuals) in the European populations and 1 % European material (2 pure AA, 3 partially admixed individuals) in the American populations. When the Icelandic individuals were treated as an unknown sample to be assigned to one of the two species, 31 individuals (11.9%) with American eel ancestry were detected in the predominantly European Icelandic samples (Table 3). Most (6.9 %) of the AR-like individuals in Iceland showed the pattern of admixture between both species (18 individuals with score < 0.90) and 5 % were considered pure AR (13 individuals with score > 0.90). To differentiate between real hybrids and a lower assignment success of a pure individual, a fully Bayesian approach was implemented in STRUCTURE. Initially, a cluster analysis was performed to test pure

individuals of both species; K = 2 showed the highest posterior probability. Assignment proportions were very high, approximating 99 % in both species (Table 3, Figure 4a). In the European eel, we found four pure AR, two completely admixed individuals and one partially admixed individual. In the American eel, five highly admixed individuals were detected. The Icelandic samples were subsequently analysed as unknown individuals, while both morphological species were given a prior of known species flag. Icelandic samples exhibited a high admixture proportion (85 % AA, 15 % AR) (Figure 4a and 4b). When tentatively defining classes, 78% pure AA ($1.0 < q < 0.80$), 1.1% pure AR ($q < 0.20$) and 21% hybrids ($0.20 < q < 0.80$) were differentiated (Figure 4b). Further geographical analysis of admixture proportions showed that *A. rostrata* material in European populations is not only present in Iceland (11.5 %), but also in Morocco (0.7%) and The Netherlands (0.2%), although only as admixed ancestor (one generation back admixture). There was also a slight difference in admixture proportion between life stage, and among geographical locations within Iceland. Sample IC05 has the highest American eel material, but this sample consisted of already sorted individuals with discordant vertebral count versus mtDNA haplotype (Table 3).

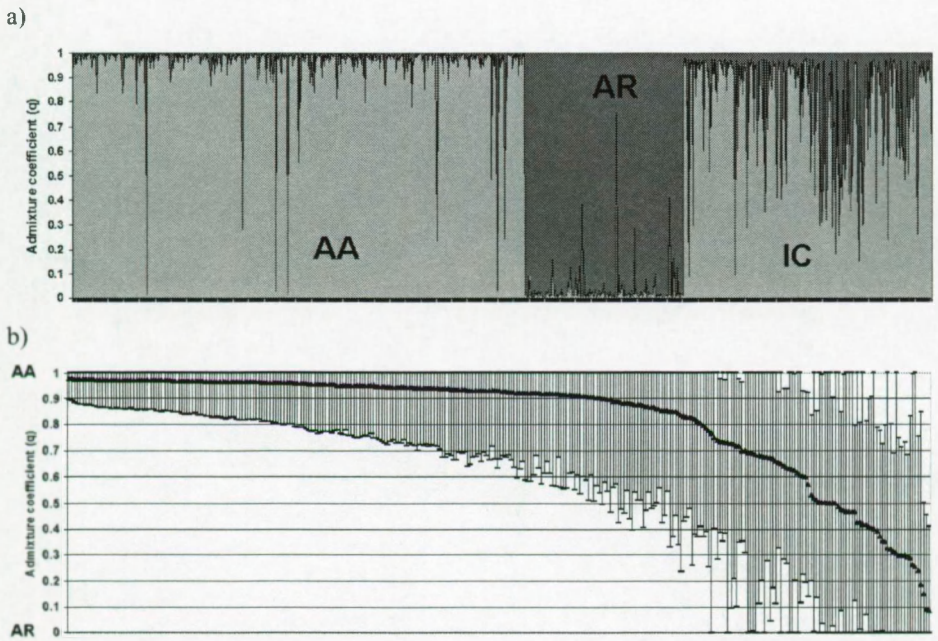

Figure 4: Admixture (STRUCTURE) results of European, Icelandic and American *Anguilla* at K= 2 (2 species). a) Each vertical bar represents the admixture proportion of an individual. Dark grey represents AR, light grey represents AA. Icelandic individuals are represented at the end of the figure. b) Admixture proportion of all Icelandic individuals at K = 2, with confidence interval bars (0.90). $q = 1$ is a pure AA, $q = 0$ is a pure AR.

