ICES Journal of Marine Science

ICES Journal of Marine Science (2018), 75(2), 675–689. doi:10.1093/icesjms/fsx160

Original Article

Seascape genetics of a flatfish reveals local selection under high levels of gene flow

Eveline Diopere[1,†], Sara G. Vandamme[1,2,‡], Pascal I. Hablützel[1], Alessia Cariani[3], Jeroen Van Houdt[4], Adriaan Rijnsdorp[5,6], Fausto Tinti[3], FishPopTrace Consortium, Filip A. M. Volckaert[1,7]*, and Gregory E. Maes[1,4,8]

[1]*Laboratory of Biodiversity and Evolutionary Genomics, University of Leuven, B-3000 Leuven, Belgium*
[2]*Institute for Agricultural and Fisheries Research (ILVO), B-8400 Oostende, Belgium*
[3]*Department of Biological, Geological and Environmental Sciences, University of Bologna, I-40126 Bologna, Italy*
[4]*Center for Human Genetics, Genomics Core, O&N 1, University of Leuven/UZ Leuven, B-3000 Leuven, Belgium*
[5]*Wageningen Marine Research, NL-1970 AB IJmuiden, The Netherlands*
[6]*Aquaculture and Fisheries Group, Wageningen University, NL-6700 AH Wageningen, The Netherlands*
[7]*Department of Marine Sciences, University of Gothenburg, CeMEB, SE-405 30 Göteborg, Sweden*
[8]*Centre for Sustainable Tropical Fisheries and Aquaculture, Comparative Genomics Centre, College of Marine and Environmental Sciences, James Cook University, Townsville 4811 QLD, Australia*

*Corresponding author: tel: +32 16 32 39 72; fax: +32 16 32 45 75; e-mail: filip.volckaert@kuleuven.be
[†]*Present address: Flemish Government, Division of Area Development, Department of the Environment, Environmental Planning and Projects, Koning Albert I-laan 1-2 box 72, B-8200 Bruges, Belgium.*
[‡]*Present address: North Western Waters Advisory Council, c/o Bord Iascaigh Mhara, Crofton Road, Dun Laoghaire Co., Dublin, Ireland.*

Diopere, E., Vandamme, S. G., Hablützel, P. I., Cariani, A., Van Houdt, J., Rijnsdorp, A., Tinti, F., FishPopTrace Consortium, Volckaert, F. A. M., and Maes, G. E. 2017. Seascape genetics of a flatfish reveals local selection under high levels of gene flow. – ICES Journal of Marine Science, 75: 675–689.

Received 28 December 2016; revised 1 July 2017; accepted 10 July 2017; advance access publication 25 August 2017.

Local adaptation is often found to be in a delicate balance with gene flow in marine species with high dispersal potential. Genotyping with mapped transcriptome-derived markers and advanced seascape statistical analyses are proven tools to uncover the genomic basis of biologically relevant traits under environmental selection. Using a panel of 426 gene-linked single nucleotide polymorphisms (SNPs), we scanned 17 samples ($n = 539$) of sole (*Solea solea* L.) from the Northeast Atlantic Ocean and applied a node-based seascape analysis. Neutral loci confirmed a clear distinction between the North Sea–Baltic Sea transition zone and the other Eastern Atlantic samples. At a more subtle level, the latter unit split in an English Channel and North Sea group, and a Bay of Biscay and Atlantic Iberian coast group. A fourth group, the Irish and Celtic Sea, was identified with 19 outlier loci. A pattern of isolation by distance (IBD) characterized the latitudinal distribution. Seascape analyses identified winter seawater temperature, food availability and coastal currents to explain a significant component of geographically distributed genetic variation, suggesting that these factors act as drivers of local adaptation. The evidence for local adaptation is in line with the current understanding on the impact of two key ecological factors, the life-history trait winter mortality and the behaviour of inshore/offshore spawning. We conclude that the subtle differentiation between two metapopulations (North Sea and Bay of Biscay) mirrors local adaptation. At least three genomic regions with strong population differentiation point to locally divergent selection. Further functional characterization of these genomic regions should help with formulating adaptive management policies.

Keywords: fish, isolation by distance, local adaptation, Northeast Atlantic Ocean, outlier locus, population genomics, SNP, sole.

Introduction

Understanding the mechanisms behind oceanic environmental variation and the emergence and maintenance of locally adapted populations is of great relevance for guiding adaptive management practices in exploited species. The strength of local adaptation is dependent on the interaction between connectivity, population size, and environmentally and human induced pressures. The initial paradigm that marine populations lacked the potential for local adaptation due to large gene flow, became gradually substituted by the observation of many potential causes for genetic discontinuities (Palumbi, 2003; Selkoe *et al.*, 2016). In addition, the large effective population size of many marine populations allows for a rich source of standing genetic variation for selection to act upon. Local adaptation is only likely to evolve faster than neutral differentiation when divergent selection is stronger than both the random effects of drift and the homogenizing effects of gene flow (Hauser and Carvalho, 2008). Hence, the presence of local marine populations has been increasingly attributed to locally adapted groups of individuals (Bernatchez, 2016).

Genomic architecture and linkage disequilibrium play an important role in balancing selection and adaptation with genetic exchange (Stern and Orgogozo, 2008). Islands of selection across the genome may be observed when gene flow is high (Guo *et al.*, 2016). After all, loci subject to strong selection will tend to diverge independently from other genes, and highly divergent genomic regions are expected under reduced gene flow due to genome hitchhiking (Nosil and Feder, 2013). In the case of marine organisms we have learned that (i) marine taxa may be highly structured despite the well-connected environment (Lamichhaney *et al.*, 2012); (ii) standing genetic variation plays an important role in adaptation (Bernatchez, 2016); (iii) ecotypes may evolve at discrete genomic regions (Pujolar *et al.*, 2014); (iv) but more commonly the genetic response might be distributed over a large number of loci (Bernatchez, 2016); and (v) local environmental factors play an important role in gene-environment responses (Berg *et al.*, 2015).

Clues on the adaptive role of particular genomic regions or outlier loci in non-model organisms can be derived by combining information on (i) the putative contribution of environmental variation to genetic variation (Fraser, 2013); (ii) genomic architecture (e.g. genomic islands) (Le Moan *et al.*, 2016); (iii) gene function (Feder and Mitchell-Olds, 2003); and (iv) experimental evolution. More specifically regions under influence of divergent selection are likely to show increased differentiation relative to regions under neutral evolution (Limborg *et al.*, 2012). Seascape genomics is an approach to analyse genome-wide variation in relation to habitat features such as temperature, currents and community parameters (Vandamme *et al.*, 2014; Selkoe *et al.*, 2016). Hence, it reveals the genomic basis of biologically relevant traits and their association with the environment. This enables the analysis of the proportional contribution of neutral differentiation and directional selection to the micro-evolution of exploited marine populations at much higher spatio-temporal resolution (Berg *et al.*, 2015).

Flatfishes are a group of benthic fishes with variable dependency on coastal regions to complete their life-cycle, hence exhibiting various levels of connectivity and environmental selection (Hemmer-Hansen *et al.*, 2007; Vandamme *et al.*, 2014). Their life-cycle is partitioned into a spawning, nursery and juvenile stage, which allows for various levels of spatial differentiation and adaptation. Hence, we expect flatfish populations to micro-evolve under dispersal limitation through models of isolation by distance (IBD) and local adaptation (Cuveliers *et al.*, 2012; Orsini *et al.*, 2013).

The flatfish sole (*Solea solea* L.; Soleidae; Pleuronectiformes) plays a crucial role in the benthic food web of shelf regions of the Northeast Atlantic Ocean; the highest sole biomass is reached in the North Sea and English Channel. Selective pressure is high due to historically high fishing mortality (Lescrauwaet *et al.*, 2013). Although most sole stocks have responded positively to adaptive management strategies over the past few years, several either remain overexploited or are recovering (ICES, 2016).

Molecular studies aiming at identifying stock structure in sole along the north eastern Atlantic coasts were implemented under the implicit assumption of neutral evolution (Kotoulas *et al.*, 1995; Guinand *et al.*, 2008; Cuveliers *et al.*, 2012). Using mitochondrial (cytochrome *b*) and nuclear (microsatellite) variation Cuveliers *et al.* (2012) identified three genetic subpopulations, namely the North Sea–Baltic Sea transition zone, the North Sea and the Bay of Biscay. The authors also suggested a fourth putative population in the Celtic and Irish Sea, although weakly supported with a low number of neutral markers. Considering the many spawning grounds and high potential for dispersal (Burt and Milner, 2008; Lacroix *et al.*, 2013), the question remained whether biologically meaningful differences existed at a finer scale if functional variation was evaluated. Such approach has increasingly often been shown to be efficient and feasible in marine fish populations (Bernatchez, 2016).

Local selection pressure may be substantial for two reasons. Environmental heterogeneity may induce selection and result in local adaptation (Yeaman and Whitlock, 2011) and environmental heterogeneity across the wide distribution range of sole is subtle but biologically meaningful. There is a large difference in ecology between northern populations living in colder environments and southern populations living in warmer environments. Temperature affects natural mortality, the onset of spawning, egg hatching, the duration of the pelagic phase, recruitment variability, metabolism, and hypoxic tolerance (Fonds, 1979; Teal *et al.*, 2008; Fincham *et al.*, 2013; Mollet *et al.*, 2013). Hence sole stocks differ phenotypically at life-history traits (LHT) such as relative fecundity, egg size, growth, and maturation rate (Rijnsdorp and Vingerhoed, 1994; Mollet *et al.* 2013). Northern populations acquire energy and invest in reproduction at higher rates than southern populations. The former also had an intrinsic tendency to mature earlier (Mollet *et al.*, 2013). Higher mortality in northern populations during cold winters might be one of the key drivers of the geographical variation in growth and maturation of sole. Mollet *et al.* (2013) attribute the population differences in phenotype to genetic factors.

Another ecological factor impacting flatfish populations is the positive relation between recruitment level and nursery area [nursery size hypothesis (Rijnsdorp *et al.*, 1992)]. Nursery grounds in the northern range are close to the spawning grounds, reducing the impact of currents and increasing larval survival relative to initial recruitment (Rijnsdorp *et al.*, 1992; Lacroix *et al.*, 2013; Heessen *et al.*, 2016). Populations in the southern range (Bay of Biscay) spawn offshore and larvae are advected inshore to the estuarine nurseries, where initial phenotypic variability tends to decrease (Amara *et al.*, 2000).

The above mentioned differences in life-history traits and behaviour raise the questions (i) whether there is support for local adaptation of sole across the continental shelves of the Northeast Atlantic Ocean under a background of IBD, (ii) which environmental factors play a role in adaptation, and (iii) which regions of the sole genome are involved.

