MOLECULAR ECOLOGY

Genome-wide single-generation signatures of local selection in the panmictic European eel

J. M. PUJOLAR,* M. W. JACOBSEN,* T. D. ALS,†‡ J. FRYDENBERG,* K. MUNCH,§ B. JÓNSSON,¶ J. B. JIAN,** L. CHENG,†† G. E. MAES,‡‡§§ L. BERNATCHEZ¶¶ and M. M. HANSEN*
*Department of Bioscience, Aarhus University, Ny Munkegade 114, Bldg. 1540, DK-8000 Aarhus C, Denmark, †National Institute of Aquatic Resources, Technical University of Denmark, Vejlsøvej 39, DK-8600 Silkeborg, Denmark, ‡Department of Biomedicine-Human Genetics, Aarhus University, DK-8000 Aarhus C, Denmark, §Bioinformatics Research Center, Aarhus University, DK-8000 Aarhus C, Denmark, ¶Biopol, Marine Biology and Biotechnology Center, Einbúastígur 2, IS545 Skagastrond, Iceland, **BGI-Shenzhen, Beishan Industrial Zone, Main Building, Yantian District, 518083 Shenzhen, China, ††BGI-Europe, Copenhagen Bio Science Park, Ole Maaløes Vej 3, DK-2200 Copenhagen, Denmark, ‡‡Laboratory of Biodiversity and Evolutionary Genomics, University of Leuven (KU Leuven), Deberiotstraat 32, B-3000 Leuven, Belgium, §§Centre for Sustainable Tropical Fisheries and Aquaculture, School of Marine and Tropical Biology, James Cook University, Townsville, Qld 4811, Australia, ¶¶IBIS (Institut de Biologie Intégrative et des Systèmes), Université Laval, Québec City, QC, Canada G1V 0A6

Abstract

Next-generation sequencing and the collection of genome-wide data allow identifying adaptive variation and footprints of directional selection. Using a large SNP data set from 259 RAD-sequenced European eel individuals (glass eels) from eight locations between 34 and 64°N, we examined the patterns of genome-wide genetic diversity across locations. We tested for local selection by searching for increased population differentiation using F_{ST}-based outlier tests and by testing for significant associations between allele frequencies and environmental variables. The overall low genetic differentiation found ($F_{ST} = 0.0007$) indicates that most of the genome is homogenized by gene flow, providing further evidence for genomic panmixia in the European eel. The lack of genetic substructuring was consistent at both nuclear and mitochondrial SNPs. Using an extensive number of diagnostic SNPs, results showed a low occurrence of hybrids between European and American eel, mainly limited to Iceland (5.9%), although individuals with signatures of introgression several generations back in time were found in mainland Europe. Despite panmixia, a small set of SNPs showed high genetic differentiation consistent with single-generation signatures of spatially varying selection acting on glass eels. After screening 50 354 SNPs, a total of 754 potentially locally selected SNPs were identified. Candidate genes for local selection constituted a wide array of functions, including calcium signalling, neuroactive ligand–receptor interaction and circadian rhythm. Remarkably, one of the candidate genes identified is PERIOD, possibly related to differences in local photoperiod associated with the >30° difference in latitude between locations. Genes under selection were spread across the genome, and there were no large regions of increased differentiation as expected when selection occurs within just a single generation due to panmixia. This supports the conclusion that most of the genome is homogenized by gene flow that removes any effects of diversifying selection from each new generation.

Keywords: anguilla, local adaptation, panmixia, spatially varying selection

Received 20 January 2014; revision received 11 April 2014; accepted 11 April 2014

Correspondence: Jose M. Pujolar, Fax: +45 89422722;
E-mail: jmartin@biology.au.dk

Introduction

Identifying which regions of the genome are under selection is essential for understanding the selective pressures acting upon natural populations and distinguishing between neutral and adaptive genetic variation (Nielsen 2005; Stapley *et al.* 2010; Radwan & Babik 2012; Bourret *et al.* 2013). Species that occupy heterogeneous environments (i.e. temperature, salinity) along their geographical distribution experience spatially varying selective pressures, often resulting in local adaptation of ecologically important traits (Kawecki & Ebert 2004; Fraser *et al.* 2011). Beginning with Levene (1953), a number of studies have shown that balancing selection due to spatial heterogeneity is an important mechanism responsible for the maintenance of genetic polymorphism (reviewed in Hedrick 2006). While polymorphism in a varying environment may be maintained even when dispersal results in complete mixing of the gene pool, in such case localities will not differentiate genetically and there will be no local adaptation (Kawecki & Ebert 2004). Recent population genomic studies predict the observation of single points of selection across the genome when gene flow is high, as loci subject to strong selection will tend to diverge independently from other genes, while highly differentiated genomic regions due to genome hitchhiking are expected with reduced gene flow (Feder & Nosil 2010; Yeaman & Otto 2011; Yeaman & Whitlock 2011; Feder *et al.* 2012). Genomic regions displaying increased differentiation referred to as genomic islands of divergence have been observed in many taxa from fungi (Ellison *et al.* 2011) and invertebrates (Nosil *et al.* 2008; Nadeau *et al.* 2012) to fishes (Hohenlohe *et al.* 2010; Jones *et al.* 2012a,b; Bradbury *et al.* 2013; Gagnaire *et al.* 2013; Hemmer-Hansen *et al.* 2013) and humans (Hoffer *et al.* 2012).

An excellent opportunity to study the interplay between spatially varying selection and gene flow exists in the European eel (*Anguilla anguilla*), a putatively panmictic species occupying a broad range of habitats from subarctic environments in Iceland, Norway and northwestern Russia to subtropical environments in North Africa and the Mediterranean Sea. The European eel is a facultative catadromous species with a complex life cycle, spawning in the remote Sargasso Sea and spending most of their life in continental (fresh, brackish and coastal) waters (Van den Thillart *et al.* 2009). After spawning, larvae are advected towards the coasts of Europe and North Africa, where they arrive following first the Gulf Stream and later the North Atlantic Drift Current. Upon reaching the continental shelf, larvae metamorphose into glass eels and move into continental growth habitats, settle and become pigmented

yellow eels. After a period of intense feeding of (on average) 7 years for males and 11 years for females, they metamorphose into partially mature silver eels that migrate back to the Sargasso Sea covering a distance of 5000–6000 km, spawn and die (Van den Thillart *et al.* 2009). In the Sargasso Sea, European eel spawns in partial spatial and temporal sympatry with its sister species, the American eel *Anguilla rostrata*, which provides opportunity to interbreed. European and American eels are known to hybridize, but hybrids have been reported almost exclusively in Iceland (Avise *et al.* 1990; Albert *et al.* 2006; Pujolar *et al.* 2014).

Despite the broad geographical distribution of the species, the European eel is regarded as a textbook example of panmixia, with the existence of a single randomly mating population. In the most comprehensive study to date genotyping over 1000 individuals at 21 microsatellite loci, Als *et al.* (2011) showed a very low and nonsignificant genetic differentiation between geographical locations across Europe and a lack of substructuring among larvae collected in the Sargasso Sea, providing very strong support for panmixia. In American eel, the recent study of Côté *et al.* (2013) genotyping over 2000 individuals representing 12 cohorts at 18 microsatellite loci over a large geographical scale in North America showed a total lack of genetic differentiation among samples, hence providing decisive evidence for panmixia in American eel. In contrast, significant geographical clines at some allozyme loci have been detected in both European (Maes & Volckaert 2002) and American eels (Koehn & Williams 1978). Moreover, Gagnaire *et al.* (2012a) found evidence for spatially varying selection at 13 of 73 candidate genes showing correlations between allele frequencies and environmental variables across the entire distribution range of American eel, particularly determined by temperature regimes. In the case of eels, owing to the apparent panmixia and random larval dispersal across habitats, any signature of spatially varying selection in a given generation is expected to be lost in the subsequent generation (Gagnaire *et al.* 2012a), hence preventing heritable trans-generational local adaptation.

This study aimed at investigating the influence of spatially varying selection on genetic diversity and potential adaptive divergence in European eel using a population genomics approach. The fact that most population genetic studies in eels have used neutral markers prompts for the investigation of the potential influence of natural selection on the genetic variation of the species.

