SIOFA | APSOI
Southern Indian Ocean Fisheries Agreement
Accord relatif aux Pêches dans le Sud de l'Océan Indien

10th Annual Meeting of the Scientific Committee (SC10)

Concarneau, France, 17-26 March 2025

SC-10-38

Stock structure of Patagonian toothfish (*Dissostichus eleginoides*) (SIOFA SER2022-TOP2)

Company for Open Ocean Observations and Logging (COOOL)

Document type	Working paper ✓ Information paper ☐
Distribution	Public ✓ Restricted [1] ☐ Closed session document [2] ☐

Abstract

This paper presents the results of the SER2022-TOP2 'Stock structure of Patagonian toothfish (Dissostichus eleginoides)'. The SER2022-TOP2 project aimed to design a genetic stock discrimination project to understand the stock structure of Patagonian toothfish in the SIOFA Area, including linkages to Patagonian toothfish in the CCAMLR Convention Area in order to advise on the delineation of management units. In the SER2022-TOP1 sister project, sampling was recommended to be undertaken in major fishing areas with flat bottoms (<0.2 radians) of <2000 m depth during the November to March presumed spawning season. A total of 251 samples from SIR (n = 65), DCR (n = 65), WR (n = 34), CR (n = 24), KER (n = 27), PEMI (n = 35) and SIR (n = 1) produced a combined dataset consisting of > 59 000 SNP loci. The bioinformatic filtering resulted in a combined dataset of 242 samples and > 2 700 SNPs ready for downstream analyses. DAPC clustering analyses of the genetic dataset identified the presence of one population cluster, i.e., a single panmictic population of D. eleginoides in the southwest Indian Ocean, with additional analysis supporting the lack of genetic structuring. Simple potential habitat distribution, informed by key environmental parameters indicate continuous potential habitat across CCAMLR and SIOFA boundaries and north of the current DCR management unit into SIR. Tagging results and data review further support the likelihood of mixing across the SIOFA/CCAMLR boundary.

[1] Restricted documents may contain confidential information. Please do not distribute restricted documents in any form without the explicit permission of the SIOFA Secretariat and the data owner(s)/provider(s).

[2] Documents available only to members invited to closed sessions.

Recommendations

Considering a single panmictic population using the potential habitat displayed, we give an adaptive series of recommendations for SIOFA management units of Patagonian toothfish including,

1) expanding the current DCR management unit northward to include the SIR potential habitat, fishing grounds and Coral Point VME,

2) harmonizing management measures in existing units, and

3) developing a joint management framework with CCAMLR.

Further genetic studies are recommended to support fisheries monitoring, including,

1) a wider population discrimination study for the Southern Ocean

2) a Close-Kin Mark-Recapture (CKMR) pilot project

3) a pilot project to determine age and sex using epigenetics to explore a potential alternative for determining age and sex, which could streamline demographic analysis, and also provide key information for stock assessment.

4) genetic projects should be supported by standard genetic sampling across the SIOFA and CCAMLR observer programs.

Stock structure of Patagonian toothfish (*Dissostichus eleginoides*) (SIOFA SER2022-TOP2)

Funded by the European Union

Anne-Elise NIEBLAS, Dominique COWART, Thomas CHEVRIER, Jérémie CHANUT

Company for Open Ocean Observations and Logging (COOOL)

Table of Contents

Executive summary

Stock structure discrimination is an important step in understanding the population dynamics of a stock, especially in cases where stocks are jointly managed by several countries or where multiple RFMOs may have overlapping mandates on a stock, as is the case with the Patagonian toothfish. The SER2022-TOP2 project aimed to design a genetic stock discrimination project to help understand the stock structure of Patagonian toothfish in the SIOFA Area, including linkages to Patagonian toothfish in the CCAMLR Convention Area. A sister project, SER2022-TOP1 reviewed the literature and existing data held by SIOFA and proposed an informed experimental design and sampling protocol for the genetic discrimination of the toothfish stock in the SIOFA area as well as detailed protocols for the collection of samples and their laboratory processing. In SER2022-TOP1, three main fishing zones were detected from SIOFA catch data including the Del Cano Rise (DCR), South Indian Ridge (SIR), and William's Ridge (WR) and we recommended that samples be collected from the three fishing zones from November to March in flat areas (<0.2 radians) of <2000 m depth. We recommended that a dataset composed of Single Nucleotide Polymorphisms (SNPs) loci be generated for *D. eleginoides*. To do so, we recommended between 30-100 individuals be sampled per fishing zone (50% female and 50% male), noting that the TOP2 project budget was limited to analyzing about 30 samples per zone, but was since increased to allow for more individuals to be analyzed.

In SER2022-TOP2, we communicated a detailed sampling protocol to onboard observers from Spanish, Australian, French and Uruguayan vessels that had planned to fish the three fishing zones and CCAMLR zones over the austral summer and fall. A Spanish vessel fished in SIR and DCR while an Australian vessel in WR. The CCAMLR zone was fished by French vessels around Crozet (CR) and Kerguelen (KER) Islands EEZs, while the Uruguayan-flagged vessel fished in the Prince Edward and Marion Islands (PEMI) EEZ. Sampling kits were dispatched to observers in August 2023 and May 2024. In March 2024, the observer on the Spanish vessel was the first to return samples and metadata for 200 individuals from SIR and DCR (100 samples each region). A preliminary assessment of the metadata indicated that the recommended sampling strategy had been achieved for these samples. The Australian vessel experienced significant delays to fishing at WR, and therefore fished this area in February and March 2024, towards the end of the presumed spawning period. The vessel returned to port in May 2024 with 36 samples. The observer aboard one of the French vessels successfully prepared 24 samples and associated metadata from CR in June 2024, outside of the spawning season that was recommended in our protocol. In early July, another observer aboard another French vessel prepared 27 samples and associated metadata from KER. Also in early July, an observer aboard a Uruguayan vessel in PEMI prepared 35 samples. Returned fin clips from DCR, SIR, WR and CR were selected and packaged for sequencing at Diversity Arrays Technology (DArT PL), which we found to have a lower price per sample and faster turnaround time than the previous company quoted in SER2022-TOP1. The remaining samples from KER and PEMI consisted of fin clips and muscle tissue that was received in August 2024, prepared and shipped to DArT. These sample sets are currently undergoing sequencing with a planned deadline of late October 2024, and the data will be incorporated into the broader sample set whose results will be described in this report as an addendum later this year.

A total of 188 samples from SIR (n = 65), DCR (n = 65), WR (n = 34) and CR (n = 24) were originally sent for sequencing to produce a dataset consisting of > 68 000 SNP loci. For the addendum project, a total of 63 samples from KER (n = 27), PEMI (n = 35) and SIR (n = 1) were added to the full dataset to produce a combined dataset of 251 samples consisting of > 59 000 SNP loci. To ensure reliable, accurate and informative results, the dataset underwent a series of bioinformatic filtering steps that include applying thresholds for sequencing depth, minor allele frequencies (MAF) and other quality control metrics. This resulted in an initial dataset of 186 samples and > 2 000 SNPs, and for the addendum project, a combined dataset of 242 samples and > 2 700 SNPs ready for downstream analyses. DAPC clustering analyses of the genetic dataset identified the presence of one population cluster, with samples from different regions showing close groupings. Accompanying STRUCTURE analyses further support the lack of genetic structuring through the identification of similar admixture in individuals obtained from the different sampling regions. Further, F_{ST} values indicated low levels of genetic differentiation, suggesting high levels of gene flow between these regions. Overall, this dataset supports the presence of a single panmictic population of *D. eleginoides* in the southwest Indian Ocean and that the different fishing regions (SIR, DCR, WR,CR, KER and PEMI) are one population that traverses SIOFA and CCAMLR-managed areas. These results were updated from the SER2022-TOP2 final report to include samples collected from KER and PEMI in the present addendum report.

Drawing on information gathered in SER2022-TOP1, we developed a simple potential habitat distribution model, informed by key environmental parameters for the species, i.e. bathymetry, slope, and bottom temperature. Potential habitat maps for adults and juveniles indicate continuous potential habitat across CCAMLR and SIOFA boundaries, including EEZs. Potential habitat extends north of the current DCR management unit into SIR, and east of the WR management unit where exploratory fishing has previously occurred. Tagging results further support the likelihood of mixing across the SIOFA/CCAMLR boundary.

Based on the assumption of a single panmictic population, as indicated by the potential habitat displayed, we present an adaptive series of recommendations for SIOFA management units of Patagonian toothfish including 1) expanding the current DCR management unit northward to include the SIR potential habitat, 2) harmonizing management measures in existing units, and 3) developing a joint management framework with CCAMLR.

Finally, we note the potential of genetic studies to support fisheries monitoring for such complex, but data-limited fisheries and we suggest that genetic sampling become a standardized part of the regular observer sampling scheme. These samples could be used to investigate the full extent of the Patagonian toothfish population in the Southern Ocean; they could support Close-Kin Mark Recapture (CKMR) studies to develop unbiased abundance indices; finally, they could also be used in epigenetics studies to determine age and sex, key inputs for stock assessment.

1. Introduction

Stock structure discrimination is an important step in understanding the population dynamics of a stock, especially in cases where stocks are jointly managed by several countries or where multiple RFMOs may have overlapping mandates on a stock, as with the Patagonian toothfish, *Dissostichus eleginoides* (Fig. 1). For example, should management measures such as catch limits be introduced, clear delineations of stock structure are key to appropriately allocating the resource, which can impact the state of the resource or the fishing activity (Avise, 1998). Furthermore, management can only be effective if the spatial scale of the measures matches that of the target population (Francis et al., 2007). Thus, a key step towards the effective management of important blue resources is to understand the population structure of the resource.

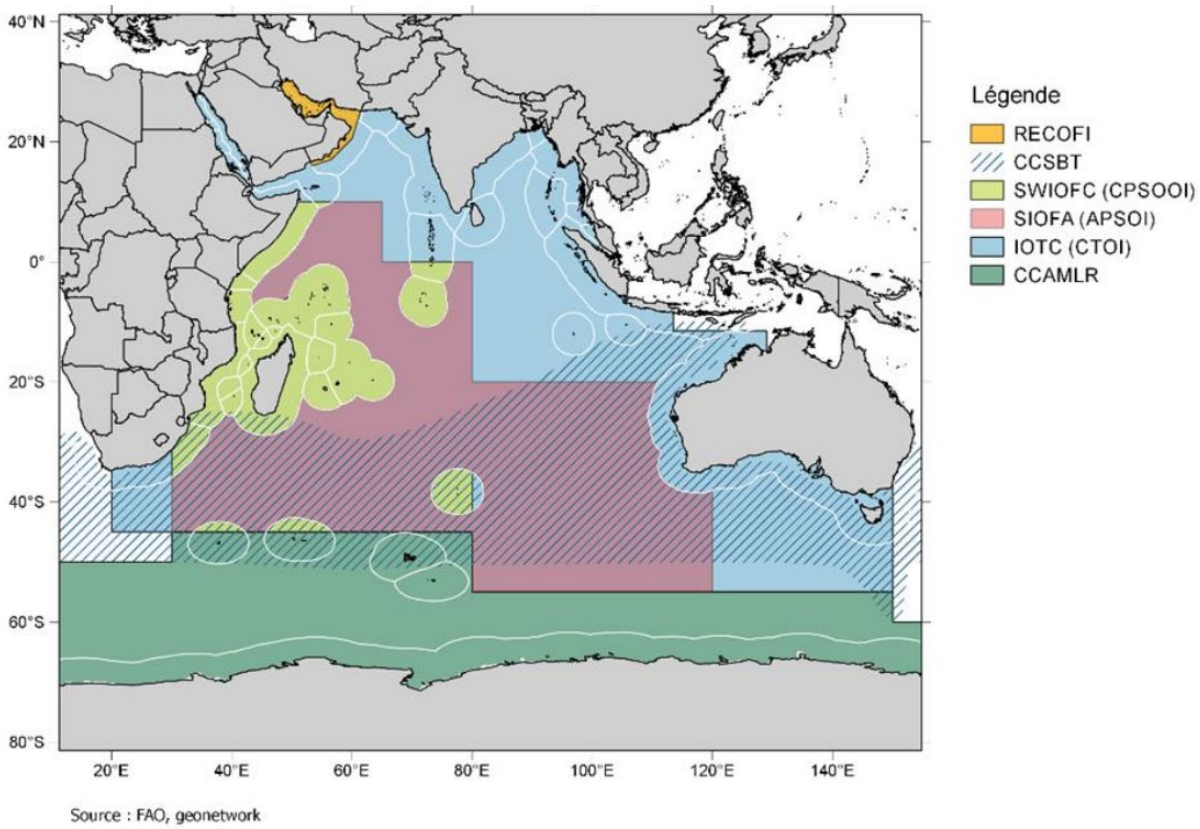


Fig. 1. Map of RFMOs and marine conservation organizations in the Indian and Southern Oceans. RECOFI = Regional Fisheries Commission, CCSBT = Commission for the Conservation of Southern Bluefin Tuna, SWIOFC = South West Indian Ocean Fisheries Commission, SIOFA = Southern Indian Ocean Fisheries Agreement, IOTC = Indian Ocean Fisheries Commission. IOTC = Indian Ocean Tuna Commission, CCAMLR = Commission for the Conservation of Antarctic Marine Living Resources.

Currently, Patagonian toothfish in the SIOFA area are not assessed, though stock assessments are performed for the adjacent CCAMLR regions and recently, SIOFA catches from Williams Ridge were

included in the Heard Island and McDonald Islands (HIMI) TOP assessment. However, stocks are likely straddling and a collaborative approach for toothfish assessments has been recommended.

Genetic studies of *D. eleginoides* around the Southern Ocean have shown that there is population structure between ocean basins, and some distinction within ocean basins separated by important oceanographic features (i.e. the Antarctic Polar Front), though little differentiation has yet been found using an array of genetic techniques in the Indian Ocean.

Genetic studies have identified populations existing on the Patagonian shelf (including the Falkland Islands: Smith and McVeagh 2000; Shaw et al. 2004; Rogers et al. 2006; Canales-Aguirre et al. 2018), the southern Atlantic Ocean (Shag Rocks, South Georgia and South Sandwich Islands: Appleyard et al. 2002; Shaw et al. 2004; Rogers et al. 2006), the western Indian Ocean (Crozet, Kerguelen, Heard, Prince Edward and Marion islands: Appleyard et al. 2004) and the southwest Pacific (Macquarie Island: Smith and McVeagh 2000; Appleyard et al. 2002) (Fig. 2). The gene flow restrictions between these regions have been attributed to the separation by deep ocean basins and the deep water trough between the South American and Antarctic peninsulas, which limit adult migration, as well as the Antarctic Polar Front (APF) (Fig. 2), a powerful jet of the Antarctic Circumpolar Current (ACC) that likely serves as a barrier to larval dispersal (Clark et al. 2005; Shaw et al. 2004; Rogers et al. 2006). In the western Indian Ocean, however, individuals may use bathymetric features such as ridges and seamounts as stepping stones to access other areas (Rogers et al. 2006). Additional genetic analyses of the species inhabiting waters of the southwest Indian Ocean may therefore provide further information on finer scale population structure of this species and how it relates to established fishing zones.

In the SER2022-TOP1 project, we reviewed the literature, catch, effort, biological and environmental data of Patagonian toothfish in the SIOFA region. We proposed a sampling strategy which aimed to sample individuals during their spawning season in the region (found to be in the austral summer between November to March), and in the habitats where most spawning individuals were found (i.e. flat areas between 800 m to < 2000 m) when and where we presumed the population would be the most mixed (Nieblas and Cowart 2023). After a review and approval by the SIOFA Scientific Advisory panel, we communicated this sampling strategy, as well as a detailed sampling protocol, to SIOFA observers, departing on Spanish-, Australian-, and French- flagged vessels over the austral summer period.

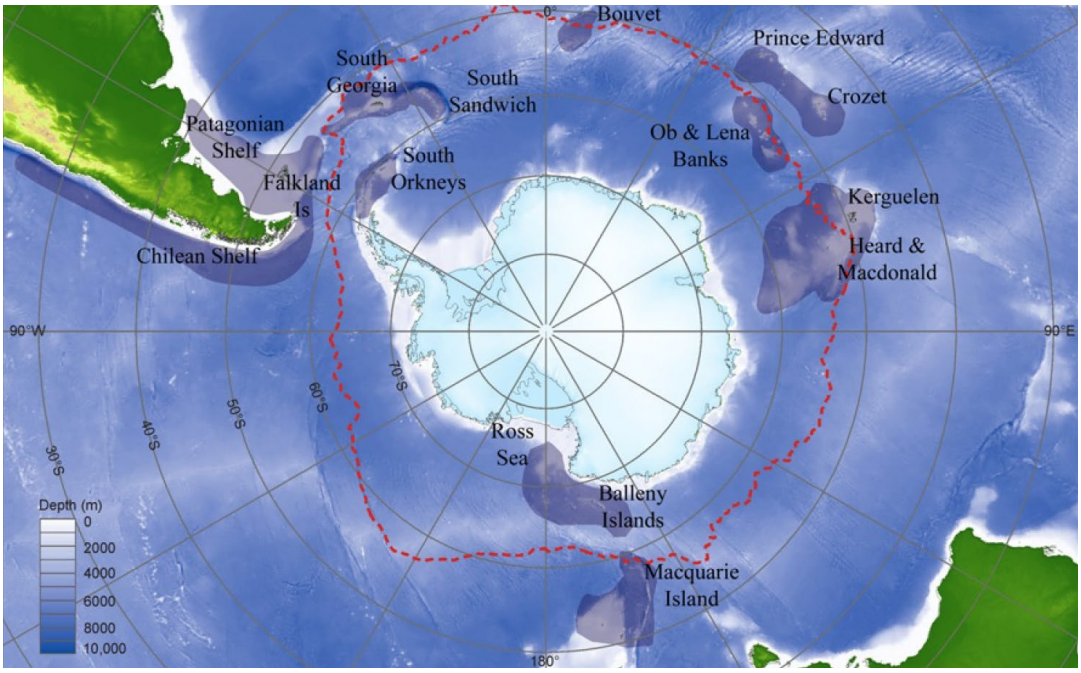


Fig. 2. Map showing the known distributions of the Patagonian toothfish (*Dissostichus eleginoides*) in gray shading with the Polar Front indicated by the red dotted line. Illustration is from Collins et al. 2010, The Patagonian toothfish: biology, ecology and fishery, Figure 4.2.

In SER2022-TOP1, the project tasks included a literature review, along with reviews of catch-effort, scientific observer, bathymetric, oceanographic, and other relevant environmental data. And the major tasks of SER2022-TOP2 were to include a genetic stock discrimination analysis; a description of the Patagonian toothfish population spatial structure in the SIOFA area; and a proposal of management units based on the stock structure. Here, we report on the progress of the SER2022-TOP2 project, including a summary of the data review, sampling strategy and protocols developed in the feasibility project SER2022-TOP1 (a detailed report of these completed tasks can be found at Nieblas and Cowart 2023), as well as progress in sample collection to undertake the major tasks of the SER2022-TOP2 project. We outline the next steps required for the genetic analyses along with an updated project timeline in Nieblas and Cowart (2024). We detail the retrieval of samples, their selection, packing and submission for sequencing analyses; we describe the results of the sequencing, potential habitat mapping based on key environmental parameters, and tagging analyses to provide details on population spatial structure and recommend management units based on these analyses.

2. Summary of initial biological data review

Based upon review of the catch, effort, and biological data held by SIOFA in SER2022-TOP1, the geographic sampling regions proposed were 1) South Indian Ridge (SIR) and 2) Del Cano Rise South (DCR), both located in SIOFA's statistical subregion 3b and 3) Williams Ridge (WR), located in SIOFA statistical subregion 7. These three sampling regions have been proposed as they represent established *D.*

eleginoides fishing grounds in SIOFA, which have regular fishing operations mostly by Spanish and Australian fleets.

Globally, both the SIR and DCR regions had substantial fishing effort starting from 2003. Fishing operations were recorded at WR beginning 2018. For all three regions, most fishing operations are set within the 800 - 1500 m depth range, and the fewest operations are set in the shallower range (< 800 m) with no shallow sets in DCR. (Fig. 3). Fishing in the deep range (> 1500 m) is considerable in DCR and SIR.

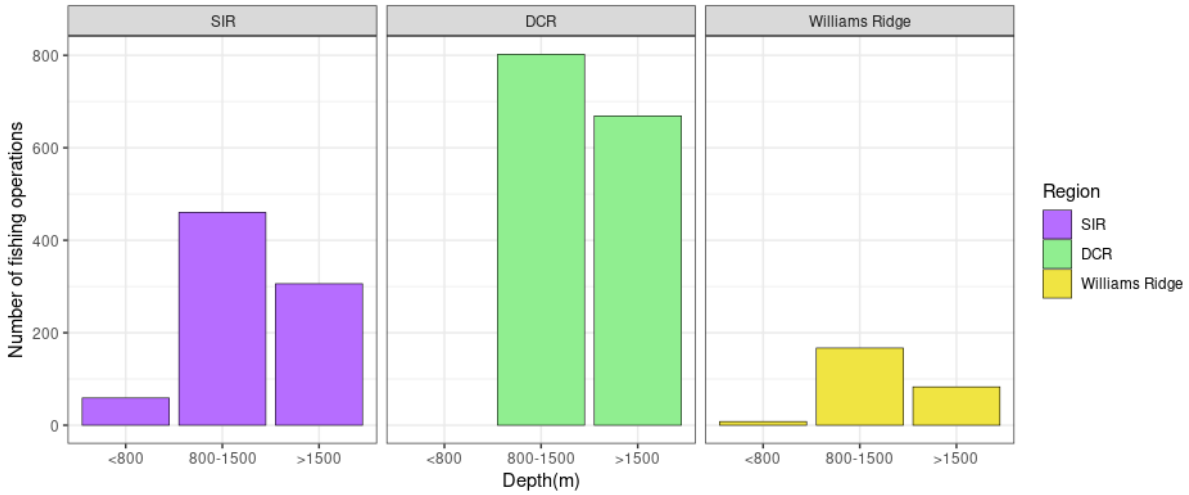


Fig. 3. Depth distribution of fishing operations for three regions.

Catch per Unit Effort (CPUE) was calculated as the catch in kg per 1000 hooks (kg/1000 hooks) for each region for the years with available records. Overall, CPUE was lower at SIR and DCR than at WR. Mid- and deep depths at SIR and DCR showed the highest effort in terms of both the number of operations and number of hooks, resulting in less variable and lower CPUEs than in WR. At WR, at depths < 800 m, there were only eight fishing operations, but more hooks set than the other regions, exceeding 10 000 hooks on average.

Females caught tended to be larger than males at each region, though males were most numerous in the catches from SIR and DCR. At all three regions, mid-sized to larger fish (70 - 100 cm) were the most frequently caught, and few larger individuals (> 130 cm) were caught at depths > 1500 m. Maturity observations, based on a 1-5 scale, indicated that spawning individuals (stage 4) were most often found between 800 to 2000 m in the austral summer months.

