KU Leuven
Faculty of Science
Department of Biology
Laboratory of Biodiversity and Evolutionary Genomics


# Connectivity and Evolution of Fishes in the Southern Ocean

## From Species to Populations

Henrik Christiansen

April 2020

Promotor: Prof. Dr. Filip A. M. Volckaert

Co-Promotor: Dr. Anton P. Van de Putte

Dissertation presented in partial fulfilment of the requirements for the degree of Doctor of Science (PhD): Biology

Uitgegeven in eigen beheer, Henrik Christiansen, Mechelsesteenweg 23/20, B-2018 Antwerpen, Belgium.

| | |
|---|---|
| Supervisory committee: | Prof. Dr. Hans Jacquemyn, KU Leuven<br>Dr. Jennifer Jackson, British Antarctic Survey |
| Additional examination committee members: | Prof. Dr. Luc De Meester, KU Leuven and IGB Berlin<br>Prof. Dr. Tom Wenseleers, KU Leuven<br>Dr. Emma Young, British Antarctic Survey |
| Chair of examination committee: | Prof. em. Dr. Frans Ollevier, KU Leuven |

For Jari.

May you still see some of the wonders of Earth’s biodiversity.

# CONTENT

## SUMMARY

Humans have become the single biggest driver of change on Earth and are largely responsible for a severe decline of wildlife. This biodiversity crisis affects all levels of biological diversity. While species diversity is a widely used biodiversity measure, intraspecific diversity, such as genetic population differentiation and local adaptation, is often neglected in conservation. Fish populations are facing a multitude of stressors in the Anthropocene, including fisheries, introduction of non-indigenous species, pollution and climate change. The fishes of the Southern Ocean in particular represent a unique ichthyofaunal assemblage with high levels of endemism and spectacular adaptations that enable them to thrive in subzero waters. Rapid environmental change and increasing anthropogenic impact are threatening the marine Antarctic ecosystems. In response, the Committee for the Conservation of Antarctic Marine Living Resources (CCAMLR) is committed to establish a network of marine protected areas (MPAs). The implementation of MPAs in the Southern Ocean, however, is challenging as CCAMLR is debating about the interpretation of "rational use" (i.e. managed fisheries) and the concept of ecosystem-based management. Against this background, I have studied the fish diversity of the Southern Ocean with special emphasis on the connectivity and evolution of Antarctic fishes. The remote nature of the Antarctic ecosystems provides an opportunity to understand effects of global change in a relatively pristine setting. At the same time, current threats to Antarctic biodiversity call for enhanced protection measures, that depend on a solid understanding of connectivity and evolution. Four main research questions are addressed here: How common is phylogeographic structure in mesopelagic Southern Ocean fish? Can high-throughput DNA sequencing methods be applied to Antarctic taxa? What are the spatial and temporal scales of genetic differentiation and local adaptation in benthic Antarctic fish? Can larval dispersal via oceanic currents facilitate long-distance connectivity?

First, the diversity of mesopelagic fishes in the Southern Ocean was investigated using DNA barcoding. Lanternfish (Myctophidae) dominate the twilight zone, but easily lose taxonomically informative characters upon capture. We used the DNA Barcode of Life Data (BOLD) Systems to verify morphological identification. This database was substantially extended, which will benefit future (meta)barcoding studies. Phylogeographic structure in myctophids is rare, possibly due to their enormous abundance and the homogenizing effect of the ACC on pelagic fishes. The overall myctophid phylogeny suggests that multiple colonization events shaped the evolutionary history of Antarctic myctophid species.

We then optimized a reduced representation sequencing (RRS) methodology to delve deeper into the genomic makeup of Antarctic notothenioids, the dominant clade in coastal Antarctic waters. Genome-wide screening of thousands of single nucleotide polymorphisms (SNPs) in parallel is possible with RRS at reasonable cost but requires careful optimization. Empirical and computational *a priori* tests were used to inform optimal experimental setups for RRS, which was tested in a range of Antarctic taxa, including fish. This pilot study provides the foundation for large-scale population genomics.

Subsequently, large and fine-scale genomic differentiation was investigated in the common coastal Antarctic fish *Notothenia coriiceps*. The data reveal previously undocumented genetic divergence between the western Antarctic Peninsula and East Antarctica. Although comparatively subtle (max. $F_{ST}$ 0.02), this genetic differentiation may have important consequences in the future as these areas of Antarctica experience contrasting effects of climate change. Long-distance connectivity is achieved via larval dispersal, but genotype-environment associations also indicate some degree of local adaptation. These results should be considered in future conservation planning to support population resilience and adaptation potential.

A related species, *Notothenia rossii*, experienced severe overexploitation in the past. Here, we combined population genomics with modelling techniques that correlate environmental data with species occurrence and simulate larval dispersal by ocean currents. This multidisciplinary approach demonstrated differing extents of evolutionary *versus* ecological connectivity. The latter relies on relatively isolated stepping stones, that deserve appropriate protection. Evolutionary connectivity in contrast is high, as evidenced by spatial genetic homogeneity. In addition, we propose a connectivity scenario that may help explain the slow recovery of this species following overfishing.

Overall, genetic differentiation in Antarctic notothenioids is a result of the combination of life history parameters, demographic history and the contemporary environmental setting. The findings presented here have important implications for the conservation and management of living resources in the Southern Ocean. This is particularly relevant given the rapid changes the marine Antarctic ecosystems are facing and the ongoing effort to establish a representative network of marine protected areas.

## SAMENVATTING

De mensheid is de voornaamste oorzaak van veranderingen op aarde en is grotendeels verantwoordelijk voor de sterke afname in biodiversiteit. Deze biodiversiteitscrisis heeft invloed op alle niveaus van biologische diversiteit. Hoewel soortendiversiteit een algemeen gebruikte maatstaf voor biodiversiteit is, wordt intraspecifieke diversiteit, zoals genetische populatiedifferentiatie en lokale adaptatie, vaak verwaarloosd bij het behoud ervan. Vispopulaties hebben af te rekenen met talrijke stressfactoren in het Antropoceen, waaronder visserij, introductie van niet-inheemse soorten, vervuiling en klimaatsverandering. Vooral de vissen van de Zuidelijke Oceaan vertegenwoordigen een unieke ichthyofaunale assemblage met een hoge graad van endemisme en spectaculaire adaptaties die hen in staat stellen om te gedijen in wateren onder nul graden Celsius. Snelle veranderingen in het milieu en toenemende antropogene effecten bedreigen Antarctische mariene ecosystemen. Als reactie hierop legt het Comité voor de Conservatie van Antarctische Mariene Levende Rijkdommen (CCAMLR) zich toe op het opzetten van een netwerk van beschermde mariene gebieden (*marine protected areas*, MPAs). De implementatie van MPAs in de Zuidelijke Oceaan is echter een uitdaging aangezien CCAMLR debatteert over de interpretatie van 'rationeel gebruik' (d.w.z. beheerde visserijen) en het concept van ecosysteembeheer.

Met dit als achtergrond heb ik de visdiversiteit van de Zuidelijke Oceaan bestudeerd met speciale nadruk op de connectiviteit en evolutie van Antarctische vissen. De afgelegen aard van de Antarctische ecosystemen biedt de mogelijkheid om de effecten van modiale verandering in een relatief ongerepte omgeving te begrijpen. Tegelijkertijd vragen huidige bedreigingen voor de Antarctische biodiversiteit om betere beschermingsmaatregelen die afhangen van goede kennis van connectiviteit en evolutie. Vier belangrijke onderzoeksvragen komen hier aan bod: Hoe algemeen komt fylogeografische structuur voor bij mesopelagische vissen in de Zuidelijke Oceaan? Kunnen *high-throughput DNA sequencing* methoden worden toegepast op Antarctische taxa? Wat zijn de ruimtelijke en temporele schaal van genetische differentie en lokale adaptatie in Antarctische benthische vissen? Kan verspreiding van larven via oceaanstromingen connectiviteit op lange afstand faciliteren?

Eerst werd de diversiteit van mesopelagische vissen in de Zuidelijke Oceaan onderzocht aan de hand van DNA barcoding. Lantaarnvissen (Myctophidae) domineren deze schemerzone, maar

verliezen gemakkelijk hun taxonomisch informatieve kenmerken bij vangst. We gebruikten de DNA Barcode of Life Data (BOLD) systemen om de morfologische identificatie te verifiëren. Deze database werd substantieel uitgebreid, wat toekomstige (meta)barcoding studies zal bevorderen. Fylogeografische structuur in lantaarnvissen is zeldzaam, mogelijks door hun zeer algemeen voorkomen en het homogeniserende effect van de ACC op pelagische vissen. De volledige fylogenie van lantaarnvissen suggereert dat meerdere kolonisatiemomenten de evolutionaire geschiedenis van Antarctische myctophidsoorten hebben gevormd.

