Microsatellite conservation and Bayesian individual assignment in four *Anguilla* species

G. E. Maes[1,*], J. M. Pujolar[1], J. A. M. Raeymaekers[1], J. Dannewitz[2], F. A. M. Volckaert[1]

[1]Laboratory of Aquatic Ecology, Katholieke Universiteit Leuven, Ch. de Bériotstraat, 32, 3000 Leuven, Belgium
[2]Institute of Freshwater Research, National Board of Fisheries, 178 93 Drottningholm, Sweden

ABSTRACT: Microsatellite flanking regions are often highly conserved in fish taxa, enabling their application in other species within or outside the source family. Moderately variable microsatellite markers may also be good candidates for species identification using multi-locus genotypes. We evaluated the degree of conservation of microsatellite flanking regions and the level of polymorphism in 4 commercially important eel species (*Anguilla anguilla*, *A. rostrata*, *A. japonica* and *A. marmorata*). Using multiplex polymerase chain reactions developed for the first 2 taxa, we assessed the discrimination power of an individual-based assignment method to differentiate all 4 species without initial species information. Detection and classification of each species was performed with high confidence (>90%), as was assignment of randomly sampled individuals to pre-defined species (>95%). Our results demonstrate the highly conserved nature of microsatellites and their level of polymorphism in *Anguilla* species. Although an inverse relationship was found between genetic diversity and differentiation estimates, likely due to homoplasy, assignment proved to be superior to multivariate and distance-based approaches for identifying the 4 species. The method enables the rapid screening of morphologically similar eel species using only 4 co-dominant nuclear loci and the detection of natural hybridisation or anthropogenic mixing between internationally highly traded species.

KEY WORDS: Admixture · Anguillids · Flanking region · Genetic variability · Homoplasy · Multiplex-PCR · Species identification

INTRODUCTION

Species identification may be based on several methods, ranging from morphology, protein blotting and allozyme scoring to PCR-based molecular techniques (Lockley & Bardsley 2000, Avise 2004). Morphological analysis, although very useful, lacks power in the case of cryptic speciation with homoplasic traits or in species with a large overlap in morphological and meristic traits due to convergent selection (Fisher et al. 2000). Similarly, protein-based techniques, widely used for species identification, lack power to discriminate between related species (Lockley & Bardsley 2000, Avise 2004). With the advent of PCR-based techniques (such as sequencing, restriction fragment length polymorphism [RFLP], amplified fragment length polymor-

phism [AFLP] and randomly amplified polymorphic DNA analysis [RAPD]), a multitude of taxonomic uncertainties as well as technical problems have been solved. In the first place, molecular analysis at the DNA level has tremendously increased the power to differentiate (cryptic) species, by focusing on the neutral signal of drift and mutation, being largely independent of environmental influences (Avise 2004). By making the distinction between homology (traits which are identical by descent) and analogy (traits which evolved independently multiple times), molecular studies have been able to discriminate between highly related species (Avise 2004). A second advantage is the small size of many PCR-based amplification products, such as PCR single strand conformational polymorphisms (SSCP), single nucleotide polymor-

phisms (SNP), or simple sequence repeats (SSR), enabling the analysis of highly degraded, processed, or archival tissue (Buonaccorsi et al. 2001).

Mitochondrial DNA (mtDNA) has until recently been the marker of choice for taxonomical questions (Avise 2004). Nuclear DNA (nDNA), however, has several advantages over maternally inherited and haploid mtDNA, such as the capacity to detect biparentally inherited polymorphisms and recent/past species admixture (Buonaccorsi et al. 2001, Ludwig et al. 2003).

Microsatellites might be good candidates for identification purposes, due to their high variability, co-dominant diploid inheritance, high intraspecific discrimination power at small and large geographical scales and, most importantly, cross-species amplification (Rico et al. 1996, Manel et al. 2005). Cross-species amplification between more-or-less distant species is a consequence of highly conserved microsatellite flanking regions (Rico et al. 1996, O'Connell & Wright 1997). Amplification success and single-locus polymorphism between species is believed to be inversely correlated with phylogenetic distance (FitzSimmons et al. 1995, Estoup et al. 2002). In fish species, however, cross-species amplification success is in general only slightly influenced by interspecific genetic distance (Rico et al. 1996; but see Holmen et al. 2005). Microsatellites may be useful to detect jointly cryptic species and intraspecific relationships, but uncertainty remains regarding their utility for phylogenetic analyses because of homoplasy (Fisher et al. 2000, Estoup et al. 2002). Due to constraints in size, microsatellite alleles will back-mutate to a former state, resulting in alleles identical in state, but not by descent. This limits the genetic distance that can accrue between reproductively isolated taxa, and may lead to spurious patterns (Nauta & Weissing 1996, Fisher et al. 2000, Estoup et al. 2002). Homoplasy is expected to be high at marine fish microsatellite loci, due to the high population sizes (low effect of drift) in marine organisms and the high mutation rates of microsatellites. By using moderately variable loci, the risk of homoplasy interfering with the signal of speciation can be reduced (Estoup et al. 2002).

Given a baseline population of each species, individual-based assignment techniques using various microsatellite markers can differentiate between related species with a high level of confidence (Hansen et al. 2001, Manel et al. 2005). Individual assignment to a population can be conducted with or without initial information on the number of sampled populations (Hansen et al. 2001). In cases where the number of populations is not known, Bayesian clustering and assignment methods search for the most likely number of populations and subsequently assign each individual with high confidence (Pritchard et al. 2000, Manel et al. 2005).

A total of 15 *Anguilla* species are officially recognised, although the morphological and meristic characteristics are highly unstable even between phylogenetically distant species and it remains difficult to use them for species determination. A recent reassessment of eel morphology resulted in the detection of only 4 unambiguous groups and the detection of much overlap in formerly accepted morphological characters (Watanabe et al. 2004). Species recognition is even more problematic at the larval stage, where essential traits such as colouring and dentition characters are lacking (Tesch 2003, Watanabe et al. 2004). Freshwater eels support important fisheries worldwide, and juveniles (glass eels) as well as adults (yellow and silver eels) are harvested (Dekker 2003, Tesch 2003). Recent fisheries data indicate a dramatic decline in the recruitment abundance of 3 important eel species *A. anguilla*, *A. rostrata* and *A. japonica* since the 1980s (Haro et al. 2000, Dekker 2003, Tseng et al. 2003). In 2004, European glass eel recruitment averaged 1% of its 1960 level, and the entire species is believed to be outside safe biological limits. Although global demands for adult eels are largely met by extensive aquaculture, glass eels for consumption and restocking are still entirely dependent on natural recruitment, as profitable artificial reproduction is still lacking (Dekker 2003). Additionally, due to far-reaching translocations of non-native species for aquaculture purposes, the natural distribution of species has become disrupted. The most striking example is the observation of European eels, which escaped from culture ponds and migrated together with Japanese eels during the spawning migration in the East China Sea (Zhang et al. 1999, Okamura et al. 2002). For these reasons, an efficient and reliable species identification toolkit is critical for eel conservation and aquaculture management.