Table 3: Assignment success (Classical method) and admixture proportion (Model based method) and Hybrid class probability (NewHybrids method) of both pure species and Icelandic eel individuals. AA: *Anguilla anguilla*, AR: *A. rostrata*, IC: Icelandic populations, F_1 : First generation hybrids, F_2 : second generation hybrids, BX AA : Backcrosses between F_1 and AA, BX AR : Backcrosses between F_1 and AR. m_Y and m_R: Estimator of Admixture proportion using gene frequencies (m_R) and coalescent information (m_Y)

	CLASSICAL ASSIGNMENT		MODEL BASED ASSIGNMENT		HYBRID CLASS ASSIGNMENT					
Presumed PURE species	**AA**	**AR**	**AA**	**AR**	**AA**	**AR**	**F_1**	**F_2**	**BX AA**	**BX AR**
AA	0.9950	0.0050	0.9890	0.0110	0.9893	0.0015	0.0079	0.0008	0.0006	0.0001
AR	0.0100	0.9900	0.0090	0.9910	0.0088	0.9649	0.0203	0.0039	0.0002	0.0018
ALL ICELANDIC EELS	0.8810	0.1190	0.8500	0.1500	0.8928	0.0380	0.0626	0.0043	0.0015	0.0008
IC01	0.9091	0.0909	0.8800	0.1197	0.9077	0.0301	0.0578	0.0027	0.0011	0.0005
IC02	0.9375	0.0625	0.8620	0.1381	0.9514	0.0207	0.0233	0.0031	0.0011	0.0004
IC03	0.8889	0.1111	0.8810	0.1188	0.9690	0.0111	0.0128	0.0046	0.0022	0.0003
IC04	0.8913	0.1087	0.8760	0.1239	0.9116	0.0137	0.0702	0.0028	0.0012	0.0005
IC05	0.3333	0.6667	0.5990	0.4013	0.5081	0.2479	0.2193	0.0170	0.0030	0.0047
IC06	0.8077	0.1923	0.7350	0.2649	0.7206	0.0998	0.1638	0.0106	0.0027	0.0024
IC07	0.7143	0.2857	0.6750	0.3250	0.7169	0.2244	0.0486	0.0066	0.0011	0.0024
IC08	0.9259	0.0741	0.8730	0.1271	0.9116	0.0214	0.0606	0.0039	0.0021	0.0004
IC09	0.9429	0.0571	0.8880	0.1116	0.9664	0.0106	0.0182	0.0033	0.0012	0.0003
JUVENILES	0.9233	0.0767	0.8720	0.1283	0.9113	0.0343	0.0492	0.0035	0.0012	0.0006
ADULTS	0.7863	0.2137	0.8350	0.1646	0.8806	0.0403	0.0714	0.0049	0.0017	0.0009
ALL 3 METHODS	**% AA**	**% AR**	***N* TOTAL**	***N* ADMIXED**	**Species contribution to Iceland:**			m_R	m_Y	
AA	0.991	0.009	548	5				0.926	0.966	
AR	0.025	0.975	199	5				0.074	0.034	
IC	0.892	0.108	259	28						

When assessing the parental contribution to the Icelandic populations by calculating the admixture proportion based on a coalescent approach and considering only gene frequencies, the contribution of AA in Iceland was high (92.6 %), while only 7.4 % of AR contributed to the gene pool in Iceland, When including molecular distance between alleles, the contribution of AR was lower (3.4%). The proportions of the six predefined hybrid classes were assessed by using the prior knowledge of both species as input. In Europe, 98.9 % was assigned to AA, 0.1 % to AR, 0.70% to the F_1 hybrid class, 0.08 % to the F_2 hybrid class and less than 0.2 % to backcrosses. In America, 96.5% was assigned to AR, 0.8 % to AA, 2 % to the F_1 hybrid class, 0.4 % to the F_2 hybrid class and less than 0.2 % to backcrosses (Table 3, Figure 5).

