Stepping up to genome scan allows stock differentiation in the worldwide distributed blue shark *Prionace glauca*

Natacha Nikolic\*1,2,3, Floriaan Devloo-Delva\*4,5, Diane Bailleul1,6, Ekaterina Noskova7, Clément Rougeux8, Chrystelle Delord1, Philippe Borsa9, Cathy Liautard-Haag1, Mohamad Hassan1,10, Amandine D. Marie3, Pierre Feutry4, Peter Grewe4, Campbell Davies4, Jessica Farley4, Daniel Fernando11, Sebastian Biton-Porsmoguer12, François Poisson1, Denham Parker13,14, Agostino Leone1, Jorden Aulich4, Matt Lansdell4, Francis Marsac1, Sophie Arnaud-Haond1

(\*) equal contribution

1. UMR MARBEC, Univ. Montpellier, IRD, Ifremer, CNRS, Sète, France
2. INRAE, Ecobiop, AQUA, St-Pée-sur-Nivelle, France
3. ARBRE, Agence de Recherche pour la Biodiversité à la Réunion, Saint-Gilles, La Réunion
4. CSIRO Oceans and Atmosphere, Hobart, TAS, Australia
5. School of Natural Sciences – Quantitative Marine Science, University of Tasmania, Hobart, 7001 TAS, Australia
6. INRAE, Castanet Tolosan, France
7. Computer Technologies Laboratory, ITMO University. St Petersburg, Russia
8. University of Calgary, Calgary, Canada
9. Institut de recherche pour le développement, UMR ENTROPIE, Montpellier, France
10. Animal Production Department, Faculty of Agriculture, Tishreen University, Syria
11. Blue Resources Trust, Colombo 00700, Sri Lanka
12. French Biodiversity Agency (OFB), Channel and North Sea Delegation, 4 rue du Colonel Fabien, BP14, 76083 Le Havre, France
13. Department of Forestry, Fisheries and the Environment, (DFFE), Cape Town, South Africa
14. Department of Biological Sciences, University of Cape Town, Rondebosch 7700, South Africa

Corresponding author: natachanikolic@hotmail.com

**Abstract**

The blue shark *Prionace glauca* is a top predator with one of the widest geographic distributions of any shark species, yet classified as critically endangered in the Mediterranean Sea, and Near Threatened globally. Previous genetic studies did not reject the null hypothesis of a single global population across the worldwide species range. Blue shark situation was proposed as a possible archetype of the ‘grey zone of population differentiation’, coined to designate cases where population structure may be too recent or too faint to be detected using a limited set of markers. Here, blue shark samples collected throughout its global range were sequenced using a specific ddRAD method (DArTseq; Georges et al. 2018), which recovered 37,655 genome-wide single nucleotide polymorphisms (SNPs). Two main groups emerged, with Mediterranean Sea and Northern Atlantic samples significantly differentiated from the Indo-west Pacific samples. Significant pairwise *F*ST values indicated further genetic differentiation within the Atlantic Ocean, and between the Atlantic Ocean and the Mediterranean Sea. Reconstruction of recent demographic history suggested the divergence between northern and southern oceanic populations emerged about 500 generations ago and revealed a drastic reduction in effective population size from a large ancestral population. Our results illustrate the power of high-density genome scans to detect population structure and reconstruct demographic history in highly migratory marine species. As the management of the blue shark fishery, either as target or as bycatch, does not account for this delineation, we strongly recommend that the results presented here be considered in future stock assessment and management plans.

**Keywords:**

Population genetics, SNP, bycatch, pelagic, stock assessment

**Introduction**

The blue shark *Prionace glauca* is widely distributed worldwide and considered as the most abundant pelagic shark (Compagno 1984; Nakano and Seki 2003). Inhabiting all oceans except polar regions (i.e., from 60°N to 50°S), its habitat extends from the surface to > 1100 m depth (Queiroz et al. 2012). Like other top predators (Estes et al. 2011; Hughes et al. 2013), the blue shark contributes to the top-down regulation of marine communities, and thus to the stability and persistence of the numerous marine ecosystems it belongs to across its vast distribution (e.g., Hernández-Aguilar et al. 2015).

The blue shark is the most frequently fished shark species worldwide (Baum and Blanchard 2010; Biton-Porsmoguer 2018a; Campana et al. 2006) and, together with shortfin mako and pelagic thresher shark, it is one of the three shark species regularly caught by pelagic fisheries (Megalofonou et al. 2005a). Although occasionally targeted for its meat on the western coast of Baja California Sur (Galván-Magana et al. 2019) and by recreational fisheries (Campana et al. 2006; Mejuto and Garcia-Cortés 2005; Panayiotou et al. 2020), the blue shark mostly is a bycatch of tuna and swordfish longline fisheries (Biton-Porsmoguer 2017, 2018b; Carvalho et al. 2015; Coelho et al. 2017). The bycatch mortality rate on pelagic longline fishing operations can be high, up to 35%, with an additional post-release mortality at 19% (Campana et al. 2009). The blue shark is also the main species involved in the shark fin trade (Clarke et al. 2006), for which the rest of the body of the shark is most often discarded because of the meat’s lower commercial value. As a result, and despite not being particularly targeted for its meat, with an estimate of 10 to 20 million individuals caught annually (Megalofonou et al. 2005b; Stevens 2009), the blue shark has been evaluated as near-threatened worldwide (Stevens 2009; Rigby et al. 2019) and critically endangered in the Mediterranean Sea (Sims et al. 2016).

Evidence of a 90% decline in blue shark abundance has been reported in the Mediterranean (Ferretti et al. 2008). Population declines have also been reported in both the Atlantic (Aires-da-Silva et al. 2008, 2009; Simpfendorfer et al. 2002; Baum and Blanchard 2010) and the Pacific (Clarke et al. 2013) oceans, although other studies indicate that the catch per unit of effort might be stable, or even slightly increasing in the same areas (Matsunaga and Nakano 1999; Nakano and Clarke 2005; Nakano and Stevens 2008; ISC 2018). These inconsistencies may point to a poor understanding of the ecology and the biology of the blue shark, or to difficulties in estimating its abundance. Intergovernmental bodies have acknowledged the need to fill important ecological and biological data gaps to improve stock assessments in the blue shark, in turn used to provide advice on fisheries management (ISC 2018; IOTC 2017). The International Commission for the Conservation of Atlantic Tunas (ICCAT) acknowledged the steady increase of blue shark catches in recent years and the large level of uncertainty in the data inputs and model assumptions about the stock and the fishery. In 2019, the ICCAT adopted recommendations to further limit blue shark catches and encourage further research to provide essential knowledge needed for its long-term management, including life-history parameters and migration (ICCAT 2019).

Considering the blue shark’s nomadic behaviour and its wide distribution (Stevens 1990), stock assessments rely on the assumption that stocks in the northern and southern Atlantic (ICCAT 2015), in the northern and southern Pacific (ISC 2018) and in the Indian Ocean (IOTC 2017) are regionally homogeneous. Electronic tags have confirmed that blue sharks swim over large distances, even crossing from one ocean to another (Maxwell et al. 2019; Vandeperre et al. 2014; Kohler et al. 2002; Queiroz et al. 2012; da Silva et al. 2010). However, non-overlapping reproductive cycles have been reported for the Northern and Southern Hemispheres (Nakano and Seki 2003; Nakano and Stevens 2008) and trans-equatorial migration is assumed to be limited (Kohler and Turner 2008).

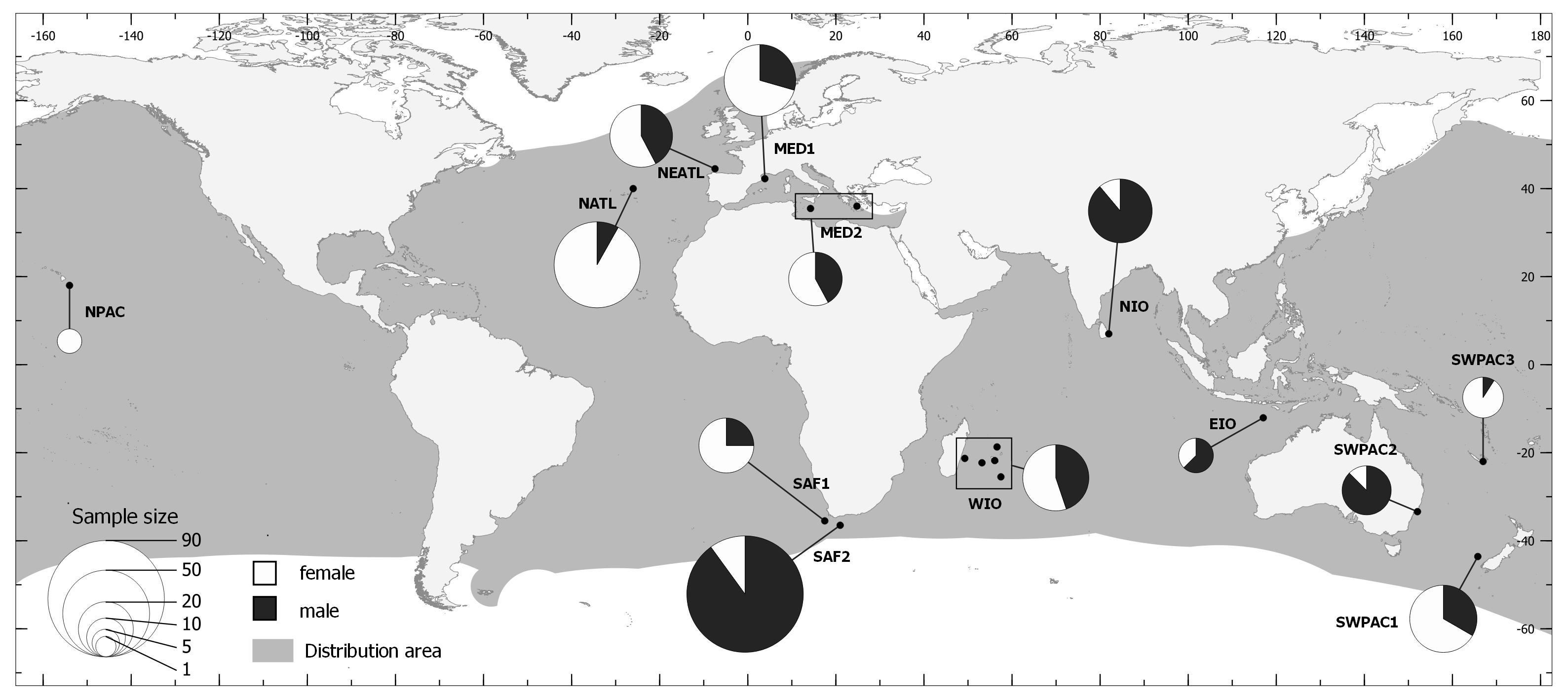
Based on mitochondrial DNA or/and microsatellite markers, no consistent pattern of genetic differentiation has been detected, even between the Northern and Southern Hemispheres (Bitencourt et al. 2019; King et al. 2015; Li et al. 2017; Taguchi et al. 2015; Verissimo et al. 2017). Only weak differentiation of the Mediterranean Sea population (Bailleul et al. 2018; Leone et al. 2017), and among the eastern South Pacific and Western Australian population (Taguchi et al. 2015) was detected. Traditional genetic approaches only detect extreme restriction to genetic exchange, far below the threshold of demographic independence (Waples 1998; Waples and Gaggiotti 2006), and they integrate migratory exchange over a number of generations that is proportional to effective population size (Hedgecock et al. 2007). Effective population size in marine species can be extremely large, a situation that may apply to the blue shark, considering its wide distribution range and relatively high population density. In fact, the blue shark has been used as a case species to illustrate the concept of ‘population grey zone’ (Bailleul et al. 2018) which refers to the often-inconclusive results on population genetic differentiation studies designed to help define management units in pelagic species. The ‘population grey zone’ effect describes the potentially very long time-lag (hundreds to thousands of generations) between the demographic split of a population into two independent demographic entities and its translation into a signal strong enough to be detectable from a handful of molecular markers (Bailleul et al. 2018). Genome scan methods such as RAD Sequencing (Davey and Blaxter 2010) or Diversity Arrays Technology sequencing (DArTseqTM; DArT Pty Ltd, Canberra: Georges et al. 2018) provide a much denser array of markers (typically thousands of SNPs - single nucleotide polymorphisms) along the genome, increasing our ability to detect much more subtle levels of genetic differentiation. The recent use of genome scan approaches in large pelagic species including yellowfin tuna (Grewe et al 2015; Pecoraro et al. 2018; Mullins et al. 2018), striped marlin (Mamoozadeh et al. 2020), spotted dolphin (Leslie and Morin 2016), and Atlantic mackerel (Rodríguez-Ezpeleta et al. 2016) revealed genetic structure where traditional genetic markers had initially detected none or very little, and sometimes incongruent, patterns of differentiation.

Here, we used the DArTseqTM sequencing approach on blue shark samples collected from the Atlantic, Pacific and Indian Oceans, as well as from the Western Mediterranean Sea, to test the null hypothesis of large-scale panmixia. Preliminary results on basic population differentiation have been presented previously (Nikolic et al. 2020a). Here, we aimed at a comprehensive and deeper analysis of blue shark population genetic structure and we inferred for the first time its demographic history. We take advantage of the power offered by genome scan analysis to provide a genetically-based delineation of distinct management units, and to investigate the present and past demographic history until the most recent decades marked by increasing fishing pressure on blue shark populations worldwide (IOTC 2017; Nakano and Seki 2003).

**Materials and Methods**

* 1. ***Sampling***

A total of 376 individual blue sharks were sampled between 2009 and 2018, and 364 individuals have been successfully genotyped (29 individuals sampled in 2009, 8 in 2010, 4 in 2011, 24 in 2012, 56 in 2013, 39 in 2014, 81 in 2015, 3 in 2017, and 120 in 2018) across several geographic regions throughout the species range (Figure 1). All were caught by longline as bycatch, except the samples from the eastern Indian Ocean where a purse seine has operated. Only dead individuals were sampled. Individual length (cm) and phenotypic sex (based on the presence or absence of claspers), geographic location (latitude and longitude), and details on operating vessel were usually recorded - except for 93 individuals missing length data and 149 individuals whose sex was not identified. Whenever applicable, curved fork length, precaudal length, and interdorsal (space on dorsal surface between first and second dorsal fins) were converted to fork length (FL) based on the equations of Cramer, Bertolino and Scott (1997). A small piece of fin, skin or muscle of dead individuals was preserved in 96% ethanol at all sampling locations except the eastern Indian Ocean where tissue samples were dried in silica gel, and Reunion Island tissues were preserved in RNAlater solution (Qiagen, Hilden, Germany).

 **Figure 1.** Location of blue shark, *Prionace glauca,* sample sites across several geographic regions (364 genotyped individuals). Circle area is proportional to the number of collected samples. Black and white colors correspond to genetic sex determination (biological sex determination was corrected for 18 individuals and completed for 149 individuals by genetic sex determination on 364 individuals. Geographic regions are Mediterranean (MED: MED1, n=34 individuals, and MED2, n=20), Northern Atlantic (NATL, n=49), Northeast Atlantic (NEATL, n=26), South Africa (SAF: SAF1, n=21, and SAF2 n=89), Eastern Indian Ocean (EIO, n=8), Northern Indian Ocean (NIO, n=27), Western Indian Ocean (WIO, n=29), Southwest Pacific (SWPAC: SWPAC1, n=30, SWPAC2, n=16, and SWPAC3, n=11) and Northern Pacific (NPAC, n=4). *Shaded*: species distribution area, from https://www.iucnredlist.org/species/39381/2915850 and Compagno (1984).

* 1. ***DArT library preparation and sequencing***

Genomic DNA was extracted from 15 mg of muscle tissue subsampled from individual biopsies on an Eppendorf EP motion 5057 liquid robotic handler using a modification of the QIAamp® 96 DNA QIAcube HT Kit (QIAGEN, Hilden, Germany). This extraction included a lysis step in the presence of proteinase K followed by bind-wash-elute QIAGEN processing. Low-quality or degraded samples were extracted using the modified CTAB method of Grewe et al. (1993).