Vervolgens hebben we de gereduceerde representatie sequenering (RRS) methodologie geoptimaliseerd om dieper in te gaan op de genomische samenstelling van Antarctische notothenioiden, de dominante clade in Antarctische kustwateren. Genoomwijde screening van duizenden enkel-nucleotide polymorfismen oftewel *single nucleotide polymorphisms* (SNPs) in parallel is mogelijk met RRS aan een redelijke kostprijs, maar vereist zorgvuldige optimalisatie. Empirische en computationele a priori testen werden gebruikt om de experimentele opstellingen voor RRS te optimaliseren, die vervolgens werden getest op een reeks Antarctische taxa, waaronder vissen. Deze pilootstudie legt de grondslag voor grootschalige populatiegenomica.

Vervolgens onderzochten we genomische differentiatie in de algemene Antarctische kustvis *Notothenia coriiceps* op grote en kleine schaal. De gegevens onthullen eerder ongedocumenteerde genetische divergentie tussen het West-Antarctische schiereiland en Oost-Antarctica. Deze genetische differentiatie, hoewel relatief subtiel (max. $F_{ST}$ 0,02), heeft mogelijks belangrijke consequenties in de toekomst aangezien deze Antarctische gebieden contrasterende effecten van klimaatsverandering ervaren. Connectiviteit op lange afstand wordt bereikt via verspreiding van larven, maar genotype-omgevingassociaties wijzen ook op een zekere mate van lokale adaptatie. Deze resultaten moeten in acht genomen worden bij toekomstige conservatieplannen om de populatieveerkracht en adaptatiepotentieel te waarborgen.

Een verwante soort, *Notothenia rossii*, kende in het verleden ernstige overexploitatie. Hier combineerden we populatiegenomica met modelleringstechnieken die omgevingsdata correleren met de aanwezigheid van de soort en die de verspreiding van larven door

oceaanstromingen simuleren. Deze multidisciplinaire aanpak toonde de verschillende mate aan van evolutionaire versus ecologische connectiviteit aan. Deze laatste hangt af van relatief geïsoleerde stepping stones die de nodige bescherming verdienen. Ecologische connectiviteit daarentegen is hoog, zoals blijkt uit ruimtelijke genetische homogeniteit. Daarnaast stellen we een connectiviteitsscenario voor dat kan helpen met het verklaren van het trage herstel van deze soort na overbevissing.

In het algemeen is de genetische differentiatie in Antarctische notothenioiden het gecombineerde resultaat van levensgeschiedenisparameters, demografische geschiedenis en de hedendaagse omgeving. De hier voorgestelde resultaten hebben belangrijke gevolgen voor het behoud en beheer van levende organismen in de Zuidelijke Oceaan. Dit is vooral relevant gezien de snelle veranderingen waarmee de Antarctische mariene ecosystemen worden geconfronteerd en de voortdurende inspanningen om een representatief netwerk van beschermende mariene gebieden tot stand te brengen.

## LIST OF ABBREVIATIONS

| | |
|---|---|
| A | Allozymes |
| AABBAA | Archives Antarctiques Belges – Belgische Antarctische Archieven |
| AAD | Australian Antarctic Division |
| AB | Agulhas Bank |
| ACC | Antarctic Circumpolar Current |
| AFGP | Antifreeze glycoproteins |
| AFLP | Amplified fragment length polymorphism |
| AIC | Akaike's information criterion |
| AMOVA | Analysis of molecular variance |
| ANOVA | Analysis of variance |
| AP | Western Antarctic Peninsula |
| APECS | Association of Polar Early Career Scientists |
| ASF | Antarctic Slope Front |
| AUC | Area under the receiver operating curve |
| AV | All-purpose vessel |
| AWI | Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research |
| B | BayeScan method |
| BAS | British Antarctic Survey |
| BELSPO | Belgian Science Policy Office |
| BI | Bouvet Island |
| BIN | Barcode index number |
| BLAST | Basic Local Alignment Search Tool |
| BNCAR | Belgian Committee on Antarctic Research |
| BOLD | Barcode of Life Data systems |
| bp | Base pair |
| BRT | Boosted regression trees |
| BS | Bootstrap support |
| CCAMLR | Commission for the Conservation of Antarctic Marine Living Resources |
| chl_a_ampli | Amplitude of annual chlorophyll *a* concentration |
| chl_a_mean | Mean annual chlorophyll *a* concentration |

| | |
|---|---|
| CI | Crozet Island |
| CI (Table 5.5) | Confidence interval |
| CNV | Copy number variants |
| CM | Conservation measure |
| *COI* | Cytochrome oxidase I subunit gene |
| COR | Point biserial correlation |
| COS | Cosmonauts Sea |
| CPS | Cooperation Sea |
| CTAB | Cetyl trimethylammonium bromide |
| D | Differentiation measure following Jost (2008) |
| DAPC | Discriminant analysis of principal components |
| DB | Dallmann Bay |
| ddRAD | Double digest restriction site associated DNA sequencing |
| DDU | Dumon d'Urville Sea |
| DI | Deception Island |
| dist | Distance to continental Antarctica |
| DNA | Deoxyribonucleic acid |
| dns | Distance, non-significant |
| ds | Distance, significant |
| ECR | Early career researcher |
| EE | Extreme events |
| EE_max_chl_a | Maximum Chlorophyll *a* concentration during extreme events |
| EE_max_chl_a_ampli | Amplitude of maximum Chlorophyll *a* concentration during extreme events |
| EE_min_chl_a | Minimum Chlorophyll *a* concentration during extreme events |
| EE_min_chl_a_ampli | Amplitude of minimum Chlorophyll *a* concentration during extreme events |
| EE_max_temp | Maximum temperature during extreme events |
| EE_min_temp | Minimum temperature during extreme events |
| EEZ | Exclusive economic zone |
| EI | Elephant Island |
| EM | Edward and Marion Islands |
| FAO | Food and Agriculture Organization of the United Nations |

| | |
|---|---|
| FI | Falkland Islands |
| $F_{IS}$ | Inbreeding coefficient of an individual relative to the subpopulation |
| $F_{IT}$ | Inbreeding coefficient of an individual relative to the total population |
| $F_{ST}$ or $G_{ST}$ | Fixation index of subpopulations compared to the total population (following Weir & Cockerham 1984 unless otherwise specified) |
| FWO | Research Foundation Flanders |
| Gbp or Gb | Giga base pair |
| GBIF | Global Biodiversity Information Facility |
| GBS | Genotyping-by-sequencing |
| GeOMe | Genomic Observatories MetaDatabase |
| GLM | Generalized linear model |
| GPZ | General Protection Zone |
| GS | Genome size |
| $H_{exp}$ or $H_E$ | Expected heterozygosity |
| $H_{obs}$ or $H_O$ | Observed heterozygosity |
| HWP | Hardy-Weinberg proportions |
| HM | Heard and MacDonald Islands |
| IBM | Individual-based modelling |
| ICR | Ice cover range |
| ID | Identity |
| IPEV | Institut Polaire Français Paul-Émile Victor |
| ITM | Mean ice thickness |
| ITR | Ice thickness range |
| IWT | Flemish Agency for Innovation by Science and Technology |
| JI | Joinville Island |
| K | Genetic clusters (as used in STRUCTURE software) |
| KI | Kerguelen Islands |
| KRZ | Krill Research Zone |
| LAS | Lazarev Sea |
| Lat | Latitude |
| LD | Linkage disequilibrium |
| LG | Linkage group |
| Lon | Longitude |

| | |
|---|---|
| M | Million |
| MAF | Minor allele frequency |
| MAS | Mawson Sea |
| maxSSS | Maximum sensitivity plus specificity threshold |
| Mb | Mega base pair |
| ML | Maximum likelihood |
| MLD | Mixed layer depth |
| MNHN | Muséum National d'Histoire Naturelle |
| MPA | Marine Protected Area |
| MQ | Macquarie Island |
| mRNA | Messenger RNA |
| mtDNA | Mitochondrial DNA |
| mya | Million years ago |
| N | Number (of samples/replicates) |
| N50 | Sequence length of the shortest contig at 50% of the total genome length |
| NC | National Committee (of APECS) |
| NCBI | National Center for Biotechnology Information |
| nDNA | Nuclear DNA |
| $N_e$ | Effective population size |
| NEMO | Nucleus for European Modelling of the Ocean |
| NGO | Non-governmental organization |
| NHM | Natural History Museum |
| NIWA | National Institute of Water and Atmospheric Research |
| NKGI | King George Island, North |
| NMDS | Non-metric multidimensional scaling (plot) |
| NSR | North Scotia Ridge |
| NTA | Adélie Land, North |
| OBIS | Ocean Biogeographic Information System |
| OL | Ob and Lena Banks |
| P | Pcadapt method |
| PC | Principal components |
| PCA | Principal component analysis |
| PCNM | Principal coordinates of neighborhood matrices |