To answer these questions, we aimed at identifying putative footprints of selection under an IBD model by applying a genome scan on sole. We sampled the spawning and feeding grounds throughout the full range in the Northeast Atlantic Ocean and genotyped all individuals with transcriptome-derived Single Nucleotide Polymorphism (SNP) markers using a genome scan. Unlike Nielsen *et al.* (2012) who attempted with the same data set to assign individual sole to their spawning population based on a few well-chosen population samples from the North Sea and Celtic/Irish Sea, this study provides a population genomic analysis in an environmental context (seascape genomics) of all populations of the Northeast Atlantic Ocean. In order to partition geographical, environmental and genetic variation, a redundancy analysis, and Similarity Network Fusion analysis were implemented. We interpret our results in view of the biological traits of sole and integrate them in the growing knowledge of functional adaptive divergence in marine species with high gene flow.

Methods

Samples

In total, 650 individual sole were collected from 17 georeferenced sampling sites in the Baltic-North Sea transition zone (TRANS), North Sea, Irish Sea, Celtic Sea, English Channel, Bay of Biscay, and Portuguese waters between 2003 and 2009 (Figure 1 and Table 1). The 17 samples were collected by contracted commercial vessels or during research surveys and used in Nielsen *et al.* (2012) for other purposes. Whole fish were immediately frozen or a fin tissue sample was collected onboard and preserved in 96% pure ethanol.

DNA extraction, genetic markers, and data quality

Extracted DNA of TRANS, North Sea, Irish Sea, Celtic Sea, and English Channel was available from Cuveliers *et al.* (2012). DNA of the other samples was extracted according to Cuveliers *et al.* (2012). Individual fish were genotyped at 426 validated muscle transcriptome-derived SNPs (see Supplementary Table S1 for accession numbers at the dbSNP database at www.ncbi.nlm.nih.gov/SNP) with the GoldenGate^{TM} (Illumina) high-throughput genotyping assay. SNP discovery and genotyping procedures are detailed in Nielsen *et al.* (2012), whose SNP panel was retained in this study. The muscle transcriptome was targeted for its relevance to growth, a major target of adaptation under latitudinal environmental variation. Genotypes were visually checked with the GenomeStudio^{TM} (Illumina) genotyping module and if necessary the clustering of homozygotes and heterozygotes was edited. Based on test analyses with various thresholds for missing values (for details see Supplementary File), all SNPs were kept but individuals with more than 10% missing values were removed, resulting in a dataset of 539 individuals. The average sample size was 30 individuals per sampling location with a minimum of 16 individuals (Table 1). Overall, we selected the same 1,536 putative SNPs as in Nielsen *et al.* (2012) from the transcriptome and 426 SNPs were used for further original analyses following validation (see Nielsen *et al.*, 2012).

Genetic diversity, Hardy–Weinberg equilibrium and linkage disequilibrium

Estimates of genetic diversity, observed, and expected heterozygosity were calculated for each sample both multi-locus and single-locus in GENETIX v4.5 (Belkhir *et al.*, 2004). Hardy–Weinberg Equilibrium (HWE) was evaluated for each SNP and sampling location, using exact tests based on heterozygote deficit or excess as implemented in the software GENEPOP v4.2 (Raymond and Rousset, 1995). To correct for multiple testing the Benjamini–Hochberg procedure, controlling for the false discovery rate (FDR; $\delta = 0.05$), was used (Benjamini and Hochberg, 1995).

Linkage among loci was inferred from (i) the presence of multiple SNPs on the same contig or linkage group (Diopere *et al.*, 2014), and (ii) statistical tests for linkage disequilibrium. Statistical significance was assessed for each SNP pair at each sampling site with GENEPOP v4.2. Significance thresholds were adjusted using a FDR approach ($\delta = 0.05$). To assess the effect of linkage on inferences of population structure, matrices of pairwise F_{ST}-values were calculated using Arlequin v3.5 (Excoffier and Lischer, 2010) based on all SNPs and only unlinked SNPs (i.e. not on the same contig or linkage group), respectively. The two F_{ST}-matrices were compared statistically with a Mantel test and visually with a Procrustes superimposition of multidimensional scaling (MDS) plots using the *vegan* package (Oksanen, 2011) in R (R Development Core Team, 2011).

Outlier analyses

There are multiple challenges to detect outlier loci that are putatively under selection: false positives, genotyping errors, population structure, variation in mutation rate and sensitivity (Narum and Hess, 2011). Since the importance of these issues is data-dependent, several methods have been developed for outlier detection. In a simulation study by (Narum and Hess, 2011), the coalescence-based FDIST2 algorithm and the Bayesian regression approach of BAYESCAN performed better than the coalescence method implemented in Arlequin (Foll and Gaggiotti, 2008; Excoffier and Lischer, 2010). We therefore limited our analysis to these two methods to identify loci not matching with neutral expectations. First, the FDIST2 method was used as implemented in the software package LOSITAN (Antao *et al.*, 2008). Here, coalescence simulations were used to simultaneously compare F_{ST} and heterozygosity of individual loci assuming neutrality (mutation-drift equilibrium in a symmetric island migration model). While being robust to non-equilibrium conditions, this method has been reported to be sensitive to demographic fluctuations (population size changes) among populations and hierarchical genetic structure, which may result in the detection of false outliers (Berg *et al.*, 2015). Outliers were considered significant after correcting for multiple testing using a FDR approach at $\delta = 0.05$. We performed 10^6 simulations on all populations to calculate a confidence interval of 95%.

Second, a Bayesian regression approach implemented in the software BAYESCAN was used as in Nielsen *et al.* (2012). It is based on locus-population F_{ST} coefficients, which are decomposed into a locus-specific component (α), shared by all populations, and a population-specific component (β), shared by all loci, using a logistic regression. When "alpha" is necessary to explain the observed pattern of diversity, departure from neutrality is assumed. All parameters that can be set in the software

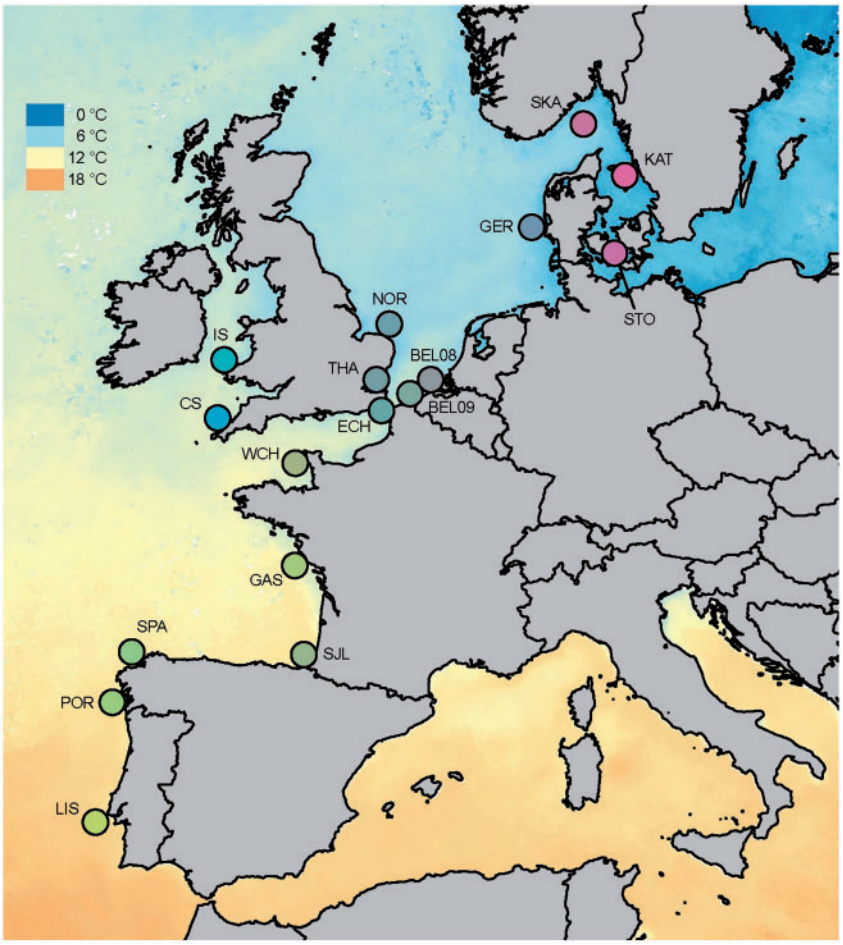

Figure 1. Location of the 17 sampling sites and mean sea surface temperature in January 2012 in the Northeast Atlantic. Colours of sampling sites are scaled according to cluster membership probabilities in the STRUCTURE analyses for the outlier dataset. Irish Sea samples have been appointed to cyan, TRANS samples to magenta and SOUTH samples to yellow. All other locations are represented as a mix of these three base colours in accordance with the individual Q-values in STRUCTURE. For site codes see Table 1.

Table 1. List of *Solea solea* samples with sampling and genetic diversity information.

Sample	Sample group	Area*	ICES area	Latitude	Longitude	Year-month	N	N_{mv}	H_e	H_o
STO	TRANS	Transition Baltic Sea (1)	III.a	55.65	10.76	2007-10	40	27	0.34	0.35
KAT	TRANS	Transition Baltic Sea (1)	III.a	57.16	11.65	2007-11	40	35	0.34	0.35
SKA	TRANS	Transition Baltic Sea (1)	III.a	57.78	9.99	2007-11	40	29	0.34	0.34
GER	NS	German Bight (2)	IV.b	54.52	7.89	2008-11	40	29	0.35	0.35
NOR	NS	Southern North Sea (2)	IV.c	52.92	2.24	2008-08	28	27	0.35	0.34
BEL8	NS	Southern North Sea (2)	IV.c	51.39	3.17	2008-05	40	35	0.35	0.35
BEL9	NS	Southern North Sea (2)	IV.c	51.22	2.83	2009-05	24	24	0.35	0.34
THA	NS	Southern North Sea (2)	IV.c	51.47	1.33	2007-08	40	40	0.35	0.35
ECH	CHANNEL	Eastern English Channel (3)	VII.d	50.78	1.48	2008-07	40	38	0.35	0.35
WCH	CHANNEL	Western English Channel (3)	VII.e	49.66	−2.13	2009-09	40	40	0.35	0.35
GAS	SOUTH	Bay of Biscay (4)	VIII.a	45.92	−1.69	2009-02	40	39	0.35	0.34
SJL	SOUTH	Bay of Biscay (4)	VIII.b	43.55	−1.57	2003-05	39	25	0.35	0.34
SPA	SOUTH	Bay of Biscay (4)	VIII.c	43.6	−8.86	2009-04	40	40	0.35	0.34
POR	SOUTH	Portugese waters (5)	IX.a	42.14	−9.28	2009-03	40	35	0.34	0.33
LIS	SOUTH	Portugese waters (5)	IX.a	38.35	−9.35	2003-05	39	16	0.34	0.33
IS	ICS	Irish Sea (6)	VII.a	52.21	−5.33	2008-03	40	33	0.34	0.35
CS	ICS	Celtic Sea (6)	VII.g/f	50.81	−5.01	2008-04	40	27	0.33	0.36

Sample acronym, sampling area, ICES area, latitude/longitude (decimal notation), sampling time, sample size before (N) and after (N_{mv}) removing individuals with >10% missing values, expected (H_e) and observed (H_o) heterozygosity are reported.
*The prior location name used in STRUCTURE is given in brackets.

were left as default, for example, the "prior odds for neutral model = 10." The FDR was set at 5%, meaning that markers with a *q*-value lower than 0.05 were considered as outlier.