Our first goal was to validate the lack of background level of neutral differentiation between locations, and for that we used a RAD-sequencing approach (Baird *et al.* 2008; Hohenlohe *et al.* 2010; Davey *et al.* 2011) to

Table 1 Sampling details including number of European eel (glass eels) individuals, geographical coordinates, sampling date and sea-surface temperature (ºC) at river mouth averaged across the 10, 30 and 90 days preceding sampling date

Location	Code	N	Coordinates	Sampling date	Temp-10 day	Temp-30 day	Temp-90 day
Vogslækur, Iceland	ICE	34	64°69′N/22°33′W	2/7/2001	9.53	9.38	8.45
Ringhals, Sweden	RHG	30	57°21′N/12°27′E	15/3/2008	4.76	4.77	4.19
Lough Erne, Northern Ireland	LG	33	54°46′N/7°77′W	1/7/2008	13.93	13.81	13.85
Burrishoole, Ireland	BG	29	53°90′N/9°58′W	14/3/2005	9.36	9.79	9.57
Gironde, France	GG	37	44°86′N/0°42′W	16/4/2008	11.55	11.78	11.26
Canet, France	CAG	32	42°70′N/3°15′E	23/1/2008	12.73	12.61	13.24
Valencia, Spain	VG	31	39°46′N/0°24′W	15/1/2010	14.06	14.16	15.05
Oved Sebou, Morocco	MOR	33	34°26′N/6°70′W	28/4/2001	17.53	17.82	17.57

identify 453 062 SNPs from 259 individuals sampled from eight locations across the species range. As part of testing the panmixia hypothesis, we were interested in testing whether genetic differentiation was consistent at nuclear and mitochondrial loci. We were also interested in testing for the presence of hybrids across Europe to assess whether hybridization could have influenced results, using diagnostic SNPs between European and American eels.

Our second goal was to contrast the outcome of neutral vs. adaptive differentiation in European eel. Using a subset of 50 354 SNPs with minor allele frequency >0.05, we tested for single-generation footprints of spatially varying selection with two main analytical approaches, one that identifies outliers as those markers with greater differentiation among all SNPs and a second based on determining positive associations between SNP frequencies and environmental variables. Following the positive associations observed by Gagnaire *et al.* (2012a) in American eel, variables used were degrees north latitude, degrees east/west longitude and sea-surface temperature at river mouths, corresponding to the sampling locations. Candidate genes were functionally annotated to assess potential functions of genes under selection. We specifically wanted to test whether genes under selection were grouped into clusters/islands as observed in other species (Nosil *et al.* 2008; Hohenlohe *et al.* 2010; Jones *et al.* 2012a,b; Nadeau *et al.* 2012; Bradbury *et al.* 2013; Gagnaire *et al.* 2013; Hemmer-Hansen *et al.* 2013) or more spread across the genome as expected to occur within just a single generation under a panmixia scenario.

## Materials and methods

Sampling

A total of 259 European eel (*Anguilla anguilla*) individuals were collected for RAD sequencing at eight locations across the geographical distribution of the species,

ranging from Iceland to Morocco (Table 1; Fig. 1). All samples were glass eels collected during January in the Mediterranean and March/April in the Atlantic, except samples from Iceland and Northern Ireland, which were collected in July. Genomic DNA was extracted using standard phenol–chloroform extraction.

RAD tag sequencing, RAD data analysis and SNP identification

Genomic DNA from each individual was digested with restriction enzyme *Eco*RI. RADs for all 259 European eel individuals were sequenced (10 individuals per lane) on an Illumina Genome Analyzer II by Beijing Genomics Institute (BGI, Hong Kong) using paired-end reads (for details see Pujolar *et al.* 2013).

Sequence reads from the Illumina runs were sorted according to barcode tag. Sequences were quality-filtered
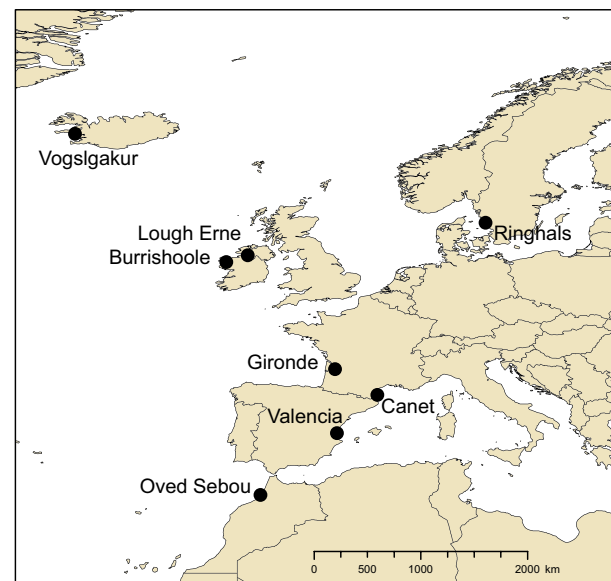
Fig. 1 Sampling locations of European eel (glass eel).

using FASTX-Toolkit (http://hannonlab.cshl.edu/fastx-toolkit), and reads with ambiguous barcodes/poor quality were removed from the analysis. Any read that presented a single nucleotide position with a Phred score lower than 10 was eliminated. This corresponds to the threshold generally used in SNP discovery studies (Van Bers *et al.* 2010; Ellison *et al.* 2011; Scaglione *et al.* 2012; Wagner *et al.* 2012). Final read length was trimmed to 75 nucleotides to reduce sequencing errors present at the tail of the sequences (Pujolar *et al.* 2013). For SNP calling, only the first (left) paired read was used due to the lower coverage of the second paired-end reads (Etter *et al.* 2011).

The un-gaped aligner BOWTIE version 0.12.8 (Langmead *et al.* 2009) was used to align sequence reads to the European eel genome draft (www.eelgenome.com), which consists of 179 Mbp of small contigs and 923 Mbp of larger contigs or scaffolds (Henkel *et al.* 2012). A maximum of two mismatches between the individual reads and the genome were allowed. Reads with alternative (two or more) alignments were excluded to avoid paralogous sequences.

Reference-aligned reads were processed using the ref_map.pl pipeline in STACKS version 0.9995 (Catchen *et al.* 2013). First, exactly matching reads were aligned together into stacks and subsequently merged to form putative loci. A minimum stack depth of 10 reads was used. Subsequently, a maximum-likelihood framework was used to call SNPs, and a catalogue was built of all existing loci and alleles against which all individuals were matched. Finally, the program Populations in Stacks was used to process all the SNP data across individuals.

Prior to the SNP analysis, loci in the catalogue were further filtered to remove paralogs and otherwise spurious loci according to the following three criteria: (i) loci with higher-than-average number of reads were excluded, because an extremely high coverage might be an indication of the presence of more than one locus (twice the standard deviation from the mean number of reads was used as threshold); (ii) loci with more than two alleles were eliminated because those might result from sequencing errors; and (iii) loci deviating from Hardy–Weinberg equilibrium (HWE), tested using GENEPOP version 4.2 (Raymond & Rousset 1995), were excluded after adjusting significance levels for multiple comparisons using the sequential Bonferroni technique (Rice 1989). As a final filtering step, minimum percentage of individuals in a population required to process a locus was set to 66.67%.

RAD analysis was also conducted separately for mitochondrial SNPs by aligning all reads against the European eel mitogenome (GenBank Accession no. NC_006531) in BOWTIE.

SNP analysis

Genome-wide measures of genetic diversity, including nucleotide diversity (Π) and observed (H_o) and expected (H_e) heterozygosities, were calculated in Stacks (Table 2). Differences in genetic diversity among samples were tested by one-way ANOVA using STATISTICA version 6.0 (StatSoft Inc). Deviations from HWE, differences in allele and genotype frequencies among samples, F-statistics for all samples and all sample pairs and isolation by distance (IBD) were tested using GENEPOP (Raymond & Rousset 1995). In all cases, significance levels were corrected for multiple comparisons using the sequential Bonferroni technique (Rice 1989). Pairwise F_{ST} values were used to conduct a multivariate ordination by multidimensional scaling (MDS) analysis using STATISTICA. IBD was tested using Mantel test (Mantel 1967) by correlating linearized genetic distance ($F_{ST}/(1-F_{ST})$) vs. geographical distance (shortest waterway distance in kilometre between sample pairs). All analyses were conducted considering (i) all nuclear DNA sequences and (ii) all mitochondrial sequences.

Finally, we tested the presence of hybrid individuals in the data set using STRUCTURE v.2.3.4 (Pritchard *et al.* 2000). For this purpose, the analysis also included a sample of 30 RAD-sequenced American eel (yellow and glass eels) *Anguilla rostrata* individuals collected in Quebec, Nova Scotia and Florida (Pujolar *et al.* 2013). A preliminary analysis was conducted using 20 random European and American eel individuals to identify diagnostic species-specific SNPs fixed at different alleles in each species ($F_{ST} = 1$) following the same approach as in Pujolar *et al.* (2014). We identified a total of 2148 diagnostic SNPs between European and American eels that were used for hybrid identification in the data set. The analysis in STRUCTURE was performed using $1 < k < 9$, with 10 replicates per k to check the consistency of results. We assumed an admixture model and uncorrelated allele frequencies, and we did not use population priors. A burn-in length of 100 000 steps followed by 1 million additional iterations was performed. The most likely k was determined using the criterion of Evanno *et al.* (2005). The presence and nature of hybrids was further tested using the Gensback setting in STRUCTURE. When using prior population information for individuals, the program tests whether each individual has an immigrant ancestor in the last G generations, where $G = 0$ corresponds to the individual being an immigrant itself. The analysis was performed with $G = 5$.

