Mar Biol (2012) 159:1239–1253
DOI 10.1007/s00227-012-1905-x

ORIGINAL PAPER

Multi-marker estimate of genetic connectivity of sole (*Solea solea*) in the North-East Atlantic Ocean

E. L. Cuveliers · M. H. D. Larmuseau · B. Hellemans · S. L. N. A. Verherstraeten · F. A. M. Volckaert · G. E. Maes

Received: 7 November 2011 / Accepted: 19 February 2012 / Published online: 3 March 2012

Abstract A thorough knowledge on the genetic connectivity of marine populations is important for fisheries management and conservation. Using a dense population sampling design and two types of neutral molecular markers (10 nuclear microsatellite loci and a mtDNA cytochrome *b* fragment), we inferred the genetic connectivity among the main known spawning grounds of sole (*Solea solea* L.) in the North-East Atlantic Ocean. The results revealed a clear genetic structure for sole in the North-East Atlantic Ocean with at least three different populations, namely the Kattegat/Skagerrak region, the North Sea and the Bay of Biscay, and with indications for a fourth population, namely the Irish/Celtic Sea. The lack of genetically meaningful differences between biological populations within the southern North Sea is likely due to a large effective population size and sufficient connection (gene flow) between populations. Nevertheless, an isolation-by-distance pattern was found based on microsatellite genotyping, while no such pattern was observed with the cytochrome *b* marker, indicating an historical pattern prevailing in the latter marker. Our results demonstrate the importance of a combined multi-marker approach to understand the connectivity among marine populations at region scales.

Communicated by T. Reusch.

E. L. Cuveliers, M. H. D. Larmuseau contributed equally to this study.

Electronic supplementary material The online version of this article (doi:10.1007/s00227-012-1905-x) contains supplementary material, which is available to authorized users.

E. L. Cuveliers · M. H. D. Larmuseau · B. Hellemans · S. L. N. A. Verherstraeten · F. A. M. Volckaert · G. E. Maes
Laboratory of Biodiversity and Evolutionary Genomics, Katholieke Universiteit Leuven, Charles Deberiotstraat 32, 3000 Leuven, Belgium

M. H. D. Larmuseau
Department of Human Genetics, Katholieke Universiteit Leuven, Campus Gasthuisberg, Herestraat 49, 3000 Leuven, Belgium

M. H. D. Larmuseau
Department of Forensic Medicine, Laboratory of Forensic Genetics and Molecular Archaeology, UZ Leuven, Kapucijnenvoer 33, 3000 Leuven, Belgium

M. H. D. Larmuseau (✉)
Laboratory of Animal Diversity and Systematics, Katholieke Universiteit Leuven, Charles Deberiotstraat 32, 3000 Leuven, Belgium
e-mail: maarten.larmuseau@bio.kuleuven.be

Introduction

Population connectivity plays a fundamental role in the dynamics of marine ecosystems (Jones et al. 2007). The spatial pattern of genetic diversity in marine species is largely determined by the effects of three fundamental evolutionary forces, namely natural selection, genetic drift and gene flow (Slatkin 1985; Slatkin 1987). Knowledge of the dynamics of these forces acting on local populations under natural highly advective conditions is a challenging prerequisite for effective and sustainable fisheries management (Hedgecock et al. 2007). Hence, the delineation of appropriate management units requires information from population genetic, ecological and demographic data, as most marine populations will probably be intermediate in structure between discrete and continuous models (Palsbøll et al. 2007). Additionally, such information is crucial for conserving intraspecific genetic diversity (Reiss et al. 2009) and for the design of marine reserves (Fogarty and Botsford 2007; Jones et al. 2007).

Marine fish with large population sizes, broad geographic distributions, high fecundity and high dispersal potential are

expected to show low genetic differentiation and high gene flow (DeWoody and Avise 2000). Although a number of population genetic studies found evidence for structure even in species with high dispersal potential (Jørgensen et al. 2005; Knutsen et al. 2003; Larmuseau et al. 2009b; Nielsen et al. 2004; Pampoulie et al. 2004), measuring genetic connectivity in marine environments without any obvious physical barriers remains a challenge. Besides high levels of gene flow, other factors such as recent recolonization events and high effective population size could explain the lack of detectable genetic differentiation at neutral loci (Buonaccorsi et al. 2001; Larmuseau et al. 2010a). On the other side, many marine species with type III survival curves show a more chaotic pattern in their genetic constitution, termed genetic patchiness, resulting from a large variance in reproductive success among parents (Pujolar et al. 2006; Selkoe et al. 2010). Many population genetic studies have used microsatellite markers for the detection of subtle population structure, because of their high levels of polymorphism (Selkoe et al. 2010). However, the effective population size of mitochondrial DNA (mtDNA) is four times smaller than that of nuclear microsatellite loci due to its haploid character and maternal inheritance, making mtDNA more susceptible to the effects of genetic drift. Estimates of population differentiation are indeed often higher for mtDNA markers than nuclear markers (Larmuseau et al. 2010a; Lukoschek et al. 2008). A combination of multiple markers is therefore ideal to study the genetic structure of marine organisms.

Sole (*Solea solea* L.), the target species of this study, is one of the few marine fish in the North-East Atlantic for which a comprehensive ecological knowledge exists on the location and connectivity between the spawning and nursery grounds. Hence, estimates of genetic connectivity can be tested against specific biological hypotheses to confirm either the presence or absence of a match between genetic and management units (Knutsen et al. 2011). Sole has a wide geographical distribution from the southern coast of Norway, the Kattegat and Skagerrak to the Northwest African coast, also including the Mediterranean Sea. On the one hand, some of its biological characteristics may prevent strong population structure, such as a high number of pelagic eggs (hatching after 7–8 days) and pelagic larvae settling some 3 weeks after hatching (Fonds 1979) in nursery areas (Horwood 2001; Rijnsdorp et al. 1992). Conversely, various characteristics point to potential subpopulations: differences in life history traits (e.g., number and size of the eggs) exist among subpopulations across a latitudinal gradient (Rijnsdorp and Vingerhoed 1994; Witthames et al. 1995) and plankton surveys indicate distinct spawning aggregations. In the North Sea, these spawning hotspots are situated in the inner German Bight, in the Eastern English Channel, off the Belgian coast, in the Thames estuary and on the Norfolk Banks (Borremans 1987; De Clerck and Van de Velde 1973; Rijnsdorp et al. 1992; Wegner et al. 2003). The peak of spawning is from March in the Bay of Biscay to May–June in the south-eastern North Sea and Irish Sea (Symonds and Rogers 1995). Third, a study modeling larval transport of sole in the southern North Sea indicated a low exchange rate of larvae among spawning populations compared to the amount of larvae retained within an area, pointing to a limited connectivity among populations (Savina et al. 2010, G. Lacroix, pers. comm.). Finally, tagging studies showed limited juvenile movement away from the spawning and nursery grounds (Burt and Millner 2008), which was recently confirmed by otolith microchemistry on juvenile sole from the same study area (Cuveliers et al. 2010). As the Southern Bight of the North Sea is currently considered a single management unit (ICES area IV) (Reiss et al. 2009), a better knowledge on the connectivity would enable the definition of biologically more meaningful management units.

Despite the commercial importance of sole, population genetic studies in the North-East Atlantic Ocean are relatively scarce, compared to studies in the Mediterranean Sea and the Bay of Biscay (Cabral et al. 2003; Garoia et al. 2007; Guarniero et al. 2002; Guinand et al. 2008). Across the NE Atlantic, a pattern of isolation by distance (IBD) was found from the English Channel to the Tunesian Coast with allozyme markers, and the geographic range of significant population structure was estimated at 100 km (Kotoulas et al. 1995). Other studies using allozyme or exon-primed intron-crossing (EPIC) markers failed to detect an IBD pattern, which was attributed to high contemporary gene flow throughout its whole distribution range in the North-East Atlantic (Exadactylos et al. 1998; Rolland et al. 2007). Genetic differences between continental Europe (Bay of Biscay and German Bight) and the British Isles (Eastern Coast of England and Irish Sea) were only found in a study using randomly amplified polymorphic DNA (RAPD) on a limited number of samples (Exadactylos et al. 2003). Nevertheless, the above-mentioned studies included only few sampling locations in the North Sea and only one study included a sample from the Kattegat (Rolland et al. 2007), while almost none of the samples were taken during the spawning season. Moreover, temporal stability of the genetic signal was investigated in only two studies on a local scale in the Bay of Biscay (Kotoulas et al. 1995; Rolland et al. 2007). Nevertheless, temporal replication is recommended to evaluate possible sources of inter-locus bias and sampling artifacts in the data, especially for marine species exhibiting high dispersal capacities (Waples 1998). Therefore, these findings should be confirmed and complemented by additional continuously distributed spatial samples from the northernmost distribution range of sole, an increased sample size, and a test of temporal stability using a combined multi-marker approach.

The main purpose of this study was to elucidate the fine-scale genetic population structure of sole in the North-East Atlantic Ocean, with a focus on the North Sea region. The combined use of microsatellite and mitochondrial markers provides the opportunity to detect subtle population structure and to evaluate the discrimination power of both markers in a highly advective marine environment. Hence, we aimed at (1) analyzing the genetic population structure of *Solea solea* in the North Atlantic over a latitudinal gradient; and (2) estimating the extent of juvenile dispersal and potential reduction in genetic diversity due to genetic patchiness by analyzing nursery samples flanking the spawning regions. Finally, we discuss the match between biologically and genetically defined populations by contrasting our results with information on life-history trait variation and ecological (microchemistry) connectivity data on dispersal. This is of importance for improved sustainable management of this highly exploited species.