Information on the genomic position of the outliers was sourced from the linkage map of sole (Diopere *et al.*, 2014). The functional significance of outlier loci was retrieved after a Basic Local Alignment Search as described in Milano *et al.* (2011).

Genetic population structure

Genetic differentiation between sampling sites was investigated using pairwise F_{ST} values calculated with the software Arlequin v3.5 at a significance level of $\alpha = 0.05$ (Excoffier and Lischer, 2010) for both neutral and outlier SNPs. Significance thresholds were adjusted using a FDR approach ($\delta = 0.05$).The resulting pairwise F_{ST} matrix were subjected to a Principal Coordinates Analysis (PCoA). The Cailliez-correction method was applied to avoid negative eigenvalues. The first and second principal coordinates were plotted to visualise genetic population structure. Furthermore, we estimated and confirmed the number of genetically distinct populations using a Bayesian individual-based clustering method as implemented in the software STRUCTURE v2.3.3 (Pritchard *et al.*, 2000). Following a burn-in of $2*10^4$, 10^5 MCMC iterations were run for a number of *K* (ranging from 1 to 10) clusters. Ten replicates were run for each *K* for the outlier loci. Admixture was used in the ancestry model, given the high degree of gene flow in sole, as well as information on the sampling location to assist clustering (see Table 1). The latter option allowed for a better performance in datasets with weak structure (Hubisz *et al.*, 2009). Allele frequencies were considered correlated among populations. To determine the most likely number of clusters of individuals based on the multi-locus genotypes, the *K* value with the largest log-likelihood using a Bayesian approach was selected (Evanno *et al.*, 2005).

To test for isolation by distance (IBD) patterns, we compared pairwise F_{ST} values with shortest waterway distances using a Mantel-test as implemented in the *mantel*-function in the R-package *vegan* (Oksanen, 2011) in R. Statistical significance was assessed with 1000 permutations of the data. We repeated the analysis after exclusion of the TRANS samples.

Environmental associations

To identify the factors that are potentially involved in the differentiation among populations, we performed a node-based seascape analysis, linking spatial and environmental variables with genetic variation (Selkoe *et al.*, 2008). This was achieved through (i) a distance-based redundancy analysis (RDA), a constrained ordination method that explores the relationship between response and explanatory matrices and (ii) a latent factor mixed model (LFMM) analysis (Frichot *et al.*, 2013). Partial RDA analyses were conducted to partition the relative contribution of space (S) and environment (E) on genetic differentiation among sampling sites. To do so, a pairwise F_{ST}-matrix was subjected to a Principal Coordinates Analysis (PCoA). The axes of this PCoA constitute the dependent variables of the RDA. Spatial explanatory variables were represented by the four most important Moran's eigenvector maps (MEMs) (Dray *et al.*, 2006) with positive eigenvalues. MEMs are eigenvectors of a Principal Coordinates Analysis, which was performed on a matrix with the geographically shortest waterway distance between sampling sites. Since dispersal is more likely among close neighbours, the original distance matrix was first truncated to retain only short distances. Analyses were performed using the *PCNM* and *vegan* packages (Oksanen, 2011) in R (R Development Team, 2011).

Environmental data was downloaded from the Environmental Marine Information System (EMIS) of the Joint Research Centre (JRC) of the European Commission. Among all parameters available, seven were selected for their known influence on the distribution of sole or other marine species, and their variance across sampling locations (Vandamme *et al.*, 2014). They include bottom (SBT) and surface (SST) temperature (°C), bottom salinity (SBS, psu), mixed layer depth (MLD, m), maximum density gradient (MDG, kg m^{-4}), chlorophyll a concentration (CHL, mg m^{-3}), surface productive layer (SPL, m) and latitude as a proxy for day length, and other seasonal features (see Supplementary File). Since there is a pronounced seasonal variation of temperature in the study area, seasonal averages were used instead of yearly averages. To account for variability between years, the data represented averages over 10 years from 1998 to 2007 for MLD, MDG, SBS, and SBT and from 2002 to 2011 for SST, CHL, and SPL. More details on the environmental variables are provided in Supplementary File. To reduce the dimensionality of the environmental data, a PCA (Principal Component Analysis) was performed. Upon visual inspection of the eigenvalues, the four most important PCs (accounting for 95.2% of the explained variation) were used for RDA and variation partitioning. Analyses were performed using the *vegan* package (Oksanen, 2011) in R (R Development Core Team, 2011). As the sampling period spans a 7-year period from 2003 to 2009, we repeated the variation partitioning including the year effect encoded as dummy-variables.

Associations between environmental variables and particular loci were investigated using LFMM (Frichot *et al.*, 2013). This method uses latent factors to correct for population structure while fitting a linear regression between allele frequencies and environmental variables. First the analysis was conducted including all 17 sites and all loci (thus including both "neutral" and outlier loci). Like for the RDA, the dimensionality of the environmental data was reduced using PCA. The number of genetic units was set at two, reflecting the genetic differentiation between the Baltic Sea and the remaining sites. The analysis was conducted using the R package *LEA* (Frichot and Francois, 2015). This method implements a MCMC algorithm which was run $5*10^3$ times after a burn-in of the same length. To avoid pseudo-replication (the number of genotypes greatly exceeds the number of independent observations of the environmental variables), one individual was randomly drawn from each location for the analysis. The process was repeated 100 times. A FDR approach was used to correct for multiple testing. Since the LFMM did not identify any candidate locus, we repeated the analysis using only the non-Baltic sites. For this analysis, the absence of major genetic discontinuities was assumed.

Similarity network fusion

Geographic clusters of sites with excess similarity of outlier loci compared to neutral loci might be taken as an indication of local adaptation, especially when this excess can be linked to local environmental conditions. Here, we used Similarity Network Fusion (SNF; Wang *et al.*, 2014) to identify such clusters. First, distance matrices were calculated for environmental data, neutral loci and outlier loci. For standardization of the environmental data, each variable was transformed to zero mean and unit variance before

calculating Euclidean distances among sites. For neutral and outlier loci, F_{ST} values were used. These distance matrices were converted into affinity matrices using the *affinityMatrix* function in the R-package *SNFtool* (Wang *et al.*, 2015). Subsequently, a non-linear message-passing method (Wang *et al.*, 2014) implemented in the SNF function (Wang *et al.*, 2015) was used to iteratively update a pair of networks (environment + neutral SNPs or environment + outlier SNPs, respectively) to resemble to one another more until they converged to a single fused network. This resulted in two fused similarity networks: one for neutral SNPs and one for outlier SNPs. To distinguish between patterns of neutral gene flow and local adaptation, the weights of the edges of these two similarity networks were compared. We are currently not aware of a statistical significance test for this comparison. Edges supported by weights 10% higher than in the other fused network were considered supported only by the focal fused network (Wang *et al.*, 2015). Edges supported only by the environment + outlier SNPs network can be assumed to indicate similar (i.e. more similar than expected under neutral gene flow) outlier allele frequencies due to similar environmental selection.

Results

Genetic diversity, Hardy–Weinberg equilibrium and linkage disequilibrium

The observed and expected multi-locus heterozygosity per sample for the 426 validated SNPs ranged between 0.33 and 0.35 and 0.33 and 0.36, respectively (Table 1). Single-locus heterozygosity values are listed in Supplementary Table S1. A total of 6,947 tests were performed to screen for deviations from HWE. After correction for multiple testing only two tests remained significant, both for the same locus (*SNP558*) at sites STO and SKA.

Among the 1,538,925 linkage disequilibrium tests that were performed to evaluate statistical linkage between loci, in total 61,692 (4%) had a *p*-value lower than 0.05. After correction for multiple testing this number decreased to 282. This panel included 61 different SNP pairs that showed linkage of which 13 had a significant test result in more than half of the sampling sites. This was not unexpected as the 426 SNPs originated from 334 different contigs; 273 were single-marker contigs and 61 contigs had multiple SNPs. Out of the 282 significant linkage disequilibrium tests, 159 (56%) were between SNPs from these multi-marker contigs and therefore physically closely linked. High congruence was observed between F_{ST}-matrices based on the dataset including all SNPs and the dataset excluding SNPs originating from multi-marker contigs (Mantel test, $r = 0.990$, $p < 0.0001$; MDS plots and results of Procrustes superimposition (Supplementary File).

Outlier analysis

We detected 19 outliers, of which 13 were suggested by both outlier methods (Table 2). Twice, two outliers were located on the same contig. Seventeen of the 19 outliers were positioned on seven different linkage groups of the sole genome map. Twelve outliers clustered in three linkage groups: LG2 (6), LG4 (3), and LG6 (3). Fifteen outliers were annotated, of which cytokeratin 13 was assigned to two linkage groups (LG2 and LG4, with LG2 including three linked loci). Two loci (LG6) were annotated to troponin C skeletal muscle (Table 2). We proceeded with all 19 outliers to maximise information.

Genetic population structure

Pairwise F_{ST} values of the neutral loci ranged from 0 to 0.0126; the overall F_{ST} value amounted to 0.007. Of the 136 site-pairs, 49 were statistically differentiated ($p < 0.05$), of which 39 involving samples from the transition zone (Supplementary File and Table S5). The TRANS samples also featured as a separate group on the first genetic principal coordinate of a PCoA. A geographical gradient suggesting isolation by distance) occurred between NS (group 2) and SOUTH (group 3) (Figure 2). Bayesian cluster analysis based on the dataset of neutral (non-outlier) markers identified two groups: (i) the TRANS samples and (ii) the remaining samples, namely the North Sea (NS), Easter English Channel (ECH), Irish and Celtic Sea (ICS), and Western English Channel + Bay of Biscay and Portuguese waters (WCH + SOUTH) (Figure 3A). The most likely number of clusters, selected by choosing K with the largest log-likelihood, was either $K = 3$ or 6, with for the $K = 3$ case the TRANS samples being unambiguously assigned to a single cluster and SOUTH samples separating weakly (Figures 1 and 3A). Here the NS/CHANNEL/ICS samples took an intermediate position, with evidence for a cline between the northern sample from German Bight (GER) and the CHANNEL samples.