Tests for local selection

Evidence of local selection was tested by searching for increased population differentiation using F_{ST}-based

Table 2 Diversity indices across European eel (glass eels) samples considering (i) nuclear DNA sequences and (ii) mitochondrial sequences

Location	N	Nuclear			Mitochondrial			
		H_o	H_e	Φ	H_e	Φ	MNA	TNA
Vogslækur, Iceland	34	0.036 (0.006)	0.040 (0.008)	0.041 (0.008)	0.056 (0.012)	0.057 (0.012)	1.34	55
Ringhals, Sweden	30	0.039 (0.007)	0.041 (0.008)	0.042 (0.008)	0.056 (0.006)	0.057 (0.005)	1.54	63
Lough Erne, Northern Ireland	33	0.033 (0.006)	0.038 (0.007)	0.039 (0.008)	0.064 (0.009)	0.065 (0.009)	1.51	62
Burrishoole, Ireland	29	0.036 (0.007)	0.039 (0.008)	0.040 (0.008)	0.050 (0.006)	0.051 (0.007)	1.44	59
Gironde, France	37	0.035 (0.006)	0.039 (0.007)	0.034 (0.008)	0.064 (0.007)	0.065 (0.008)	1.56	64
Canet, France	32	0.036 (0.006)	0.039 (0.007)	0.040 (0.008)	0.068 (0.010)	0.070 (0.010)	1.46	60
Valencia, Spain	31	0.035 (0.006)	0.039 (0.008)	0.034 (0.008)	0.069 (0.008)	0.070 (0.008)	1.46	60
Oved Sebou, Morocco	33	0.036 (0.006)	0.040 (0.008)	0.041 (0.008)	0.067 (0.010)	0.068 (0.010)	1.59	65

H_o, observed heterozygosity; H_e, expected heterozygosity; Φ, nucleotide diversity; MNA, mean number of alleles; TNA, total number of alleles.
Standard deviation in parentheses.

outlier analyses implemented in LOSITAN (Antao *et al.* 2008) and BAYESCAN (Foll & Gaggiotti 2008) and by testing for significant associations between allele frequencies and environmental variables using BAYENV (Coop *et al.* 2010). Loci showing either significantly greater population differentiation or significant covariance with environmental variables relative to reference SNP distributions were considered candidates for being under local selection. In all three approaches, only SNPs with a minor allele frequency >0.05 were included in the analysis.

The selection detection workbench LOSITAN (http://popgen.eu/soft/lositan/) uses a coalescent-based simulation approach to identify outliers based on the distributions of heterozygosity and F_{ST} (Beaumont & Nichols 1996). First, LOSITAN was run using all SNPs to estimate the mean neutral F_{ST} as recommended by Antao *et al.* (2008). After the first run, the mean neutral F_{ST} was recomputed by removing those SNPs outside the confidence interval to obtain a better approximation of the mean neutral F_{ST}. This mean was then used to conduct a second and final run of LOSITAN using all SNPs. An estimate of *P*-value was obtained for each SNP. We used a strict threshold of 0.995 and a false discovery rate of 0.1 to minimize the number of false positives.

Outlier SNPs were also detected using the Bayesian test of Foll & Gaggiotti (2008) implemented in BAYESCAN (http://cmpg.unibe.ch/software/bayescan/). The method is based on a logistic regression model that separates locus-specific effects of selection from population-specific effects of demography. Loci under selection are detected after comparing the posterior probabilities of a neutral model considering only population-specific F_{ST} parameters and a model including selection via a

locus-specific F_{ST} component to describe the observed allele frequencies. Outlier analysis was conducted on the whole data set divided according to sampling location. BAYESCAN runs were implemented using default values for all parameters, including a total of 100 000 iterations after an initial burn-in of 50 000 steps. A *q*-value of 10% was used.

As an alternative to these F_{ST}-based outlier tests, we also searched for SNP–environment associations using BAYENV (Coop *et al.* 2010), which tests for covariance between candidate SNP allele frequencies and environment variables that exceed the expected covariances under genetic drift. In the first step, the program is run to estimate the covariate matrix using an MCMC algorithm. In the second step, the program is run to estimate the Bayes Factors (BF) for the environmental variables of interest. A BF >3 was considered indicative of an allele frequency correlation with an environmental variable. Results were compared over five independent runs for consistency. Environmental variables used included degrees north latitude, degrees east/west longitude and sea-surface temperature at river mouth averaged across the 10 days, 30 days and 3 months preceding the sampling date. Temperature data were obtained from the IRI (International Research Institute for Climate and Society) Climate Data Library (http://iridl.ldeo.columbia.edu/SOURCES/.NOAA/.NCDC/.OISST).

Gene predictions for the European eel genome (http://www.zfgenomics.org/sub/eel) were used to establish the genomic position of the candidate SNPs for local selection using a custom-made script. SNPs were considered to be located in a gene when included in complete coding sequences (CDS) and exonic and intronic regions. Functional annotation of

these genes was obtained using the BLAST2GO suite (Götz *et al.* 2008), which conducts BLAST similarity searches and maps Gene Ontology (GO) terms to the homologous sequences found. Only ontologies with *E*-value < 1E-6, annotation cut-off >55 and a GO weight >5 were considered for annotation. A more systematic functional interpretation of the set of candidate genes was obtained using the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway approach for higher-order functional annotation implemented in the Database for Annotation, Visualization and Integrated Discovery (DAVID) web-server v6.7 (http://david.abcc. ncifcrf.gov). Zebrafish *Danio rerio* was used as reference genome for annotation. Prior to the analysis in DAVID, a local BLAST was conducted for significant matches directly against zebrafish Ensembl proteins using BLASTX. Zebrafish Ensembl Gene IDs were obtained from the corresponding Ensembl protein entries using the BIOMART data mining tool in the Ensembl website (http://www.ensembl.org/biomart/). Gene functional analysis in DAVID was conducted defining the zebrafish IDs corresponding to those genes including a locally selected SNP as 'Gene list' and the zebrafish IDs corresponding to all genes as 'Background'. Standard settings of gene count = 2 and ease = 0.1 were used.

Finally, the patterns of differentiation across genome regions were characterized to test whether genes putatively under selection were grouped into clusters (genomic islands of differentiation) or more scattered across the genome. We estimated the levels of genetic differentiation between populations by calculating average F_{ST} for 50-kb genomic sliding windows. Alternative sliding windows (100 and 200 kb) were also tested. F_{ST} was calculated between the two most ecologically and geographically differing locations, Iceland and Morocco. Windows were restricted to the 30 longest scaffolds (903 936–2 025 234 bp) from the European eel draft genome.

Results

RAD sequencing

Sequencing of the RAD libraries generated an average of 9.59 million reads of 90 bp per individual. After trimming sequences to 75 bp and quality filtering, on average, 7.90 million (82.19%) reads per individual were retained. Mean quality score of the retained reads was 38.65, with a GC content of 41.1% (Table 3). An average of 69.96% of the quality-filtered reads aligned to the European eel draft genome and 4.46% of the reads were discarded due to alternative (two or more) alignments.

Aligned reads were assembled into an average of 526 821 stacks and subsequently into a set of 335 343 loci (Table 3). A total of 118 016 (35.5%) reads per individual were discarded due to insufficient depth of coverage. Loci with a minimum stack depth of 10 reads were retained to construct a catalogue of 527 504 loci, using a total of 80 individuals (10 individuals per location). Prior to SNP discovery, loci were further filtered to minimize errors in sequencing, alignment or assembly. First, 815 loci showing >57.35 reads (twice the standard deviation from the mean number of reads, 19.16 ± 19.09) were discarded, as a higher-than-average number of reads suggest the presence of more than one locus. Second, 40 270 loci showing three alleles per individual were eliminated. Third, 1749 loci not adjusting to Hardy–Weinberg proportions after Bonferroni correction were also removed, representing either loci at which all individuals were heterozygotes or loci at which one single individual was homozygote for an allele not observed in the rest of the individuals. Finally, after a filtering step selecting only loci genotyped in >66.7% of individuals in all sampling locations, a total of 72 932 loci were retained for SNP discovery. Using the program Population in Stacks, a total of 453 062 SNPs were discovered. After aligning against

Table 3 Statistics describing the distribution of different properties of RAD sequences after each step of filtering (FASTX-Toolkit), alignment to the eel draft genome (BOWTIE) and assembly into loci (Ref_map.pl)

FASTX										
Raw reads	Filtered reads	% Eliminated	Mean *Q*	*Q*1	Med	*Q*3	%A	%C	%G	%T
9593701	7899505	17.81	38.65	37.97	39.40	40.22	29.5	20.7	20.4	29.3

BOWTIE						
Reads	Aligned	% Aligned	Nonaligned	% Nonaligned	Discarded	% Discarded
7899505	5527660	69.96	2018833	25.59	353042	4.46

REF_MAP						
Reads	Stacks	Loci	Loci used	% Loci used	Loci discarded	% Loci discarded
5527660	526821	335343	217326	64.50	118016	35.50

the European eel mitogenome, a total of 11 loci and 41 SNPs were identified as mitochondrial.