Eel taxa have been identified using various molecular techniques (Comparini & Rodino 1980, Rehbein et al. 2002, Hwang et al. 2004, Minegishi et al. 2005), and several reliable molecular tests now exist to identify *Anguilla* spp. simultaneously in processed, historical, or alcohol-preserved samples without sequencing (Lin et al. 2002, Itoi et al. 2005 and references therein). Nevertheless, nuclear markers have never been tested in eel for species identification, despite their great potential to detect pure species and their hybrids (Hansen et al. 2001, Manel et al. 2005). Microsatellites have been developed for the 3 commercially most fished species (*A. anguilla*, *A. rostrata* and *A. japonica*), and several of these markers are known to cross-amplify in related eel species (Wirth & Bernatchez 2001, Mank & Avise 2004, Dannewitz et al. 2005). Less-

related species were never considered, however, and no microsatellite-based species identification has yet successfully been tested using such markers.

Here, we report on the development and validation of microsatellite DNA multiplex PCRs on 4 eel taxa, the 2 North Atlantic species (*Anguilla anguilla*, *A. rostrata*) and the 2 Indo-Pacific species (*A. japonica*, *A. marmorata*). Our aims were: (1) to evaluate cross-species amplification, the degree of conservation and variability of available microsatellite markers in phylogenetically distant eel species; (2) to test the influence of homoplasy by evaluating the relation between allele frequency (F_{ST})- and allele size (R_{ST})-based differentiation estimators and single-locus diversity indices; (3) to screen for the most suitable microsatellite markers for a routine PCR assay; and finally (4) to test the discrimination power of a Bayesian individual assignment technique with and without prior species knowledge, in comparison to more classical discrimination techniques.

MATERIALS AND METHODS

Sampling. Three American eel (*Anguilla rostrata*, AR) samples and 3 European eel (*A. anguilla*, AA) samples were collected from different geographical locations. One Japanese eel (*A. japonica*, AJ) sample and 1 giant mottled eel (*A. marmorata*, AM) sample were collected in Taiwan (see Table 1). Sample sizes ranged from 30 to 60 individuals. Fin or liver tissue was preserved in 100% ethanol. Initial species identification was based on geographical sampling location and specific morphological traits following Tesch (2003).

DNA purification and microsatellite amplification. Purification of genomic DNA and amplification of microsatellite loci were performed following methods described in Dannewitz et al. (2005). We analysed the following 8 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). Loci designated 'AAN' or 'ANG' have *Anguilla anguilla* as source species, while all 'ARO' loci originate from *A. rostrata*. Electrophoresis and size determination of alleles was made on an LI-COR 4200 automated sequencer (LI-COR) 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).

Genetic data analysis. Genetic diversity was evaluated based on genotype and allele frequencies, using the level of polymorphism (0.95 criterion, where a locus is considered polymorphic when the frequency of the most common allele does not exceed 0.95), observed and expected heterozygosity (H_o and H_e)

and allelic richness (*R*) as criteria. Deviations from Hardy-Weinberg equilibrium (HWE) were tested using the inbreeding coefficient F_{IS} (Weir & Cockerham 1984) values calculated in GENETIX Version 4.05 (Belkhir et al. 1999). Genetic differentiation was characterised using hierarchical *F*-statistics, theta (θ, Weir & Cockerham 1984) and R_{ST}-values as implemented in the FSTAT software package Version 2.9.3.2 (Goudet et al. 1995). Significance of F_{IS} and multilocus F_{ST} was assessed with permutation tests (1000 replicates). To test for the influence of homoplasy on inter-species genetic differentiation, a regression analysis between single-locus genetic differentiation estimators (R_{ST} and F_{ST}) and expected heterozygosity (H_e) was performed using STATISTICA 6.1 (StatSoft 2004). A factorial correspondence analysis was performed in GENETIX (Belkhir et al. 1999) to plot multilocus genotypes of the 4 species in 3 dimensions. To compare the performance of distance-based and Bayesian-assignment techniques, an allele sharing distance (D_{AS}, Bowcock et al. 1994) matrix was computed from a random subset of 20 individuals per species using the microsatellite toolkit (available at http://animalgenomics.ucd. ie/sdepark/ms-toolkit/), and a neighbour-joining tree was constructed in PHYLIP Version 3.6 (Felsenstein 1995). Genetic variability was partitioned into a within-species (F_{SC}) and between-species component (F_{CT}) using an analysis of molecular variance (AMOVA) in ARLEQUIN Version 2.0 (Schneider et al. 2001). The WHICHLOCI 1.0 software (Banks et al. 2003) was used to target the most discriminating loci for species identification, by simulating 1000 new populations (N = 100) based on allele frequencies data for each species. Through trial assignments with loci, one at a time, this software ranks loci in terms of their efficiency for correct population assignment and, conversely, their propensity to cause false assignments. Subsequent trials with increasing numbers of loci determine the minimum number of specific loci needed to attain userdefined power for species assignment. We used the standard 95% stringency option for all populations, as well as the critical population option, where 1 specific population has to meet the stringency level (95% correct assignment) instead of the complete dataset. The latter option enables a high assignment success of both highly and weakly differentiated samples in 1 single test. We used a model-based clustering algorithm in order to search for the most likely number of groups (species) in the data, as implemented in the software STRUCTURE Version 2.1 (Pritchard et al. 2000). This method initially assumes no prior knowledge about population structure or species identity. The software organises individuals into a predefined number of clusters (*K*), which may represent putative populations or species, and provides loglikelihood val-

