12-13-2018

Population genomics and phylogeography of a benthic coastal shark (Scyliorhinus canicula) using 2b-RAD single nucleotide polymorphisms

Alice Manuzzi

Lorenzo Zane

Antonio Munoz-Merida

Andrew M. Griffiths

Ana Verissimo
Virginia Institute of Marine Science

Recommended Citation

Manuzzi, Alice; Zane, Lorenzo; Munoz-Merida, Antonio; Griffiths, Andrew M.; and Verissimo, Ana, Population genomics and phylogeography of a benthic coastal shark (Scyliorhinus canicula) using 2b-RAD single nucleotide polymorphisms (2018). *Biological Journal of the Linnean Society*, 126(2), 289-303. doi: 10.1093/biolinnean/bly185

Biological Journal of the Linnean Society, 2019, **126**, 289–303. With 6 figures.

Population genomics and phylogeography of a benthic coastal shark (*Scyliorhinus canicula*) using 2b-RAD single nucleotide polymorphisms

ALICE MANUZZI[1,2], LORENZO ZANE[3,4], ANTONIO MUÑOZ-MERIDA[1], ANDREW M. GRIFFITHS[5] and ANA VERÍSSIMO[1,6,*]

[1]*CIBIO – U.P. – Research Center for Biodiversity and Genetic Resources, Campus Agrário de Vairão, 4485–661 Vairão, Portugal*
[2]*National Institute of Aquatic Resources, Technical University of Denmark, Vejlsøvej 39, 8600 Silkeborg, Denmark*
[3]*Department of Biology, University of Padova, via U. Bassi 58/B, 35131 Padova, Italy*
[4]*Consorzio Nazionale Interuniversitario per le Scienze del Mare (CoNISMa), Piazzale Flaminio 9, 00196 Roma, Italy*
[5]*School of Biosciences, University of Exeter, Devon EX4 4PS, UK*
[6]*Virginia Institute of Marine Science, College of William and Mary, Route 1208, Greate Road, Gloucester Point, VA 23062, USA*

Received 24 July 2018; revised 30 October 2018; accepted for publication 30 October 2018

The existence of strong genetic structure is expected in species with limited ability to disperse and philopatric behaviour. These life-history traits are found in many small benthic elasmobranchs, such as in the small-spotted catshark (*Scyliorhinus canicula*). However, no evidence of genetic structure was found across its northeastern Atlantic (NEA) range using traditional molecular markers. Here, fine-scale genetic differentiation was detected between the British Isles and southern Iberia using 2674 single nucleotide polymorphism loci generated using 2b-restriction site-associated DNA (2b-RAD). Geographical distance and historical demography were two major drivers shaping the distribution of genetic diversity of *S. canicula* along the NEA. Significant positive spatial autocorrelation of allelic frequencies was detected, with genetic differentiation generally increasing with geographical distance. However, marked genetic divergence of the Celtic Sea and South Portugal collections from their closest neighbours resulted in geographically constrained genetic breaks south of the British Isles and off southwestern Iberia. Historical demographic reconstruction of population pairs across these genetic breaks suggested a scenario of historical isolation before secondary contact, probably related to distinct northern and southern glacial refugia. These results provide new insights into the population structure of *S. canicula* along the NEA and serve as a reference for benthic elasmobranchs with similar distribution ranges.

ADDITIONAL KEYWORDS: catsharks – genetic breaks – glacial refugia – isolation by distance – local populations.

INTRODUCTION

Uncovering the patterns of genetic diversity distribution within species can provide important insights into the patterns and drivers of differentiation among populations. Population genetic studies can also contribute valuable information on the biology, ecology and behaviour of species, particularly for species where direct observations are limited, such as in many marine species, including Elasmobranchs (Dudgeon *et al.*, 2012; Portnoy & Heist, 2012; Feldheim *et al.*, 2014; Domingues *et al.*, 2018). Indeed, the last two decades have seen significant progress in our knowledge of the genetic diversity and population structure of sharks and rays (Dudgeon *et al.*, 2012; Portnoy & Heist, 2012). These studies have largely relied on traditional molecular markers, such as

*Corresponding author. E-mail: averissimo@cibio.up.pt

nuclear microsatellite loci and mitochondrial genes. However, previous studies focused mostly on highly mobile, pelagic sharks (particularly carcharhinids), whereas limited knowledge is available for less vagile, benthic species (Domingues *et al.*, 2018).

The small-spotted catshark *Scyliorhinus canicula* (Linnaeus, 1758) is a small coastal benthic shark distributed from Norway to Senegal and across the Mediterranean Sea (Ebert *et al.*, 2013). It shows low dispersal (< 300 km) and philopatric behaviour (Sims *et al.*, 2001; Rodríguez-Cabello *et al.*, 2004). Reproduction is oviparous with direct development, where the egg cases are attached to a variety of substrates (Ellis & Shackley, 1997; Rodríguez-Cabello *et al.*, 1998). Thus, the species lacks dispersive egg or larval stages and exhibits limited movement in adults. In line with expectations from its life-history traits, previous studies showed strong genetic divergence over small spatial scales in *S. canicula* from the Mediterranean Sea, using mitochondrial sequences and nuclear microsatellite loci (Barbieri *et al.*, 2014; Gubili *et al.*, 2014; Ramírez-Amaro *et al.*, 2018). Interestingly, a contrasting pattern was observed across similar spatial scales along the northeast Atlantic (NEA) coast, because no significant genetic differentiation was identified (Gubili *et al.*, 2014), perhaps reflecting increased connectivity between catsharks in this region.

The detection of genetic population structure using a few neutral markers (e.g. nuclear microsatellites) may be difficult. Instead, large numbers of markers are desirable to reduce inter-locus sampling variance and increase the power for detecting genetic differentiation among populations (André *et al.*, 2011). Single nucleotide polymorphisms (SNPs) are rapidly becoming the preferred markers in genetic studies of natural populations owing to their high abundance across the genome (approximately one or two SNPs per 1000 bp), their ability to track both neutral and non-neutral variation, and the potential for direct cross-study comparisons (Morin *et al.*, 2004).

One approach for simultaneous SNP calling and genotyping is based on restriction enzyme digestion to reduce genome complexity (i.e. restriction site-associated DNA sequencing, RADseq; for a review, see Andrews *et al.*, 2016), which allows screening hundreds to tens of thousands of SNP loci in population-level studies. These methods thus increase the power and the resolution of population-level comparisons (Luikart *et al.*, 2003; Nielsen *et al.*, 2009; Hemmer-Hansen *et al.*, 2014; Andrews *et al.*, 2016). Among the many approaches derived from the original RAD protocol (Miller *et al.*, 2007; Baird *et al.*, 2008), the 2b-RAD method (Wang *et al.*, 2012) uses a IIB-type restriction enzyme that produces fragments of uniform length (~33–36 bp) that are sequenced in next-generation sequencing (NGS) platforms. This method is simple and cost-effective, being easy to implement and requiring no intermediate purification steps or size selection of restriction fragments (Wang *et al.*, 2012; Andrews *et al.*, 2014). It also allows screening of the full range of restriction fragments generated, and the number of loci/marker density can be adjusted by using selective adaptors (Wang *et al.*, 2012).

Here, a 2b-RAD method for simultaneous SNP calling and genotyping is applied to a small benthic coastal shark, *S. canicula*, to investigate the population structure with a large panel of SNPs across part of its NEA range, i.e. from Scotland to southern Portugal. It is hypothesized that the increased power gained with the large number of genetic markers will reveal significant population structure in *S. canicula* across this region. This is in line with its life history, showing limited ability to disperse and philopatric behaviour, which will act to limit gene flow. The results will also allow for a direct comparison with those obtained with traditional markers (e.g. mitochondrial control region and nuclear microsatellites). Furthermore, the results can provide important insights into the patterns and drivers of genetic diversity distribution in many other elasmobranchs sharing similar life-history traits with *S. canicula*, e.g. highly diverse families, such as Scyliorhinidae and Rajidae.

MATERIAL AND METHODS

Sample collection and DNA extraction

Sampling aimed at covering the western European coast in the NEA and included five different locations (Table 1): the North Sea and NW Scotland (hereafter North Sea), the Irish Sea, the Celtic Sea, northwestern Portugal (NW Portugal) and southern Portugal (S Portugal) (Fig. 1). Tissue sample collections (muscle and fin clips) of *S. canicula* (N = 88) were obtained during scientific research surveys or from locally operating commercial fishing vessels and preserved in 96% ethanol. Genomic DNA (gDNA) was extracted using the EasySpin Genomic DNA Tissue Kit (Citomed, Lisbon, Portugal), according to the manufacturer's protocol, and checked for quality and quantity (see Supporting Information, S1 for details). All samples were standardized to a final gDNA concentration of 25 ng/μL before library preparation.

2b-RAD library preparation

Single nucleotide polymorphisms were simultaneously identified and genotyped *de novo* on the five sample collections, following the 2b-RAD protocol of Paterno *et al.* (2017), with modifications (for details, see

Table 1. Genetic diversity indices and levels of missing data across sample collections of the small-spotted catshark *Scyliorhinus canicula* from the northeast Atlantic

Parameter	Celtic Sea	Irish Sea	North Sea	NW Portugal	S Portugal
Sample size (number of individuals)	17	15	13	19	7
Number of loci with no missing genotypes (% of total)	1500 (56)	2573 (96)	2481 (93)	1645 (62)	2004 (75)
Average percentage of missing genotypes across loci	4.4	0.3	0.6	2.8	4.5
Average percentage of missing genotypes in loci with missing data	10	7	9	7	18
Number of alleles scored	5225	5320	5309	5332	4897
Average R_S	1.79	1.88	1.86	1.85	1.72
Sum of R_S across loci	4794.0	5024.5	4978.2	4959.7	4603.7
Average H_O	0.58	0.73	0.69	0.66	0.53

Abbreviations: H_O, observed heterozygosity; R_S, allelic richness across loci.