Several aspects of the data reviewed indicate a close connection between the CCAMLR and SIOFA areas. We find that the bathymetric features straddle areas north ofCCAMLR and DCR and WR of SIOFA, that would facilitate ontogenetic migration. Furthermore, spawning individuals appear to concentrate near the border of the SIOFA and CCAMLR areas in DCR and WR, indicating that the population straddles the two zones. Tagging studies indicate some movement between the Kerguelen Plateau and DCR, and between the Kerguelen Plateau and WR, and though rare, long distance displacements have been found.

The data review points to the conclusion that the fishing areas in SIOFA are likely fishing the same population as in the adjacent CCAMLR region.

The three fishing regions are distinctly defined by their bathymetry, with DCR being relatively flat throughout the management area, SIR defined by relatively shallow ridgelines separated by deep canyons, and WR extending from CCAMLR into SIOFA as a narrow, shallow ridge. Mixed layer depth was found to have patterns in line with deep winter convective mixing in the WR area that then also showed a peak in chlorophyll in October, as described by Song et al. 2016. Mesoscale features appear to be influential to local productivity for SIR. We found, however, that stable physical features such as bottom depth, slope, and bottom temperature are significantly related to both length and sex ratio, while more seasonal and temporary features (e.g. productivity patterns and mesoscale features) do not appear to be of influence.

3. Summary of sampling strategy, protocols and initial recommendations

In the SIOFA SER2022-TOP1 report, we defined a sampling strategy targeting mixed-sex spawning grounds and seasons to ensure the highest likelihood of sampling the fully-mixed population at each site and thus enabling us to genetically discriminate between populations in the three fishing zones (SIR, DCR, and WR). Based on this strategy, we recommended sampling from November to March in flat areas (< 0.2 radians) of > 800 m and < 2000 m depth. While samples from spawners would have the highest likelihood of capturing a single population, targeting the habitat and season where and when spawners occur should also lead to an increased probability of representative sampling for a single population.

Toomey et al. (2016) found evidence that toothfish in the southwest Indian Ocean may not be fully panmictic suggesting that a finer level of differentiation may exist in this region, which would require additional study to confirm. We recommended that a dataset composed of Single Nucleotide Polymorphism (SNP) loci should be generated for *D. eleginoides* in the southwest Indian Ocean, as hundreds to thousands of SNPs genotyped across many individuals could provide a clearer resolution of population dynamics (Morin et al. 2004). We requested observers collect between 30 to 100 samples per region, with as even a sex ratio as possible. We also requested that trip data (vessel, registration number, date, latitude and longitude of catch) and biometric data (weight, length, maturity, sex) be collected along with two fin clips per individual.

A sampling protocol was communicated to the onboard observers, and we also developed a sampling protocol and a laboratory protocol for preparing the samples for shipping to the initially selected sequencing company, as well as the shipping protocol for sending the samples between facilities (Nieblas and Cowart 2023).

4. Sample and data collection

4.1 Observer contact and delivery of samples

The SIOFA Scientific Advisory panel advised COOOL on the contact details of the vessels that were known to be fishing in the SIOFA convention area on toothfish, during the targeted sampling period (November to March, i.e. the presumed spawning period). This included a Spanish-flagged vessel, the *Ibsa Quinto*, and an Australian-flagged vessel, *Cape Arkona.* Further, we were advised on the vessels that fish in the CCAMLR areas, including two French flagged vessels, the *Ile de la Reunion II* and the *Saint Rose*, as well as a Uruguayan-flagged vessel, the *Ocean Azul*, operating in the Prince Edward and Marion Islands (PEMI) area. The Spanish vessel was planned to fish in South Indian Ridge (SIR) and Del Cano Rise (DCR), while the Australian vessel was planned to fish in the Williams Ridge (WR). The French vessels were scheduled to fish in the vicinity of Crozet (CR) and Kerguelen islands (KER) in May through July and the Uruguayan vessel was scheduled to fish around Prince Edward and Marion Island (PEMI) in late June 2024, well after the targeted sampling periods (Fig. 4).

The observer managers for each of the vessels were contacted to introduce them to the project. For Australia and South Africa, the COOOL team was advised to contact Observers@afma.gov.au and Capfish@mweb.co.za, and interacted with Justine Johnston at Australian Fisheries Management Authority, Rhys Arragio of Austral Fisheries, and Captain Emmanuel Le Roy. For Spain, the COOOL team contacted Roberto Sarralde Vizuete of IEO-CSIC, who assisted in the interaction with the Spanish observers. For South Africa, the COOOL team contacted Sobahle Somhlaba (Department of Agriculture, Forestry and Fisheries, South Africa) and Melanie Williamson of Capfish. The COOOL team organized sampling kits (labeled tubes, data sheets, printed and plasticized observer sampling protocols) for each boat with enough material for the collection of 100 duplicate samples for each of the regions (SIR, DCR, WR, CCAMLR).

4.2 Sample collection

Each individual fish sampled was measured to the nearest centimeter for total length and to the nearest gram for weight, while the location and date of capture, vessel name and registration number, sex, and maturity stage were all noted. Two sets of fin clips were sampled for each individual (i.e. two slices of the same fin were placed into separate tubes). Observers donned gloves prior to decontaminating workspaces using a 10% bleach solution, followed by wiping the area dry. This step was followed by thorough rinsing with 70% ethanol and wiping dry. Next, transfer pipettes were used to load approximately 0.5 - 1 ml of absolute ethanol into clean and pre-labeled 1.5 ml tubes. Cutting tools such as forceps, scalpels and scissors were also disinfected by rinsing with the 10% bleach solution, followed by rinsing with 70%

ethanol and wiped dry using a clean paper towel. For each fish, fin samples were carefully cut, then submerged in the ethanol of a prepared 1.5 ml tube. The tube was then capped tightly and placed in a cryobox that was stored on ice. Cutting tools were sterilized between the sampling of each fish. Once all samples had been collected, the cryoboxes were closed and secured with thick rubber bands looped over the box. The boxes were sealed in ziplock bags and placed into -20º C until ready for shipment. All spaces and tools were decontaminated once completed.

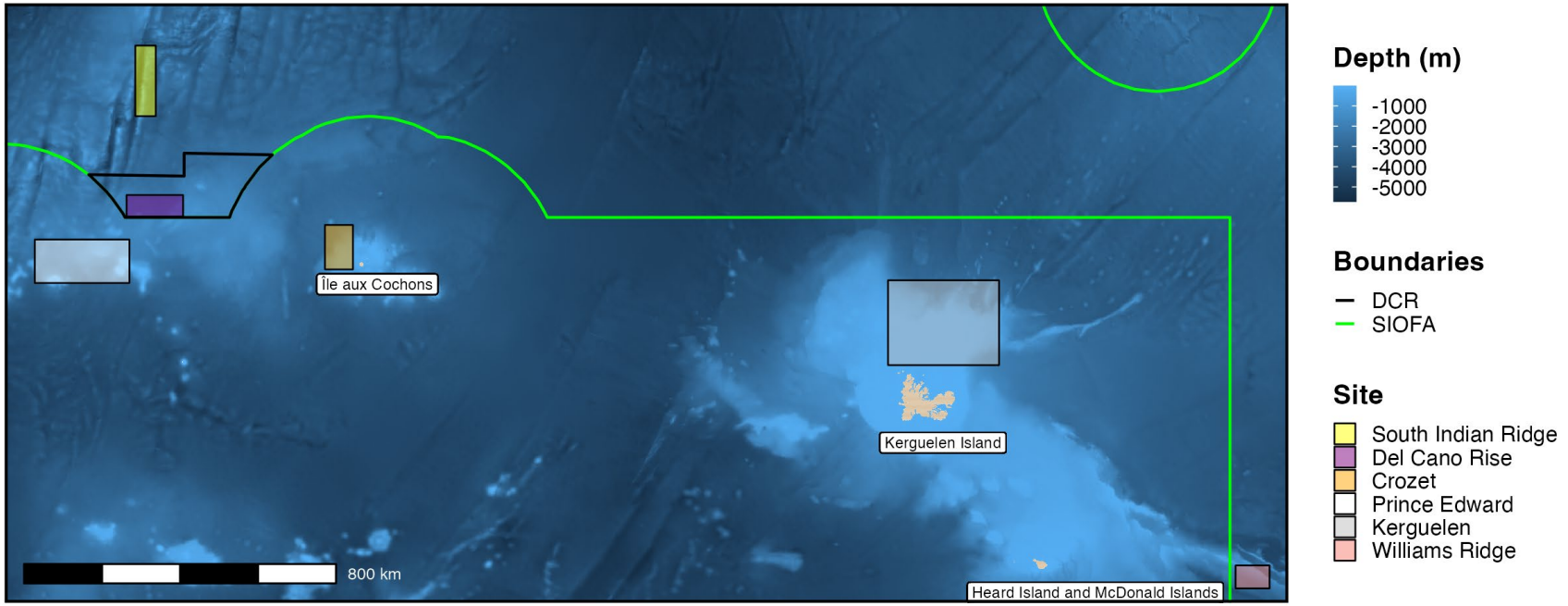


Fig. 4. Six sampling regions and their sites of collection: Prince Edward and Marion Island (PEMI), South Indian Ridge (SIR), Del Cano Rise (DCR), Crozet (CR), Kerguelen (KER) and Williams Ridge (WR). The green line marks the boundary of the SIOFA Area including the DCR management area, which is denoted by the black outline. The inset map in the upper right corner shows the entire SIOFA Area, as well as the red outline indicating the area covering all sampling regions. The scale bar may not accurately represent distances due to the Coordinate Reference System (CRS) EPSG:4326 used to generate the map with the aid of *ggplot2* and *sf* packages in R.

4.3 Sample summary and review

Samples were initially collected from 260 individuals that were retrieved from SIR, DCR, WR and CR (Fig. 4). Sampling at SIR and DCR occurred between two separate 8-day periods; from 20 to 28 October 2023, 100 individuals were sampled at SIR and from 14 - 22 November 2023, 100 individuals were sampled at DCR. At WR, 36 individuals were sampled over six days from 2 - 7 March 2024, while at CR, 24 individuals were sampled over four days from 31 May - 2 June 2024. Later in the year, samples from 27 individuals were obtained from three locations around KER from 1 - 4 July 2024, while 33 individuals were sampled from 15 locations near PEMI from 1 - 6 July 2024 (Fig. 4). As requested, metadata files contain the name and the registration number of the associated vessel, dates of capture and sampling coordinates as well as weight, total length (LT), sex and maturity stage of each individual sampled. For all sampling regions except PEMI, metadata also included the depth of capture.

We began by reviewing the metadata with regards to the length and depth distribution of the samples, as well as examining sex ratio and maturity of the catches. Two samples from WR were removed from consideration due to uncertainties in the metadata associated with these two samples, leaving a total of 318 samples to be assessed. Across all regions where depth was recorded, depths ranged from 800 – 1783 m and fish lengths from 51.0 – 167.0 cm. Both females and males were captured at each region and all maturity stages (1 – 5) were represented in the total dataset.

4.3.1 Length distributions

Lengths of individuals, based on total length (LT) ranged from 53.0 cm – 167.0 cm at SIR, 64.0 cm – 123.0 cm at DCR, 51.9 – 97.5 cm at WR, 51.6 – 95.7 at CR, 60.5 – 113.5 cm at KER and 51.0 – 162.0 at PEMI. Each region showed similar distributions, with overall total lengths skewing smaller at SIR, WR, and CR (Fig. 5). At SIR, individuals were obtained from all three depth ranges; however, the majority of the individuals were recovered from the mid-water depths (801 m to 1499 m) as requested. Individuals across all size classes were recovered from the mid-water depths and those caught at shallower depths were smaller (L_T < 100 cm, ≤ 800 m), while those caught deeper were larger (L_T > 90 cm, depth ≥ 1500 m). At DCR, individuals were only retrieved from depths > 800 m. Deep water individuals were more numerous, and while these individuals had a similar length distribution as mid-water individuals, there were more from the larger size classes (L_T > 100 cm) obtained from ≥ 1500 m. At WR and CR, all individuals were caught between 801 – 1499 m and skewed smaller than those retrieved from SIR and DCR at the same depth range (Fig. 5). At KER, all samples were retrieved from ≥ 1500 m and skewed smaller. Samples from PEMI had no depth recorded, but were measured across all size classes, with more individuals measuring between 70 – 74 cm (Fig. 5).

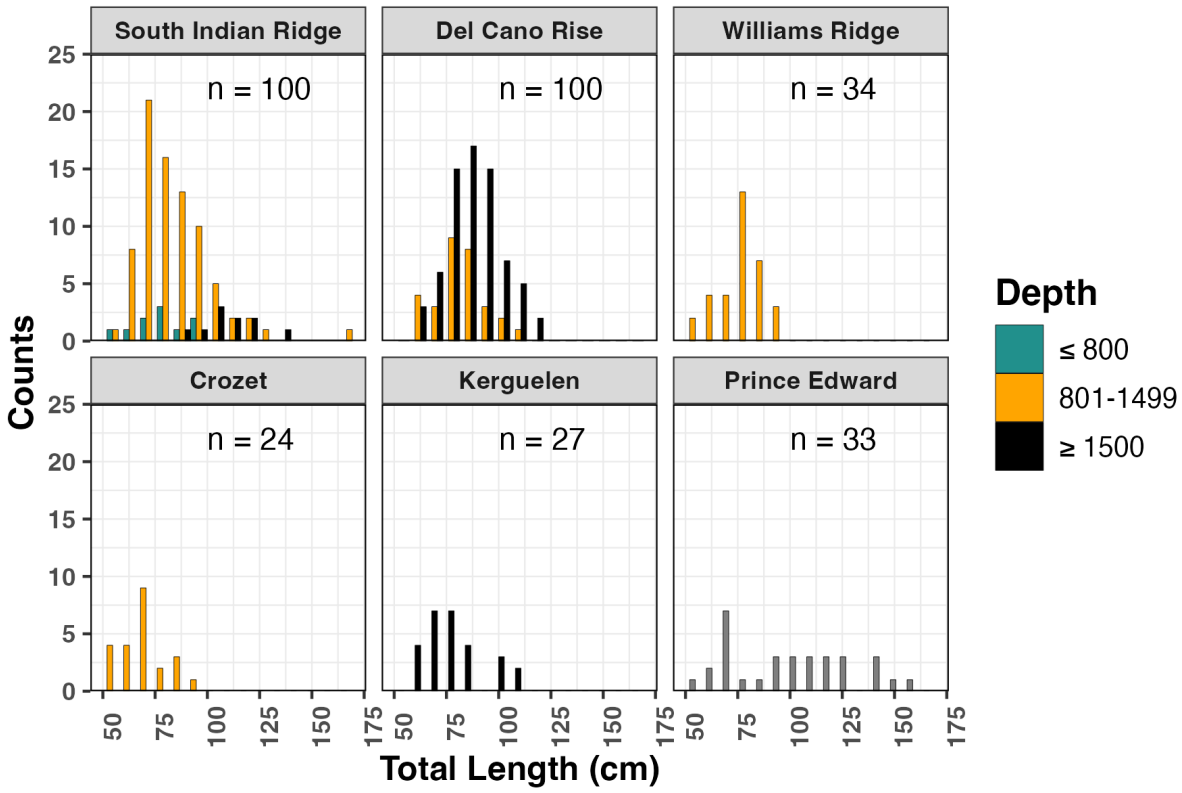


Fig. 5. Total Length distributions across depth ranges for all sampling regions except Prince Edward and Marion Islands, for which capture depth was not recorded (gray). Number of samples for each region is denoted in the plot areas.

The sex ratio varied for each sample set, identifying male-dominated collections for DCR, CR, and PEMI and female-dominated collections for SIR, WR, and KER (Fig. 6). Despite the sex imbalance, the length distributions for females and males were similar for SIR, DCR and KER. WR, CR and PEMI show evidence of similar distributions but as few males from WR and few females from CR and PEMI were collected, the distributions are less well defined (Fig. 6).

A relationship of increasing weight with length was observed across all sample sets (Fig. 7). The largest individuals were females captured at ≥ 950 m in sample sets from SIR (950 m), DCR (1550 m), WR (1150 m) and PEMI (unknown depth), while the largest individuals were males captured at > 1000 m for CR (1087 m) and KER (1783 m) (Fig. 7).

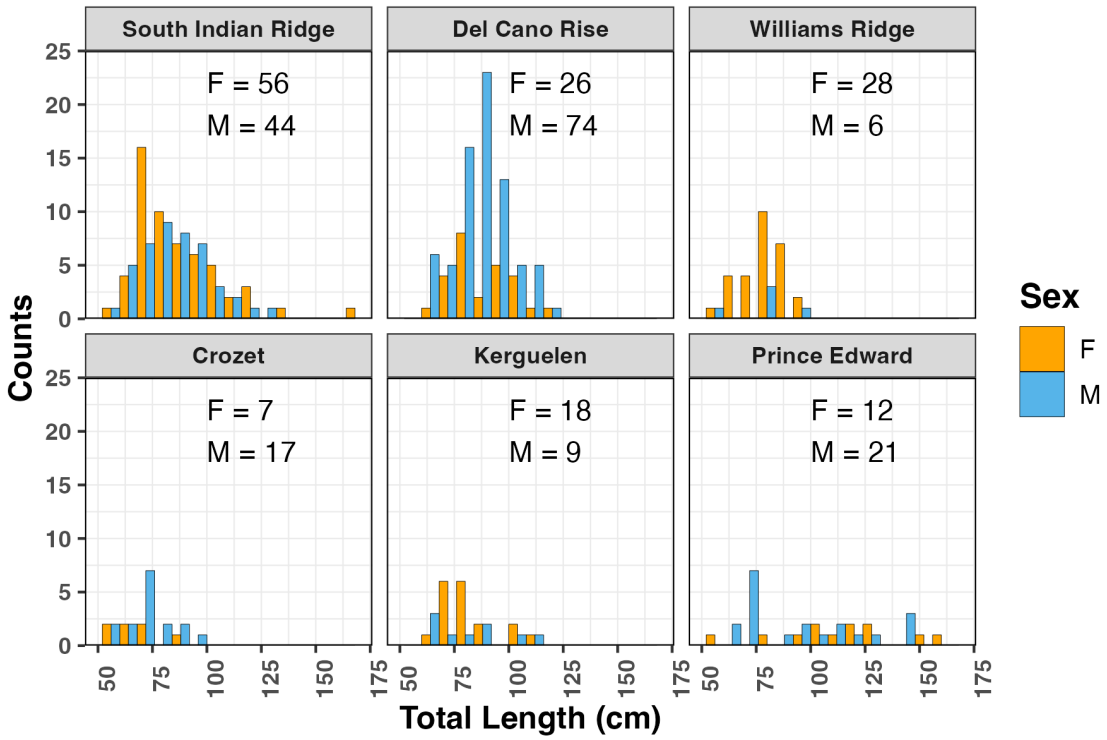


Fig.6. Length distributions across regions for sex. F = female and M = male.

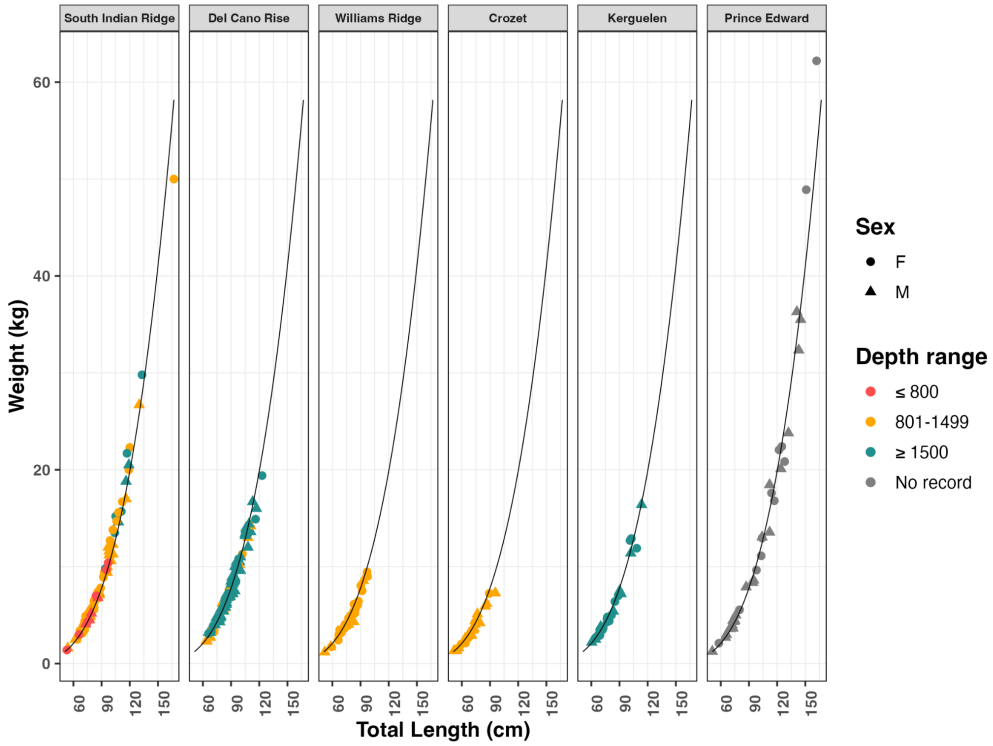


Fig.7. Relationship between total length and weight across regions by sex and depth (m) except for Prince Edward and Marion Islands, for which capture depth was not recorded (gray).

4.3.2 Maturity

Length at maturity is given in Figure 8, noting that all five maturity stages were only present in the sample sets obtained from SIR and DCR regions (Fig. 8, Fig. 9). At SIR, the majority of individuals were categorized at stages 1 and 2, for both males and females, leaving few individuals categorized as stages 4 and 5 (Fig. 9). At DCR, the majority of males were categorized at stages 2 or 3, whereas females were more evenly distributed amongst the stages (Fig. 9). At WR, only maturity stages 1 and 2 were sampled, with 80% of the samples obtained from females categorized as stage 1 (Fig. 8, Fig. 9).