| | |
|---|---|
| PCR | Polymerase chain reaction |
| PF | Antarctic Polar Front |
| PKGI | King George Island, Potter Cove |
| POC | Particulate organic carbon |
| Q&A | Question and answer |
| qPCR | Quantitative polymerase chain reaction |
| QTL | Quantitative trait locus |
| RADseq | Restriction site associated DNA sequencing |
| RAPD | Randomly amplified polymorphic DNA |
| RE | Restriction enzyme |
| RECTO | Refugia and Ecosystem Tolerance in the Southern Ocean |
| RDA | Redundancy Analysis |
| RFLP | Restriction fragment length polymorphism |
| *rh1* | Rhodopsin gene |
| RNAseq | High throughput sequencing of ribonucleic acid (RNA) |
| ROS | Ross Sea |
| RRS | Reduced representation sequencing |
| RV | Research vessel |
| sACCF | Southern Antarctic Circumpolar Current Front |
| SAF | Sub-Antarctic Front |
| SAO | South Atlantic Ocean |
| SB | Skiff Bank, Kerguelen Islands |
| SBF | Southern Boundary Front |
| SCAR | Scientific Committee on Antarctic Research |
| SCARBio17 | SCAR Biology Symposium 2017 |
| SCS | Scotia Sea |
| SDM | Species distribution modelling |
| SFCS | Sea floor current speed |
| SFS | Mean sea floor salinity |
| SFS_ampli | Amplitude of mean sea floor salinity |
| SFT | Mean sea floor temperature |
| SG | South Georgia |
| SNP | Single nucleotide polymorphism |

| | |
|---|---|
| SO | South Orkney Islands |
| SOISS | South Orkney Islands Southern Shelf MPA |
| SP | Stacks parameters |
| SR | Shag Rocks |
| SRA | Sequence read archive |
| SRZ | Special Research Zone |
| SSa | South Sandwich Islands |
| SSCS | Sea surface current speed |
| SSD | South Shetland Islands, Deception Island |
| SSh | South Shetland Islands |
| SSI | South Sandwich Islands (Chapter 2) |
| SSK | South Shetlands Islands, King George Island |
| SSR | Simple sequence repeat |
| SSW | Size selection window |
| STA | Adélie Land, South |
| STF | Subtropical Front |
| STR | Short tandem repeat (microsatellite) |
| TSS | True skill statistic |
| vERSO | Ecosystem Responses to global change – a multiscale approach in the Southern Ocean |
| VIF | Variation inflation factor |
| VLAIO | Flanders Innovation & Entrepreneurship |
| WGA | Whole genome amplification |
| WS | Weddell Sea |

## ACKNOWLEDGEMENTS

This work would not have been possible without the help of many people. I would like to thank everyone that – by being who and how they are – has helped this thesis come about:

**Filip**, thanks for admitting me to your group and believing in me. Always having the feeling that you support me in my endeavours and having the freedom to pursue side projects and shape my research path in my own way has been a tremendous help. **Anton**, thanks for supporting me from start to finish and especially for introducing me to your extensive polar network and for being my partner-in-crime for all logistic issues. **Hans** and **Jen**, thank you so much for joining my supervisory and examination committee and always being very responsive and providing thoughtful, timely and helpful feedback. **Luc**, **Tom**, **Emma**, and **Frans**, many thanks for immediately agreeing to be part of the jury, for going through all the corresponding efforts and for reading my work in detail. I truly appreciate the time you have invested, and the work has certainly improved thanks to you.

**Franz**, I was very happy when Filip (finally ;-)) announced that he would hire another Antarctic researcher (actually I wanted it to be a postdoc). I can say, it couldn't have turned out better for me. Not only are you a close friend, but we've also scientifically worked together extremely well. You've kept me focused and helped to set the priorities straight and I have hugely profited from your determination and lab skill. Thank you. **Bart**, you should in fact be considered a co-supervisor of this thesis. Your official job title is senior research technician, but you're much more than that. Thank you for all your scientific input (regarding methods, lab work, data interpretation etc.) as well as your personal input. **Conny** and **Eddy**, you are indispensable, and I am very thankful for all your help throughout the years!

**Sarah**, what should I say, I could have never wished for a better master student. I am grateful for being able to still work with you and deeply impressed by your dedication, commitment and positive attitude. **Jasmien**, thanks for sharing all your ddRAD experiences and expertise and always being supportive. **Sophie**, I am very grateful for all your unconditional support, your good faith and important reminders of life beyond work. **Leo**, merci for being such a nice and very bright colleague. **Els**, you've now taken over as most important guardian of social cohesion in the lab – thank you for that, but also for your always insightful perspective on my work. **Nele**, for me you're one of the most missed ex-LBEGers, thank you for your always thoughtful opinions and feedback. **Ilaria & Allan**, thanks for helping in the early days of RADseq at LBEG and for always lending an open ear. **Federico**, thanks for playing the "old grumpy dude" of the lab, while actually being really supportive. **Io**, **Tim**, **Andreas**, **Thomas** (Kerkhove), **Anurag**, **Pascal**,

**Nikol**, and everyone else from LBEG that I haven't mentioned yet: thank you for making this a very enjoyable work group! Thanks also to the other thesis students/interns that I supervised for your time and efforts: **Thomas** (Smit), **Ana**, **Stijn**, **Aurelie**, **Valerie**.

Then there's people outside of our little lab environment. I'd like to gratefully acknowledge:

**Jen** (Jackson), you get a special mention here again: I met you as a very (!) fresh PhD student and without much context asked if you'd join my committee. Rightaway you agreed and I have deeply valued you as a researcher since. Thank you for making the time to support me. **Agnés** (Dettaï), I approached you during the same meeting and you were extremely welcoming and approachable. That has never changed, even after I took years to finish what was supposed to be a quick side project. I am grateful for your kindness and support during and after my visit to Paris. **Emma** (Young), we've been in touch for a quite a while now and only in the past year I really started working on what should have happened earlier. But you remained patient and always offered immediate, highly useful and sympathetic feedback whenever I asked for it. I was fortunate to be able to work with you! **Isa** (Schön), even though I'm still not officially (as in "paid") part of RECTO, you have always treated me as a valuable collaborator and not least helped made our upcoming research project happen. Many thanks!

**Hauke** (Flores), thanks a lot for being a genius networker or "people-together-bringer", who has also welcomed and supported me from the first moment I suddenly entered your office, scared that I'd be disturbing you… **Bruno** (Frédérich), **Loïc** (Michel), **Gilles** (Lepoint) – you're cooperation at it's best! **Camille** (Moreau), **Quentin** (Jossart), merci beaucoup for scientific and personal support, beers, jokes and football (and baby name suggestions). **Charlène** (Guillaumot), you're are an ingenious modeler with a very thorough commitment to help others - thank you so much. **Bruno** (Danis), I am most grateful for you integrating me into the vERSO-RECTO and B120-121 world. The B121 expedition was an extraordinary – likely once-in-lifetime – experience, thank you for organizing that! Many thanks also to everyone else from this expedition: **Ben** (Wallis), **Katy** (Lucas), **Ryan** (Houston), **Camille** (Moreau), **Quentin** (Jossart), **Franz** (Heindler), **Charlène** (Guillaumot), **Francesca** (Pasotti), **Henri** (Robert), **Thomas** (Saucède). I would also like to thank **Lili** (Hess) for all her tremendous efforts to create a beautiful documentary.

**Felix** (Mark), during Polarstern expedition PS96 I basically approached you saying, "Hi I want to have fish, too!". That was rather upfront, but you were totally welcoming and supportive, so thanks a lot for that and continuos support. **Magnus** (Lucassen), **Chiara** (Papetti), same goes for you, I don't take it for granted that you so openly welcomed a new person on the scene of

Antarctic fish genetics. **Nils** (Koschnick), **Emi** (Riginella), most importantly you've shared one of the greatest experiences of my PhD with me and I am truly grateful (and have also learnt a lot form you!). **Cheps** (Sands), special thanks to you for being exactly how you are! **Christoph** (Held), without whom I wouldn't have gone on Polarstern in 2015, thank you so much for this opportunity and your trust in me. **Mike** (Schröder), **Santi** (Pineda Metz), **Dieter** (Piepenburg), **Kevin** (Kocot), **Daniel** (Kersken), **Rebeca** (Zapata), and all the others of PS96, thanks for the nice atmosphere and an amazing cruise.

**Huw** (Griffiths), **Steve** (Parker), **Malte** (Damerau), **Guy** (Duhamel), **Cheps** (Sands), **Dave** (Barnes), **Gary** (Carvalho), **Ben** (Wallis), **Thomas** (Desvignes), many many thanks for giving samples/fishes to us and supporting the research I do. In fact there are still many others I should thank for support in one way or another, for example **APECS** (especially **Gerlis** Fugmann, Yi **Ming** Gan, **Igor** Pessi, **Valentina** Savaglia, and everyone of APECS Belgium), **CeMEB** (especially **Carl** André, **Per** Jonsson & **Eva-Marie** Rodström), **FSBI** (that really were some of the best conferences I so far attended!) and the wider network of Filip's lab (**Joost** Raeymakers, **Greg** Maes, **Maarten** Vanhove, **Sara** Vandamme, …). Additional thanks to everyone I forgot and to friendly colleagues all over the world for advice and support: Charlotte Havermans, Chiara Papetti, Marie Verheye, Valérie Dulière, Anne Gro Salvanes, Nicolas Fournier, Martin Collins, Antonio Aguera, Marc Eleaume, Marc Kochzius, Fredrik Leliaert, Tasnim Patel, Dirk Steinke, Ann Vanreusel.

If you, dear reader, have made it to this point, congratulation, **the most important support** is usually acknowledged at the very end:

Thomas, Bärbel, Hannes, I am what I am because of you. And I hope you can take that as a compliment. I am forever grateful for you being my main socialization.