Supplemental Information, Supplemental document S1). Briefly, high-quality gDNA was digested with the IIB-type restriction enzyme *CspC*I to generate a pool of fragments of uniform length (~33 bp). The fragments were ligated to non-selective adapters before incorporation of individual barcodes with Illumina-compatible adaptors via the polymerase chain reaction. Individual barcoded libraries were pooled in equimolar amounts, run on agarose gel electrophoresis and excised from the gel to remove primer dimers. The final pool was purified with Agencourt AMPure XP beads (Beckman Coulter, Indianapolis, IN, USA) to retain only the targeted restriction fragments, which have a final length of ~160 bp. Final validation and quantification of the pooled library concentration was performed with the KAPA Library Quantitation kit (Kapa Biosystems, Boston, MA, USA), according to the manufacturer's instructions. The cleaned pooled library was sequenced in one lane of an Illumina HiSeq 1500 (San Diego, CA, USA) rapid run of 101 paired-end cycles.

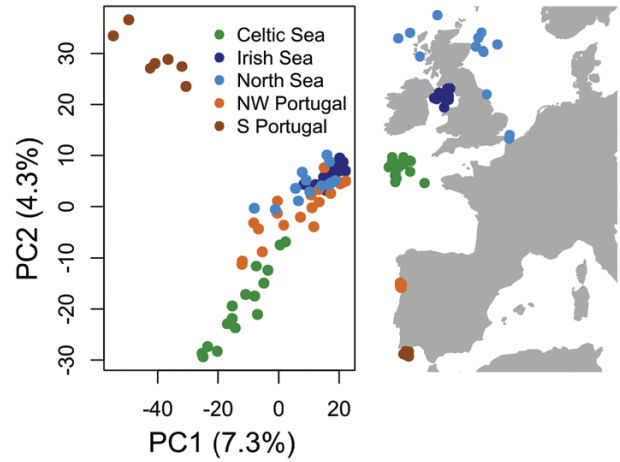

Figure 1. Principal components analysis of 2674 single nucleotide polymorphism loci of the small-spotted catshark *Scyliorhinus canicula* (left), considering 20% missing data at the locus and individual levels. The locations of the 71 individuals included in the final dataset are shown in the map (right). Abbreviation: PC, principal component.

Data processing

Sequence reads were demultiplexed using a custom-made Perl script, and only the forward reads were kept for subsequent analyses, because forward and reverse reads had a complete overlap of each restriction fragment. Preliminary quality checks were made using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc, last accessed on Nov 13, 2018.), followed by adaptor removal and quality trimming using PRINSEQ Lite Software (Schmieder & Edwards, 2011), with the following options: 3′-end trimming of positions with quality scores < 28; removal of reads with minimum mean quality scores < 28; trimming of poly-N tails at the 5′ end; and trimming of all reads to 40 bp of total length (which includes the restriction site of *CspC*I, located in the middle of the fragment).

De novo assembly, read mapping, SNP calling and genotyping were performed using the *dDocent* pipeline (Puritz *et al.*, 2014) with default parameters, and including a minimum of three supporting reads per variant and a minimum frequency of base call of 0.35 across the whole sample. The resulting raw SNP dataset from *dDocent* was filtered further using VCFtools (Danecek *et al.*, 2011), using the sequential filters as outlined in the *dDocent* tutorial (https://github.com/jpuritz/dDocent/blob/master/tutorials/Filtering%20Tutorial.md): (1) filter loci with > 25% missing data and read quality < 30; (2) filter loci with minor allele frequency < 0.05; (3) filter loci at the 5′ and 3′ ends (i.e. positions 1–4 and 37–40), to avoid artefacts introduced by preferential sequencing errors at the read ends; (4) keep only loci with a read balance of alleles at heterozygous sites between 0.25 and 0.75;

and finally, (5) filter loci with excess depth to exclude possible paralogous and multicopy loci (Table 2).

Additional filtering steps were performed in the R package *adegenet* v.2 (Jombart & Ahmed, 2011) to assess the effect of missing data, because VCFtools assumes missing data only when there is no read depth for a given locus and not when a genotype call is missing (i.e. when read depth < 3). Different thresholds of missing data at the locus and individual levels were evaluated to provide a global perspective of the genetic diversity among genotypes within the sampled range using Principal Components Analysis (PCA) on centred and scaled individual allelic frequencies. Data are shown only for the following combinations: (1) the dataset obtained with VCFtools; and the datasets filtered with *adegenet* for (2) 20% and (3) 5% of missing data at the locus and individual levels.

Genetic data analysis

Genetic diversity per sample collection was assessed by calculating the total number of alleles across loci, average allelic richness (Rs) per locus, summed allelic richness across loci and average observed (H_O) heterozygosity, using the R packages *diveRsity* (Keenan *et al.*, 2013) and *adegenet*. In addition, the number of loci with missing genotypes and the average percentage of missing genotypes across loci were presented to allow assessment of potential bias in the dataset.

No outlier loci tests were performed owing to the combination of overall small sample size (< 100 individuals) and a total number of SNPs < 3000, two conditions shown greatly to limit robust (low false-positive) inferences of outlier loci (Ahrens *et al.*, 2018). Thus, the full dataset was used in all genetic data analyses. Genetic differentiation among sample collections was assessed by means of Weir and Cockerham pairwise F_{ST} tests (Weir & Cockerham, 1984), which have been shown to provide unbiased estimates at low levels of genetic differentiation (i.e. $F_{ST} < 0.05$) when samples sizes are small ($N < 10$) and unbalanced between populations (Willing *et al.*, 2012). Jost's D estimates of genetic differentiation (Jost, 2008) were also calculated to allow for comparison between differentiation metrics. Both statistics and their corresponding P-values (based on 1000 bootstrap replicates) were calculated with the R package *STRATAG* (Archer *et al.*, 2017). The resulting P-values were corrected for false positives using Benjamini & Hochberg (1995) false-discovery rate (FDR) correction. To visualize the relative genetic differentiation among sampling collections, Edwards' genetic distances based on allelic frequencies (Cavalli-Sforza & Edwards, 1967; Edwards, 1971) were estimated in *adegenet* and used to construct an unrooted neighbour-joining (NJ) tree using the R package *ape* (Paradis *et al.*, 2004; Popescu *et al.*, 2012). Bootstrap support values for the NJ branches were estimated based on 1000 sampling replicates, using the R package *poppr* (Kamvar *et al.*, 2014, 2015).

Several individual-based analyses were performed to explore the genetic structure of *S. canicula* along the NEA. The spatially explicit Bayesian inference method TESS3 (Caye *et al.*, 2016) was used to estimate individual ancestry coefficients assuming admixture of K ancestral populations. The TESS3 algorithm performs better than similar algorithms (e.g. STRUCTURE; Pritchard *et al.*, 2000) when the levels of ancestral population divergence are low (Durand *et al.*, 2009) and is well suited for species with continuous distributions and where individuals geographically close to each other are more likely to share ancestral genotypes compared with individuals far apart (Caye *et al.*, 2016). TESS3 was run in R using the package *tess3r* (Caye *et al.*, 2016), with K values between one and ten. Five replicate runs per K value, 1000 iterations and a tolerance value of 10^{-7} were used to choose the best K value, using a cross-validation method with 10% of masked data in the test set to assess consistency of results.

The spatial genetic structure was assessed by means of an individual-based spatial principal components analysis (sPCA) as implemented in *adegenet*, using a Gabriel graph (least-squares adjacency graph) to build a connection network from each individual's spatial coordinates. Global (i.e. allelic frequencies at population i are positively correlated with those of its neighbours) and local (i.e. allelic frequencies at population i are negatively correlated with those of its neighbours) structures were evaluated based upon inspection

Table 2. Summary of filtering steps applied to the original single nucleotide polymorphism dataset obtained from *dDocent*

	Loci retained	Loci filtered
Initial SNP dataset	15 411	0
Filter applied:		
1. > 25% missing genotypes, Qual < 30	14 181	1230
2. MAF < 0.05	11 717	2464
3. Filter loci at 5′ and 3′ ends	9703	2014
4. AB < 0.25 and > 0.75	6016	3687
5. Excess read depth	4656	1360
6. > 20% missing data	2674	1982

Loci meeting the different conditions were sequentially removed (loci filtered), and those retained (loci retained) were used in the subsequent filtering step. Abbreviations: AB, allele balance in heterozygotes; MAF, minimum allele frequency; Qual, minimum phred quality score.

of the sPCA eigenvalues and their corresponding Moran's *I* value. Monte Carlo *r*-tests were performed to assess significance of the global and local components of the sPCA using 1000 permutations of individuals in the dataset while preserving the original connection network, as implemented in *adegenet*.