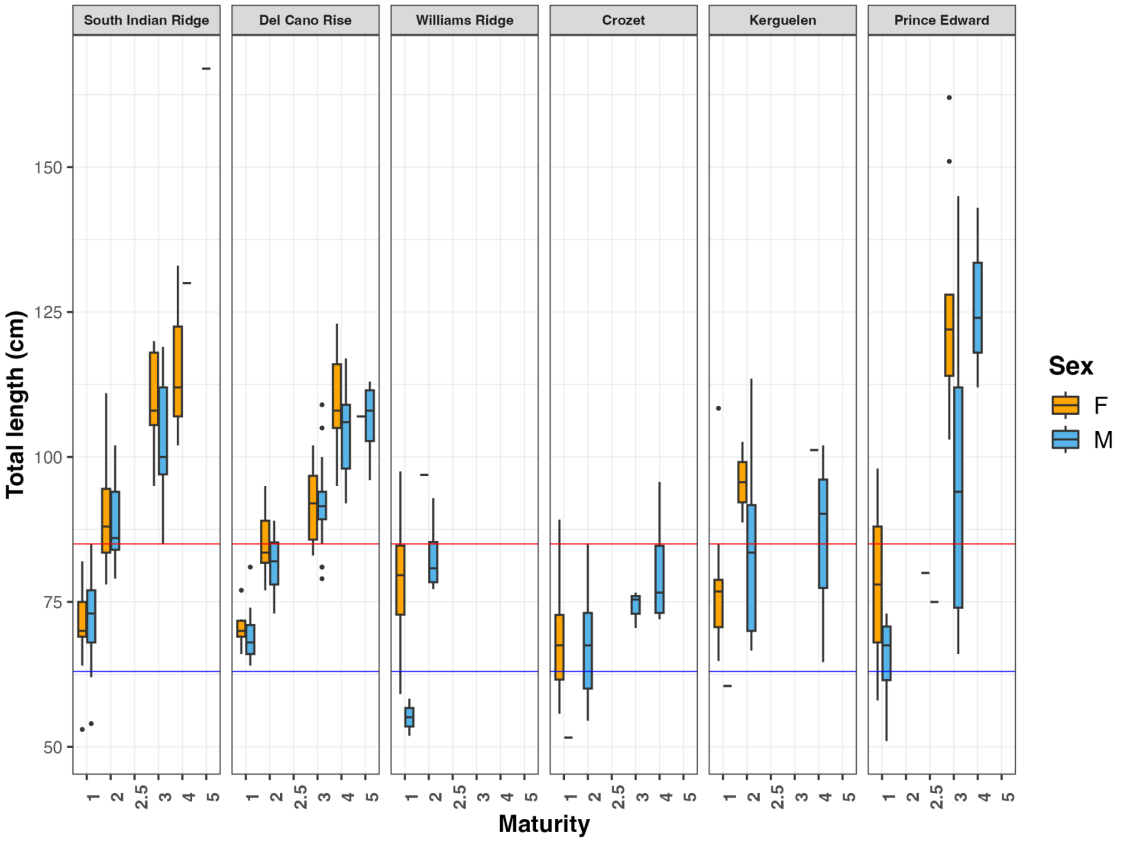


Fig.8. Maturity stage (1 - 5) by total length (cm) for each sex as grouped by sampling region. Horizontal lines indicate the length at first maturity for females (Lm50 = 85 cm) and males (Lm50 = 63 cm).

Though fewer individuals were collected from CR compared to the other regions, maturity stages 1-4 were represented in this sample set and the only females collected from CR were categorized at stage 1 (Fig. 9). At KER maturity stages 1, 2 and 4 were present in the sample set, with maturity stage 1 individuals being the most numerous (Fig. 9). At PEMI, maturity stages 1, 2, 3 and 4 were present in the sample set, with maturity stage 3 individuals being the most numerous (Fig. 9). Further, one female and one male were classified as maturity “2/3” and were denoted as “2.5”, while there was a wider distribution of male lengths for maturity stages 3 and for females at stage 1 (Fig. 8).

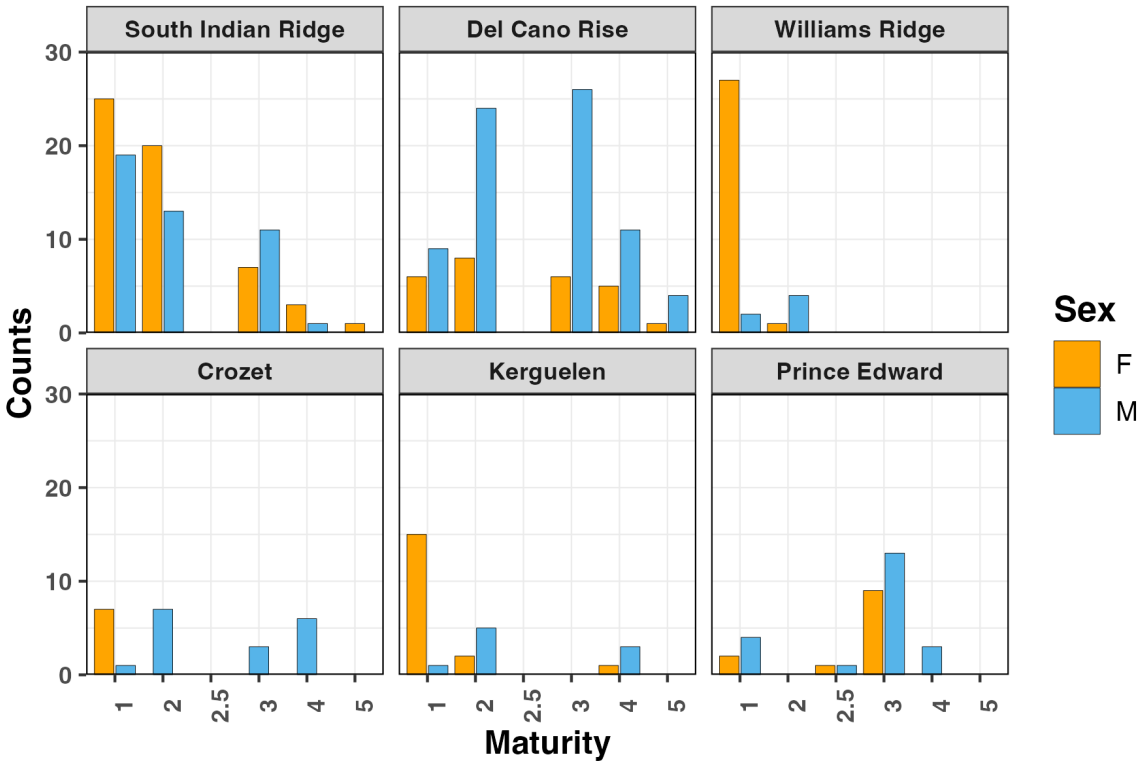


Fig.9. Count of individual maturity stages across regions for each sex.

When combining samples across regions, the sex ratio based on total length varied by depth (Fig. 10). At the shallowest depth (≤ 800 m), males measuring from L_T = 66 cm to 94 cm dominated the sampled individuals, while females dominated all other size classes. At depths between 801 and 1499 m, where the majority of all individuals were fished, the sex ratio was more balanced, around 0.5, until L_T > 125 cm, where the ratio skewed towards males until reaching the largest sizes (Fig. 10). Across all depth ranges, males dominate the middle size classes, while the sex ratio is skewed towards females at larger lengths. Capture depth for samples from PEMI was not recorded in the metadata ("No record", Fig. 10, far right); however, males still dominated the smaller sized individuals (< 90 cm), and above this size, females dominated.

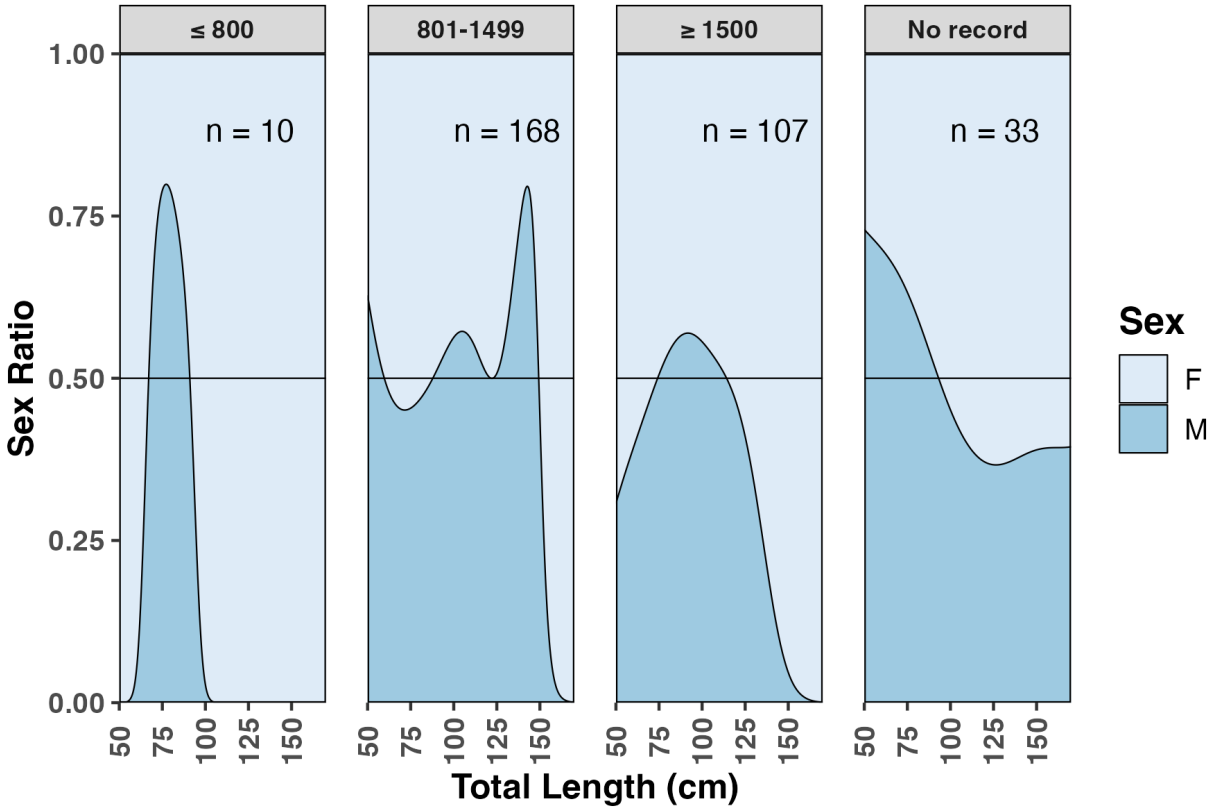


Fig.10. Sex ratio for *D. eleginoides catch* by depth where F = female and M = male. "No record" refers to the Prince Edward and Marion Islands samples, whose capture depth was not recorded in the metadata. These illustrations were produced using the R geom_density plot of the ggplot2 package with an adjust value=2 for the smoothing effect.

4.4 Sample selection

The total number of samples retrieved from the observers was 318 across six regions. Samples from KER and PEMI were collected and sequenced outside of the time frame of the current project, and we focused on samples from SIR, DCR, WR, and CR to prepare two 94-well plates. The number of samples available for SIR, DCR, WR ,and CR was 260, thus to fill the allotted space on two plates, we had to exclude some of these samples. Below, we detail the final selection where we employ a slightly modified version of the previous strategy to fit with the characteristics of each region and the space allotted on two plates.

The total number of samples from the four regions are as follows: SIR (n = 100), DCR (n = 100), CR (n = 24) and WR (n = 36) (Fig. 6). Two samples from WR were removed from selection consideration due to uncertainty regarding their sample names and linkage to the remaining metadata. As samples from both of these regions contain far fewer individuals with more skewed sex ratios (Fig. 6), we chose to select all samples from CR and all remaining samples from WR, which totaled 58. Next, we made a selection of 130 samples amongst the SIR and DCR sample sets, in which 65 of 100 samples from each region were chosen. Increasing the number of samples originating from SIR and DCR was ideal as these regions are geographically close to one another, and the increase in the number of individuals is likely to provide

higher power for discrimination between two sub-populations, if they are indeed distinct (*see* Toomey et al. 2016). While we cannot sequence equal numbers of individuals for WR or CR since fewer individuals were sampled from these regions, we note that ≥ 30 individuals were obtained from WR as initially requested from the observers. Further, we are awaiting sequencing data from samples from KER (n = 27) and PEMI (n = 35), which are scheduled to arrive in October 2024. In an addendum report, we will integrate sequencing data obtained from both these sites into the larger dataset containing 188 samples. It should be noted that 62 samples for KER and PEMI includes two samples from PEMI that did not have associated metadata. In total, 250 samples from six sites will be sequenced for this project.

As stated in the SER2022-TOP1 report, samples from SIR and DCR should have a sex ratio as close as possible to 1:1 as there is evidence that males and females may exhibit differing migratory behaviors and that this can result in specific patterns of dispersal depending on sex, possibly influencing the genetic structure of the species (Appleyard et al. 2002, Shaw et al. 2004, Rogers et al. 2006, Toomey et al. 2016). The sequencing of a 1:1 ratio between females and males may thus help determine the presence of genetic structuring linked to sex. We also prioritize spawning individuals (i.e. maturity stage 4) sampled at depths between 801 to 2000 m, to aid the selection of mix-sex aggregations for sequencing. The selection criteria are listed below, with additional focus on the depth range the fish were caught, maturity stages and lengths as additional criteria.

<u>Prioritized selection criteria:</u>

1. n = 65 individuals for SIR and DCR each (n = 130)
2. Priority is as close to a 1:1 sex ratio as possible
3. Priority depth > 800 m where spawning individuals were most likely to be present
4. Maturity stage 4
5. Maturity stages 3 and 5
6. Maturity stage 2 individuals with $L_T \geq 83$ cm found at > 800 m
7. Remaining maturity stages 3 and 5 found at ≤ 800 m
8. $L_T \geq 83$ cm for females and ≥ 63 cm for males at ≤ 800 m regardless of maturity stage remaining
9. If needed, finalize with largest females and males ≤ 800 m of smaller maturity stages

We prioritized selecting all individuals classified at maturity level 4. Next, stage 3 has previously been employed as the threshold for which both female and male fish reach maturity and ability to spawn (Everson and Murray 1999, Lord et al. 2006, Arana 2009), therefore stages 3 and 5 were also selected, as needed. The remaining individuals were classified at stage 1 and 2. Using knowledge of the average size-at-first maturity (Lord et al. 2006), we chose the threshold of $L_T \geq 83$ cm for females and ≥ 63 cm for males, that usually places the individual at stage 2. We selected these stages and lengths as needed to reach the target number of samples. Our ability to follow these criteria depended on the available samples in a dataset.

4.4.1 South Indian Ridge (SIR)

Thirty-three females and 32 males were selected to achieve a sex ratio close to 1 (0.97) at 33:32. An additional female was selected over a male to represent the female dominance in the overall sample set (Fig. 6). The selected samples for females retrieved from SIR contained individuals that measured ≥ 78 cm, nearly all from > 800 m. The female sample set therefore contained three samples classified as stage 4, eight samples from stages 3 and 5, 20 samples from stage 2 and two samples from stage 1. The male sample set contained individuals that measured at least ≥ 76 cm, nearly all from > 800 m. The male dataset therefore contained one sample classified as stage 4, 11 samples from stages 3 and 5, 13 samples from stage 2 and seven samples from stage 1 (Appendix table 1).

4.4.2 Del Cano Rise (DCR)

All females (n = 26) and 39 males were selected to obtain 65 samples and a sex ratio close to 1 (0.66) as possible at 2:3. The female sample set contained individuals that measured ≥ 66 cm, with five samples classified as stage 4, seven samples from stages 3 and 5, eight samples from stage 2, and six samples from stage 1. The male sample set contained individuals that measured at least ≥ 79 cm and had four samples classified as stage 5, 11 samples from stage 4, and 24 samples from stage 3 (Appendix table 1).

4.4.3 Williams Ridge (WR)

Considered samples retrieved from WR totaled 34. The sex ratio was skewed in favor of females at 3:14. Female lengths measured from 59.1 - 97.5 cm, while males measured 51.9 - 92.9 cm. All samples from this region were selected for sequencing (Appendix table 1).

4.4.4 Crozet (CR)

Samples retrieved from CR totaled 24. The sex ratio skewed in favor of males at 7:17. Female lengths measured from 55.7 - 89.2 cm, while males measured 51.6 - 95.7 cm. All samples from this region were selected for sequencing (Appendix table 1).

4.4.5 Kerguelen (KER)

Samples retrieved from KER totaled 27. The sex ratio skewed in favor of females at 2:1. Female lengths measured from 64.8 - 108.0 cm, while males measured 60.5 - 114.0 cm. All samples from this region were selected for sequencing (Appendix table 1). These sample numbers are lower than what was originally requested for processing (n = 30).

4.4.6 Prince Edward and Marion Islands (PEMI)

Samples retrieved from PE totaled 35. The sex ratio skewed in favor of males at 4:7. Female lengths measured from 58.0 - 162 cm, while males measured 51 - 145 cm. All samples from this region were selected for sequencing, which includes two without metadata (Appendix table 1).

4.4.7 Summary of the selected sample dataset

Once sample selection was completed, selected samples from SIR and DCR were combined into a single large list containing samples from WR and CR. This list was verified that there were no duplicated samples. Samples from KER and PEMI were not obtained until later (they were sent at a later date) are shown in Appendix table 1. The distribution of selected samples by sex and maturity for each region is plotted (Fig. 11).

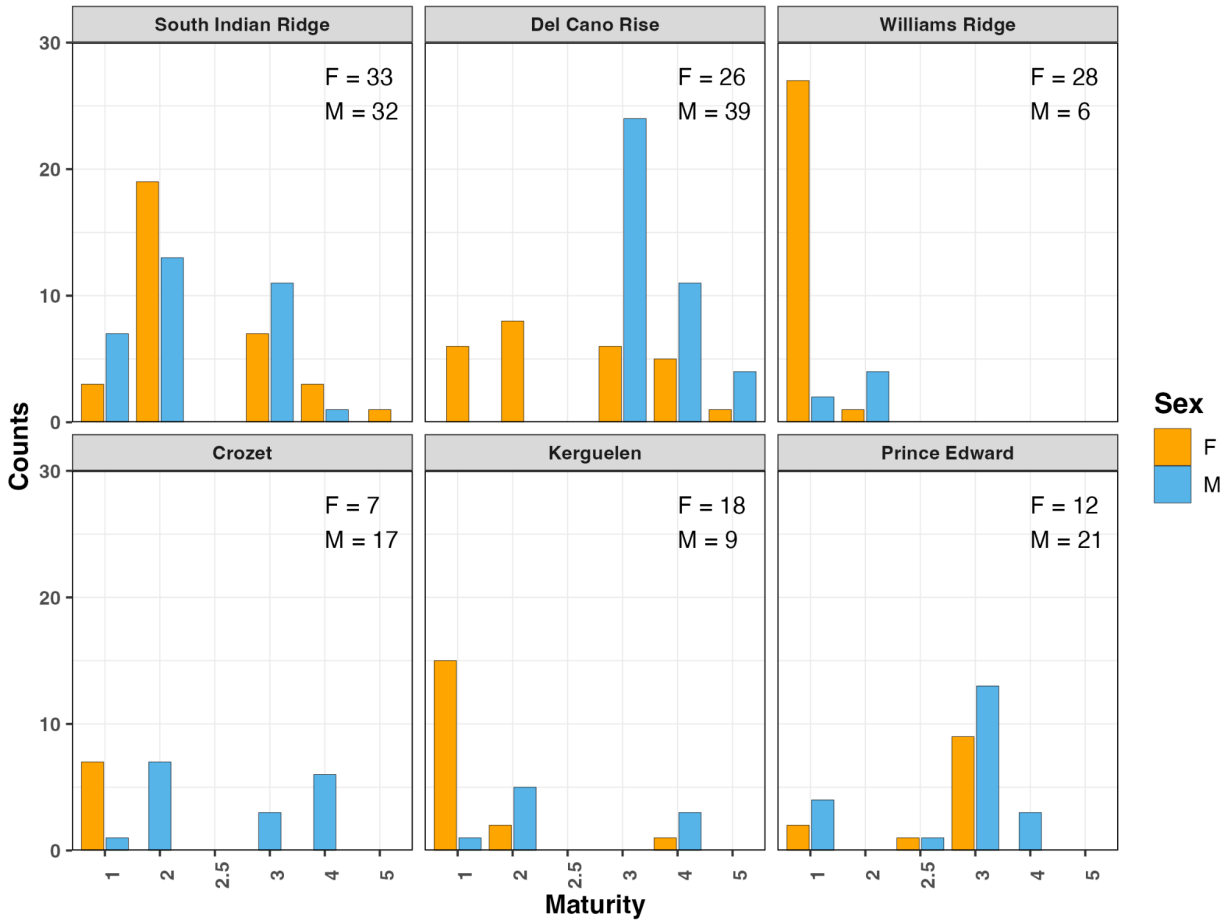


Fig.11. Maturity stage distributions for selected samples of each sex, across regions. F = female and M = male. Note that PEMI (Prince Edward) shows only 33 samples, as metadata was present for these individuals, while 35 samples will be sequenced. The total number of samples to be sequenced is 250.

5. Methods

5.1 Sample packing protocol

Fin clips were prepared for shipment to Diversity Arrays Technology (DArT PL) in Canberra, Australia. In accordance with DArT packing specifications, approximately 10 - 15 mg of each fin was placed in 1.1 ml

tubes, each containing 150 µl absolute ethanol. The racked tubes were grouped in strips of 8, which are arranged in the format of a 96-well plate. Each tube has its own sealing cap.

Gloves were donned prior to decontaminating workspaces using a 10% bleach solution, followed by wiping the area dry. This step was followed by thorough rinsing with water and wiping dry. Each strip of rack tubes was arranged in the orientation of a plate and labeled by sample name using a permanent marker. Next, a pipette was used to load 150 µl of absolute ethanol into each tube. Forceps and scalpels were disinfected by rinsing in a 10% bleach solution, followed by rinsing with 70% ethanol and wiped dry using a clean paper towel. As each sample had to be weighed, sterilized forceps were used to remove the first fin clip sample to place it carefully on a clean parafilm square residing on a balance. The sample name and weight of the original sample were recorded, and if the sample weighed over 15 mg, it was cut to a smaller size. The sub-sample was then securely placed in the rack tube denoting 'well A1' of the plate format, ensuring that it was completely submerged.

Next, all handling tools were decontaminated, and the used parafilm was discarded prior to sub-sampling the next sample. Once the first full column of racked tubes was filled with the fin clips and ethanol, each tube was securely closed by snapping a strip of barrette caps onto the column and placing it on ice. The process was completed for the selected 188 samples, filling two plates of racked tubes. A folded paper towel was then placed over the top of the capped tube plates, then each tube rack plate was enclosed in its associated box. Each box was secured with rubber bands, then sealed in its own ziplock bag labeled either "PLATE 1" or "PLATE 2" and stored in a -20º C freezer until shipment. The full sample packing protocol is described as Appendix 1. A sample tracking file was created and submitted to DArT's online order interface prior to packing and sending the samples through DHL.

5.2 Sample processing at DArT

DNA was extracted from the samples using the NucleoMag kit (Macherey-Nagel, Dueren, Germany); for the lysis step, samples were overlaid with 50 µL of T1 Buffer and 6.25 µL of proteinase K. The plate was centrifuged for 30-60 sec at 1000 rpm to ensure that the tissue samples were completely submerged in the solution. Next, samples were digested at 60° C overnight prior to being centrifuged for 10 min at 3000 rpm. The resulting lysate was clear and transferred to a new well plate. DNA was then bound to NucleoMag B-beads using a suspension of 6 µL beads in 90 µL MB2. The plates were continuously agitated to prevent the beads from settling. Samples were then transferred to the Tecan T100 robot (T100) to perform the final extraction steps (washing and elution into Elution Buffer).

DArT implemented their proprietary DarTseq™ high density sequencing procedure, which is a genome complexity reduction-based technology typically involving the use of next generation platforms to sequence a subset of locations spread throughout the genome (Sansaloni et al, 2011; Kilian et al, 2012; Courtois et al, 2013; Raman et al. 2014; Cruz et al. 2013). Details of the approach are considered proprietary by DArT; however, the DArT method is based on array hybridisations optimized for each

organism and application by selecting the most appropriate complexity reduction method, in terms of the size of the representation and the genome fraction.