ues for different *K*s. These analyses were performed with $1 < K < 10$ to account for population structure within species, initially using the non-admixture model in combination with the non-correlated allele frequency model (high differentiation expected between 'pure' species). To avoid bias, a burn in length of 10^4 iterations and, subsequently, 10^5 additional Monte Carlo Markov chain (MCMC) iterations were performed, as recommended by Pritchard et al. (2000). Each assessment of *K* was repeated 3 times to check the repeatability of the results. The most likely *K* was then used to assign each individual. Assignment scores of individuals to their respective clusters/populations (*q*) were plotted, and the proportion of correctly assigned individuals (with $q > 0.80$) was calculated. Finally, a random set of 100 individuals was drawn from the original dataset; the remaining individuals (N = 299) were given a species tag based on morphological and genetic criteria and used as a prior group in the STRUCTURE analysis (discarding <75% complete genotypes). Subsequently, the unknown samples were assigned probabilistically to 1 of the 4 predefined species clusters by using the non-admixture as well as the admixture model to account for possible hybridisation. Individuals with $q < 80\%$ were considered as misclassified.

RESULTS

Cross-species amplification and genetic diversity within *Anguilla* spp.

All 8 microsatellites developed for the North Atlantic species *Anguilla anguilla* (AA) and *A. rostrata* (AR), amplified successfully in *A. japonica* (AJ) and *A. mar-morata* (AM). Only 1 locus was monomorphic in AJ (AAN 01). Most diversity indices (allelic richness, observed heterozygosity and mean allele length) were concordant between species for each locus (Fig. 1). Genetic variability was high at all 8 loci, ranging from 9 (AAN 03) to 71 alleles (ARO 063). Observed heterozygosity per population and per locus ranged from 2 (AAN 03 in *A. japonica*) to 93% (ARO 063 in *A. anguilla*), and H_e per population over all loci ranged from 0.59 to 0.80 (Table 1; see also Appendix 1, available as supplementary material at www.int-res.com/articles/suppl/m319p251_app.pdf). Allelic richness (*R*) ranged from 10.42 in AM to 13.42 in AR. The difference between observed and expected heterozygosity was mainly due to locus ANG 075, showing strong deviation from HWE and known to exhibit null alleles (Dannewitz et al. 2005). Due to its low influence on genetic differentiation (see following subsection), we initially decided to keep this locus in the genetic variability and multilocus dataset. On a large scale, the Indo-Pacific taxa (AJ and AM) exhibited a significantly lower heterozygosity than the North Atlantic taxa (permutation test FSTAT: p = 0.03).

Genetic differentiation between *Anguilla* spp. and assignment power

Genetic differentiation was high and significant between all species (overall $F_{ST} = 0.186$, $R_{ST} = 0.315$). A locus-by-locus AMOVA confirmed the high interspecific component of genetic variation (18 versus 0.6% between populations within species) (Table 2). The loci exhibiting most differentiation among species were moderately polymorphic, namely AAN 01, AAN 03, AAN 05 and ANG 151 (Table 2). A factorial analysis

Table 1. *Anguilla* spp. Sampling details and summary of genetic variability at 8 microsatellite loci for 4 *Anguilla* taxa. Lat: latitude; Lon: longitude; life stages (LS) — G: glass eel; S: silver eel; N: number of samples; H_e: expected heterozygosity; H_o: observed heterozygosity; p$_{(0.95)}$: level of polymorphism at the 95% level; *R*: allelic richness. For an explanation of codes see 'Materials and methods'

Sampling site and country	Lat.	Long.	Code	LS	Year	N	H_e	H_o	p$_{(0.95)}$	*R*
A. anguilla										
Burrishoole system (Ireland)	53° 55′ N	09° 55′ W	AA1	G	2001	60	0.748	0.688	1	11.40
St. Nazaire, Loire (western France)	47° 12′ N	01° 44′ W	AA2	G	2001	60	0.773	0.725	1	11.86
Tour-Du-Valat, Rhône (southern France)	43° 33′ N	04° 38′ E	AA3	G	2001	60	0.763	0.648	1	11.86
A. rostrata										
St. Johns River, Florida (USA)	27° 12′ N	80° 13′ W	AR1	S	1999	30	0.779	0.740	1	13.42
Musquash River, New Brunswick (Canada)	45° 11′ N	66° 19′ W	AR2	G	1995	47	0.783	0.690	1	12.22
West Harbour Pond, Maine (USA)	43° 59′ N	69° 50′ W	AR3	G	2003	61	0.796	0.702	1	13.27
A. japonica										
River Tanshui (Taiwan)	25° 10′ N	121° 25′ E	AJ	G	1990	51	0.590	0.467	0.75	12.08
A. marmorata										
Shuang-Hsi (Taiwan)	25° 01′ N	121° 51′ E	AM	G	1994	30	0.609	0.454	0.87	10.42
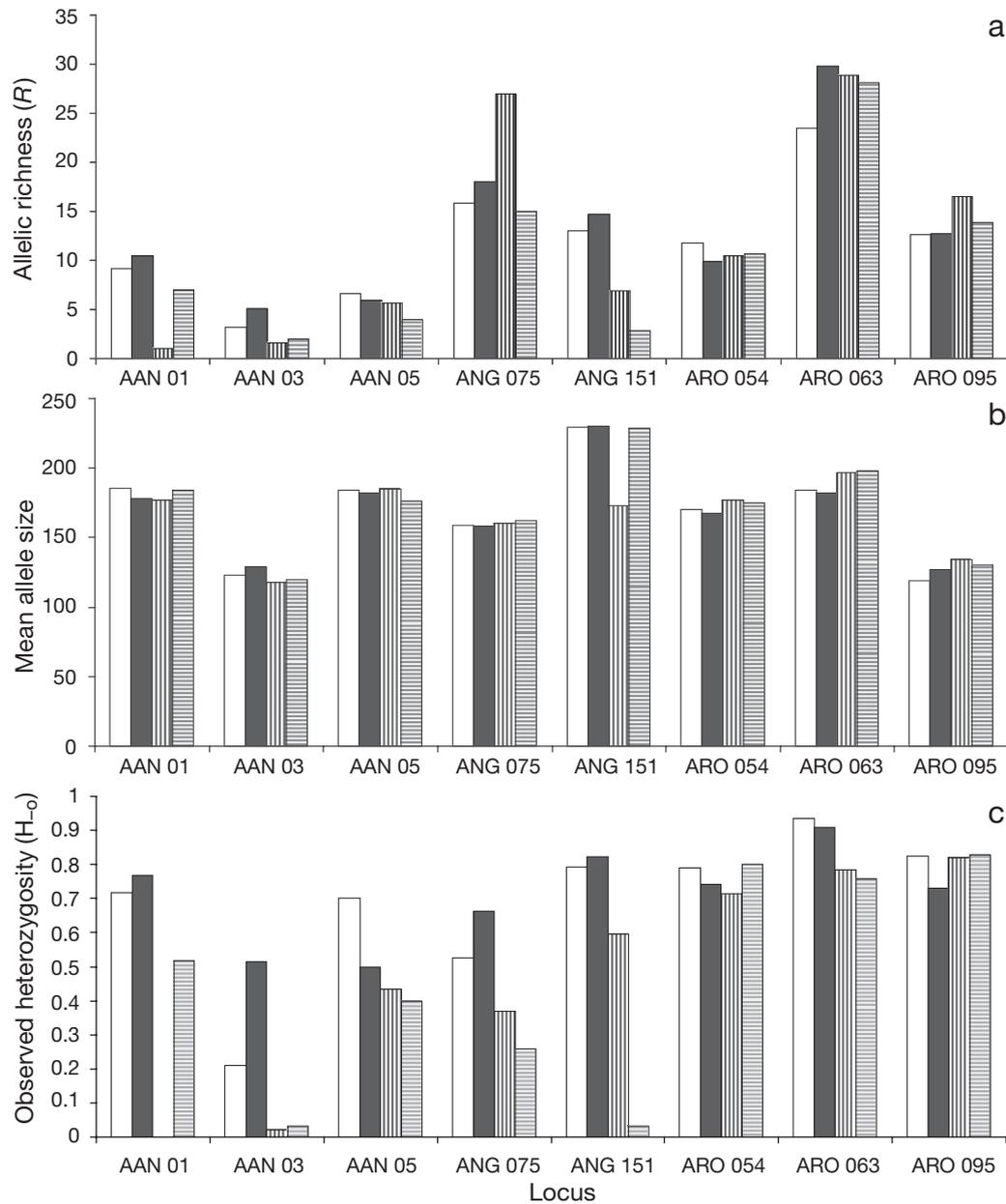
Fig. 1. *Anguilla* spp. Comparison of loci among 4 *Anguilla* species: (a) allelic richness, (b) mean allele size and (c) observed heterozygosity. White bars: *A. anguilla*; black bars: *A. rostrata*; vertically striped bars: *A. japonica*; horizontally striped bars: *A. marmorata*. For an explanation of locus abbreviations see 'Materials and methods'