Relative migration rates among all populations were estimated based on allele frequency data according to the method of Sundqvist *et al.* (2016). This method allows detection and estimation of relative migration rates and the direction of migration, testing for asymmetric gene flow. Thus, it is informative in reconstructing source–sink population dynamics and the evolutionary processes leading to the present genetic diversity distribution. Moreover, the outputs are plotted in a network where populations are nodes connected by links (if gene flow is detected), with closely related nodes being plotted closer in the network space. Computations were performed using the *divMigrate* function implemented in the R package *diveRsity* (Keenan *et al.*, 2013), implementing all three measures of genetic differentiation available (i.e. Jost's D, Nei's G_{ST} and effective number of migrants Nm), with asymmetric gene flow between all pairs of populations being tested using 10 000 bootstrap replicates.

Past demographic events influencing the distribution of genetic diversity of *S. canicula* along the NEA were inferred using the composite likelihood approach implemented in the δaδi method (Gutenkunst *et al.*, 2009). This method uses the joint allele frequency spectrum (JAFS) to infer the demographic history of populations, and can be used to run several theoretical models representing likely demographic events (e.g. Tine *et al.*, 2014; Le Moan *et al.*, 2016). Here, a modified version of δaδi was used as described by Tine *et al.* (2014), which includes seven plausible demographic models: strict isolation (SI); isolation with migration (IM); isolation with ancient migration (AM); isolation with recent secondary contact (SC); and three additional models considering heterogeneous migration rates along the genome to account for the effect of selection (AM2m, IM2m and SC2m). The models were run for two subsets of the data, focusing on population pairs across the genetic breaks found with the sPCA, namely NW Portugal vs. S Portugal, and Irish Sea vs. Celtic Sea. The JAFS of each population pair was built by using 12 chromosomes (i.e. six individuals) per population per SNP for the NW Portugal–S Portugal models (given the small sample size in the latter), and 24 chromosomes (i.e. 12 individuals) per population per SNP for the Irish Sea–Celtic Sea models. The original SNP dataset was re-filtered to comply with δaδi requirements, including removal of the minimum allele frequency filter, masking of singletons in each population to minimize the influence of genotyping errors, and retention of a single SNP per RAD tag to eliminate linked loci, resulting in 1452 SNPs for the NW Portugal–S Portugal dataset and 1304 SNPs for the Irish Sea–Celtic Sea dataset. Monomorphic (221 and 365, respectively) and triallelic SNPs (nine and 12, respectively) were also removed from the two datasets. Fifty independent runs were performed in the optimization of each demographic model, and the resulting spectrum was compared with the observed JAFS obtained from the data. Model fitting was evaluated using the Akaike information criterion (AIC), whereby the lowest AIC value obtained with each demographic model was compared across models, and the best demographic model for each population pair was chosen based on the lowest overall AIC value.

RESULTS

Upon adaptor removal and quality trimming of raw reads using PRINSEQ, the number of cleaned reads was on average 1.68 million per individual (minimum–maximum: 265 949–4 968 097). Cleaned reads were clustered into 10 144 contigs, and 27 833 variants were identified with *dDocent*, from which 15 411 occurred in 90% of the sequenced individuals and were kept for further processing. After filtering with VCFtools, a total of 4656 SNPs were retained (Table 2), showing a total read depth per locus > 500 for the vast majority of markers (Supporting Information, Fig. S1). This dataset was subsequently filtered in *adegenet* for different levels of missing genotype data at the locus and individual levels, but visual inspection of the overall genetic diversity and structure in exploratory PCAs showed generally consistent patterns across datasets regardless of the missing data threshold used (Supporting Information, Fig. S2). The dataset including 20% missing data at the locus and individual levels was used for all subsequent analyses, because it was deemed a good compromise between minimizing the bias in the data and maximizing the number of markers and individuals kept. Thus, the final dataset included 2674 SNPs genotyped in 71 individuals from all five sampling locations (Table 2). All markers were biallelic except for 24 SNPs, which were triallelic.

The genetic diversity indices per sample collection, namely average and summed R_S, and H_O showed higher values in the northern-most collections from the Irish Sea and North Sea, and in NW Portugal (Table 1). The lowest diversity was found for the S Portugal collection, followed by the Celtic Sea collection.

The first three axes of the PCA explained 14% of total inertia. Principal component (PC) 1 separated S Portugal from the remainder of the collections, whereas PC2 separated both S Portugal and the Celtic Sea from the other sample collections (Fig. 1). Celtic

Table 3. Weir and Cockerham pairwise F_{ST} estimates (lower diagonal) and Jost's D (upper diagonal), for the small-spotted catshark *Scyliorhinus canicula* from the northeast Atlantic

	Irish Sea	North Sea	Celtic Sea	NW Portugal	S Portugal
Irish Sea		0.0003*	0.0021*	0.0004*	0.0064*
North Sea	0.009*		0.0012*	0.0003	0.0046*
Celtic Sea	0.031*	0.026*		0.0011*	0.0028*
NW Portugal	0.011*	0.009*	0.020*		0.0045*
S Portugal	0.067*	0.057*	0.050*	0.056*	

*Comparisons yielding significant P-values after false-discovery rate correction.

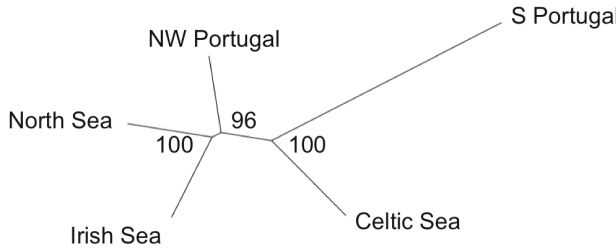

Figure 2. Neighbour-joining tree of Edwards' distances among sample collections. Bootstrap support values are shown, based on 1000 sampling replicates.

Sea individuals were on the opposite extreme of the PC1 axis from their geographically closer North Sea and Irish Sea counterparts and showed no overlap with the latter on the PC2 axis. The NW Portugal sample collection occupied an intermediate position between the two PC1 extremes but had a higher overlap with the two northern-most collections of individuals, i.e. Irish Sea and North Sea.

Levels of pairwise F_{ST} and Jost's D genetic differentiation ranged from a minimum value of 0.009 and 0.0003 between the Irish Sea and the North Sea, respectively, to a maximum of 0.067 and 0.0064 between the Irish Sea and S Portugal, respectively (Table 3). Significant P-values were found for all pairwise comparisons using both F_{ST} and Jost's D after FDR correction, with the exception of the North Sea–NW Portugal comparison using Jost's D statistic (P-value 0.012), indicating significant genetic differentiation among collections (Table 3). The genetic relationships among sample collections inferred from the NJ tree of Edwards' distances had high bootstrap support values (> 95%), with the largest genetic distances existing between S Portugal and all other collections, followed by those including the Celtic Sea. The Irish Sea, North Sea and NW Portugal were equally distant from each other (Fig. 2).

The TESS3 results indicated the most likely number of ancestral populations/genetic clusters as three (Supporting Information, Fig. S3) and showed that sample collections had varied but distinct admixture compositions (Fig. 3; Supporting Information, Table S1). Individuals from S Portugal and the Irish Sea showed > 85% ancestry proportions of two different clusters. The Celtic Sea was predominantly composed of a third cluster (65–93%), with ~20% contribution of the dominant Irish Sea cluster. The North Sea and NW Portugal individuals showed the most admixed compositions, also predominantly composed of the dominant Irish Sea cluster and, to a lesser extent, of the dominant Celtic Sea cluster, but also having 5–9% average contribution of the dominant S Portugal cluster (Fig. 3). On a different note, North Sea individuals showed the largest among-individual variability in ancestry proportions, ranging between 42 and 95% proportion of a given cluster.

The sPCA provided additional insight into the spatial pattern of population structure of *S. canicula*, showing the largest eigenvalues on the positive axis (Supporting Information, Fig. S4) and significant global structure (P-value 0.001). In contrast, no local structure is apparent in the data (P-value 0.084). The first two positive eigenvalues were clearly distinct from the remainder in terms of the variance and spatial autocorrelation components (i.e. had the highest Moran's I values; Supporting Information, Fig. S5), and the corresponding global scores were explored in more detail. The first global scores showed clear north–south separation, with the northern group including the North Sea and Irish Sea individuals, and the southern group including those from the Celtic Sea, NW and S Portugal (Fig. 4A). S Portugal showed larger negative scores compared with the remaining southern collections. The second global scores highlighted the differentiation of the Celtic Sea individuals from all the remaining collections (Fig. 4B). The sPCA highlighted the existence of two strong genetic breaks, i.e. areas of strong genetic divergence, along the sampled range: one slightly south of the British Isles, separating the Celtic Sea from its northern neighbours; and a second one off the SW Iberian coast, separating S Portugal from all other collections.