On the other hand, the dataset of 19 outlier loci showed evidence of a more subtle genetic structure. Pairwise F_{ST} values ranged from −0.008 to 0.247 with 99 significant F_{ST} values (Supplementary File and Table S6), most of them between samples from different groups (i.e. TRANS vs. NS, TRANS vs. ICS, TRANS vs. SOUTH, NS vs. ICS, NS vs. SOUTH, and ICS vs. SOUTH). The overall F_{ST} value amounted to 0.046. Assignment analysis with STRUCTURE based on the outlier loci showed a most probable number of clusters at $K = 2$ with the TRANS samples representing a distinct cluster (Figures 1 and 3; Supplementary File). Upon closer inspection, within the NS group, the southernmost samples (i.e. ECH, THA, BEL8, and BEL9) were somewhat more associated with the SOUTH group than the northern North Sea samples (GER and NOR), suggesting a population break in the North Sea (Figure 3B). In the PCoA plot, the TRANS and ICS samples positioned as distinct clusters (Figure 2). The NS, CHANNEL, and SOUTH samples represented a north-south cline with the German Bight (GER) and Portuguese sample (LIS) at each end. The Eastern English Channel sample (ECH) clustered with the NS samples, whereas the Western English Channel sample (WCH) was found south of this gradient. In addition, the clusters were more separated from each other than on the PCoA plot of the neutral dataset, hence four clusters can be proposed: TRANS, NS + ECH, ICS, and SOUTH + WCH.

We found a significant IBD pattern with both the full dataset ($r = 0.746$, $p = 0.001$) and after exclusion of the TRANS samples ($r = 0.790$, $p = 0.001$), with an excess of genetic differentiation in comparison with an IBD model between TRANS and Atlantic samples. A similar excess can be observed among regional clusters outside the TRANS area for outliers but not for neutral SNPs (Figure 4).

Associations between genetic and environmental variation

We used RDA to associate genetic variation with environmental and spatial variability. Environment and space explained 65.0% of the genetic variation at neutral SNPs among sites (Table 3). Unlike the environment, the contribution of space for neutral variation is highly significant. The combined contribution of environment and space was much larger (42.2%) than their unique contributions (8.6% for

Table 2. Outlier SNP loci detected with LOSITAN and BAYESCAN (**) in the global genome scan of sole; *l: outlier detected in LOSITAN. *b: outlier detected in BAYESCAN.

Locus	Outlier	F_{ST}	LG	cM	Contig	Annotation
SNP1199	*l	0.012	LG1	38.7		Phosphodiesterase 4D interacting protein
SNP1354	**	0.132	LG2	89.2	C13543	Cytokeratin 13
SNP1353	**	0.101	LG2	89.2	C13543	Cytokeratin 13
SNP573	**	0.042	LG2	89.2		Cytokeratin 13
SNP1129	**	0.041	LG2	89.2		Histidine-rich calcium binding protein
SNP1478	**	0.027	LG2	89.2		Nucleoside diphosphate kinase b
SNP1432	*b	0.027	LG2	89.2		Unknown
SNP609	**	0.042	LG4	29.5	C78890609	Valosin-containing protein
SNP789	**	0.041	LG4	29.5	C78890609	Valosin-containing protein
SNP933	**	0.046	LG4	29.5		Cytokeratin 13
SNP116	**	0.029	LG5	0		Unknown
SNP678	*l	0.008	LG5	37.3		Unknown
SNP228	**	0.054	LG6	5.9		Troponin C skeletal muscle
SNP200	**	0.041	LG6	5.9		Troponin T skeletal muscle
SNP737	*b	0.027	LG6	5.9		Troponin C skeletal muscle
SNP1347	**	0.078	LG12	0		Unknown
SNP1168	*b	0.029	LG21	21.9		Adenovirus E1B 19 kDa protein-interacting protein 3-like
SNP1466	**	0.068	n/a	n/a		Gelsolin
SNP1200	*b	0.021	n/a	n/a		Synaptic glycoprotein 2

Individual F_{ST} value. LG (linkage group) and cM (centiMorgan) refers to the position on the linkage map of sole (Diopere *et al.*, 2014); loci on the same contig are listed by number. n/a: SNPs not included in the linkage map. Loci were annotated with BLASTN (NCBI).

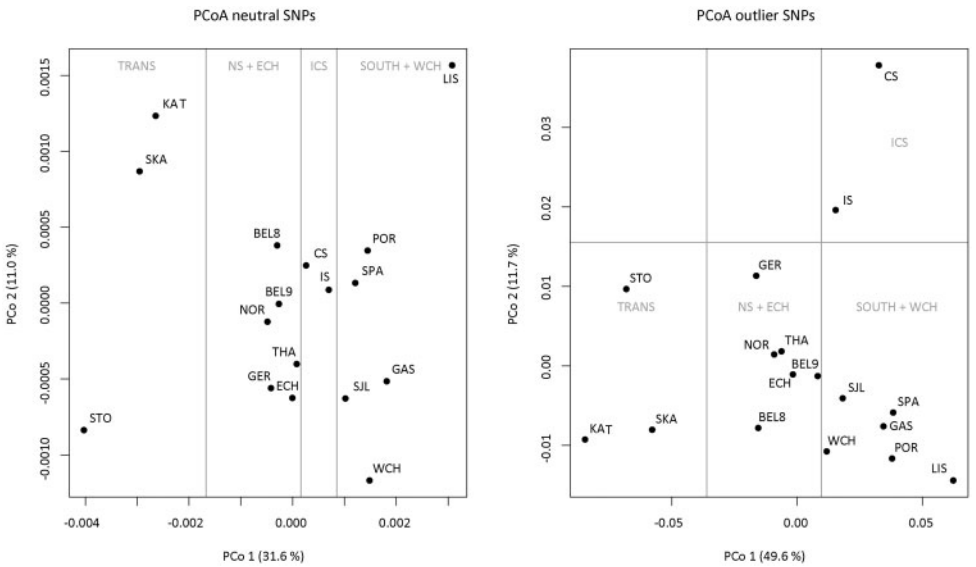

Figure 2. Plots of Principal Coordinates Analysis (PCoA) illustrating similarities among sampling locations in allele frequencies of neutral (left panel) and outlier (right panel) SNP loci.

space (significant) and 14.3% for environment (not significant), indicating that the effect of space and environment is highly intertwined). Genetic variation at outlier SNPs might have an even better fit with space and environment, but test results were not significant (Table 3). Sampling year accounted for 37.9% and 59.0% of the genetic variation at neutral and outlier SNPs, respectively. However, these effects were highly intertwined with space and environment and the pure effect of sampling year accounted for only 0.12% and 0%, respectively.

The LFMM identified two candidates, other than those identified in the outlier analysis, potentially under environmental selection (*SNP62*: p-value $= 1.42^{*}10^{-5}$, z-score $= 1.717$; and *SNP267*: p-value $= 2.88^{*}10^{-5}$, z-score $= -0.251$). A BLAST-search identified a 92% similarity of the *SNP267* sequence with the gene encoding fructose-bisphosphate aldolase A in Japanese rice fish (*Oryzias latipes*). The other locus remains unidentified.

Similarity network fusion

The neutral SNPs did not support clear clusters of sites except for the three southernmost sites along the Iberic coasts. Strong similarity tended to be supported between geographically close sites,

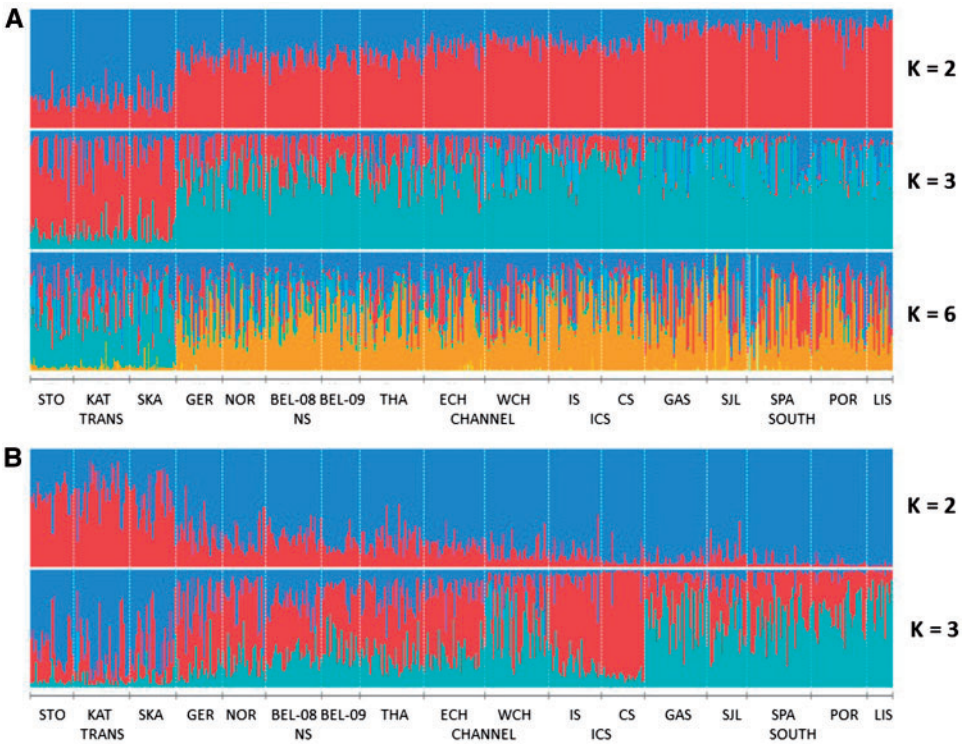

Figure 3. Results from individual based clustering analyses with STRUCTURE for two data sets of *Solea solea* based on (a) neutral ($K = 2$, 3, and 6) and (b) outlier SNP markers ($K = 2$ and 3). Each bar represents an individual with its probability of membership to one the hypothetical clusters. See Table 1 for code assignment and more information on the samples.

whereas frequent weak connections were found between the southern (SOUTH) and northern sites (NS, CHANNEL, and ICS; Figure 5A). In contrast to the neutral SNPs, the outliers supported a tight clustering among southern (SOUTH) and northern sites (NS, CHANNEL and ICS; Figure 5B). The absence of edges between these realms indicated an excess of environmentally mediated differentiation for outlier SNPs compared to neutral ones. The NS and TRANS sites do also form distinct clusters, but outlier loci followed the pattern of neutral divergence (Figure 5C), and the effect of adaptive divergence could therefore not be distinguished from dispersal limitation.