Based on testing, the PstI-SphI enzyme combination was selected for *Dissostichus* spp. Therefore, 308 assays of 247 samples were processed with PstI and SphI compatible adaptors with two different restriction enzyme overhangs (Sansaloni et al, 2011). The PstI and SphI compatible adaptors were designed to include the Illumina flow cell attachment sequence, sequencing primer sequence and a "staggered", varying length barcode region, similar to the sequence reported by Elshire et al (2011). The reverse adaptor contained the flow cell attachment region and SphI compatible overhang sequence. Only "mixed fragments" (PstI-SphI) were amplified through PCR using the following reaction conditions: 94° C for 1 minute followed by 30 cycles of 94° C for 20 seconds, 58° C for 30 seconds, 72° C for 45 seconds, and ending with an extension at 72° C for 7 minutes. Following PCR assays, equimolar amounts of amplification products from each sample were pooled for sequencing on the Illumina NovaSeq X+(single read) for 100 cycles.

5.3 Bioinformatic analyses

5.3.1 Raw sequence data processing

Sequencing data produced were stored as FASTQ files that were processed through DArT's proprietary analytic pipeline. The pipeline included quality control steps such as the removal of poor-quality sequences by applying more stringent selection criteria to the barcode region compared to the rest of the sequence. Filtering was performed on the raw sequences using the following parameters: Min Phred pass score 30, Min pass percentage 75 for the barcode region, and Min Phred pass score 10, Min pass percentage 50 for the whole read.

Upon arrival, DArT performed DNA extraction, quality control, library preparation, sequencing, and SNP genotyping on the samples. Approximately 278 389 unique sequences per sample were used for SNP calling. Identical sequences were collapsed into "fastqcoll" files, which were "groomed" using DArT PL's proprietary algorithm which corrects low quality base from a singleton tag into a correct base using collapsed tags with multiple members as a template. The "groomed" fastqcoll files were used in the secondary pipeline for DArT PL's proprietary SNP and SilicoDArT (presence/absence of restriction fragments in representation) calling algorithms (DArTsoft14). For SNP calling, all tags from all libraries included in the DArTsoft14 analysis are clustered using DArT PL's C++ algorithm at the threshold distance of 3, followed by parsing of the clusters into separate SNP loci using a range of parameters, including the balance of read counts for the allelic pairs. Additional selection criteria were added to the algorithm based on analysis of approximately 1000 controlled cross populations. Testing for Mendelian distribution of alleles in these populations facilitated selection of technical parameters discriminating true allelic variants from paralogous sequences. In addition, technical replicates were processed for scoring consistency, used as the main selection criteria for high quality/low error rate markers and to estimate reproducibility of reported markers.

To call SNP loci, remaining sequences were aligned to the most recently published *Dissostichus eleginoides* genome found under GenBank assembly GCA_031216635.1 (Korea University, 2023) using NCBI BLAST algorithm (Altschul et al. 1990). E-value parameters were set at 5e-7, and a minimum sequence identity of 80% was implemented.

DArT provided final reports which contained data specifying the SNP loci identified from each sequenced and genotyped sample, as well as call rates and the codominant status of each sample (Appendix table 1). Three reports in total, plus associated metadata, were provided. The first report is in DArT's "2 Rows Format", whereby each allele is scored in a binary fashion ("1"=Presence and "0"=Absence). Heterozygotes are therefore scored as 1/1 (presence for both alleles/both rows). The second report is in DArT's "1 Row Format". This report displays two states for each individual sample at each locus and genotype codes were as follows: "0" for the homozygous reference allele or the allele most represented across individuals, "1" for the homozygous alternative allele, "2" for heterozygotes, and missing data were coded as "NA" or "-". The third report is the "SilicoDArT Format" for SilicoDArTs which are scored in a binary fashion, representing genetically "dominant" markers, with "1" as presence and "0" as absence of a restriction fragment with the marker sequence in genomic representation of the sample. In this report, "-" represents calls with non-zero counts, but too low counts to score confidently as "1".

For this study, the 1-row format was used to perform population structure analyses in R (R Core Team, 2024) given its smaller file size that is easier to process using a local computer. A metadata key was provided by DArT to provide corresponding information and explanations for each report column, and is available in Appendix Table 2.

5.3.2 Dataset filtering steps

Data analysis was completed following the computational strategy detailed in Chevrier et al. 2024. First, we created a separate metadata file which contained information for each individual sample (n = 251) to be used in downstream analyses. In this file, columns were named as follows: "id" to denote the individual sample name, "pop" which listed the population origin of the individual without referencing any geographic region, "sex", "maturity" to include the individual's maturity stage and "depth" to include the depth range where each sample was fished. Downstream analyses were performed with the aid of R packages specialized to handle output data files from DArT and population genetic analyses, most notably ***dartR*** v.2.9.7 (Gruber et al. 2018) and ***adegenet*** v.2.1.10 (Jombart, 2008).

Data filtering steps required the use of *dartR*, in which the 1-row SNP report and the sample metadata were imported using the gl.read.dart() function to retain the data as a genlight object through *adegenet*. Details for each function are described in the *dartR* manual.

The first step was to examine the sequencing depth of the dataset, then remove loci with exceptionally low or high numbers of sequence reads to avoid missing data or biases in downstream analyses. The filtering threshold therefore kept loci having between 20 - 145 reads.

SNP loci datasets can include DNA fragments that contain more than one SNP, and as multiple SNPs within a fragment are likely to be linked, these secondary loci are frequently removed in favor of the loci more representative across the samples. Therefore, the second step was to remove secondary SNPs from the dataset.

DArT calculates "Average Reproducibility" as a fraction of allele calls (n = 30) which are consistent among the technical replicates generated from the same DNA samples, in a fully independent manner. The reproducibility fraction was calculated for each of the two alleles and averaged for the marker to provide a repeatable result. Accordingly, the data were then filtered based on a threshold for the average repeatability of 0.95 (default = 0.99) chosen as a more flexible threshold, whereby loci below this value were removed.

Within the dataset, loci have a call rate, which in *dartR* is calculated as the proportion of uncalled SNPs, (i.e., missing values). The next filtering step was then to remove loci/samples that had proportional call rates below 0.95 for both loci and samples, i.e., if a locus was called only 20% of the time, it was removed. The following step was examining and removing Minor Allele Frequencies (MAF). MAFs are the frequencies of the less common alleles occurring at each loci. These rarer variants could be due to genotyping errors, but also, they are not present at high enough frequencies to differentiate between populations. Therefore, the threshold to filter MAFs was set at 0.05 (default = 0.01), as is commonly applied to favor more common variants (Anderson et al. 2010; Montes et al. 2013).

Following the filtering of MAFs, a plot was generated to visualize the observed heterozygosities (H_o, proportion of heterozygous loci observed), and unbiased expected heterozygosities (uH_e, proportion of heterozygous loci expected under Hardy-Weinberg equilibrium), as well as the inbreeding coefficient (F_{IS}, measuring inbreeding within a "subpopulation" with respect to the local "population") for each sampled geographic region. The scale for H_o and uH_e is frequencies between 0 - 1 while for F_{IS} is a range extending from -1 to +1, providing information on the deficit of heterozygotes.

Next, loci that showed significant departure from Hardy-Weinberg (HW) proportions were filtered from the dataset at threshold 3 (default = 1), i.e., a locus has to be out of HW proportions at a minimum of three populations before it is removed. This value was chosen as it represents one third of the total population. Following this, locus metrics such as call rate, were recalculated prior to continuing filtering the dataset. Further, any monomorphic loci (when individuals have the same allele at a locus) remaining following the previous filtering steps were also removed, prior to another recalculation of locus metrics.

Finally, diagnostic plots were generated through a Pearson Principal Component analysis (PCA). This allowed the identification of anomalous patterns in the samples to be addressed by filtering. Following this last filtering step, remaining monomorphics were removed and metrics were again recalculated before proceeding to downstream statistical analyses.

5.3.3 Population assignment (DAPC analysis)

The number of population clusters was identified through the Discriminant Analysis of Principal Components method (DAPC, Jombart et al. 2010). The 'optimal' or best supported number of clusters (k) is identified without *a priori* population information through the lowest associated Bayesian Information Criterion (BIC) value, which is produced after applying the K-means algorithm used to separate individuals into clusters based on genetic similarity (Jombart et al. 2010). Further, DAPC combines the principal component analysis (PCA) and Linear Discriminant Analysis (LDA) approaches to define and distinguish between clusters of individuals and is useful for disentangling datasets that may have overlapping or closely related groups. DAPC first implements PCA to reduce the dimensionality of the data by transforming it into principal components (PCs) that capture the main axes of genetic variation across all individuals and loci. The PCs that capture the most significant variation are retained and LDA is used to maximize separations between clusters.

The first step was to determine the best number of 'k' clusters in the dataset by implementing the find.clusters() function of *adegenet* used to calculate BIC values, setting the estimated maximum number of potential clusters at 10 (n > 6 potential populations) for 10 000 iterations and the number of k = 2 (n > 1 populations). Next, the cross-validation step is implemented through the xvalDapc() function for identifying the optimal number of PCAs to retain. Here, the dataset is divided into two sets: a training set which comprises the vast majority of the data and a validation set which contains the rest. Cross validation is run on the training set having variable numbers of PCs retained, and the analysis is able to predict the group membership of the individuals in the validation set. Finally, the dapc () function was implemented with parameters n.pca = 100 (number of axes to retain for the PCA step) and n.da = 2 (number of axes to retain for the DA step).

5.3.4 Population assignment (STRUCTURE analysis)

In addition to PCA-type approaches to examine the potential presence of population structure amongst *D. eleginoides*, we also implemented STRUCTURE v.2.3.4 (Pritchard et al. 2000; Porras-Hurtado et al. 2013), a widely applied Bayesian clustering method used for population structure inferences where individuals (samples) are placed into groups that are share similar patterns of variation.

To run STRUCTURE, the data was output from R in .vcf format to be converted into a STRUCTURE input file using the PGDSpider v.3.0.0.0 (Lischer and Excoffier 2012). STRUCTURE was implemented using the admixture model, with K =1 to K = 7 at a repetition of five, including a burn-in period of 5 000 and 50 000 MCMC repetitions. STRUCTURE analyses were run keeping *a priori* information on sampling regions as well as without information on sampling regions. The resulting data were visualized using *ggplot2* in R.

5.3.5 Pairwise analysis

Pairwise F_{ST} values were calculated through the Weir-Cockerham method (Weir and Cockerham 1984) using the gl.fst.pop() function from *dartR*, which implements StAMPP v1.6.3 (Pembleton et al. 2013).

Additionally, 95% confidence intervals and associated p-values were also computed through the test at 10 000 bootstraps.

5.3.6 Additional analyses

To visualize and highlight potential SNPs under selection and their locations in the genome, the .vcf file was uploaded to vcf tools v.0.0.16 (Danecek et al. 2011) to calculate F_{ST} values for individual SNPs using the Weir-Cockerham method, which were then visualized in relation to SNP chromosomal positions on a Manhattan plot generated in the **qqman** v.0.1.9 package (Ehret 2010). Further, we investigated if there were loci under selection through OutFLANK (Whitlock and Lotterhos 2015). OutFLANK helps identify loci through the calculation of F_{ST} values for each locus using the Whitlock and Lotterhos method, plotting them on a distribution and calculating a false discovery rate for q-values at the default 0.05, which is the threshold for identifying loci as under selection. Lastly, we examined whether or not any loci linked to sex on sex chromosomes via the filter.sexlinked() function of dartR, where we chose to keep putative sex-linked loci, while the sex for each sample was provided in the sample metadata file.

5.4 Potential habitat distribution towards management unit delineation

Considering firstly the outcome of the population discrimination analysis, the management units are proposed based on biological and ecological considerations, geographic and environmental factors, fisheries data and current management boundaries. To investigate for biological and environmental factors, we used the generalized additive model (GAM) developed in the SIOFA-TOP1 project (see Nieblas and Cowart 2023 for details) to develop potential habitat distribution maps at 0.0833° x 0.0833° for adults (maturity stage ≥2 (Everson and Murray 1999, Yates et al. 2018) and juveniles (maturity stage 1) based on bathymetry, slope, and bottom temperature. These habitat maps were overlaid with Vulnerable Marine Environments (VMEs; FAO, 2024) closed areas in the SIOFA and CCAMLR zones to consider additional ecological aspects for the proposed management units, catch distribution, as well as current management boundaries.

The preference range for the environmental variables informing potential habitat distribution were derived from the literature search and the GAM outputs of SER2022-TOP1 (Table 1). We define juvenile habitat in shallow zones (50-800 m; Welsford et al. 2011, Péron et al. 2016) and adult habitat from 800 - 2000 m (Nieblas and Cowart 2023). Both juveniles and adults prefer relatively flat bottoms (<0.2 radians; Nieblas and Cowart 2023), and bottom temperature should not be lower than 2° C (Duhamel et al. 1982), with an upper limit of around 3.2 (2.2° C ± 1.4; Figure 12). Bottom temperatures < 2° C found in our study may be due to a spatial/temporal mismatch between the catch data and the ocean model reanalysis used as the source of the temperature data, or an underestimation of temperature by the ocean model. We base the lower threshold on the literature (Dumel et al. 1982) and the upper threshold on the data from our study. Further, we note that the fisheries-dependent catch data are constrained by targeted depths of the fishing sets and potentially size-selective fishing gear (hook size), which may impact the depth range observed for juveniles.

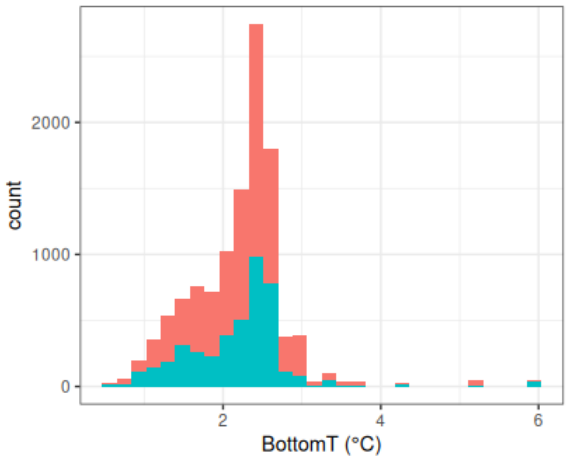


Fig. 12 Bottom temperature (°C) of catch of adults (red, maturity stages ≥2) and juveniles (blue, maturity stage = 1) across all sites, where temperatures are derived from the CMEMS global reanalysis product (see Nieblas and Cowart 2023).

Table 1. The habitat preferences of adults and juveniles of the three key environmental variables, with the justification for the preference derived either/or/both the observed mean and information from the literature.

Variable	**Adult preference**	**Justification**	**Juvenile preference**	**Justification**
Bathymetry (m)	800 - 2000	Nieblas and Cowart 2023	50 - 800	Welsford et al. 2011, Péron et al. 2016
Slope (radians)	0 - 0.2	Nieblas and Cowart 2023	0 - 0.11	Nieblas and Cowart 2023
Bottom temperature (°C)	2 - 3.2	Duhamel et al. 1982; Present study (Fig. 12)	2-3.4	Duhamel et al. 1982; Present study (Fig. 12)

Environmental variables were overlaid according to these habitat preferences to produce a potential habitat distribution of adults and juveniles. Climatology across all years was used to represent the average bottom temperature.

6. Results

6.1 Raw data summary

A total of 308 FASTQ files were produced, each one ascribed to a specific sample. Of these, 61 sequence files were considered “technical replicates”, or samples that were sequenced twice or thrice for

assessment of the sequencing run (described above). Technical replicate samples were removed by DArT prior to issuing the final reports. The final report therefore consisted of data for 247 (of 251) samples, as one sample from Crozet (CRO_008) and three samples from PEMI (ZAF_013, ZAF_026, ZAF_028) all failed sequencing, as well as a total of 59 343 SNPs available for downstream processing.

6.2 Dataset filtering

The sequencing depth of the dataset is represented in Fig. 13, with the pre-filtered dataset shown on top with red bars denoting the 20 -145 sequence read range chosen for filtering. The post-filtered dataset resulted in 247 samples and 23 234 SNP loci.

After removing the secondary SNPs, filtering the dataset based on the average repeatability and call rate thresholds, the remaining dataset contained 247 samples and 10 583 loci. Next, filtering MAF frequencies at 0.05 removed very rare alleles below this threshold (Fig. 14), further reducing the dataset to 243 samples and 2 835 loci.

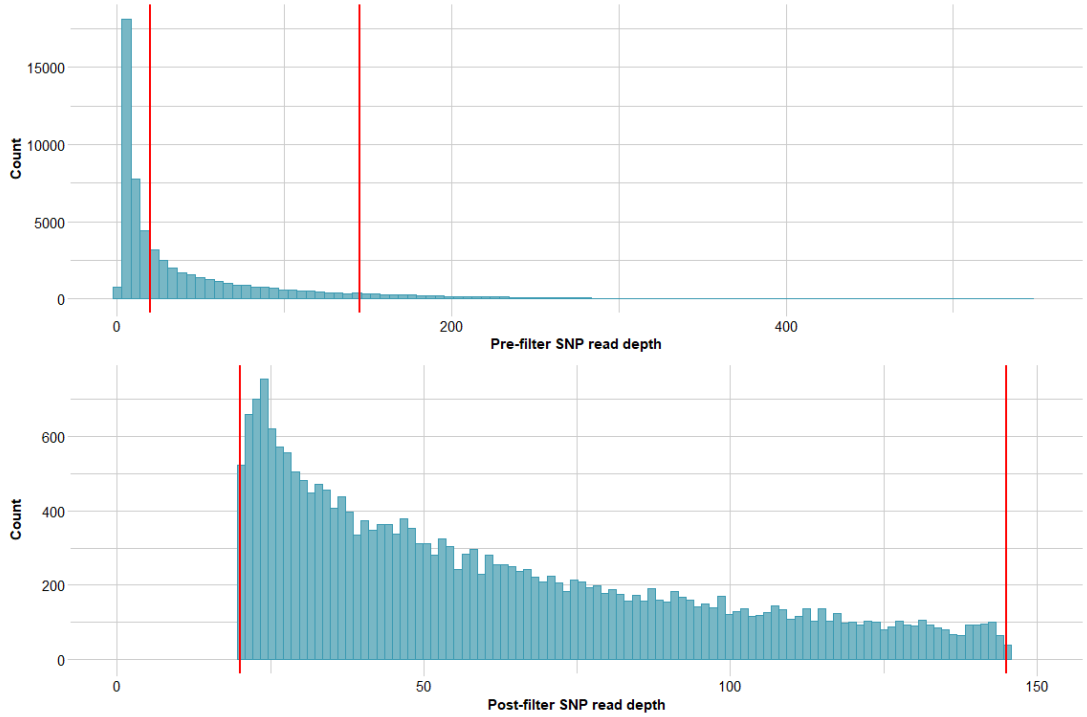


Fig.13. Dataset filtering for sequencing depth by the number of sequencing reads on the x-axis and the number of SNP loci on the y-axis. The top plot shows the full dataset, while the bottom plot shows the dataset remaining after filtering where SNPs having numbers of sequences outside of the 20 -145 range (red lines) were removed.

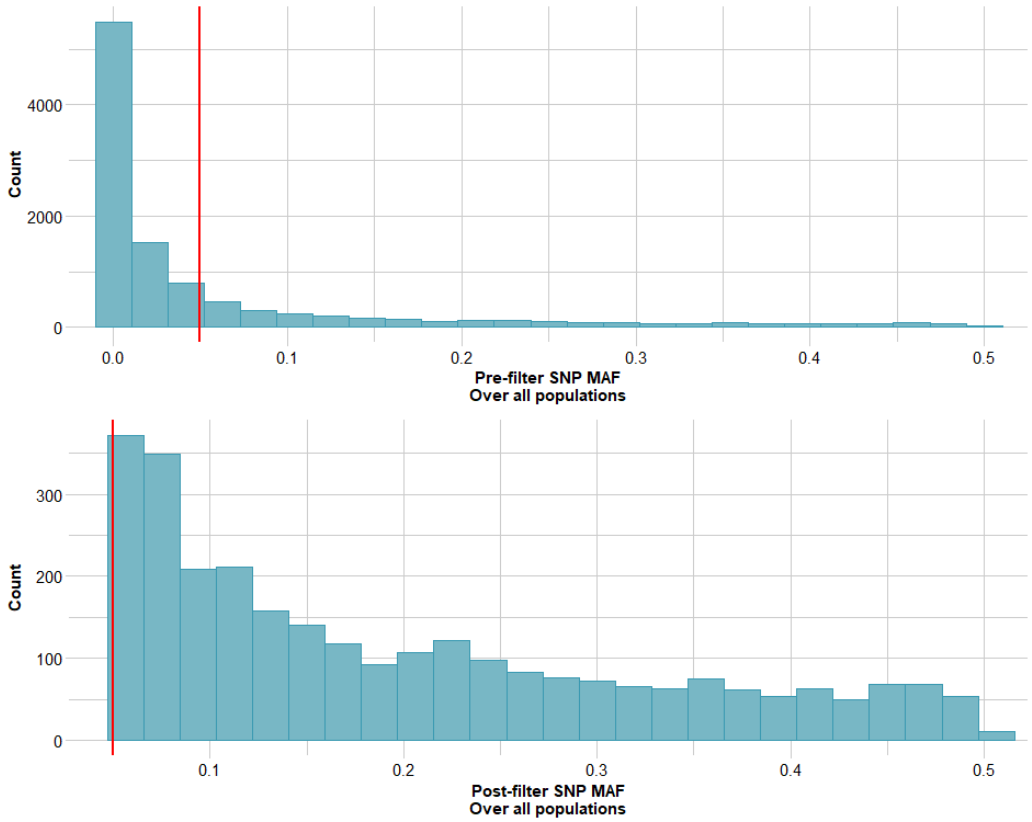


Fig. 14. Dataset filtering of minor allele frequencies (MAF) by the frequency of MAFs over all populations on the x-axis and the number of SNP loci on the y-axis. The top plot shows the full dataset, while the bottom plot shows the dataset remaining after filtering, and the MAFs below a frequency of 0.05 (red line) were removed. Those loci at 0.5 identify that both alleles are equally frequent.

Heterozygosities and F_{IS} values by sampling region, after the removal of secondary loci, filtration by call rates, and MAFs, are shown in Fig. 15. At the loci examined, individuals belonging to each sampling region exhibited comparable H_o and uH_e values near 0.3 (on a scale from 0 to 1), supporting that observed populations are largely consistent with Hardy-Weinberg equilibrium (i.e., random mating, constant allele frequencies). Each sampling region also exhibited F_{IS} values near zero, in accordance with no significant difference between observed and expected heterozygosity, and indicating very little inbreeding within these regions. These findings suggest that individuals mate across different sampling regions rather than being limited to their own.