Relative migration rates varied among pairs of populations sampled, but the patterns were consistent

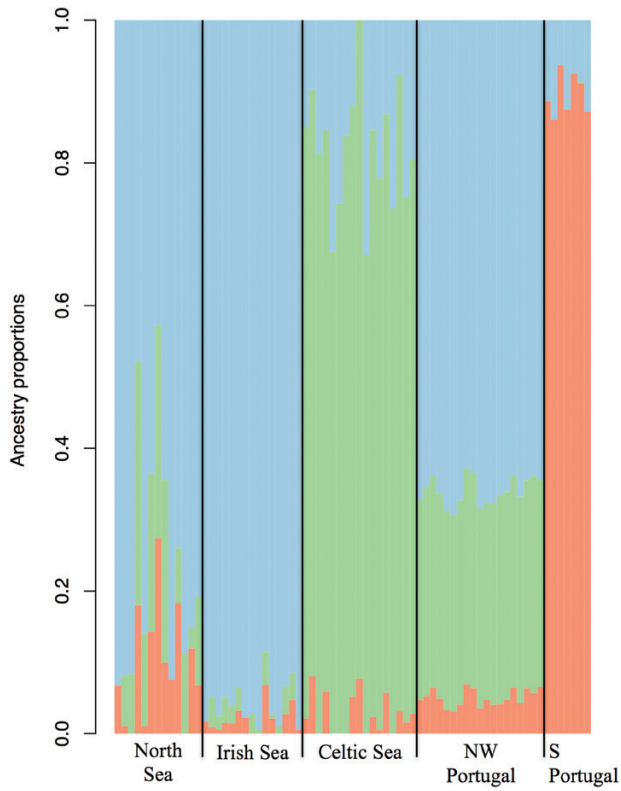

Figure 3. Individual admixture compositions based on K = 3 ancestral populations (marked in blue, green and orange) reconstructed using TESS3, based on 2674 single nucleotide polymorphisms from the small-spotted catshark *Scyliorhinus canicula* sampled along the northeastern Atlantic coast. The collections are separated by black vertical lines. Each individual is represented by a narrow vertical column broken into K = 3 coloured segments, with lengths proportional to each of the inferred K clusters. The results refer to the best run of K = 3 when using 10% of masked loci.

across the three genetic differentiation metrics used (Supporting Information, Tables S2–S4), thus results are shown only for Jost's D (Fig. 5). The highest relative migration rates were found among the Irish Sea, North Sea and NW Portugal, whereas S Portugal showed lower relative migration rates overall and was placed furthest apart in the population network. The Celtic Sea showed intermediate relative rates of migration and also occupied a peripheral position in the network. Only S Portugal and the Celtic Sea showed statistically significant asymmetric gene flow favouring emigration over immigration (Fig. 5; Supporting Information, Table S2).

Similar historical demographic scenarios were found for the population pairs across the two genetic breaks detected with the sPCA, with the SC model having the lowest AIC values for both NW Portugal–S Portugal and Irish Sea–Celtic Sea population pairs (Table 4). The theoretical spectrum of the SC model provided a good fit to the observed JAFS of the NW Portugal–S Portugal divergence (Fig. 6), and the corresponding AIC value was much lower than those obtained with the other models (ΔAIC > 10). Regarding the Irish Sea–Celtic Sea pair, the two models of secondary contact (SC and SC2m) provided the best fit to the observed JAFS, although the simpler SC model performed better (ΔAIC = 6.5). In both cases, the δaδi analyses indicated that populations across the two genetic breaks probably went through a long isolation period before a recent secondary contact, as indicated by the ratio of secondary contact to divergence times for NW Portugal–S Portugal and Irish Sea–Celtic Sea, respectively (Table 4). Also, in both cases, splitting of the ancestral population resulted in much smaller descendant populations, although the effective sizes were similar between NW and S Portugal, whereas the Irish Sea was three times larger than the Celtic Sea (Table 4). Migration rates upon secondary contact were similar between NW Portugal and S Portugal, albeit slightly higher northwards, and they were higher southwards between the Irish Sea and the Celtic Sea (Table 4).

DISCUSSION

Implementation and cost-efficiency of the 2b-RAD protocol

Here, the 2b-RAD protocol was implemented successfully for *de novo* calling and genotyping of thousands of SNP markers in a sample collection of ~100 individuals of a species with a relatively large genome size (3.5 Gb; Wyffels *et al.*, 2014). Studies using RAD-based methods for *de novo* SNP identification and genotyping on elasmobranch taxa are still scarce but have also produced thousands of SNPs (range, 2674–8103), regardless of the RAD method applied (i.e. 2b-RAD in the present study; ddRAD, Portnoy *et al.*, 2015; and DArTseq, Momigliano *et al.*, 2017; Pazmiño *et al.*, 2017). However, the 2b-RAD method implemented here could be a more cost-efficient approach to obtain individual genotypes for population genetic studies, at a relatively low cost (< €3000). This is particularly relevant when the main goal is to obtain a high-resolution perspective on the genetic population structure of a given taxon for which no genetic resources are available (as is the case for the vast majority of elasmobranchs), compared with *de novo* development and genotyping of ten to 20 microsatellite loci.

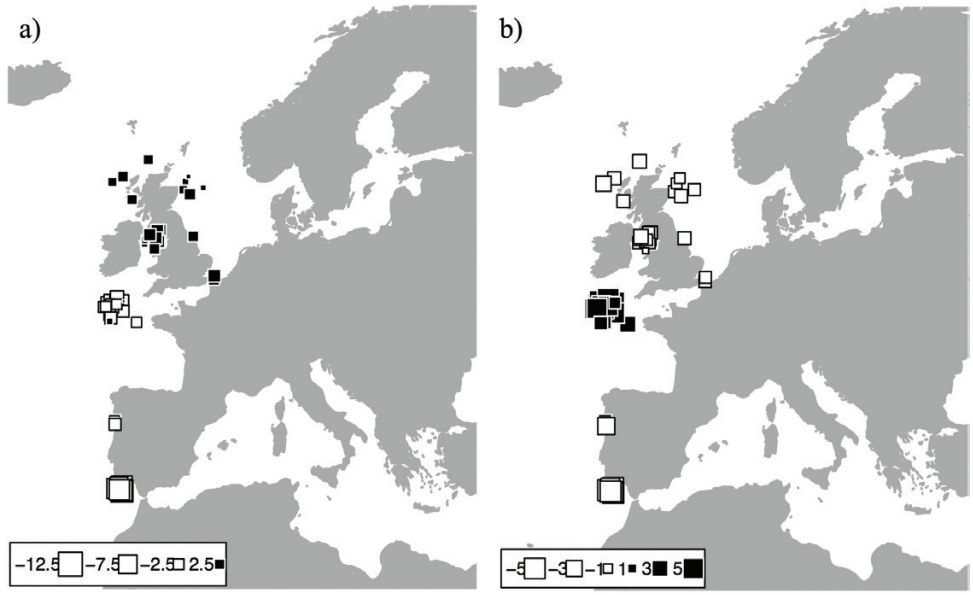

Figure 4. Global spatial principal components analysis (sPCA) scores based on 2674 single nucleotide polymorphism loci and 71 individuals of the small-spotted catshark *Scyliorhinus canicula*, along the northeastern Atlantic coast. A, first global sPCA lagged scores. B, second global sPCA lagged scores.

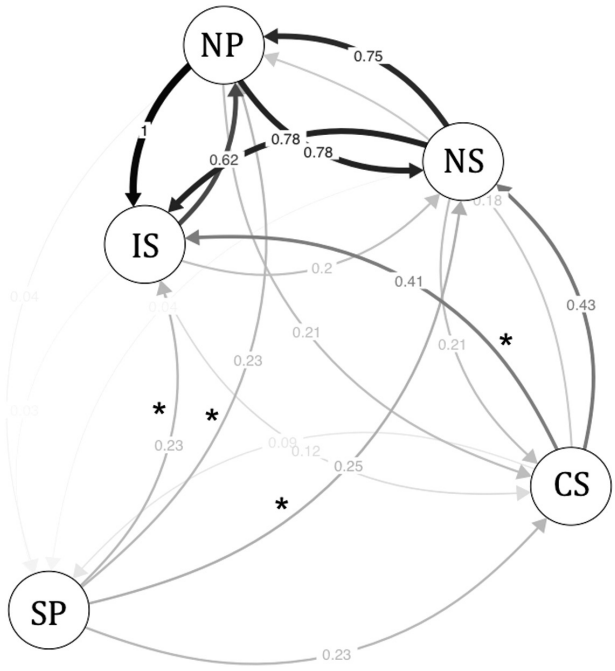

Figure 5. Population network and relative migration rates based on Jost's *D* estimates of genetic differentiation. The thickness of connecting lines is proportional to the relative rate of migration. Abbreviations are as follows: CS, Celtic Sea; IS, Irish Sea; NP, NW Portugal; NS, North Sea; SP, S Portugal. Statistically significant asymmetric rates are marked with an asterisk.

Genetic diversity and structure of *S. canicula* along the northeastern Atlantic

Levels of genetic diversity were moderate across all collections but were generally higher in the northern locations and lower in S Portugal. The estimated diversity indices for S Portugal might have been affected by its small sample size, although estimates based on allelic richness were still lowest at this location (Table 1). The genetic diversity of *S. canicula* was spatially structured along the NEA coast, consistent with the limited dispersal and philopatric behaviour previously described for the species (Rodríguez-Cabello *et al.*, 1998, 2004; Sims *et al.*, 2001). Indeed, significant genetic differentiation was found among all sampled collections (except between North Sea and NW Portugal using Jost's *D*), with each showing distinct admixture coefficients (Fig. 3). The spatial distances at which differentiation was detected here (~500–3500 km) are in line with those reported for the species in the Mediterranean Sea (Gubili *et al.*, 2014; Cardoso, 2015; Kousteni *et al.*, 2015; Ramírez-Amaro *et al.*, 2018). Overall, these results support the hypothesis that *S. canicula* forms local populations across its range.