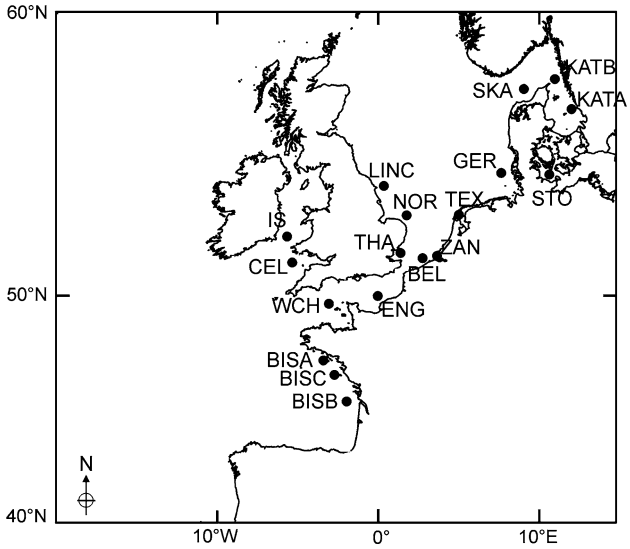

Fig. 1 *Solea solea.* Map of study area with sampling locations indicated. See Table 1 for sample codes

Materials and methods

Biological samples

A total of 1,499 adult sole (*Solea solea* L.) were caught at 17 locations across the North Sea, Irish Sea, Celtic Sea, Bay of Biscay, English Channel, Kattegat, Skagerrak and Belt Sea, during the period 2006–2008 (Fig. 1). During the same period, 312 juvenile sole were sampled at four nursery grounds: the Wadden Sea near Texel, the Scheldt estuary near Zandvliet, off the Thames and along the Belgian coast (Table 1). Adult fish were caught by commercial vessels or during research surveys. Most adult sole were sampled in spring or summer, except for the samples from the Kattegat, Skagerrak and Belt Sea, which were sampled during an autumn survey. Juvenile fish were caught in autumn with fyke nets or beam trawl. Temporal replicas were available for six sampling locations within the North Sea to test for temporal stability between years or between seasons. Sample codes refer to the sampling location and sampling year (Table 1). Fish were either immediately stored frozen or length measurements were taken onboard, and a fin tissue sample was preserved in 96% ethanol.

Molecular methods

Total genomic DNA was isolated from a piece of pectoral fin tissue using the Nucleospin Tissue Extraction Kit according to the manufacturer's guidelines (Macherey–Nagel GmBH, Düren, Germany). Individual fish were initially genotyped for 15 microsatellite loci (F8-ICA9, F8-ITG11, F13-II8/4/7, F8-IIGT15, F14-IIGT16 (Iyengar et al. 2000); (Sos(AC)6, Sos(AC)20, Sos(AC)40, Sos(AC)30, Sos(AC)45, (Garoia et al. 2006); SolCA13, SolGA12, Sol19A (Porta and Alvarez 2004); SseGATA26, SseCA28 (Funes et al. 2004)) on an automated capillary sequencer ABI 3130 AVANT (Applied Biosystems). MtDNA data have been generated for 23 samples; a total of 651 individuals were sequenced for a 590-bp fragment of the mitochondrial cytochrome *b* (cyt *b*) locus (Table 1). The cyt *b* gene was amplified by PCR using the primers CB1bis-F (5′-TACGTCCTCCCCTGAGGACAGATATC-3′) and SolCytb1-R (5′-GGCGCTCTAACACTGAGCTAC-3′). Sequences of 590 bp were evaluated with SeqScape v. 2.1 (Applied Biosystems) and aligned with BIOEdit v.7.0.5 (Hall 1999). Details on the polymerase chain reaction (PCR) conditions of microsatellite and mitochondrial markers are presented in the Supplementary Material (Table S1).

Statistical analyses

To identify potential genotyping errors in the microsatellite data (i.e., stuttering, large allele dropout or null alleles), the software Microchecker v.2.2.3 (van Oosterhout et al. 2006) was used. Null allele frequencies were also estimated for each locus and population using the Expectation Maximization algorithm of Dempster et al. (1977), carried out by the software FreeNA (Chapuis and Estoup 2007). The software Lositan (Antao et al. 2008), based on the fdist F_{ST} outlier method of Beaumont and Nichols (1996), was used to assess whether any of the microsatellite loci deviated from selective neutrality. A total of 95 000 simulations were computed for SMM with the options 'neutral mean F_{ST}' and 'force mean F_{ST}'. Deviations from Hardy–Weinberg equilibrium were tested per locus and sample using

Table 1 Sampling information for *Solea solea* including sample code, life stage (A = adult; J = juvenile), location (see Fig. 1), area, sampling date (including attribution to the spawning season *), latitude, longitude, number of samples (N), number genotyped (G), number sequenced (S)

Sample	A/J	Location	Area	Date	Latitude	Longitude	N	G	S
STO07	A	STO	Belt Sea	Oct/2007	55°10′29″N	11°02′44″E	48	45	0
KATA07	A	KATA	Kattegat	Nov/2007	57°08′91″N	11°38′52″E	48	44	0
KATB07	A	KATB	Kattegat	Nov/2007	56°25′31″N	12°11′21″E	48	44	39
SKA07	A	SKAS	Kagerrak	Nov/2007	58°09′43″N	9°30′32″E	48	39	32
GER07	A	GER	North Sea	May/2007*	54°31′12″N	7°53′23″E	60	54	31
LINC07	A	LINC	North Sea	Aug/2007	53°19′96″N	0°25′63″E	96	47	0
LINC08	A	LINC	North Sea	Aug/2008	53°19′96″N	0°25′63″E	52	51	0
NOR07	A	NOR	North Sea	Aug/2007	53°00′70″N	1°33′62″E	46	45	14
NOR08	A	NOR	North Sea	Aug/2008	53°00′70″N	1°33′62″E	28	23	23
THA07	A	THA	North Sea	Aug/2007	51°27′80″N	1°20′00″E	96	94	17
THA08	A	THA	North Sea	Aug/2008	51°27′80″N	1°20′00″E	63	58	57
BEL07 s	A	BEL	North Sea	May/2007*	51°23′22″N	3°10′01″E	96	96	10
BEL07f	A	BEL	North Sea	Aug/2007	51°21′14″N	2°55′45″E	80	75	3
BEL08 s	A	BEL	North Sea	May/2008*	51°23′22″N	3°10′01″E	96	71	45
BEL08f	A	BEL	North Sea	Aug/2008	51°21′14″N	2°55′45″E	60	58	52
CEL08	A	CEL	Celtic Sea	Apr/2008*	50°49′00″N	5°01′00″W	96	81	44
IS08	A	IS	Irish Sea	Mar/2008*	52°13′00″N	5°20′00″W	96	91	44
ENG08	A	ENG	English Channel	Jul/2008	50°46′54″N	1°29′04″E	58	52	48
WCH09	A	WCH	English Channel	Aug/2009	49°39′41″N	2°07′38″W	80	37	0
BISA07	A	BISA	Bay of Biscay	Mar/2007*	46°53′00″N	2°47′00″W	95	88	3
BISB07	A	BISB	Bay of Biscay	Mar/2007*	45°36′00″N	1°24′00″W	61	55	5
BISC07	A	BISC	Bay of Biscay	Mar/2007*	46°20′00″N	1°53′00″W	48	45	26
TEX06	J	TEX	Wadden Sea	Aug/2006	52°58′11″N	4°56′35″E	58	46	21
TEX07	J	TEX	Wadden Sea	May/2007*	52°58′11″N	4°56′35″E	48	43	48
THA07j	J	THA	North Sea	Aug/2007	51°27′80″N	1°20′00″E	35	34	17
ZAN06	J	ZAN	Scheldt estuary	Sep/2006	51°23′50″N	4°06′59″E	74	70	15
ZAN07	J	ZAN	Scheldt estuary	Oct/2007	51°23′50″N	4°06′59″E	61	61	30
BEL08j	J	BEL	North Sea	May/2008*	51°23′22″N	3°10′01″E	36	32	27
Total							1,811	1,579	651

exact tests (Guo and Thompson 1992) implemented in Genepop v.2.3.4 (Raymond and Rousset 1995). The statistical significance was adjusted using Bonferroni correction to correct for multiple testing (Rice 1989).

Genetic diversity and population structure

For the microsatellite data, the observed and unbiased expected heterozygosity and the number of alleles were analyzed in Genetix v. 4.05 (Belkhir et al. 2004). Single- and multilocus F_{IS} (Weir and Cockerham 1984) and allelic richness (AR) per locus (El Mousadik and Petit 1996) were calculated with Fstat v.2.9.3 (Goudet 2001). Haplotype diversity (h), nucleotide diversity (Π) and their standard deviation, and the number of polymorphic sites were calculated for the mtDNA using DnaSP v.5 (Librado and Rozas 2009).

Genetic differentiation of the microsatellite data was analyzed with several methods. First, global and pairwise F_{ST} (Weir and Cockerham 1984) were estimated using Fstat. Second, the unbiased estimator of divergence 'D_{est}' of Jost (2008) was calculated with Smogd (Crawford 2010) because it is expected to perform better for populations with high allelic diversity and with alleles that are alternatively fixed in different populations (Jost 2008). A principal component analysis based on allele frequencies was performed with Pcagen v.1.2 (Goudet 2004) to visualize the relationships among samples; 10^4 randomizations were performed to test for the significance of axes.