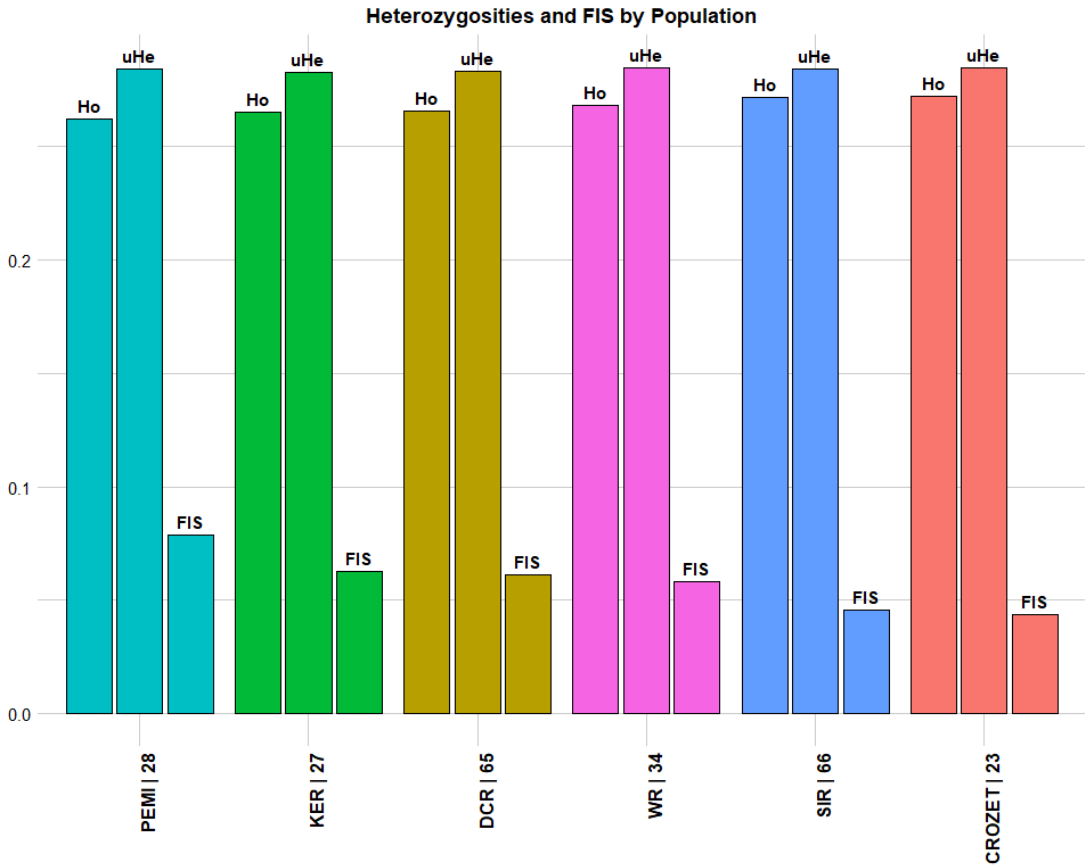


Fig.15. Heterozygosities and FIS by sampling regions, or "populations". The x-axis shows the sampling region names and number of individuals: PEMI = Prince Edward and Marion Islands, KER = Kerguelen, DCR = Del Cano Rise, WR = Williams Ridge, SIR = South Indian Ridge and CROZET = Crozet. The y-axis exhibits ranges of values, which for Ho and uHe would be frequencies between 0 - 1 and for FIS would be a range extending from -1 to +1.

Following the removal of loci based on significant departures from Hardy-Weinberg (HW) proportions, the PCA diagnostic plot identified two outlier samples from SIR clustering very closely to one another (Fig.16, top). Further inspection identified these samples as SIR_003 and SIR_008, suggesting potential cross-contamination/sample carryover between these samples given their similar values and proximity on the well plate. Once we chose to remove SIR_008 from the dataset and no further outliers were identified (Fig.16, bottom). Following the removal of SIR_008, monomorphic loci were filtered again and metrics recalculated, resulting in a final filtered dataset of 242 samples and 2 729 loci for population structure analyses.

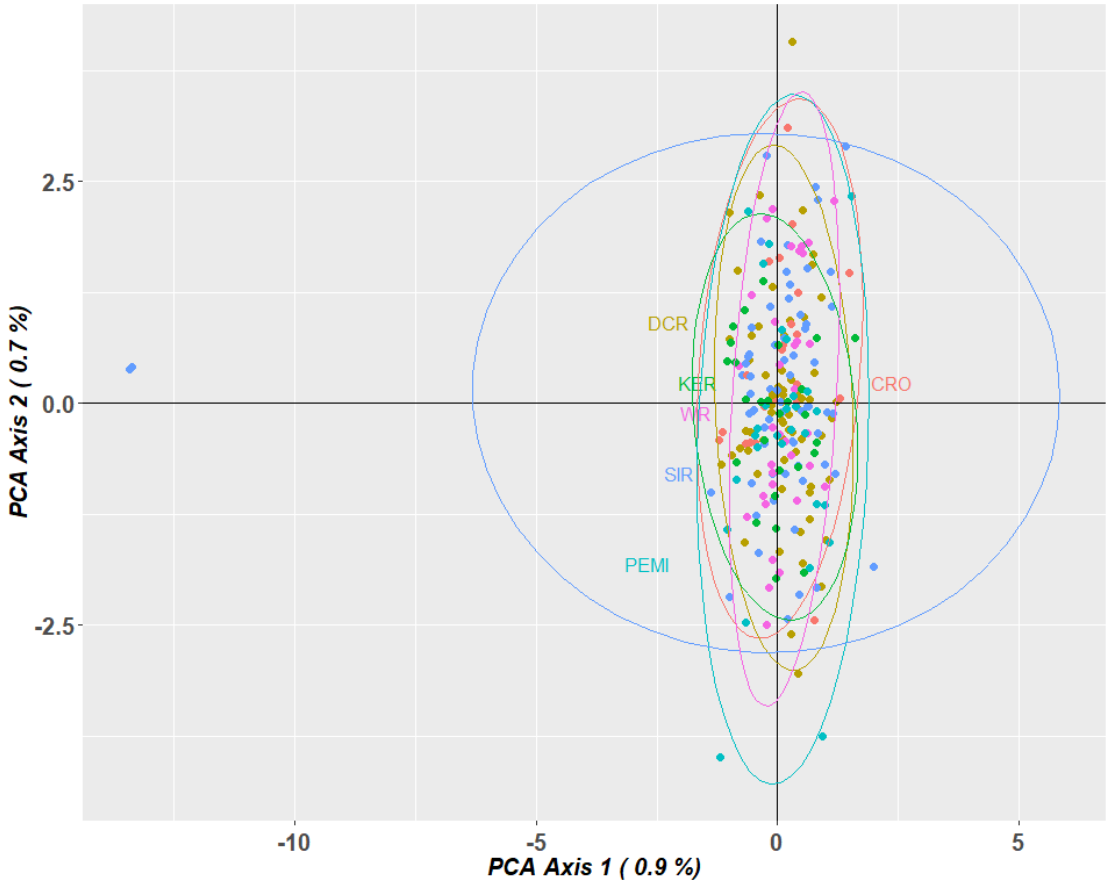
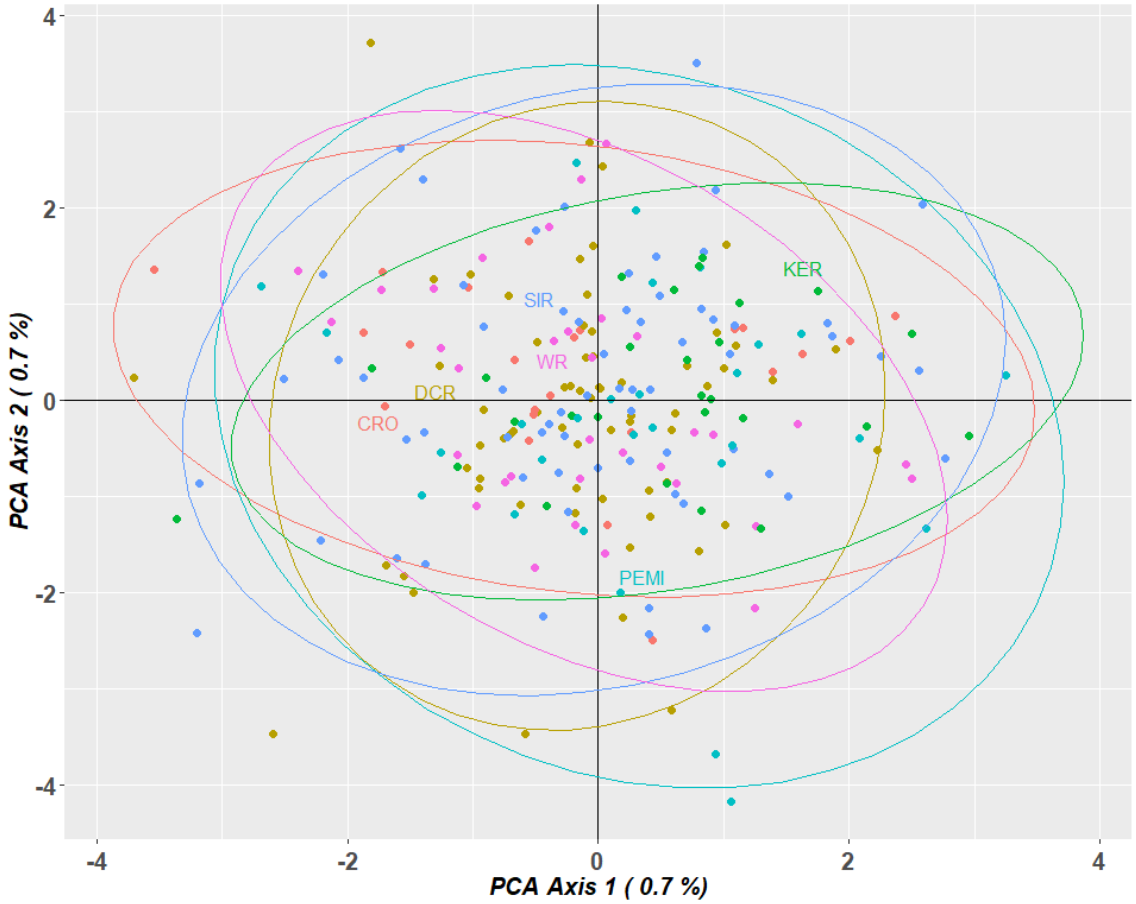


Fig.16. Diagnostic PCAs before (top) and after (bottom) the removal of sample SIR_008. The top figure shows two blue points on the far right which are SIR_003 and SIR_008. These samples are similar due to carryover.



6.3 Genetic stock discrimination analysis

6.3.1 Population assignment (DAPC analysis)

The best supported number of clusters (k) identified through the k-means clustering method of DAPC was one (K = 1), as identified by the lowest associated Bayesian Information Criterion (BIC) value (Fig.17).

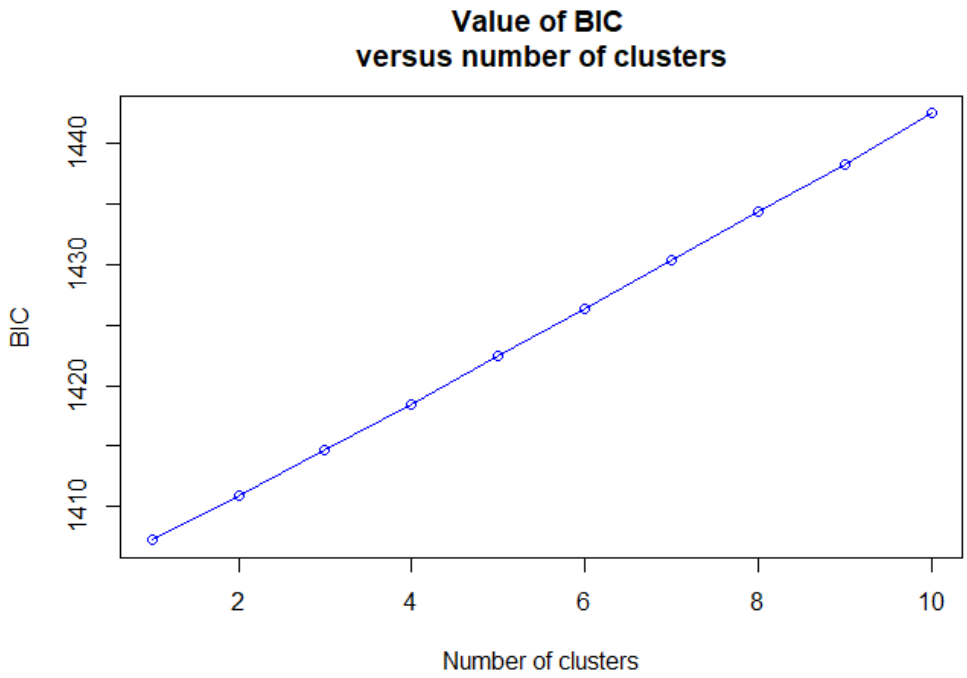


Fig.17. Output depicting the optimal number of population clusters (k) through Bayesian Information Criterion (BIC) values, assessed through the K-means clustering method. The optimal number of clusters is that which the BIC value is lowest (Jombard et al. 2010).

The cross-validation step identified the optimal number of PCs at 100 , out of a total 200 to retain (Fig.18), as 100 was the number of PCs achieving the lowest root Mean Squared Error (MSE).

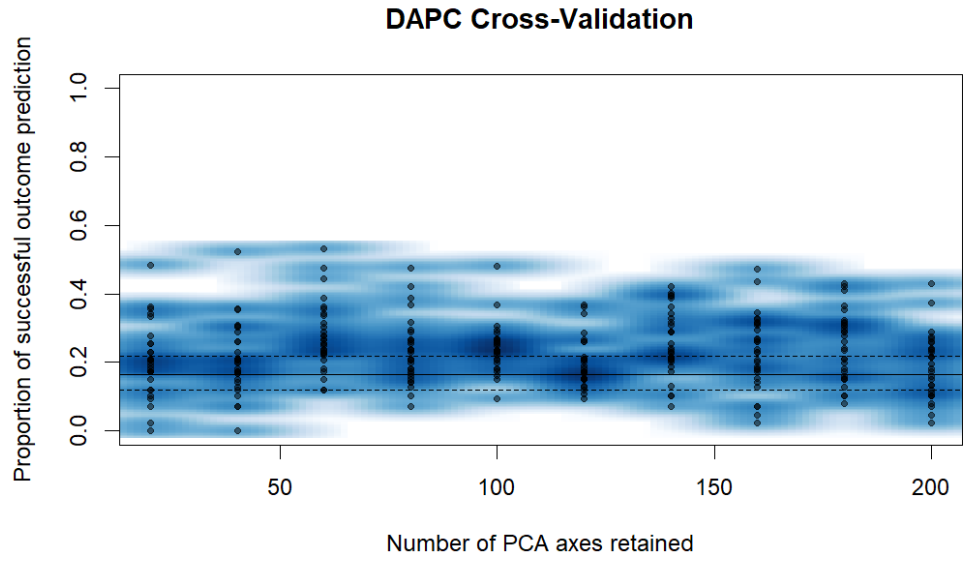


Fig.18. DAPC cross-validation plot with the number of PCs retained by the proportion of successful outcome predictions in DAPC. The optimal number of PCs was 100 , having a root MSE of 0.700.

The clustering visualization of the DAPC analyses identifies relationships between the different sampling regions, with the first two discriminant axes representing nearly half of the genetic variation found within the total dataset (Fig.19). Eigenvalues (not shown) identify that a third axis represents 19.4 % of variation, while axes 4 and 5 represent the remaining variation (32.4 % combined). While Fig 19 illustrates the clusters as grouped by sampling regions, the samples (colored circles) of different regions are close to one another and in some cases, overlap. This is evident with South Indian Ridge (SIR), Del Cano Rise (DCR) and Crozet (CR) clusters, supporting that the genetic variation between these individuals is small.

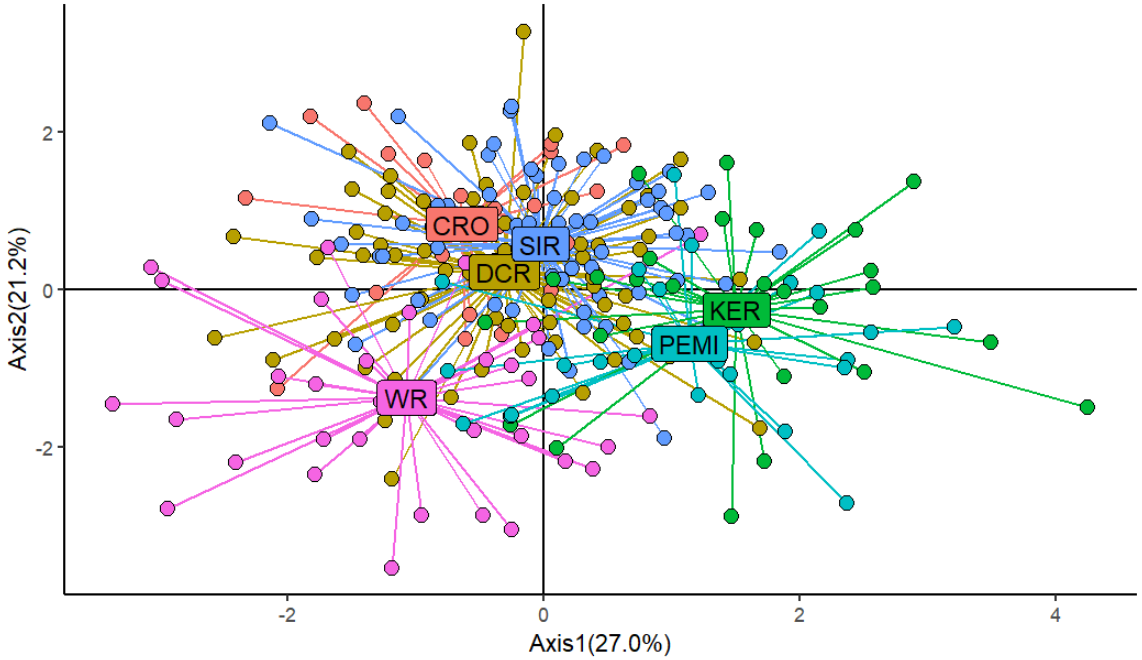


Fig.19. Visualization of clusters assessed through the DAPC approach, focusing on the first two discriminant axes, which represent the majority of the genetic variation. Individual samples are represented as dots and grouped by their sampling regions as follows: WR = Williams Ridge, DCR = Del Cano Rise, SIR = South Indian Ridge, CR = Crozet, Kerguelen (KER) and Prince Edward and Marion Islands (PEMI).

6.3.2 Population assignment (STRUCTURE analysis)

The occurrence of a single population of D. eleginoides across the four sampling regions was supported by STRUCTURE analyses (Fig. 20). When conducting admixture analyses for clusters K = 2 to K = 7 , individuals from the six regions showed high levels of admixture, with no distinct color (i.e. genetic cluster) predominating for any individual. This supports a lack of well-defined population groups in the dataset, whether a priori sampling information is added or not (Fig. 20).

a priori sampling analysis

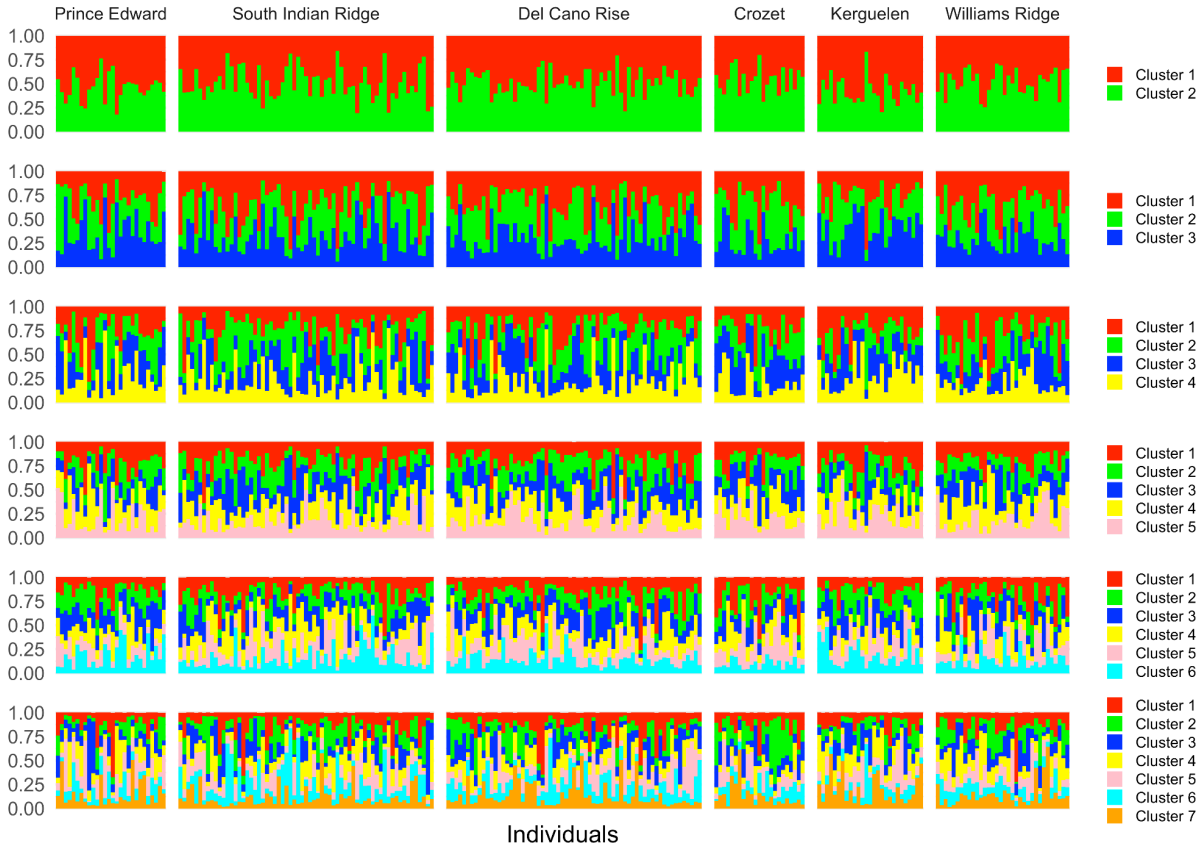


Fig. 20. STRUCTURE results (admixture model, K = 2 to K = 7 , five replicate runs) for D. eleginoides data run with a priori sampling information (top) and no a priori sampling information (bottom) based on 242 individuals. Each vertical bar represents an individual fish, while the y-axis shows the proportion of each individual's genotype belonging to a cluster, defined in the legend.

6.3.3 Pairwise analysis

The pairwise FST values between each sampling region are shown in Fig. 21. While there were significant differences between CRO and four other regions (PEMI, SIR, KER and WR), as well as SIR and three other regions (CRO, KER and WR) , actual FST values were very close to zero, indicating minimal genetic diversity present across the six regions (Fig. 21). Darker-toned boxes in the figure suggest slightly lower genetic variation between DCR and SIR which are geographically closer to one another, and DCR and WR, as well as PEMI and KER, despite the wider geographic distances between these sampling regions, which is assumed to impact gene flow.