These results contrast with those obtained previously for *S. canicula* from the same geographical area, whereby the mitochondrial control region (mtDNA CR) and 12 nuclear microsatellite loci showed evidence of only weak genetic differentiation (Gubili *et al.*, 2014; Cardoso, 2015). This discrepancy might be attributable to a combination

Table 4. Comparison of the seven demographic models tested with δaδi for two population pairs of *Scyliorhinus canicula* across two genetic breaks: (1) NW Portugal and S Portugal; and (2) Irish Sea and Celtic Sea, and their corresponding demographic parameters

1. NW Portugal (Pop 1) vs. S Portugal (Pop 2)

Model	Log likelihood	AIC	theta	Pop 1	Pop 2	m12	m21	Ts	Tsc/Tam
SI	−616.31	1238.61	493.01	61.40	0.01	–	–	1.50×10^{-3}	–
SC	**−224.56**	**461.12**	**121.77**	**0.40**	**0.48**	**23.1**	**26.8**	**7.58**	**0.063**
AM	−521.12	1054.25	732.56	0.23	0.08	0.0	59.7	0.01	0.001
IM	−548.13	1106.26	533.96	14.44	0.06	38.2	0.0	1.70×10^{-3}	–
SC2m	−448.25	914.5	757.08	0.15	0.08	0.0	149.8	1.07×10^{-7}	0.090
AM2m	−448.78	915.57	738.2	0.18	0.14	0.0	145.1	0.11	0.000
IM2m	−495.37	1006.75	706.08	0.23	0.21	26.7	0.0	0.15	–

2. Irish Sea (Pop 1) vs. Celtic Sea (Pop 2)

Model	Log likelihood	AIC	theta	Pop 1	Pop 2	m12	m21	Ts	Tsc/Tam
SI	−1199.35	2404.69	387.33	42.30	0.01	–	–	7.90×10^{-4}	–
SC	**−544.801**	**1101.601**	**128.34**	**0.60**	**0.18**	**40.1**	**24.8**	**7.81**	**0.041**
AM	−822.62	1657.23	705.79	16.90	0.08	57.3	42.8	0.13	0.000
IM	−874.56	1759.12	685.16	39.10	0.18	39.8	0.5	0.15	–
SC2m	−545.06	1108.11	122.59	0.41	0.17	43.4	35.3	8.53	0.048
AM2m	−794.01	1606.02	792.08	7.88	0.08	0.0	0.0	0.06	0.000
IM2m	−777.24	1570.47	789.81	8.74	0.07	147.7	2.4	0.06	–

Abbreviations: AIC, Akaike information criterion; m12 and m21, migration rate from Pop 1 to Pop 2, and vice versa, respectively; Pop 1 and Pop 2, effective population size for each of the two descendant populations; theta, effective population size for the ancestral population; Ts, time since split of the ancestral population into the two descendant populations; Tsc/Tam, duration of secondary contact (for SC and SC2m models) and of the ancestral migration (for models AM and AM2m). Abbreviations for model designations follow those described in the Material and Methods section. The best demographic model in for each population pair is marked in bold.

of unbalanced spatial coverage of sample collections along the NEA, particularly the under-representation of Iberian collections, combined with differences owing to the type and number of markers. The large number of SNPs screened here coupled with a more even representation of samples along the study area might have increased the power to detect genetic divergence among populations and increased the resolution of the spatial genetic structure pattern.

The detection of fine-scale population structure in *S. canicula* shown here is also in contrast to findings from studies on several other elasmobranchs along the NEA using mitochondrial gene sequences and nuclear microsatellites. Specifically, no genetic differentiation was found among samples of coastal and deepwater elasmobranchs, including several squaloid sharks (Veríssimo *et al.*, 2010, 2011, 2012; Cunha *et al.*, 2012; Gubili *et al.*, 2016) and skates (Griffiths *et al.*, 2010, 2011). The use of large numbers of SNPs can thus provide increased power for detecting genetic differentiation in elasmobranch taxa, including highly mobile species, such as in the bonnethead *Sphyrna tiburo* (Portnoy *et al.*, 2015), the Galapagos shark *Carcharhinus galapagensis* (Pazmiño *et al.*, 2017) or the grey reef shark *Carcharhinus amblyrhynchos* (Momigliano *et al.*, 2017).

Two main drivers of genetic diversity distribution in this small benthic shark were apparent in the data: geographical distance and historical demography. The influence of geographical distance was suggested by the significant global structure (positive spatial autocorrelation) detected with the sPCA (Fig. 4) and by generally increased genetic distance among increasingly distant locations (Fig. 2). The exceptions to this pattern are the marked genetic divergence of the Celtic Sea and S Portugal collections from their closest neighbours, resulting in geographically constrained genetic breaks south of the British Isles and off southwestern Iberia, respectively.

The genetic break located south of the British Isles separates a northern and a southern group of collections (Fig. 3). This separation was suggested in the sPCA, but is also apparent in other analyses performed: in the PCA, PC1 separates individuals from S Portugal and Celtic Sea (i.e. southern group) from those in Irish Sea (northern group), with North Sea and N Portugal individuals occupying intermediate positions (Fig. 1). Likewise, the NJ tree also shows the northern-most locations being the most divergent from those included in the southern group (S Portugal and Celtic Sea), with NW Portugal again having an intermediate position in the tree (Fig. 2). Another (stronger) genetic break was detected within

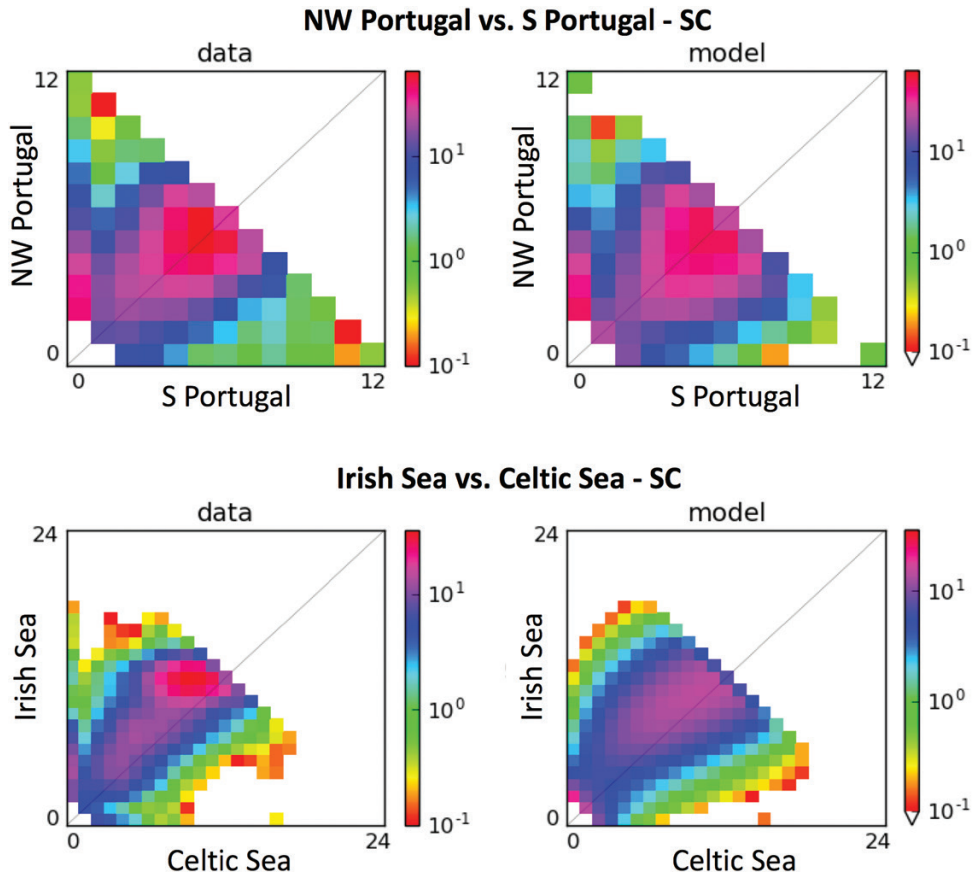

Figure 6. Joint allele frequency spectrum based on observed data (left panel) and on data obtained under the best-fitting demographic model, SC (secondary contact) using δaδ (right panel). Results are shown for the population pairs associated with the two genetic breaks observed in the spatial principal components analysis: NW Portugal vs. S Portugal (upper panel) and Irish Sea vs. Celtic Sea (lower panel). The coloured scale to the right of each graph shows the number of single nucleotide polymorphisms per pixel of minor allele counts in six and 12 individuals from each of the two population pairs as indicated above, respectively.

the southern group of *S. canicula* collections, separating S Portugal from the rest (Figs. 3, 4). The strong genetic divergence of S Portugal is particularly striking because the distribution of *S. canicula* is continuous throughout Atlantic Iberian waters, and no apparent physical (e.g. depth) or biogeographical barriers to individual dispersal and gene flow are thought to exist around the southwestern coast of the Iberian Peninsula.

The existence of such stark genetic breaks suggests levels of gene flow not conforming to a strict isolation-by-distance model, particularly those involving the Celtic Sea and S Portugal collections. Indeed, estimates of relative migration rate based on allele frequency data support the existence of restricted gene flow into and out of S Portugal, and into the Celtic Sea (albeit to a lower extent), regardless of the geographical proximity to other populations. In contrast, higher levels of bi-directional gene flow were estimated among the Irish Sea, North Sea and NW Portugal (Fig. 5).

The demographic inferences focusing on the population pairs across each of the two genetic breaks suggest that historical demographic events also contributed to the current pattern of genetic population structure of *S. canicula* along the NEA. Specifically, the JAFS simulated under the model of isolation followed by secondary contact provided the best fit for the observed JAFS in both population pairs across the reported genetic breaks. The models consistently showed that the period of isolation since the split from an ancestral population was much longer than the time since secondary contact in both population pairs analysed (Table 4). Hence, the genetic signal of such long-term divergence between spatially adjacent populations has not yet been erased by recent gene flow, leading to the observed deviations from an isolation-by-distance pattern of genetic structure in *S. canicula*.