Discussion

Our study provided evidence for subtle, but nevertheless biologically relevant genetic structure in sole of the Northeast Atlantic Ocean, contributing to the growing knowledge on environmentally driven genetic differentiation. We showed that a combination of neutral (dispersal limitation) and adaptive processes (temperature and dispersal behaviour) shape the genetic structure of an important member of the demersal flatfish community.

The subtle population structure of sole

We confirmed and provided more in depth understanding on spatial population structure of sole in the study area than suggested by Cuveliers *et al.* (2012). Our analysis with a set of neutral SNP markers revealed two levels of structure. The North Sea–Baltic Sea transition zone is clearly distinguished from other north eastern Atlantic samples. At a more subtle level the North Sea, English Channel, Irish Sea, and Celtic Sea are distinguished from a southern unit including the Bay of Biscay and Portuguese waters. The combination of space and environment were important factors influencing the genetic pattern, with space explaining most of the genetic variation. Much of it correlated with a gradient from the north eastern North Sea along the English Channel to the Bay of Biscay–Portuguese waters, confirming an isolation-by-distance (IBD) model of population differentiation (Kotoulas *et al.*, 1995, Cuveliers *et al.*, 2012). However, a subset of 19 outlier SNPs sourced from an extensive set of gene-linked markers revealed a fourth population of sole inhabiting the Irish and Celtic Sea, first suggested by Cuveliers *et al.* (2012) unit. Also, the seascape analysis pointed to an important contribution by the environment, which suggests local adaptation among sole populations.

Our study extends previous knowledge on sole population biology and dynamics. First, we identified the geographically isolated sole of the Irish and Celtic Sea as a distinct population. This matches with its discrete spawning sites (Fox *et al.*, 2000) and the genetically discrete populations of dab *Limanda limanda* (Tysklind *et al.*, 2013) and turbot *Scophthalmus maximus* (Vandamme *et al.*, 2014) in the Irish Sea and Bristol Channel. Biologically significant is the tide-driven coastal flow and baroclinic currents in the Irish and Celtic Sea. For example dispersal of cockle larvae (*Cardium edule*) from southern Wales is more likely in a western than eastern direction (Coscia *et al.*, 2013), while selective tidal current transport was modelled to retain juvenile plaice *Pleuronectes platessa* (Fox *et al.*, 2006). The distinctness of the stocks around the British Isles made it feasible with the SNP genotypes used in this study to trace individual sole of the southern North Sea and Irish and Celtic Sea (Nielsen *et al.*, 2012).

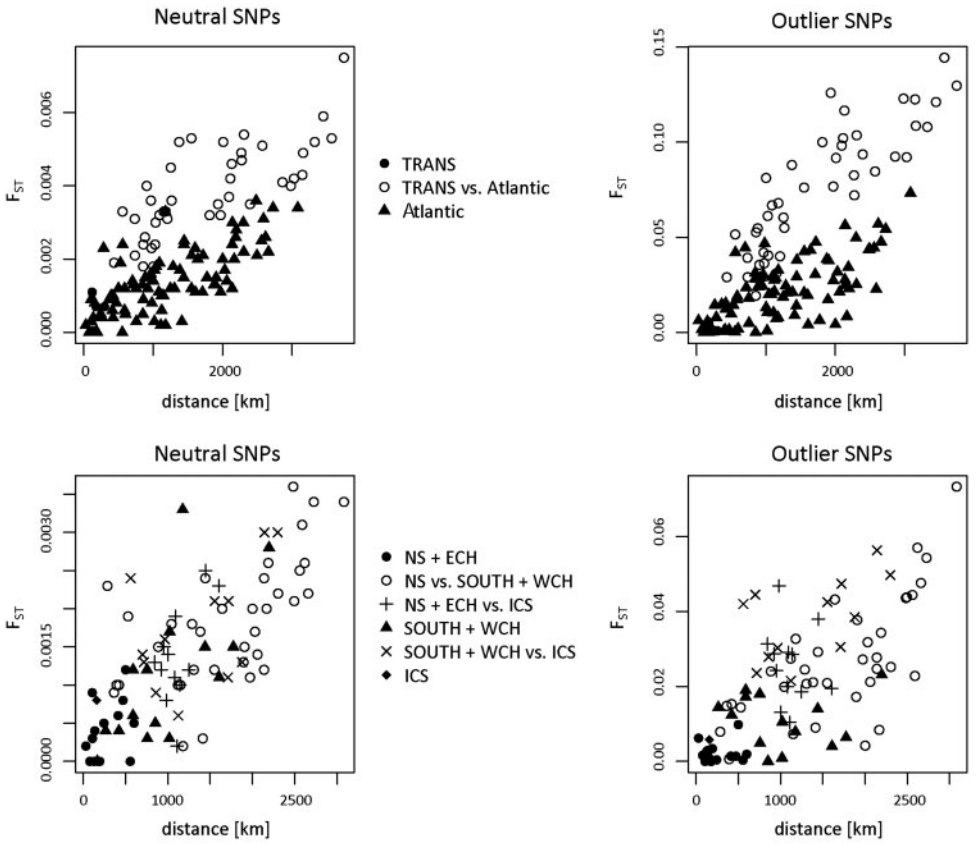

Figure 4. Isolation by distance for neutral and outlier SNP datasets. Filled circles, triangles and rhombuses depict distances within clusters, while other symbols (open circles, triangles, and crosses) indicate distances between clusters. Upper panel: an excess of genetic differentiation in comparison with an IBD model can be identified between the TRANS and the remaining sample sites (Atlantic) as between-cluster differentiation exceeds within-cluster differentiation independent of geographic distance. Lower panel: a similar excess of genetic differentiation can be seen among regional clusters outside the TRANS area for outliers, but not for neutral SNPs.

Table 3. RDA analysis of neutral (407 SNPs) and outlier (19 SNPs) datasets, partitioning the explained variation by space (S), environment (E), both space and environment (S + E), space corrected for environment (S|E), environment corrected for space (E|S), the shared effects between environment and space (S ∩ E), and the residual effects.

	Neutral SNPs					Outlier SNPs				
Effect	**Num DF**	**Den DF**	**adj R^2**	***F*-value**	***P*-value**	**Num DF**	**Den DF**	**adj R^2**	***F*-value**	***P*-value**
Space	4	12	0.508	1.578	**0.005**	4	12	0.741	0.914	0.750
Environment	4	12	0.564	1.246	0.125	4	12	0.757	0.992	0.410
S + E	8	8	0.650	1.522	**0.010**	8	8	0.798	0.918	0.760
S|E	4	12	0.086	1.564	**0.035**	4	12	0.041	0.883	0.580
E|S	4	12	0.143	1.306	0.190	4	12	0.058	0.941	0.600
S ∩ E			0.422					0.700		
Residuals			0.350					0.202		

The columns point to the degrees of freedom in the numerator (Num DF), the degrees of freedom in the denominator (Den DF), the adjusted correlation coefficient (adj R^2), the *F*-value and the *P*-values (when significant they are indicated in bold).

Second, the genetic profile of the Western English Channel was transitional with evidence for admixture between the Gulf of Biscay and Eastern English Channel. The Ushant tidal front separates the tidally mixed Channel waters from the stratified Celtic Sea. It is also the boundary between the Boreal and Lusitanian biogeographical provinces (Ayata *et al.*, 2010). Here, coastal waters are highly dynamic; cyclonic eddies, tidally induced mixing, and a north eastern residual current influence connectivity (Savina and Ménesguen, 2008). But history might also have influenced the isolation and subsequent discreteness of sole populations. Similar to other taxa (Maggs *et al.*, 2008), sole recolonised the northern range after the Last Glacial Maximum (LGM) following the retreating ice cap and the rising sea level, hence inducing a pattern of isolation by distance with subsequent local adaptation.

Third, sole inhabiting the north eastern North Sea (German Bight) segregated spatially from those of the southern North Sea.

Figure 5. Fused similarity networks of environmental data and neutral and outlier SNPs. The three panels show edges supported by (a) a fused similarity network of environment + neutral SNPs, (b) a fused similarity network of environment + outlier SNPs, or (c) both fused similarity networks, respectively. Thick and intensely coloured lines indicate high similarity among site-pairs. The thickness and intensity of the edges corresponds to similarities in a fused network of environment and both neutral and outlier SNP datasets. Networks were drawn using the *qgraph* package (Epskamp *et al.*, 2012) in R.

It matches with the pattern observed in sole but genotyped with microsatellites, turbot *Scophthalmus maximus* (Vandamme *et al.*, 2014), shore crab *Carcinus maenas* (Moksnes *et al.*, 2014) and Individual Based Modelling (L. Barbut, pers. comm.). These populations respond to the absence or presence of stratification along the Frisian hydrographic front located west of Texel (NL). An alternative explanation is the close association of sole with estuaries and coastal waters (Kostecki *et al.*, 2012). Topography (the shallow Wadden Sea) and freshwater discharge from several large rivers might be conducive to the discreteness of the German Bight population. The distribution of nursery grounds and biophysical modelling of larval dispersal supports the presence of such regional groups (Lacroix *et al.*, 2013).

Fourth, the strongest signal of neutral and putative adaptive genetic structure was observed in the North Sea–Baltic Sea transition zone. The pattern might be attributed to three factors, either single or in combination (Johannesson and André, 2006; Vandamme *et al.*, 2014). First, geographically and hydrodynamically, the Skagerrak-Kattegat population (TRANS) is isolated. It occupies the easternmost edge of the distribution without entering the brackish Baltic Sea. That edge populations typically have a genetic make-up with a lack of rare alleles and reduced genetic variability (Swaegers *et al.*, 2014) is not substantiated here (Supplementary Table S1). Second, local population density is low and census population size small compared to the North Sea ecoregion. Spawning stock biomass (SSB) of sole has dropped by 30% from historically high levels of 3,880 tonnes in 1992 to 2,749 tonnes in 2016 and the stock has been exploited at unsustainable levels of fishing mortality (0.558 in 1993 and 0.113 in 2015) (ICES 2016). During the past 100 years sole catches have varied and were, proportionally to plaice catches, considerably influenced by the Northern Hemisphere Temperature Anomalies (Sparrevohn *et al.*, 2013). Third, the unique environment of the salinity transition between the North Sea and Baltic Sea may have led to local adaptation in sole, similar to many taxa (Johannesson and André 2006; Limborg *et al.*, 2012).