My grandparents, Lilo & Heinz, who can't read this anymore, but surely would be extremely proud: you helped me throughout life in many ways, I am thankful for that and will not forget you.

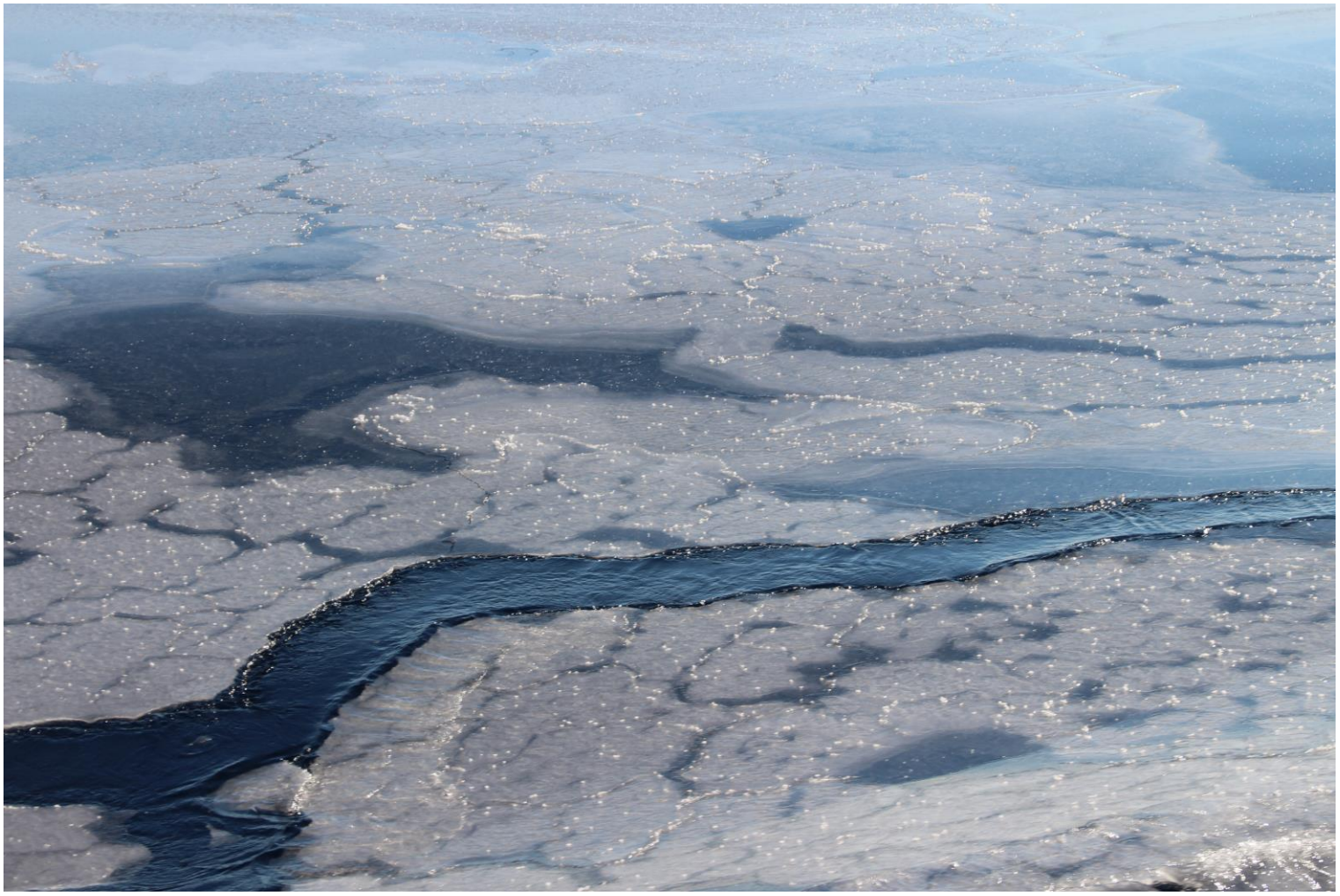
Sandra, we are the two-body problem. But we turn it from a perceived disadvantage to an advantage. You are the one that makes me complete. Personally, but in fact also scientifically.

Thank you.

Henrik, April 2020

**MAIN TEXT**

# CHAPTER 1: Evolution in a rapidly changing marine ecosystem



**Evolution in a rapidly changing marine ecosystem**

Henrik Christiansen

## 1. Fishes in the Anthropocene

The last geological epoch is now often referred to as the Anthropocene, a period where humans have become the single biggest driver of change on Earth. Human presence is inevitably shaping the biological world. On land, the biomass of humans and livestock is now an order of magnitude higher than that of wild animals (Bar-On *et al.* 2018). Anthropocene defaunation has driven especially large, apex predators to or near extinction (Dirzo *et al.* 2014). Defaunation is less advanced in the oceans, but an era of intense marine wildlife decline is dawning (McCauley *et al.* 2015). Biological diversity in general is in a state of rapid decline, with no signs of slowing of this alarming trend (Butchart *et al.* 2010). The biodiversity crisis affects all realms including major alterations to marine ecosystems, particularly through fisheries, pollution, warming, and ocean acidification (e.g. Hofmann *et al.* 2010; Cheung *et al.* 2013; Eriksen *et al.* 2014). Global fisheries have reached a plateau, where catch levels remain relatively steady, despite increasing fisheries efforts driven by growing demands (Watson *et al.* 2013). The fish biodiversity of planet Earth is facing a multitude of challenges in the Anthropocene that call for integrative, cross-disciplinary research responses (Obregón *et al.* 2018). The research presented here is a small step towards that goal, specifically focusing on the inter- and intraspecific genetic diversity of Antarctic fishes. Before delving deeper into the research methodology, study system and scientific questions, however, we need to define a number of key terms and concepts (see also Box 1.1).

Clearly, the global biodiversity of our planet is threatened. But what exactly is biodiversity in general and specifically in fishes? One may define biodiversity as the biological variability that occurs naturally in organisms across three major levels, namely "within species, between species and of ecosystems", as recognized by the Convention on Biological Diversity (United Nations 1992). It is important to emphasize that biodiversity in this sense does not merely reflect the number of species present in a system, but also compartments within these entities and the community they form together. The species concept itself is not unequivocal. It is rooted in the work of John Ray (1686) and others, and the subsequent biological classification of Linnaeus (1753), which formalized the ranking of sorts of organism in the binomial taxonomic

nomenclature still used today. By now, however, at least 34 different (but not necessarily mutually exclusive) concepts of how a species can or should be defined have appeared (Zachos 2018). The reason for this overwhelming number of definitions and a long history of debate lies in the inherent problem of species definition. In reality, organisms do not necessarily correspond to a categorical classification system of biological entities. Instead species and populations are likely forming a continuum (Carvalho *et al.* 2017; Coates *et al.* 2018; Galtier 2019) with all sorts of states between more or less reproductively, spatially, genetically or ecologically separated units that may have experienced vicariance and speciation, but also secondary contact, hybridization, or even speciation reversal (Kearns *et al.* 2018). Defining a unit such as "species" can therefore be complicated and will always depend on the specific organism and the definition employed. Probably the most popular species concept is still the biological species concept, as developed in the early twentieth century in works by Mayr (1942) and Dobzhansky (1950; see De Queiroz 2007). This concept emphasizes the role of reproductive isolation where a species is formed by a group of organisms that can reproduce with one another. Extensions of this concept add that the organisms should not only be able to reproduce, but also need to produce viable and fertile offspring. In an evolutionary sense, one may add that species are not only reproductively isolated groups (or populations or metapopulations) of organisms, but evolve over time and hence can be viewed as lineages (De Queiroz 2007). Furthermore, additional species concepts can be viewed as complementary and helpful for species delimitation in specific situations (e.g. asexual organisms). In that sense, many species concepts can essentially be united through a common element, but differ in the properties they additionally emphasize (De Queiroz 2007). Yet, some of the properties necessary to decide whether a given group of animals belongs to one or two species are difficult or nearly impossible to test in marine species as they are often challenging to manipulate experimentally. Even if species criteria can be tested, most of these will be continuous and therefore cannot objectively answer a categorical question. Some authors consequently argue that completely non-arbitrary delimitation of biological units is impossible (Zachos 2018). Cataloging and ascribing species or sub-species status to specific groups of organisms is, however, absolutely crucial when it comes to management and conservation measures. Only recognized and defined units can be effectively protected. A recent example includes the flatfish *Platichthys flesus* in the Baltic Sea, which has evolved through parapatric, ecological speciation into two species (*P. flesus* and *P. solemdali*) with currently low genome-wide divergence (Momigliano *et al.* 2017, 2018). This divergence was previously unrecognized and consequently the fishery was managed as one stock. Retrospective molecular analysis to disentangle the two units helped explain the decline in commercial landings, and future management measures should be refined (Momigliano *et al.* 2019). Readily available DNA

sequence data has revolutionized the detection of species in basically all biological domains. In fishes, we stand now at 35,122 valid species with a few hundred being added each year (Fricke *et al.* 2019). The fact that so many new fish species are still being described underlines the unabated importance of systematics and taxonomy, even if less and less research funding is available for such fundamental endeavors. Given the present and upcoming challenges for fish species to survive in a human dominated world, it is necessary to continue to identify, delineate and extensively describe the extant fish (species) diversity.