A north–south split of genetic groups along the NEA has been found in other coastal benthic marine

taxa, including polychaete tubeworms (Jolly *et al.*, 2005, 2006), bivalves (Krakau *et al.*, 2012), gobies (Gysels *et al.*, 2004; Larmuseau *et al.*, 2009), flatfish (Hemmer-Hansen *et al.*, 2007; Diopere *et al.*, 2018) and the thornback ray *Raja clavata* Linnaeus, 1758 (Chevolot *et al.*, 2006). The geographical location of group separation varies among studies but generally separates populations around the British Isles from those off the Iberian Peninsula. This large-scale pattern of genetic diversity distribution has also been attributed to historical isolation of populations within distinct northern and southern glacial refugia during the last 10 000 yr up to the Last Glacial Maximum (Maggs *et al.*, 2008). However, samples from southern Iberia are generally lacking in genetic diversity studies of benthic taxa along the NEA; therefore, limited comparisons are available from the literature. Nonetheless, a marked genetic break between western and southern Iberian coastal waters has been found in the seaweed *Fucus vesiculosus* Linnaeus, 1753 (Assis *et al.*, 2014), the thornback ray *R. clavata* (Chevolot *et al.*, 2006), the European hake *Merluccius merluccius* (Linnaeus, 1758) (Lundy *et al.*, 1999; Castillo *et al.*, 2005) and the boarfish *Capros aper* (Linnaeus, 1758) (Farrell *et al.*, 2016). The differentiation of southern Iberian populations across taxa with distinct life strategies and ecological requirements is still poorly understood, but might involve a combination of range-edge effects on populations subjected to extreme climatic events (Assis *et al.*, 2014) and environmental factors exerting strong local selection even in the presence of gene flow among populations (e.g. Nielsen *et al.*, 2009, 2012). This represents an interesting area for future study in *S. canicula*, especially given that there are suggestions of morphological differences present in African populations closer to the southern range limit of the species (Litvinov, 2003).

Phylogeographical reconstruction

In addition to improving the spatial resolution of the genetic structure pattern, the data generated here also helped in elucidating the phylogeography of the target species. Specifically, the collections of *S. canicula* sampled along the NEA appear to have been derived from three putative ancestral populations, as suggested by TESS3 and δaδi analyses: two within the southern group and dominant in the S Portugal and Celtic Sea ancestry coefficients, respectively; and a third in the northern group of collections and dominant in the Irish Sea and North Sea ancestry coefficients.

The current spatial distribution of these ancestral populations generally matches previously proposed and generally accepted glacial refugia of benthic marine taxa in the NEA, namely the southern coast of the Iberian Peninsula, the southwest of Ireland and the western English Channel (reviewed by Maggs *et al.*, 2008), respectively. In species with limited dispersal, such as *S. canicula*, gene flow may be slow in breaking down the genetic differences accumulated during previous and long-lasting vicariant distributions. Indeed, the most likely model of demographic history retrieved for *S. canicula* along the NEA included long-term isolation followed by recent secondary contact (Table 4). Thus, the species might have sustained populations in these three putative refugia during glaciations.

The direction of recolonization of *S. canicula* from these putative glacial refugia can also be inferred from the data. The much lower migration rates estimated for S Portugal and the little evidence of admixture of its dominant cluster northwards strongly suggest that recolonization of the NEA from a putative glacial refugia in southern Iberia (or southwards) is unlikely. In turn, recolonization of the NEA by *S. canicula* might generally have occurred southwards, and possibly from distinct glacial refugia. This hypothesis is supported by the admixture of the Irish Sea and Celtic Sea clusters into the NW Portugal genetic make-up (Fig. 3), and by the relative increase in admixture percentages from north to south (excluding S Portugal).

Secondary contact and varying levels of admixture between northern and southern clades have also been found in many marine taxa along western Europe, either following an isolation-by-distance model and/or tracing the recolonization direction from glacial refugia (Chevolot *et al.*, 2006; Milano *et al.*, 2014; Vandamme *et al.*, 2014; Diopere *et al.*, 2018). However, the pattern described for *S. canicula* is opposite to the generally described northwards recolonization from southern glacial refugia (Hewitt, 2000; Maggs *et al.*, 2008), as observed in another benthic elasmobranch with similar distribution, the thornback ray *R. clavata* (Chevolot *et al.*, 2006). However, the limited contribution of southern refugia to recolonization of the NEA coast has also been proposed previously for some marine coastal taxa, such as the brown seaweed *Fucus serratus* Linnaeus, 1753 (Hoarau *et al.*, 2007), the cockle *Cerastoderma edule* (Linnaeus, 1758) (Krakau *et al.*, 2012) or the sand goby *Pomatoschistus minutus* (Pallas, 1770) (Larmuseau *et al.*, 2009).

Management and conservation implications

The spatial genetic structure of *S. canicula* uncovered here supports the management approach defined for the species in NEA waters, which assumes the existence of multiple local populations (ICES, 2015). However, the sharp genetic discontinuity and reduced gene flow found between S Portugal and the remaining northward collections of *S. canicula* suggest the species might have distinct stocks along the Portuguese coast.

Populations exhibiting low genetic diversity and a high degree of genetic differentiation within the species range, such as *S. canicula* of S Portugal, should be considered as discrete and be conserved as such (Stephenson, 1999). This result contrasts with the currently assumed single stock unit in the Portuguese coast (i.e. ICES area IXa; ICES, 2015), but further confirmation of this genetic break should be obtained using larger sample sizes and increased spatial resolution of sample collections along western Iberia.

These results have wider implications for fisheries management of other exploited elasmobranchs, because a similar pattern of multiple local populations might also exist for species of similar habit and habitat (e.g. many skate species). Nevertheless, species-specific studies should be conducted whenever possible to assess the spatial population structure adequately and delimit putative stocks correctly.

ACKNOWLEDGEMENTS

Funding from the Erasmus+ programme (European Commission) was granted to A.M. and is very much appreciated. The authors are thankful to L. Pascal, C. Gubili and to P. McCorriston for help in sample collection. Alan Le Moan is also acknowledged for his help with the δaδi software. The authors are thankful to the three anonymous reviewers for providing helpful comments to the manuscript, which have greatly improved its quality. A.V. was funded by Fundação para a Ciência e Tecnologia (SFRH/BPD/77487/2011). L.Z. was funded by University of Padova grant CPDA148387/14.

REFERENCES

Ahrens CW, **Rymer PD**, **Stow A**, **Bragg J**, **Dillon S**, **Umbers KDL**, **Dudaniec RY. 2018.** The search for loci under selection: trends, biases and progress. *Molecular Ecology* **27:** 1342–1356.

André C, **Larsson LC**, **Laikre L**, **Bekkevold D**, **Brigham J**, **Carvalho GR**, **Dahlgren TG**, **Hutchinson WF**, **Mariani S**, **Mudde K**, **Ruzzante DE**, **Ryman N. 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.

Andrews KR, **Good JM**, **Miller MR**, **Luikart G**, **Hohenlohe PA. 2016.** Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics* **17:** 81–92.

Andrews KR, **Hohenlohe PA**, **Miller MR**, **Hand BK**, **Seeb JE**, **Luikart G. 2014.** Trade-offs and utility of alternative RADseq methods: reply to Puritz *et al*. *Molecular Ecology* **23:** 5943–5946.

Archer FI, **Adams PE**, **Schneiders BB. 2017.** STRATAG: an R package for manipulating, summarizing and analysing population genetic data. *Molecular Ecology Resources* **17:** 5–11.

Assis J, **Serrão EA**, **Claro B**, **Perrin C**, **Pearson GA. 2014.** Climate-driven range shifts explain the distribution of extant gene pools and predict future loss of unique lineages in a marine brown alga. *Molecular Ecology* **23:** 2797–2810.

Baird NA, **Etter PD**, **Atwood TS**, **Currey MC**, **Shiver AL**, **Lewis ZA**, **Selker EU**, **Cresko WA**, **Johnson EA. 2008.** Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE* **3:** e3376.

Barbieri M, **Maltagliati F**, **Roldán MI**, **Castelli A. 2014.** Molecular contribution to stock identification in the small-spotted catshark, *Scyliorhinus canicula* (Chondrichthyes, Scyliorhinidae). *Fisheries Research* **154:** 11–16.

Benjamini Y, **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.

Cardoso PLNN. 2015. *Variability in life history and population structure in a model shark species, Scyliorhinus canicula (Linnaeus 1758)*. MSc. Thesis, Faculdade de Ciências da Universidade do Porto.

Castillo AGF, **Alvarez P**, **Garcia-Vazquez E. 2005.** Population structure of *Merluccius merluccius* along the Iberian Peninsula coast. *ICES Journal of Marine Science* **62:** 1699–1704.

Cavalli-Sforza LL, **Edwards AWF. 1967.** Phylogenetic analysis: models and estimation procedures. *Evolution* **21:** 550–570.

Caye K, **Deist TM**, **Martins H**, **Michel O**, **François O. 2016.** TESS3: fast inference of spatial population structure and genome scans for selection. *Molecular Ecology Resources* **16:** 540–548.

Chevolot M, **Hoarau G**, **Rijnsdorp AD**, **Stam WT**, **Olsen JL. 2006.** Phylogeography and population structure of thornback rays (*Raja clavata* L., Rajidae). *Molecular Ecology* **15:** 3693–3705.