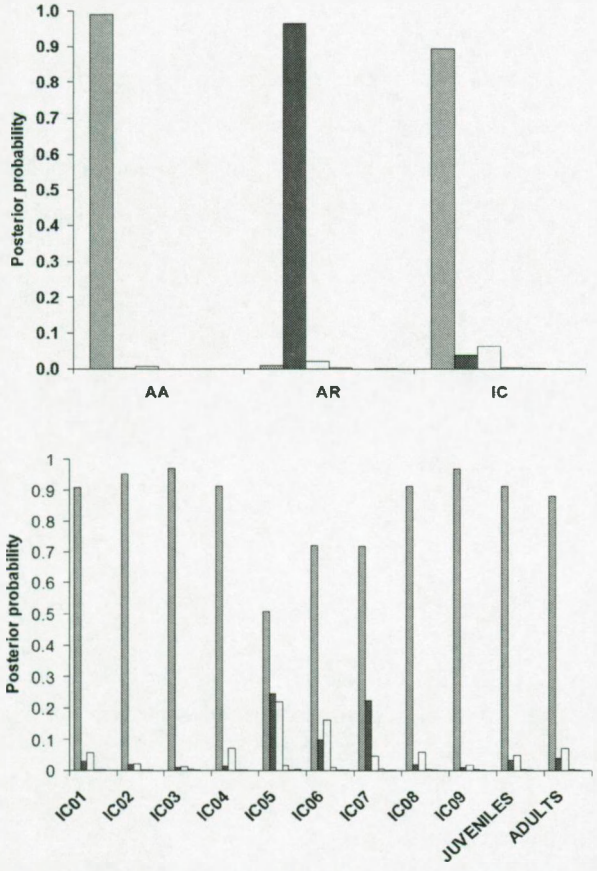

Figure 5: NEWHYBRIDS hybrid class probabilities of European, American and Icelandic *Anguilla* showing the proportion of pure AA (light grey), pure AR (dark grey), F_1 hybrids (white) in the pure species populations and in the Icelandic samples grouped (upper panel) or all Icelandic samples separately (lower panel). No F_2 hybrids, backcrosses with AA or backcrosses with AR could be detected.

A different pattern was observed in the Icelandic samples: in total 89 % of the individuals were pure AA, 3.8 % were pure AR and 6.3 % were F_1 (Table 3). No posterior probability was the highest for F_2 and backcross hybrid classes in any population. On average, there were more admixed individuals in adults than in juveniles. When testing the discriminative power of four moderately variable versus all seven loci, there was virtually no differentiation in outcome between both analyses (data not shown). When combining the results of all three assignment methods by selecting only admixed individuals detected by at least two techniques, an admixture proportion of 10.8 % American eel genomic material was detected within Icelandic populations, while only 1% and 2.5% of foreign genetic material was detected in the European and American eel, respectively (Table 3).

Relationship between meristic and genetic markers - A regression analysis between the mean number of vertebrae and the Bayesian admixture coefficient (q) yielded a significant positive correlation ($r = 0.46$; $p < 0.001$) (Figure 6). The regression improved even more when only admixed individuals ($q < 0.90$) were used for the analysis ($r = 0.58$ and $p < 0.001$) (Figure 6).

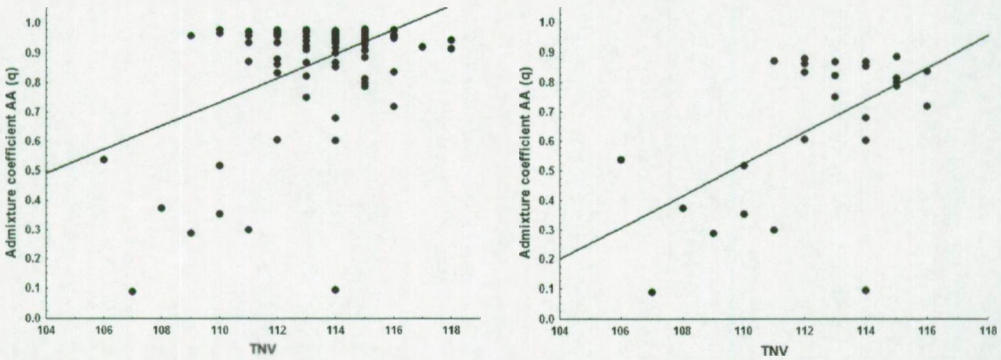

Figure 6: Correlation between the Total Number of vertebrae (TNV) and the admixture coefficient (q) of all Icelandic glass eels (left) and only Icelandic individuals exhibiting $q < 0.90$ (right).