Genetic structure based on the microsatellite data was further described using a Bayesian clustering method, implemented in the program Structure v.2.3 (Hubisz et al. 2009; Pritchard et al. 2000). This method estimates the number of genetic clusters in the data without making

a priori assumptions about population structure. The 'no-admixture' algorithm was used with information on sampling location included to assist the clustering, allowing for better performance for data with weak structure (Hubisz et al. 2009). Only adult samples were included to estimate the number of clusters in the data (K). For each simulation of K ($K = 1 - 10$), ten replicate runs were performed. The most likely number of clusters given the data was selected by choosing K with the largest log-likelihood according to Evanno et al. (2005). In total, 10^4 runs were used as burn-in and 10^5 MCMC repeats after burn-in, assuming correlated allele frequencies among populations. The proportional coefficients (Q) of individuals from STRUCTURE were plotted with the software DISTRUCT 1.1 (Rosenberg 2004) to visualize the patterns in clustering of predefined populations.

For the mitochondrial sequences, population differentiation was estimated using global (Weir and Cockerham 1984) and with pairwise F_{ST} and genetic distance of Tamura and Nei (1993), with ARLEQUIN v. 3.11 (Excoffier et al. 2005). Jost's D estimator ('D_{est}') was calculated using samples with more than 20 sequences, with SPADE (Chao and Shen 2010). Intraspecific relationships between the haplotypes were visualized with a median joining network, implemented in the software NETWORK v. 4.5.1.6 (www.fluxus-engineering.com). A multidimensional scaling analysis (MDS) was conducted with STATISTICA v.9 (Statsoft) to visualize genetic relationships among samples. This analysis was done based on the genetic distance of Tamura and Nei (1993). Stress values below 0.20 provide interpretable information regarding intersite relationships (Clarke 1993).

To compare the degree of population differentiation between both types of markers, the pairwise F_{ST} (Rousset 1997) matrix of the microsatellite data was correlated with the genetic distances of Tamura and Nei (1993) from the mtDNA, using a Mantel procedure (Mantel 1967) in Genetix. Significance was tested with 2000 permutations.

Isolation by distance

To test for a population differentiation pattern following an isolation-by-distance (IBD) model, a Mantel test (Mantel 1967) with the log of geographical distance and $\theta/(1 - \theta)$ for the microsatellite data and genetic distance for the mitochondrial data was carried out in GENETIX. Only adult samples were included, and for the sequence data, only those samples with more than 20 sequences were included. The correlation was tested first including all samples and secondly, using only the samples along the European continental shoreline. Significance was tested with 2000 permutations. Geographical distance between sampling locations was calculated using Microsoft ENCARTA World Atlas 2001. Distances were measured as the shortest coastal distances between sampling locations because adult sole remain restricted to waters less than 50 m deep.

Temporal stability of spawning populations, juvenile dispersal and genetic patchiness

To evaluate the degree of temporal variation, we performed a hierarchical analysis of molecular variance (AMOVA) in ARLEQUIN v.3.1 (Excoffier et al. 2005), comparing temporal replicates of the microsatellite data from the same geographical locations in the North Sea (2007 samples versus 2008 samples; namely LINC07, LINC08, NOR07, NOR08, THA07, THA08, BEL07 s, BEL07f, BEL08 s and BEL08f). We evaluated jointly the degree of spatial variation with AMOVA for those locations wherefore temporal replicates were available. Significance levels were determined after 10,100 permutations.

Genetic variation (AR, Ho, He, relatedness) and differentiation (F_{ST}) were compared between adult samples and juvenile samples in FSTAT and the confidence level tested using 10^4 permutations. To investigate whether we could assign juvenile sole to an adult population, an individual assignment test was done with the software GENECLASS2 (Piry et al. 2004) using the Bayesian method of Rannala and Mountain (1997). For this analysis, we only included the following adjacent populations as baseline samples: Thames (THA07 + THA08), Norfolk Banks (NOR07 + NOR08), Belgian Coast (BEL07s + BEL08s), German Bight (GER07), Eastern English Channel (ENG08), Western English Channel (WCH09) and Skagerrak (SKA07).

Results

The MICROCHECKER analysis indicated that loci F8-IIGT15, F14-IIGT16, Sos(AC)30 and Sos(AC)40 might be affected by null alleles or stuttering in the majority of the samples. Therefore, these loci were excluded from further statistical analyses. The average null allele frequency estimated by the EM algorithm in FREENA (Dempster et al. 1977) was lower than 10% in each of the 10 remaining loci. The test with LOSITAN confirmed selective neutrality for all 10 microsatellite loci. Significant deviations from Hardy–Weinberg equilibrium were found for 8 of the 280 tests conducted, after Bonferroni correction. Jackknifing over the loci did not alter the multilocus estimates of F_{IS} or F_{ST}.

Genetic variability

All loci were polymorphic, with the number of alleles per microsatellite locus within samples varying from 3 at locus F8-ICA9 to 27 at locus Sos(AC)20 (Supplementary Material, Table S2). The mean number of alleles per

locus per sample varied from 8.9 in NOR08 to 15.1 in the BEL07 s sample. The mean allelic richness based on 20 individuals did not differ much among samples, ranging from 8.19 in the Skagerrak (SKA07) to 9.51 in the Texel sample (TEX07) (Table 2). Average observed heterozygosities varied from 0.668 in the sample from Norfolk banks (NOR08) to 0.804 in the juvenile sole from the Belgian coast (BEL08j). Expected heterozygosity was uniformly high, ranging from 0.725 in NOR08 to 0.775 in the Thames sample of 2007 (THA07) (Table 2). F_{IS} values ranged from −0.06 in the sample from the Belgian coast juveniles (BEL08j) to 0.105 in the sample from the Belgian spawning and feeding population (BEL08 s and BEL08f). Twenty-one of the 28 samples showed low but significant F_{IS} values (Table 2).

The 651 mtDNA sequences of 590 bp contained 79 variable sites with 46 parsimony informative sites and 33 singleton variable sites (see Supplementary Material, Table S3 for details on variable nucleotide positions). The diversity was high, with in total 107 haplotypes detected (GenBank accession numbers JN571603–JN571709). Haplotype H2 (JN571604) was the most abundant haplotype, occurring in 40% of the specimens. Haplotypes H3 (JN571605), H5 (JN571607), H8 (JN571610) and H14 (JN571616) were shared between multiple sites, occurring in more than 20% of the specimens. There were 63 unique haplotypes; 6 haplotypes occurred only in the Skagerrak/Kattegat region; 5 haplotypes were unique to the Bay of Biscay, and 9 were unique to the Irish/Celtic Sea (Supplementary Material, Table S4). The overall level of genetic variability was high, with haplotype diversity and nucleotide diversities of 0.823 and 0.0049, respectively. The highest diversity was observed in the Skagerrak and the lowest in THA08 and BISC07 (excluding BISA07 with only 3 sequences) (Table 3).

Table 2 Estimates of genetic diversity of all samples based on 10 microsatellite markers

Sample code	He	Ho	MNA	AR	F_{IS}
Adults					
STO07	0.744	0.748	11.3	8.68	−0.004
KATA07	0.743	0.706	10.9	8.29	**0.050**
KATB07	0.743	0.761	10.9	8.33	−0.024
SKA07	0.730	0.710	10.0	8.19	0.028
GER07	0.746	0.703	11.8	8.55	**0.058**
LINC07	0.767	0.727	11.8	9.09	**0.053**
LINC08	0.761	0.732	12.6	9.23	**0.039**
NOR07	0.772	0.722	11.4	8.88	**0.065**
NOR08	0.725	0.668	8.9	8.51	**0.080**
THA07	0.775	0.740	14.2	9.03	**0.045**
THA08	0.747	0.739	12.7	8.95	0.011
BEL07s	0.769	0.718	15.1	9.09	**0.067**
BEL07f	0.767	0.710	12.9	8.83	**0.074**
BEL08s	0.771	0.691	14.5	9.43	**0.105**
BEL08f	0.771	0.691	12.5	9.07	**0.105**
CEL08	0.748	0.741	13.7	8.73	0.009
IS08	0.766	0.762	14.8	9.15	0.006
ENG08	0.760	0.702	12.4	8.98	**0.077**
WCH09	0.747	0.700	13.6	8.86	**0.064**
BISA07	0.766	0.743	14.2	9.15	**0.030**
BISB07	0.751	0.700	12.0	8.77	**0.068**
BISC07	0.763	0.726	11.5	8.80	**0.049**
Juveniles					
TEX06	0.758	0.743	11.0	8.42	0.020
TEX07	0.755	0.723	12.7	9.51	**0.043**
THA07j	0.750	0.686	11.5	9.43	**0.086**
ZAN06	0.756	0.706	12.9	8.88	**0.067**
ZAN07	0.773	0.732	13.1	9.18	**0.054**
BEL08j	0.753	0.804	9.4	8.43	**−0.069**

Non-biased expected heterozygosity (He), observed heterozygosity (Ho), mean number of alleles (MNA), allelic richness (AR, based on 20 individuals) and F_{IS} values. See Table 1 for sample codes. Significant F_{IS} values ($\alpha = 0.05$, Bonferroni corrected) are listed in bold

Population structure

Microsatellite DNA

The principal component analysis based on microsatellite data showed three groups (Fig. 2). The first group consisted of the samples of Kattegat, Skagerrak and the Belt Sea. The second group contained the North Sea samples, also including Irish and Celtic Sea and English Channel, but located at the periphery of the cluster. The sample from Kattegat B 2007 (KATB07) was located at the periphery zone of the second group, however, close to the first cluster. The third group clustered together the samples from the Bay of Biscay. The first axis explained 14.97% of the genetic variation ($F_{ST} = 0.002$; $P = 0.005$) and separated group 1 and group 3 entirely. The second PCA axis explained 9.57% of the total genetic variation ($F_{ST} = 0.001$; $P = 0.546$). The sample from Norfolk 2008 (NOR08) is an outlier in the PCA plot, although this might be an artifact resulting from the fact that this sample has less than 25 genotyped individuals, which is too low to analyze the genetic relationship with other samples based on the high variable microsatellite markers in *Solea solea* (Table 1).