split all genotypes into 4 distinct clusters based on all 8 loci (Fig. 2) and on only the 4 most discriminating ones (data not shown). Both North Atlantic species are adjacent, while the other 2 species are much more distinct. The additional information of the 4 remaining highly polymorphic loci is limited, as the inertia on the first axis is 41 and 38% for 8 and 4 loci, respectively. The allele sharing distance tree shows 4 clusters, 3 fairly pure species clusters (AA, AJ and AR), but 1 cluster containing individuals from AM and AR (Fig. 3). Using only 4

loci, similar clusters are formed (data not shown). There was a strong negative correlation between inter-species pairwise F_{ST} and expected heterozygosity per locus ($r = -0.71$ to -0.95; $p < 0.05$), and inter-species F_{ST} and allelic richness per locus ($r = -0.67$ to -0.82; $p < 0.05$). No significant relationship was found between R_{ST} and any diversity indices. Two loci were shown to be highly discriminative between all 4 species after simulating 1000 new populations using observed allele frequencies. Loci AAN 05 and AAN 01 enabled high

assignment success at the 95% stringency level, with ranking scores of 88.5 and 86.5%, respectively. When defining *A. marmorata* as the critical population (stringency of 95% for the least well assigned population instead of over all populations), locus ANG 151 proved to be necessary for correct assignment. The other loci showed a lower discrimination power, locus AAN 03 being highly discriminative between the North Atlantic species, but not for other species (between 46 and 73% accuracy). Further analyses were performed with only the 4 most discriminating loci (highest F_{CT}) from the AMOVA and the WHICHLOCI analysis.

Cluster and assignment analysis

The Bayesian MCMC cluster analysis showed a continuous increase in likelihood, but levelled off at $K = 4$, splitting the dataset into 4 homogeneous clusters, using 4 loci. Pritchard et al. (2000) advise choosing the lowest value when likelihood values are approaching a maximum. At $K = 5$, the method splits 1 species' (AA) genotypes into 2 clusters of individuals with equal *q*-values (0.50), which is highly unlikely. After removal of incomplete genotypes (>25% missing), assignment proportions of morphological species to 1 of the 4 clusters ranged from 0.96 to 1 (Table 3a, Fig. 4). On average >96% of all individuals were assigned with high confidence to 1 specific cluster (further called genetic species), except for some AR individuals (9%). Two AR individuals were

Table 2. *Anguilla* spp. Locus-by-locus analysis of molecular variance among *Anguilla* species (AS, F_{CT}), within species among populations (WSAP, F_{SC}) and within populations (WP). Global fixation indices per locus are given as allele frequency (F_{ST}) and allele size (R_{ST}). Most differentiating loci are listed in **bold** (**p < 0.01, *p < 0.05). For an explanation of locus abbreviations see 'Materials and methods'