Genomic DNA was processed for the construction of a reduced representation library, sequenced, and genotyped by using the DArTseqTM technique. DNA sample libraries were created through digestion and ligation reactions using two methylation-sensitive restriction enzymes, *PstI* and *SphI*. The *PstI* site was compatible with a forward adapter that included a flow cell (Illumina, San Diego) attachment sequence and a sequencing primer sequence incorporating a staggered, barcode region of varying length. The *SphI* digestion generated a compatible overhang sequence that was ligated to a reverse adapter containing a flow cell attachment region and a reverse-priming sequence. Only mixed *PstI*-*SphI* restriction fragments were amplified by PCR. PCR conditions consisted of an initial denaturation at 94°C for 1 min followed by 30 cycles of 94°C for 20 sec, 58°C for 30 sec and 72°C for 45 sec, with a final extension step at 72°C for 7 min. After PCR, equimolar amounts of amplification products from each sample of the 96-well microtiter plate were bulked and subjected to cBot bridge PCR (Illumina), followed by sequencing on an Illumina Hiseq2000 automated sequencing system. The sequencing (single-end) was run for 77 cycles. The DNA fragments selected by this process were ca. 75 bp long. Further details on this sequencing method can be found in Sansaloni et al. (2011), Kilian et al. (2012), and Georges et al. (2018).

***2.1. SNP genotyping process***

For initial assessment of read quality and sequence representation, raw reads (raw data, Figure 2) were processed using the Illumina CASAVA v.1.8.2 software. Genotypes were generated from sequencing runs completed at DArT using a proprietary DArTseq analytical pipeline (DArT-Soft14 version). The DArT toolbox was then used to filter away poor-quality sequences, apply more stringent selection criteria to the barcode region, and generate the final genotypes (Kilian et al. 2012), here called one-row data (Figure 2). Details of the bioinformatic steps leading to SNP genotyping can be found in Georges et al. (2018).

***2.2. SNP filtering***

Multiple steps were performed in the data analysis, each using the appropriate software. Details of the scripts for SNP filtering and data analyses are presented in an R Markdown report (https://cloud.r-project.org/package=rmarkdown; Allaire et al. 2020) as Supplementary material (S1). Here, we provide an overview of the process followed for SNP filtering, which was performed with individual samples grouped per geographic region (Filtration 1) or per sample sites (Filtration 2).

First, metadata with all information on the individuals including sample location, sex and length, and one-row data with both SNP and reference alleles where a zero (0) score denotes a homozygote, one (1) a homozygote, and two (2) a heterozygote, were merged to built the dataset called DATA 1 (Figure 2). Then, SNPs from DATA 1 were filtered for low reproducibility based on technical replicate libraries, monomorphism across all individuals, low minor allele count, low coverage per locus, high missing data per SNP, short-linkage disequilibrium by keeping one SNP per locus, and SNPs out of Hardy-Weinberg equilibrium (HWE) using the *radiator* package v1.2.0 (Gosselin 2018; Gosselin et al. 2020) in R v4.1.0 (R Development Core Team 2020) (see details in Supplementary material S1). Individuals were filtered based on high missing data, high heterozygosity and duplicate individuals. At this stage, the dataset is called DATA 2 (Figure 2).

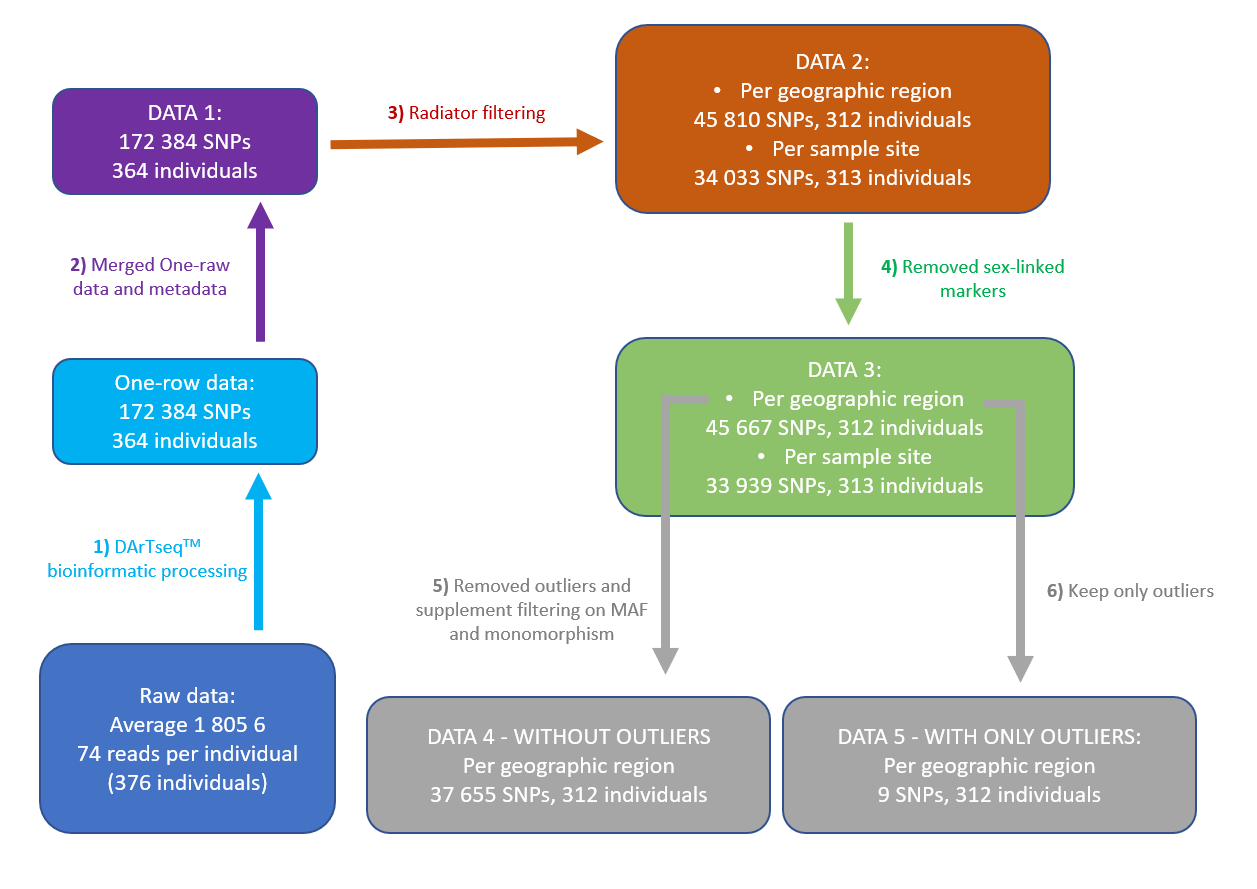
Sex-linked marker identification:

The unfiltered data (DATA 1) were tested for the presence of sex-linked markers using the *sexy\_markers* function in the *radiator* package (Gosselin et al. 2020). To eliminate erroneous detections, the DATA 1 dataset was filtered for high individual missing data and heterozygosity, as well as monomorphic markers and short-distance linkage of SNPs. Subsequently, we identified markers on the Y or W chromosomes by their presence in one sex, but absence in the other. Similarly, X- or Z-linked markers were detected based on the heterozygosity and coverage patterns between sexes. All identified sex-linked markers were eliminated from the DATA 2 dataset to obtain a modified dataset (DATA 3) containing all other loci (Figure 2, see Supplementary material S1 for details on sex-linked analysis). Markers on the Y or W chromosomes were used to assign the genetic sex to individuals with missing phenotypic sex.

Outlier identification:

Two complementary algorithms were used to identify putative outliers: *PCAdapt* v4.3.3 (Luu et al. 2017) and *OutFLANK* v0.2 (Whitlock et al. 2015). We removed the outliers detected by both *PCAdapt* and *OutFLANK* from DATA 3 with individual samples grouped per geographic region. The last step of filtering consisted of removing SNP loci with overall allele frequency lower than 0.01 (“minor” allele frequency; MAF) using dartR v1.9.9 (Georges et al. 2018). Thus, we finally obtained the DATA 4 dataset where we kept only non-outlier loci, and the DATA 5 dataset where we kept only outlier loci (Figure 2, Supplementary material S1).

All downstream analyses were run on DATA 3 with all loci including outliers (with individual samples grouped per geographic region or per sample sites), and on DATA 4 and DATA 5 with only non-outlier and only outlier loci, respectively (with individual samples grouped per geographic region) (Figure 2, see Results section and Supplementary material S1).



**Figure 2.** Summary of the genetic data filtering process for *Prionace glauca* from the Raw data (DArT-sequencing read output) to the final genotype datasets (DATA 3, DATA 4 and DATA 5). In step 5, minor allele frequency is abbreviated as MAF.

* 1. ***Population genetic analysis***

All group stratification from geographic regions (Figure 1), as Mediterranean merged and not, was tested in population genetic analysis (see Supplementary material S1).Indices of genetic diversity including heterozygosity, allelic richness (Ar), *F*IS, and their probabilities under the null hypothesis of HWE were estimated using the *diveRsity* package v1.9.90 (Keenan et al. 2013).

Pairwise *F*ST, and average sample-pairwise differences together with their significance estimated by bootstrap resampling were obtained using the *strataG* v2.5.01 (Archer et al. 2017) and *StAMPP* (Pempleton et al. 2013) packages. Principal component analysis (PCA) on allelic frequencies was run under *adegenet* v2.1.5 (Jombart 2008; Jombart et Ahmed 2011). Discriminant analysis of principal components (DAPC; Jombart et al. 2010) was run using the *adegenet* package. Principal components were selected from cross-validation in DAPC according to the putative origin of individuals and with 1,000 replicates (see Supplementary material S1 for more details). Based on the retained discriminant functions, we inferred group membership probabilities to assess the extent of admixture in the inferred clusters.

We estimated the partition of molecular variance within individual samples, between individual samples from sample sites (or geographic regions), and between sample sites (or geographic regions) using an analysis of molecular variance (AMOVA; Excoffier et al. 1992) implemented in the R packages *poppr* v2.9.3 (Kamvar et al. 2014) and *Pegas* v1.1 (Paradis 2010). To test for significance, we did a randomization test. The matrices of samples were randomly permuted as described in Excoffier et al. (1992).

Hierarchical genetic clustering was done using *ADMIXTURE* v1.3 (Alexander et al. 2009) assuming a number (K) of ancestral populations between two and six. The value of K with lowest associated error value was identified using the cross-validation procedure of *ADMIXTURE*. Then, the R package *stockR* v1.0.74 (Foster 2020) was used with K values (designed in *stockR* to correspond to a number of differentiated groups) from two to six and following the approach outlined in Foster et al. (2018), which has been designed to discriminate groups with no contemporary mixture, based on assignment probability.

The correlation between pairwise genetic (*F*ST and Euclidean Edwards’ distance) and geographic distances was tested using a Mantel test with 1,000 permutations between individuals and geographic regions using the R packages *adegenet* and *ade4* v1.7-18 (Dray and Dufour 2007; Bougeard and Dray 2018; Chessel and Dufour 2004; Dray et al. 2007) on the dataset without outlier only (DATA 4).

* 1. ***Demographic analysis***

We used the software *GADMA* (Genetic Algorithm for Demographic Model Analysis; Noskova et al. 2020)to rebuild the evolutionary trajectory of the two main genetic clusters detected from the previous analyses (see the Results section). Our main objective was the reconstruction of the past and recent demographic trends of these populations, considering the global conservation status of the blue shark. Besides, we also expected demographic reconstruction to allow a better understanding of the position of one particular geographic region (namely, South Africa or SAF) relative to the two detected clusters.

*GADMA* implements a global genetic algorithm and different existing “*engines”* of local optimization (e.g., δaδi (Gutenkunst et al. 2009), moments (Jouganous et al. 2017) or momi (Kamm et al. 2020)) to simulate the expected joint allele-frequency spectrum (JAFS) from multiple populations under various demography scenarios, and compare it with the observed JAFS. More precisely, the genetic algorithm implemented in *GADMA* enables to delineate time structures (epochs) showing particular trends in effective population size (*Ne*) through time (e.g., linear or exponential growth or decline for a given period of time). Local optimization algorithm supports finer inference of demographic parameter values in those predefined structures. This favors the identification of the demographic scenario most likely to explain the observed genetic variation, even with limited *a priori* knowledge, implying the exploration of a wide parameter space.

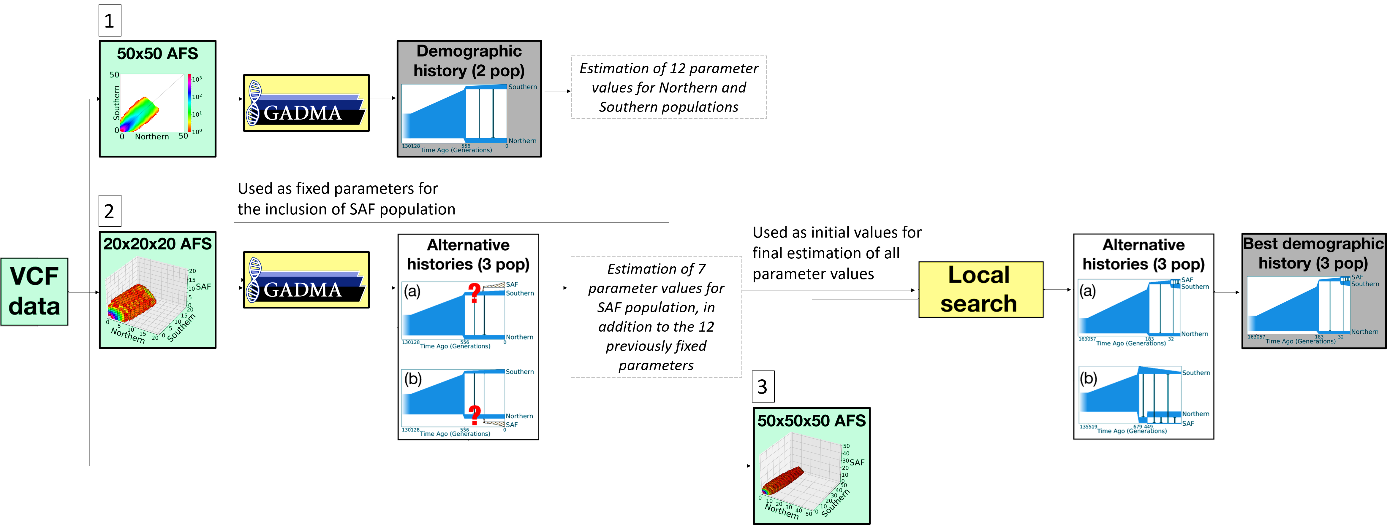
First, we converted the DATA 4 dataset in Variant Call Format (VCF file) using the function *genomic\_converter* in *radiator* (see Supplementary material S1). Then from these converted data, we built the observed JAFS by grouping individual samples according to the populations delineated from the previous structure and clustering analyses. However, considering the computational costs of the analysis of high-density JAFS, we subsampled the original JAFS using *easySFS* (https://github.com/isaacovercast/easySFS) to allow the analysis with realistic computing time and resources**.** *EasySFS* builds the JAFS on this subset, while selecting the number of individual samples and variants needed to maintain the characteristics of the initial JAFS. The *EasySFS* approach implies a subsampling that has an effect similar to the reduction of population sample size, still allowing a reasonable trade-off between data resolution and computational limitations (see Fraisse et al. 2018).