Even if one group of fish is well-established as one and only one species, this taxon may come in further variations. Many organisms display a metapopulation structure with several more or less distinct populations, often identified genetically. Metapopulations are a set of individual, independent and spatially separated populations, or in other words a "population of populations" (Levins 1969). The populations in this framework are in theory independent, but need not remain so. They can for instance exchange individuals through migration and/or temporal spatial overlap (e.g. spawning aggregations). Exchange of individuals can be characterized as the extent of connectivity (Fig. 1.1), an important measure to determine colonization patterns and the resilience to harvest. Alternatively, connectivity can be defined as a state, that is the level of connectedness between organism aggregations in space. If one considers this state as dynamic and depending on successful dispersal, one could also think of it as a process, as we will throughout this thesis. Dispersal is the movement of individuals (typically offspring/propagules) between populations or spatially separated habitats. In contrast, connectivity is the overall rate of successful dispersal over short (ecological connectivity, shorter than generation time) or long time scales (evolutionary connectivity; longer than generation time). Variable connectivity can lead to extinction-recolonization-cycles, or, less definitively, source-sink dynamics of individual populations and habitats, while the overall metapopulation remains relatively stable. Environmental heterogeneity or otherwise fragmented habitats are typically prerequisites for metapopulation structure. Spatial scales in general are the defining feature of metapopulations, especially from a fisheries science perspective (Kritzer & Sale 2004). In the marine realm, environmental heterogeneity and physical barriers were in fact for a long time considered absent or negligible compared to terrestrial systems. This notion implied that metapopulation structure (or in other words the presence of distinct sub-populations) would be largely absent in the ocean. A panmictic overall population, with random mating between all individuals of the population, is the classical null model for population genetic investigations of spatial genetic structure. However, as the resolution of genetic markers and other techniques to

assess spatial intraspecific variation improved, it became clear that marine species do regularly exhibit local variations (Hauser & Carvalho 2008). Nevertheless, many marine species may rather occur as "patchy populations" with high migration rather than as metapopulations (Fig. 1.1), because the risk of local extinction is virtually absent (Smedbol *et al.* 2002).

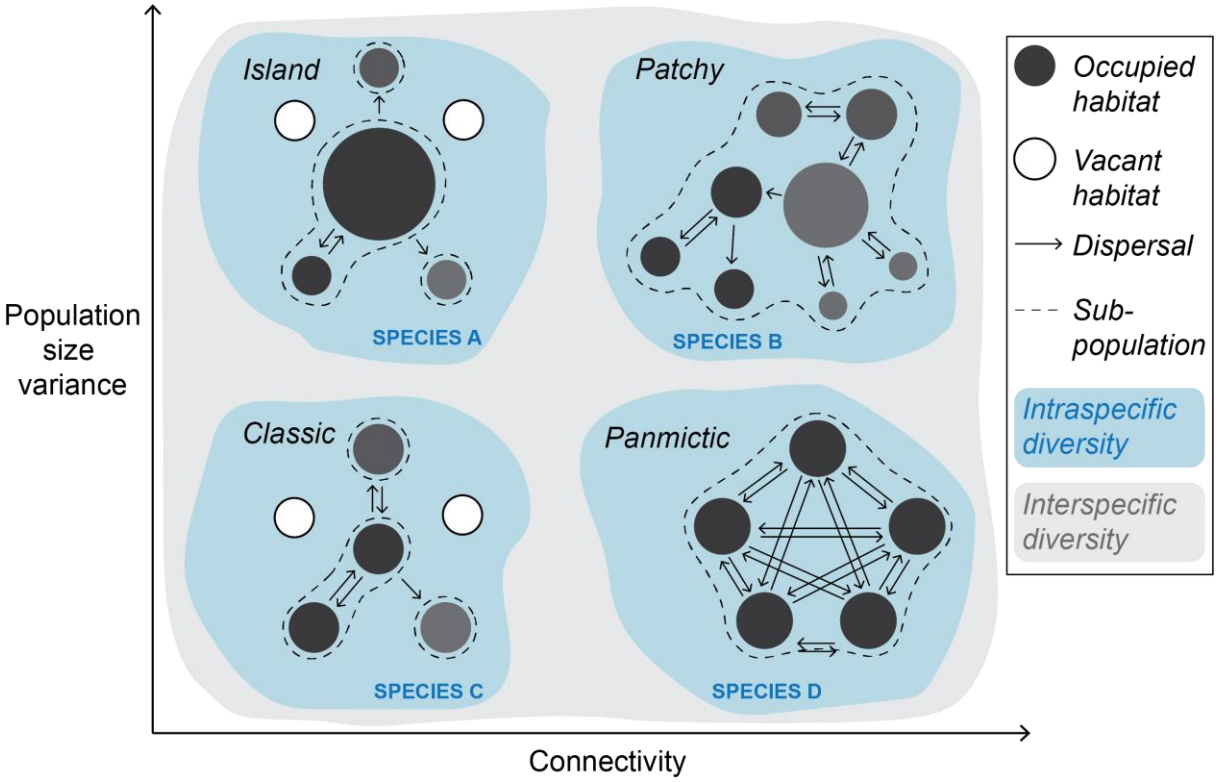



**Fig. 1.1.** Biological diversity as a function of variance in population size (represented by circle size), physical distance and dispersal (which together determine connectivity). Interspecific diversity comprises variability between assemblages of different species (such as the hypothetical species A-D in this case), while intraspecific diversity reflects the variability within one species, which is higher in situations of classic or island metapopulation structure as opposed to patchy or panmictic populations. Adapted from figures in Harrison & Taylor (1997) and Fullerton et al. (2011).

Irrespective of the overall population terminology, it remains relevant to determine the level of uniqueness, environmental specialization or local adaptation in putative, genetic sub-populations. Genetic clusters are often identified with genome-wide data and population geneticists may be satisfied with identifying such clusters irrespective of the systematic classification (species, subspecies, metapopulation, subpopulation) applied to them (Galtier 2019). It is important, however, to either classify genetic clusters as species, or integrate subspecies and population as important levels in conservation in order not to risk extinction of diversity that was recognized but not formally named and thus ignored or given a lower priority.

When genetic clusters can be identified, one should cautiously assume that they may contain important, possibly further, still unknown variation. This is crucial under rapid environmental change or anthropogenic disturbance, because genetically diverse groups of animals are expected to be more resilient against change and have higher potential to adapt to new conditions (Reusch *et al.* 2005; Harrisson *et al.* 2014). Most evolutionary adaptation responses are indeed derived from standing genetic variation (Bernatchez 2016). Three-spined sticklebacks (*Gasterosteus aculeatus*) are the prime example in fishes, where a similar genetic basis has repeatedly led to adaptations for freshwater habitats (e.g. Hohenlohe *et al.* 2012; Raeymaekers *et al.* 2017). Genetic variability thus serves as evolutionary potential, which should be protected for most successful long-term conservation management. Management units, however, often reflect (historical) political rather than biological boundaries, in particular in marine fisheries (Waples *et al.* 2008). In addition, fisheries management traditionally relies on the definition of stocks rather than (sub-)populations that are then separately managed, assuming them to be discrete and amenable to independent exploitation (FAO 2002). Mismatches between management units such as stocks and biological realities can lead to severe misassessments and unsustainable practices (Reiss *et al.* 2009; Kerr *et al.* 2016). Such problems are exacerbated through additional pressure on fish diversity from a variety of global anthropogenic influences.

**Box 1.1.** Important terms and concepts and their definition as used throughout this thesis.

| | |
|---|---|
| **Biodiversity** | The biological variation (genetic, ecological or both) between different groups of organisms from the same species, or between different species, or between different species assemblages (communities). |
| **Connectivity** | The successful exchange of individuals between regional groups of organisms of the same species over short (ecological connectivity) or long (evolutionary connectivity) time scales. |
| **Dispersal** | The movement of individuals (often as eggs or larvae) in space, leading to gene flow between (sub-)populations. A special form of migration, that usually considers the movement from birth to settlement or maturation. |
| **Gene flow** | The transfer of genetic variation (alleles) from one (sub)population to another, through dispersal or migration (see dispersal). |
| **Genetic population structure** | Genetic differences (in allele frequencies) within a population (or between subpopulations) caused by restricted gene flow. |
| **Metapopulation** | A set of (sub)populations that occupy different areas and may be linked through exchange of individuals (emigration/immigration). |
| **Local adaptation** | A subpopulation that is particularly fit to survive in its (local) habitat. A group of individuals must have evolved in response to selective pressure to attain local adaptation. |
| **Panmixia** | Panmixia occurs in a random mating population, where gene flow is high because all individuals are able to interbreed with one another (at the same likelihood; due to a lack of geographic or ecological barriers). |
| **Population** | Organisms of the same species that occupy the same (possibly large) area and (can) interbreed (see also Hartl & Clark 2007). |
| **Species** | A group of organisms that form a "separately evolving metapopulation lineage" (De Queiroz 2007). |
| **Stock** | A subpopulation in a fisheries context, that is individuals of one fish species that occupy a clearly defined geographical area and may have characteristic life history parameters (but the connectivity between subpopulations is often ignored). |
| **Subpopulation** | A population in a metapopulation context, that is the individuals within this subpopulation regularly interbreed, while they may be connected to other subpopulations through less common interbreeding events (enabled via dispersal/migration). |