Locus	% AS	% WSAP	% WP	F_{CT}	F_{SC}	F_{ST}	R_{ST}
AAN 01	29.865	0.370	69.764	0.299**	0.0052	0.302**	0.714
AAN 03	62.387	0.183	37.430	0.624**	0.0049	0.626**	0.210
AAN 05	36.871	–0.224	63.353	0.369**	–0.0036	0.366**	0.639
ANG 075	0.975	1.078	97.947	0.009*	0.0109**	0.020**	0.343
ANG 151	15.635	0.069	84.295	0.156**	0.0008	0.157**	0.066
ARO 054	3.711	0.174	96.115	0.037**	0.0018	0.039**	0.003
ARO 063	0.849	1.733	97.418	0.008	0.0175**	0.026**	0.290
ARO 095	1.304	0.606	98.088	0.013*	0.0061**	0.019**	0.088
Total	0.181	0.005	0.814	0.180**	0.006**	0.186**	0.315
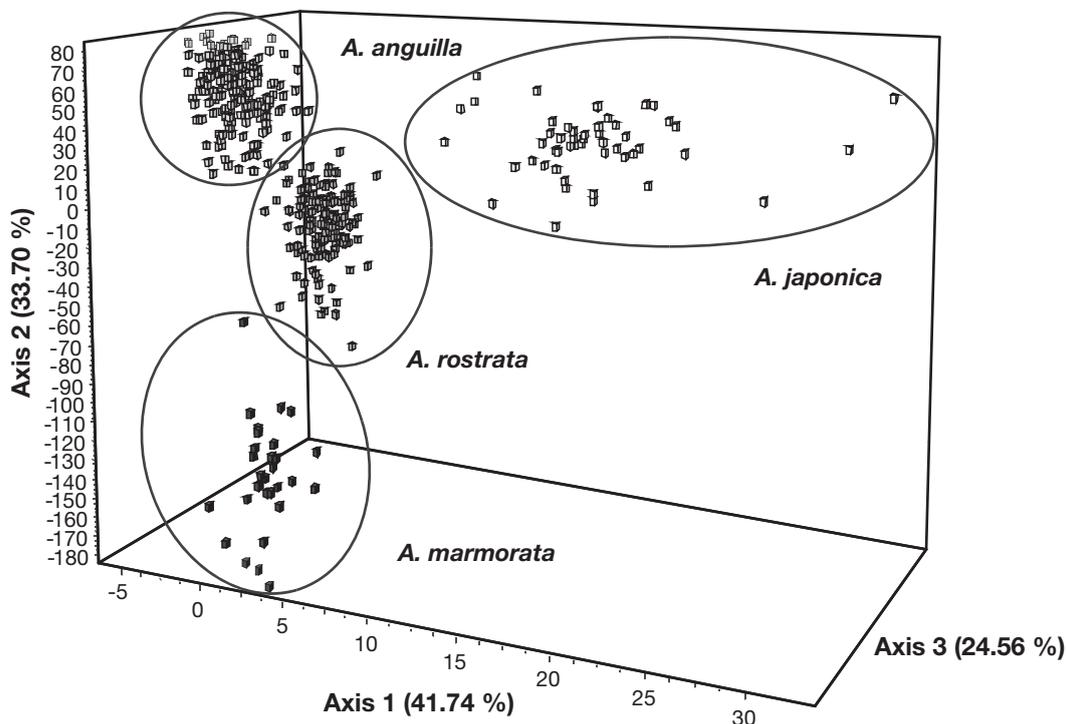
Fig. 2. *Anguilla* spp. Factorial analysis of correspondence of *Anguilla* individuals from 4 species using 8 loci. Each point represents a multilocus genotype and each type of shading a morphological species tag. Light grey: *A. anguilla*; dark grey: *A. rostrata*; white: *A. japonica*; black: *A. marmorata*; clusters have been circled for clarity

completely or partially assigned to AA, while some individuals showed a signal of admixture with AM (Table 3). The accuracy of species identification was tested using as baseline data individuals with matching morphological and genetic identity from the initial assignment analysis, while assigning unknown individuals to 1 of the 4 genetic species. Most unknown

randomly drawn individuals (N = 100) were correctly assigned to their putative species of origin (Table 3b). Only AR individuals were not all assigned to their own species (success rate of 97%). Using the admixture or non-admixture model, 1 individual showed a *q* score of 0.50 to AR and AM; 2 other individuals showed a *q* score of 80%.
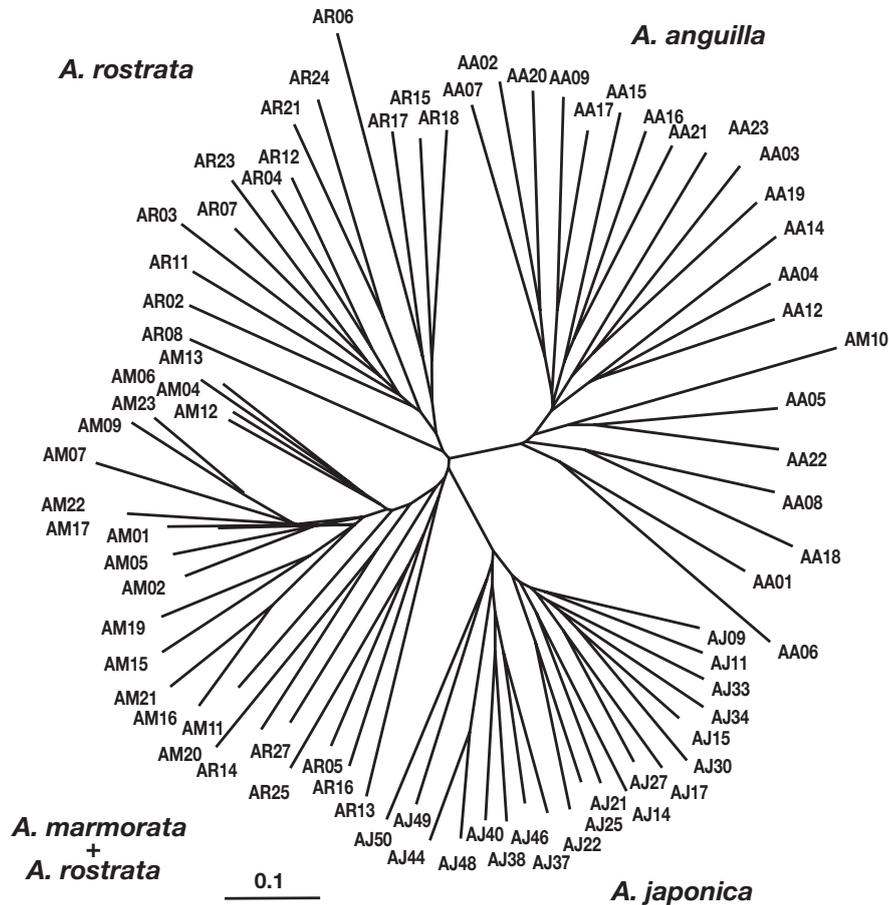
Fig. 3. *Anguilla* spp. Allele sharing distance (D_{AS}) neighbour-joining tree of a subset of 20 random individuals per *Anguilla* species. AA: *A. anguilla*; AR: *A. rostrata*; AJ: *A. japonica*; AM: *A. marmorata*