Genetic diversity and differentiation

Measures of genetic variability across locations considering all identified SNPs are summarized in Table 2. No differences in values of observed heterozygosity (H_o = 0.038–0.039), expected heterozygosity (H_e = 0.038–0.041) and nucleotide diversity (Π = 0.034–0.042) were found across locations ($P > 0.05$ in all comparisons). No differences were observed when H_o, H_e and Π were calculated using all (fixed and variable) positions. Similarly, genetic diversity was similar across locations (H_e = 0.056–0.069; Π = 0.051–0.070) when considering only mitochondrial RADs (41 SNPs at 11 loci).

When investigating the genetic structure among locations, all pairwise F_{ST} comparisons were not significant ($P > 0.05$), with an average pairwise F_{ST} value of 0.0007. Comparison of allele frequencies across locations showed highly significant differences at only 26 of 453 062 SNPs after Bonferroni correction for multiple testing. An MDS plotting the first and second coordinates obtained from pairwise genetic distances showed all samples clustering together fitting no apparent geographical pattern (Fig. S1a, Supporting information). A Mantel test showed no correlation between genetic and waterway distances ($r = -0.232$; $P = 0.250$), which suggests no IBD pattern (Fig. 2a).

When considering only mitochondrial markers (41 SNPs), genetic differentiation was low (F_{ST} = 0.0002), and no significant differences were found when comparing allele frequencies across locations after Bonferroni correction. However, pairwise F_{ST} values were significant at five of 28 comparisons, all involving Iceland. Concordantly, Iceland was the most distant sample in the MDS analysis (Fig. S1b, Supporting information) and also contributed to the observed positive but nonsignificant pattern of IBD ($r = 0.311$; $P = 0.107$; Fig. 2b).

Using 2148 species-diagnostic nuclear SNPs with an F_{ST} value of 1 between the two species, STRUCTURE was used to investigate the possible presence of admixed individuals in the data set. A scenario with two groups ($K = 2$) corresponding to the two species was inferred to be the most likely (Fig. 3). Occurrence of hybrids was low, with a total of three admixed individuals found within European eel: one from Gironde with an admixture proportion of 0.053 (0.037–0.070) and two from Iceland with admixture proportions of 0.190 (0.176–0.218) and 0.192 (0.176–0.217), respectively. Additionally, three individuals (two from Canet and one from Ringhals) presented an admixture proportion of 0.03. Within American eel, one single admixed individ-
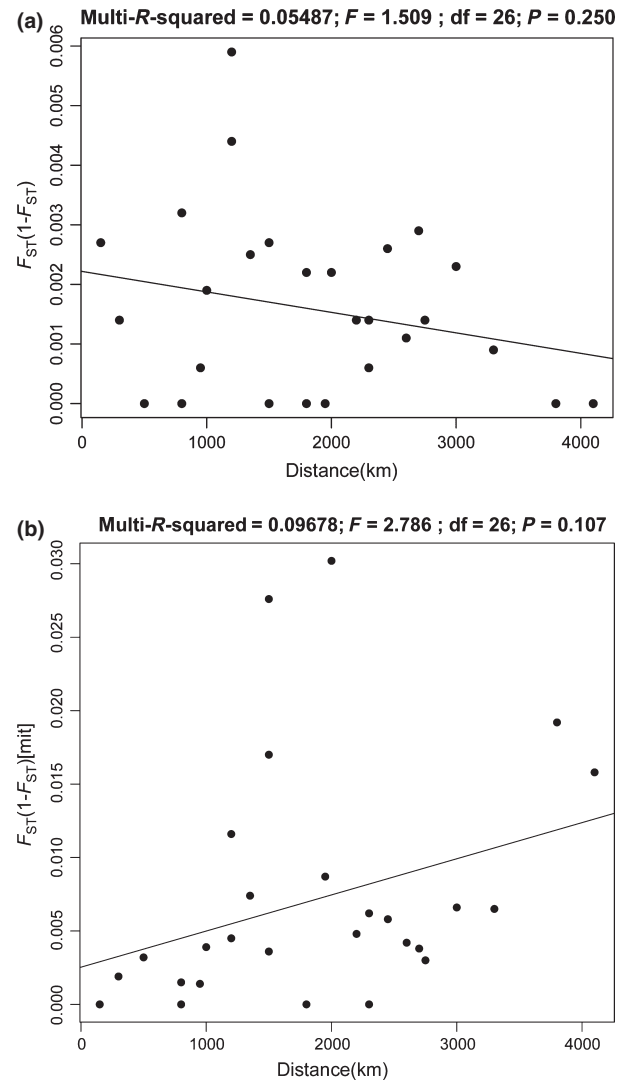

Fig. 2 Mantel test correlating pairwise linearized genetic distances ($F_{ST}/(1-F_{ST})$) vs. geographical distances (shortest waterway distance in kilometre) between all glass eel sampling locations considering (a) all RAD sequences and (b) mitochondrial RAD sequences only.

ual was found, with an admixture proportion of 0.056 (0.041–0.073). Using the Gensback setting in STRUCTURE, the two admixed individuals from Iceland were classified as first-generation backcrosses, while the admixed individual from Gironde was classified as a third-generation backcross, same as the admixed American eel individual.

Finally, all individuals were identified as European or American eel on the basis of the mitochondrial RADs (Fig. 3). Five of 41 SNPs were diagnostic, that is, fixed at different alleles. All European eels plus the three hybrids from Iceland and Gironde showed the European eel-diagnostic mitochondrial alleles, except one individ-
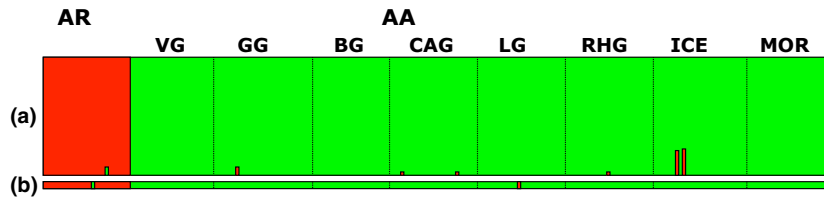
Fig. 3 Admixture analysis using STRUCTURE. (a) Individuals were assigned on the basis of the most likely K (in this case, $K = 2$). Each vertical line represents one individual, partitioned into segments according to admixture proportion of European eel (AA; green) and American eel (AR; red). Locations are labelled as in Table 1. (b) Identification of all individuals as European or American eel on the basis of the mitochondrial RADs is also included (bottom panel).

ual from Lough Erne identified as pure European eel on the basis of all nuclear loci presenting the five American eel-diagnostic mitochondrial alleles. All American eels plus the admixed American eel individual showed the American eel-diagnostic mitochondrial alleles, except one individual identified as pure American eel on the basis of all nuclear loci that presented the five European eel-diagnostic mitochondrial alleles.

Local selection

In all tests for local selection, a total of 50 354 SNPs with a minor allele frequency >0.05 were included in the analysis. Using LOSITAN, a total of 670 SNPs representing 639 unique loci were identified as outliers possibly under directional selection, after applying a significance level of 0.995. A smaller number of outlier loci were identified using BAYESCAN: 33 SNPs representing 30 unique loci, all of which were part of the outliers identified using LOSITAN. As expected when considering that both LOSITAN and BAYESCAN are F_{ST}-based methods, all outliers showed high F_{ST} values (0.04–0.12) compared with the background F_{ST}.

Bayesian tests for SNP–environment associations in BAYENV identified a total of 87 candidate SNPs representing 74 unique loci: 12 SNPs representing 12 unique loci associated with latitude, 29 SNPs representing 28 unique loci associated with longitude and 51 SNPs representing 49 unique loci associated with temperature, with a few SNPs being correlated with more than one variable. Highly similar results were obtained when using the three temperature data sets (last 10 days, last 30 days and last 90 days prior to sampling), with only eight nonshared candidate SNPs across data sets.