The global F_{ST} based on the microsatellite data was very low 0.0024 (95% C.I., 0.001–0.003) but highly significant ($P < 0.0001$). Significant spatial differentiation based on pairwise F_{ST} values was observed between the samples from the most Northern region (Skagerrak, Kattegat and Belt Sea) and the rest of the samples. Within the North Sea, patterns are less clear, with temporal instability and several

Table 3 Molecular diversity indices of mtDNA sequences: sample code, number of sequences (N), number of polymorphic sites (S), number of haplotypes (h), haplotype diversity (Hd) (standard deviation) and nucleotide diversity (Π) (standard deviation)

Sample code	N	S	h	Hd	Π
Adults					
KATB07	39	18	13	0.825 (0.051)	0.0061 (0.00059)
SKA07	32	21	17	0.950 (0.019)	0.0066 (0.00056)
GER07	31	18	13	0.815 (0.066)	0.0044 (0.00081)
NOR07	14	18	12	0.967 (0.044)	0.0064 (0.00114)
NOR08	23	15	11	0.806 (0.079)	0.0055 (0.00080)
THA07	17	8	7	0.824 (0.064)	0.0032 (0.00046)
THA08	57	24	19	0.751 (0.060)	0.0043 (0.00062)
BEL07 s	10	12	6	0.867 (0.085)	0.0069 (0.00090)
BEL07f	3	7	3	1.000 (0.272)	0.0079 (0.00324)
BEL08 s	45	25	20	0.782 (0.065)	0.0047 (0.00067)
BEL08f	52	30	25	0.853 (0.044)	0.0055 (0.00051)
BEL08j	27	13	13	0.840 (0.059)	0.0039 (0.00075)
CEL08	44	17	16	0.813 (0.053)	0.0031 (0.00046)
IS08	44	21	21	0.794 (0.064)	0.0048 (0.00063)
ENG08	48	24	18	0.778 (0.060)	0.0053 (0.00064)
BISA07	3	6	2	0.667 (0.314)	0.0068 (0.00320)
BISB07	5	5	4	0.900 (0.161)	0.0041 (0.00117)
BISC07	26	14	12	0.754 (0.090)	0.0041 (0.00073)
Juveniles					
TEX06	21	14	11	0.819 (0.082)	0.0046 (0.00084)
TEX07	48	20	16	0.795 (0.056)	0.0045 (0.00062)
THA07j	17	13	10	0.875 (0.070)	0.0036 (0.00085)
ZAN06	15	13	9	0.886 (0.069)	0.0053 (0.00091)
ZAN07	30	19	16	0.871 (0.051)	0.0049 (0.00077)
Total data set	651	79	107	0.823 (0.015)	0.0049 (0.00017)

See Table 1 for sample codes

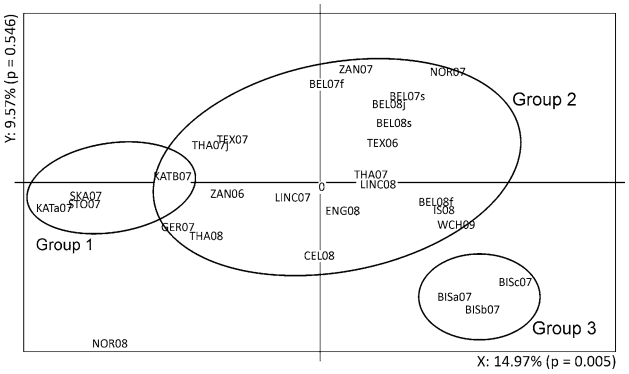

Fig. 2 Principal component analysis based on microsatellite data of adult and juvenile samples. For sample codes, see Table 1

genetic differences that did not remain significant after Bonferroni correction. Especially the sample from Norfolk Banks (2007) showed significant differentiation from other North Sea samples such as the German Bight, Thames and

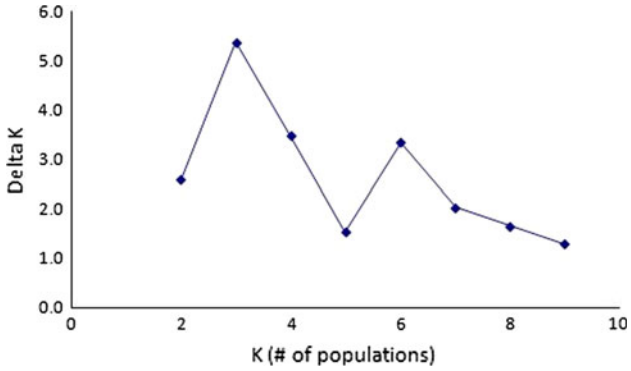

Fig. 3 Plot of delta *K* values obtained with STRUCTURE v.2.3 and according to Evanno et al. (2005). Results are based on the mean delta *K* of 10 replicates for each *K*

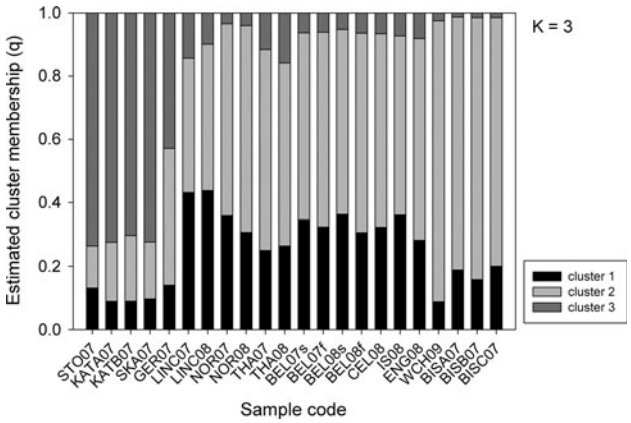

Fig. 4 *Solea solea*. Estimated probability of cluster membership for each sample, based on microsatellite data from STRUCTURE (*K* = 3). *x*-axis indicates samples codes (See Table 1)

Belgian Coast. The Bay of Biscay samples differed from the Baltic region and from many of the North Sea samples. Only the genetic differences between NOR07 and KATA07 (F_{ST} = 0.001) and between BEL07f and BISA07 (F_{ST} = 0.004) remained significant after Bonferroni correction (Supplementary Material, Table S5). The harmonic mean of Jost D_{est} was 0.006, and pairwise D_{est} was generally slightly higher than F_{ST} estimates but showed similar patterns (Supplementary Material, Table S6). No significant differences were observed in the results with or without the markers Sos(AC)45 and Sos(AC)6, which showed high F_{IS} estimates across many samples.

The clustering approach in STRUCTURE resulted in the assignment of the individuals to one of three hypothetical clusters (highest ΔK for $K = 3$) (Figs. 3, 4, 5). The lowest proportion of assignment to a particular cluster was 0.43 to cluster 2 for the sample from the German Bight (GER07). An equally large proportion (0.43) was assigned to cluster 3 for this specific sample. The highest proportion of assign-

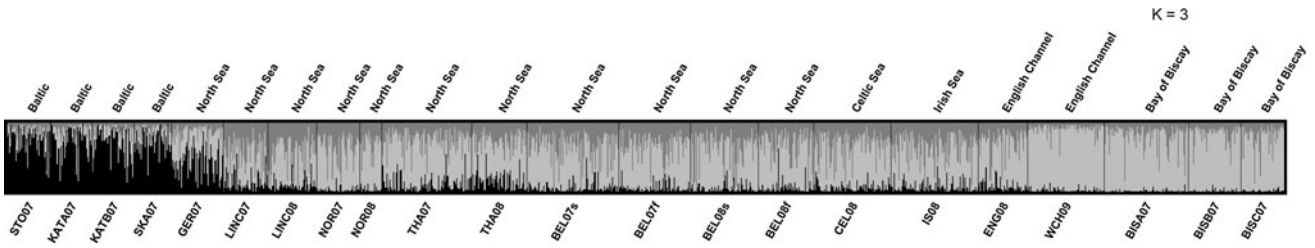

Fig. 5 *Solea solea.* Individual assignment based on Bayesian clustering method based on microsatellite data from STRUCTURE. *Each bar* represents an individual with its probability of membership to one of the hypothetical clusters ($K = 3$). Labels on top indicate sampling regions; labels at the bottom indicate samples codes. See Table 1 for more information on samples

ment to a cluster was 0.887 for the Western Channel sample to the second cluster. The majority of the samples from the Belt Sea, Kattegat and Skagerrak had the highest assignment value for the third cluster. Most of the other individuals from the North Sea samples showed highest assignment to the first or the second cluster. Although there are no obvious population assignments to cluster 1, visual inspection suggests that samples from the Bay of Biscay differ slightly from samples from the North Sea (Fig. 5). The western English Channel seems more affiliated with the Biscay cluster, while the Eastern English Channel resembles most to the North Sea cluster. The Irish and Celtic Sea are very similar to samples from the North Sea. Given the presence of two strongly geographically defined clusters and the challenge of clustering methods under an IBD model, we forced $K = 2$ to evaluate the geographical distribution of cluster members over a latitudinal gradient and to identify potential directional dispersal. The samples from the Belt Sea, the Kattegat and the Skagerrak clustered well together, while all other samples showed the highest probability for the other cluster. The German Bight consisted of a mixture of both clusters. Leaving out the samples of the Baltic region did not improve the spatial resolution within the North Sea region (data not shown). The Baltic cluster clearly showed a restricted distribution outside the Baltic region, while the Biscay area cluster was much broader distributed along the North-Eastern Atlantic coast.