Table 3. *Anguilla* spp. (a) Proportion of individuals assigned to each cluster (*K* = 4) without prior species information and (b) assignment success of randomly drawn eel individuals to genetically defined eel species. An individual is considered correctly assigned if *q* > 0.80. AA: *A. anguilla*; AR: *A. rostrata*; AJ: *A. japonica*; AM: *A. marmorata*; U: undefined samples

	N	Proportion ind. assigned to each cluster				Assignment success (%)
(a) No prior species information		Cluster 1	Cluster 2	Cluster 3	Cluster 4	
AA	180	1.000	0.000	0.000	0.000	98.25
AR	138	0.012	0.961	0.000	0.026	90.98
AJ	51	0.000	0.000	1.000	0.000	100
AM	30	0.000	0.000	0.000	1.000	96.67
(b) Random samples		AA	AR	AJ	AM	
AA-U	36	1.000	0.000	0.000	0.000	100
AR-U	34	0.005	0.973	0.000	0.022	97.10
AJ-U	20	0.000	0.000	1.000	0.000	100
AM-U	10	0.000	0.000	0.000	1.000	100
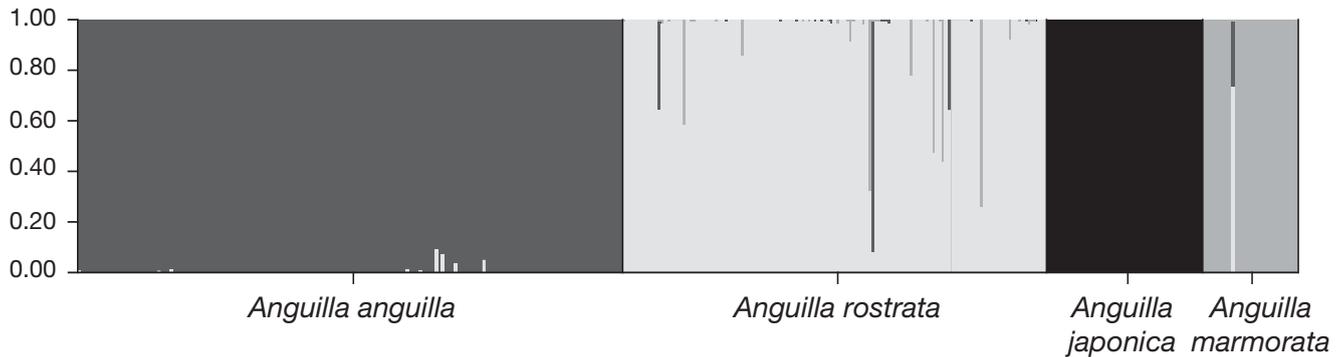
Fig. 4. *Anguilla* spp. Results of STRUCTURE clustering and assignment analysis (Pritchard et al. 2000). Bar plot of assignment score for 8 populations from 4 morphologically defined eel species (4 types of shading). Each bar represents 1 individual (dark grey: *A. anguilla*; light grey: *A. rostrata*; black: *A. japonica*; medium grey: *A. marmorata*)

DISCUSSION

Conservation of microsatellite flanking region and polymorphism in *Anguilla* spp.

This study shows the highly conserved nature of microsatellite flanking regions in *Anguilla* species, despite their evolutionary distance. All 8 microsatellites amplified in the 4 eel species, with only 1 locus (in Japanese eel) being monomorphic. The 100% success in our study is concordant with the observations of Rico et al. (1996), who were able to amplify 17 loci (94%) in various fish species based on primers for 18 microsatellite loci originally isolated from whiting *Merlangius merlangus*, threespine stickleback *Gasterosteus aculeatus* and cod *Gadus morhua*, each representing the 3 main superclasses and superorders of fish. Other studies have reported lower cross-species amplification success. For example, Holmen et al. (2005) tested 120 microsatellite primer combinations developed for *Danio rerio* and *Campostoma anomalum* in 7 cyprinids. Amplification products were recorded in only 29 to 41% of the loci, and 6 to 23% of the loci exhibited polymorphism.

The high cross-species amplification success reported in *Anguilla* is in line with phylogenetic studies, suggesting that the genus is either of recent origin or is characterised by a slow evolutionary rate (Minegishi et al. 2005). Given the discrepancy between eel divergence time, ranging from 60 to 20 Mya (million years ago), but the existence of fossil records of Anguilliformes dating from the Upper Cretaceous (Aptian, about 113 to 119 Mya), a slow evolutionary rate thus seems more likely as an explanation of these results (Minegishi et al. 2005).

In our study, locus polymorphism was very similar between phylogenetically related or distant species (genetic distances ranging from 9 to 26% based on D-loop sequences; Ishikawa et al. 2004), enabling the direct comparison of variability estimates between loci and species. Genetic variability was high in all species; however, there was a clear difference between the Indo-Pacific and North Atlantic species. Although the Indo-Pacific species showed similar diversity values to the North Atlantic species, the observed and expected heterozygosity were significantly lower (24 and 17% differences, respectively) in the former group. A likely explanation for this is a lower effective population size than North Atlantic species, due to strong overfishing, a lower habitat size and population structuring. The first 2 reasons most likely play a role in *Anguilla japonica*, which is known to have declined dramatically over the last decades (Tesch 2003, Tseng et al. 2003). The European eel has indeed the largest 'migration' loop among anguillids (Ishikawa et al. 2004), enabling the maintenance of a larger population size and higher genetic variability than other anguillid species. In the case of the giant mottled eel *A. marmorata*, its peculiar geographical distribution (from eastern Africa to southern Japan) and multiple population structure suggest that the species' genetic variability is distributed among various habitats (Ishikawa et al. 2004).

The power of individual assignment versus classical methods in *Anguilla* spp.

Our study shows the potential of moderately variable microsatellite DNA markers in discriminating highly traded species with a high level of confidence. Here, we provide a powerful tool to identify *Anguilla* species using only 4 microsatellite loci: AAN 01, AAN 03, AAN 05 and ANG 151. Although the discrimination power for genetic assignment increases with the number of loci used (Hansen et al. 2001), some loci may fail to provide additional information due to the lack of polymorphism, significant measurement errors, high correlation with other loci (e.g. linkage disequilibrium), or

homoplasy. In our study, by removing the 4 loci showing low divergence values (ARO 095, ARO 054, ARO 063 and ANG 075 with $F_{ST} < 5\%$), a false signal of admixture between distant species can be avoided, while retaining a high assignment power.