Collectively, a total of 754 potentially locally selected SNPs were identified, only three of which were common to both F_{ST}-based and SNP–environment association approaches. These three SNPs showed both strong SNP–environmental correlations (BAYENV) and increased F_{ST} relative to the reference-based threshold (LOSITAN and BAYESCAN). By contrast, 84 candidate loci exceeded the reference-based significance threshold solely for the

BAYENV test of environmental association, without being significant outliers in F_{ST}.

When genomic position of the locally selected SNPs was investigated, a hit with a gene was obtained for 39.8% of the SNPs. Thus, no hits were obtained for 60.2% of the SNPs, although 10.2% were located in upstream (5000 bp) regions of the genes. Most hits represented intronic regions (86.9%), with CDS and exons representing only 7.8% and 5.3% of the hits, respectively.

Among hits, 255 (65.4%) were associated with one or more among 1476 unique GO terms, for a total of 2628 term occurrences. Subsequently, the KEGG pathway approach for higher-order functional annotation was implemented using the tool DAVID. Using zebrafish as reference genome, a total of 335 zebrafish genes homologous to European eel were mapped to KEGG pathways. Enriched KEGG pathways using a standard setting of gene count = 2 are summarized in Table 4. The pathway with the highest number of genes was neuroactive ligand–receptor interaction (eight genes), including several genes implicated in behaviour such as dopamine receptor, glutamate receptor subunit AMPA3 and alpha-1D adrenoreceptor. Other pathways of particular interest were calcium signalling pathway (six genes) and circadian rhythm (two genes), including PERIOD.

Finally, average F_{ST} values calculated using a 50-kb sliding window were plotted for the 30 longest scaffolds (Fig. 4). F_{ST} was low throughout the scaffolds, with just a few narrow peaks. No regions of the scaffolds with pronounced divergence peaks were observed, consistent with panmixia removing any effect of diversifying selection from each new generation. Similar results were obtained when using alternative (100 and 200 kb) sliding windows.

Discussion

Further evidence for genomic panmixia

This study represents the first high-density SNP-based genome scan of genetic diversity and differentiation in

Table 4 KEGG pathways and genes identified using the F_{ST} outlier approach implemented in LOSITAN/BAYESCAN and by testing for associations with environmental variables in BAYENV

KEGG pathway	Gene	Approach
Calcium signalling pathway	ATPase, Ca^{++} transporting, cardiac muscle, slow twitch 2a	F_{ST} outlier
	Calcium channel, voltage-dependent, L type, alpha 1S subunit	F_{ST} outlier
	Calcium channel, voltage-dependent, L type, alpha 1S subunit, a	F_{ST} outlier
	Similar to *N*-methyl-D-aspartate receptor channel subunit epsilon 1	F_{ST} outlier
	Similar to Neuromedin-K receptor (NKR)	F_{ST} outlier
	Similar to alpha-1D adrenoreceptor	F_{ST} outlier
Neuroactive ligand–receptor interaction	Dopamine receptor D1	F_{ST} outlier
	Gamma-aminobutyric acid (GABA) A receptor, alpha 1	F_{ST} outlier
	Glutamate receptor, ionotropic, AMPA 4b; glutamate receptor, ionotropic, AMPA 4a; glutamate receptor, ionotropic, AMPA 2a; glutamate receptor, ionotropic, AMPA 1b; glutamate receptor, ionotropic, AMPA 1a; similar to glutamate receptor, ionotropic, AMPA 4; glutamate receptor, ionotropic, AMPA 3b; glutamate receptor, ionotropic, AMPA 3a	F_{ST} outlier
	Pyrimidinergic receptor P2Y, G-protein coupled, 4-like	F_{ST} outlier
	Similar to adenosine A1 receptor	F_{ST} outlier
Focal adhesion	Integrin, alpha 5	F_{ST} outlier
	Phosphatase and tensin homolog B	BAYENV
	Talin 1	F_{ST} outlier
	V-crk sarcoma virus CT10 oncogene homolog (avian)-like	F_{ST} outlier
	V-raf murine sarcoma viral oncogene homolog B1	F_{ST} outlier
	Calpain 2, (m/II) large subunit, like	F_{ST} outlier
Wnt signalling	C-terminal binding protein 2	F_{ST} outlier
	MAD homolog 2 (Drosophila)	F_{ST} outlier
	Secreted frizzled-related protein 1	F_{ST} outlier
	Wingless-type MMTV integration site family, member 7b	F_{ST} outlier
	Wingless-type MMTV integration site family, member 8b	F_{ST} outlier
Mucin-type *O*-glycan biosynthesis	UDP-*N*-acetyl-alpha-D-galactosamine: polypeptide *N*-acetylgalactosaminyltransferase 7	BAYENV
	WD repeat domain 51B, like	BAYENV
Circadian rhythm	Period circadian protein homolog 2	BAYENV
	Period circadian protein homolog-like 2	F_{ST} outlier

KEGG, Kyoto Encyclopedia of Genes and Genomes.

the European eel. All analyses of genetic diversity, genetic differentiation and isolation by distance are consistent with the interpretation of genomic panmixia that European eels sampled along the coasts of Europe and northern Africa belong to a single spatially homogenous population. The low levels of genetic differentiation in our study, revealed by both nuclear (F_{ST} = 0.0007) and mitochondrial loci (F_{ST} = 0.0002), are concordant with the recent study of Als *et al.* (2011), the most extensive study to date in terms of spatial sampling, which showed nonsignificant genetic differentiation among samples using 21 microsatellite loci, with an F_{ST} value of 0.00076 among larvae in the Sargasso Sea and 0.00024 among glass eels samples across Europe. Very low mean F_{ST} values have also been reported in all other studies on European eel that included large numbers of sampling sites and individuals (e.g. Wirth & Bernatchez 2001: F_{ST} = 0.0017; Dannewitz *et al.* 2005: F_{ST} = 0.0014; Maes *et al.* 2006: F_{ST} = 0.0099), the only

exception being a recent study showing a microsatellite F_{ST} value of 0.02 (this about 10 times more than reported in other studies) and a mitochondrial F_{ST} of 0.11 (Baltazar-Soares *et al.* 2014).

Random mating and larval dispersal may lead to the lack of spatial structure found in European eel despite its broad geographical distribution from Iceland to Morocco. While congregating at a single dominant spawning site could lead to panmixia, the vastness (*c.* 3 × 10^6 km^2) and heterogenous hydrographic structure of the Sargasso Sea plus changes in oceanographic conditions caused by ocean-atmospheric regime shifts may affect the location of spawning areas by silver eels (Friedland *et al.* 2007). Based on the occurrence of eel larvae, it has been proposed that particular thermal fronts within the subtropical convergence zone characterized by steep temperature and salinity gradients are used by eels to locate spawning areas and facilitate mating (McCleave 1993; Munk *et al.* 2010).
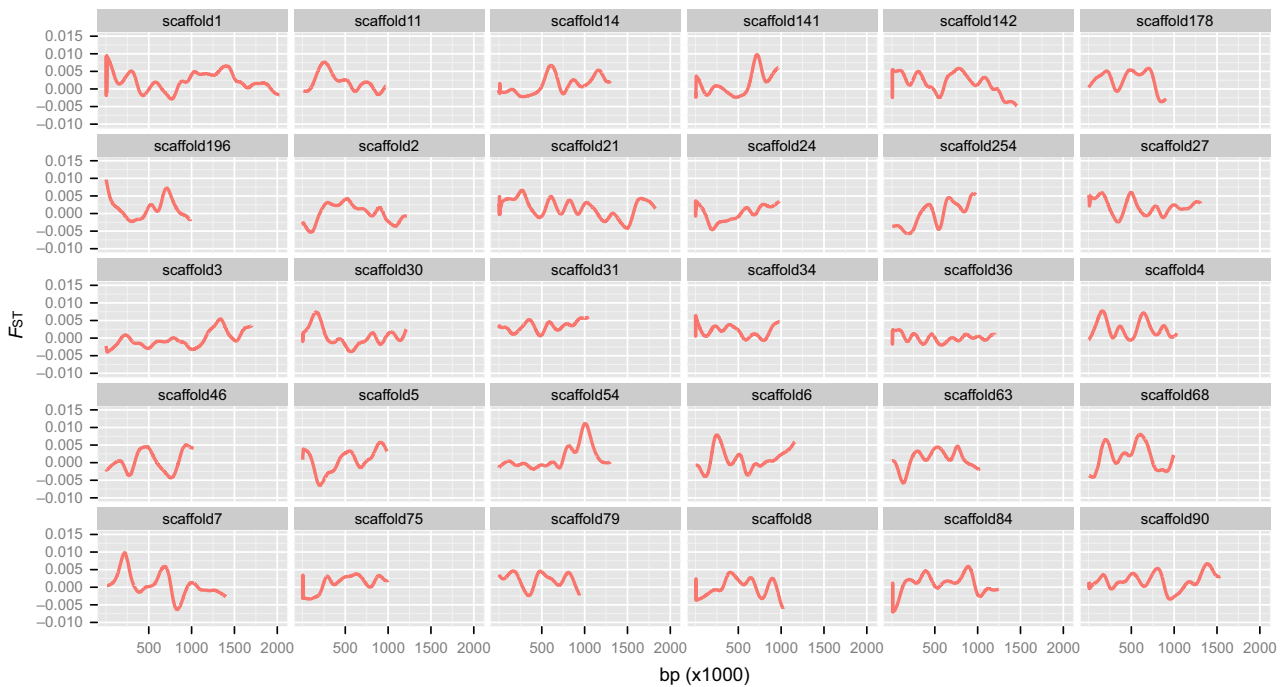
Fig. 4 Plots of average F_{ST} calculated using a 50-kb sliding window for the 30 longest scaffolds in the European eel genome.