Anthropogenic stressors threaten the intraspecific diversity of fishes by causing local depletion or extinction (Ryman *et al.* 1995). Many stressors may act upon fish populations concomitantly (Crain *et al.* 2008), which can lead to complex patterns that are still poorly understood. A straightforward example is local extinction caused by interference or blocking of migration pathways, as evidenced in salmonids where the physical alteration of rivers (dams and other hydrological features) has extirpated many local varieties (Gustafson *et al.* 2007). Less obvious changes in the environment can have substantial effects as well. Introduction of non-indigenous species may lead to competition that threatens local fish diversity, although this threat is considered relatively uncommon in marine (fish) species with the notable exception of lionfish (*Pterois volitans* and *P. miles*) in the West Atlantic Ocean (Green *et al.* 2012). Pollution and increased nutrient loads constitute major threats for fish populations, even when their impact on individual specimens is not obvious (Hamilton *et al.* 2015). Consequently, pollution can alter the genomic make-up of wild fish and thus impact intraspecific genetic diversity (Reid *et al.* 2016). Hybridization, i.e. the interbreeding of different varieties, threatens the genetic integrity of natural populations, by eroding rare genetic diversity (Ryman *et al.* 1995). When domesticated or translocated individuals mix with a native population, low frequency – but nevertheless important – variation (of alleles, genes, or phenotypes) may be lost or reproduction may be reduced through maladaptive hybrids. This is of special concern in species that are farmed in aquaculture, where wild individuals may hybridize with farm escapees (Hillen *et al.* 2017; Faust *et al.* 2018). Climate change exposes natural fish populations to a multitude of selective pressures (FAO 2018), that are often similar to existing natural selection forces, but at accelerated rates. Consequences include large scale distribution shifts (Pinsky *et al.* 2013; Fossheim *et al.* 2015) and reduced overall productivity (Free *et al.* 2019), which may alter the degree of connectivity or isolation between subpopulations (Young *et al.* 2018). Finally, harvest itself imposes artificial selection on the exploited (meta-)population. The reduction of individuals in a given metapopulation leads to an overall reduction in genetic diversity (Hauser *et al.* 2002; Pinsky & Palumbi 2014). In addition, the typically size-selective nature of fisheries may lead to rapidly changing phenotypes in the fished species (Conover & Munch 2002; Swain *et al.* 2007). These observed parallel phenotypic responses, however, can result from complex genomic changes that are inconsistent among populations (Therkildsen *et al.* 2019). Importantly, these results indicate how much is still unknown about the interaction of fisheries-imposed selection and genomic diversity. Nevertheless, the genetic effects of fisheries can occur well before population collapse or dramatic reductions become noticeable (Olsen *et al.* 2004). For example, reduced effective population size may be a direct consequence of fishing pressure that could therefore serve as an early warning system (Hauser *et al.* 2002; Allendorf *et al.* 2008).

As global threats to fish diversity at all levels increase, it is crucial to strengthen our understanding of fish diversity in time and space and from populations to species. This can partly be achieved using molecular tools, although approaches from other disciplines, such as traditional morphology-based taxonomy, environmental and biodiversity data science, and modelling techniques, are highly useful complements. The Southern Ocean is a fascinating, remote and – compared to heavily used marine regions such as the North Sea – pristine marine ecosystem (Halpern *et al.* 2008), where the study of species diversity, population structure and connectivity is highly topical. In the following, we will first review current trends in some of the methodologies used within this thesis and then explore the background of the study system, that is the Southern Ocean and its fish fauna.

## 2. Molecular tools to investigate fish diversity and connectivity

The study of DNA has been pivotal for our modern understanding of biological diversity, species identification and delimitation, phylogenetics, phylogeography, population genetics, connectivity and more. Modern taxonomy and systematics ideally integrate traditional, morphological and modern molecular tools to clearly identify species and update previous systematic classifications (e.g. Baeza & Behringer 2017; Seefeldt *et al.* 2017; Glover *et al.* 2018). At times, it may be arduous for ecologists to follow the confusion that ensues when molecular phylogenetics indicate that one species actually comprises two or *vice versa* (e.g. Addamo *et al.* 2016; Dornburg *et al.* 2016; Momigliano *et al.* 2018). However, to fully acknowledge biological diversity it is indispensable to update classifications when molecular work uncovers previously unknown relationships (Miya *et al.* 2010). Genetic markers are ubiquitously used in that sense and have revolutionized much of our understanding of the tree of life. Solid phylogenetic inferences are nowadays derived from the simultaneous screening of several mitochondrial (mt) and nuclear genetic loci, or phylogenomic data sets with thousands of markers (Moriarty Lemmon & Lemmon 2013; Eytan *et al.* 2015). Before high-throughput DNA sequencing became accessible most studies used short, informative loci to detect species and their relationships with each other. Mitochondrial DNA loci such as the *cytochrome oxidase I subunit* (COI) gene are applied as "barcoding" markers, which rapidly identify many animals to species level. This approach has also enhanced basic ecology, as difficult to identify specimens can be ascribed a species status without expert knowledge. DNA barcoding was developed in the early 21$^{st}$ century and caused a leap in species discoveries and readily available databases and systems to determine species status of specimens (Hebert *et al.* 2003; Ratnasingham & Hebert 2007, 2013). DNA

barcoding campaigns have substantially increased our knowledge, of fish diversity in particular (Ward *et al.* 2005, 2009; Dettai *et al.* 2011; Kenchington *et al.* 2017). Notwithstanding, DNA barcoding was also subject to criticism regarding its utility and robustness (Deagle *et al.* 2014). Clear guidelines specify the appropriate implementation of DNA barcoding (Collins & Cruickshank 2013), which can then be a valuable initial step for further phylogenetic and population genetics assessments (Hajibabaei *et al.* 2007). In addition, DNA barcoding can be used to determine large scale phylogeographic patterns (Bucklin *et al.* 2011), uncover trophic relationships (Pompanon *et al.* 2012; McInnes *et al.* 2017; Clarke *et al.* 2018), microbiome communities (Lluch *et al.* 2015; Heindler *et al.* 2018), or food labelling issues (Wong & Hanner 2008; Pardo *et al.* 2016; Christiansen *et al.* 2018).

**Box 1.2.** Molecular tools to study fish diversity.

| | |
|---|---|
| **Allozymes** | Variant forms of enzymes that are separated through gel electrophoresis and used as population genetic data. |
| **AFLPs** | Amplified Fragment Length Polymorphisms; a method similar to microsatellites, but less costly because no locus-specific primers have to be developed. |
| **DNA barcoding** | Few or single DNA loci such as the cytochrome oxidase I subunit (COI) gene that are sequenced to identify the taxonomic rank of a given sample. |
| **DNA metabarcoding** | DNA barcoding material from several organisms at the same time to identify the taxonomic composition; for example to identify a fish's diet by sequencing its stomach content. |
| **Microsatellites** | Tandem repeat sequences in the genome. Their size can be determined after amplification and gel electrophoresis and used as population genetic data. |
| **Reduced representation sequencing (RRS)** | Approaches yielding SNP genotypes after genomic complexity reduction using restriction enzymes, such as restriction site associated DNA sequencing (RADseq) or genotyping by sequencing (GBS). More costly than microsatellite genotyping or DNA barcoding. |
| **SNPs** | Single Nucleotide Polymorphisms; point mutations in the genome. Large numbers of SNPs can be obtained using RRS or whole genome re-sequencing. |
| **Whole genome (re-)sequencing** | Sequencing of the entire genome of several individuals to gain population genomic data, such as SNPs. Highest level of information, but also highest cost. |

Contrastingly, in depth studies of population genetics rely on techniques that provide higher resolution than a single mtDNA locus such as COI (see also Fig. 1.6). Decades ago, typical markers of choice were allozymes or amplified fragment length polymorphisms (AFLPs), but more recent investigations usually employ microsatellites or single nucleotide polymorphisms (SNPs, Box 1.2). Both marker types offer advantages and disadvantages (Table 1.1), that need to be considered before deciding which approach is most promising to reach a particular objective.