Mitochondrial DNA

MtDNA showed similar patterns of spatial differentiation but higher absolute values with $G_{ST} = 0.0037$ and pairwise F_{ST} values ranging from -0.020 between Eastern Channel and Norfolk Banks to 0.089 between the Kattegat and Celtic Sea. Spatial differentiation was most obvious between the samples from Skagerrak/Kattegat and all other samples. Also, the Celtic Sea was genetically distinct, with 6 out of 22 significant pairwise F_{ST} comparisons (Supplementary Material, Table S7). Differentiation remained significant after Bonferroni correction for the comparisons of SKA07 and KATB07 with CEL08.

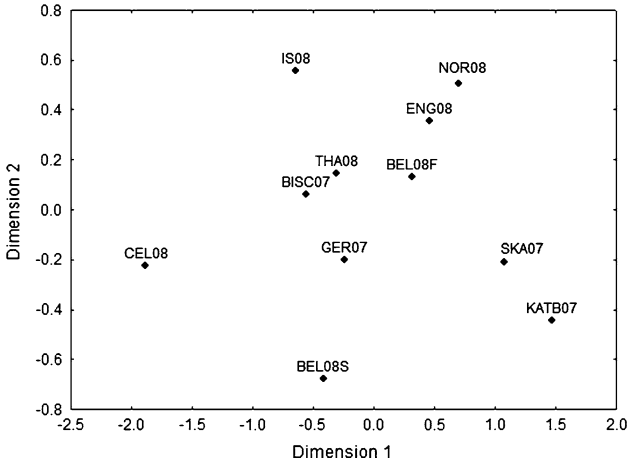

Fig. 6 Multidimensional scaling plot based on pairwise genetic distance (Tamura and Nei 1993) of mitochondrial data of sole. See Table 1 for sample codes

The haplotype network based on cyt *b* illustrates the large number of haplotypes with only one mutation difference from the central haplotype (Fig. S1). Besides one central haplotype, there are three other haplotypes present in high frequency, with many derived haplotypes. There are no distinct geographical groupings of samples. However, the multidimensional scaling analysis based on mitochondrial data (stress value = 0.018) also showed the clear separation of the Kattegat/Skagerrak samples from the rest of the North Sea samples. The Celtic and Irish Sea were also separated from the other samples (Fig. 6). No correlation was evidenced between the pairwise genetic distances of the mtDNA data and the genetic distances of the microsatellite data (Mantel $r = 0.026$, $P = 0.384$).

Isolation by distance

A strong and significant IBD pattern was found with the microsatellite data including only the adult samples along the continental shoreline ($r = 0.697$; $P = 0.002$) (Fig. 7). However, by including non-continental samples from the British Isles and Ireland, a positive correlation between geographic and genetic distance disappeared. As for the cyt *b*

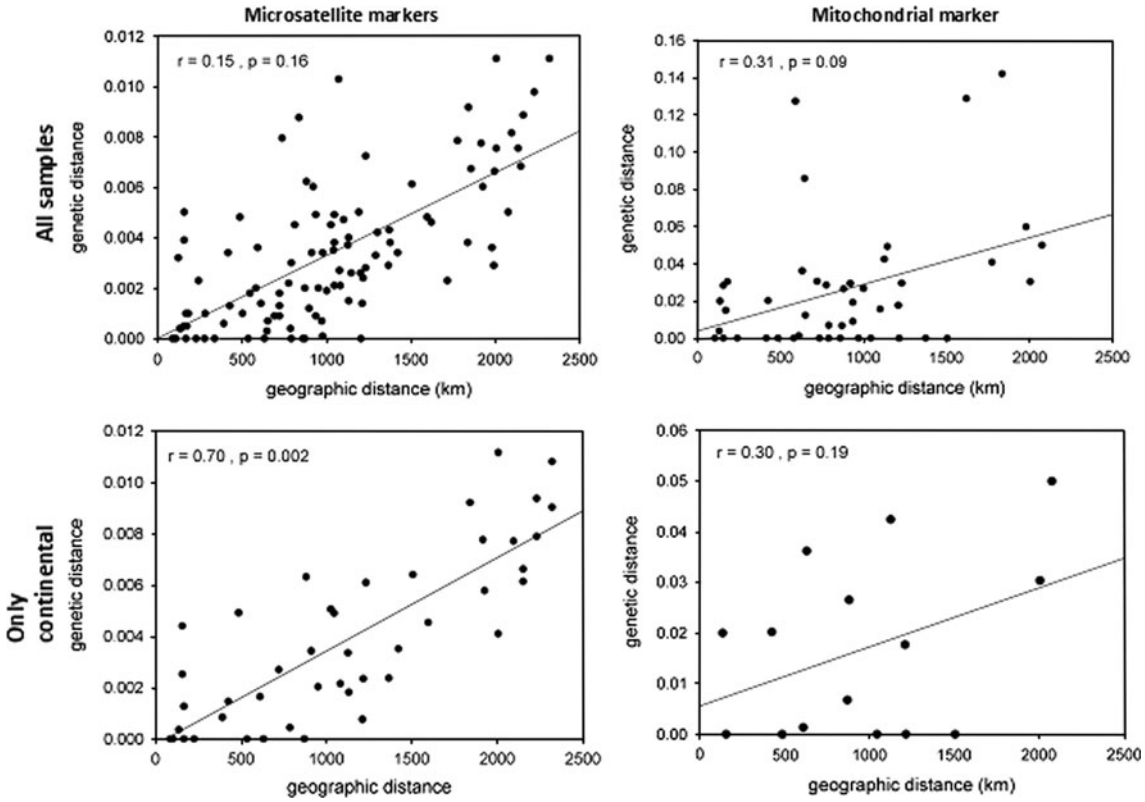

Fig. 7 Relationship between genetic distance and geographic distance of sole for microsatellite markers (*left panels*) and for mitochondrial marker (*right panels*). Upper graphs include all samples and lower graphs include only samples along the continental shoreline. See text for more details

sequences, no significant pattern of IBD could be detected, although a positive linear trend was obvious (all samples: $r = 0.307$; $P = 0.09$; only samples along the continental shoreline: $r = 0.298$, $P = 0.19$) (Fig. 7).

Temporal stability of spawning populations

The AMOVA analysis indicated that almost all genetic variation (99.9%) was attributed to the 'within sample' component ($F_{ST} = 0.002$, $P = 0.01$), while the variation among sampling years was only 0.21% and not significant ($F_{CT} = 0.0002$, $P = 0.93$). Because of the low differentiation between samples of locations within the North Sea, no significant variation among locations could be detected with the AMOVA analysis. Except for the samples from Norfolk Banks (NOR07 and NOR08), no pairwise comparisons (F_{ST}) of temporally replicated samples from the same location were significant (Supplementary Material, Table S5). Pairwise F_{ST} among the temporal replicates ranged from −0.0002 between the feeding samples of the Belgian Coast (BEL07f vs. BEL08f) to 0.0117 between the temporal samples from the Norfolk Banks. The significant F_{ST} at Norfolk Banks ($F_{ST} = 0.0117$; $P = 0.01$) might, however, be attributed to the low sample size in 2008 (N = 23). There was no significant genetic differentiation between the Belgian sole samples caught during the spawning season (BEL07 s and BEL08 s) and the samples caught during late summer of the same year (BEL07f and BEL08f) (Supplementary Material, Table S5).

Juvenile dispersal and genetic patchiness

Genetic variation in juvenile and adult samples did not differ significantly from each other. Mean observed heterozygosity was 0.721 for adult sole and 0.720 for juvenile sole ($P = 0.920$). The mean sample F_{ST} values among both adult and juvenile samples were 0.002. There was no indication for an increase in relatedness in the juvenile samples, which would be the consequence of genetic patchiness (relatedness (Queller and Goodnight 1989): adults = 0.005, juveniles: 0.003; $P = 0.470$). No within-sample relatedness values differed from random mating either.

There was a marginally significant overall differentiation among the juvenile samples from the different locations ($F_{ST} = 0.0018$; $P = 0.058$). Pairwise F_{ST} estimates ranged from −0.001 to 0.008. Some differentiation was apparent between the juveniles from Texel (both years) and the Belgian coast and between the Scheldt (ZAN06) and the Belgian Coast (BEL08j), but these differences did not remain significant after Bonferroni correction. There was

Table 4 Assignment percentages of juveniles to adult baselines based on microsatellite analysis with GENECLASS. See Table 1 for sample codes

	Assigned population						
	SKA	GER	NOR	THA	BEL	ENG	WCH
Juvenile sample							
TEX06	15.2	15.2	8.7	15.2	26.1	10.9	8.7
TEX07	4.7	2.3	20.9	14.0	37.2	11.6	9.3
THA07j	17.6	0.0	17.6	14.7	29.4	11.8	8.8
ZAN06	12.9	5.7	14.3	21.4	24.3	8.6	12.9
ZAN07	6.6	11.5	11.5	13.1	36.1	13.1	8.2
BEL08j	18.8	15.6	12.5	9.4	15.6	9.4	18.8

no temporal differentiation present between the juvenile samples from the Wadden Sea (TEX) or from the Scheldt (ZAN) during two successive years (Supplementary Material, Table S6). Based on the PCA plot, all juvenile samples are clustered within the 'North Sea' adult group. Temporal replicates of the same juvenile nursery location are not necessarily lying closely together on the graph (Fig. 2).