Genetic evidence of isolation by distance

Isolation by distance (IBD), which reflects the limited exchange of individuals between geographically adjacent populations, is supported by extensive evidence (Johannesson and André, 2006; Wright *et al.*, 2015). Evidence of IBD for sole is convincingly based on allozyme markers (Kotoulas *et al.*, 1995), nuclear markers (microsatellites), hundreds of neutral SNP markers (this study), but not cytoplasmic (mitochondrial DNA) markers along the coasts of the Northeast Atlantic Ocean (Cuveliers *et al.*, 2012). Our statistical evidence is supported by empirical dispersal distances of up to 150 km by adults (Burt and Milner, 2008) and up to 30 km by juveniles on the nursery grounds (Le Pape and Cognez, 2016), modelled larval dispersal of maximally 140–300 km depending on the spawning ground (Lacroix *et al.*, 2013) and modelled gene flow of two to five times the mean distance of larval dispersal (Palumbi, 2003). In addition, local coastal features such as fronts at the contact of mixed and stratified waters off Texel and Brittany, and the patchy distribution of feeding grounds, influence dispersal (Rijnsdorp *et al.*, 1992; Burt and Milner, 2008). Moreover, IBD might occur in a metapopulation context, where populations exchange genes at a rate proportional to the geographic distance between them. Evidence for a metapopulation structure is available from temporally variable connectivity (Lacroix *et al.*, 2013), a patchy distribution of habitats, a life-cycle closely associated with the dynamics of benthic communities, and in the Bay of Biscay and Mediterranean Sea with estuarine environments (Le Pape *et al.*, 2003; Darnaude *et al.*, 2004). The annual variation in recruitment (Rijnsdorp *et al.*, 1992; Engelhard *et al.*, 2011; Fincham *et al.*, 2013) points to source-sink dynamics (Kritzer and Sale, 2004). The characteristic migration dynamics of sole linked to a triangular life-cycle with distances between spawning grounds in the order of 100 km, leaves opportunities for the exchange of individuals between population units.

There might be concerns for those cases where population units don't match with stocks, the units considered in management (Reiss *et al.*, 2009). The management units of sole along the

Northeast Atlantic Ocean have a finer resolution than the four (meta)population units (which might be called stocks) we suggest (Heessen *et al.*, 2015; ICES 2016). This evidence combined with the IBD pattern would suggest that management advice should systematically take into account adjacent management units. Some of the metapopulations are too small and isolated to be fished intensively (e.g. Irish Sea and Skagerrak), with no evidence from adjacent stocks supplementing them. However there is no clear evidence of genetic diversity loss judged from earlier indices of genetic diversity and effective population size in sole (Cuveliers *et al.*, 2011; but see Cuveliers *et al.* (2012) for possible explanations, although overall evidence of genetic erosion in overfished species is still growing worldwide (Pinsky and Palumbi, 2014).

Growing evidence of local adaptation

Our network analysis pointed to an excess of neutral gene flow between the English Channel and the Bay of Biscay compared to gene flow at outlier loci when environmental parameters are considered simultaneously. Temperature, typically varying on a latitudinal gradient, and the surface productive layer, a proxy for food availability and the presence of predators, appear to be the main factors discriminating the populations.

Support for local adaptation in sole was provided by several lines of environmental and biological evidence. First, the key abiotic factor temperature varies on a latitudinal gradient, tracing the coasts of the Northeast Atlantic Ocean. It influences in many regards the biology of sole, for example its life-cycle (Teal *et al.*, 2008), winter survival (Rijnsdorp *et al.*, 1992), initiation of spring spawning (Fonds, 1979; Rijnsdorp and Vingerhoed, 1994), timing of peak spawning (Fincham *et al.*, 2013) and regional variability in spawning (Rijnsdorp *et al.*, 1992). Temperature affects the wind regime and hence mixing of the water column, which is crucial for the survival of ichthyoplankton (Lacroix *et al.*, 2013). Also larval development time (Fonds, 1979), juvenile growth rate and habitat quality of the nursery grounds (Teal *et al.*, 2012) are influenced by temperature. Thus, temperature has a direct impact on the fundamental niche and hence geographical distribution. Local evidence might attribute the differences in life-history traits of sole to acclimation, but given the geographical consistency adaptation is also contributing (Fincham *et al.*, 2013; Mollet *et al.*, 2013). Molecular support is provided by the association of allelic variation in the gene encoding fructose-bisphosphate aldolase, part of the Calvin cycle, with a latitudinal gradient. Expression of this gene is seasonally up-regulated in rainbow smelt (*Osmerus mordax*) in winter and believed to play an important role in cold-adaptation (Richards *et al.*, 2010).

Second, coastal ecosystems are differentiated latitudinally and in an inshore-offshore direction, impacting the habitats occupied by the various life-stages of sole (Kostecki *et al.*, 2012). Southern habitats are classified as temperate warm (Lusitanian) biogeographical provinces while more northern habitats as temperate cold (Boreal); transition between both occurs in the western English Channel. Most significant is that sole spawn offshore in the Bay of Biscay and are advected inshore (Amara *et al.*, 2000), while nurseries in the English Channel and North Sea are in proximity of the spawning grounds (Rijnsdorp *et al.*, 1992; Heessen *et al.*, 2015). Plankton community composition, differing in both provinces, determines first feeding success of the larvae (Fonds, 1979). Similarly, sandy and sandy-muddy benthic communities determine survival of the newly settled larvae (Eriksson *et al.*, 2010). Hence, the southern and northern early life history patterns match exactly with the two ecographical groups appearing from our adaptive genetic analysis.

Third, changes in North-South phenology influence larval survival. The surface productive layer correlates with fish larval survival (Pepin, 1993), larval growth rate (Le Pape and Bonhommeau, 2015) and (sub)adult growth rate (Teal *et al.*, 2012). However, the plankton community and its productivity have changed since the 1970s (Beaugrand *et al.*, 2009), affecting on their turn fish productivity. For example the growth rate of sole and plaice has accelerated in the 1960s and 1970s and dropped in the 1980s (Rijnsdorp *et al.*, 2004). Since 1970 the date of recruitment and hence of first feeding has advanced in four of the seven sole stocks in the Northeast Atlantic Ocean (Fincham *et al.*, 2013). Overlap varies among years as phytoplankton dynamics are linked to wind mixing and irradiation (Siegel *et al.*, 2003), while sole spawning is triggered by temperature. Hence climate change may affect the match between prey and fish larva [match-mismatch hypothesis (Cushing, 1969)]. The latter represents the ecological equivalent of the genetic sweepstake hypothesis, where some parents have a higher stake in determining cohort strength either through fertility or survival (Hedgecock, 1994). High mortalities, when there is a poor match with the prey field, are equivalent to enhanced selection pressure and hence survival of the most adapted, with some families contributing more than others. This might contribute to the variable cohort structure of sole (Heessen *et al.*, 2015; ICES 2016) and explain the correlation of Virtual Population Analysis estimates of recruitment at age 1 with modelled connectivity under historical and future climate scenarios (Lacroix *et al.*, 2013; Lacroix *et al.*, in press).

Fourth, functional annotation of genetic markers showing footprints of selection provides opportunities for molecular characterisation and functional understanding (Feder and Mitchell-Olds, 2003). Here, among the 15 annotated outlier loci, all isolated from a muscle transcriptome library, most genes were not unexpectely growth-related. However, growth together with gonadal maturation features among the most selection-responsive metabolic traits, which in poikilotherms is under the environmental control of temperature (see above). The troponin T fast skeletal muscle isoform, a regulatory protein involved in the contraction of skeletal muscle, has been attributed a role in the growth of Atlantic halibut *Hippoglossus hippoglossus* (Campinho *et al.*, 2007), in the spermatogenesis of Senegalese sole *Solea senegalensis* (Forné *et al.*, 2011) and in the metamorphosis of common sole (Focant *et al.*, 2003). The cytoskeletal protein cytokeratin 13 is expressed in internal stratified and mucus-secreting epithelia of fish (Chua and Lim, 2000). The thus far poorly explored functional role of genomic architecture in sole is expected to be crucial in understanding adaptation. A first step has been made with the construction of a linkage map (Diopere *et al.*, 2014) and the characterization of the transcriptome (Benzekri *et al.*, 2014). However, a full understanding of single-locus (hard sweep) and polygenic (soft sweep) adaptation requires high resolution genome mapping (Bernatchez, 2016).

The above mentioned evidence points to the importance of adaptive responses under changing conditions and their significance for fisheries management. However, exploited stocks are subjected to additional anthropogenic pressures such as pollution, habitat modification and fishing pressure. The latter has led to the observation of reduced growth and earlier maturation in fish (Jørgensen *et al.*, 2007), including in sole (Mollet *et al.*,

2007). This suggests that the sole genome adapts directionally through the sorting of alleles, recombination, and epigenetics. Therefore, functional impacts might be suspected in the most heavily exploited sole stocks, such as the North Sea. Most noticeably, the introduction of heavy demersal fishing gear in the 1960 s had a major impact on the benthic communities and exerted a strong selection pressure (Thurstan *et al.*, 2010; Diopere, 2014).

Conclusions

Although sole of the Northeast Atlantic Ocean inhabit an environment with high potential for gene flow, a combination of isolation by distance and local adaptation along a latitudinal gradient structures the populations. Each of the metapopulations of the North Sea and the Bay of Biscay match with a distinct spawning behaviour and winter mortality. We identified a few of the loci potentially involved in adaptation, but genotyping at a higher marker density, combined with experimental testing in common garden settings (de Villemereuil *et al.*, 2016), modelling with integrated genetic dynamic energy budgets (DEB) (Teal *et al.*, 2012) and/or molecular physiological experiments should elucidate other standing questions. They include population-specific timing of the spawning season in relation to local plankton production, the offshore/inshore spawning strategy and the impact of intensive resource exploitation.

Supplementary data

Supplementary material is available at the *ICESJMS* online version of the manuscript.

Acknowledgements

The Research Foundation–Flanders (FWO) supported research through a PhD scholarship to E.D., a postdoctoral fellowship to G.E.M. and the Scientific Research Network "Eco-evolutionary dynamics in natural and anthropogenic communities" (grant W0.037.10 N). S.G.V. received a PhD scholarship from the Institute for Agricultural and Fisheries Research (ILVO). The European Community's 7th Framework Programme (FP7/2007–2013) project FishPopTrace (The structure of fish populations and traceability of fish and fish products) funded the research. The authors thank the crew of R.V. *Zeeleeuw* and *Belgica* and numerous colleagues for sample collection and E. Cuveliers for DNA extraction. They also thank F. Panitz (University of Aarhus, DK) for RNA sequencing, R. Ogden (Trace Wildlife Forensics, Edinburgh, UK) for advice on genotyping, J. Martinsohn and E. Mac Aoidh (European Commission, JRC, Ispra, I) for advice on the mapping of genetic clusters, and F. Calboli for plotting the Structure figures. Finally, sincere thanks to J. Raeymaekers, K. Cottenie, N. Boon and F. Van den Broeck for comments and help with the statistical analyses, C. André, I. Coscia, P. Jonsson, editor L. Hauser, and four anonymous reviewers for commenting on various drafts.