DISCUSSION

Genetic diversity of North Atlantic Eels

Overall genetic variability estimates in European and American eel were similar to the values observed in other marine species with large effective population sizes (Ruzzante, 1998, Shaw, 1999; Roques *et al.*, 2001; Hutchinson *et al.* 2003; Knutsen *et al.* 2003) and correspond with previous microsatellite surveys in eel (Daemen *et al.*, 2001; Wirth and Bernatchez, 2001,

2003; Dannewitz *et al.*, 2005). The American eel shows a significantly higher gene diversity and allelic richness than the European eel, although no differences were found in observed and expected heterozygosities or polymorphism. Both species exhibit a similar life history with a long spawning migration and larval drift of almost a year, resulting in a high chance for dispersal, generation overlap and a large geographical distribution (Tesch, 2003). Nevertheless, the European eel is expected to maintain a larger effective population size, as it colonizes a much larger habitat and shows a higher age at maturity (Tesch, 2003). Additionally, previous studies showed more evidence for heterogeneity within the European eel compared to the American eel (Maes & Volckaert, 2002; Wirth & Bernatchez, 2003), a pattern expected to maintain genetic diversity in a species (Frankham, 2002). Accordingly, we only observed genetic heterogeneity among European locations but not among American locations. The slightly higher genetic diversity in the American eel may be attributable to the stronger decline in European eel recruitment in the last decades, partly due to the high fishing pressure and the numerous obstructions of the migration routes in Europe (Dekker, 2003). Alternatively, a lower historical effective population size within the European eel during glaciations has been suggested (Wirth & Bernatchez, 2003). Finally, the European eel might be a sister species of the American eel (Williams & Koehn, 1984; Avise *et al.*, 1990; Tsukamoto *et al.*, 2002), only retaining a subset of the original genetic diversity. In the larger scale study of Dannewitz *et al.* (2005), the American eel (N = 260) also showed a higher gene diversity and allelic richness than the European eel (N = 2626) at six out of eight loci, indicating that sample size are not the cause of our observation.

Genetic differentiation between the North Atlantic eels

As expected in marine species, the pattern of genetic differentiation within species was low ($F_{ST} < 0.01$) (Waples, 1998), and congruent with the hypothesis of panmixia in the American eel (Wirth & Bernatchez, 2003) and a subtle spatio-temporal heterogeneity within the European eel (Dannewitz *et al.*, 2005). Our study shows a high genetic differentiation between American and European eel, with multilocus pairwise differentiation (F_{ST}) values as high as 16.3 % and monolocus values of up to 69 % (AAN 03). Such high values contrast with earlier microsatellite surveys, where a maximal F_{ST} of 5% was observed (Mank & Avise, 2003; Wirth & Bernatchez, 2003). The discrepancies among studies are probably due to the fact that different sets of microsatellites were used. Our study included a combination of four moderately polymorphic and four highly polymorphic loci. Using the latter, F_{ST} values were

in the same range as detected in previous studies, which might be related to homoplasy (Mank & Avise, 2003). Estoup *et al.* (2002) recently showed that homoplasy is expected to reach the highest values when population sizes and mutation rates are high, characteristics that apply to marine fish and microsatellite loci, respectively. Species discrimination with moderately polymorphic markers is preferable (Manel *et al.*, 2002), as it lowers the probability to score Identical-In-State (IIS) alleles due to irregular mutation patterns or length restriction. The strategy of using only moderately variable SSR loci proved successful in the discrimination of various eel species (Maes *et al.*, 2005, in prep), where F_{ST} was negatively correlated to microsatellite polymorphism. The most discriminating loci between the American and the European eel (and contributing the most to the high F_{ST} observed) were the three lowly polymorphic loci (AAN 01, AAN 03, AAN 05). Locus AAN 05 was the most discriminating between the two species, indicating that less polymorphic loci exhibiting a lower mutation rate could better reflect long-term historical divergence (Estoup *et al.*, 2002).

Evidence for hybridisation in Icelandic samples

The Icelandic samples clearly exhibited patterns of mixed ancestry, detected by an intermediate level of genetic variability at three diversity estimators (heterozygosities, genetic diversity and allelic richness), intermediate allelic frequencies between American and European eel at loci AAN 01, AAN 03 and AAN 05, Hardy-Weinberg deviations, significant gametic disequilibria, and especially by a higher proportion of American genomes than in European populations (11-15% versus 1%). All methods (the classical and two Bayesian methods) yielded congruent results, suggesting that the Icelandic samples consist largely of European eel (89%), but with a high admixture (10.8%) of American eel material. By comparison, Avise *et al.* (1990) observed only 2-4% of American eel material in Icelandic samples. The more than double amount of admixture observed in our study might be attributed to the multiple polymorphic unlinked nuclear loci, which increases the power of hybrid detection (Boecklen & Howard, 1997; Anderson & Thompson, 2002). Our estimates of admixture within Iceland (10-15%, depending on the method used) were similar to other hybridizing marine taxa. (redfish: Roques *et al.*, 2001; flatfishes: She *et al.*, 1987). To discriminate between hybridisation and accidental recruitment of pure American eels in Iceland, admixture analyses enable the classification of genotypes as originating from one (pure) or two (hybrid) species. Icelandic genotypes consisted largely of parental species (89.3% AA and 3.8% AR), but strong evidence for the presence of admixed individuals with

common ancestry from both species (6.3 %) was also detected. A pure mechanical common distribution of both species is not concordant with such data.