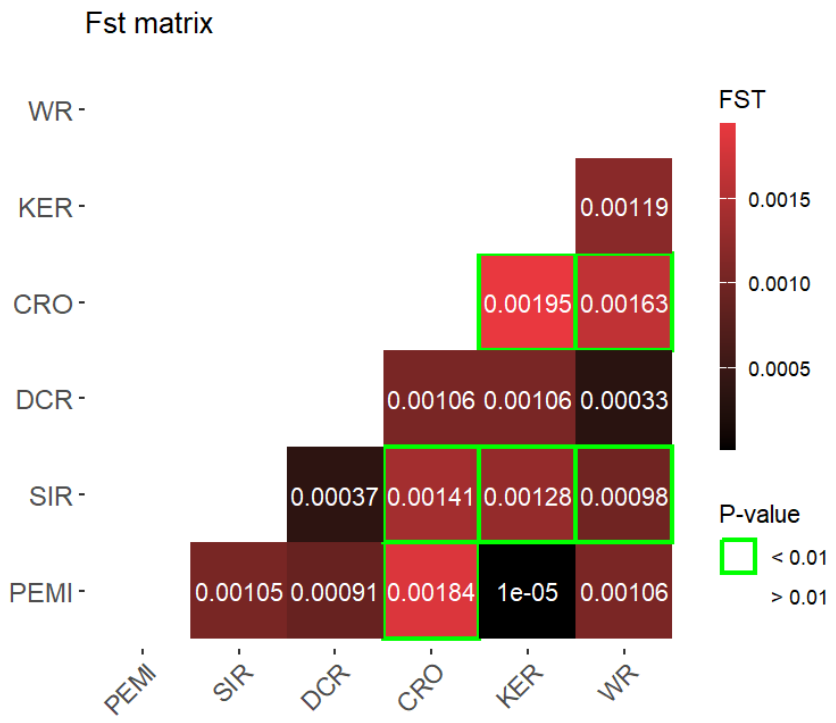


Fig.21. Pairwise FST matrix and Weir-Cockerham testing at p-value ≤ 0.01 for the four sampling regions: Prince Edward and Marion Islands (PEMI), SIR = South Indian Ridge, DCR = Del Cano Rise, , , CR = Crozet, Kerguelen (KER) and WR = Williams Ridge.

6.3.4 Additional analyses

Examining the dataset for individual SNPs potentially under selection, the Manhattan plot showed that most SNPs had FST values close to zero, with the exception of two loci (Fig. 22). These included 'SNP 86' (FST = 0.167), whose true location is unknown (Chromosome 0), and 'SNP 2622' (FST = 0.478) located on Chromosome 7 at position 20 664 386 (Fig. 22).

An investigation of 'SNP 2622' on Chromosome 7 using the NCBI Genome Data Viewer indicates that it is located in an intergenic region; however, due to limitations in the genome annotations, its functional significance is unclear. Observing allele frequencies shows that KER and PEMI have frequencies of ~50% for this locus, while the other populations are close to 0%, which is likely contributing to the high FST value seen for this particular locus. The differentiation of this particular locus could be due to genetic differentiation or technical issues such as differences in DNA quality and tissue type, but additional examination is needed to ascertain which.

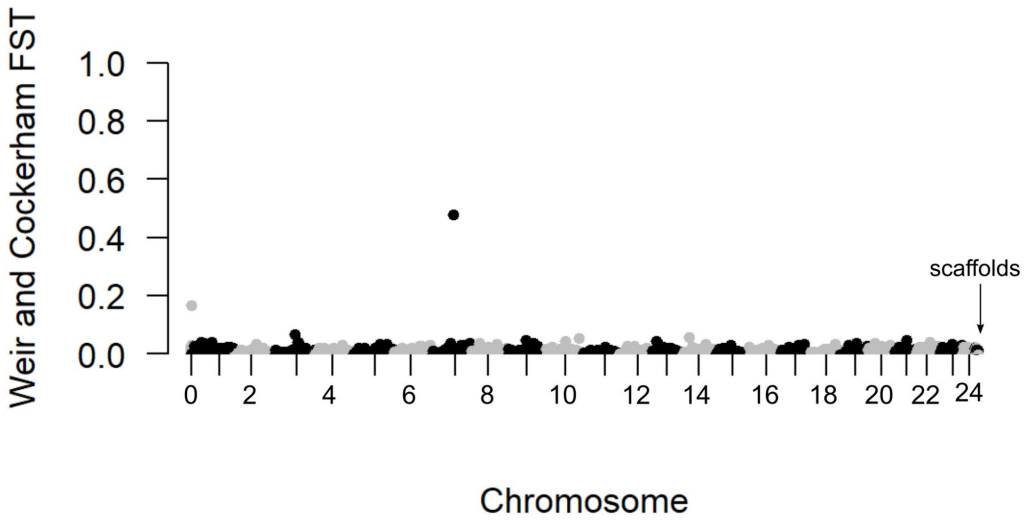


Fig.22. Manhattan plot for the entire D. eleginoides dataset, with chromosome number by FST (Weir and Cockerham). Each SNP is represented by a dot, and alternating black and gray colors and chromosomes are for ease of viewing SNP positioning.

The distribution of differentiation is visualized in Fig. 23, with FST values for the loci showing a peak near zero, supporting the presence of little population structure. Despite very few loci having higher FST values, at a threshold of 0.05, none of the loci were significantly different from the expected distribution and therefore are not considered under selection.

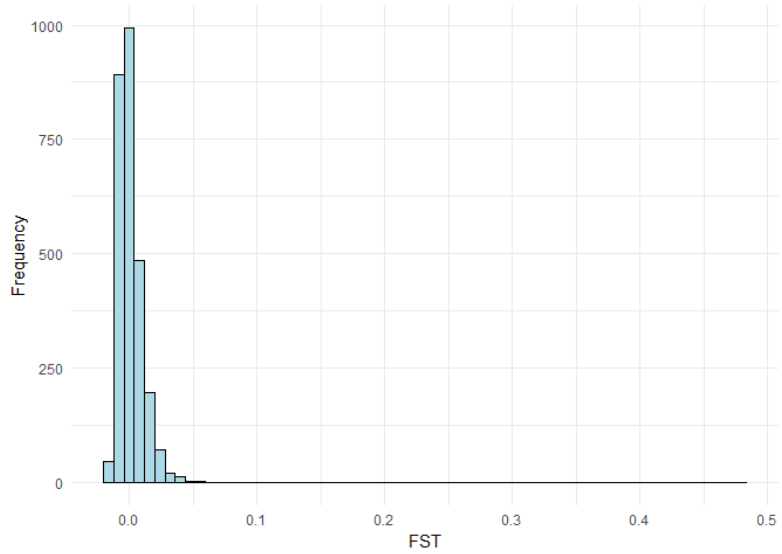


Fig.23. Distribution of differentiation illustrating FST values calculated using the Whitlock and Lotterhos method on the x-axis and frequency of loci on the y-axis.

The sex-ratio across the entire dataset (all sampling regions) was 0.46 (nearly 1:1). The positive trend between female and male heterozygous loci (blue dots) supports that all loci are autosomal, with similar levels of heterozygosity seen amongst females and males (Fig. 24). The sex-linked analysis therefore did not identify any loci that are sex-linked (Fig. 24, yellow XX/XY box).

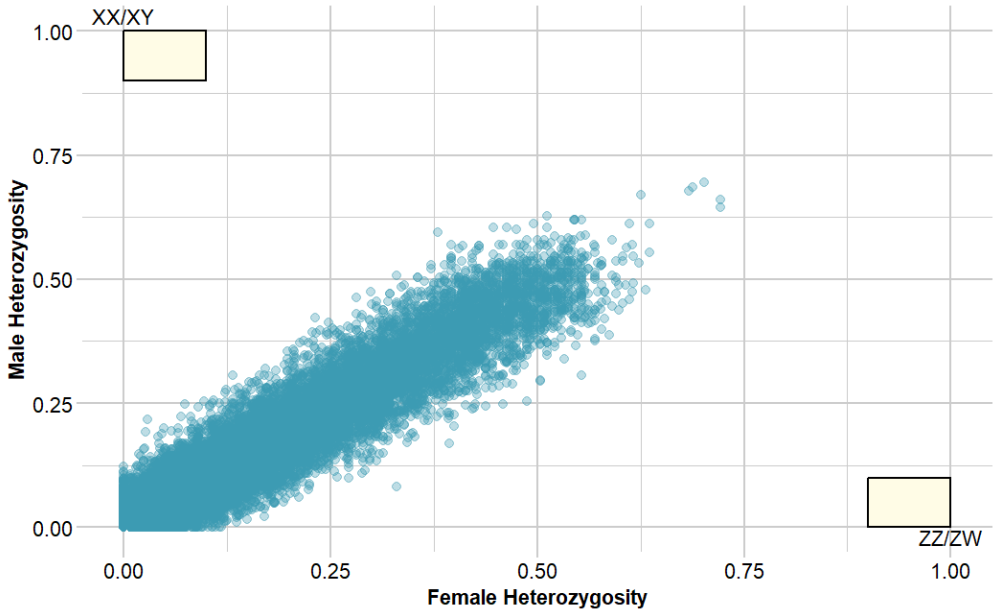


Fig.24. Sex-linkage plot identifying no presence of sex-linked loci. Based on metadata indicating sex, axes explain the proportion of individuals that are heterozygous at each locus (identified as the blue dots) for each sex, respectively, where 0 = all individuals are homozygous at the locus and 1 = all individuals are heterozygous at the locus. Loci are not present in the yellow XX/XY box, which would indicate complete heterozygosity at 1 for males (Y-linked loci). The yellow ZZ/ZW box is not applicable as D. eleginoides is an XX/XY species.

6.4 Potential habitat distribution

The habitat maps (Fig. 25) indicate that the potential habitat distribution for Patagonian toothfish extends across the CCAMLR/SIOFA border at DCR and WR and into the EEZs of the southern Indian Ocean at KER, CR, and PEMI. The habitat also extends further north beyond the current DCR management unit along the SIR.

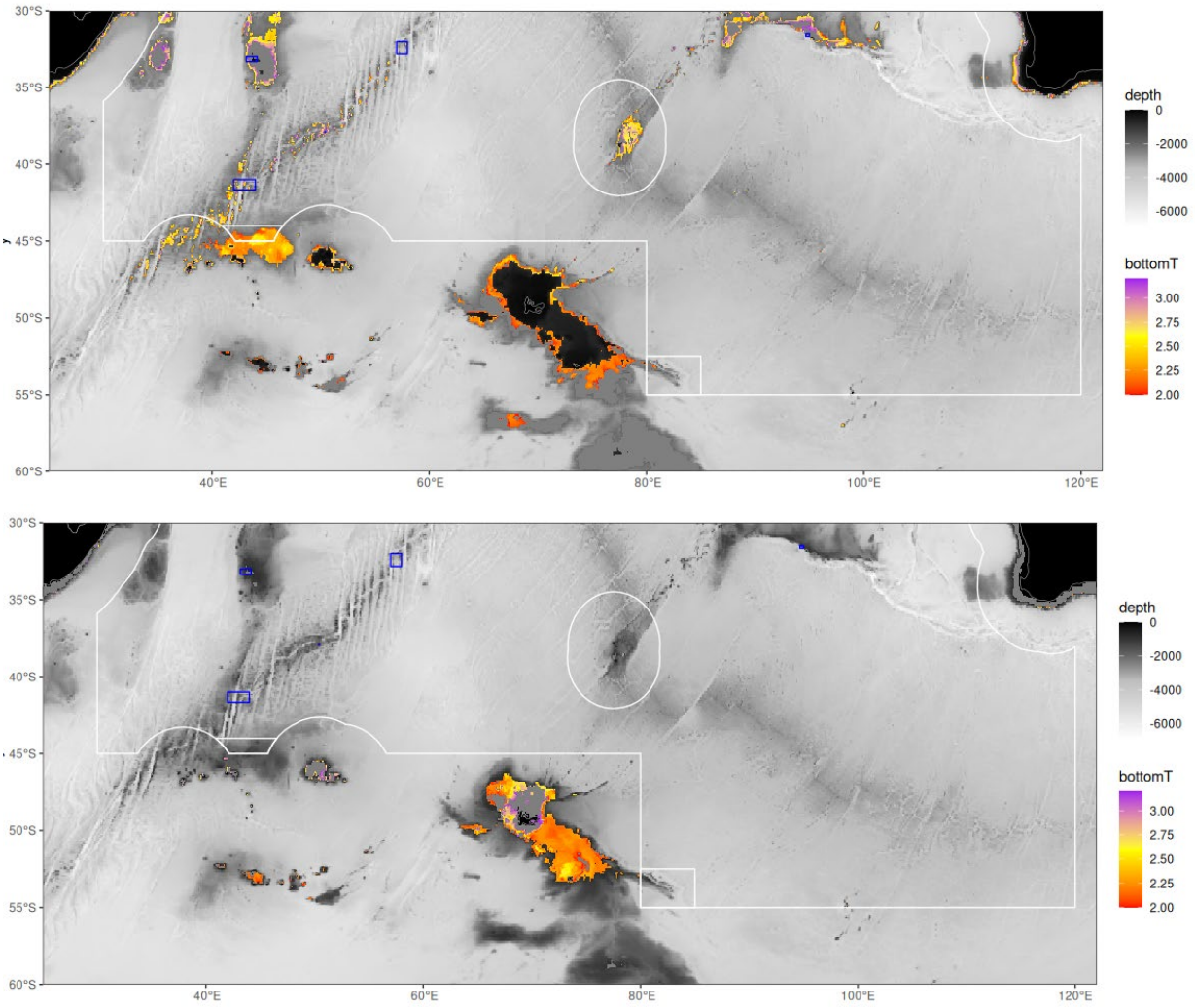


Fig.25. Potential habitat of adults (top panel) and juveniles (bottom panel) based on bathymetry (adults: 800-2000 m, juveniles: 50-800 m), slope (< 0.2 radians), and bottom temperature (2-3.2° C). All colored areas (red to purple) reflect the bottom temperature constrained within the potential habitat. SIOFA area and current management units of DCR and WR are indicated by the white polygons. VMEs within the SIOFA area are indicated by blue polygons.

7. Discussion

7.1 Population spatial structure

Multiple analyses conducted on a dataset consisting of > 2 700 SNP markers genotyped for 242 individuals from South Indian Ridge (n = 65), Del Cano Rise (n = 65), Williams Ridge (n = 34), as well as Crozet (n = 23), Kerguelen (m = 27) and Prince Edward and Marion Island (n = 28) in the CCAMLR region support the presence of a single panmictic population of D. eleginoides in the southwest Indian Ocean. We therefore conclude that these different fishing regions are most likely one population that traverses SIOFA and CCAMLR-managed areas.

DAPC clustering analyses of the SNP dataset identified the presence of one population cluster, with samples from different regions grouping closely together (Figs. 17 and 19). Accompanying STRUCTURE analyses further support the lack of genetic

structuring through the identification of high levels of admixture in individuals obtained from the different sampling regions (Fig. 20). With the addition of KER and PEMI samples into the dataset, there were significant differences in Fixation index values (FST) values between CR and several other regions (PEMI, SIR, KER and WR), as well as SIR and three other regions (CR, KER and WR), (Fig. 21) which was not seen previously (Nieblas et al. 2024, Fig. 21). While these slight differences could be linked to geographic distance, they are most likely related to variation generated by the addition of the samples from two sites KER and PEMI, particularly those from PEMI.

PEMI samples were muscle tissue, as opposed to fin, and their preparation varied from the protocol used by observers in the other regions (see Appendix 1); this difference in handling and storage possibly impacted DNA quality as there were several samples from PEMI that had to be discarded due to low quality (failing sequencing, not passing filtering thresholds). Differences in DNA quality and source tissue is known to heavily influence allele frequencies (Montgomery et al. 2005), possibly inflating FST estimates. Nonetheless, genetic distinctness of KER and PEMI are not supported by other analyses (DAPC, STRUCTURE) and the significant FST values remained close to zero, identifying low levels of genetic differentiation in further support of high levels of gene flow between the regions (Fig. 21).

Altogether, these results support the absence of genetic structuring across six regions, regardless of maturity stage or sex of the individuals. We suggest, however, that performing separate analysis excluding PEMI samples may provide further insight into how much variation these samples may introduce into the dataset. If the significant pairwise differences disappear without these samples, this will support that sample quality, rather than population differentiation, influenced these results. Since DArT generates the final sequencing reports, we will need to request reports excluding the PEMI samples, to complete downstream analyses. Given the current time constraints, we are unable to conduct this comparison before the present deadline; however, we plan to revisit it in the future and compare the results before and after removing these samples.

Previous studies also found limited support for population structuring among D. eleginoides in the southwest Indian Ocean, regardless of genetic-types (Appleyard et al. 2002, 2004; Toomey et al. 2016). Appleyard et al. 2002 reported low levels of differentiation through mitochondrial DNA (mtDNA). This suggested that female philopatry and male dispersal could lead to different patterns of genetic structuring as alleles associated with females might persist in the population, while alleles associated with males would show less structure due to dispersal driven gene flow. In our analyses of SNP marker data, however, we found no evidence in support of population structuring, whether driven by female-relatedness or otherwise. Appleyard et al. 2004 implemented mtDNA and microsatellite markers to screen samples from Crozet, Kerguelen, Prince Edward and Marion Islands (PEMI), and Heard and McDonald Islands (HIMI), finding weak evidence of population structuring depending on the locus, as well as non-significant associations between geographic and genetic distances. More recently, Toomey et al. 2016 analyzed mtDNA and nuclear marker data of samples retrieved from HIMI, Kerguelen, and Crozet, revealing some evidence of population structuring between HIMI and Crozet (mtDNA), possibly due to the differences in sample numbers between the two sites (30 vs. > 100) as well as specificity of some haplotypes at Crozet shared by only a few individuals. However, no differentiation between HIMI and Kerguelen, nor Kerguelen and Crozet was seen in either mtDNA or nuclear datasets (Toomey et al. 2016). Findings from Toomey and colleagues underscore that maternally inherited mtDNA is more sensitive to genetic drift than nuclear markers, while also supporting the necessity of having higher sample sizes to detect subtle differentiation in the form of rare haplotypes in populations. As such, we suggest increasing sample sizes for Crozet, as SNP loci were genotyped for only 24 individuals obtained from this region.

The genetic homogeneity seen in western Indian Ocean toothfish is likely linked to both the movement of adults and the dispersal of early life stages (Appleyard et al. 2002, 2004). Tagging studies have shown the increasing vertical (ontogenetic) migration of adult individuals with increasing age and short-distance (< 25 km) horizontal migration once fish settle at depth, with some individuals making short forays around slopes (<130 km for 6 months) (Brown et al. 2013, Welsford et al. 2014; Williams et al., 2002; Marlow et al.,2003; Tuck et al., 2003). While D. eleginoides is known to be a more residential species, long distance migrations (> 2 000 km) have also been recorded, likely facilitated by ridges, seamounts and island chains that serve as 'stepping stones' that allow the gradual movement between regions and more isolated areas of the southwest Indian Ocean (Rogers et al. 2006, Péron et al. 2016). One such adult migration was recorded between HIMI and Crozet (Williams et

al. 2002), which is a comparable distance between WR and CR in our dataset (~ 2 400 km), and very little genetic differentiation was identified across this spatial distance (Fig. 21). Although adult movement over vast distances is apparently rare, the absence of known population structure in the southwest Indian ocean could suggest that even a small number of individuals migrating between regions can help facilitate gene flow across the area, enabling the connection between D. eleginoides sampling regions.

In contrast to adults, early life stages are thought to have broader dispersal potential due to the lengthy egg and larval stages that could also facilitate long distance dispersal through oceanic processes (Evseenko et al. 1995; Belchier and Collins 2008). D. eleiginoides adults typically spawn at depths > 800 m, releasing positively buoyant eggs that drift to shallower depths (< 500 m) during an estimated incubation time of 3 months. Following hatching, larvae remain in the upper water column (< 250 m) over a period of at least 3 months until metamorphosis (Koubbi et al. 1990, Kock and Kellerman 1991, Evseenko et al. 1995, Collins et al. 2010, Harte 2020). The survival and recruitment success of early life stages relies on the availability of food during dispersal among 'stepping stones', with increased productivity enhanced by localized upwelling around these areas (Shelton and Hutchings 1982; Bakun 1996; Tolimieri et al. 2018; Mori 2013; Park et al. 2014). With regards to dispersal, Mori et al. 2013 modeled egg and early larvae transport in the southwest Indian Ocean (Kerguelen Plateau), identifying that most of the eggs spawned in western regions of the plateau were successfully settled larvae east of the plateau, suggesting frequent cross plateau transport dynamics. The updated dataset containing samples from six regions that showed small significant differences between WR and some northern regions (SIR and CR), but not others (DCR and PEMI) suggests that connectivity between these areas is possibly linked to the Kerguelen plateau, with Kerguelen Island possibly serving as an intermediate point of connectivity. The integration of samples from KER into the dataset supported the presence of connectivity between WR and KER through the lack of overall genetic differentiation between these two regions (Fig 21).

7.2 Satellite tagging

In SER2022-TOP1, we observed that the fish travel both short and long distances. Fish that moved relatively short distances likely traveled along the same or similar isobaths without needing to cross deep water areas to reach recovery sites. We found that the continuity of the depth range indicated that these are likely individuals of the same population. Those fish that traveled long distances must have crossed a deep water divides (> 4000 m). As noted above, though rare, in the absence of known population structure, even a small number of individuals migrating between regions can help facilitate gene flow across the area.

Further studies could employ satellite tagging of Patagonian toothfish to gain a better understanding of migration pathways, providing insights into connectivity within the population.

7.4 SIOFA area and connections to CCAMLR

We note several points of evidence that indicate a close connection between the CCAMLR and SIOFA areas. Bathymetric features and potential habitat span the CCAMLR and neighbouring SIOFA zones, potentially facilitating ontogenetic migration. Length distribution data indicate that juveniles and smaller individuals from the adjacent, shallower CCAMLR regions appear to migrate into the deeper SIOFA waters as they age (Fig.25). Tagging studies show some movement between the Kerguelen Plateau and Del Cano Rise, as well as between the Kerguelen Plateau and Williams Ridge (Nieblas and Cowart 2023). Assuming a single population, the data review and the potential habitat mapping suggest that SIOFA fishing areas likely target the same population as the adjacent CCAMLR region. Similarly, fishing across SIOFA in SIR, DCR, WR (and even Eastern, see Nieblas and Cowart 2023) are likewise fishing the same population.

8. Recommendations and conclusions

8.1 Proposed management units

Given the findings of the genetic population analysis suggesting a single population of the target species across the SIOFA region, extending into CCAMLR areas and EEZs in the southern Indian Ocean, the following recommendations are proposed to provide SIOFA with an adaptable approach to facilitate the necessary groundwork for a fully integrated, collaborative management program with CCAMLR.

Option 1. Expand Management Units to Encompass Potential Habitat to the North (Strategic Expansion)

Extend SIOFA management units northward to incorporate known habitat areas of both juvenile and adult populations, aiming to account for the stock's full range and fishing effort within the SIOFA-managed area. Expansion into potential habitat areas, particularly along SIR (proposed management unit extension in Fig.26, black polygon), which are also coincident with significant effort and catches since 2003 (Nieblas and Cowart 2023) and includes the Coral Point closed-area VME, would increase SIOFA's capacity to address the full spatial extent of the stock, capturing vital life stages and habitat areas that are currently unprotected, as well as potentially improve management for a key SIOFA VME. While expanding units may increase administrative and enforcement requirements, it enables a more precautionary approach by mitigating potential range-edge exploitation.