Author contributions

This study is part of the PhD thesis of E.D focusing on local adaptation in marine fishes. F.A.M.V. and G.E.M. supervised E.D. and designed the study with input from A.C. and F.T.; E.D., S.G.V, and F.T. contributed samples. E.D., P.I.H., S.G.V., P.I.H., and J.V. analysed the data. E.D., P.I.H., and F.A.M.V. wrote the paper with contributions from all authors. All authors read and approved the manuscript.

Data submission

Genotypes have been submitted to the dbSNP database at www.ncbi.nlm.nih.gov/SNP under accession numbers ss1026565497 to ss1026566439 and ss503772144 to ss503772273.

References

Amara, R., Lagardere, F., Desaunay, Y., and Marchand, J. R. 2000. Metamorphosis and estuarine colonisation in the common sole, *Solea solea* (L.): implications for recruitment regulation. Oceanologica Acta, 23: 469–484.

Antao, T., Lopes, A., Lopes, R. J., Beja-Pereira, A., and Luikart, G. 2008. LOSITAN: A workbench to detect molecular adaptation based on a F(st)-outlier method. BMC Bioinformatics, 9: 323.

Ayata, S. D., Lazure, P., and Thiebaut, E. 2010. How does the connectivity between populations mediate range limits of marine invertebrates? A case study of larval dispersal between the Bay of Biscay and the English Channel (North-East Atlantic). Progress in Oceanography, 87: 18–36.

Beaugrand, G., Luczak, C., and Edwards, M. 2009. Rapid biogeographical plankton shifts in the North Atlantic Ocean. Global Change Biology, 15: 1790–1803.

Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N., and Bonhomme, F. 2004. GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier (France).

Benjamini, Y., and Hochberg, Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B-Methodological, 57: 289–300.

Benzekri, H., Armesto, P., Cousin, X., Rovira, M., Crespo, D., Merlo, M. A., Mazurais, D., *et al.* 2014. *De novo* assembly, characterization and functional annotation of Senegalese sole (*Solea senegalensis*) and common sole (*Solea solea*) transcriptomes: integration in a database and design of a microarray. BMC Genomics, 15: 18.

Berg, P. R., Jentoft, S., Star, B., Ring, K. H., Knutsen, H., Lien, S., Jakobsen, K. S., *et al.* 2015. Adaptation to low salinity promotes genomic divergence in Atlantic Cod (*Gadus morhua* L.). Genome Biology and Evolution, 7: 1644–1663.

Bernatchez, L. 2016. On the maintenance of genetic variation and adaptation to environmental change: considerations from population genomics in fishes. Journal of Fish Biology, 89: 2519–2556.

Burt, G. J., and Milner, R. S. 2008. Movement of sole in the southern North Sea and eastern English Channel from tagging studies (1955–2004). Science Series Technical Report, 44: 1–44.

Campinho, M. A., Silva, N., Sweeney, G. E., and Power, D. M. 2007. Molecular, cellular and histological changes in skin from a larval to an adult phenotype during bony fish metamorphosis. Cell and Tissue Research, 327: 267–284.

Chua, K. L., and Lim, T. M. 2000. Type I and type II cytokeratin cDNAs from the zebrafish (*Danio rerio*) and expression patterns during early development. Differentiation, 66: 31–41.

Coscia, I., Robins, P. E., Porter, J. S., Malham, S. K., and Ironside, J. E. 2013. Modelled larval dispersal and measured gene flow: seascape genetics of the common cockle *Cerastoderma edule* in the southern Irish Sea. Conservation Genetics, 14: 451–466.

Cushing, D. H. 1969. The regularity of the spawning season of some fishes. Journal Du Conseil International Pour L'Exploration De La Mer, 33: 81–92.

Cuveliers, E. L., Larmuseau, M. H. D., Hellemans, B., Verherstraeten, S. L. N. A., Volckaert, F. A. M., and Maes, G. E. 2012. Multi-marker estimate of genetic connectivity of sole (*Solea solea*) in the North-East Atlantic Ocean. Marine Biology, 159: 1239–1253.

Cuveliers, E. L., Volckaert, F. A. M., Rijnsdorp, A. D., Larmuseau, M. H. D., and Maes, G. E. 2011. Temporal genetic stability and high

effective population size despite fisheries-induced life-history trait evolution in the North Sea sole. Molecular Ecology, 20: 3555–3568.

Darnaude, A. M., Salen-Picard, C., Polunin, N. V. C., and Harmelin-Vivien, M. L. 2004. Trophodynamic linkage between river runoff and coastal fishery yield elucidated by stable isotope data in the Gulf of Lions (NW Mediterranean). Oecologia, 138: 325–332.

de Villemereuil, P., Gaggiotti, O. E., Mouterde, M., and Till-Bottraud, I. 2016. Common garden experiments in the genomic era: new perspectives and opportunities. Heredity, 116: 249–254.

Diopere, E. 2014. Genetic signatures of selection in sole. A spatio-temporal approach. p. 249. University of Leuven, Leuven.

Diopere, E., Maes, G. E., Komen, H., Volckaert, F. A. M., and Groenen, M. A. M. 2014. A genetic linkage map of sole (*Solea solea*): A tool for evolutionary and comparative analyses of exploited (flat)fishes. PLoS One, 9: e115040.

Dray, S., Legendre, P., and Peres-Neto, P. R. 2006. Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). Ecological Modelling, 196: 483–493.

Engelhard, G. H., Pinnegar, J. K., Kell, L. T., and Rijnsdorp, A. D. 2011. Nine decades of North Sea sole and plaice distribution. ICES Journal of Marine Science, 68: 1090–1104.

Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., and Borsboom, D. 2012. qgraph: Network visualizations of relationships in psychometric data. Journal of Statistical Software, 48: 1–18.

Eriksson, B. K., van der Heide, T., van de Koppel, J., Piersma, T., van der Veer, H. W., and Olff, H. 2010. Major changes in the ecology of the Wadden Sea: Human impacts, ecosystem engineering and sediment dynamics. Ecosystems, 13: 752–764.

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

Excoffier, L., and Lischer, H. E. L. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources, 10: 564–567.

Feder, M. E., and Mitchell-Olds, T. 2003. Evolutionary and ecological functional genomics. Nature Reviews Genetics, 4: 651–657.

Fincham, J. I., Rijnsdorp, A. D., and Engelhard, G. H. 2013. Shifts in the timing of spawning in sole linked to warming sea temperatures. Journal of Sea Research, 75: 69–76.

Focant, B., Vandewalle, R., and Huriaux, F. 2003. Expression of myofibrillar proteins and parvalbumin isoforms during the development of a flatfish, the common sole *Solea solea*: comparison with the turbot *Scophthalmus maximus*. Comparative Biochemistry and Physiology B-Biochemistry and Molecular Biology, 135: 493–502.

Foll, M., and 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.

Fonds, M. 1979. A seasonal fluctuation in growth rate of young plaice (*Pleuronectes platessa*) and sole (*Solea solea*) at the laboratory at constant temperatures and a natural daylight cycle. Proceedings of the 13th European Marine Biology Symposium: 151–158.

Forné, I., Castellana, B., Marin-Juez, R., Cerda, J., Abian, J., and Planas, J. V. 2011. Transcriptional and proteomic profiling of flatfish (*Solea senegalensis*) spermatogenesis. Proteomics, 11: 2195–2211.

Fox, C. J., McCloghrie, P., Young, E. F., and Nash, R. D. M. 2006. The importance of individual behaviour for successful settlement of juvenile plaice (*Pleuronectes platessa* L.): a modelling and field study in the eastern Irish Sea. Fisheries Oceanography, 15: 301–313.

Fox, C. J., O'Brien, C. M., Dickey-Collas, M., and Nash, R. D. M. 2000. Patterns in the spawning of cod (*Gadus morhua* L.), sole (*Solea solea* L.) and plaice (*Pleuronectes platessa* L.) in the Irish Sea as determined by generalized additive modelling. Fisheries Oceanography, 9: 33–49.

Fraser, H. B. 2013. Gene expression drives local adaptation in humans. Genome Research, 23: 1089–1096.

Frichot, E., and Francois, O. 2015. LEA: an R package for landscape and ecological association studies. Methods in Ecology and Evolution, 6: 925–929.

Frichot, E., Schoville, S. D., Bouchard, G., and François, O. 2013. Testing for associations between loci and environmental gradients using latent factor mixed models. Molecular Biology and Evolution, 30: 1687–1699.

Guinand, B., Rolland, J. L., and Bonhomme, F. 2008. Genetic structure of the common sole (*Solea solea*) in the Bay of Biscay: nurseries as units of selection? Estuarine Coastal and Shelf Science, 78: 316–326.

Guo, B. C., Li, Z. T., and Merilä, J. 2016. Population genomic evidence for adaptive differentiation in the Baltic Sea herring. Molecular Ecology, 25: 2833–2852.

Hauser, L., and Carvalho, G. R. 2008. Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. Fish and Fisheries, 9: 333–362.

Hedgecock, D. 1994. Does Variance in Reproductive Success Limit Effective Population Sizes of Marine Organisms? Chapman & Hall, London, Bangor, Wales. 122–134 pp.

Heessen, H. J. L., Daan, N., and Ellis, J. R. 2015. Fish Atlas of the Celtic Sea, North Sea, and Baltic Sea Based On International Research-Vessel Surveys. Wageningen Academic Publishers. 572 pp.

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

Hubisz, M. J., Falush, D., Stephens, M., and Pritchard, J. K. 2009. Inferring weak population structure with the assistance of sample group information. Molecular Ecology Resources, 9: 1322–1332.

ICES 2016. ICES Advice on fishing opportunities, catch and effort Greater North Sea and Celtic Seas ecoregions. International Council for the Exploration of the Sea, Copenhagen.

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

Jørgensen, C., Enberg, K., Dunlop, E. S., Arlinghaus, R., Boukal, D. S., Brander, K., Ernande, B., *et al.* 2007. Ecology: managing evolving fish stocks. Science, 318: 1247–1248.

Kostecki, C., Roussel, J. M., Desroy, N., Roussel, G., Lanshere, J., Le Bris, H., and Le Pape, O. 2012. Trophic ecology of juvenile flatfish in a coastal nursery ground: contributions of intertidal primary production and freshwater particulate organic matter. Marine Ecology Progress Series, 449: 221–232.

Kotoulas, G., Bonhomme, F., and Borsa, P. 1995. Genetic structure of the common sole *Solea vulgaris* at different geographic scales. Marine Biology, 122: 361–375.