The genetic identification of hybrids was confirmed by vertebral counts. While the number of vertebrae in American eels ranged from 104 to 110, 99.4% of European eels exhibited 111 vertebrae or more, with only 2 out of 296 European eels exhibiting 110 vertebrae. In agreement with microsatellite data, the number of vertebrae in Icelandic individuals showed an intermediate position (106-120 vertebrae) between the American and the European eel. A higher proportion of low vertebral count individuals (7.4 % with TNV < 110) than in Europe was observed. Furthermore, a strong positive correlation was observed between genetic admixture coefficients and the total number of vertebrae, which indicates that the basis of this trait is genetic and not environmental. Avise (1990) suggested that Icelandic individuals exhibiting less than 110 vertebrae either were pure *A. rostrata* expatriates migrating together with *A. anguilla* individuals, rare *A. anguilla* individuals showing environment-related ontogenic abnormalities or hybrids between the two species. The detection of F_1 hybrids besides pure American eel individuals in our study clearly supports the latter explanation. The second explanation is unlikely, as the strong correlation between the genetic admixture coefficient q and the total number of vertebrae constitutes evidence in favour of a genetically based number of vertebrae.

Our data can only be interpreted as evidence for a hybrid zone in the Sargasso Sea, still detectable at about 6,000 km from its origin. Interestingly, some admixed individuals were also detected in Morocco, in the southernmost distribution area. Hybrids are known to colonize and survive better in intermediate niches or challenging habitats (Dowling & Secor, 1997), the extremely disjunct distribution of admixed eels and their survival until adulthood (silver eels) may point to a similar pattern. The occurrence of admixed and absence of pure American eel in Morocco rules out an anthropogenic contribution to this pattern. In Iceland, the admixture proportion was similar across life stages, although we detected more American-like genomes in adults, which might point to temporal stability and a better survival of hybrids in marginal habitats. We have a common sampling site with the study of Avise *et al.* (1990) (Reykholar), where a higher proportion of hybrids and pure AR was also found (IC6 and IC7). This suggests either differential survival of hybrids in certain habitats or the importance of current patterns advecting American like larvae to Iceland (Avise *et al.*, 1990; Dowling & Secor, 1997). Within Icelandic samples exhibiting hybrids, an unequal contribution of the two species was detected, pointing to the genetic introgression of American eel material into European eel. While a small proportion of European eel material

was found in American eel samples (2.5 %), American eel material accounted for only 1% in Europe, but 11 % in Iceland.

A primary or secondary tension zone in the Sargasso Sea ?

The nature of hybrid zones in marine organisms remains difficult to define due to difficulties in assessing the selection forces acting on genotypes at sea and due to the sympatric reproduction of several species, blurring the signal of primary or secondary contact (Barton & Hewitt, 1995; Nielsen *et al.*, 2003). The low number of hybrid genotypes observed in this study may have two causes. First, both species spawn at different times and places, although partly in sympatry (McCleave, 1993;Tesch, 2003). One-week-old larvae are recorded in the Sargasso Sea with a huge distributional overlap, suggesting a similar overlap in spawning time/site in both species. No adult eel has ever been observed spawning, so other pre-zygotic barriers like spawning depth, assortative mating and gametic incompatibility are highly speculative and untestable at this time. The reinforcement of differential spawning location or behaviour, through reproductive character displacement is, however, a likely explanation (Servedio, 2004). Reinforcement is known to be strongest in sympatric zones; hence, hybrid avoidance might well be achieved through differences in spawning depth in the zone of overlap. Secondly, the distribution of genotypes in Icelandic samples fits a tension zone model, in which selection acts against hybrid individuals, which present a lower fitness in comparison with parental species. Icelandic genotypes were made up mostly by parental species (89.3 % AA and 3.8 % AR), while hybrids showed no superiority, with only occasional F_1 and no F_2 hybrids (6.3 %). In a tension zone, selection must act strongly on hybrids in order to lower their numbers (Barton & Hewitt, 1985). We detected a significant gametic phase disequilibrium in the Icelandic samples, which may be the consequence of genomic stress within hybrids, concordant with a tension zone. Further, no backcrosses were observed using the NewHybrids admixture approach (Anderson & Thompson, 2002), pointing either to a lower reproductive fitness of hybrids, to their very low number or to a lack of discrimination power using our set of markers. In the model based clustering method of Pritchard *et al.* (2000), however, we arbitrarily discriminated three genotype classes, albeit with high confidence interval. Only 78% was assigned with high confidence to the European eel, while 21 % showed the signal of admixture with one or the other species. Combining both results, we conclude that only the use many almost diagnostic markers may enable the detection of more complicated hybrid classes, such as backcrosses and F_2's (Anderson &