Short tandem repeats or simple sequence repeats (SSRs), commonly called DNA microsatellites, are tandem repeat motifs of 1 to 6 base pairs, which are repeated for on average 100 times. When such a repeat sequence is flanked by conserved regions, primers can be designed to bind to that region allowing for subsequent amplification via polymerase chain reaction (PCR). Microsatellites are putatively neutral markers with elevated mutation rate offering high allelic diversity, which makes them versatile to assess for instance genetic differentiation (Chistiakov *et al.* 2006). Due to the large number of studies using microsatellites, these are well understood and accepted but some drawbacks prevail. The *de novo* development of primers can take significant time and effort (Zane *et al.* 2002), but microsatellites of many species have been published, including Southern Ocean fish (Van Houdt *et al.* 2006; Agostini *et al.* 2013; Papetti *et al.* 2016a). If flanking regions are conserved the same primers may be used for related species (cross-amplification), albeit with caution. In addition, standard microsatellite analysis has comparably low requirements regarding the input material (DNA quality and quantity) and can even be used on semi-degraded, historical or ancient DNA (Cuveliers *et al.* 2011). After amplification of target regions, PCR products are run on automatic capillary sequencers. Here, different platforms sometimes yield slightly different results regarding fragment size. Likewise, genotyping within platforms is not always without problems. Mutations in the primer annealing sites may lead to amplification failure, the so-called “null alleles”. Stutter peaks, allele dropout and insertions/deletions can obscure results, therefore measures such as repeated genotype scoring by independent readers or re-runs need to be implemented. Selkoe & Toonen (2006) and Pompanon *et al.* (2005), for example, provide guidelines and suggested quality control measures for microsatellite analysis. Recent progress in the field is achieved by employing high throughput DNA sequencing to expedite microsatellite identification and/or genotyping (Vartia *et al.* 2016; Pimentel *et al.* 2018). Microsatellite development can thus be achieved in three months (e.g. Kerkhove *et al.* 2019). Multiplexing (i.e. co-amplification of several loci in a single PCR reaction mix) and poolplexing (running several PCR products in parallel on e.g. automated capillary sequencing systems) are common to decrease costs and time spent. The amount of

information depends on the number of loci (usually from a couple to dozens) and the level of polymorphism at each locus. The cost per locus per sample amounts to about 1 euro in consumables; typically, 15 to 20 markers would be used. In conclusion, microsatellites are still versatile and cost-effective molecular markers, but the use of quality controls and multiplexing should be emphasized (Guichoux *et al.* 2011).

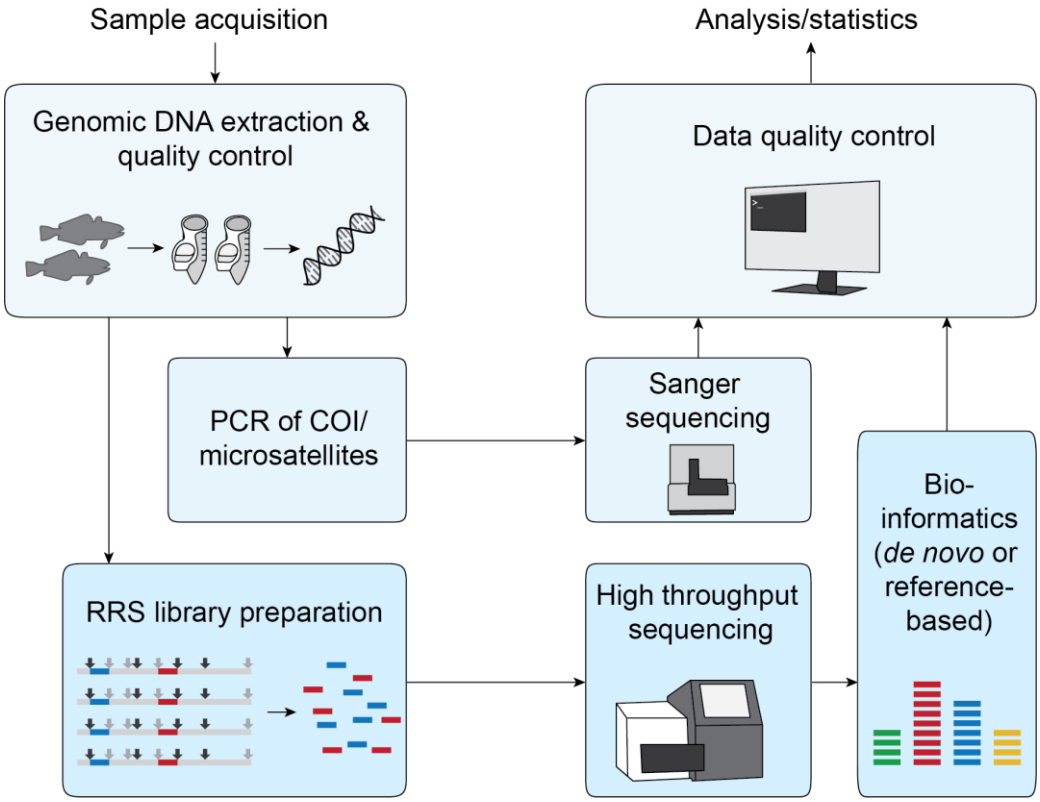



**Fig. 1.2.** Workflow of molecular methods (classic or high throughput) to investigate genetic diversity in/between fishes. The high throughput approach (bottom) with reduced representation sequencing (RRS) requires considerably more work during all phases, but provides higher genomic resolution (see also Fig. 1.6).

In recent years, the rise of massively parallel high throughput sequencing (often described by the rather blurry term "next generation sequencing") has led to an increasing use of point mutations, i.e. SNPs, as the marker of choice. Their biallelic information content, however, is lower than the multi-allelic microsatellites, with on average five SNPs equaling one microsatellite locus . SNP loci are not necessarily evolutionary neutral as they might be part of or linked to genes. Therefore, outlier analysis can detect polymorphisms that might be affected by selection, because of their linkage to genes or other coding DNA. Genotyping of many individuals at many SNPs with comparably low costs is achieved by micro-array platforms (Narum *et al.* 2013). These so-called SNP-chips require substantial prior knowledge of the

genome (i.e. ideally a sequenced reference genome) and are designed specifically for each project. Ascertainment bias, however, will hamper the detection of rare population-specific alleles (Helyar *et al.* 2011). Additionally, a designed SNP-chip may not be applicable to any other organisms but the target species.

Some difficulties of SNP-chips can be overcome by sequencing and genotyping large parts of the genome in parallel. Using restriction enzymes genomic DNA is cut into smaller fragments and subsequently the multiple copies of these fragments are sequenced (Fig. 1.2). Such an approach offers reduced complexity compared to whole genome sequencing, while yielding high resolution with thousands of markers across many individuals (Davey & Blaxter 2010; Davey *et al.* 2011). A variety of methods with that principle have been developed, that can all be summarized under the umbrella term "reduced representation sequencing" (RRS; Campbell *et al.* 2018). The most common protocols are restriction site associated DNA sequencing (RADseq; Baird *et al.* 2008) and genotyping-by-sequencing (GBS; Elshire *et al.* 2011). A very popular variety has introduced an additional restriction enzyme and a subsequent DNA size selection step (ddRAD; Peterson *et al.* 2012). The advantage of this approach is the great flexibility and scalability, providing a relatively inexpensive method to identify SNPs and population genotypes. It is suitable for organisms with and without reference genomes and costs amount to about 35 euro in consumables (including sequencing) per individual. Because RRS methods are relatively recent, they must be used with caution. For instance, it is important to attain good coverage (i.e. number of reads per locus per sample) and to distinguish true SNPs from sequencing errors (Puritz *et al.* 2014a). RRS data has several sources of bias relating to uneven PCR amplification or genotyping error, that need to be considered, but can be identified and excluded (Davey *et al.* 2013). Furthermore, the computational and statistical analyses in general are more demanding and involve substantial bioinformatics efforts (Andrews & Luikart 2014).

While more expensive and labor-intense than microsatellites, screening large amounts of SNPs throughout the genome provides information about all evolutionary forces (migration, drift, selection, mutation). Such an approach allows us to obtain the most information possible from difficult to collect samples, such as from environmental or ancient DNA collection or from remote ecosystems. Using RRS to obtain SNP genotypes creates vast possibilities to answer more specialized questions about e.g. genomic adaptation (Gagnaire *et al.* 2015) and structure (Corander *et al.* 2013).

**Table 1.1.** Summarized advantages (+) and disadvantages (-) of microsatellite and SNP loci as genetic markers to study population structure and connectivity.

| | Microsatellites | SNPs |
|---|---|---|
| **Development/ isolation** | + already available in many species (especially widely studied fish; Van Houdt *et al.* 2006); relatively few loci are sufficient (Guichoux *et al.* 2011) | + depending on the method, but once set up thousands of loci can be isolated quickly (Davey & Blaxter 2010) |
| | - relatively lengthy development phase; ascertainment bias may occur (Zane *et al.* 2002; Guichoux *et al.* 2011) | - depending on method, ascertainment bias may occur (Helyar *et al.* 2011) |
| **Characteristics** | + potentially highly polymorphic; transfer between related species possible; unproblematic regarding DNA quality & quantity (Chistiakov *et al.* 2006) | + due to the usually large number of loci, information about different evolutionary histories (selection, drift, etc.) can be found (Guichoux *et al.* 2011; Peterson *et al.* 2012) |
| | - subjectivity of allele scoring; genotyping errors due to stutter, large allele dropout etc.; comparisons between platforms difficult (Selkoe & Toonen 2006; Guichoux *et al.* 2011) | - usually only two alleles per locus; high quality DNA extracts necessary; sequencing errors and genotyping errors due to allele dropout, PCR duplicates etc.; difficult to detect errors(Guichoux *et al.* 2011; Davey *et al.* 2013; Puritz *et al.* 2014a) |
| **Costs** | + low cost (Selkoe & Toonen 2006, Guichoux *et al.* 2011; Puckett 2016) | + high throughput (Elshire *et al.* 2011) |
| | - rather low throughput (Guichoux *et al.* 2011) | - more expensive (but costs are dropping steadily; Peterson *et al.* 2012; Puckett 2016) |
| **Interpretation** | + well-studied and accepted; good power to detect divergence (Selkoe & Toonen 2006) | + most comprehensive information next to whole genome sequencing, better e.g. to estimate population history, genomic diversity (Davey & Blaxter 2010, Guichoux *et al.* 2011; Bohling *et al.* 2019) |
| | - limited amount of information, not genome wide, low power at low differentiation levels (Guichoux *et al.* 2011; Bohling *et al.* 2019) | - possible for non-model species, but involves some extra effort; stochastic effects; advanced statistics/ bioinformatics (Andrews & Luikart 2014) |