Thus, we ran demographic inferences using a JAFS with 50x50 cells, composed by 31,901 and 32,387 segregating sites for Northern (Northern Atlantic Ocean and Mediterranean Sea) and Southern (Indian and Southwest Pacific oceans) populations, respectively. Indeed, the population structure analysis revealed a clear differentiation of those two clusters of samples (see the Results section). From now on, this will be referred as the 2-population model. The predefined 2-populations demographic scenario allowed for *Ne* of an ancestral population to vary, before that ancestral populations splits into two daughter populations, with steady *Ne* variation allowed and continuous gene flow occurring during divergence. This, overall, corresponded to an Isolation-with-Migration (IM) model integrating *Ne* variations. This 2-population model relied on 12 parameters, nine of which were quantitative (Figure 3.1, Table 4). Three qualitative parameters described trends and dynamics in population effective size (constant, linear or exponential). Five quantitative parameters described the effective size of each population at the start and the end of each time period. Two quantitative parameters described bidirectional migration rates between the two daughter populations. Finally, two quantitative parameters described the duration of the different time periods, specifically, before and after the split of the ancestral population.

The South African geographic region was excluded from the 2-population model, due to the inconsistent results obtained depending on the methods employed for detecting differentiation (see Results section). Thus, we included SAF as a possible undetected third demographic entity in what we will be referred, from now on, as a 3-population demographic model. The 3-population model thus distinguished the South African , Northern (Northern Atlantic Ocean and Mediterranean Sea) and southern (Indian and Southwest Pacific oceans) sampling regions. In the 3-population model, we tested two tree models: (a) first tree model - the ancestral population was split into northern population and a common population made up of SAF and the Southern population, which was split later. (b) second tree model - the ancestral population was split into a southern population and a common population made up of SAF and the northern population, which was split later (Figure 3).

In a first step, all parameter values (both qualitative and quantitative) for the northern and southern populations were fixed as pointed in the 2-population models. A number of 7 quantitative parameters related to the SAF geographic region were left to be inferred. Two were related to the effective population size of SAF at the beginning and at the end of its time period. One was related to its divergence time from either the “Southern (tree model (a))” or the “Northern (tree model (b))” population. Four were related to migration rates with the northern and the southern populations. For fast inference of those seven parameters, we used a downsized JAFS of 20x20x20 cells obtained using *EasySFS* (Figure 3.2).

In a further step, we re-estimated all parameters (16 of which were quantitative, Supplementary material S7) from the 3-population model using the local search Powell's algorithm and a bigger JAFS of size 50x50x50 with 32387, 31901, 31557 segregating sites for northern, southern and SAF populations (Figure 3.3). Fixed and inferred parameters values from the previous step were used as initial points of run.



**Figure 3.** A scheme of demographic inference for two and three populations of the blue shark, *Prionace glauca*. The 2-population model (upper part, (1)) distinguished Northern (Atlantic Ocean and Mediterranean Sea) from Southern (Indian and Southwest Pacific oceans) populations. The 3-population models (mid and lower part, (2) and (3)) distinguished South African, Northern (Atlantic Ocean and Mediterranean Sea) and Southern (Indian and Southwest Pacific oceans) populations, with the two tree models differing depending on the late divergence of SAF from the Northern (a) or Southern (b) group. Light green color corresponds to the data used (Size and dimensions of the observed JAFS, obtained from DATA 4 converted into a VCF file). Light yellow refers to the algorithms used. White and grey refer to the evaluated scenarios, the latter representing the scenario identified as the most likely to explain the observed genetic variation.

For both the 2- and 3-populations models, we converted the inferred parameter values into biologically meaningful values, *i.e.*, in genetic units, using a mutation rate (*µ*) of 10-8 (Karl et al. 2010; Hara et al. 2018) and the effective sequence length (*L*) of 2,598,195 bp as described in Rougeux et al. (2017)[. We first estimated the ancestral effective population size](#_ENREF_15) () from flux Theta :

Then, we defined the effective population size of contemporary populations ():

where corresponds to the inferred effective population size for population *i* from the model. The migration rate () between populations (*i* and *j)* was estimated as the proportion of individuals per generation migrating from population j to population *i*:

where corresponds to the migration parameter estimated from the model.

Finally, we converted time parameter values (i.e., duration of different epochs) in number of generations and inferred the duration of divergence in years ( by integrating a generation time (*tg*) of nine years round value -as previously published estimates were 8.1 years (IOTC 2007)[, and 8.2 and 9.8 years for South African and North Atlantic populations, respectively (Cortés et al. 2015)](#_ENREF_5) [(giving a mean generation time of 8.70 ~ 9 years) :](#_ENREF_2)

For demographic inferences, we used 50 independent GADMA runs with the engine *moments* (Jouganous et al. 2017) to observe convergence in models and infer parameter values, as well as statistical robustness.

As all model parameter values were positive, their logarithms were used to calculate confidence intervals (CIs) for the inferred parameter values. Each bootstrap was performed 100 times under the assumption that all SNPs are independent.

**Results**

* + 1. ***Samples***

The body length of individuals ranged from 74.5 cm to 330.0 cm FL with a mean of 140.4 cm FL. Females were on the average shorter than males (Supplementary material S1, S2) but the opposite trend was observed in the Northern Atlantic, Western Indian Ocean and Eastern Indian Ocean geographic regions (Supplementary material S1). The smallest individuals were sampled in the Northern Atlantic and the largest in the Mediterranean Sea (Supplementary material S1). As only 23 of the 54 individuals sampled from the Mediterranean were measured, caution should be taken when interpreting these results on body size. Indeed, a study with a three times higher sample size (https://fishreg.jrc.ec.europa.eu/web/medbluesgen/sampling-data) revealed a difference in the range of sizes between East and West Mediterranean samples[.](#_ENREF_9)

* + 1. ***Sequencing and quality control***

DArT sequencing yielded 545,764 to 2,702,952 reads per individual with an average of 1,805,674 reads per individual (Figure 2). We obtained 364 genotypes and 172,384 SNPs after applying the DArTseq analytical pipeline (DATA 1; Figure 2). The different filtering steps using the R package *radiator* (detailed in Supplementary material S1) resulted in a dataset of 45,810 SNPs (one SNP per *de novo* assembled fragment) from 312 individuals when considering a stratification per geographic region (Filtration 1), and 34,033 SNPs from 313 individuals when considering a stratification per sample site (Filtration 2) (DATA 2) (Figure 2, Supplementary material S1). Individual samples from the geographic region Northeast Pacific (NPAC, Hawaii) were discarded during this process, as NPAC consisted of only eight individuals with high genotype missingness (considered as a potential result of long-term storage and low DNA quality). The next steps of filtering removed 143 sex-linked SNPs from DATA 2 when considering a stratification per geographic region (45,810 SNPs initially), and 94 sex-linked SNPs from DATA 2 when considering a stratification per sample site (34,033 SNPs initially). From this stage, further filtering led to identical results regardless of a stratification per sample sites or per geographic region, indicating robust data (Supplementary material S1). Thus, search for outliers from DATA 3 was performed considering a stratification per geographic region only (Filtration 1), and not per sample site. We detected 9 outliers in common with both *OutFLANK* and *PCAdapt*, 2,832 loci that were monomorphic at this stage, and 5,171 SNPs with allele frequency below 0.01, yielding to a final dataset of 37,655 non-outlier SNPs on a total of 312 individual samples (DATA 4) (Figure 2). Analyses on the 9 outlier SNPs (DATA 5) revealed results very similar to those detailed below (but see section 4 for an exception in pairwise *F*ST values). Thus, unless stated otherwise, in the section below, we focus on results from the DATA 4 dataset including non-outlier SNPs with a stratification per geographic region. Results for analyses on other datasets or with a stratification per sample site are presented in Supplementary material S1.

* + 1. ***Diversity***

Genetic diversity values are presented in Table 1 and Supplementary material S1. The heterozygosity was low in all geographic regions (observed around 14% and expected around 16% to 17%). These levels of genetic diversity, very similar among all sampled area, are consistent with previous reports (Leone et al. 2017; Bailleul et al. 2018). We did not detect significant Hardy-Weinberg disequilibrium, except for South Africa and Southwest Pacific (SWPAC) where observed heterozygosity was lower than expected. The positive *FIS* values obtained (0.031-0.115) differed from zero in all areas except the Eastern Indian Ocean (EIO); possibly due to low statistical power linked to low sample size (N = 8).

**Table 1.** Genetic diversity estimates for *Prionace glauca* per geographic region and without outliers (from DATA 4: 37,655 non-outlier SNPs and 312 individuals). Allelic richness (Ar) with the low and high 95% CI, Number of individual samples (Nb), Observed heterozygosity (Hobs), Expected heterozygosity (Hexp), Unbiased expected heterozygosity (Nei 1978) (Hexp unbiased), inbreeding coefficient (*FIS*) with the low and high 95% CI on *F*IS, p-values from chi-square test for goodness-of-fit to Hardy-Weinberg equilibrium (HWE), test significance on homozygote and heterozygote deficiency. *MED* Mediterranean; *NATL* Northern Atlantic; *NEATL* Northeastern Atlantic; *SAF* South Africa; *EIO* Eastern Indian Ocean; *NIO* Northern Indian Ocean; *SWPAC* Southwest Pacific; *WIO* Western Indian Ocean

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Pop** | **Ar** | **Nb** | **Hobs** | **Hexp** | **Hexp unbiased** | ***F*IS** | **Test global HWE** | **Test significance HWE (homozygote deficiency)** | **Test significance HWE (heterozygote deficiency)** |
|
| NATL | 1.594 (1.54-1.62) | 42 | 0.145 | 0.168 | 0.170 | 0.091 (0.074-0.088) | 1.000 | 1.000 | 1.000 |
| NEATL | 1.568 (1.50-1.61) | 21 | 0.142 | 0.166 | 0.170 | 0.102 (0.069-0.094) | 1.000 | 1.000 | 1.000 |
| MED | 1.583 (1.53-1.61) | 45 | 0.142 | 0.166 | 0.168 | 0.099 (0.083-0.096) | 1.000 | 1.000 | 0.000 |
| SAF | 1.583 (1.53-1.60) | 105 | 0.140 | 0.166 | 0.166 | 0.115 (0.103-0.113) | 0.000 | 1.000 | 0.000 |
| WIO | 1.562 (1.49-1.60) | 22 | 0.141 | 0.163 | 0.167 | 0.088 (0.059-0.081) | 1.000 | 1.000 | 1.000 |
| NIO | 1.547 (1.48-1.60) | 16 | 0.143 | 0.161 | 0.166 | 0.067 (0.028-0.057) | 1.000 | 1.000 | 1.000 |
| EIO | 1.506 (1.44-1.58) | 8 | 0.145 | 0.157 | 0.167 | 0.031 (-0.067-0.021) | 1.000 | 1.000 | 1.000 |
| SWPAC | 1.581 (1.53-1.61) | 53 | 0.139 | 0.166 | 0.167 | 0.113 (0.096-0.110) | 0.000 | 1.000 | 0.000 |

* + 1. ***Genetic differentiation***

Differences between geographic regions, between individual samples within a geographic region, and within individual samples explained 0.09%, 11.83%, 88.08% of the molecular variance, respectively. Despite *ϕ* statistic of population differentiation suggested a low amount of differentiation between geographic regions (Supplementary material S1), the variance estimates within and between samples, and between geographic regions were all significant (*p* < 0.001).

Pairwise *F*ST values were of the order of 10-3 to 10-4 both between geographic regions (Table 2) and between sampling sites (Table 3). *F*ST values were significant among all geographic regions, with South African populations differentiated from the North Atlantic but not from Indo-Pacific ones ; Table 2), while no genetic differentiation was detected within regions (Table 3), with two exceptions involving the South African site SAF2 differing from SWPAC2 and SWPAC3 in the Indo-Pacific (Table 3). Similar pattern emerged on the dataset containing only outlier loci and a stratification per geographic region (DATA 5) (*F*ST Tables 12 and 13 from Supplementary material S1).

Clustering analyses using PCA (Supplementary material S1, S3), DAPC (Figure 4), *ADMIXTURE* (Figure 5A, Supplementary material S1) and *stockR* (Figure 5B, Supplementary material S1) revealed two distinct genetic groups, one including geographic regions from Northern Atlantic and the Mediterranean (MED, NATL and NEATL), the other including geographic regions from the Indo-West Pacific (EIO, NIO, WIO and SWPAC). Individual samples from the SAF geographic region showed an admixed profile with *ADMIXTURE* analysis, while *stockR* clustered of SAF individuals with Indo-West Pacific ones (in concordance with *F*ST results) with both «non-outlier» (DATA 4) and «outlier» (DATA 5) datasets.

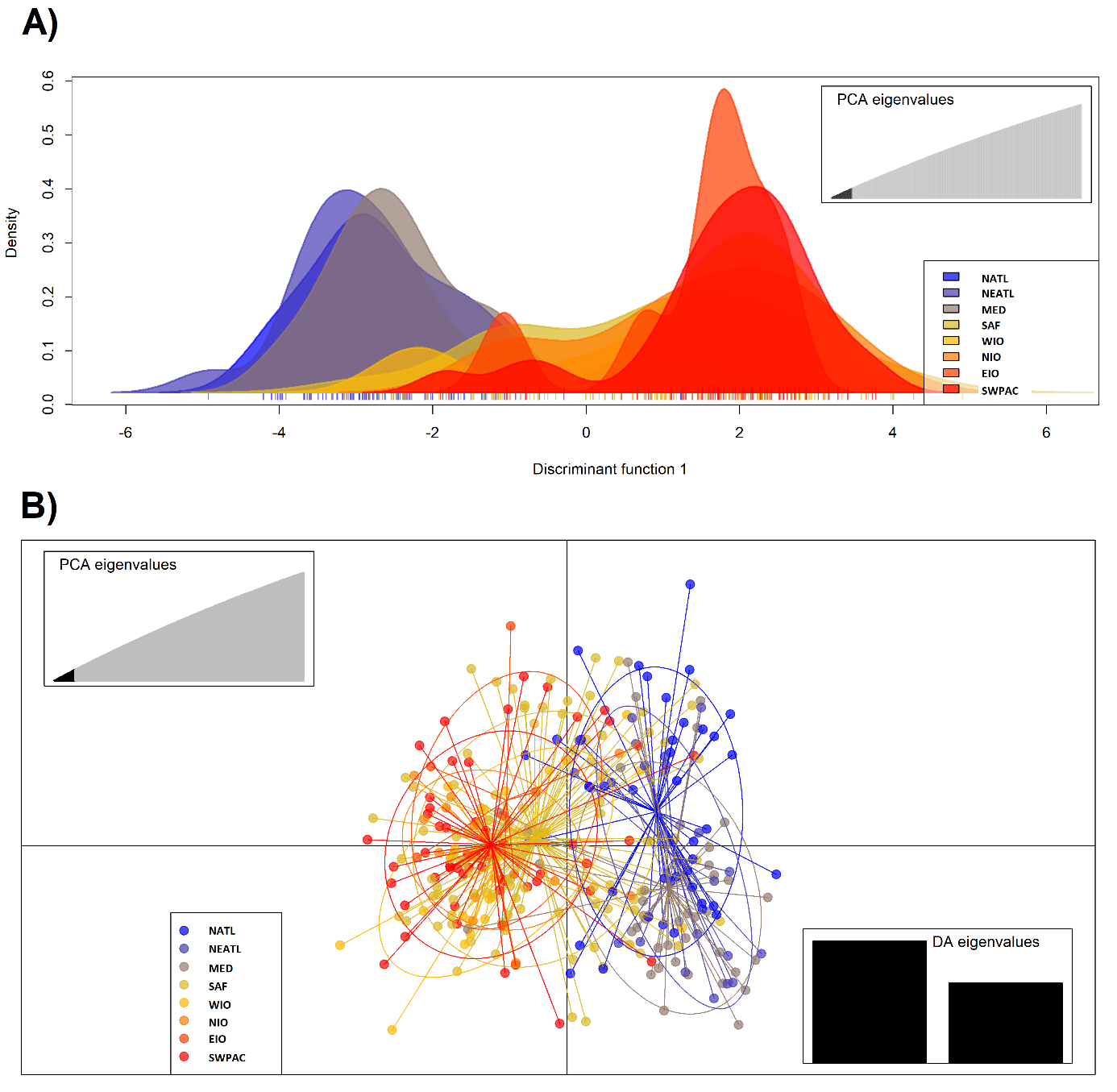
Finally, no significant correlation between geographic and genetic distance (either expressed as pairwise *F*ST or Euclidean distance) was detected through the Mantel test performed on geographic regions (p > 0.05) (Supplementary material S1).