Cunha RL, **Coscia I**, **Madeira C**, **Mariani S**, **Stefanni S**, **Castilho R. 2012.** Ancient divergence in the trans-oceanic deep-sea shark *Centroscymnus crepidater*. *PLoS ONE* **7:** e49196.

Danecek P, **Auton A**, **Abecasis G**, **Albers CA**, **Banks E**, **DePristo MA**, **Handsaker RE**, **Lunter G**, **Marth GT**, **Sherry ST**, **McVean G**, **Durbin R**; **1000 Genomes Project Analysis Group**. **2011.** The variant call format and VCFtools. *Bioinformatics* **27:** 2156–2158.

Diopere E, **Vandamme SG**, **Hablützel PI**, **Cariani A**, **Van Houdt J**, **Rijnsdorp A**, **Tinti F**, **FishPopTrace Consortium, Volckaert FAM, Maes GE**. **2018.** Seascape genetics of a flatfish reveals local selection under high levels of gene flow. *ICES Journal of Marine Science* **75:** 675–689.

Domingues RR, **Hilsdorf AWS**, **Gadig OBF. 2018.** The importance of considering genetic diversity in shark and ray conservation policies. *Conservation Genetics* **19:** 501–525.

Dudgeon CL, **Blower DC**, **Broderick D**, **Giles JL**, **Holmes BJ**, **Kashiwagi T**, **Krück NC**, **Morgan JA**, **Tillett BJ**, **Ovenden JR. 2012.** A review of the application of molecular genetics for fisheries management and conservation of sharks and rays. *Journal of Fish Biology* **80:** 1789–1843.

Durand E, **Jay F**, **Gaggiotti OE**, **François O. 2009.** Spatial inference of admixture proportions and secondary contact zones. *Molecular Biology and Evolution* **26:** 1963–1973.

Ebert D, **Fowler S**, **Compagno L. 2013.** *Sharks of the world: a fully illustrated guide*. Wild Nature Press, Plymouth, UK.

Edwards AW. 1971. Distances between populations on the basis of gene frequencies. *Biometrics* **27:** 873–881.

Ellis JR, **Shackley SE. 1997.** The reproductive biology of *Scyliorhinus canicula* in the Bristol Channel, U.K. *Journal of Fish Biology* **51:** 361–372.

Farrell ED, **Carlsson JEL**, **Carlsson J. 2016.** Next Gen Pop Gen: implementing a high-throughput approach to population genetics in boarfish (*Capros aper*). *Royal Society Open Science* **3:** 160651.

Feldheim KA, **Gruber SH**, **Dibattista JD**, **Babcock EA**, **Kessel ST**, **Hendry AP**, **Pikitch EK**, **Ashley MV**, **Chapman DD. 2014.** Two decades of genetic profiling yields first evidence of natal philopatry and long-term fidelity to parturition sites in sharks. *Molecular Ecology* **23:** 110–117.

Griffiths AM, **Sims DW**, **Cotterell SP**, **El Nagar A**, **Ellis JR**, **Lynghammar A**, **McHugh M**, **Neat FC**, **Pade NG**, **Queiroz N**, **Serra-Pereira B**, **Rapp T**, **Wearmouth VJ**, **Genner MJ. 2010.** Molecular markers reveal spatially segregated cryptic species in a critically endangered fish, the common skate (*Dipturus batis*). *Proceedings of the Royal Society B: Biological Sciences* **277:** 1497–1503.

Griffiths AM, **Sims DW**, **Johnson A**, **Lynghammar A**, **McHugh M**, **Bakken T**, **Genner MJ. 2011.** Levels of connectivity between longnose skate (*Dipturus oxyrinchus*) in the Mediterranean Sea and the north-eastern Atlantic Ocean. *Conservation Genetics* **12:** 577–582.

Gubili C, **Macleod K**, **Perry W**, **Hanel P**, **Batzakas I**, **Farrell ED**, **Lynghammar A**, **Mancusi C**, **Mariani S**, **Menezes GM**, **Neat F**, **Scarcella G**, **Griffiths MJ. 2016.** Connectivity in the deep: phylogeography of the velvet belly lanternshark. *Deep-Sea Research Part I: Oceanographic Research Papers* **115:** 233–239.

Gubili C, **Sims DW**, **Veríssimo A**, **Domenici P**, **Ellis J**, **Grigoriou P**, **Johnson AF**, **McHugh M**, **Neat F**, **Satta A**, **Scarcella G**, **Serra-Pereira B**, **Soldo A**, **Genner MJ**, **Griffiths AM. 2014.** A tale of two seas: contrasting patterns of population structure in the small-spotted catshark across Europe. *Royal Society Open Science* **1:** 140175–140175.

Gutenkunst RN, **Hernandez RD**, **Williamson SH**, **Bustamante CD. 2009.** Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genetics* **5:** e1000695.

Gysels ES, **Hellemans B**, **Pampoulie C**, **Volckaert FA. 2004.** Phylogeography of the common goby, Pomatoschistus microps, with particular emphasis on the colonization of the Mediterranean and the North Sea. *Molecular Ecology* **13:** 403–417.

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

Hemmer-Hansen J, **Therkildsen NO**, **Pujolar JM. 2014.** Population genomics of marine fishes: next-generation prospects and challenges. *The Biological Bulletin* **227:** 117–132.

Hewitt G. 2000. The genetic legacy of the Quaternary ice ages. *Nature* **405:** 907–913.

Hoarau G, **Coyer JA**, **Veldsink JH**, **Stam WT**, **Olsen JL. 2007.** Glacial refugia and recolonization pathways in the brown seaweed *Fucus serratus*. *Molecular Ecology* **16:** 3606–3616.

ICES. **2015.** *Report of the Working Group on Elasmobranch Fishes (WGEF), 17–23 June 2015, Lisbon, Portugal. ICES CM 2015/ACOM:19*. Lisbon, Portugal, 711.

Jolly MT, **Jollivet D**, **Gentil F**, **Thiébaut E**, **Viard F. 2005.** Sharp genetic break between Atlantic and English Channel populations of the polychaete *Pectinaria koreni*, along the North coast of France. *Heredity* **94:** 23–32.

Jolly MT, **Viard F**, **Gentil F**, **Thiébaut E**, **Jollivet D. 2006.** Comparative phylogeography of two coastal polychaete tubeworms in the Northeast Atlantic supports shared history and vicariant events. *Molecular Ecology* **15:** 1841–1855.

Jombart T, **Ahmed I. 2011.** *adegenet 1.3-1*: new tools for the analysis of genome-wide SNP data. *Bioinformatics* **27:** 3070–3071.

Jost L. 2008. G_{ST} and its relatives do not measure differentiation. *Molecular Ecology* **17:** 4015–4026.

Kamvar ZN, **Brooks JC**, **Grünwald NJ. 2015.** Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Frontiers in Genetics* **6:** 208.

Kamvar ZN, **Tabima JF**, **Grünwald NJ. 2014.** *Poppr*: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* **2:** e281.

Keenan K, **McGinnity P**, **Cross TF**, **Crozier WW**, **Prodöhl PA. 2013.** DiveRsity: an R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution* **4:** 782–788.

Kousteni V, **Kasapidis P**, **Kotoulas G**, **Megalofonou P. 2015.** Strong population genetic structure and contrasting demographic histories for the small-spotted catshark (*Scyliorhinus canicula*) in the Mediterranean Sea. *Heredity* **114:** 333–43.

Krakau M, **Jacobsen S**, **Jensen KT**, **Reise K**, **2012.** The cockle *Cerastoderma edule* at Northeast Atlantic shores: genetic signatures of glacial refugia. *Marine Biology* **159:** 221–230.

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

Le Moan A, **Gagnaire PA**, **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.

Litvinov FF. **2003.** Sexual dimorphism as an index of the isolation of West African populations of the cat shark Scyliorhinus canicula. *Journal of Ichthyology* **43:** 81–85.

Luikart G, **England PR**, **Tallmon D**, **Jordan S**, **Taberlet P. 2003.** The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews Genetics* **4:** 981–994.

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. *Molecular Ecology* **8:** 1889–1898.

Maggs CA, **Castilho R**, **Foltz D**, **Henzler C**, **Jolly MT**, **Kelly J**, **Olsen J**, **Perez KE**, **Stam W**, **Väinölä R**, **Viard F**, **Wares J. 2008.** Evaluating signatures of glacial refugia for North Atlantic marine organisms. *Ecology* **89:** S108–S122.

Milano I, **Babbucci M**, **Cariani A**, **Atanassova M**, **Bekkevold D**, **Carvalho GR**, **Espiñeira M**, **Fiorentino F**, **Garofalo G**, **Geffen AJ**, **Hansen JH**, **Helyar SJ**, **Nielsen EE**, **Ogden R**, **Patarnello T**, **Stagioni M**, **FishPopTrace Consortium**, Tinti F, Bargelloni L. **2014.** Outlier SNP markers reveal fine-scale genetic structuring across European hake populations (*Merluccius merluccius*). *Molecular Ecology* **23:** 118–135.

Miller MR, **Dunham JP**, **Amores A**, **Cresko WA**, **Johnson EA. 2007.** Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Research* **17:** 240–248.

Momigliano P, **Harcourt R**, **Robbins WD**, **Jaiteh V**, **Mahardika GN**, **Sembiring A**, **Stow A. 2017.** Genetic structure and signatures of selection in grey reef sharks (*Carcharhinus amblyrhynchos*). *Heredity* **119:** 142–153.