The influence of homoplasy on genetic differentiation has been shown in several studies on various organisms (Estoup et al. 2002), including marine fish (O'Reilly et al. 2004) and eels (Mank & Avise 2004). For instance, O'Reilly et al. (2004) demonstrated an inverse relationship between overall F_{ST} and single-locus heterozygosity values, a pattern attributed to homoplasy at highly variable loci due to length restriction (Estoup et al. 2002). We found a significant negative correlation between inter-species pairwise F_{ST} and both expected heterozygosity and allelic richness per locus, which suggests that moderately variable loci suffer less from interconversions between a limited number of alleles than highly variable ones, with a higher mutation rate (Mank & Avise 2004). This explains the results of Mank & Avise (2004), who used highly variable loci (30 to 57 alleles) but failed to discriminate hybrids between the 2 North Atlantic eel species. The high F_{ST} values in our study enabled the discrimination of each species without any initial geographical or species information. No relationship was found between the allele-size-based differentiation estimator (R_{ST}) and genetic variability, possibly due to the higher variance associated with this estimator (Estoup et al. 2002). Classical multivariate and distance methods generally performed less well than Bayesian methods. Using a factorial correspondence analysis (FCA), 4 clusters corresponding to the 4 species were clearly visible, although the first-axis inertia increased by only 3% when using all 8 loci. A graphical view, however, does not provide a usable assignment success proportion that would enable power assessment. The performance of the allele sharing distance tree was even poorer, and, although a rough clustering of the 4 species was provided, part of the AR individuals clustered into the AM group. This illustrates some shortcomings of these methods, such as the absence of distinctive groups without information on their origin and the lack of a statistical framework for individuals misclassified by chance (Pritchard et al. 2000). Another important shortcoming is that it is impossible to use baseline information gathered on a species/population to help classify unknown individuals into pre-defined clusters (Pritchard et al. 2000).

Multilocus assignment methods versus single-locus identification in *Anguilla* spp.

Our assignment approach seems as fast and reliable as recent eel identification protocols based on mtDNA

(Lin et al. 2002, Itoi et al. 2005 and references therein), but enables the detection of additional information, such as sex-biased interbreeding or hybridisation, including a power analysis on species discrimination. No sequencing or cloning is needed, and, given a baseline population, assignment scores are extremely high (>97%). Earlier studies based on reliable PCR-SSCP or RT-PCR techniques required a complete diagnostic marker for each species to classify individuals, and focused mainly on 2 species (Lin et al. 2002, Rehbein et al. 2002, Hwang et al. 2004, Itoi et al. 2005 and references therein). The crucial requirements for accurate assignment scores to populations or species are a moderate genetic differentiation ($F_{ST} > 5\%$) and some initial knowledge of genetic variability at each locus, to choose loci less prone to homoplasy (Hansen et al. 2001, Estoup et al. 2002, Manel et al. 2005).

Interestingly, *Anguilla rostrata* seemed to show some affinities with *A. anguilla* and *A. marmorata*, as 1.2 and 2.6% of the individuals showed a signal of admixture, respectively. Possible reasons can be either natural hybridisation between adjacently distributed species or the anthropogenic introduction of a foreign species in the range of indigenous eel species. The first reason is very likely between AA and AR, as genetic leakage between both species has been suggested to occur (Avise et al. 1990, Maes 2005). The second reason is unlikely for the AR–AM affinity, because of their distant distribution and the unlikely introduction of AM in American waters. An alternative explanation may be the lack of power to discriminate both species without prior knowledge. However, when using baseline species information and testing undefined samples (u), we obtained an assignment score >95% (see Table 3), confirming the strength of the Bayesian cluster method when using a baseline population.

Microsatellite assignment techniques for species identification are very promising, although a case-to-case assessment should be performed to extrapolate the present results to other species. In eel, despite some indications of introgression, our results were in agreement with morphological and geographical information (97% correct identification). For species showing no external morphological differences at the adult or larval stage, which is typical for many anguillids, the reliable discrimination of such species opens many possibilities to study hybridisation and translocations. Highly traded species are often released accidentally or intentionally in the natural environment. The example of the European eel occurring at 31% in some Japanese rivers (Zhang et al. 1999, Okamura et al. 2002, Tesch 2003) highlights the need for identification of possible hybrids between morphologically almost indistinguishable eel species. As both species are even known to co-migrate (Okamura et al. 2002) and form

viable hybrids in aquaculture (Okamura et al. 2004), the method proposed here may help eel conservation and aquaculture management, by rapidly and reliably detecting translocated individuals and possible genetic admixture with indigenous species.

LITERATURE CITED

Avise JC (2004) Molecular markers, natural history and evolution. Sinauer Associates, Sunderland

Avise JC, Nelson WS, Arnold J, Koehn RK, Williams GC, Thorsteinsson V (1990) The evolutionary genetic status of Icelandic eels. Evolution 44:1254–1262

Banks MA, Eichert W, Olsen JB (2003) Which genetic loci have greater population assignment power? Bioinformatics 19:1436–1438

Belkhir K, Borsa P, Goudet J, Bonhomme F (1999) Genetix 4.05: logiciel sous Windows pour la génétique des populations. Laboratoire Génome & Population, CNRS-UPR, Université de Montpellier II

Bowcock AM, Ruizlinares A, Tomfohrde J, Minch E, Kidd JR, Cavalli-Sforza LL (1994) High-resolution of human evolutionary trees with polymorphic microsatellites. Nature 368:455–457

Buonaccorsi VP, Mcdowell JR, Graves JE (2001) Reconciling patterns of inter-ocean molecular variance from four classes of molecular markers in blue marlin (*Makaira nigricans*). Mol Ecol 10:1179–1196

Comparini A, Rodino E (1980) Electrophoretic evidence for two species of *Anguilla* leptocephali in the Sargasso Sea. Nature 287:435–437

Daemen E, Cross T, Ollevier F, Volckaert FAM (2001) Analysis of the genetic structure of European eel (*Anguilla anguilla*) using microsatellite DNA and mtDNA markers. Mar Biol 139:755–764