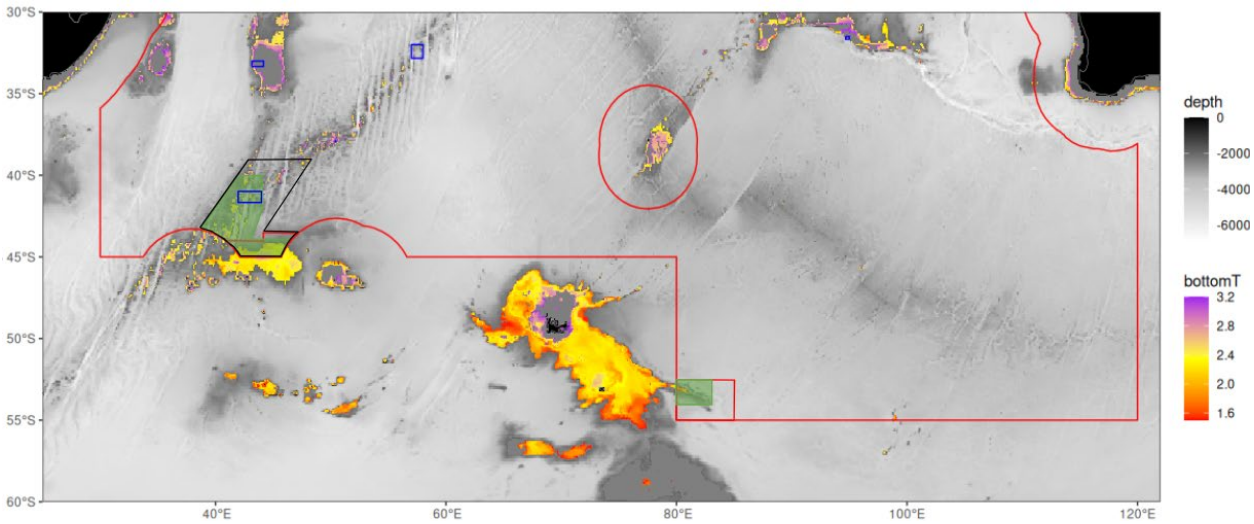


Fig.26. Proposed extension of the current DCR management unit (black polygon) to include DCR and SIR fishing grounds (green polygon) and the Coral Point VME (overlaid blue polygon). Other VMEs (blue) within the SIOFA area (red) are also indicated. Current DCR and WR management units are also indicated in red. The color gradient red to purple indicates potential habitat for D. eleginoides.

Option 2. Harmonize Management Measures Across Existing Units (Interim Solution)

Implement consistent management measures across the two existing SIOFA management units, treating the stock as a single population to improve coherence and sustainability. Harmonizing measures across units supports a unified approach to management that better reflects the stock's biology, potentially reducing local depletion risks. Aligning regulatory frameworks within SIOFA provides a foundation for future joint management if collaboration with CCAMLR is pursued. SIOFA should enhance coordination with adjacent jurisdictions to track the impact of harmonized measures on stock health and ecosystem stability.

Option 3. Develop a Joint Management Program with CCAMLR (Long-Term Solution)

Establish a collaborative management framework between SIOFA and CCAMLR, recognizing the stock as a transboundary population and ensuring cohesive management across its entire range. A joint management approach directly aligns with the biological distribution of the stock, promoting sustainable harvest practices and reducing fragmentation in management strategies. A coordinated framework with CCAMLR could provide consistent monitoring, shared data collection, and harmonized management responses, thereby enhancing both conservation and compliance outcomes. This long-term collaborative approach offers the best alignment with the stock's biological and ecological realities, supporting sustainable harvest practices and enhancing regulatory efficacy across the SIOFA and CCAMLR regions.

Recommended Steps:

1) **Memorandum of Understanding (MoU)**

Initiate a formal MoU between SIOFA and CCAMLR to outline cooperative intent, shared objectives, and agreed-upon principles for managing this transboundary stock.

2) **Establish a Joint Working Group**

Form a Joint Working Group consisting of scientific, management, and legal representatives from both SIOFA and CCAMLR. The group would focus on aligning stock assessment, research priorities, and the development of compatible management measures.

3) **Create a data-sharing agreement**

Develop a data-sharing protocol to ensure that both SIOFA and CCAMLR have access to up-to-date information on stock assessments, genetic analyses, and habitat models, supporting a comprehensive view of the stock's status.

4) **Develop pilot joint management measures**

As a first phase, the Joint Working Group could establish pilot measures (e.g., shared catch limits or area-based restrictions) within overlapping jurisdictional areas, allowing both organizations to trial joint initiatives. Based on pilot outcomes, a full-scale, jointly managed framework could be implemented, subject to adaptive management reviews and feedback from both SIOFA and CCAMLR.

8.1.1 Summary recommendations of management units

Recognising that CCAMLR, as a conservation organisation and SIOFA, as an RFMO, have similar objectives towards the sustainable use of marine resources, and in light of the stock's shared population structure and habitat across SIOFA, CCAMLR, and adjacent EEZs, Option 3, developing a joint management framework with CCAMLR, represents the most sustainable and biologically sound solution. However, recognizing the complexity of establishing a transboundary management system, it is recommended that Option 2, harmonizing existing units, be pursued in the interim to improve alignment within SIOFA' s jurisdiction. Finally, Option 1, expanding management units northward, should be implemented regardless of whether Options 2 and 3 are implemented, considering that a large area of potential habitat is currently being fished outside any management area.

8.2 Future genetic studies

The genetic study conducted in SER2022-TOP2 identified a single population in the southern Indian Ocean, suggesting that this stock could extend beyond the SIOFA and CCAMLR zones and potentially represent a single, connected population across the entire Southern Ocean. This is especially possible given the potential for larval connectivity through the Antarctic Circumpolar Current system, which may facilitate the transport of larvae across vast distances and could play a critical role in maintaining genetic flow and population structure. Given this possibility, it is recommended to expand genetic studies into additional Southern Ocean regions and increase sampling efforts, particularly across the CCAMLR boundary, to confirm population connectivity on a larger scale. This continued research would provide valuable insights into stock structure and inform more cohesive and adaptive management strategies across jurisdictions.

The comprehensive evaluation of stock structure and population size for management of a species requires drawing on genetic, demographic, ecological and life history studies (Waples 1998). Given its economic importance and knowledge of a single panmictic population in the southwest Indian Ocean, Dissotichus eleginoides is a natural candidate for subsequent Close Kin Mark and Recapture (CKMR) and epigenetic aging studies targeting the estimation of its population size (Bravington and Carroll et al. 2023).

CKMR analysis applies the principles of genetic relatedness between individuals, which can be identified from SNP-loci datasets, to derive a population size estimate thus providing abundance indices unbiased by fisheries dependent calculations (e.g., catch and effort data, CPUE). This approach may also help clarify the complexities of demographic traits of a species such as fecundity and mortality (Bravington et al. 2016; Hillary et al. 2018; Wacker et al. 2021; Trenkel et al. 2022). Notably, CKMR studies typically require well defined sampling programs to amass a large sample repository consisting of female and male individuals of differing ages (adults and juveniles) to recover familial links between individuals.

Accurate age information is therefore critical to identify clear distinctions between age-cohorts and could be obtained through an 'epigenetic clock'. Epigenetic aging tools, which have previously been developed in other fish species (Bonhommeau et al, in prep), rely on the examination of changes in DNA methylation— an epigenetic modification where a methyl group is added to cytosine-guanine loci—which correlates with aging (Piferrer and Anastasiadi 2023). Future studies implementing auxiliary genetics-based methods for assessing individual relatedness and age prediction can aid the management and conservation of the co-administered Patagonian toothfish stocks.

Further research using genetic techniques is recommended to clarify population structure, improve demographic assessments, and enhance management strategies.

1. Continue genetic sampling for a pan-Southern Ocean population discrimination study

The initial genetic study has indicated a single population within the southern Indian Ocean, with a possibility of a connected stock across the entire Southern Ocean. Further sampling across different regions, including the CCAMLR boundary and adjacent areas, is required to confirm connectivity and population structure.

- **Establish a clearer understanding of the stock structure of Patagonian toothfish across the Southern Ocean.**
- **Develop a sampling protocol with standardized genetic markers to allow for consistent data comparison and reliable population discrimination analyses.**
- **Collaborate with CCAMLR and CPs to coordinate sampling efforts and data sharing to maximize the efficiency and reach of the study.**

2. Pilot project to identify sampling strategy for Close-Kin Mark-Recapture (CKMR) study

CKMR is a powerful tool for assessing population size and demographic parameters without requiring large-scale physical mark-recapture efforts. A pilot project is essential to determine the feasibility, optimal sampling rates, and potential coverage required for a full CKMR study on Patagonian toothfish. A pilot CKMR study should focus on identifying the necessary sample sizes, geographic range, and logistical requirements to scale up to a comprehensive CKMR study.

3. Pilot project to identify age and sex of Patagonian toothfish using epigenetics

Traditional aging methods, such as otolith analysis, are time-consuming, costly, and biased by the observer. Sex can only be determined by emptying the fish and examining the gonads. Epigenetic markers offer a potential alternative for determining age and sex, which could streamline demographic analysis, and also provide key information for stock assessment.

- **Select a sample cohort representing various ages and both sexes to validate epigenetic assays and refine methodologies for accurate demographic assessment.**
- **Develop a protocol for integrating epigenetic age and sex determination into routine assessments.**

4. Implement standard genetic sampling by observers

Integrate routine genetic sampling into observer programs to build a large, geographically diverse genetic dataset over time.

- **Systematic collection of genetic samples through observer programs offers an efficient and cost-effective means of gathering samples necessary for future genetic studies. Routine genetic sampling will support both current and future research efforts by creating a repository of genetic information from across the species' range.**
- **Standardize genetic sampling protocols for observers, setting a target sampling frequency, such as collecting a genetic sample from every 50th or 100th fish caught. Provide training to observers on sample collection, handling, and data recording to ensure high-quality, usable genetic samples. Sampling materials are relatively low cost, requiring squirt bottles of alcohol and bleach, gloves, and sampling tubes. Current sampling time required to properly (hygienically) sample a fin clip is about 3-5 minutes. We are testing alternative strategies to reduce this sampling time.**
- **Coordinate with CCAMLR, CPs and observer management companies to implement standardized sampling across jurisdictions, enhancing the comparability and utility of genetic data.**
- **Establish a centralized database for collected samples, accessible to relevant research and management entities to support ongoing and future studies.**

9. Acknowledgments

We thank the SIOFA Secretariat and the SER2022-TOP2 Scientific Committee Advisory Panel for their support and guidance throughout the development of this report.

We would like to thank the observer managers and the observers that assisted in the collections of the samples, including Rhys Arangio of Austral Fisheries, Justine Johnston of AFMA, Captain Emmanuel Le Roy and Tim Lamb of the Cape Arkona, Louis Wamberge and Edward Caput of TAAF, Christian Engelbrecht of CapMarine, Roberto Sarralde Vizuete of IEO-CSIC, Humberto Aguirre Figueras, and Melanie Williamson of Capfish.

This study was funded under the SIOFA-SEAs project (EU Grant n. 101078892).

10. References

1. **Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. Journal of molecular biology, 215(3), 403-410.**
2. **Anderson, C. A., Pettersson, F. H., Clarke, G. M., Cardon, L. R., Morris, A. P., & Zondervan, K. T. (2010). Data quality control in genetic case-control association studies. Nature protocols, 5(9), 1564-1573.**
3. **Appleyard, S. A., Ward, R. D., & Williams, R. 2002. Population structure of the Patagonian toothfish around Heard, McDonald and Macquarie Islands. Antarctic Science, 14(4), 364-373.**
4. **Appleyard, S. A., Williams, R., & Ward, R. D. 2004. Population genetic structure of Patagonian toothfish in the West Indian Ocean sector of the Southern Ocean. CCAMLR Science, 11, 21-32.**
5. **Arana, P. (2009). Reproductive aspects of the Patagonian toothfish (Dissostichus eleginoides) off southern Chile. Latin American Journal of Aquatic Research, 37(3), 381-393.**
6. **Avise, J.C., 1998. Conservation genetics in the marine realm. J. Hered. 89, 377–382.**
7. **Bakun, A. (1996) Patterns in the Ocean: Ocean Processes and Marine Population Dynamics, Univ. California Sea Grant, San Diego, CA.**
8. **Belchier, M., & Collins, M. A. (2008). Recruitment and body size in relation to temperature in juvenile Patagonian toothfish (Dissostichus eleginoides) at South Georgia. Marine Biology, 155, 493-503.**
9. **Brown, J., Brickle, P., & Scott, B. E. (2013). Investigating the movements and behaviour of Patagonian toothfish (Dissostichus eleginoides Smitt, 1898) around the Falkland Islands using satellite linked archival tags. Journal of Experimental Marine Biology and Ecology, 443, 65-74.**
10. **Bravington, M., Grewe, P. & Davies, C. Absolute abundance of southern bluefin tuna estimated by close-kin mark-recapture. Nat Commun 7, 13162 2016.**
11. **Bravington, M. V., & Carroll, E. L. (2023). 11 The practical magic of close-kin mark-recapture. Applied Environmental Genomics, 120.**
12. **Canales-Aguirre, C. B., Ferrada-Fuentes, S., Galleguillos, R., Oyarzun, F. X., & Hernández, C. E. 2018. Population genetic structure of Patagonian toothfish (Dissostichus eleginoides) in the Southeast Pacific and Southwest Atlantic Ocean. PeerJ, 6, e4173.**
13. **Chevrier, T., Cowart, D.A., Nieblas, A-E, Charrier, G., Bernard, S., Evano, H., Brisset, B., Chanut, J., Bonhommeau, S. Population structure of the swordfish, Xiphias gladius, across the Indian Ocean using next-generation sequencing. ICES Journal of Marine Science (2024): fsae179.**
14. Clarke, A., Barnes, D. K., & Hodgson, D. A. (2005). How isolated is Antarctica?. *Trends in Ecology & Evolution*, *20*(1), 1-3.
15. Collins, M. A., Brickle, P., Brown, J., & Belchier, M. 2010. The Patagonian toothfish: biology, ecology and fishery. *Advances in marine biology*, *58*, 227-300.

16. Courtois B, Audebert A, Dardou A, Roques S, Ghneim- Herrera T, et al. (2013) Genome-Wide Association Mapping of Root Traits in a Japonica Rice Panel. PLoS ONE 8(11): e78037
17. Cruz VM, Kilian A, Dierig DA (2013) Development of DArT marker platforms and genetic diversity assessment of the U.S. collection of the new oilseed crop lesquerella and related species. PLoS One 8(5): e64062.
18. Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., ... & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156-2158.
19. Elshire RJ, Glaubitz JC , Sun Q, Poland JA, Kawamoto K, et al. (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One, 6(5):e19379.
20. Everson, I., & Murray, A. (1999). Size at sexual maturity of Patagonian toothfish. CCAMLR Sci, 6, 37-46.
21. Evseenko, S. A., Kock, K. H., & Nevinsky, M. M. 1995. Early life history of the Patagonian toothfish, *Dissostichus eleginoides* Smitt, 1898 in the Atlantic sector of the Southern Ocean. *Antarctic Science*, *7*(3), 221-226.
22. FAO 2024 Vulnerable Marine Ecosystems DatabaseA global inventory of fisheries measures to protect vulnerable marine ecosystems in areas beyond national jurisdiction.https://www.fao.org/in-action/vulnerable-marine-ecosystems/vme-database/en/vme.html. Accessed 21 November 2024.
23. Francis, R., Hixon, M., Clarke, M., Murawski, S., Ralston, S., 2007. Ten commandments for ecosystem-based fisheries scientists. Fisheries 32, 217–233.
24. Harte, E. (2021). *Testing the connectivity of Dissostichus eleginoides (Patagonian toothfish) between the Pacific coast of southern Chile and the Patagonian Shelf in the southwest Atlantic* (Doctoral dissertation, University of Otago).
25. Gruber, B., Unmack, P. J., Berry, O. F., & Georges, A. (2018). dartr: An r package to facilitate analysis of SNP data generated from reduced representation genome sequencing. *Molecular ecology resources*, *18*(3), 691-699.
26. Hillary, R.M., Bravington, M.V., Patterson, T.A. et al. Genetic relatedness reveals total population size of white sharks in eastern Australia and New Zealand. *Sci Rep* 8, 2661 (2018).
27. Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: a new
28. method for the analysis of genetically structured populations. *BMC genetics*, *11*, 1-15.
29. Kilian, A., Wenzl, P., Huttner, E., Carling, J., Xia, L., Blois, H., ... & Uszynski, G. (2012). Diversity arrays
30. technology: a generic genome profiling technology on open platforms. *Data production and analysis in*
31. *population genomics: Methods and protocols*, 67-89.
32. Kock, K. H., & Kellermann, A. 1991. Reproduction in Antarctic notothenioid fish. *Antarctic Science*, *3*(2), 125-150.
33. Koubbi, P., Duhamel, G., & Camus, P. (1990). Early life stages of Notothenioidei from the Kerguelen Islands. *Cybium*, *14*(3), 225-250.
34. Lischer, H. E., & Excoffier, L. (2012). PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, *28*(2), 298-299.
35. Lord, C., Duhamel, G., & Pruvost, P. (2006). The patagonian toothfish (Dissostichus eleginoides) fishery in the Kerguelen Islands (Indian Ocean sector of the Southern Ocean). CCAMLR Science, 13, 1-25.
36. Marlow TR, Agnew DJ, Purves MG, Everson I (2003) Movement and growth of tagged *Dissostichus eleginoides* around South Georgia and Shag Rocks (Subarea 48.3). CCAMLR Sci 10:101–112
37. Montgomery, G. W., Campbell, M. J., Dickson, P., Herbert, S., Siemering, K., Ewen-White, K. R., ... & Martin, N. G. (2005). Estimation of the rate of SNP genotyping errors from DNA extracted from different tissues. Twin Research and Human Genetics, 8(4), 346-352.
38. Montes, I., Conklin, D., Albaina, A., Creer, S., Carvalho, G. R., Santos, M., & Estonba, A. (2013). SNP discovery in European anchovy (Engraulis encrasicolus, L) by high-throughput transcriptome and genome sequencing. *PLoS One*, *8*(8), e70051.
39. Mori, M., Corney, S. P., Melbourne-Thomas, J., Welsford, D. C., Klocker, A., & Ziegler, P. E. (2016). Using satellite altimetry to inform hypotheses of transport of early life stage of Patagonian toothfish on the Kerguelen Plateau. *Ecological Modelling*, *340*, 45-56.
40. Morin, P. A., Luikart, G., Wayne, R. K., & SNP Workshop Group. (2004). SNPs in ecology, evolution and conservation. Trends in ecology & evolution, 19(4), 208-216.

41. Park, Y.-H., Durand, I., Kestenare, E., Rougier, G., Zhou, M., d' Ovidio, F., Cotté, C., Lee,J.-H., 2014. Polar Front around the Kerguelen Islands: an up-to-date determination and associated circulation of surface/subsurface waters. Journal of Geophysical Research: Oceans 119, 6575–6592.
42. Shaw, P. W., Arkhipkin, A. I., & Al-Khairulla, H. 2004. Genetic structuring of Patagonian toothfish populations in the Southwest Atlantic Ocean: the effect of the Antarctic Polar Front and deep-water troughs as barriers to genetic exchange. Molecular Ecology, 13(11), 3293-3303.
43. Shelton PA, Hutchings L (1982) Transport of anchovy, *Engraulis capensis* Gilchrist, eggs and early larvae by a frontal jet current. J Cons Int Explor Mer 40:185–198.
44. Smith, P., & McVeagh, M. 2000. Allozyme and microsatellite DNA markers of toothfish population structure in the Southern Ocean. Journal of Fish Biology, 57, 72-83
45. Pembleton, L. W., Cogan, N. O., & Forster, J. W. (2013). St AMPP: An R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Molecular ecology resources*, *13*(5), 946-952.
46. Péron, C., Welsford, D. C., Ziegler, P., Lamb, T. D., Gasco, N., Chazeau, C., Sinègre, R., Duhamel, G. (2016). Modelling spatial distribution of Patagonian toothfish through life-stages and sex and its implications for the fishery on the Kerguelen Plateau. *Progress in Oceanography*, *141*, 81-95.
47. Piferrer, F., & Anastasiadi, D. (2023). Age estimation in fishes using epigenetic clocks: Applications to fisheries management and conservation biology. *Frontiers in Marine Science*, *10*, 1062151.
48. Porras-Hurtado, L., Ruiz, Y., Santos, C., Phillips, C., Carracedo, Á., & Lareu, M. V. (2013). An overview of STRUCTURE: applications, parameter settings, and supporting software. *Frontiers in genetics*, *4*, 98.
49. Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945-959.
50. R Core Team (2023) R: A Language and Environment for Statistical Computing. R Foundation for
51. Statistical Computing.
52. Raman H, Raman R, Kilian A, Detering F, Carling J, et al. (2014) Genome-Wide Delineation of Natural Variation for Pod Shatter Resistance in Brassica napus. PLoS ONE 9(7). e101673.
53. Rogers, A. D., Morley, S., Fitzcharles, E., Jarvis, K., & Belchier, M. 2006. Genetic structure of Patagonian toothfish (Dissostichus eleginoides) populations on the Patagonian Shelf and Atlantic and western Indian Ocean Sectors of the Southern Ocean. Marine Biology, 149(4), 915-924.
54. Sansaloni, C., Petroli, C., Jaccoud, D., Carling, J., Detering, F., Grattapaglia, D., & Kilian, A. (2011, September). Diversity Arrays Technology (DArT) and next-generation sequencing combined: genome-wide, high throughput, highly informative genotyping for molecular breeding of Eucalyptus. In *BMC proceedings* (Vol. 5, No. Suppl 7, p. P54). London: BioMed Central.
55. Song, H., Marshall, J., Follows, M. J., Dutkiewicz, S., & Forget, G. (2016). Source waters for the highly productive Patagonian shelf in the southwestern Atlantic. Journal of Marine Systems, 158, 120-128.
56. Toomey, L., Welsford, D., Appleyard, S. A., Polanowski, A., Faux, C., Deagle, B. E., Belchier, M., Marthick, J., Jarman, S. (2016). Genetic structure of Patagonian toothfish populations from otolith DNA. Antarctic Science, 28(5), 347-360.
57. Tolimieri, M. A. Haltuch, Q. Lee, M. G. Jacox, S. J. Bograd (2018) Oceanographic drivers of sablefish recruitment in the California Current. Fisheries Oceanography. 27(5):458-474
58. Trenkel V, G. Charrier, P. Lorance, M. Bravington. Close-kin mark–recapture abundance estimation: practical insights and lessons learned. *ICES Journal of Marine Science*, 2022, 79 (2), pp.413-422. ff10.1093/icesjms/fsac002ff. Ffhal-03760486f
59. Tuck GN, de la Mare WK, Hearn WS, Williams R, Smith ADM, He X, Constable A (2003) An exact time of release and recapture stock assessment model with an application to Macquarie Island Patagonian toothfish (*Dissostichus eleginoides*).
60. Wacker S., Hans J. Skaug, Torbjørn Forseth, Øyvind Solem, Eva M. Ulvan, Peder Fiske, Sten Karlsson (2021) *Ecology and Evolution*. 11(9):3917-3932. doi/10.1002/ece3.7279
61. Waples, R. 1998. Identifying stock structure and resolving stock mixtures in tuna. *Col. Vol. Sci. Pap. ICCAT* 50, 207-208

62. Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *evolution*, 1358-1370.
63. Welsford, D.C., Péron, C., Ziegler, P.E., Lamb, T., 2014. Development of the Patagonian toothfish *(Dissostichus eleginoides)* Tagging Program in Division 58.5.2, 1997–2014. CCAMLR Conference Paper WG-FSA-14/43.
64. Williams, R., & Tuck, G. N. 2002. Movement, growth and available abundance to the fishery of *Dissostichus eleginoides* Smitt, 1898 at Heard Island, derived from tagging experiments. *CCAMLR Science*, *9*, 33-48.
65. Whitlock, M. C., & Lotterhos, K. E. (2015). Reliable detection of loci responsible for local adaptation: Inference of a null model through trimming the distribution of F ST. *The American Naturalist*, *186*(S1), S24-S36.