Hydrographic profiling studies show that the position and strength of the thermal fronts are highly variable and unstable (Munk *et al.* 2010), which suggests no possibilities of philopatry of adults to the original birthplace within the Sargasso Sea and hence random mixing of individuals. In regard to larval dispersal, despite circulation patterns and oceanic currents in the Atlantic Ocean being complex, it is well established that European eel larvae are advected towards Europe following the Gulf Current and the North Atlantic Current (Lecomte-Finiger 1994). An alternative shorter transatlantic journey following the Subtropical Counter Current towards the Azores and Europe has been proposed for southern European eels (Munk *et al.* 2010) that would explain the observed highly significant heterogeneity in size and condition between Mediterranean and North European glass eels (Pujolar *et al.* 2007), with Mediterranean samples showing smaller size and condition. A shorter transatlantic migration route for Mediterranean larvae was also indicated in the modelling study of Kettle & Haines (2006). The putative existence of different larval migratory routes contrasts with the lack of spatial genetic structure in our study, in which European eel samples were compared with the largest number of markers to date. In accordance with Als *et al.* (2011), this seems to indicate that larval migration is random and that any segregation of individuals that might occur in the Sargasso Sea (either spatial or temporal) is not reinforced by larval homing to the parental original freshwater habitat.

Low occurrence of hybrids in Europe

Our study identified a total of 2148 SNP markers that were diagnostic between European and American eels (F_{ST} = 1 as they were fixed for different alleles), which were used for hybrid identification. Admixture analysis showed a low level of hybridization, with three hybrids among European glass eels (two in Iceland and one in Gironde) and one single hybrid among American eels.

We observed a hybrid occurrence of 5.9% in Iceland that fits the values found in previous studies, ranging from 2–4% to 15.5% (Avise *et al.* 1990; Albert *et al.* 2006; Pujolar *et al.* 2014). By comparison, four out of 225 (1.8%) individuals were identified as hybrids in mainland Europe, one individual with a 0.05 admixture proportion and three individuals with a 0.03 admixture proportion. The low occurrence of hybrids in mainland Europe is in agreement with the recent study of Als *et al.* (2011) that identified 0.2% of hybrids (one in Ireland and one in Belgium) using 21 microsatellites. Collectively, our data corroborate previous findings that a moderate percentage of Icelandic eels have American eel ancestry, showing that a much lower percentage of eels in mainland Europe is of admixed origin.

The low occurrence of hybrids in mainland Europe also suggests strong natural selection against hybrids. As a consequence of reduced hybrid viability and/or fertility, hybridizing individuals usually experience several costs that result in selection against heterospecific pairing (i.e. development instability; Coyne & Orr

2004). Recently, Gagnaire *et al.* (2012b) identified a possible cytonuclear incompatibility between North Atlantic eels after showing that positive selection has operated on both the mitochondrial atp6 gene and its nuclear interactor atp5c1. However, our data show that individuals presenting the nuclear genome of one species and the mitogenome of the other species are viable. One individual from Lough Erne (Northern Ireland) presented a pure European nuclear genome, but an American mitogenome. The reverse pattern was also observed, and European eel mtDNA was found in an individual with a pure American nuclear genome. This suggests that, albeit rare and sporadic, hybridization over several generations is possible and can occur in both directions.

Single-generation signatures of selection

Considering the high historical effective population size estimated for the European eel (from 100 000 to 1×10^6 individuals; Pujolar *et al.* 2013), random drift is expected to be negligible, while natural selection (together with gene flow) would be the major force determining allele frequency differences, in a pattern that is expected to vary from locus to locus. Effectively, the great majority of SNPs in our study showed no appreciable differences in allele frequencies among sampling locations, whereas a small set of SNPs showed significantly high genetic differentiation, consistent with the action of natural selection.

The low levels of baseline differentiation (F_{ST} of 0.0007) found in the European eel with no highly differentiated genome regions except a few narrow peaks suggest that most of the genome is homogenized by gene flow. When gene flow is high, strong divergence selection acting directly on individual target genes that diverge independently from other genes is expected (i.e. direct selection with no hitchhiking), while genomic islands of divergence due to genome hitchhiking are expected in the case of reduced gene flow (Feder *et al.* 2012). The observation of single points of selection rather than regions (islands) under selection in our study is in agreement with panmixia in the European eel.

After screening over 50 000 SNPs with a minor allele frequency >0.05, we identified several candidate genes possibly undergoing divergent selection associated with the highly variable environmental conditions used by the European eel along its geographical range. Our results are in accordance with the study of Gagnaire *et al.* (2012a) in American eel, which showed significant correlations between allele frequencies at 13 loci and the same variables used in our study (temperature, latitude and longitude).

We found large variation between the approaches used (F_{ST} outlier vs. environment association methods), and few SNPs were identified as targets of local selection by both methods. This is in agreement with recent studies in which genes that showed positive SNP–environment associations were not outliers in analyses of population structure (Hancock *et al.* 2010; Ma *et al.* 2010; Keller *et al.* 2012), which suggests that the SNP–environment association approach is more sensitive to subtle/gradual shifts in allele frequencies and may be complementary (rather than concordant) to F_{ST}-based outlier methods when testing for selection (Coop *et al.* 2010). Nevertheless, some of the putative targets of local selection in our study showed both high F_{ST} and SNP–environment associations, including genes in two important pathways, calcium signalling and circadian rhythm.

Calcium serves a number of functions in fish and controls processes as diverse as fertilization, development, learning and memory, mitochondrial function, muscle contraction and secretion (Berridge *et al.* 2000). It is also recognized as very important in ion exchange and osmoregulation. Spatially varying selection in regard to osmoregulation seems plausible in the European eel, because the species is facultatively catadromous and can occupy either freshwater or salt water (i.e. marine and brackish) habitats (Daverat *et al.* 2006).

Within the circadian rhythm pathway, the gene that showed the strongest and most consistent evidence for local selection across all analyses was the central circadian gene PERIOD (per), showing both a highly increased population differentiation in the outlier F_{ST} approach and a strong pattern of covariance with two environmental variables, temperature and latitude (allele frequencies ranged from 0.12 in Ringhals to 0.02 in the Mediterranean locations and Morocco). Many organisms exhibit circadian rhythms based on 24-h intervals regulated by endogenous biochemical oscillators or clocks that enable organisms to organize physiological and behavioural processes to occur at biologically advantageous times in a day (Hastings *et al.* 1991). Changing photoperiod lengths can also be an important environmental cue controlling seasonal activities. Clocks can be synchronized with environmental factors, most notably changes in temperature and light intensity, which is consistent with the statistical association in our study between PERIOD allele frequencies and temperature and latitude. PERIOD could be under local selection, in this case related to differences in local photoperiod associated with the >30° difference in latitude between the geographically most extreme locations (Iceland vs. Morocco).

The remaining genes showed high F_{ST} but no SNP–environment associations, including many genes associated

with behaviour, namely dopamine receptor DRD1, glutamate receptor subunit AMPA3 and alpha-1D adrenoreceptor. Dopamine and glutamate are neurotransmitters mediating a wide range of actions, including stress and aggressiveness and also locomotion and exploring behaviour. Social behaviour related to aggressiveness (agonistic behaviour) could be important for food accessibility and size-related sex determination in eels. Alpha-1D adrenoreceptors are also involved in a variety of stimulus-induced changes in locomotor behaviours. It should be noted that due to the methodological approach used (i.e. RAD sequencing), SNPs under selection could be underestimated because only a portion of the restriction fragments will be located in actual gene-coding regions.