**Table 2.** *Prionace glauca*. Pairwise *F*ST values and their level of significance after correction with q-value (\* p < 0.01, and \*\*p < 0.001) (from DATA 4: 37,655 non-outlier SNPs and 312 individuals) per geographic region.*MED* Mediterranean; *NATL* Northern Atlantic; *NEATL* Northeastern Atlantic; *SAF* South Africa; *EIO* Eastern Indian Ocean; *NIO* Northern Indian Ocean; *SWPAC* Southwest Pacific; *WIO* Western Indian Ocean.

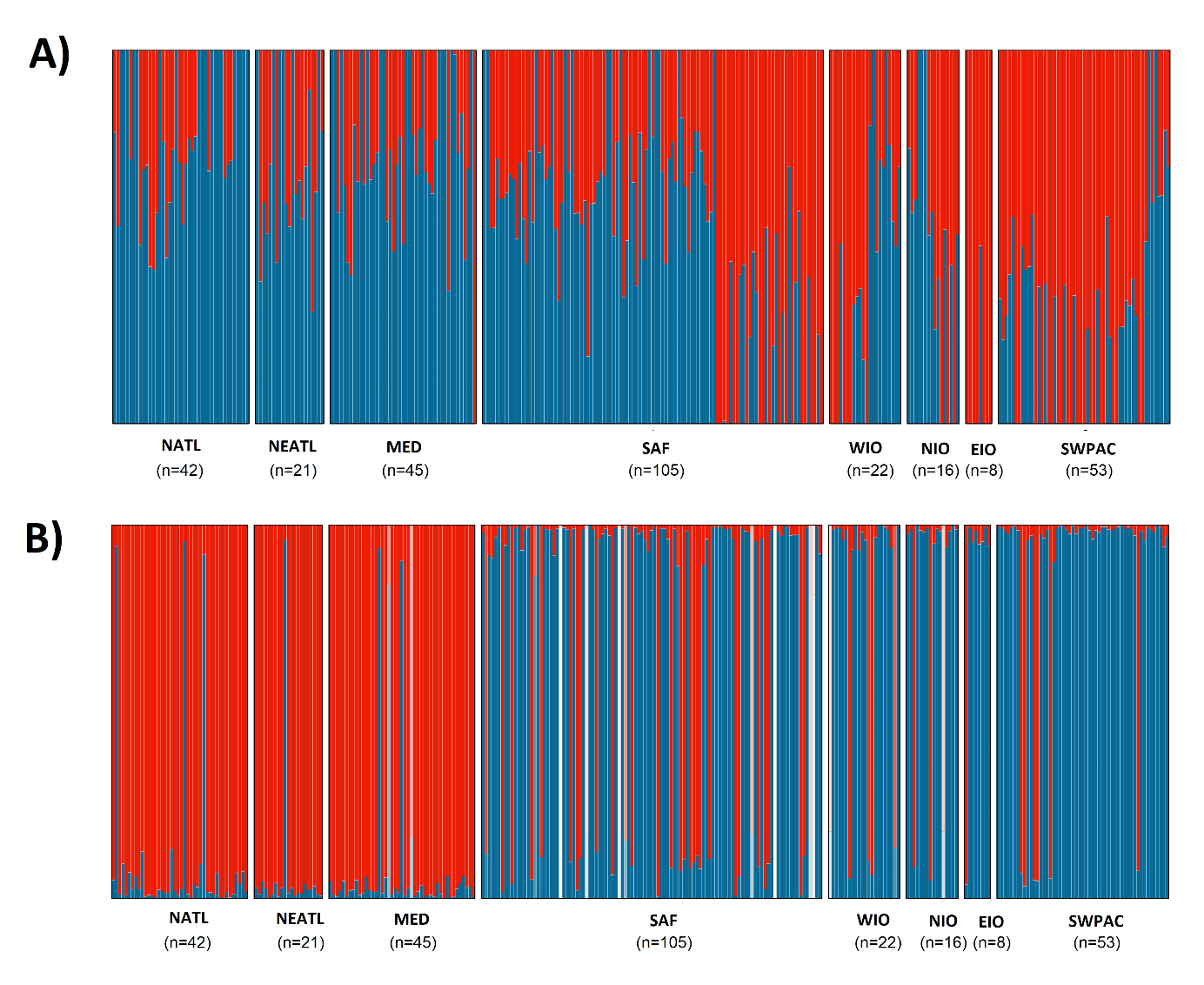
|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **MED (45)** | **NATL (42)** | **NEATL (21)** | **SAF (105)** | **EIO (8)** | **NIO (16)** | **WIO (22)** | **SWPAC (53)** |
| **MED** |  | 0.0007\*\* | 0.0010\*\* | 0.0015\*\* | 0.0023\* | 0.0017\*\* | 0.0017\*\* | 0.0022\*\* |
| **NATL** |  |  | 0.0005\* | 0.0012\*\* | 0.0010\* | 0.0011\*\* | 0.0016\*\* | 0.0017\*\* |
| **NEATL** |  |  |  | 0.0017\*\* | 0.0015\* | 0.0015\*\* | 0.0015\*\* | 0.0020\*\* |
| **SAF** |  |  |  |  | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **EIO** |  |  |  |  |  | 0.0004 | 0.0000 | 0.0000 |
| **NIO** |  |  |  |  |  |  | 0.0000 | 0.0000 |
| **WIO** |  |  |  |  |  |  |  | 0.0001 |
| **SWPAC** |  |  |  |  |  |  |  |  |

**Table 3.** *Prionace glauca*. Pairwise *F*ST values and their level of significance after correction with q-value (\* p < 0.01, and \*\*p < 0.001) (from DATA 3: 33,939 SNPs and 313 individuals) per sampling site. *MED1, MED2* Mediterranean; *NATL* Northern Atlantic; *NEATL* Northeastern Atlantic; *SAF1, SAF2* South Africa; *EIO* Eastern Indian Ocean; *NIO* Northern Indian Ocean; *SWPAC1-3* Southwest Pacific; *WIO* Western Indian Ocean

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **NATL** | **NEATL** | **MED1** | **MED2** | **SAF1** | **SAF2** | **WIO** | **NIO** | **EIO** | **SWPAC1** | **SWPAC2** | **SWPAC3** |
| **(33)** | **(20)** | **(29)** | **(14)** | **(14)** | **(76)** | **(20)** | **(15)** | **(7)** | **(30)** | **(10)** | **(9)** |
| **NATL** |  |  |  |  |  |  |  |  |  |  |  |  |
| **NEATL** | 0.0005 |  |  |  |  |  |  |  |  |  |  |  |
| **MED1** | 0.0004\*\* | 0.0008\* |  |  |  |  |  |  |  |  |  |  |
| **MED2** | 0.0013\*\* | 0.0012\*\* | 0.0005 |  |  |  |  |  |  |  |  |  |
| **SAF1** | 0.0014\*\* | 0.0015\*\* | 0.0015\*\* | 0.0028\*\* |  |  |  |  |  |  |  |  |
| **SAF2** | 0.0021\*\* | 0.0025\*\* | 0.0019\*\* | 0.0036\*\* | 0.0001 |  |  |  |  |  |  |  |
| **WIO** | 0.0022\*\* | 0.0024\*\* | 0.0021\*\* | 0.0036\*\* | 0.0002 | 0.0000 |  |  |  |  |  |  |
| **NIO** | 0.0024\*\* | 0.0028\*\* | 0.0025\*\* | 0.0039\*\* | 0.0004 | 0.0003 | 0.0000 |  |  |  |  |  |
| **EIO** | 0.0025\* | 0.0031\*\* | 0.0031\*\* | 0.0049\*\* | 0.0008 | 0.0008 | 0.0000 | 0.0007 |  |  |  |  |
| **SWPAC1** | 0.0028\*\* | 0.0030\*\* | 0.0029\*\* | 0.0044\*\* | 0.0003 | 0.0002 | 0.0000 | 0.0002 | 0.0002 |  |  |  |
| **SWPAC2** | 0.0030\*\* | 0.0032\*\* | 0.0032\*\* | 0.0046\*\* | 0.0006 | 0.0007\* | 0.0001 | 0.0000 | 0.0000 | 0.0000 |  |  |
| **SWPAC3** | 0.0034\*\* | 0.0029\*\* | 0.0033\*\* | 0.0051\*\* | 0.0006 | 0.0012\* | 0.0000 | 0.0012 | 0.0006 | 0.0007 | 0.0000 |  |



**Figure 4.** *Prionace glauca*. Results of genetic clustering using DAPC (from DATA 4: 37,655 non-outlier SNPs and 312 individuals). (A) density of individuals on discrimination function with K = 2. (B) Scatter plot with K=3, with colours corresponding to geographic regions.*MED* Mediterranean; *NATL* Northern Atlantic; *NEATL* Northeastern Atlantic; *SAF* South Africa; *EIO* Eastern Indian Ocean; *NIO* Northern Indian Ocean; *SWPAC* Southwest Pacific; *WIO* Western Indian Ocean.

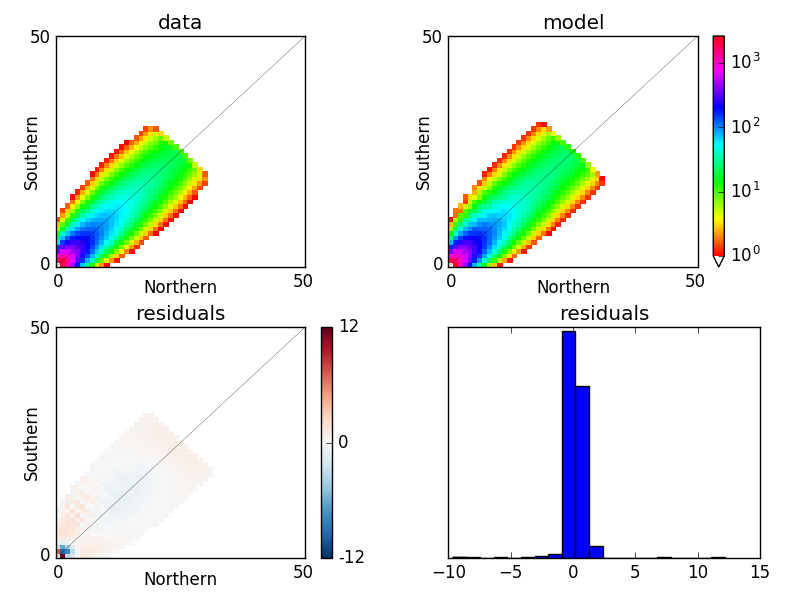


**Figure 5.** *Prionace glauca*. Results of genetic clustering using *ADMIXTURE* (A) and *stockR* (B) (from DATA 4: 37,655 non-outlier SNPs and 312 individuals). Individuals with admixed background are indicated in white, as *stockR* does not test admixture. *MED* Mediterranean; *NATL* Northern Atlantic; *NEATL* Northeastern Atlantic; *SAF* South Africa; *EIO* Eastern Indian Ocean; *NIO* Northern Indian Ocean; *SWPAC* Southwest Pacific; *WIO* Western Indian Ocean.

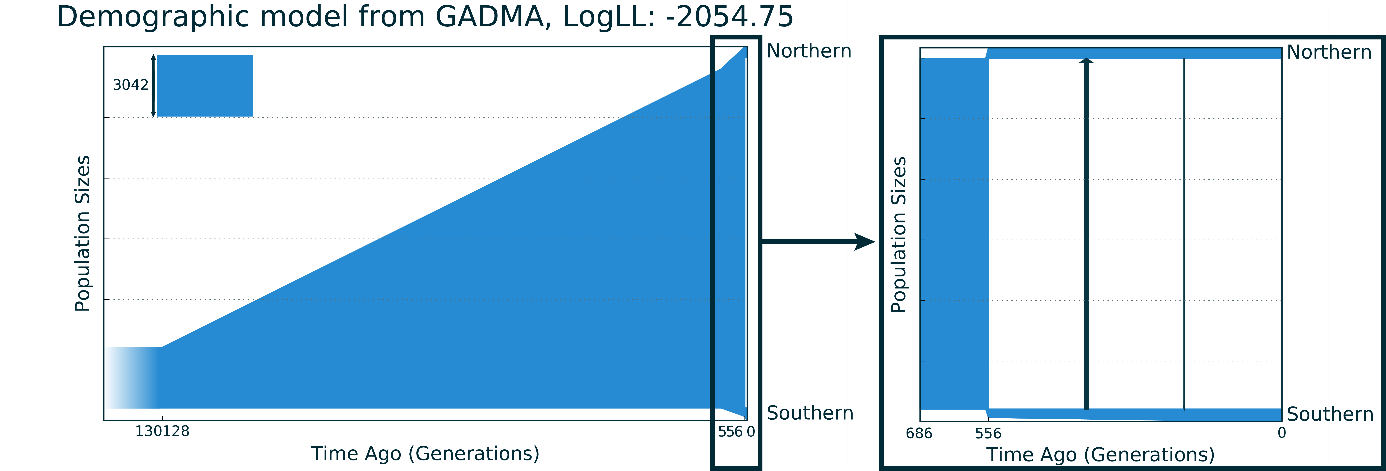
* 1. ***Demographic inference***

Scenario with northern and southern populations (2-population model):

The demographic scenario showing the best fit to the observed genetic variation (log-likelihood = -2054.75) (Figure 6, Supplementary material S4) corresponded to an ancestral population that started to linearly increase ca. 130,000 generations ago (between ca. 1 and 1.275 My ago) reaching an effective population size from 30,000 to up to 170,000 individuals, before splitting into two populations (Figure 7). During the split that would have occurred ca. 550 generations ago (i.e., ca. 5,000 [4,503-5,449] years ago), the two populations may have experienced important bottlenecks, as suggested by the inferred contemporary effective population sizes of only about 5,300 and 6,400 in northern and southern populations, respectively (Table 4). The constant effective population size of the northern population was estimated to be approximately 5,300 and the Southern population may have experienced a marginal linear growth of ca. 2000 (ca. 4,200 - ca. 6,300) between the time of split and present day. In addition, an important asymmetry in gene flow was estimated, with preferential migration from southern to northern population (MSou-Nor = 0.168 vs. MNor-Sou = 0.004) (Table 4).



**Figure 6.** Projection of the joint allele frequency spectrum (JAFS) observed from empirical data (top left) (50x50 cells JAFS subsampled from DATA 4 – 37,655 non-outlier SNPs and 312 individuals), and expected under the best demographic scenario inferred by *GADMA* (top right). The bottom line represents the Anscombe residuals (left) and the histogram of the residuals (right) from the comparisons between the observed and expected JAFS.



**Figure 7.** Schematic plot of the best demographic scenario inferred by *GADMA* for the demographic history of southern and northern populations. Past time is measured in number of generations, zero generation being present time. Vertical arrows in the right panel represent asymmetric migration after the split.