As expected from the low genetic differentiation between adult samples, the juvenile assignment test in GENECLASS resulted in almost random assignments of juveniles to the adult baselines (Table 4). Even if the baseline data were pooled into larger groups, there were no consistent regional patterns (data not shown).

Discussion

Overall, our study showed the most distinct pattern of genetic divergence when proposing an IBD model at the regional scale, namely between sole populations from the Bay of Biscay and the Baltic transition region. This pattern was supported by the strong correlation between geographic and genetic distance based on microsatellite markers and by individual-based cluster analyses identifying two main overlapping population clusters. No evidence for closed populations or genetic patchiness could be observed in the North Sea, indicating a large effective population size of spawning aggregations (Cuveliers et al. 2011) and/or sufficient genetic connectivity between North Sea populations.

Population structure

The strongest signal for population genetic structure in sole was found between populations from the Baltic transition region (Skagerrak/Kattegat/Belt Sea) and the North Sea, as well based on microsatellite as on mitochondrial data. This was in contrast to the previous population genetic study of sole by Rolland et al. (2007), which used only three nuclear EPIC markers. Our results are concordant with observations for several other marine fish species in the region, such as for turbot (Nielsen et al. 2004), herring (André et al. 2011; Bekkevold et al. 2005), plaice (Hoarau et al. 2002), hake (Lundy et al. 1999), cod (Nielsen et al. 2003) and sand goby (Larmuseau et al. 2010a). The significant differentiation between the North Sea and the Baltic transition zone is mainly attributed to geographic isolation, a bottleneck, selection on adaptive traits in Baltic populations or a combination thereof (Johannesson and André 2006; Larmuseau et al. 2010b). However, a recent study on plaice (*Pleuronectes platessa* L.), a flatfish species sharing similar ecological niches with sole, did not find any significant population differentiation between the Baltic and the North Sea (Was et al. 2010). The authors attributed this contrasting pattern between plaice and other (flat) fishes from this region to a small increase in gene flow between plaice populations of the two basins, leading to a perceived absence of genetic differentiation. Even when significant differentiation was observed between North Sea and Baltic Sea populations of the other species, the pairwise F_{ST} values were mostly always quite low (<0.006) for neutral markers. The Skagerrak, Kattegat and Belt Sea provide a transition zone between the saline North Sea and the brackish Baltic proper. Based on the hydrodynamics of the area (Rodhe 1996), it is expected that larval exchange is mainly directed into the Skagerrak (Knutsen et al. 2004; Stenseth et al. 2006). For sole, this region represents the north-eastern limit of its distribution, as the species has only occasionally been spotted in the Baltic proper (Bacevičius and Karalius 2008). Nevertheless, further research on the interaction between the Baltic and North Sea sole populations is required as one Kattegat sample was located at the periphery of the North Sea cluster on the PCA plot based on the microsatellite markers, suggesting a more subtle pattern of genetic exchange between the two basins (Fig. 2).

Allozyme data on sole indicated near panmixia between the North Sea, the Bay of Biscay and the Irish Coast populations and suggested movement of individuals through the English Channel (Exadactylos et al. 2003; Exadactylos et al. 1998), with a pattern of IBD (Kotoulas et al. 1995). In our study, significant genetic differentiation based on microsatellite markers between the Bay of Biscay and the North Sea was apparent from the pairwise F_{ST} values, while the Bayesian clustering approach identified a Biscay cluster extending to the north, indicating an IBD pattern. In contrast to earlier studies on sole using allozyme and RAPD markers, we found small genetic differences between the Irish/Celtic Sea and some of the North Sea samples. However, these differences were not temporally stable, likely due to transient dispersal events into the Channel and subsequently the North Sea. A lack of strong and stable

differentiation between the Irish Sea and North Sea has also been documented for other North Atlantic fish such as plaice (Was et al. 2010), turbot (Hemmer-Hansen et al. 2007a) and whiting (Charrier et al. 2007).

Finally, the overall very low genetic divergence among sole populations within the North Sea is consistent with other marine fish (Borsa et al. 1997; Hemmer-Hansen et al. 2007a, b; Hoarau et al. 2002; Nielsen et al. 2004). Weak population structure may result from the recent colonization histories following postglacial sea level rise and high effective population size (N_e), reducing the effects of genetic drift detected by microsatellite markers. For sole in the North Sea, effective population size estimates based on the analysis of archived otolith samples with microsatellite markers were large (Cuveliers et al. 2011). Therefore, low F_{ST} values do not necessarily imply high gene flow (Hedgecock et al. 2007; Larmuseau et al. 2010a; Whitlock and McCauley 1999), but could be due to the long time span needed before F_{ST} reaches equilibrium. The effects of genetic drift should appear faster in mtDNA due to its four times smaller effective size than microsatellite markers, enabling the detection of population structure at an earlier stage (Buonaccorsi et al. 2001). In populations with high effective population sizes, such as in sole, mitochondrial markers might therefore show higher resolution power even for recently diverged populations in comparison with microsatellites, although the latter type of genetic markers has a higher mutation rate (Larmuseau et al. 2010a). Nevertheless, even though the mitochondrial signal of differentiation (pairwise F_{ST}) was an order of magnitude higher than differentiation based on microsatellite markers, the power of spatial resolution and general differentiation patterns were similar in our study, pointing to the Baltic as most differentiated region, followed by the Biscay area and the Irish–Celtic Sea region. No evidence of IBD was found using the mitochondrial marker, but there was a positive trend between genetic distance and geographic distance. The limited resolution with the mitochondrial marker in our study could be due to the single-locus effect of cyt *b*, compared to the use of multiple independent microsatellite loci, sex-biased dispersal or the lower sampling size used for the mtDNA marker (Grant and Waples 2000).

The clear IBD pattern between the continental sole samples indicates that genetic drift might be strong enough to be detectable and that the population structure of sole is influenced by subtle levels of gene flow as well, next to low genetic drift due to high effective population sizes. Gene flow between sole populations could occur either through passive drift during the larval stage or through active migration of (sub)adults. On the one hand, passive mixing during the egg or larval stages might directly cause an exchange of genetic material originating from different spawning aggregations. Although a particle tracking model coupled to a hydrodynamic model has shown that the exchange of sole larvae among the main spawning sites is much lower than the number of larvae retained, some connections among spawning sites are possible with the level of exchange mainly depending on hydrodynamic variability (Savina et al. 2010). An integrative analysis of genetic, biological and hydrodynamic analysis is a next step toward disentangling the proportional effect of all three factors in the population structure of sole. On the other hand, a limited number of active subadult or adult migrants between populations across the Channel would suffice to prevent strong genetic differentiation (Waples 1998). Although sole undertake relatively short migrations, active movements of tagged sole across the English Channel have been observed (Burt and Millner 2008). Average migration distances of mature fish were estimated at 75 km in the spawning season and up to 150 km during the rest of the year (Burt and Millner 2008). The English Channel may thus act as a route for gene flow between the British Isles and the southern North Sea through passive drift of larvae following the residual current flowing eastward or through active migration of (sub)adults. On the other hand, the clear IBD pattern for continental samples disappeared when including the British Isles samples, suggesting that gene flow events across the English Channel are not as evident as expected solely based on the relative small geographic distance between the British Isles and the European continent. Nevertheless, only few North Sea samples in this study included fully spawning individuals, a common logistic issue in many other studies on commercial marine fish. Therefore, subtle differentiation patterns may still be overlooked across the North Sea, and a cautious interpretation of the results is recommended because of the sampling design.

Sampling design and temporal stability

Besides the above-mentioned biological reasons for the observed genetic patterns, one should be careful when interpreting genetic results from highly dispersive marine taxa. Technical issues such as scoring errors and limited resolution power may blur the subtle differentiation signal (Waples 1998). We have carefully checked the genotypes and removed all markers with a high occurrence of null alleles. Still, estimates of population differentiation and the chances in detecting significant differentiation diminish when heterozygosity and allelic richness are very high. This appears to be one of the consequences of size homoplasy at microsatellite markers (O'Reilly et al. 2004). A reanalysis of the data without the two markers that gave quite high F_{IS} estimates across many samples, mainly Sos(AC)45 and Sos(AC)6 (Table S2), showed that both markers did not affect significantly the observed patterns based on data of all ten microsatellites (results not shown). Furthermore,

Cuveliers et al. (2011) showed that the statistical power was largely sufficient to detect even small genetic differentiations, as a true F_{ST} of 0.001 could be detected with a probability of 98% based on the observed allele frequencies of a subset (11 loci) of the 15 microsatellite loci used in this study. Finally, sampling did not always occur within the spawning period, the moment when the spawning populations have their most distinct signature. Even though the summer samples taken along the English coast were caught at the spawning grounds, they might represent mixed feeding aggregations. Nevertheless, the sole caught off the Belgian coast in the spawning season in spring were not significantly different from those sampled in late summer. This was the case in two consecutive years, suggesting temporal stability of the local spring and summer population. None of the pairwise tests between temporal samples taken from the same location in both years were significant.