Impossibility of local adaptation

High gene flow associated with the particular life history characteristics of eels prevents heritable trans-generational local adaptation. Genomic panmixia suggests that larval dispersal is random and there is no larval homing to the parental original freshwater habitat. As a consequence, the offspring of surviving individuals experiencing specific local conditions (e.g. high temperature) have no chance to return to the parental habitat in which the phenotype was originally advantageous and selected for. Despite evidence for single-generation footprints of spatially varying selection, local adaptation is impossible in the case of eels and locally adapted SNPs in a given generation may not be favoured by selection in the next generation. In fish, a similar scenario in which strong selection occurs in every generation was suggested for lagoon populations of the European sea bass *Dicentrarchus labrax* (Lemaire *et al.* 2000). As spawning only occurs at sea, any adaptation to the lagoon environment is lost as surviving young adults migrate from the lagoons to the open sea to mate with individuals from marine populations. Spatially and temporally varying selection has also been reported previously in several marine invertebrates, including blue mussel (*Mytilus edulis*) (Koehn *et al.* 1976) and acorn barnacle (*Semibalanus balanoides*) (Véliz *et al.* 2004).

The resilience of a species depends on its vulnerability in the face of environmental changes, which can lead to either genetic (local adaptation) or plastic (phenotypic plasticity) responses (Hoffmann & Willi 2008). Heritable local adaptation can allow species to cope with environmental variability. This has been shown in the mummichog *Fundulus heteroclitus*, in which local populations have evolved pollution tolerance allowing them to survive concentrations of contaminants that are lethal to populations from clean environments, in a response that is not plastic but adaptive and heritable

(Nacci *et al.* 2010; Whitehead *et al.* 2011). Gagnaire *et al.* (2012a) argued that locally selected polymorphisms are not easily maintained by spatially varying selection in the panmictic American eel and that under such conditions phenotypic plasticity provides a more functionally adaptive response to spatial environmental variation. As such, high phenotypic plasticity could explain the presence of eels in extremely heterogenous environments across Europe in terms of salinity, temperature, substrate, depth or productivity, as well as the large variance observed in life history traits such as age at maturity or growth rate.

Conclusions

Analysis of hundreds of thousands of SNPs provides compelling support for the notion that European eel is a panmictic species. Panmixia at the genomic level agrees with previous genetic studies using neutral markers (microsatellites and AFLP) and is also concordant with the hybrid ecological-genetic model of Andrello *et al.* (2011). Mitochondrial SNPs in our study also showed no differentiation (41 SNPs: $F_{ST} = 0.0002$), which contrasts with the significant differences recently reported by Baltazar-Soares *et al.* (2014), with a mitochondrial F_{ST} of 0.11. Reasons for such a high discrepancy between studies are unclear; they might be related to differences in sample sizes, but this clearly deserves further investigation.

The results also demonstrate that the occurrence of hybrids between European and American eel is rare outside Iceland. Nevertheless, the extensive number of SNPs allows for tracking hybridization several generations back in time and indicates the presence of introgressed individuals that may contribute to genuine gene flow between the species. Despite panmixia, some genes are identified that are candidates for being under local selection within generations. By considering their functions and the ecological and geographical diversity covered by the sampled sites, it is also biologically plausible that they are indeed under selection. Although phenotypic plasticity undoubtedly plays a major role in the persistence of eels in a range of highly variable environments, the present study and the related study by Gagnaire *et al.* (2012a) of American eel draw attention to the importance of spatially and temporally varying selection in high gene flow organisms, among which European and American eel represent extreme cases. Future work directions include testing whether the same genes are under spatially varying selection in glass eels and yellow/silver eels or whether the response involves a different set of genes. The latter option seems plausible considering the different selective pressures that affect eels in early vs. late life stages.

References

Albert V, Jónsson B, Bernatchez L (2006) Natural hybrids in Atlantic eels (*Anguilla anguilla*, *A. rostrata*): evidence for successful reproduction and fluctuating abundance in space and time. *Molecular Ecology*, **15**, 1903–1916.

Als TD, Hansen MM, Maes GE *et al.* (2011) All roads lead to home: panmixia of European eel in the Sargasso Sea. *Molecular Ecology*, **20**, 1333–1346.

Andrello M, Bevacqua D, Maes GE, De Leo GA (2011) An integrated genetic-demographic model to unravel the origin and genetic structure in European eel (*Anguilla anguilla*). *Evolutionary Applications*, **4**, 517–533.

Antao T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G (2008) LOSITAN- a workbench to detect molecular adaptation based on a F_{ST}-outlier method. *BMC Bioinformatics*, **9**, 323.

Avise JC, Nelson WS, Arnold J, Koehn RJ, Williams GC, Thorsteinsson V (1990) The evolutionary status of Iceland eels. *Evolution*, **44**, 1254–1262.

Baird NA, Etter PD, Atwood TS, *et al.* (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE*, **3**, e3376.

Baltazar-Soares M, Biastoch A, Harrod C *et al.* (2014) Recruitment collapse and population structure of the European eel shaped by local ocean current dynamics. *Current Biology*, **24**, 1–5.

Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society B-Biological Sciences*, **263**, 1619–1626.

Berridge MJ, Lipp P, Bootman MD (2000) The versatility and universality of calcium signalling. *Nature Reviews Molecular Cell Biology*, **1**, 11–21.

Bourret V, Dionne M, Kent MP, Lien S, Bernatchez L (2013) Landscape genomics in Atlantic salmon (*Salmo salar*): searching for gene-environment interactions driving local adaptation. *Evolution*, **67**, 3469–3487.

Bradbury IR, Hubert S, Higgins B *et al.* (2013) Genomic islands of divergence and their consequences for the resolution of spatial structure in an exploited marine fish. *Evolutionary Applications*, **6**, 450–461.

Catchen JM, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool set for population genomics. *Molecular Ecology*, **22**, 3124–3140.

Coop G, Witonsky D, Di Rienzo A, Pritchard JK (2010) Using environmental correlations to identify loci underlying local adaptation. *Genetics*, **185**, 411–1423.

Côté C, Gagnaire PA, Bourret V, Verrault G, Castonguay M, Bernatchez L (2013) Population genetics of the American eel (*Anguilla rostrata*): F_{ST} = 0 and North Atlantic Oscillation effects on demographic fluctuations of a panmictic species. *Molecular Ecology*, **22**, 1763–1776.

Coyne JA, Orr HA (2004) *Speciation*. Sinauer Associates, Sunderland, Massachusetts.

Dannewitz J, Maes GE, Johansson L, Wickström H, Volckaert FAM, Jarvi T (2005) Panmixia in the European eel: a matter of time. *Proceedings of the Royal Society B-Biological Sciences*, **272**, 1129–1137.

Daverat F, Limburg K, Thibaut I *et al.* (2006) Phenotypic plasticity of habitat use by three temperate eel species *Anguilla anguilla*, *A. japonica* and *A. rostrata*. *Marine Ecology Progress Series*, **308**, 231–241.

Davey JW, Hohenlohe PA, Etter PD *et al.* (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics*, **12**, 499–510.

Ellison CE, Hall C, Kowbel D *et al.* (2011) Population genomics and local adaptation in wild isolates of a model microbial eukaryote. *Proceedings of the National Academy of Science of the USA*, **108**, 2831–2836.

Etter PD, Preston JL, Bassham S, Cresko WA, Johnson EA (2011) Local *de novo* assembly of RAD paired-end contigs using short sequencing reads. *PLoS ONE*, **6**, e18561.

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

Feder JL, Nosil P (2010) The efficacy of divergence hitchhiking in generating genomic islands during ecological speciation. *Evolution*, **64**, 1729–1747.

Feder JL, Egan SP, Nosil P (2012) The genomics of speciation-with-gene-flow. *Trends in Genetics*, **28**, 342–350.

Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics*, **180**, 977–993.

Fraser DJ, Weir LK, Bernatchez L, Hansen MM, Taylor EB (2011) Extend and scale of local adaptation in salmonid fishes: review and meta-analysis. *Heredity*, **106**, 404–420.

Friedland KD, Miller MJ, Knights B (2007) Oceanic changes in the Sargasso Sea and declines in recruitment of the European eel. *ICES Journal of Marine Science*, **64**, 519–530.

Gagnaire PA, Normandeau E, Côté C, Hansen MM, Bernatchez L (2012a) The genetic consequences of spatially varying selection in the panmictic American eel (*Anguilla rostrata*). *Genetics*, **190**, 725–736.

Gagnaire PA, Normandeau E, Bernatchez L (2012b) Comparative genomics reveals adaptive protein evolution and a possible cytonuclear incompatibility between European and American eels. *Molecular Biology and Evolution*, **29**, 2909–2919.

Gagnaire PA, Pavey SA, Normandeau E, Bernatchez L (2013) The genetic architecture of reproductive isolation during speciation-with-gene-flow in lake whitefish species pairs assessed by RAD sequencing. *Evolution*, **67**, 2483–2497.

Götz S, Garcia-Gomez JM, Terol J *et al.* (2008) High throughput functional annotation and data mining with the Blast2Go suite. *Nucleic Acids Research*, **36**, 3420–3435.