**Table 4.** Demographic parameter values inferred by *GADMA* from 50x50 JAFS with 2-population model including northern (Northern Atlantic Ocean and Mediterranean Sea) and southern (Indian and Southwest Pacific oceans) populations. Nanc: Ancestral population size. NA\_split: Population size after the initial split. NNor0: Present population size of the northern population. NNor: Population size of the northern population after the split. NSou0: Present population size of the southern population. NSou: Population size of the southern population after the split. TA: Beginning time of ancestral population growth. Tsplit: Time from the initial split. MSou-Nor: Migration rate from southern to northern population. MNor-Sou: Migration rate from northern to southern population. *CI* confidence interval

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Parameter** | **Value** | **CI** | **Unit** | **Parameters’ definition** |
| Nanc | 30,422 | [12,745 – 29,119] | Individuals | Population size of the ancestral population |
| NA\_split | 176,072 | [152,916 – 172,826] | Individuals | Population size of the ancestral population after linear growth |
| NNor0 | 5,255 | [1,315 – 7,144] | Individuals | Population size of the northern population now after linear growth |
| NNor | 5,255 | [1,315 – 7,144] | Individuals | Population size of the northern population after the split and till now (constant size from NNor0) |
| NSou0 | 4,220 | [979 – 7,382] | Individuals | Population size of the southern population after split of ancestral population |
| NSou | 6,356 | [1,067 – 10,171] | Individuals | Population size of the southern population now after linear growth |
| TA | 130,128 | [141,491 – 197,366] | Number of generations | Time of the beginning of ancestral population linear growth |
| Tsplit | 556 | [139 - 705] | Number of generations | Time of ancestral population split |
| MSou-Nor | 0.0168 | [0.0127 - 0.0853] | Proportion per generation | Migration rate from southern population to northern |
| MNor-Sou | 0.0004 | [0.0001 - 0.0173] | Proportion per generation | Migration rate from northern population to southern |

Scenario with northern and southern populations, and South Africa (3-population model):

Previous population structure and clustering analysis suggested that the SAF geographic region may not be considered as a separate population. This 3-population demographic was constructed to evaluate the most likely membership of SAF individuals favoured the first tree model (a) (i.e., with SAF population originating from the southern cluster) (Supplementary material S5, S6, S7, S8). Nevertheless, non-negligible exchange through migration was detected over time (Supplementary material S6, S7) between the northern and southern areas in particular, through SAF which seemingly acts as a stepping stone between the Atlantic and the Indo-West Pacific regions.

**Discussion**

In contrast to previous reports of large scale panmixia in the blue shark (within oceanic basins: Queiroz et al. 2012; King et al. 2015 and among oceanic basins: Verissimo et al. 2017; Ovenden et al. 2009; Taguchi et al. 2015; Bailleul et al. 2018), the work presented here demonstrates for the first-time clear evidence of population genetic structure both within and among oceanic basins. These signatures of population differentiation were obtained with genome-wide markers. This work also highlights hints of bottleneck in a recent past.

***Genetic diversity, population structure and demography history of blue shark***

Genetic diversity levels (ex. heterozygosity around 14% - 17%) are consistent with previous reports on this species (Leone et al. 2017; Bailleul et al. 2018) and similar to those reported in the closely related genus *Carcharinus* (Pazmiño et al. 2017; Momigliano et al. 2017; Green et al. 2019). The limited but significant heterozygotes deficiency, from *FIS* values, was comparable to results obtained in previous studies (e.g., *FIS* ranging from −0.073 to 0.054 in Bailleul et al. 2018; −0.01 to 0.07 in Verissimo et al. 2017). It is most likely due to null alleles (and/or Wahlund effects, although limited, if existing, in view of the low level of population differentiation) considering the absence of any behaviour known to favour inbreeding in blue.

Significant pairwise *F*ST values, together with genetic clustering analyses, revealed two distinct genetic clusters across the studied area: (1) the Northern Atlantic Ocean region, including the Mediterranean Sea, and (2) the Indo-West Pacific region, seemingly acting as independent demographic units. In fact, most analyses here pointed toward a clear split between samples collected in northern (Mediterranean and Northern Atlantic) and southern (South African Atlantic, Indian and Pacific Oceans) areas. When considering the two groups with scrutiny, significant *F*ST values also showed a subtle differentiation between the Atlantic and the Mediterranean (Table 3 & 4).

Reconstruction of past demographic history using the site frequency spectrum suggested that the two main clusters diverged from an ancestral population after it experienced substantial increase in effective population size. The estimate of the time periods of population growth for the ancestral population, as well as divergence time between ancestors of the contemporary populations, required postulating the average generation time. We chose nine years as an intermediate value between the previous estimates of 8.1, 8.2 and 9.8 years (IOTC 2017, Cortés et al. 2015). Under this scenario, the linear growth of the ancestral population would have started in the early Pleistocene, between 1.05 Mya and 1.28 Mya, and the split into northern and southern populations would have happened 4,503-5,449 years ago ( 550 generations ago), during the Holocene.

Paleoclimatic events could have triggered the divergence between northern and southern populations: Holocene tropical sea-surface temperatures (SSTs) followed a warming trend until 5,000 years ago and a global stabilisation until present, albeit some local variations. Leduc et al. (2010) revealed a warming of up to ~2°C (SSTs) in the western tropical Atlantic and eastern tropical Pacific from the early Holocene to present. Meanwhile, a global Northern Hemisphere cooling happened about 5,000 years ago (Masson-Delmotte et al. 2013). Lake sediment records from Greenland “suggest that around 4,500 and 650 years ago variability associated with the North Atlantic Oscillation changed from generally positive to variable, intermittently negative conditions” (Olsen et al. 2012; Masson-Delmotte et al. 2013). Furthermore, this cooling of SST occurred concomitantly in the Southern Hemisphere. In the Australian-New Zealand region, Holocene SST followed a cooling trend (Bostock et al. 2013), as in the high-latitude Southern Ocean from the early to the late Holocene (Kaiser et al. 2008; Shevenell et al. 2011). These past variations in SST may have favored a split between northern and southern populations. Shifts in blue shark reproductive seasons between the two hemispheres might have contributed to maintain that scission. In fact, while reproduction occurs in the summer (July, August) in the Northern hemisphere (Fujinami et al. 2017), it is reported as more likely from December to July in the Southwestern Equatorial Atlantic Ocean (Coelho et al. (2017), and from October to December in the Indian ocean (Druon et al. 2022)

Historical demography analysis suggested this past divergence event was associated with a strong bottleneck followed by rather constant population size in both the northern and the southern populations, despite suggestion of an increase of the latter, albeit moderate. The *GADMA* analysis suggested just after the split, the effective population size of both daughter populations represented only c. 2 to 3% of the ancestral effective population size. Furthermore, the analysis suggested rather low contemporary effective population sizes (Ne  4,000-6,500 based on the 2-population model). These estimates are of the same order of magnitude as the estimates ( 4,000 to 5,000) previously proposed by Verissimo et al. (2017) in the Atlantic, and King et al. (2015) in the Pacific, using distinct estimation methods based on microsatellite datasets. Historical demography analysis also suggested limited and strongly asymmetric gene flow, most of it being directed from the southern to the northern population. This asymmetrical direction of gene flow might be due to the increase in population size of the southern population (e.g., Rougeux et al. 2017). Besides, gene flow towards the northern area might explain the difficult assignment of the individual samples from the geographical region South Africa (SAF) to one or the other of the two genetic clusters detected with clustering analysis.

In fact, the relative positioning of the SAF sample compared to the two clusters slightly varied between different genetic clustering methods. For instance, PCA and *ADMIXTURE* analyses suggested an intermediate position of the SAF sample between the Atlantic and the Indo-West Pacific. In contrast, *StockR* (designed to explicitly ignore contemporary mixture) and *F*ST values pointed toward the inclusion of the SAF sample in the southern cluster, in line with the three-population GADMA demographic. Yet, *F*ST also showed hints of genetic differentiation between the Southwest Atlantic and the Indo-West Pacific, similar to earlier reports using mitochondrial DNA markers (Bitencourt et al. 2019). South Africa might be an area of occasional reproductive mixing between northern and southern oceanic populations or be under the influence of an east-west migration road. Indeed, currents are particularly strong and structured in eddies in this region. Off the southern tip of Africa, the Agulhas current retroflects eastward, into the Indian Ocean. The retroflection takes the form of an unstable jet that can shed warm eddies into the Atlantic Ocean (Nikolic et al. 2020b). Studies on bigeye tuna reported the lack of a clear physical barrier preventing mutual exchange between the Atlantic and Indo-West Paciﬁc regions and the co-occurrence of two separate populations off the southern tip of Africa (Chow et al. 2000; Durand et al. 2005).

It is unclear whether the differentiation detected between the two main clusters reflects a northern-southern split (possibly due to the shift in reproductive season between both hemispheres: Nakano and Seki 2003; Nakano and Stevens 2008), or to a Western/Eastern split between Mediterranean-Atlantic and Indo-West Pacific regions. Samples from the southwest Atlantic such as the Brazilian samples discussed by Verissimo et al (2017) and Biton-Porsmoguer et al. (2019) and southeast Atlantic, such as Gulf of Guinea, may help ascertaining the localization and the origin of the split between the two main clusters identified here.

Despite this apparent split dominating the clustering analysis, results also revealed structure at a finer spatial scale, with differentiation between the Mediterranean and the Northern Atlantic. These results confirm the previous suggestions of differentiation of the Mediterranean populations based on mitochondrial (Leone et al. 2017) and microsatellite data (Bailleul et al. 2018). These thus support the existence of two demographic units in the Northern Hemisphere, although this differentiation is subtler and less marked than the split between the two main clusters. Further analyses by Leone et al. (in prep.) are ongoing to investigate the connectivity level between the North East Atlantic and the Mediterranean Sea.

***Implications for fisheries***

Fisheries management requires the identification of demographically independent units (Carvalho and Hauser 1995; Waples et al. 2008). Bailleul et al. (2018) showed through simulations the extensive time-lag between demographic change and its detectable imprint in population genetic structure (the ‘grey zone’ effect), and suggested genome scans may offer the necessary power to detect patterns of population structure and escape this ‘grey zone’ effect. Results obtained here support this prediction, showing that denser locus coverage enables detecting subtle genetic differences among blue shark populations from distinct ocean basins. They support the use of genome scan as a useful tool to define management units and assess demographic trends for conservation proposes.

The ICCAT provided recommendations to fill gaps in our understanding of population differentiation using genetic data (ICCAT 2015; ICCAT 2019). Results obtained here should feed future models and management plans by the ICCAT, particularly regarding the minimum number of management units to be considered. In fact, genetic differentiation between the Mediterranean and the Northern Atlantic indicates the occurrence of at least two independent demographic entities for the blue shark, supporting the need for a revision of the management units recognized thus far. We failed to detect very recent bottlenecks (i.e., less than 10 generations ago) using the GADMA analysis: the detection of such recent events is complex and varies depending on the type of genetic data and approaches (Sovic et al. 2019).

No genetic differentiation was detected between the Indian Ocean and the Southwest Pacific, which is consistent with tagging results that reveal long migrations by individuals from the Southwest Pacific, across the Indian Ocean to South Africa (West et al. 2004). Our easternmost samples came from southeast Australia, New Zealand, and New-Caledonia (Figure 1), the eastern boundary of the Indian Ocean blue shark stock could therefore be set at 170°E. A complementary assessment would require joint work between the IOTC and the Western and Central Pacific Fishery Commission. In the Indian Ocean, the Indian Ocean Tuna Commission (IOTC) blue shark assessment currently assumes a single stock (IOTC, 2017). Our results did not provide evidence for multiple stocks at the scale of the Indian Ocean. However, considering the low number of migrants sufficient to erase population differentiation (Bailleul et al. 2018), this lack of evidence for demographic independence within the Indian Ocean and between the Indian Ocean and Southern Pacific should still be interpreted cautiously.

We encourage additional genetic studies including more locations, with a focus on the South Atlantic, and the Southeast and North Pacific to refine our understanding of connectivity within the Indian Ocean and around South Africa. More importantly, considering the possible implication of shifts in reproductive period among locations, we encourage a targeted sampling in reproductive zones. Such additional data would help ascertaining in genetic differentiation among oceanic basins, and possibly detecting finer-scale differentiation. Moreover, we encourage pursuing thorough population monitoring within blue shark management units. Indeed, an anthropogenic-driven decrease in blue shark effective population sizes could actually be occurring, but undetectable at present. Such a phenomenon would be worthy of interest, considering the relatively low effective population sizes estimated in this study.

**Conclusion**

Knowledge of population structure, dynamics and connectivity is essential for the management and conservation of exploited and threatened species. Indirect methods are essential in the vast majority of cases where information cannot be obtained through direct observations, a situation exemplified in the marine realm. Great expectations emerged with the availability of molecular markers enabling the estimation of geographic structure and demographic trends in the theoretical framework of population genetics, yet results have sometimes been inconclusive, particularly for those species with large population sizes and/or potential for large-scale migration. The results obtained here support Bailleul et al.’s (2018) prediction that the “grey zone effect” may be circumvented thanks to the enhanced power offered by genome scans. The use of several tens of thousands of SNPs enabled the detection of weak levels of population differentiation, demonstrating for the first time the existence of distinct demographic units in the blue shark. It also enabled the reconstruction of past fluctuations in effective population size. The present results should lead to revising stock assessment in the blue shark and taking appropriate management measures in the dedicated fisheries commissions. They also should stimulate dedicated genetic sampling of reproductive aggregations, rather than opportunistic sampling through fisheries, in still poorly represented areas of the blue shark global range. Beyond the blue shark case, our results highlight the interest of the increased power of genome scan to describe demographically independent populations and stocks. For those marine species in which the “population grey zone” had thus far hampered such essential discrimination, genome scan could help distinguish between true panmixia and insufficient statistical power from limited molecular markers.

**Acknowledgements**

We are grateful to skippers Morgan Le Guernic, Rudy Levian, Mathias Hoarau, Samuel Kazambo, Sébastien Laffont, Rudy Levian, Leopold Corbrejaud of the domestic longline fleet for their continued support of the “PSTBS-IO” project. We also thank Reunimer and Enez, particularly Hubert Chenede and Frédéric Payet for facilitating access to work plans for sampling in good condition. We thank Estelle Crochelet and Loïc Le Foulgoc for their participation in the sampling. Many thanks to Steward Norman (CapFish) for their huge help to collect samples in South Africa. We thank all the CSIRO team for their help on production of DArT genotype data and collection of samples in the Pacific and eastern Indian Ocean. We also thank Thierry Gosselin for very interesting exchanges on genetic analysis and help on filter\_rad. Finally, we thank Arthur Georges for his advice on the *dartR* package and contributors to the dartR forum (https://groups.google.com/forum/#!topic/dartr/).

**Funding**

This work is part of the PSTBS-IO project supported by funding from CSIRO Oceans and Atmosphere, AZTI Tecnalia, Institut de recherche pour le développement (IRD), and Research Institute for Tuna Fisheries (RITF) and financial assistance of the European Union (GCP/INT/233/EC – Population structure of IOTC species in the Indian Ocean), and POPSIZE project supported by FEAMP (2014-2020 UE N°508/2014).

**Authorizations**

Permit: 0001500212. This permit was issued under Biosecurity Act 2015 Section 179 (1).

**Data Access**

Metadata, raw data and final data are freely available from the IFREMER Sextant repository:

<https://doi.org/10.12770/f0ad76a4-a9d3-4cd7-aaf5-7a1a43dc1e99>.

**Author Contributions Statement**

The study was designed by NN, SA, FM, and FDD. Samples were obtained by NN, DB, CLH, MH, PB, FP, DF, SBP, MF, PG, and all people acknowledged in the paper. DNA was extracted by FDD, JA, and ML. NN, FDD, SA, CD, and EN did the genetic analysis. NN, FDD, EN, and SA prepared the figures and tables. NN, SA, and FDD wrote the first version of the manuscript. All authors (NN, FDD, DB, EN, CR, CD, CLH, MH, AM, PB, PF, PG, CD, JF, DF, SBP, FP, DP, AL, JA, ML, FM, and SA) revised and improved the manuscript. All authors (NN, FDD, DB, EN, CR, CD, CLH, MH, AM, PB, PF, PG, CD, JF, DF, SBP, FP, DP, AL, JA, ML, FM, and SA) reviewed the manuscript.

**References**

Aires-da-Silva, A.M., Hoey, J.J., Gallucci, V.F. 2008. A historical index of abundance for the blue shark (*Prionace glauca*) in the western North Atlantic. *Fisheries Research*, 92: 41-52.

Aires-da-Silva, A.M., Maunder, M.N., Gallucci, V.F., Kohler, N.E., Hoey, J.J. 2009. A spatially structured tagging model to estimate movement and fishing mortality rates for the blue shark (*Prionace glauca*) in the North Atlantic Ocean. *Marine and Freshwater Research, 60*: 1029-43.

Alexander, D.H, Novembre, J., Lange, K. 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome research, 19*: 1655-64.