Thompson, 2002). Icelandic individuals have nevertheless an increased chance to reproduce with European eel, as they migrate with the same currents to the Sargasso Sea (Tsukamoto *et al.*, 2002; Tesch, 2003).

The low number of F_1 hybrids observed indicates that the hybrid zone between the American and the European eel is most likely a secondary tension zone, maintained by a balance between natural selection and extensive dispersal to safeguard species integrity. Both American and European eel share the same spawning grounds (the Sargasso Sea between 20-35°N), which provides an opportunity for hybridisation (Palumbi, 1994). In North Atlantic eels, spawning site migration seems to have evolved together with oceanic current shifts under the influence of glaciations (Tsukamoto *et al.*, 2002). After an initial sympatric speciation some 3-10 million years ago (Aoyama *et al.*, 2001; Lin *et al.*, 2001) under the influence of ice sheets inducing a disjunct distribution of eels in the North-Atlantic, assortative mating and spawning diachrony probably maintained reproductive isolation. Due to the unpredictability in current patterns combined with an increasing migration distance and duration after recolonisation, a secondary contact between the species may have arisen. In this sense, eels cover an above average distance (6,000 km) to reach the spawning grounds, which is likely to increase the chance for temporal overlap between species. On the other hand, a recent re-evaluation of sympatric speciation alleged that reproductive isolation mechanisms could arise with a significant amount of dispersal/gene flow between diverging species (Mallet, 2005). Given the active component of larval migration of both North-Atlantic species, reinforcement may be very strong to maintain the East-West migrational cue of larvae, with incidental hybridisation occurring.

Despite the peak spawning season being in February for the American eel and in April for the European eel, larval surveys indicate a large overlap in hatching larvae (7 mm) of American and European eel in the Sargasso Sea (McCleave *et al.*, 1993). The observation of a hybrid zone with a bimodal pattern of dominating parental genotypes and only occasional hybrids points to a combination of spatio-temporal disjunct spawning, pre-zygotic barriers (behavioural control increases assortative mating between conspecifics) and post-zygotic barriers (strong selection against hybrids).

Do Icelandic hybrid eels represent an evolutionary dead end ?

The paradigm of why hybrids are mainly found in Iceland remains. Although the North Atlantic Drift passes west of the British Isles, no hybrids were found in Ireland, which might

point to an active process of trans-Atlantic dispersal. Studies on birds show that intermediate migratory behaviour of hybrids is mostly maladaptive and decreases overall fitness (Berthold, 1988). This explanation holds for the Icelandic eels, possibly showing an active intermediate migration by bilateral genetic imprinting. Admixed individuals found elsewhere in Europe are defined as second-generation hybrids, most likely originating from backcrosses with Icelandic F_1 hybrids. Selection must thus act on pre-zygotic barriers and subsequent hybrids fitness to maintain the species integrity, doomed to be lost if intermediate genotypes survive equally well.

ACKNOWLEDGMENTS

Research was funded by the European project EELREP (Q5RS-2001-01836). Samples were kindly provided by B. Jonsson, R. Poole, J. Van Willigen, P. Elie, A. Crivelli, A. Yahyaoui, G. Wippelhauser (S. Zink) and J. Aoyama. Helpful comments on earlier versions were given by T. Wirth, E. Nielsen, P. Galbusera, L. De Meester and F. Ollevier. G.E. Maes was supported by an IWT (Institute for the Promotion of Innovation by Science and Technology in Flanders) PhD scholarship. J.M. Pujolar received a postdoctoral fellowship from the Ministerio de Educación, Cultura y Deporte (Spain).