Dannewitz J, Maes GE, Johansson L, Wickström H, Volckaert FAM, Jarvi T (2005) Panmixia in the European eel: a matter of time. Proc R Soc Lond B Biol Sci 272:1129–1137

Dekker W (2003) Did lack of spawners cause the collapse of the European eel, *Anguilla anguilla*? Fish Manage Ecol 10:365–376

Estoup A, Jarne P, Cornuet JM (2002) Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. Mol Ecol 11:1591–1604

Felsenstein J (1995) Phylogeny inference package (PHYLIP) 3.57c. University of Washington, Seattle

Fisher MC, Koenig G, White TJ, Taylor JW (2000) A test for concordance between the multilocus genealogies of genes and microsatellites in the pathogenic *Fungus coccidioides immitis*. Mol Biol Evol 17:1164–1174

FitzSimmons NN, Moritz C, Moore SS (1995) Conservation and dynamics of microsatellite loci over 300 million years of marine turtle evolution. Mol Biol Evol 12:432–440

Goudet J (1995) FSTAT (Version 1.2) a computer program to calculate *F*-statistics. J Hered 86:485–486

Hansen MM, Kenchington E, Nielsen EE (2001) Assigning individual fish to populations using microsatellite DNA markers. Fish Fish 2:93–112

Haro A, Richkus W, Whalen K, Hoar A, Busch WD, Lary S, Brush T, Dixon D (2000) Population decline of the American eel: implications for research and management. Fisheries 25:7–16

Holmen J, Vollestadt LA, Jakobsen KS, Primmer CR (2005) Cross-species amplification of zebrafish and central stoneroller microsatellite loci in six other cyprinids. J Fish Biol 66:851–859

Hwang DF, Jen HC, Hsieh YW, Shiau CY (2004) Applying DNA techniques to the identification of the species of dressed toasted eel products. J Agric Food Chem 52:5972–5977

Ishikawa S, Tsukamoto K, Nishida M (2004) Genetic evidence for multiple geographic populations of the giant mottled eel *Anguilla marmorata* in the Pacific and Indian Oceans. Ichthyol Res 51:343–353

Itoi S, Misako N, Kaneko G, Kondo H, Sezaki K, Watabe S (2005) Rapid identification of eels *Anguilla japonica* and *A. anguilla* by polymerase chain reaction with single nucleotide polymorphism-based specific probes. Fish Sci 71(6):1356–1364

Lin YS, Poh YP, Lin SM, Tzeng CS (2002) Molecular techniques to identify freshwater eels: RFLP analyses of PCR-amplified DNA fragments and allele-specific PCR from mitochondrial DNA. Zool Stud 41:421–430

Lockley AK, Bardsley RG (2000) DNA-based methods for food authentication. Trends Food Sci Technol 11:67–77

Ludwig A, Congiu L, Pitra C, Fickel J, Gessner J, Fontana F, Patarnello T, Zane L (2003) Non-concordant evolutionary history of maternal and paternal lineages in Adriatic sturgeon. Mol Ecol 12:3253–3264

Maes GE (2005) Evolutionary consequences of a catadromous life-strategy on the genetic structure of European eel (*Anguilla anguilla* L.). PhD thesis, Katholieke Universiteit, Leuven

Manel S, Gaggiotti OE, Waples RS (2005) Assignment methods: matching biological questions with appropriate techniques. Trends Ecol Evol 20:136–142

Mank JE, Avise JC (2004) Individual organisms as units of analysis: Bayesian-clustering alternatives in population genetics. Genet Res 80: 135–143

Minegishi Y, Aoyama J, Inoue JG, Miya M, Nishida M, Tsukamoto K (2005) Molecular phylogeny and evolution of the freshwater eel genus *Anguilla* based on the whole mitochondrial genome sequences. Mol Phylogenet Evol 34:134–146

Nauta MJ, Weissing FJ (1996) Constraints on allele size at microsatellite loci: implications for genetic differentiation. Genetics 143:1021–1032

O'Connell M, Wright JM (1997) Microsatellite DNA in fishes. Rev Fish Biol Fish 7:331–363

Okamura A, Yamada Y, Mikawa N, Tanaka S, Oka HP (2002) Exotic silver eels *Anguilla anguilla* in Japanese waters: seaward migration and environmental factors. Aquat Living Resour 15:335–341

Okamura A, Zhang H, Utoh T, Akazawa A and 5 others (2004) Artificial hybrid between *Anguilla anguilla* and *A. japonica*. J Fish Biol 64:1450–1454

O'Reilly PT, Canino MF, Bailey KM, Bentzen P (2004) Inverse relationship between *f*-st and microsatellite polymorphism in the marine fish, walleye pollock (*Theragra chalcogramma*): implications for resolving weak population structure. Mol Ecol 13:1799–1814

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Rehbein H, Sotelo CG, Perez-Martin RI, Chapela-Garrido MJ and 7 others (2002) Differentiation of raw or processed eel by PCR-based techniques: restriction fragment length polymorphism analysis (RFLP) and single strand conformation polymorphism analysis (SSCP). Eur Food Res Technol 214:171–177

Rico C, Rico I, Hewitt G (1996) 470 million years of conservation of microsatellite loci among fish species. Proc R Soc Lond B Biol Sci 263:549–557

Schneider S, Kueffer JM, Roessli D, Excoffier L (2001) Arlequin: a software for population genetics data analysis, Version 2.0.

StatSoft (2004) STATISTICA (data analysis software system), Version 6.0. Available at www.statsoft.com

Tesch FW (2003) The eel. Blackwell Science, Oxford

Tseng MC, Tzeng WN, Lee SC (2003) Historical decline in the Japanese eel *Anguilla japonica* in northern Taiwan inferred from temporal genetic variations. Zool Stud 42: 556–563

Watanabe S, Aoyama J, Tsukamoto K (2004) Reexamination of Ege's (1939) use of taxonomic characters of the genus *Anguilla*. Bull Mar Sci 74:337–351

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population-structure. Evolution 38:1358–1370

Wirth T, Bernatchez L (2001) Genetic evidence against panmixia in the European eel. Nature 409:1037–1040

Zhang H, Mikawa N, Yamada Y, Horie N, Okamura A, Utoh T, Tanaka S, Motonobu T (1999) Foreign eel species in the natural waters of Japan detected by polymerase chain reaction of mitochondrial cytochrome *b* region. Fish Sci 65:684–686