Juvenile dispersal and genetic patchiness

With limited larval dispersal and a large variance in reproductive success, we expected 1) the genetic composition of juveniles to reflect neighboring adult spawning aggregations and 2) juvenile individuals to show reduced genetic variation and/or a chaotic genetic structure. Although juvenile sole dispersal is presumed to be limited (Cuveliers et al. 2010), no self-assignment could be achieved to nearby spawning populations; similar levels of genetic diversity in juvenile sole compared to adult fish were found. The absence of distinct genetic differentiation among juvenile samples from different nurseries might indicate either that adult fish form a single population or that juveniles arriving and settling at the nursery grounds represent admixture of progeny from different populations. Since we could not detect stronger genetic differentiation among adult samples within the North Sea, we believe the first hypothesis is more likely. The genetic differentiation observed among adult samples within the North Sea with the current set of neutral markers might be just too low for an accurate assignment of juveniles (Hedgecock et al. 2007). Although no significant temporal differentiation was observed between juveniles from the same location in successive years, these samples did not always cluster together, suggesting some temporal variation in genetic composition. Such temporal differences might arise due to chaotic genetic patchiness (Larson and Julian 1999; Pujolar et al. 2006). Also for juvenile sole in the Bay of Biscay (Guinand et al. 2008) and for juvenile plaice in the Irish Sea (Watts et al. 2004), no genetic differentiation was found among nurseries with neutral genetic markers.

Even if neutral genetic differentiation is small in the presence of gene flow, local adaptation might still characterize the different sole aggregations and should be the next step in sole research. It is well established that in large populations, differentiation at selective traits may occur faster than the effects of genetic drift (Hauser and Carvalho 2008; Larmuseau et al. 2009a; Nielsen et al. 2009), so that a large battery of genetic markers under selection (e.g., gene-linked SNPs; Helyar et al. 2011) might show more power to detect fine-scale and biologically relevant population structure in high gene flow–high *Ne* species.

Conclusion

Sole populations in the North-East Atlantic Ocean are characterized by a clear IBD pattern and genetic homogeneity within basins. Genetic differentiation was highest between the Skagerrak/Kattegat region and the rest of the North Atlantic. Some differentiation also existed between the North Sea and the Bay of Biscay and subtle differences between the North Sea and the Celtic Sea/Irish Sea. Juvenile genetic variability was very similar to adult diversity pointing to large spawning aggregations, and spatial genetic differences among nursery grounds were very small. We believe that dispersal of fish larvae away from local spawning aggregations may lead to a mixing of juveniles from different origins at nearby nursery grounds. Once settled, movement of juveniles and (sub)adults is probably limited. The observed genetic structure seems to be relatively stable on a short temporal timescale, but further research, incorporating postlarval sampling, hydrodynamic modeling and functional markers linked to relevant biological traits, might help to fine-tune the dispersal model applicable for sole.

Acknowledgments This study was part of the WESTBANKS project funded by the Belgian Science Policy (BELSPO; contract no. SD/BN/01A). We thank all scientists and fishermen who contributed to the collection of samples: E. Nielsen and colleagues (DTU-AQUA), M.L. Bégout and K. Mahé (Ifremer), L. Bolle (IMARES), H. van der Veer (NIOZ); U. Damm and N. Rohlf (ISH-BFAFI), ILVO-Fisheries, the crew of RV *Zeeleeuw* and RV *Belgica*, E. Diopere, K. Vancampenhout, S. Geldof and J. Guelinckx (K.U.Leuven). We also thank the Editor and two anonymous referees for their useful corrections and comments on an earlier version of this manuscript. E.C. acknowledged a grant of the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen). M.H.D.L. and G.E.M. are postdoctoral researchers funded by the Fund for Scientific Research (FWO Vlaanderen).

References

André C, Larsson LC, Laikre L et al (2011) Detecting population structure in a high gene-flow species, Atlantic herring (*Clupea harengus*): direct, simultaneous evaluation of neutral vs putatively selected loci. Heredity 106:270–280

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

Bacevičius E, Karalius S (2008) A common sole (*Solea solea* Linnaeus, 1758): Actinopterygii: Pleuronectiformes: Soleidae) caught in the coastal zone of Lithuania. Acta Zoologica Lituanica 18:174–179

Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. Proc R Soc B Biol Sci 263:1619–1626

Bekkevold D, Andre C, Dahlgren TG et al (2005) Environmental correlates of population differentiation in Atlantic herring. Evolution 59:2656–2668

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2004) GENETIX 4.05, Logiciel Sous Windows™ Pour la Génétique Des Populations. Université de Montpellier II, Montpellier (France)

Borremans C (1987) North Sea spawning grounds of the sole (*Solea solea*) located from the 1984 Belgian plankton survey. In: Kullander SO, Fernholm B (eds) Proceedings of the V congress of european ichthyologists. Swedish Museum of Natural History, Stockholm, pp 187–191

Borsa P, Blanquer A, Berrebi P (1997) Genetic structure of the flounders *Platichthys flesus* and *P. stellatus* at different geographic scales. Mar Biol 129:233–246

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

Burt GJ, Millner RS (2008) Movement of sole in the southern North Sea and eastern English Channel from tagging studies (1955–2004). Science Series Technical, Report 144

Cabral HN, Castro F, Linhares D, Alexandrino P (2003) Genetic differentiation of *Solea solea* (Linnaeus, 1758) and *Solea senegalensis* Kaup, 1858, (Pisces: Pleuronectiformes) from several estuarine systems of the Portuguese coast. Scientia Marina 67:43–52

Chao A, Shen T-J (2010) Program SPADE (Species prediction and diversity estimation). Program and user's guide published at http://chao.stat.nthu.edu.tw

Chapuis MP, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. Mol Biol Evol 24:621–631

Charrier G, Coombs SH, McQuinn IH, Laroche J (2007) Genetic structure of whiting *Merlangius merlangus* in the northeast Atlantic and adjacent waters. Mar Ecol Prog Ser 330:201–211

Clarke KR (1993) Nonparametric multivariate analyses of changes in community structure. Aust J Ecol 18:117–143

Crawford NG (2010) SMOGD: software for the measurement of genetic diversity. Mol Ecol Resour 10:556–557

Cuveliers EL, Geffen AJ, Guelinckx J et al (2010) Microchemical variation in juvenile *Solea solea* otoliths as a powerful tool for studying connectivity in the North Sea. Mar Ecol Prog Ser 401:211–220

Cuveliers EL, Volckaert FAM, Rijnsdorp AD, Larmuseau MHD, Maes GE (2011) Temporal genetic stability and high effective population size despite fisheries-induced life-history trait evolution in the North Sea sole. Mol Ecol 20:3555–3568

De Clerck R, Van de Velde J (1973) A study of the spawning and nursery areas of soles along the Belgian coast. In: ICES report of demersal fish committee

Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via Em Algorithm. J R Stat Soc B Methodol 39:1–38

DeWoody JA, Avise JC (2000) Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. J Fish Biol 56:461–473

El Mousadik A, Petit RJ (1996) High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L) Skeels] endemic to Morocco. Theor Appl Genet 92:832–839

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

Exadactylos A, Geffen AJ, Thorpe JP (1998) Population structure of the Dover sole, *Solea solea* L., in a background of high gene flow. J Sea Res 40:117–129

Exadactylos A, Geffen AJ, Panagiotaki P, Thorpe JP (2003) Population structure of Dover sole *Solea solea*: RAPD and allozyme data indicate divergence in European stocks. Mar Ecol Prog Ser 246:253–264

Excoffier L, Laval G, Schneider S (2005) ARLEQUIN ver.3.0: an integrated software package for population genetics data analysis. Evol Bioinform Online 1:47–50

Fogarty MJ, Botsford LW (2007) Population connectivity and spatial management of marine fisheries. Oceanography 20:112–123

Fonds M (1979) Laboratory observations on the influence of temperature and salinity on development of the eggs and growth of the larvae of *Solea solea* (Pisces). Mar Ecol Prog Ser 1:91–99

Funes V, Zuasti E, Catanese G, Infante C, Manchado M (2004) Isolation and characterization of ten microsatellite loci for Senegal sole (*Solea senegalensis* Kaup). Mol Ecol Notes 4:339–341

Garoia F, Marzola S, Guarniero I, Trentini M, Tinti F (2006) Isolation of polymorphic DNA microsatellites in the common sole *Solea vulgaris*. Mol Ecol Notes 6:144–146

Garoia F, Guarniero I, Grifoni D, Marzola S, Tinti F (2007) Comparative analysis of AFLPs and SSRs efficiency in resolving population genetic structure of Mediterranean *Solea vulgaris*. Mol Ecol 16:1377–1387

Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3.)

Goudet J (2004) PCAGEN [Computer program]. Available from http://wwww2.unil.ch/popgen/softwares/pcagen.html

Grant WS, Waples RS (2000) Spatial and temporal scales of genetic variability in marine and anadromous species: implications for fisheries oceanography. In: Harrison PJ, Parson TR (eds) Fisheries oceanography: an integrative approach to fisheries oceanography. Blackwell Science, Oxford, pp 61–93

Guarniero I, Franzellitti S, Ungaro N et al (2002) Control region haplotype variation in the central Mediterranean common sole indicates geographical isolation and population structuring in Italian stocks. J Fish Biol 60:1459–1474

Guinand B, Rolland JL, Bonhomme F (2008) Genetic structure of the common sole (*Solea solea*) in the Bay of Biscay: nurseries as units of selection? Estuar Coast Shelf Sci 78:316–326

Guo SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 48:361–372

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Window 95/98/NT. Nucleid Acids Symp Ser 41:95–98

Hauser L, Carvalho GR (2008) Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. Fish Fish 9:333–362

Hedgecock D, Barber PH, Edmands S (2007) Genetic approaches to measuring connectivity. Oceanography 20:70–79

Helyar SJ, Hemmer-Hansen J, Bekkevold D, Taylor MI, Ogden R, Limborg M, Cariani A, Maes GE, Diopere E, Carvalho GR, Nielsen EE (2011) Application of SNPs for population genetics of non-model organism: new opportunities and challenges. Mol Ecol Resour 11:123–136