11. Appendices

Appendix 1: Sample packing protocol for submission to DArT

DArT provides instructions for packing samples for shipment on their website https://help.diversityarrays.com/v1/docs/how-to-pack-and-ship-samples which also includes a YouTube video that is linked in this document. The number of samples is n = 188. Approximately 10 - 15 mg of each sample will need to be placed in 150 µl of absolute ethanol, distributed amongst two plates of racked tubes sealed with caps.

Supplies

- Nitrile gloves
- Absolute alcohol
- 70% ethanol
- 10 % bleach
- Water (distilled or MilliQ)
- Plastic weigh boats or parafilm squares
- Balance
- Metal forceps
- Metal scalpels
- Metal scissors
- Ice bucket
- Paper towels
- Double zip sandwich bags (3 l)
- Marker
- Racked tubes and barrettes
- Pipette and pipette tips
- Pipette or tube boxes for racked tubes
- Large and thick rubber bands or tape
- Sample list

Notes before beginning

- Sample names from the sample list should be written on the side of each tube.
- Each tube should contain approximately 150 µl of absolute ethanol.
- Sample size should measure 10 - 15 mg.

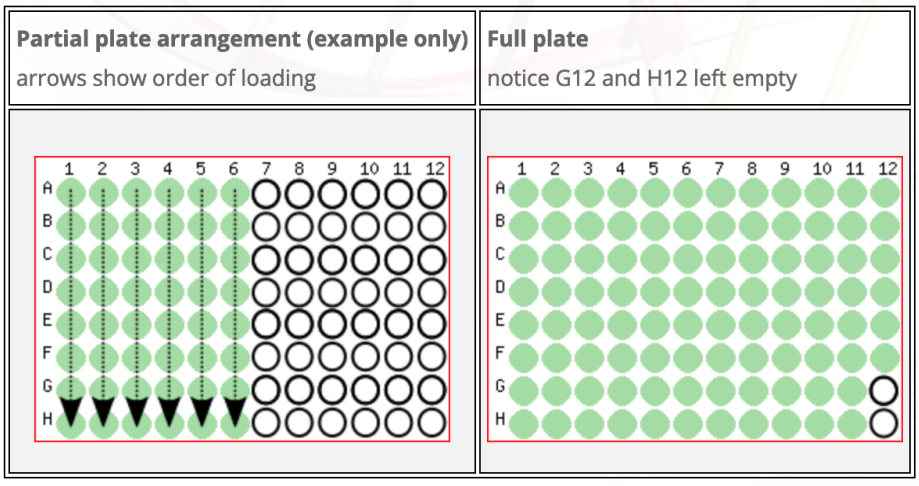
Sample preparation video

https://www.youtube.com/watch?v=fs2Y-FU_iWI

Plate preparation

1. Don gloves.
2. Decontaminate lab bench or hood surface using 10% bleach solution and wipe dry, followed by thorough rinsing with water and wipe dry.
3. Remove a strip of racked tubes and place them in a tube rack, arranged in a plate orientation (e.g., first column is A1 → H1, the second column is A2 → H2, etc, **Figure 1**).
4. Label each tube clearly using a marker, in accordance with the sample list file provided.
5. Proceed to add 150 µl volume of absolute ethanol to every well of the plate, **except wells G12 and H12**, (the last two weeks) using a pipette.
    1. Note that the volume of ethanol in each tube should be consistent and enough to submerge the entire sample.
    2. It is not necessary to fill the entire tube full of ethanol.

- Make sure you comply with requirement **Empty wells** outlined above related to creation of the sample tracking file .



**green circle represents a well filled with the sample, white circle represents an empty well*

Figure 1: Orientation for loading samples into the wells as requested by DArT. Note the difference between the partially-filled plate (left) and full plate (right). Wells G12 and H12 must always remain empty.

Adding tissues to ethanol-filled strips/plates

1. Decontaminate the surface of the balance surface 10% bleach solution and wipe dry, followed by thorough rinsing with water and wipe dry.
2. Prior to handling the samples, disinfect cutting equipment.
    1. Sterilizing forceps and scissors can be done by using 10% bleach solution, followed by rinsing with water or 70% ethanol.
    2. Please be sure to use a clean paper towel or kim wipe to wipe dry tools.
    3. Please change the tools between sampling sites.

    Remove samples from the freezer and place them in an ice bucket to avoid over-warming.
    Consult the sample list to identify the first fin-clip to be placed in well A1.
    Use the sterilized forceps to remove the first fin-clip sample and place it carefully in a clean parafilm square on the balance.
    Record sample name and weight of original sample.
    If the weight of the sample is over 15 mg, cut the sample using sterilized scissors so that it is a smaller size.
    Once the sub-sample is about 15 mg, place it securely in well A1. Ensure that it is completely submerged.
    If any portion of the fin-clip remains (i.e.,the sample weighed more than 15 mg and had to be cut), place the remaining fin-clip back in its original 1.5ml sample tube on ice.
    Decontaminate all handling tools and discard used weight boats/parafilm and paper towels/kim wipes before moving onto the next sample using steps 5 - 8.

Note: Weighing samples may only need to be done for the first column (A1 – H1) to identify the approximate size of 15mg of tissue. After weighing these samples, we can estimate the sizes for the remaining samples based on photographs and memory; this method is recommended by DArT to help reduce contamination of samples and save time during packaging.

Once the first full column of racked tubes is filled with the tissue and ethanol, securely close each well by snapping a strip of barrette caps onto the column (**Figure 2**).
Move on to the next column A2 – H2 using sterilized forceps to remove the fin-clip sample and a scalpel blade to cut the approximate size of 15mg.

1. Even if samples will no longer need to be weighed, it is still recommended to sterilize forceps and scalpels in between samples.
1. Place the tissue piece in well A2 and make sure that it is submerged in ethanol as described previously.

Repeat this process above for the next wells of the column and plate, being sure to maintain clean handling procedures between samples.
The use of two pairs of forceps and scalpels is recommended, as one pair can be used alternatively for the next sample, while the other pair is cooling from the flame.
Avoid working with more than one sample at a time to prevent cross-contamination and sample mix-up.
Change your gloves if they have been soiled.
Be sure to correctly match the sample name and its associated well on the sample list. This information will be uploaded as a .csv file to the DArT website for ordering.
A full plate should include fin-clips and ethanol in 94 wells, making sure that **G12 and H12 wells remain empty**.

Once the first plate is completed, place a paper towel over the top of the capped tubes, then close the tube rack with its associated box (**Figure 3**).
Secure the box lid with tape (as seen in the video) or rubber bands.
Place the box in a ziplock bag labeled “PLATE 1” and place the bag into the -20 ºC freezer until ready for shipment.
Perform the previous steps for the second plate.

Figure 3: Photos provided by DArT to illustrate approximately racked tubes of tissue being readied for shipping by placing them in a tube rack, covering them with paper towel and securing the rack lid.

Shipping instructions

Information on shipping instructions is found here. It is extremely important to precisely follow the instructions for filling and packing the documents and the documents needed are linked below.

1. Print the full contact details of the sender and provide them inside the package
2. Make sure all sender and DArT contact details are correct and legibly written on outside the package, and include DArT phone number (**Figure 4**).
3. Include sender's Service Specification and Sample Tracking File(s) in the package.
    a. The Service Specification form will be sent by email once the order is placed.

   Package should be sent in a rigid box/container with ample packing material to allow for rough handling during shipment.

Figure 4: DArT PL shipping addresses. ***Use both addresses when shipping***. DArT PL phone numbers are Tel: + 61 2 6122 7300, Fax: + 61 2 6122 7333 and ABN No.: 47 097 662 514

Shipping documents (international shipments)

All documents can be found in the "Documents_for_shipment" folder.

Can print three copies of the following five (5) documents. One copy will be for DHL, one copy is for Australian Customs and another copy is for our safekeeping.

Place a set of the documents colored in **RED** in an envelope attached to the outside of the package, clearly labeled **"Australian Customs".** A set of these documents will also be given to DHL, which also may request them to be scanned/uploaded for faster facilitation of the samples at customs.

The documents in **BLUE** are placed inside of the box.

1. Manufacturers Declaration_Animal Tissues & Fluid.docx
    1. Be printed on organization letterhead paper (COOOL)
    1. Be in English
    1. Prominently quote the Air WayBill (AWB) number on all pages of declaration

1. Be issued in and dated in the last six months
1. Be signed by the sender
1. Describe samples accurately.
1. The declaration MUST also clearly state that the samples are sent for destructive analysis in the PC1 laboratory of Diversity Arrays Technology Pty. Ltd.

Pro Forma Invoice Template

1. On the organization's letterhead (COOOL) stating the value at 20 euros (the cost of consumables).

Permit 8468921 Animal Fluids and Tissues.pdf

1. Condition 1 is relevant and marked on the Manufacturers Declaration_Animal Tissues & Fluid.docx

Air WayBill (AWB) (from the shipping company)

1. The samples may need to be accurately described for the WayBill: **"Preserved fin-clip in absolute ethanol from species** *Dissostichus eleginoides* **for in vitro use only"**.
1. Do not use the words ANIMAL or ANIMAL SAMPLES without also mentioning PRESERVED IN Ethanol.

Service Specification form

1. This will be provided after submitting an order online
1. Placed in the box with the samples

When the package is shipped:

Complete the sample shipment notification form:
https://www.diversityarrays.com/contact-us/sample-shipment-notification/

and email samples@DiversityArrays.com the

- Name of the courier company;
- The tracking number from the courier; and
- The Service Number

Appendix 2: Sample selection

Appendix table 1. Count of selected samples, based on individual sex and maturity stages. The total number of samples is 248; however, two from PE had no metadata associated and were sent for sequencing, while one originally selected from SIR but not found in the sample inventory was located and was also sent. This brought the total to 251 samples .

Region	Sex	Maturity stage	Count
South Indian Ridge (SIR)	F	1	3
	F	2	19
	F	3	7
	F	4	3
	F	5	1
	M	1	7
	M	2	13
	M	3	11
	M	4	1
Del Cano Rise (DCR)	F	1	6
	F	2	8
	F	3	6
	F	4	5
	F	5	1
	M	3	24
	M	4	11
	M	5	4

Williams Ridge (WR)	F	1	27
	F	2	1
	M	1	2
	M	2	4
Crozet (CR)	F	1	7
	M	1	1
	M	2	7
	M	3	3
	M	4	6
Kerguelen (KER)	F	1	15
	F	2	2
	F	4	1
	M	1	1
	M	2	5
	M	4	3
Prince Edward and Marion Islands (PEMI)	F	1	2
	F	2.5	1
	F	3	9
	M	1	4
	M	2.5	1
	M	3	13
	M	4	3
Total			**248***

Appendix list 1. List of final reports names provided by DArT

Report_DDiss24-9609_SNP_2.csv
Report_DDiss24-9609_SNP_mapping_2.csv
Report_DDiss24-9609_SilicoDArT_1.csv
metadata.json
metadata.xlsx

Appendix Table 2. List of DArT metadata, including BLAST alignment column information.

Metadata	
Column	**Description**
SNP	Contains the base position and base variant details
SNP Position	The position (zero indexed) in the sequence tag at which the defined SNP variant base occurs
Trimmed Sequence	Sequence containing SNP(s), but with removed adapters in short marker tags
Call Rate	The proportion of samples for which the genotype call is present (0,1,2) and not missing ("-")
OneRatioRef	The proportion of samples for which the genotype score is "0"
OneRatioSNP	The proportion of samples for which the genotype score is "1"
FreqHomRef	The proportion of samples which score as homozygous for the Reference allele
FreqHomSNP	The proportion of samples which score as homozygous for the SNP allele
FreqHets	The proportion of samples which score as heterozygous
PICRef	The polymorphism information content (PIC) for the Reference allele row
PICSnp	The polymorphism information content (PIC) for the SNP allele row
AvgPIC	The average of the polymorphism information content (PIC) of the Reference and SNP allele rows
AvgCountRef	The sum of the tag read counts for all samples, divided by the number of samples with non-zero tag read counts, for the Reference allele row
AvgCountSnp	The sum of the tag read counts for all samples, divided by the number of samples with non-zero tag read counts, for the SNP allele row
RepAvg	The proportion of technical replicate assay pairs for which the marker score is consistent
BLAST results	
Column	**Description**
AlnCnt_	Total count of aligning markers / tags with selection criteria described below
AlnEvalue_	E value of the best alignment to an existing model genome
ChromPosSnp_	Calculated position(s) of the SNP for best aligned marker on a contig(s) to an existing model genome

ChromPosTag_	Position(s) on contig(s) with the best alignment of marker / tag to an existing model genome
Chrom_	Contig(s) with the best alignment of marker / tag to an existing model genome
Strand_	Strand of the marker alignment - Plus for forward and Minus for reverse

Appendix 3 Terms of Reference

Terms of Reference (ToR) for the provision of scientific services to SIOFA Scientific Committee

Project title: Genetic analysis to inform the stock structure of Patagonian toothfish (*Dissostichus eleginoides*)

Project Code: SER2022-TOP1

Introduction

SIOFA CMM2018/01 (paragraph 6a) requires the SIOFA Scientific Committee to provide advice to the Meeting of Parties on the status of stocks of deep-sea fishery resources, including Patagonian toothfish (*Dissostichus eleginoides*). In 2020, the SIOFA Scientific Committee (SC3) conducted the first preliminary analysis of the Patagonian toothfish fishing data from the Del Cano Rise in the SIOFA Area. Those approaches were in early stages and to estimate stock structure in the SIOFA Area, more robust approaches and data would be needed.

This document describes the project Terms of Reference (ToR), milestones, and administrative matters for a consultancy to undertake Patagonian toothfish stock assessments. Once appointed, the Consultant should direct any questions and clarifications to the SIOFA Science Officer (Marco Milardi, marco.milardi@siofa.org) who will coordinate the project and its interactions with the project advisory panel, the relevant SC HoDs and the SIOFA Scientific Committee Chair, as appropriate.

This project aims to design a genetic stock discrimination project. Note that the collection of samples, analysis, and a full review of the stock structure of Patagonian toothfish will be conducted under SIOFA Project SER2022-TOP2.

Terms of Reference

The project objectives and tasks are described below. The Consultant shall undertake these tasks and consult with the project coordinator, to ensure that the project objectives are met.

A project advisory panel consisting of the SIOFA Scientific Committee Chair, selected members of the SIOFA Scientific Committee, and the SIOFA Secretariat will meet periodically with the consultant to assist the consultant access and interpret reports, data, and to provide advice on relevant analyses or data interpretation for the project.

Overall objectives

Objective 1: Provide advice to the SIOFA Scientific Committee on the design of a genetic stock discrimination project to understand the stock structure of Patagonian toothfish in the SIOFA Area, including linkages to Patagonian toothfish in the CCAMLR Convention Area.

Task 1: Literature review

Review the previous stock assessments, SIOFA reports and publications, CCAMLR scientific papers and reports, the general scientific literature, and other relevant information sources, including Patagonian toothfish stocks in other areas, to design and a genetic analysis of Patagonian toothfish stock structure in the SIOFA Area. The

outcomes of this project will be used to support SIOFA project SER2022-TOP2: Stock structure of Patagonian toothfish.

Task 2: Review of catch-effort and other relevant data
Review the relevant catch-effort and scientific observer data (e.g., age, length, and other biological data) held by SIOFA, and available bathymetric, oceanographic, and other relevant environmental drivers to design a genetic analysis sampling project of Patagonian toothfish in the SIOFA Area. This will also include consideration of potential linkages with Patagonian toothfish stocks in the Indian Ocean sector of the CCAMLR Convention Area[3].

Task 3: Genetic stock discrimination
Evaluate the feasibility of genetic stock discrimination for Patagonian toothfish, including the development and design of a genetic stock discrimination project to improve the understanding of stock structure in the SIOFA Area, by:

(i) evaluating the feasibility of a genetic stock discrimination project, and
(ii) develop and design a genetic sampling project including specifications of the number of samples, locations and timing for the collection of samples using commercial fishing operations, the contents of a genetic sampling kit for observers and/or vessels, timelines, and costs for the project.
(iii) describe the contents of genetic sampling kits and collection protocols for distribution to SIOFA vessels and observers to enable them to collect samples

Reporting requirements

1. Provide updates and engage with the project advisory panel that will assist the consultant access and interpret reports, data, and to provide advice on relevant analyses or data interpretation for the project

2. Provide a draft report detailing the methods, outcomes of reviews, conclusions, and recommendations to the SIOFA project advisory panel for review by 31 January 2022.

3. Update the draft report in (2) by considering any comments and advice from the project advisory panel and submit this report to SIOFA Secretariat for submission to the SIOFA Scientific Committee meeting in 2023 by 15 February 2023

4. Present the draft report in (3) to the SIOFA Scientific Committee to its meeting in March 2023 by videoconference.

5. Provide an amended final report to the SIOFA Secretariat, considering any comments made at the SIOFA Scientific Committee meeting in March 2023, by 15 April 2023

6. Provide all the information collected to the SIOFA Secretariat (including that sourced from the Secretariat) before the final payment of the contract is made to the consultant. Such information includes electronic data files, analysis codes, biological samples, and other relevant data if applicable.

Confidentiality and distribution of project outcomes

[3] CCAMLR Convention Area includes the South African, French and Australian management areas

The Consultant shall not release confidential data provided for conducting this study to any persons nor any organisations, other than SIOFA Secretariat. The consultant shall delete all the confidential data after the completion of the contract. Any arrangements for ownership, storage, or disposal of physical samples shall be agreed by SIOFA as a part of the contract.

All Intellectual Property generated as a part of this contract shall become the property of SIOFA unless otherwise excluded in the proposal and agreed by SIOFA in the contract.

All reports and presentations will be reviewed by the SIOFA Secretariat prior to any form of further distribution. The Consultant will revise the report according to comments received from the review process before the report or presentation is accepted as a submission against the requirements in the Terms of Reference.

Relevant SIOFA information

1. SIOFA data (provided by the SIOFA Secretariat upon request)
2. SIOFA reports:
    a. SIOFA SC reports and National Reports. Scientific Committee Meeting | SIOFA (siofa.org)
    b. MoP reports. Meeting of the Parties | SIOFA (siofa.org)
    c. SIOFA technical and scientific reports (public reports available from siofa.org, and restricted reports available from the SIOFA Secretariat to the project consultant)

Relevant CCAMLR information

1. CCAMLR papers and reports that consider linkages with Patagonian toothfish stocks in the Indian Ocean sector of the CCAMLR Convention Area
2. Previous studies on the genetic stock structure of Patagonian toothfish in the CCAMLR and adjacent areas
3. Patagonian toothfish management options currently in use for these stocks in the CCAMLR Convention Area

Work plan and payment schedule

The funds for this project are budgeted under General Objective 1 of the SIOFA/EU Grant Agreement SI2837681 - Scientific Work Support, for a total allocated budget of 8,333 euro (including all costs and including any travel related expenses). Any report and/or presentation, in paper or electronic form, must indicate that this task has received EU funding and display the EU emblem.

The consultant shall follow the timeline described in Table 1 below.

Table 1: Timeline for payments, milestones, and report submission

Milestone	Date	Activities
Initiation of contract	6 January 2022	First instalment payment (30% of the total contract sum)
Delivery of draft report	15 February 2023	Submission of draft report to SC8
Delivery of final report	15 April 2023	Submission of final report and project information to SIOFA.

		Final instalment payment (70% of the total contract sum) on acceptance of the final report and the submission of project information

Submission of applications

The applicants should have appropriate experience and knowledge of developing stock structure hypotheses and preferably on the stock dynamics and life cycle of Patagonian toothfish. The applicants should submit a proposal to the project coordinator (SIOFA Science Officer - Marco Milardi, marco.milardi@siofa.org) containing the following items:

1. A current CV that summarises the applicant(s) relevant educational background and professional experience
2. A brief proposal (indicatively 1-2 pages) outlining the proposed methods and analyses, including a description of how the objectives of the ToRs will be achieved
3. Any proposed exclusions to the intellectual property clause
4. The proposed consultancy price (including all consultant expenses and project related costs), noting that the available budget for this work is a maximum of €8,333
5. Identification of any project risks and associated mitigation and management required to successfully complete the project
6. A statement that identifies any perceived, potential, or actual conflicts of interest of the applicant(s), including those described in paragraph 4 of the SIOFA recruitment procedure (see Box 1), and
7. Any additional relevant information the applicant(s) wish to submit.
8. We note that similar projects for alfonsino and orange roughy in the SIOFA Area are also available, and we encourage consultants to submit combined proposals for these projects if appropriate.

Applications received before 12 AM (9 AM UTC) on Monday the 2nd of January 2023, Reunion Island time, will be considered in the following selection process.

EVALUATION CRITERIA FOR THE SELECTION OF CANDIDATES

The selection criteria will be developed by the evaluation panel along with the project manager, the Secretariat, and the Chairpersons of the relevant subsidiary bodies. The criteria may include following items:

1. Adequate submission of information to allow the panel to evaluate the candidate
2. Evaluation of the proposal from the candidate, including the proposed contract price
3. Ability to undertake and complete the analyses or work required in the ToR
4. The candidate's agreement with confidentiality provisions required for the project
5. Acceptable conflict of interest statement
6. Agreement with the data submission and intellectual property terms required in this ToR, and
7. Financial and resourcing considerations.

Conflicts of interest. Paragraph 4 of SIOFA's Recruitment Procedure
To ensure that situations relating to potential and actual conflict of interests are avoided, persons falling into the following categories may not normally be considered for SIOFA consultancy: (i). any person designated as a designated representative or alternate representative of a CCP to the Meeting of Parties (MOP) as per Rule

3.1 of the Rules of Procedure, and to the SC and any other subsidiary bodies of the MOP, as per Rule 21.3 of the Rules of Procedure; (ii). Any person fulfilling the function of Chair or Vice-Chair of the MOP or Chair or Vice-Chair of a SIOFA subsidiary body or working group; (iii). Any person acting as a member of a delegation involved in the SIOFA decision-making process resulting in recommendations and/or approval for the SIOFA work requiring the engagement of a consultant; and (iv). Individuals who were SIOFA Secretariat staff members at the time when the recommendations and/or approval for the SIOFA works were adopted or who are members of immediate family (e.g., spouse or partner, father, mother, son, daughter, brother, or sister) of any Secretariat staff member or of the persons identified in 4 (i), (ii), and (iii).

CONTACTS

Project Coordinator – SIOFA Science Officer (Marco Milardi, marco.milardi@siofa.org)

Administration – SIOFA Executive Secretary (Thierry Clot, thierry.clot@siofa.org)