Allaire, J. J., Xie, Y., McPherson, J., Luraschi, J., Ushey, K., Atkins, A., Wickham, H., Cheng, J., Chang, W., Iannone, R. 2020. Rmarkdown: dynamic documents for R. https://github.com/rstudio/rmarkdown.

Archer, F.I., Adams, P.E., Schneiders, B.B. 2017. stratag: An r package for manipulating, summarizing and analysing population genetic data. *Molecular Ecology Resources, 17*: 5-11.

Bailleul, D., Mackenzie, A., Sacchi, O., Poisson, F., Bierne, N., Arnaud-Haond, S. 2018. Large-scale genetic panmixia in the blue shark (*Prionace glauca*): A single worldwide population, or a genetic lag-time effect of the "grey zone" of differentiation?. *Evolutionary Applications, 11*: 614-30.

Baum, J. K., Blanchard, W. 2010. Inferring shark population trends from generalized linear mixed models of pelagic longline catch and effort data. *Fisheries Research*, *102*: 229-39.

Bitencourt, A., Silva, D.A., Carvalho, E.F., Loiola, S., Amaral, C.R.L. 2019. Study of genetic variability of the Blue Shark *Prionace glauca* (Linnaeus, 1758). *Forensic Science International Genetics Supplement Series, 7*: 594-96.

Biton-Porsmoguer, S. 2017. Análisis de la explotación del pez espada *Xiphias gladius* y de la tintorera Prionace glauca por la flota palangrera catalana durante el periodo 2010-2015 en el Mediterráneo occidental. *Revista de Biología Marina y Oceanografía, 52*(1):175-179.

Biton-Porsmoguer, S. 2018a. Explotación intensiva de la tintorera *Prionace glauca* y del marrajo *Isurus oxyrinchus* en el Atlántico noreste entre 2001 y 2016. *Revista de Biología Marina y Oceanografía*, 53(1):27-38.

Biton-Porsmoguer, S. Lloret, J. 2018b. Potentially unsustainable fisheries of a critically-endangered pelagic shark species: The case of the blue shark (*Prionace glauca*) in the Western Mediterranean Sea. *Cybium: International Journal of Ichthyology*, *42*(3): 299-302.

Bostock, H. C., et al. 2013. A review of the Australian–New Zealand sector of the Southern Ocean over the last 30 ka (Aus-INTIMATE project*). Quat. Sci. Rev., 74*:35-57.

Bougeard, S., Dray, S. 2018. Supervised Multiblock Analysis in R with the *ade4* Package. *Journal of Statistical Software, 86*: 1-17.

Campana, S.E., Joyce, W., Manning, M.J. 2009. Bycatch and discard mortality in commercially caught blue sharks *Prionace glauca* assessed using archival satellite pop-up tags. *Marine Ecology Progress Series, 387*: 241-53.

Campana, S.E., Marks, L., Joyce, W., Kohler, N. E. 2006. Effects of recreational and commercial fishing on blue sharks (*Prionace glauca*) in Atlantic Canada, with inferences on the North Atlantic population. *Canadian Journal of Fisheries and Aquatic Sciences*, 63: 670-82.

Carvalho, F., Ahrens, R., Murie, D., Bigelow, K., Aires-Da-Silva, A., Maunder, M.N., Hazin, F. 2015. Using pop-up satellite archival tags to inform selectivity in fisheries stock assessment models: a case study for the blue shark in the South Atlantic Ocean. *ICES Journal of Marine Science, 72*: 1715-30.

Carvalho, G.R., Hauser, L. 1995. Molecular genetics and the stock concept in fisheries. *Molecular Genetics in Fisheries* (eds. Carvalho, G. R. & Pitcher, T. J.): 55–79.

Chessel, D., Dufour, A.T. 2004. The *ade4* package – I: One-table methods. *R News, 4*: 5-10.

Chow, S., Okamoto, H., Miyabe, N., Hiramatsu, K., Barut, N. 2000. Genetic divergence between Atlantic and Indo-Pacific stocks of bigeye tuna (*Thunnus obesus*) and admixture around South Africa. *Molecular Ecology 9*:221–227.

Clarke, S.C., Harley, S.J., Hoyle, S.D., Rice, J.S. 2013. Population trends in Pacific oceanic sharks and the utility of regulations on shark finning. *Conservation Biology, 27*: 197-209.

Clarke, S.C., Magnussen, J.E., Abercrombie, D.L., McAllister, M.K., Shivji, M.S. 2006. Identification of shark species composition and proportion in the Hong Kong shark fin market based on molecular genetics and trade records. *Conservation Biology, 20*: 201-11.

Coelho, R., Mejuto, J., Domingo, A., Yokawa, K., Liu, K.M., Cortés, E., Romanov, E.V., da Silva, C., Hazin, F., Arocha, F., Mwilima, A.M., Bach, P., Ortiz de Zárate, V., Roche, W., Lino, P.G., García-Cortés, B., Ramos-Cartelle, A.M., Forselledo, R., Mas, F., Ohshimo, S., Courtney, D., Sabarros, P.S., Perez, B., Wogerbauer, C., Tsai, W.-P., Carvalho, F., Santos, M.N. 2017. Distribution patterns and population structure of the blue shark (*Prionace glauca*) in the Atlantic and Indian Oceans*. Fish and Fisheries,* *19*(1), 90–106. http://dx.doi.org/10.1111/faf.12238.

Compagno, L.J.V. 1984. FAO species catalogue. Vol. 4. Sharks of the world. An annotated and illustrated catalogue of shark species known to date. Part 2. Carcharhiniformes. *FAO Fish. Synop.* (125, Vol. 4, Part 2), 655 p. http://www.fao.org/3/ad123e/ad123e00.htm

Cortés, E., Domingo, A., Miller, P., Forselledo, R., Mas, F., Arocha, F., Campana, S., Coelho, R., Da Silva, C., Hazin, F.H.V., Holtzhausen, H., Keene, K., Lucena, F., Ramirez, K., Santos, M.N., Semba-Murakami, Y., Yokawa, K. 2015. Expanded ecological risk assessment of pelagic sharks caught in Atlantic pelagic longline fisheries. *Collective Volume of Scientific Papers ICCAT, 1*: 2637–88.

Cramer, J., Bertolino, A., Scott, G.P. 1997. Estimates of recent shark bycatch by U.S. vessels fishing for Atlantic tuna and tuna-like species. *ICCAT Working Document*., SCRS/97/58.

da Silva, C., Kerwath, S.E., Wilke, C.G., Meyer, M., Lamberth, S.J. 2010. First documented southern transatlantic migration of a blue shark *Prionace glauca* tagged off South Africa. *African Journal of Marine Science, 32*: 639-642.

Davey, J.W., Blaxter, M.L. 2010. RADSeq: next-generation population genetics. *Briefings in Functional Genomics, 9*: 416-23.

Dray, S., Dufour, A. 2007. The *ade4* package: Implementing the duality diagram for ecologists. *Journal of Statistical Software, 22*: 1-20.

Dray, S., Dufour, A., Chessel, D. 2007. The *ade4* Package – II: Two-table and K-table methods. R News, 7: 47–52.

Druon J-N., Campana, S., Vandeperre, F., Hazin, F.H.V., Bowlby, H., Coelho, R., Queiroz, N., Serena, F., Abascal, F., Damalas, D., Musyl, M., Lopez, J., Block, B., Afonso, P., Dewar, H., Sabarros, P.S., Finucci, B., Zanzi, A., Bach, P., Senina, I., Garibaldi, F., Sims, D.W., Navarro, J., Cermeño, P., Leone, A., Diez, G., Zapiain, M.T.C., Deflorio, M., Romanov, E.V., Jung, A., Lapinski, M., Francis, M.P., Hazin, H., Travassos, P. 2022. Global-Scale Environmental Niche and Habitat of Blue Shark (Prionace glauca) by Size and Sex: A Pivotal Step to Improving Stock Management. Frontiers in Marine Science, 9:828412.

Durand, J.-D., Collet, A., Chow, S., Guinand, B., Borsa, P. 2005. Nuclear and mitochondrial DNA markers indicate unidirectional gene flow of Indo-Pacific to Atlantic bigeye tuna (*Thunnus obesus*) populations, and their admixture off southern Africa. Marine Biology,

Estes, J.A., Terborgh, J., Brashares, J.S., Power, M.E., Berger, J., Bond, W.J., Carpenter, S.R., Essington, T.E., Holt, R.D., Jackson, J.B.C., Marquis, R.J., Oksanen, L., Oksanen, T., Paine, R.T., Pikitch, E. K., Ripple, W.J., Sandin, S.A., Scheffer, M., Schoener, T.W., Shurin, J.B., Sinclair, A.R.E., Soulé, M.E., Virtanen, R., and Wardle, D.A. 2011. Trophic downgrading of planet Earth. *Science*, *333*: 301-306.

Excoffier, L., Foll, M. 2011. fastsimcoal: a continuous-time coalescent simulator of genomic diversity under arbitrarily complex evolutionary scenarios. *Bioinformatics, 27*(9):1332–1334.

doi: 10.1093/bioinformatics/btr124.

Excoffier, L., Smouse, P.E., Quattro, J.M. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics, 131*: 479‐91.

Ferretti, F., Myers, R.A., Serena, F., Lotze, H.K. 2008. Loss of large predatory sharks from the Mediterranean Sea. *Conservation Biology, 22*: 952-964.

Foll, M., Gaggiotti, O.E. 2008. A genome scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics, 180*: 977-93.

Foster, S.D. 2020. stockR: Identifying Stocks inGenetic Data. R package version, 1.0.74.

Foster, S.D., Feutry, P., Grewe, P.M., Berry, O., Hui, F.K.C., Campbell, R.D. 2018. Reliably discriminating stock structure with genetic markers: Mixture models with robust and fast computation. *Molecular Ecology Resources, 18*: 1310-25.

Fujinami Yuki, Semba Yasuko, Okamoto Hiroaki, Ohshimo Seiji, Tanaka Sho. 2017. Reproductive biology of the blue shark (Prionace glauca) in the western North Pacific Ocean. *Marine and Freshwater Research*, 68, 2018-2027.

Fraisse, C., Roux, C., Gagnaire, P.A., Romiguier, J., Faivre, N., Welch, J.J., Bierne, N. 2018. The divergence history of European blue mussel species reconstructed from Approximate Bayesian Computation: the effects of sequencing techniques and sampling strategies. *Peerj, 6*: e5198.

Galván-Magaña, F., Castillo-Geniz, J.L., Hoyos-Padilla, M., Ketchum, J., Klimley, A.P., Ramírez-Amaro, S., Torres-Rojas, Y.E., and Tovar-Ávila, J. 2019. Shark ecology, the role of the apex predator and current conservation status in *Sharks in Mexico: Research and Conservation Part A*, Chapter Three, 61–114. DOI:10.1016/bs.amb.2019.08.005.

Georges, A., Gruber, B., Pauly, G.B., White, D., Adams, M., Young, M. J., Kilian, A., Zhang, X., Shaffer, H.B., Unmack, P.J. 2018. Genome wide SNP markers breathe new life into phylogeography and species delimitation for the problematic short-necked turtles (Chelidae: *Emydura*) of eastern Australia. *Mol. Ecol., 27*: 5195-213.

Gosselin, T. 2018. radiator: RADseq data exploration, manipulation and visualization using R. In. R package version 0.0.11. Retrieved from https://github.com/thierrygosselin/radiator.

Gosselin, T., Lamothe, M., Devloo-Delva, F., Grewe, P. 2020. radiator: RADseq Data Exploration, Manipulation and Visualization using R. https://thierrygosselin.github.io/radiator/.

Gutenkunst, R. N., Hernandez, R. D., Williamson, S. H., & Bustamante, C. D. (2009). Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data*. Plos Genetics*, 5, e1000695.

Green, M.E., Appleyard, S.A., White, W., Tracey, S., Devloo-Delva, F., Ovenden, J.R. 2019. Novel multimarker comparisons address the genetic population structure of silvertip sharks (*Carcharhinus albimarginatus*). *Marine and Freshwater Research*. doi:10.1071/mf18296

Grewe, P., Grueger, C., Aquadro, C.F., Berminghmam, E. 1993. Mitochondrial DNA Variation among Lake Trout (*Salvelinus namaycush*) Strains Stocked into Lake Ontario. *J. Fish. Aquat. Sci. Can., 50*: 2397–403.

Grewe, P.M. et al. 2015. Evidence of discrete yellowfin tuna (*Thunnus albacares*) populations demands rethink of management for this globally important resource. *Sci. Rep.* 5, 16916; doi: 10.1038/srep16916.

Hara, Y., Yamaguchi, K., Onimaru, K., Kadota, M., Koyanagi, M., Keeley, S.D., Tatsumi, K., Tanaka, K., Motone, F., Kageyama, Y., Nozu, R., Adachi, N., Nishimura, O., Nakagawa, R., Tanegashima, C., Kiyatake, I., Matsumoto, R., Murakumo, K., Nishida, K., Terakita, A., Kuratani, S., Sato, K., Hyodo, S., Kuraku, S. 2018. Shark genomes provide insights into elasmobranch evolution and the origin of vertebrates. *Nat. Ecol. Evol., 2:* 1761-71

Hedgecock, D., Barber, P.H., Edmands, S. 2007. Genetic approaches to measuring connectivity. *Oceanography, 20*: 70-79.

Hernández-Aguilar, S.B., Escobar-Sánchez, O., Galván-Magaña, F., Abitia-Cárdenas, L. A. 2015. Trophic ecology of the blue shark (*Prionace glauca*) based on stable isotopes (δ13C and δ15N) and stomach content. *Journal of the Marine Biological Association of the United Kingdom*. 96 (7*)*.

Hughes, B. B., Eby, R., Van Dyke, E., Tinker, M.T., Marks, C.I., Johnson, K.S., Wasson, K. 2013. Recovery of a top predator mediates negative eutrophic effects on seagrass. *Proceedings of the National Academy of Sciences*, 110 (38): 15313-15318.

ICCAT, International Commission for the conservation of Atlantic Tunas - 2015. Report of the 2015 ICCAT Blue shark stock assessment session. Retrieved from www.iccat.int. retrieved from www.iccat.in on March 10th 2020.

ICCAT, International Commission for the conservation of Atlantic Tunas - 2019. Management Measures for the Conservation of the north Atlantic Blue Shark caught in association with ICCAT fisheries. In, edited by retrieved from www.iccat.in on March 10th 2020.

IOTC, Indian Ocean Tuna Commission 2007. Compilation of information on blue shark (*Prionace glauca*), silky shark (*Carcharhinus falciformis*), Oceanic whitetip shark (*Carcharhinus longimanus*), Scalloped hammerhead (*Sphyrna lewini*) and shortfin mako (*Isurus oxyrinchus*) in the Indian Ocean. A working paper ed. GTEPA.

IOTC, Indian Ocean Tuna Commission 2017. Stock assessment blue shark (*Prionace glauca*) in the Indian Ocean using Stock Synthesis. In, edited by Working Party on Ecosystems and Bycatch (WPEB).

ISC, International Scientific Committee for Tuna and Tuna-like Species in the north Pacific Ocean. 2018. Stock assessment and future projections of Blue Shark in the North Pacific Ocean through 2015. SC14-SA-IP-13. https://meetings.wcpfc.int/node/10722

Jombart, T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics, 21*: 1403‐05.

Jombart, T., Devillard, S., Balloux, F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC* *Genetics, 11*: 94.

Jombart, T., Ahmed, I. 2011. *adegenet* 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics, 27*: 3070-71.

Jouganous, J., Long, W., Ragsdale, A. P., & Gravel, S. 2017. Inferring the joint demographic history of multiple populations: beyond the diffusion approximation. Genetics, 206(3), 1549-1567.

Kaiser, J., Schefuß, E., Lamy, F., Mohtadi, M., Hebbeln, D. 2008. Glacial to Holocene changes in sea surface temperature and coastal vegetation in north central Chile: high versus low latitude forcing. *Quat. Sci. Rev., 27*, 2064–2075.