Hemmer-Hansen J, Nielsen EE, Frydenberg J, Loeschcke V (2007a) Adaptive divergence in a high gene flow environment: *Hsc70* variation in the European flounder (*Platichthys flesus* L.). Heredity 99:592–600

Hemmer-Hansen J, Nielsen EE, Gronkjaer P, Loeschcke V (2007b) Evolutionary mechanisms shaping the genetic population structure of marine fishes; lessons from the European flounder (*Platichthys flesus* L.). Mol Ecol 16:3104–3118

Hoarau G, Rijnsdorp AD, Van der Veer HW, Stam WT, Olsen JL (2002) Population structure of plaice (*Pleuronectes platessa* L.) in northern Europe: microsatellites revealed large-scale spatial and temporal homogeneity. Mol Ecol 11:1165–1176

Horwood J (2001) Population biology and ecology of the sole. Nat Resour Model 14:233–256

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. Mol Ecol Resour 9:1322–1332

Iyengar A, Piyapattanakorn S, Stone DM et al (2000) Identification of microsatellite repeats in Turbot (*Scophthalmus maximus*) and Dover sole (*Solea solea*) using a RAPD-based technique: characterization of microsatellite markers in Dover sole. Mar Biotechnol 2:49–56

Johannesson K, André C (2006) Life on the margin: genetic isolation and diversity loss in a peripheral marine ecosystem, the Baltic Sea. Mol Ecol 15:2013–2029

Jones GP, Srinivasan M, Almany GR (2007) Population connectivity and conservation of marine biodiversity. Oceanography 20:100–111

Jørgensen HBH, Hansen MM, Bekkevold D, Ruzzante DE, Loeschcke V (2005) Marine landscapes and population genetic structure of herring (*Clupea harengus* L.) in the Baltic Sea. Mol Ecol 14:3219–3234

Jost L (2008) G(ST) and its relatives do not measure differentiation. Mol Ecol 17:4015–4026

Knutsen H, Jorde PE, Andre C, Stenseth NC (2003) Fine-scaled geographical population structuring in a highly mobile marine species: the Atlantic cod. Mol Ecol 12:385–394

Knutsen H, André C, Jorde PE et al (2004) Transport of North Sea cod larvae into the Skagerrak coastal populations. Proc R Soc B Biol Sci 217:1337–1344

Knutsen H, Olsen EM, Jorde PE et al (2011) Are low but statistically significant levels of genetic differentiation in marine fishes 'biologically meaningful'? A case study of coastal Atlantic cod. Mol Ecol 20:768–783

Kotoulas G, Bonhomme F, Borsa P (1995) Genetic structure of the common sole *Solea vulgaris* at different geographic scales. Mar Biol 122:361–375

Larmuseau MHD, Raeymaekers JAM, Ruddick KG, Van Houdt JKJ, Volckaert FAM (2009a) To see in different seas: spatial variation in the rhodopsin gene of the sand goby (*Pomatoschistus minutus*). Mol Ecol 18:4227–4239

Larmuseau MHD, Van Houdt JKJ, Guelinckx J, Hellemans B, Volckaert FAM (2009b) Distributional and demographic consequences of Pleistocene climate fluctuations for a marine demersal fish in the north-eastern Atlantic. J Biogeogr 36:1138–1151

Larmuseau MHD, Raeymaekers JAM, Hellemans B, Van Houdt JKJ, Volckaert FAM (2010a) Mito-nuclear discordance in the degree of population differentiation in a marine goby. Heredity 105:532–542

Larmuseau MHD, Vancampenhout K, Raeymaekers JAM, Van Houdt JKJ, Volckaert FAM (2010b) Differential modes of selection on the rhodopsin gene in coastal Baltic and North Sea populations of the sand goby, *Pomatoschistus minutus*. Mol Ecol 19:2256–2268

Larson RJ, Julian RM (1999) Spatial and temporal genetic patchiness in marine populations and their implications for fisheries management. CALCOFI Report 40:94–99

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–1452

Lukoschek V, Waycott M, Keogh JS (2008) Relative information content of polymorphic microsatellites and mitochondrial DNA for inferring dispersal and population genetic structure in the olive sea snake, *Aipysurus laevis*. Mol Ecol 17:3062–3077

Lundy CJ, Moran P, Rico C, Milner RS, Hewitt GM (1999) Macrogeographical population differentiation in oceanic environments: a case study of European hake (*Merluccius merluccius*), a commercially important fish. Mol Ecol 8:1889–1898

Mantel N (1967) The detection of disease clustering and a generalised regression approach. Cancer Res 27:209–220

Nielsen EE, Hansen MM, Ruzzante DE, Meldrup D, Gronkjaer P (2003) Evidence of a hybrid-zone in Atlantic cod (*Gadus morhua*) in the Baltic and the Danish Belt Sea revealed by individual admixture analysis. Mol Ecol 12:1497–1508

Nielsen EE, Nielsen PH, Meldrup D, Hansen MM (2004) Genetic population structure of turbot (*Scophthalmus maximus* L.) supports the presence of multiple hybrid zones for marine fishes in the transition zone between the Baltic Sea and the North Sea. Mol Ecol 13:585–595

Nielsen EE, Hemmer-Hansen J, Larsen PF, Bekkevold D (2009) Population genomics of marine fishes: identifying adaptive variation in space and time. Mol Ecol 18:3128–3150

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

Palsbøll PJ, Berube M, Allendorf FW (2007) Identification of management units using population genetic data. Trends Ecol Evol 22:11–16

Pampoulie C, Gysels ES, Maes GE et al (2004) Evidence for fine-scale genetic structure and estuarine colonisation in a potential high gene flow marine goby (*Pomatoschistus minutus*). Heredity 92:434–445

Piry S, Alapetite A, Cornuet JM et al (2004) GENECLASS2: a software for genetic assignment and first-generation migrant detection. J Hered 95:536–539

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

Pujolar JM, Maes GE, Volckaert FAM (2006) Genetic patchiness among recruits in the European eel *Anguilla anguilla*. Mar Ecol Prog Ser 307:209–217

Queller DC, Goodnight KF (1989) Estimating relatedness using genetic markers. Evolution 43:258–275

Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. Proc Nat Acad Sci USA 94:9197–9201

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J Hered 86:248–249

Reiss H, Hoarau G, Dickey-Collas M, Wolff WJ (2009) Genetic population structure of marine fish: mismatch between biological and fisheries management units. Fish Fish 10:361–395

Rice WR (1989) Analyzing tables of statistical tests. Evolution 43:223–225

Rijnsdorp AD, Vingerhoed B (1994) The ecological significance of geographical and seasonal differences in egg size in sole *Solea solea* (L.). Neth J Sea Res 32:255–270

Rijnsdorp AD, van Beek FA, Flatman S et al (1992) Recruitment of sole stock, *Solea solea* (L) in the Northeast Atlantic. Neth J Sea Res 29:173–192

Rodhe J (1996) On the dynamics of the large-scale circulation of the Skagerrak. J Sea Res 35:9–21

Rolland JL, Bonhomme F, Lagardere F, Hassan M, Guinand B (2007) Population structure of the common sole (*Solea solea*) in the Northeastern Atlantic and the Mediterranean Sea: revisiting the divide with EPIC markers. Mar Biol 151:327–341

Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. Mol Ecol Notes 4:137–138

Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. Genetics 145:1219–1228

Savina M, Lacroix G, Ruddick K (2010) Modelling the transport of common sole larvae in the southern North Sea: influence of hydrodynamics and larval vertical movements. J Mar Syst 81:86–98

Selkoe KA, Watson JR, White C et al (2010) Taking the chaos out of genetic patchiness: seascape genetics reveals ecological and oceanographic drivers of genetic patterns in three temperate reef species. Mol Ecol 19:3708–3726

Slatkin M (1985) Gene flow in natural populations. Annu Rev Ecol Syst 16:393–430

Slatkin M (1987) Gene flow and the geographic structure of natural populations. Science 236:787–792

Stenseth NC, Jorde PE, Chan KS et al (2006) Ecological and genetic impact of Atlantic cod larval drift in the Skagerrak. Proc R Soc B Biol Sci 273:1085–1092

Symonds DJ, Rogers SI (1995) The influence of spawning and nursery grounds on the distribution of sole *Solea solea* (L) in the Irish Sea, Bristol Channel and adjacent areas. J Exp Mar Biol Ecol 190:243–261

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial-DNA in humans and chimpanzees. Mol Biol Evol 10:512–526

van Oosterhout C, Weetman D, Hutchinson WF (2006) Estimation and adjustment of microsatellite null alleles in nonequilibrium populations. Mol Ecol Notes 6:255–256

Waples RS (1998) Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. J Hered 89:438–450

Was A, Gosling E, Hoarau G (2010) Microsatellite analysis of plaice (*Pleuronectes platessa* L.) in the NE Atlantic: weak genetic structuring in a milieu of high gene flow. Mar Biol 157:447–462

Watts PC, Nash RDM, Kemp SJ (2004) Genetic structure of juvenile plaice *Pleuronectes platessa* on nursery grounds within the Irish Sea. J Sea Res 51:191–197

Wegner G, Damm U, Purps M (2003) Physical influences on the stock dynamics of plaice and sole in the North Sea. Scientia Marina 67:219–234

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

Whitlock MC, McCauley DE (1999) Indirect measures of gene flow and migration: F_{ST} not equal 1/(4Nm + 1). Heredity 82:117–125

Witthames PR, Walker MG, Dinis MT, Whiting CL (1995) The geographical variation in the potential annual fecundity of Dover sole *Solea solea* (L) from European shelf waters during 1991. Neth J Sea Res 34:45–58