Kritzer, J. P., and Sale, P. F. 2004. Metapopulation ecology in the sea: from Levins' model to marine ecology and fisheries science. Fish and Fisheries, 5: 131–140.

Lacroix, G., Barbut, L., and Volckaert, F. A. M. in press. Complex response of projected sea temperature and wind change on flatfish dispersal. Climate Change Biology,.

Lacroix, G., Maes, G. E., Bolle, L. J., and Volckaert, F. A. M. 2013. Modelling dispersal dynamics of the early life stages of a marine flatfish (*Solea solea* L.). Journal of Sea Research, 84: 13–25.

Lamichhaney, S., Barrio, A. M., Rafati, N., Sundstrom, G., Rubin, C.-J., Gilbert, E. R., Berglund, J., *et al.* 2012. Population-scale sequencing reveals genetic differentiation due to local adaptation in Atlantic herring. Proceedings of the National Academy of Sciences of the United States of America, 109: 19345–19350.

Le Moan, A., Gagnaire, P. A., and Bonhomme, F. 2016. Parallel genetic divergence among coastal-marine ecotype pairs of European anchovy explained by differential introgression after secondary contact. Molecular Ecology, 25: 3187–3202.

Le Pape, O., and Bonhommeau, S. 2015. The food limitation hypothesis for juvenile marine fish. Fish and Fisheries, 16: 373–398.

Le Pape, O., and Cognez, N. 2016. The range of juvenile movements of estuarine and coastal nursery dependent flatfishes: estimation from a meta-analytical approach. Journal of Sea Research, 107: 43–55.

Le Pape, O., Holley, J., Guerault, D., and Desaunay, Y. 2003. Quality of coastal and estuarine essential fish habitats: estimations based on the size of juvenile Common Sole (*Solea Solea* L.). Estuarine Coastal and Shelf Science, 58: 793–803.

Lescrauwaet, A. K., Fockedey, N., Debergh, H., Vincx, M., and Mees, J. 2013. Hundred and eighty years of fleet dynamics in the Belgian sea fisheries. Reviews in Fish Biology and Fisheries, 23: 229–243.

Limborg, M. T., Helyar, S. J., de Bruyn, M., Taylor, M. I., Nielsen, E. E., Ogden, R., Carvalho, G. R., *et al.* 2012. Environmental selection on transcriptome-derived SNPs in a high gene flow marine fish, the Atlantic herring (*Clupea harengus*). Molecular Ecology, 21: 3686–3703.

Maggs, C. A., Castilho, R., Foltz, D., Henzler, C., Jolly, M. T., Kelly, J., Olsen, J., *et al.* 2008. Evaluating signatures of glacial refugia for North Atlantic benthic marine taxa. Ecology, 89: S108–S122.

Milano, I., Babbucci, M., Panitz, F., Ogden, R., Nielsen, R. O., Taylor, M. I., Helyar, S. J., *et al.* 2011. Novel tools for conservation genomics: comparing two high-throughput approaches for SNP discovery in the transcriptome of the European hake. PLoS One, 6: 1–13.

Moksnes, P. O., Corell, H., Tryman, K., Hordoir, R., and Jonsson, P. R. 2014. Larval behavior and dispersal mechanisms in shore crab larvae (*Carcinus maenas*): local adaptations to different tidal environments? Limnology and Oceanography, 59: 588–602.

Mollet, F. M., Engelhard, G. H., Vainikka, A., Laugen, A. T., Rijnsdorp, A. D., and Ernande, B. 2013. Spatial variation in growth, maturation schedules and reproductive investment of female sole *Solea solea* in the Northeast Atlantic. Journal of Sea Research, 84: 109–121.

Mollet, F. M., Kraak, S. B. M., and Rijnsdorp, A. D. 2007. Fisheries-induced evolutionary changes in maturation reaction norms in North Sea sole *Solea solea*. Marine Ecology-Progress Series, 351: 189–199.

Narum, S. R., and Hess, J. E. 2011. Comparison of F_{ST} outlier tests for SNP loci under selection. Molecular Ecology Resources, 11: 184–194.

Nielsen, E. E., Cariani, A., Mac Aoidh, E., Maes, G. E., Milano, I., Ogden, R., Taylor, M., *et al.* 2012. Gene-associated markers provide tools for tackling illegal fishing and false eco-certification. Nature Communications, 3: 1–6.

Nosil, P., and Feder, J. L. 2013. Genome evolution and speciation: toward quantitative descriptions of patterns and processes. Evolution, 67: 2461–2467.

Oksanen, J. 2011. Vegan: R functions for vegetation ecologists. https://cran.r-project.org/web/packages/vegan/index.html.

Orsini, L., Vanoverbeke, J., Swillen, I., Mergeay, J., and De Meester, L. 2013. Drivers of population genetic differentiation in the wild: isolation by dispersal limitation, isolation by adaptation and isolation by colonization. Molecular Ecology, 22: 5983–5999.

Palumbi, S. R. 2003. Population genetics, demographic connectivity, and the design of marine reserves. Ecological Applications, 13: S146–S158.

Pepin, P. 1993. Application of empirical size-dependent models of larval fish vital-rates to the study of production-accuracy associated with adult stock dynamics in a comparison among species. Canadian Journal of Fisheries and Aquatic Sciences, 50: 53–59.

Pinsky, M. L., and Palumbi, S. R. 2014. Meta-analysis reveals lower genetic diversity in overfished populations. Molecular Ecology, 23: 29–39.

Pritchard, J. K., Stephens, M., and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. Genetics, 155: 945–959.

Pujolar, J. M., Jacobsen, M. W., Als, T. D., Frydenberg, J., Munch, K., Jonsson, B., Jian, J. B., *et al.* 2014. Genome-wide single-generation signatures of local selection in the panmictic European eel. Molecular Ecology, 23: 2514–2528.

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

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

Richards, R. C., Short, C. E., Driedzic, W. R., and Ewart, K. V. 2010. Seasonal changes in hepatic gene expression reveal modulation of multiple processes in rainbow smelt (*Osmerus mordax*). Marine Biotechnology, 12: 650–663.

Rijnsdorp, A. D., Van Keeken, O. A., and Bolle, L. J. 2004. Changes in the productivity of the southeastern North Sea as reflected in the growth of plaice and sole. ICES, 13: 15.

Rijnsdorp, A. D., Vanbeek, F. A., Flatman, S., Millner, R. M., Riley, J. D., Giret, M., and Declerck, R. 1992. Recruitment of sole stocks, *Solea solea* (L), in the Northeast Atlantic. Netherlands Journal of Sea Research, 29: 173–192.

Rijnsdorp, A. D., and Vingerhoed, B. 1994. The ecological significance of geographical and seasonal differences in egg size in sole *Solea solea* (L). Netherlands Journal of Sea Research, 32: 255–270.

Savina, M., and Ménesguen, A. 2008. A deterministic population dynamics model to study the distribution of a benthic bivalve with planktonic larvae (*Paphia rhomboides*) in the English Channel (NW Europe). Journal of Marine Systems, 70: 63–76.

Selkoe, K. A., D'Aloia, C. C., Crandall, E. D., Iacchei, M., Liggins, L., Puritz, J. B., von der Heyden, S., *et al.* 2016. A decade of seascape genetics: contributions to basic and applied marine connectivity. Marine Ecology Progress Series, 554: 1–19.

Selkoe, K. A., Henzler, C. M., and Gaines, S. D. 2008. Seascape genetics and the spatial ecology of marine populations. Fish and Fisheries, 9: 363–377.

Siegel, D. A., Kinlan, B. P., Gaylord, B., and Gaines, S. D. 2003. Lagrangian descriptions of marine larval dispersion. Marine Ecology-Progress Series, 260: 83–96.

Sparrevohn, C. R., Lindegren, M., and Mackenzie, B. R. 2013. Climate-induced response of commercially important flatfish species during the 20th century. Fisheries Oceanography, 22: 400–408.

Stern, D. L., and Orgogozo, V. 2008. The loci of evolution: How predictable is genetic evolution? Evolution, 62: 2155–2177.

Swaegers, J., Janssens, S. B., Ferreira, S., Watts, P. C., Mergeay, J., McPeek, M. A., and Stoks, R. 2014. Ecological and evolutionary drivers of range size in Coenagrion damselflies. Journal of Evolutionary Biology, 27: 2386–2395.

Teal, L. R., de Leeuw, J. J., van der Veer, H. W., and Rijnsdorp, A. D. 2008. Effects of climate change on growth of 0-group sole and plaice. Marine Ecology Progress Series, 358: 219–230.

Teal, L. R., van Hal, R., van Kooten, T., Ruardij, P., and Rijnsdorp, A. D. 2012. Bio-energetics underpins the spatial response of North Sea plaice (*Pleuronectes platessa* L.) and sole (*Solea solea* L.) to climate change. Global Change Biology, 18: 3291–3305.

Team, R. D. C. 2011. R: A language and environment for statistical computing. Foundation for Statistical Computing Vienna, Austria.

Thurstan, R. H., Brockington, S., and Roberts, C. M. 2010. The effects of 118 years of industrial fishing on UK bottom trawl fisheries. Nature Communications, 1: 1–6.

Tysklind, N., Taylor, M. I., Lyons, B. P., Goodsir, F., McCarthy, I. D., and Carvalho, G. R. 2013. Population genetics provides new insights into biomarker prevalence in dab (*Limanda limanda* L.): a key marine biomonitoring species. Evolutionary Applications, 6: 891–909.

Vandamme, S. G., Maes, G. E., Raeymaekers, J. A. M., Cottenie, K., Imsland, A. K., Hellemans, B., Lacroix, G., *et al.* 2014. Regional environmental pressure influences population differentiation in turbot (*Scophthalmus maximus*). Molecular Ecology, 23: 618–636.

Wang, B., Mezlini, A., Demir, F., Fiume, M., Tu, Z., Brudno, M., Haibe-Kains, B., *et al.* 2015. Similarity Network Fusion - Package 'SNFtool' 2.2 edn. https://cran.r-project.org/web/packages/SNFtool/index.html.

Wang, B., Mezlini, A. M., Demir, F., Fiume, M., Tu, Z. W., Brudno, M., Haibe-Kains, B., *et al.* 2014. Similarity network fusion for aggregating data types on a genomic scale. Nature Methods, 11: 333–319.

Wright, D., Bishop, J. M., Matthee, C. A., and von der Heyden, S. 2015. Genetic isolation by distance reveals restricted dispersal across a range of life histories: implications for biodiversity conservation planning across highly variable marine environments. Diversity and Distributions, 21: 698–710.

Yeaman, S., and Whitlock, M. C. 2011. The genetic architecture of adaptation under migration-selection balance. Evolution, 65: 1897–1911.

Handling editor: Lorenz Hauser