Kamm, J., Terhorst, J., Durbin, R., and Song, Y.S. 2020. Efficiently inferring the demographic history of many populations with allele count data. *Journal of the American Statistical Association*, 115(531), 1472-1487.

Kamvar, Z.N., Tabima, J.F., Grünwald, N.J. 2014. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ, 2*: e281-e81.

Karl, S. A., Castro, A.L.F., Lopez, J.A., Charvet, P., Burgess, G.H. 2010. Phylogeography and conservation of the bull shark (*Carcharhinus leucas*) inferred from mitochondrial and microsatellite DNA. *Conservation Genetics, 12:* 371-82.

Keenan, K., McGinnity, P., Cross, T.F., Crozier, W.W., Prodöhl, P.A. 2013. diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution, 4*: 782-88.

Kilian, A., Wenzl, P., Huttner, E., Carling, J., Xia, L., Blois, H., Caig, V., Katarzyma, H-U., Jaccoud, D., Hopper, C., Aschenbrenner-Kilian, M., Evers, M., Peng, K., Cayla, C., Hok, P., Uszynski, G. 2012. Diversity arrays technology: A generic genome profiling technology on open platforms. *Methods in Molecular Biology (Methods and Protocols), 888*: 67–88.

King, J.R., Wetklo, M., Supernault, J., Taguchi, M., Yokawa, K., Sosa-Nishizaki, O., Withler, R.E. 2015. Genetic analysis of stock structure of blue shark (*Prionace glauca*) in the north Pacific Ocean. *Fisheries Research, 172*: 181-89.

Kohler, N.E, Turner, P.A. 2008. Stock structure of the blue shark (*Prionace glauca*) in the North Atlantic Ocean based on tagging data. *Sharks of the Open Ocean: Biology, Fisheries and Conservation*: 339-350.

Kohler, N.E., Turner, P.A., Hoey, J.J., Natanson, L.J., Briggs, R. 2002. Tag and recapture data for three pelagic shark species: Blue Shark (*Prionace glauca*), Shortfin Mako (*Isurus oxyrinchus*), and Porbeagle (*Lamna nasus*) in the North Atlantic Ocean. *Collective Volume of Scientific Papers, 54*: 1231-60.

Leduc, G., Schneider, R., Kim, J.H., Lohmann, G. 2010. Holocene and Eemian sea surface temperature trends as revealed by alkenone and Mg/Ca paleothermometry. *Quat. Sci. Rev., 29*, 989–1004.

Leone, A., Urso, I., Damalas, D., Martinsohn, J., Zanzi, A., Mariani, S., Sperone, E., Micarelli, P., Garibaldi, F., Megalofonou, P., Bargelloni, L., Franch, R., Macias, D., Prodohl, P., Fitzpatrick, S., Stagioni, M., Tinti, F., Cariani, A. 2017. Genetic differentiation and phylogeography of Mediterranean-North Eastern Atlantic blue shark (*Prionace glauca*, L. 1758) using mitochondrial DNA: panmixia or complex stock structure?. *Peerj, 5*: 18.

Leslie, M.S., Morin, P.A. 2016. Using genome-wide SNPs to detect structure in high-diversity and low-divergence populations of severely impacted Eastern Tropical Pacific spinner (Stenella longirostris) and pantropical spotted dolphins (*S. attenuata*). *Frontiers in Marine Science, 3*:253. DOI: 10.3389/fmars.2016.00253.

Li, W.W., Dai, X.J., Zhu, J.F., Tian, S.Q, He, S., Wu, F. 2017. Genetic differentiation in blue shark, *Prionace glauca*, from the central Pacific Ocean, as inferred by mitochondrial cytochrome b region. *Mitochondrial DNA Part A, 28*: 575-78.

Luu, K., Bazin, E., Blum, M.G.B. 2017. pcadapt: an R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources, 17*: 67-77.

Mamoozadeh, N.R., Graves, J.E., McDowell, J.R. 2020. Genome-wide SNPs resolve spatiotemporal patterns of connectivity within striped marlin (*Kajikia audax*), a broadly distributed and highly migratory pelagic species. *Evolutionary Applications, 13*:677–698. https://doi.org/10.1111/eva.12892

Masson-Delmotte, V., Schulz, M., Abe-Ouchi, A., Beer, J., Ganopolski, A., González Rouco, J.F., Jansen, E., Lambeck, K., Luterbacher, J., Naish, T., Osborn, T., Otto-Bliesner, B., Quinn, T., Ramesh, R., Rojas, M., Shao, X., Timmermann, A. 2013. Information from Paleoclimate Archives. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Matsunaga, H., Nakano, H. 1999. Species composition and CPUE of pelagic sharks caught by Japanese longline research and training vessels in the Pacific Ocean. *Fisheries Science, 65*: 16-22.

Maxwell, S.M., Scales, K.L., Bograd, S.J., Briscoe, D.K., Dewar, H., Hazen, E.L., Lewison, R.L., Welch, H., Crowder, L.B. 2019. Seasonal spatial segregation in blue sharks (*Prionace glauca*) by sex and size class in the Northeast Pacific Ocean. *Diversity and Distributions, 25*: 1304-17.

Megalofonou, P., Yannopoulos, C., Damalas, D., de Metrio, G., Deflorio, M., de la Serna, J.M. 2005a. Incidental catch and estimated discards of pelagic sharks from the swordfish and tuna fisheries in the Mediterranean Sea. *Fishery Bulletin*, 103:620–634.

Megalofonou, P., Dimitris D., de Metrio, G. 2005b. Size, age and sexual maturity of the blue shark, Prionace glauca, in the Mediterranean Sea. *Biology*. CM 2005/N:09.

Mejuto, J., Garcia-Cortés, B. 2005. Reproductive and distribution parameters of the Blue Shark *Prionace glauca*, on the basis of on-board observations at sea in the Atlantic, Indian and Pacific oceans. *Col. Vol. Sci. Pap. ICCAT,* 58(3):951-973.

Momigliano, P., Harcourt, R., Robbins, W.D., Jaiteh, V., Mahardika, G.N., Sembiring, A., Stow, A. 2017. Genetic structure and signatures of selection in grey reef sharks (*Carcharhinus amblyrhynchos*). *Heredity (Edinb), 119*: 142-53.

Mullins, R.B., McKeown, N.J, Sauer, W.H.H, Shaw, P.W. 2018. Genomic analysis reveals multiple mismatches between biological and management units in yellowfin tuna (*Thunnus albacares*). *ICES Journal of Marine Science, 75*: 2145-52.

Nakano, H., Clarke, S. 2005. Standardized CPUE for blue sharks caught by the Japanese longline fishery in the Atlantic Ocean, 1971–2003. *ICCAT Collective Volume of Scientific Papers, 58*: 1127-34.

Nakano, H., Seki, M.P. 2003. Synopsis of biological data on the blue shark, *Prionace glauca* Linnaeus. *Bulletin-fisheries research agency Japan*: 18-55.

Nakano, H., Stevens, J.D. 2008. The Biology and Ecology of the Blue Shark, Prionace Glauca. In *Sharks of the Open Ocean* (Blackwell Publishing Ltd.).

Nikolic, N., Devloo-Delva, F., Bailleul, B., Noskova, E., Rougeux, C., Liautard-Haag, C., Hassan, M., Marie, A., Borsa, P., Feutry, P., Grewe, P., Davies, C., Farley, J., Fernando, D., Biton Porsmoguer, S., Poisson, F., Parker, D., Aulich, J., Lansdell, M., Marsac, F., Arnaud-Haond, S. 2020a. Genome scans discriminate independent populations of the blue shark *Prionace glauca*. IOTC-2020-WPEB16-14.

Nikolic, N., Montes, I., Lalire, M., Puech, A., Bodin, N., Arnaud-Haond, S., Kerwath, S., Corse, E., Gaspar, P., Hollanda, S., Bourjea, J., West, W., Bonhommeau, S. 2020b. Connectivity and population structure of albacore tuna across southeast Atlantic and southwest Indian Oceans inferred from multidisciplinary methodology. *Scientific Reports*, 10:15657.

Noskova, E., Ulyantsev, V., Koepfli, K.P., O'Brien, S.J., Dobrynin, P. 2020. GADMA: Genetic algorithm for inferring demographic history of multiple populations from allele frequency spectrum data. *GigaScience*, 9.

Olsen, J., Anderson, N.J., Knudsen, M.F. 2012. Variability of the North Atlantic Oscillation over the past 5,200 years. *Nature Geosci., 5*:808–812.

Ovenden, J. R., Kashiwagi, T., Broderick, D., Giles, J., Salini, J. 2009. The extent of population genetic subdivision differs among four co-distributed shark species in the Indo-Australian archipelago. *BMC Evol Biol, 9:* 40.

Panayiotou, N., Biton-Porsmoguer, S., Moutopoulos, D.K., Lloret, J. 2020. Offshore recreational fisheries of large vulnerable sharks and teleost fish in the Mediterranean Sea: First information on species caught. *Mediterranean Marine Science*, 21(1), 222-227. DOI: 10.12681/mms.21938

Paradis, E. 2010. pegas: an R package for population genetics with an integrated–modular approach. *Bioinformatics, 26*: 419-20.

Pazmiño, D.A., Maes, G.E., Simpfendorfer, C.A., van Herwerden, L. 2017. Genome-wide SNPs reveal small scale conservation units in the highly vagile Galápagos shark (*Carcharhinus galapagensis*). *Conservation Genetics, 18*: 1151–1163.

Pecoraro, C., Babbucci, M., Franch, R., Rico, C., Papetti, C., Chassot, E.l, Bodin, N., Cariani, A., Bargelloni, L., Fausto, T. 2018. The population genomics of yellowfin tuna (*Thunnus albacares*) at global geographic scale challenges current stock delineation. *Scientific Reports, 8*: 13890.

Pembleton, L., Cogan, N., Forster, J. 2013. StAMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Molecular Ecology Resources*, *13*: 946-952. doi: [10.1111/1755-0998.12129](https://doi.org/10.1111/1755-0998.12129).

Queiroz, N., Humphries N.E., Noble, L.R., Santos, A.M., Sims, D.W. 2012. Spatial dynamics and expanded vertical niche of blue sharks in oceanographic fonts reveal habitat targets for conservation. *Plos One*, 7.

Rigby, C.L., Barreto, R., Carlson, J., Fernando, D., Fordham, S., Francis, M.P., Herman, K., Jabado, R.W., Liu, K.M., Marshall, A., Pacoureau, N., Romanov, E., Sherley, R.B. and Winker, H. 2019. *Prionace glauca*. *The IUCN Red List of Threatened Species, 2019*: e.T39381A2915850. https://dx.doi.org/10.2305/IUCN.UK.2019-3.RLTS.T39381A2915850.en.

Downloaded on 31 March 2021.

Rodríguez-Ezpeleta, N., Bradbury, I.R., Mendibil, I., Álvarez, P., Cotano, U., Irigoien, X. 2016. Population structure of Atlantic mackerel inferred from RAD-seq-derived SNP markers: effects of sequence clustering parameters and hierarchical SNP selection. *Molecular Ecology Resources, 16*: 991-1001.

Rougeux, C., Bernatchez, L., Gagnaire, PA. 2017. Modeling the multiple facets of speciation-with-gene-flow toward inferring the divergence history of Lake whitefish species pairs (*Coregonus clupeaformis*). *Genome Biology and Evolution, 9*: 2057-74.

Sansaloni, C., Petroli, C., Jaccoud, D., Carling, J., Detering, F., Grattapaglia, D., Kilian, A. 2011. Diversity Arrays Technology (DArT) and next-generation sequencing combined: genome-wide, high throughput, highly informative genotyping for molecular breeding of Eucalyptus. *BMC Proceedings, 5*: 54.

Shevenell, A.E., Ingalls, A.E., Domack, E.W., Kelly, C. 2011. Holocene Southern Ocean surface temperature variability west of the Antarctic Peninsula. *Nature, 470*, 250–254.

Simpfendorfer, C.A., Hueter, R.E., Bergman, U., Connett, S.M.H. 2002. Results of a fishery-independent survey for pelagic sharks in the western North Atlantic, 1977-1994. *Fisheries Research, 55*: 175-92.

Sims, D., Fowler, F.L., Ferretti, F., Stevens, J. 2016. *Prionace glauca*. *The IUCN Red List of Threatened Species, 2016*: e.T39381A16553182.

Sovic, M., Fries, A., Martin, S.A., Lisle, H. Gibbs. 2019. Genetic signatures of small effective population sizes and demographic declines in an endangered rattlesnake, *Sistrurus catenatus*. *Evol Appl, 12*: 664-78.

Stevens, J. 2009. *Prionace glauca*. *The IUCN Red List of Threatened Species, 2009*: e.T39381A10222811.

http://web.archive.org/web/20180412085123/http://www.iucnredlist.org/details/39381/0. Downloaded on 12 April 2021.

Stevens, J.D. 1990. Further results from a tagging study of pelagic sharks in the north-east Atlantic. *Journal of the Marine Biological Association of the United Kingdom, 70*: 707-20.

Taguchi, M., King, J.R., Wetklo, M., Withler, R.E., Yokawa, K. 2015. Population genetic structure and demographic history of Pacific blue sharks (Prionace glauca) inferred from mitochondrial DNA analysis. *Marine and Freshwater Research, 66*: 267-75.

Vandeperre, F., Aires-da-Silva, A., Fontes, J., Santos, M., Santos, R.S., Afonso, P. 2014. Movements of Blue Sharks (Prionace glauca) across Their Life History. *Plos One, 9*: 14.

Verissimo, A., Sampaio, I., McDowell, J.R., Alexandrino, P., Mucientes, G., Queiroz, N., da Silva, C., Jones, C.S., Noble, L.R. 2017. World without borders: genetic population structure of a highly migratory marine predator, the blue shark (Prionace glauca). *Ecology and Evolution, 7*: 4768-81.

Waples, R.S. 1998. Separating the wheat from the chaff: Patterns of genetic differentiation in high gene flow species. *Journal of Heredity, 89*: 438-50.

Waples, R.S., Gaggiotti, O.E. 2006. What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology, 15*: 1419-39.

Waples, R.S., Punt, A.E., Cope, J.M. 2008. Integrating genetic data into management of marine resources: how can we do it better?. *Fish and Fisheries, 9*: 423–49.