Hancock AM, Witonsky DB, Ehler E *et al.* (2010) Colloquium paper: human adaptations to diet, subsistence, and ecoregion are due to subtle shifts in allele frequency. *Proceedings of the National Academy of Sciences of the USA*, **107**, 8924–8930.

Hastings JW, Rusak B, Boulos Z (1991) Circadian rhythms: the physiology of biological timing. In: *Neural and Integrative Animal Physiology* (ed Prosser CL), pp. 435–546. Wiley-Liss, New York.

Hedrick PW (2006) Genetic polymorphism in heterogeneous environments: the age of genomics. *Annual Review of Ecology Evolution and Systematics*, **37**, 67–93.

Hemmer-Hansen J, Nielsen EE, Therkildsen NJ *et al.* (2013) A genomic island linked to ecotype divergence in Atlantic cod. *Molecular Ecology*, **22**, 2653–2667.

Henkel CV, Burgerhout E, Danielle L *et al.* (2012) Primitive duplicate hox clusters in the European eel's genome. *PLoS ONE*, **7**, e32231.

Hoffer T, Foll M, Excoffier L (2012) Evolutionary forces shaping genomic islands of population differentiation in humans. *BMC Genomics*, **13**, 107.

Hoffmann AA, Willi Y (2008) Detecting genetic responses to environmental change. *Nature Reviews Genetics*, **9**, 421–432.

Hohenlohe PA, Basshan S, Etter PD, Stiffler N, Johnson EA, Cresko WA (2010) Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genetics*, **6**, e1000862.

Jones FC, Chan YF, Schmutz J *et al.* (2012a) A genome-wide SNP genotyping array reveals patterns of global and repeated species-pair divergence in sticklebacks. *Current Biology*, **22**, 83–90.

Jones FC, Grabherr MG, Chan YF *et al.* (2012b) The genomic basis of adaptive evolution in threespine sticklebacks. *Nature*, **484**, 55–61.

Kawecki TJ, Ebert D (2004) Conceptual issues in local adaptation. *Ecology Letters*, **7**, 1225–1241.

Keller SR, Levsen N, Olson MS, Tiffin P (2012) Local adaptation in the flowering-time gene network of balsam poplar, *Populus balsamifera*. *Molecular Biology and Evolution*, **29**, 3143–3152.

Kettle A, Haines K (2006) How does the European eel (*Anguilla anguilla*) retain its population structure during its larval migration across the Atlantic Ocean? *Canadian Journal of Fisheries and Aquatic Sciences*, **63**, 90–106.

Koehn RK, Williams GC (1978) Genetic differentiation without isolation in American eel, *Anguilla rostrata*. 2. Temporal stability of geographic patterns. *Evolution*, **32**, 624–637.

Koehn RK, Milkman R, Mitton JB (1976) Population genetics of marine pelecypods. IV. selection, migration, and genetic differentiation in the blue mussel *Mytilus edulis*. *Evolution*, **30**, 2–30.

Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, **10**, R25.

Lecomte-Finiger R (1994) The early life of the European eel. *Nature*, **370**, 424.

Lemaire C, Allegrucci G, Naciri M, Bahri-Sfar L, Kara H, Bonhomme F (2000) Do discrepancies between microsatellite and allozyme variation reveal differential selection between sea and lagoon in the sea bass (*Dicentrarchus labrax*). *Molecular Ecology*, **9**, 457–467.

Levene H (1953) Genetic equilibrium when more than one ecological niche is available. *American Naturalist*, **87**, 331–333.

Ma XF, Hall D, Onge KR, Jansson S, Ingvarsson PK (2010) Genetic differentiation, clinal variation and phenotypic associations with growth cessation across the *Populus tremula* photoperiodic pathway. *Genetics*, **186**, 1033–1044.

Maes GE, Volckaert FAM (2002) Clinal genetic variation and isolation by distance in the European eel *Anguilla anguilla*. *Biological Journal of the Linnean Society*, **77**, 509–522.

Maes GE, Pujolar JM, Hellemans B, Volckaert FAM (2006) Evidence for isolation by time in the European eel (*Anguilla anguilla*). *Molecular Ecology*, **15**, 2095–2107.

Mantel N (1967) The detection of disease clustering and generalised regression approach. *Cancer Research*, **27**, 209–220.

McCleave JD (1993) Physical and behavioral controls on the oceanic distribution and migration of leptocephali. *Journal of Fish Biology*, **43**, 243–273.

Munk P, Hansen MM, Maes GE *et al.* (2010) Oceanic fronts in the Sargasso Sea control the early life and drift of Atlantic eels. *Proceedings of the Royal Society B-Biological Sciences*, **277**, 3593–3599.

Nacci DE, Champlin D, Jayaraman S (2010) Adaptation of the estuarine fish *Fundulus heteroclitus* (Atlantic killifish) to polychlorinated biphenyls (PCBs). *Estuaries and Coasts*, **33**, 853–864.

Nadeau NJ, Whible A, Jones RT *et al.* (2012) Genomic islands of divergence in hybridizing *Heliconius* butterflies identified by large scale targeted sequencing. *Philosophical Transactions of the Royal Society B-Biological Sciences*, **367**, 343–353.

Nielsen R (2005) Molecular signatures of natural selection. *Annual Review of Genetics*, **39**, 197–218.

Nosil P, Egan SP, Funk DJ (2008) Heterogenous genomic differentiation between walking-stick ecotypes: "isolation by adaptation" and multiple roles for divergent selection. *Evolution*, **62**, 316–336.

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

Pujolar JM, Maes GE, Volckaert FAM (2007) Genetic and morphometric heterogeneity among recruits of the European eel *Anguilla anguilla*. *Bulletin of Marine Science*, **81**, 297–308.

Pujolar JM, Jacobsen MW, Frydenberg J *et al.* (2013) A resource of genome-wide single-nucleotide polymorphisms by RAD tag sequencing in the critically endangered European eel. *Molecular Ecology Resources*, **13**, 706–714.

Pujolar JM, Jacobsen MW, Als TD *et al.* (2014) Assessing patterns of hybridization between North Atlantic eels using diagnostic single nucleotide polymorphisms. *Heredity*. in press.

Radwan J, Babik W (2012) The genomics of adaptation. *Proceedings of the Royal Society B-Biological Sciences*, **279**, 5024–5028.

Raymond M, Rousset F (1995) GENEPOP (version 1.2): a population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Rice WR (1989) Analyzing tables and statistical tests. *Evolution*, **43**, 223–225.

Scaglione S, Acquadro A, Portis E, Tirone M, Knapp SJ, Lanteri S (2012) RAD tag sequencing as a source of SNP markers in *Cynara cardunculus*. *BMC Genomics*, **13**, 3.

Stapley J, Reger J, Feulner PGD *et al.* (2010) Adaptation genomics: the next generation. *Trends in Ecology and Evolution*, **25**, 705–712.

Van Bers NEM, Van Oers K, Kerstens HHD (2010) Genome-wide SNP detection in the great tit *Parus major* using high throughput sequencing. *Molecular Ecology*, **19**, 89–99.

Van den Thillart G, Rankin JC, Dufour S (2009) *Spawning Migration of the European Eel: Reproduction Index, a Useful Tool for Conservation Management*. Springer, Dordecht, The Netherlands.

Véliz D, Bourget E, Bernatchez L (2004) Regional variation in the spatial scale of selection at MPI* and GPI* in the acorn

barnacle *Semibalanus balanoides* (Crustacea). *Journal of Evolutionary Biology*, **17**, 953–966.

Wagner CE, Keller I, Wittwer S *et al.* (2012) Genome-wide RAD sequence data provide unprecedented resolution in species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. *Molecular Ecology*, **22**, 787–798.

Whitehead A, Galvez F, Zhang S, Williams LM, Oleksiak MF (2011) Functional genomics of phenotypic plasticity and local adaptation in killifish. *Journal of Heredity*, **102**, 499–511.

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

Yeaman S, Otto SP (2011) Establishment and maintenance of adaptive genetic divergence under migration, selection and drift. *Evolution*, **65**, 2123–2129.

Yeaman S, Whitlock MC (2011) The genetic architecture of adaptation under migration-selection balance. *Evolution*, **65**, 1897–1911.

Data accessibility

Sequence reads have been deposited in the NCBI Sequence Read Archive (Project Number PRJNA195555). Raw SNP data and a custom-made script are available from Dryad database (http://datadryad.org) under doi:10.5061/dryad.s8v7q.

Supporting information

Additional supporting information may be found in the online version of this article.

Fig. S1 Plots from Multi-Dimensional Scaling analysis based on pairwise F_{ST} values considering (a) all RAD sequences and (b) mitochondrial RAD sequences only.