Morin PA, **Luikart G**, **Wayne RK. 2004.** SNPs in ecology, evolution and conservation. *Trends in Ecology & Evolution* **19:** 208–216.

Nielsen EE, **Cariani A**, **Mac Aoidh E**, **Maes GE**, **Milano I**, **Ogden R**, **Taylor M**, **Hemmer-Hansen J**, **Babbucci M**, **Bargelloni L**, **Bekkevold D**, **Diopere E**, **Grenfell L**, **Helyar S**, **Limborg MT**, **Martinsohn JT**, **McEwing R**, **Panitz F**, **Patarnello T**, **Tinti F**, **Van Houdt JK**, **Volckaert FA**, **Waples RS**; **FishPopTrace consortium**, **Carvalho GR. 2012.** Gene-associated markers provide tools for tackling illegal fishing and false eco-certification. *Nature Communications* **3:** 851.

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

Paradis E, **Claude J**, **Strimmer K. 2004.** APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20:** 289–290.

Paterno M, **Schiavina M**, **Aglieri G**, **Ben Souissi J**, **Boscari E**, **Casagrandi R**, **Chassanite A**, **Chiantore M**, **Congiu L**, **Guarnieri G**, **Kruschel C**, **Macic V**, **Marino IAM**, **Papetti C**, **Patarnello T**, **Zane L**, **Melià P. 2017.** Population genomics meet Lagrangian simulations: oceanographic patterns and long larval duration ensure connectivity among *Paracentrotus lividus* populations in the Adriatic and Ionian seas. *Ecology and Evolution* **7:** 2463–2479.

Pazmiño DA, **Maes GE**, **Simpfendorfer CA**, **Salinas-de-León P**, **van Herwerden L. 2017.** Genome-wide SNPs reveal low effective population size within confined management units of the highly vagile Galapagos shark (*Carcharhinus galapagensis*). *Conservation Genetics* **18:** 1151–1163.

Popescu AA, **Huber KT**, **Paradis E. 2012.** Ape 3.0: new tools for distance-based phylogenetics and evolutionary analysis in R. *Bioinformatics* **28:** 1536–1537.

Portnoy DS, **Heist EJ. 2012.** Molecular markers: progress and prospects for understanding reproductive ecology in elasmobranchs. *Journal of Fish Biology* **80:** 1120–1140.

Portnoy DS, **Puritz JB**, **Hollenbeck CM**, **Gelsleichter J**, **Chapman D**, **Gold JR. 2015.** Selection and sex-biased dispersal in a coastal shark: the influence of philopatry on adaptive variation. *Molecular Ecology* **24:** 5877–5885.

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

Puritz JB, **Hollenbeck CM**, **Gold JR. 2014.** *dDocent*: a RADseq, variant-calling pipeline designed for population genomics of non-model organisms. *PeerJ* **2:** e431.

Ramírez-Amaro S, **Picornell A**, **Arenas M**, **Castro JA**, **Massutí E**, **Ramon MM**, **Terrasa B. 2018.** Contrasting evolutionary patterns in populations of demersal sharks throughout the western Mediterranean. *Marine Biology* **165:** 3.

Rodríguez-Cabello C, **Sánchez F**, **Fernández A**, **Olaso I. 2004.** Is the lesser spotted dogfish (*Scyliorhinus canicula*) population from the Cantabrian Sea a unique stock? *Fisheries Research* **69:** 57–71.

Rodríguez-Cabello C, **Velasco F**, **Olaso I. 1998.** Reproductive biology of lesser spotted dogfish *Scyliorhinus canicula* (L., 1758) in the Cantabrian Sea. *Scientia Marina (Barcelona)* **62:** 187–191.

Schmieder R, **Edwards R. 2011.** Quality control and preprocessing of metagenomic datasets. *Bioinformatics* **27:** 863–864.

Sims D, **Nash J**, **Morritt D. 2001.** Movements and activity of male and female dogfish in a tidal sea lough: alternative behavioural strategies and apparent sexual segregation. *Marine Biology* **139:** 1165–1175.

Stephenson RL. 1999. Stock complexity in fisheries management: a perspective of emerging issues related to population sub-units. *Fisheries Research* **43:** 247–249.

Sundqvist L, **Keenan K**, **Zackrisson M**, **Prodöhl P**, **Kleinhans D 2016.** Directional genetic differentiation and relative migration. *Ecology and Evolution* **6:** 3461–3475.

Tine M, **Kuhl H**, **Gagnaire PA**, **Louro B**, **Desmarais E**, **Martins RS**, **Hecht J**, **Knaust F**, **Belkhir K**, **Klages S**, **Dieterich R**, **Stueber K**, **Piferrer F**, **Guinand B**, **Bierne N**, **Volckaert FA**, **Bargelloni L**, **Power DM**, **Bonhomme F**, **Canario AV**, **Reinhardt R. 2014.** European sea bass genome and its variation provide insights into adaptation to euryhalinity and speciation. *Nature Communications* **5:** 5570.

Vandamme SG, **Maes GE**, **Raeymaekers JAM**, **Cottenie K**, **Imsland AK**, **Hellemans B**, **Lacroix G**, **Mac Aoidh E**, **Martinsohn JT**, **Martínez P**, **Robbens J**, **Vilas R**, **Volckaert FA. 2014.** Regional environmental pressure influences population differentiation in turbot (*Scophthalmus maximus*). *Molecular Ecology* **23:** 618–636.

Veríssimo A, **McDowell JR**, **Graves JE. 2010.** Global population structure of the spiny dogfish *Squalus acanthias*, a temperate shark with an antitropical distribution. *Molecular Ecology* **19:** 1651–1662.

Veríssimo A, **McDowell JR**, **Graves JE. 2011.** Population structure of a deep-water squaloid shark, the Portuguese dogfish (*Centroscymnus coelolepis*). *ICES Journal of Marine Science* **68:** 555–563.

Veríssimo A, **McDowell JR**, **Graves JE. 2012.** Genetic population structure and connectivity in a commercially exploited and wide-ranging deepwater shark, the leafscale gulper (*Centrophorus squamosus*). *Marine and Freshwater Research* **63:** 505–512.

Wang S, **Meyer E**, **McKay JK**, **Matz MV. 2012.** 2b-RAD: a simple and flexible method for genome-wide genotyping. *Nature Methods* **9:** 808–810.

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

Willing EM, **Dreyer C**, **van Oosterhout C. 2012.** Estimates of genetic differentiation measured by F_{ST} do not necessarily require large sample sizes when using many SNP markers. *PLoS ONE* **7:** e42649.

Wyffels J, **King BL**, **Vincent J**, **Chen C**, **Wu CH**, **Polson SW. 2014.** SkateBase, an elasmobranch genome project and collection of molecular resources for chondrichthyan fishes. *F1000Research* **3:** 191.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Figure S1. Read depth per locus for the set of 4656 single nucleotide polymorphism (SNP) loci filtered with VCFtools, for the 88 individuals of *Scyliorhinus canicula* sampled.

Figure S2. Principal components analysis of multilocus single nucleotide polymorphism (SNP) genotypes of the lesser spotted dogfish *Scyliorhinus canicula*, considering: A, 4656 SNP loci and 88 individuals [filtered dataset obtained from VCFtools; principal component (PC)1 6.3%, PC2 3.3%]; B, 2674 SNP loci and 71 individuals, considering 20% missing data at the locus and individual levels (PC1 7.3%, PC2 4.3%); and C, 1243 SNP loci and 79 individuals, considering 5% missing data at the locus and individual levels (PC1 12.2%, PC2 4.6%). Legend in the top graph is the same for all graphs.

Figure S3. Plot of cross-validation error values, based on five replicate runs per K value, and 5% (A) and 10% (B) of masked data in the test set, using 2674 single nucleotide polymorphism (SNP) loci. The most likely K corresponds to the lowest value where the first plateau of error value is reached, and corresponds to K = 3 in the graphs.

Figure S4. Eigenvalues of the spatial principal components analysis based on 2674 single nucleotide polymorphisms (SNPs) and 71 individuals of the lesser spotted dogfish *Scyliorhinus canicula*, along the northeastern Atlantic coast. See Material amd methods and Results sections for details.

Figure S5. Spatial and variance components of the eigenvalues from the spatial principal components analysis based on 2674 single nucleotide polymorphisms (SNPs) and 71 individuals of the lesser spotted dogfish *Scyliorhinus canicula*, along the northeastern Atlantic coast.

Table S1. Average values of individual ancestry coefficients for the different sample collections of Scyliorhinus canicula along the northeastern Atlantic.

Table S2. Relative migration rates among sample collections based on 2674 single nucleotide polymorphism (SNP) loci, estimated based on Jost's *D* genetic differentiation measure. Values in bold showed statistically significant asymmetric gene flow. Source populations are in rows; receiving populations are in columns.

Table S3. Relative migration rates among sample collections based on 2674 single nucleotide polymorphism (SNP) loci, estimated based on the G_{ST} genetic differentiation measure. Values in bold showed statistically significant asymmetric gene flow. Source populations are in rows; receiving populations are in columns. Note that estimates based on G_{ST} are identical to those based on *Nm*.

Table S4. Relative migration rates among sample collections based on 2674 single nucleotide polymorphism (SNP) loci, estimated based on *Nm* estimate. Values in bold showed statistically significant asymmetric gene flow. Source populations are in rows; receiving populations are in columns. Note that estimates based on *Nm* are identical to those based on G_{ST}.

Supplemental document S1 - Detailed 2b-RAD protocol used for de novo SNP calling and genotyping in Scyliorhinus canicula.