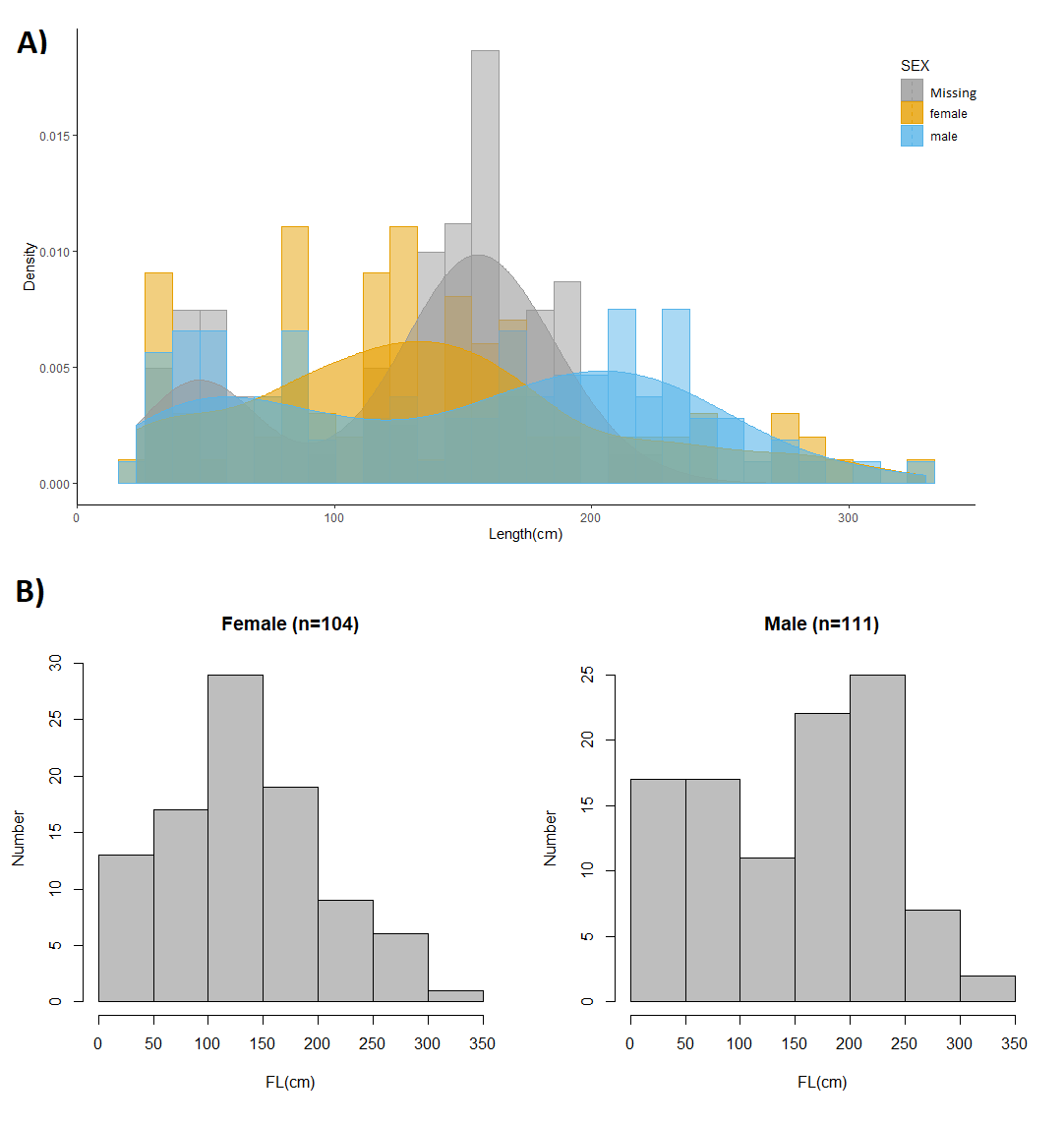
West, G., Stevens, J., Basson, M. 2004. Assessment of Blue Shark population status in the Western South Pacific. AFMA projectR01/1157. *CSIRO Marine Research*, Hobart, Tasmania, Australia, 139 pp.

Whitlock, M.C., Lotterhos, K.E., Bronstein, L. 2015. Reliable Detection of Loci Responsible for Local Adaptation: Inference of a Null Model through Trimming the Distribution of FST. *The American Naturalist, 186*: S24-S36.

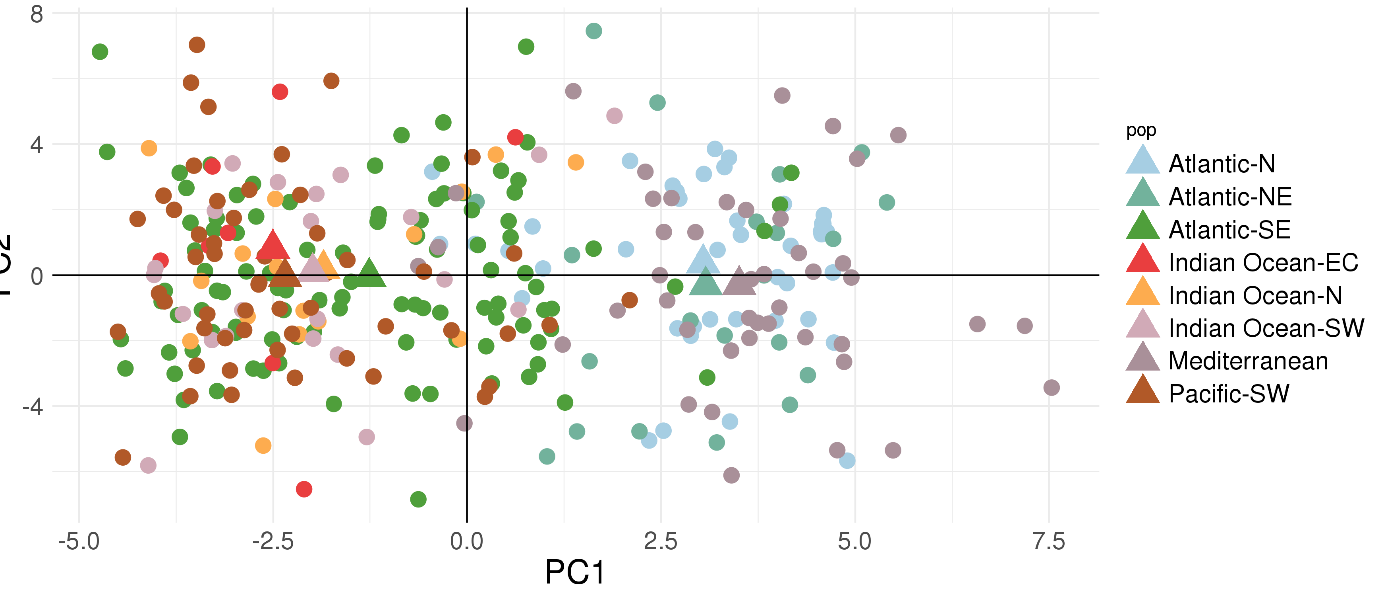
**Supplementary Material**

**S1.** R Markdown report of DarTseq data processing and population genetic analyses on R 4.1.0.

See the PDF file.



**S2.** Distribution of fork length (FL, in cm) per sex. Male (blue), female (yellow) and undetermined sex (missing, total of 149 individuals) (grey)) (A). Distribution of the total length (FL, in cm) for females and males (B).



**S3.** Principal Component Analysis on allelic frequencies on blue shark with 37,655 SNPs and 312 individuals on the eight geographic regions (DATA 4).

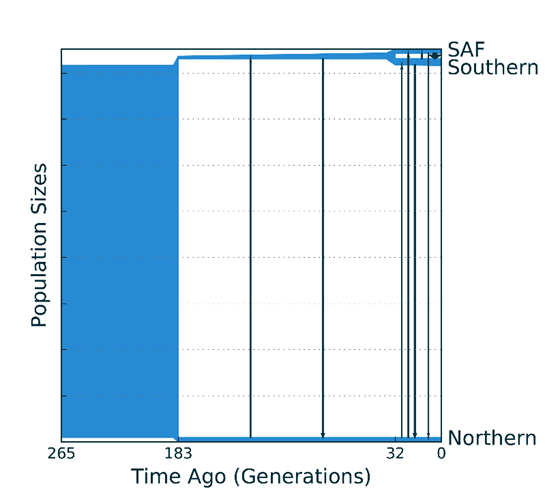
**S4.** Top three best demographic models from 50 independent demographic inference runs for the 2-population model (including northern and southern populations). Runs are sorted by the value of log-likelihood (logLL). See Table 4 for the variable names.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | logLL | Nanc | Nanc2 | Dyn(Nanc->Nanc2) | Nsou0 | Nsou | Dyn(Nsou0->Nsou) | Nnor0 | Nnor | Dyn(Nnor0->Nnor) | Msou->nor | Msou->nor | Tanc | Tsplit |
| Units |  | ind | ind | Dynamic | ind | ind | Dynamic | ind | ind | Dynamic | frac per gen | frac per gen | years | years |
| Model #1 | -2054.75 | 30422 | 176072 | Linear | 4220 | 6356 | Linear | NA | 5255 | Constant size | 0.0168 | 0.0004 | 1171152 | 5004 |
| Model #2 | -3239.55 | 76659 | 767 | Exponential | NA | 4 | Constant size | 1 | 31 | Exponential | 31.5 | 0.00003 | 653 | 0.2 |
| Model #3 | -3241.37 | 76651 | 767 | Sudden change | 7 | 18396 | Exponential | 767 | 1534 | Exponential | 0.0001 | 0.65 | 168 | 9 |

**S5.** Detailed comments on the 3-population demographic inference:

The best model with three populations was obtained with the First tree model (i.e., with SAF population originating from the Southern cluster). Estimates from the 3-population model with 50x50x50 JAFS displayed similar demographic qualitative trends than 2-population scenario (i.e., including only northern and southern populations), in spite of showing slightly more contrasted estimated values when compared to the 2-population 50x50 JAFS model (and also showing a much lower likelihood, in which it could be considered as less efficient or representative than the 2-populations model).

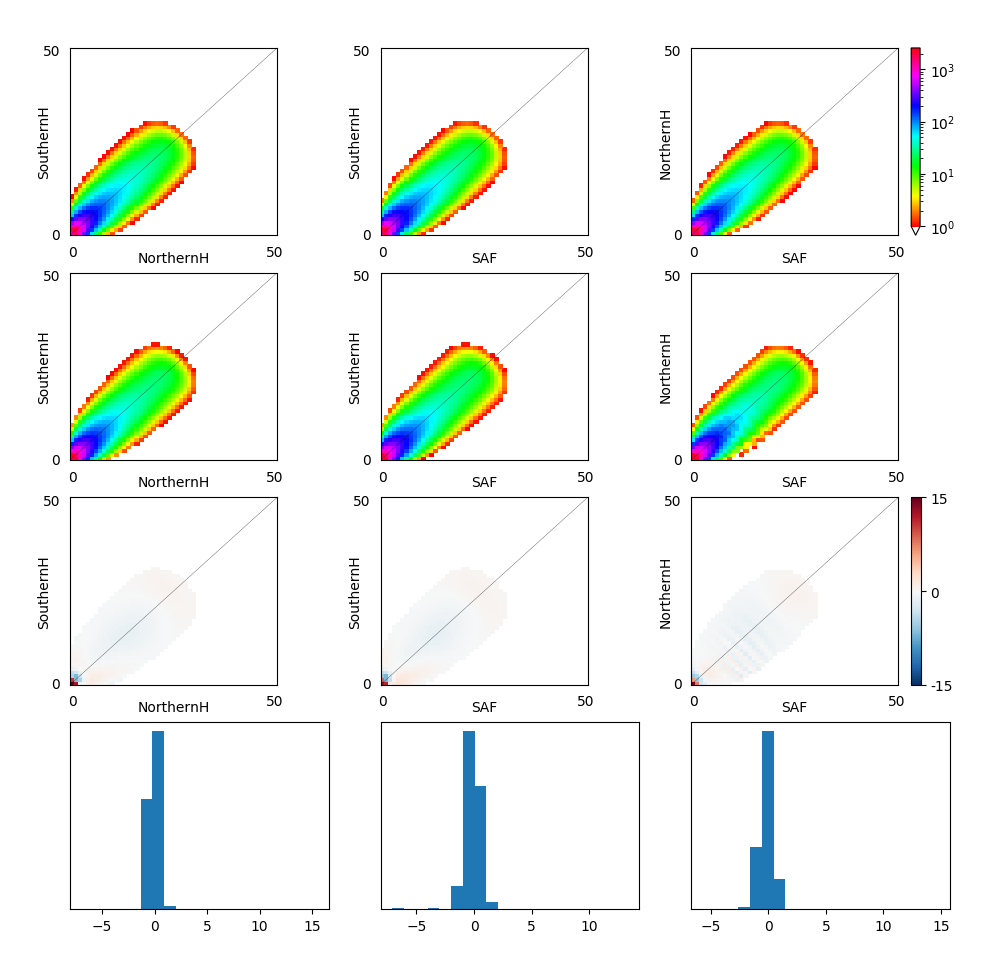
Overall, the 3-population model using 50x50x50 JAFS suggested the following. First, the ancestral population growth in its size linearly began in the same period than with the 2-population model (Figure 4). The split of the ancestral population from this model is later (ab. 183 generations ago, S7 and S8) forming the two new northern and southern sub-populations (S7). The northern population had a constant size and the southern population experienced linear growth after the split (S8), as in the previous 2-population scenario (Table 4). About 32 generations ago (corresponding to ab. 288 years ago [259-314 years ago]) Southern African (SAF) subpopulation would have split from the Southern population (S8). Two migration rates (SAF -> northern; northern -> southern) are very small and close to zero (S8). Otherwise, migration rate from SAF to Southern population was estimated the strongest (S7).



**S6.** Schematic plot of the best demographic history for the 3-population model (northern, southern and South Africa (SAF)) from now to the past time split in term of number of generations (inferred from the 50x50x50 JAFS and the First tree model). The vertical arrows represent asymmetric migration after the split.

**S7.** Demographic parameters with 3-population models (including northern, southern, and South Africa (SAF) populations) inferred by *GADMA* from 50x50x50 JAFS. Demographic history for the 2-population model was used for inference: SAF (South Africa) population was added to the model according to one of two different trees (i.e., First and Second tree\*), and its additional parameters were inferred by *GADMA*. All other parameters were extracted from the 2-population model and fixed during demography reconstruction. Log-likelihood is presented for 50x50x50 JAFS and both tree models, with best values marked in bold. \*(First tree – Figure 3 picture a) the ancestral population is splitted into northern population and common population of SAF and southern that is splitted later. (Second tree – Figure 3 picture b) the ancestral population is splitted into southern population and common population of SAF and northern that is splitted later. CI below the values of first tree model (Confidence intervals).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **First tree model** | **Second tree model** |  | |
| Log-likelihood for 50x50x50 JAFS | **-10714.8** | -10890.8 |
| **Parameters** | **Values first tree** | **Values second tree** | **Units** | **Parametrs’ definition** |
| NA | 20,448  [15,976 – 21,733] | 28,064 | Individuals | Size of the ancestral population. |
| NA\_split | 165,077  [158,382 – 169,528] | 153,743 | Individuals | Size of the ancestral population after linear growth. |
| NNor0 | 1,920  [1,366 – 2,131] | 2,669 | Individuals | Size of northern population after split of ancestral population. |
| NNor | 1,920  [1,366 – 2,131] | 2,669 | Individuals | Size of northern population now. |
| NSou0 | 1,120  [937 – 1,239] | 3,835 | Individuals | Size of southern population after split of ancestral population. |
| NSou | 2,917  [2,054 – 3,258] | 1,044 | Individuals | Size of southern population now (after linear growth). |
| NSAF0 | 1,708  [938 – 2,232] | 479 | Individuals | Size of SAF population after split from either northern or southern population. |
| NSAF | 1,593  [888 – 1,821] | 570 | Individuals | Size of SAF population now (after exponential growth). |
| TA | 163,057  [157,567 – 176,887] | 135,519 | Number of generations | Time of the beginning of ancestral population linear growth |
| Tsplit | 183  [136 - 201] | 679 | Number of generations | Split of ancestral population into northern and southern populations. |
| TSAFsplit | 32  [24 - 35] | 449 | Number of generations | Split of SAF population either from northern or southern population. |
| MSou-Nor | 0.0588  [0.0396 - 0.1054] | 0.1027 | Proportion per generation | Migration rate from southern population to northern |
| MNor-Sou | ~0  [0 - 0] | ~0 | Proportion per generation | Migration rate from northern population to southern |
| MSou-SAF | 0.0019  [0.0016 - 0.0026] | 0.1178 | Proportion per generation | Migration rate from southern population to SAF |
| MSAF-Sou | 0.2245  [0.0342 - 0.5090] | 0.1781 | Proportion per generation | Migration rate from SAF population to southern |
| MNor-SAF | 0.0290  [0 - 0.0784] | 0.1317 | Proportion per generation | Migration rate from northern population to SAF |
| MSAF-Nor | ~0  [0 - 0] | 0.1781 | Proportion per generation | Migration rate from SAF population to northern |



**S8**. Comparison of the joint allele frequency spectrum used for demographic analysis and simulated JAFS for best demographic history. (Top, line 1) Pairwise projection of 50x50x50 JAFS for northern, southern and SAF populations. Demographic history was inferred for this spectrum. (Line 2) Pairwise projection of simulated under best demographic model (log-likelihood -10754.7) JAFS. (Line 3) Anscombe residuals of the projections on first line and second. (Bottom, line 4) The histogram of the Anscombe residuals.