Finally, several additional data sources and analytical approaches can be used in conjunction with molecular data. Such multidisciplinary studies are powerful to reveal complex biological realities. Occurrence data from animals, for example, is accessible at online databases such as the Global Biodiversity Information Facility (GBIF), the Ocean Biogeographic Information System, or through the SCAR Biogeographic Atlas of the Southern Ocean (De Broyer et al. 2014) and the Antarctic Biodiversity Information Facility (data.biodiversity.aq). With a suitably extensive occurrence and environmental data set it is possible to use species distribution modelling (SDM), sometimes referred to as ecological niche modelling, to predict occurrence probability in space or time (Elith & Leathwick 2009). Environmental data can be mined for example from the World Ocean Database (Boyer et al. 2013), or – for future predictions – from the Bio-ORACLE database (Tyberghein et al. 2012) based on IPCC scenarios. The combination of these two data sources (species occurrence and environmental variables) with appropriate model algorithms can then support ecological or evolutionary inferences (Elith & Leathwick 2009). Model realism and robustness should be carefully checked and several tools to ensure appropriate parameterization, especially in data-limited situations, exist (Guillaumot et al. 2018a, 2019). In the context of population differentiation and connectivity, SDMs can be particularly useful to provide an extended view of the entire (predicted) species distribution landscape. In addition, environmental data from the sites samples for genetic analyses is often used to search for correlations between genotypes and environmental variables. Uncovering genotype-environment associations can help determining loci that are underlying local adaptation (Rellstab et al. 2015). A powerful tool for genotype-environment associations is redundancy analysis, which outperforms univariate statistical methods when used with large genomic data sets (Forester et al. 2018). This method can be used in conjunction with outlier tests based on genetic differentiation (Hoban et al. 2016) to detect even weak signatures of multilocus selection (Forester et al. 2018). Lastly, individual-based modelling (IBM) is widely applied to characterize connectivity, particularly of marine organisms. As dispersal is very challenging or virtually impossible to directly observe in the ocean (but see Pinsky et al. 2017), a modelling framework based on oceanography models can be employed to simulate the dispersal of eggs and/or larvae by ocean currents (Cowen & Sponaugle 2009). The resulting spatially explicit prediction of connectivity between habitat patches can be interpreted in conjunction with estimates based on genomic data. The combination of environmental data, IBM and genomics is often referred to as landscape or seascape genomics (Riginos et al. 2016). Clearly, molecular tools are most powerful when used in integrative, multidisciplinary studies to answer questions about fish diversity between species and populations.

### 3. The Southern Ocean – a biodiversity sanctuary?

The Southern Ocean can be defined as the ocean south of the Subtropical Convergence (Gon & Heemstra 1990), or south of the Antarctic Convergence (Clarke *et al.* 2005), or using a mix of physical and socio-economic borders (Kock 1992; CCAMLR 2018). These definitions are naturally arbitrary to some extent and the arguing for or against certain definitions is here left to oceanographers. Instead, for the remainder of this thesis, we will use the definition as employed by the relevant fisheries management organization, i.e., the Commission for the Conservation of Antarctic Marine Living Resources (CCAMLR; CCAMLR 2018; Fig. 1.4). This demarcation roughly follows the Antarctic Convergence, but also includes islands near it. Several sub-divisions are generally accepted within the Southern Ocean, most notably the Weddell Sea and Ross Sea that also separate West and East Antarctica (Fig 1.3). Both seas cover extensive deep-sea basins and have pronounced regional hydrodynamics, including the large Weddell and Ross Sea gyres that are important for deep and bottom water formation (Dotto *et al.* 2018; Vernet *et al.* 2019). The Southern Ocean can also be divided into the Atlantic, Pacific, and Indian Ocean sectors according to the respective ocean basin (Kock 1992). Other seas include (in clockwise order, starting off Patagonia): the Scotia, Lazarev, Cosmonauts, Cooperation, Davis, Mawson, D'Urville, Amundsen, and Bellingshausen Sea (Orsi *et al.* 1995; Thompson *et al.* 2018; Vernet *et al.* 2019; Fig. 1.3). The dominant oceanographic feature in the entire Southern Ocean is the Antarctic Circumpolar Current (ACC), which circulates clockwise around Antarctica (Orsi *et al.* 1995; Rintoul & Naveira Garabato 2013). In fact, it is the world's largest ocean current system facilitating exchange between deep and shallow water layers of the global ocean (Rintoul & Naveira Garabato 2013). The ACC comprises a series of ocean fronts, that is sharp transition zones between water masses, and jets, that is intense, narrow currents (Chapman *et al.* 2020). These jets are dominating the ACC's flow, while mesoscale eddies may form alongside meandering fronts. Traditionally, three main fronts are considered to form the ACC, the Sub-Antarctic Front (SAF), the Polar Front (PF) and the Southern ACC Front (sACCF; Orsi *et al.* 1995; Fig. 1.3). North of these fronts lies the Subtropical Front (STF) and to the South the Southern Boundary Front (SBF) near the Antarctic continent (Fig. 1.3). These fronts were considered circumpolar, eastward-flowing and with geographically variable extent, owing largely to bathymetric forcing (Orsi *et al.* 1995; Pollard *et al.* 2002). The Polar Front, also sometimes referred to as Antarctic Convergence (cf. definitions of the Southern Ocean above), is a central element of the ACC. Warmer sub-Antarctic water meets cold Antarctic water in this zone, generating a sudden drop in surface seawater temperature (on average 1.7° C; Freeman *et al.* 2016). The PF is a dynamic barrier that limits faunal exchange between the Southern Ocean and

other marine regions (Clarke *et al.* 2005). In the Weddell and Ross Sea, the ACC forms the northern boundary of the respective gyres. Ross and Weddell Sea gyres also facilitate exchange between the ACC and the contintental shelf through westward near-slope flow (Ross Sea, Dotto et al. 2018) and mesoscale eddies and the Antarctic Slope Front (Weddell Sea, Vernet *et al.* 2019).

It is now clear that the ACC is not always (both spatially and temporally) comprised of a well-defined, continuous frontal system as depicted in Figure 1.3. Instead, the ACC is a dynamic, rearranging current system with changing properties, including meandering filaments, jets, eddies, and subfronts, that appear and disappear and change position (Thompson 2008; Langlais *et al.* 2011; Rintoul & Naveira Garabato 2013; Chapman 2017; Chapman *et al.* 2020). We will use the frontal positions as defined by Orsi *et al.* (1995) in their seminal work as guidelines for the remainder of this thesis (Fig. 1.3), but bear in mind that these absolute front positions are merely approximations of a highly dynamic feature and that the PF in fact passes south of the Kerguelen Islands (Park *et al.* 2014). In order to link individual animal behavior to oceanography it would be crucial to regard fine-scale dynamics of the Southern Ocean fronts (Chapman *et al.* 2020). However, for mainly large-scale, evolutionary questions that explicity do not focus on small scale variability the approximations after Orsi *et al.* (1995) should be informative and can be complemented with more accurate individual-based modelling to clearly link biological dispersal and oceanography (Young *et al.* 2015). Additional current systems operate in the Southern Ocean, such as the Antarctic Peninsula Coastal Current (APCC) and Antarctic Slope Current (ASC), which flow in the opposite direction to the ACC but close to the continent (Fig. 1.3; Moffat *et al.* 2008; Thompson *et al.* 2018). Like the ACC, the ASC is also a large and coherent, but spatially and temporally variable current system. Generally, the ASC forms between Antarctic shelf waters and warmer circumpolar deep water with temperature differences between 1° - 3° C (Thompson *et al.* 2018). At the tip of the western Antarctic Peninsula, the ASC meets the Southern Boundary, while a costal current continues towards the APCC into the Bransfield Strait, resulting in a complex and multifaceted current system with strong cross-shore gradients (see Moffat & Meredith 2018 for a detailed overview figure). At the central western Antarctic Peninsula, the APCC flows almost parallel to the ACC, but in opposite direction, leading to a strong cross-shelf density gradient (Moffat & Meredith 2018). Circulation patterns at the Antarctic Peninsula are likely strongly influenced by ocean eddies, but further modelling and observational studies are needed to resolve this. The latter statement is true for our understanding of Southern Ocean physical oceanography in general, albeit tremendous progress over the past decades due to increased capacities for autonomous sampling and modelling.

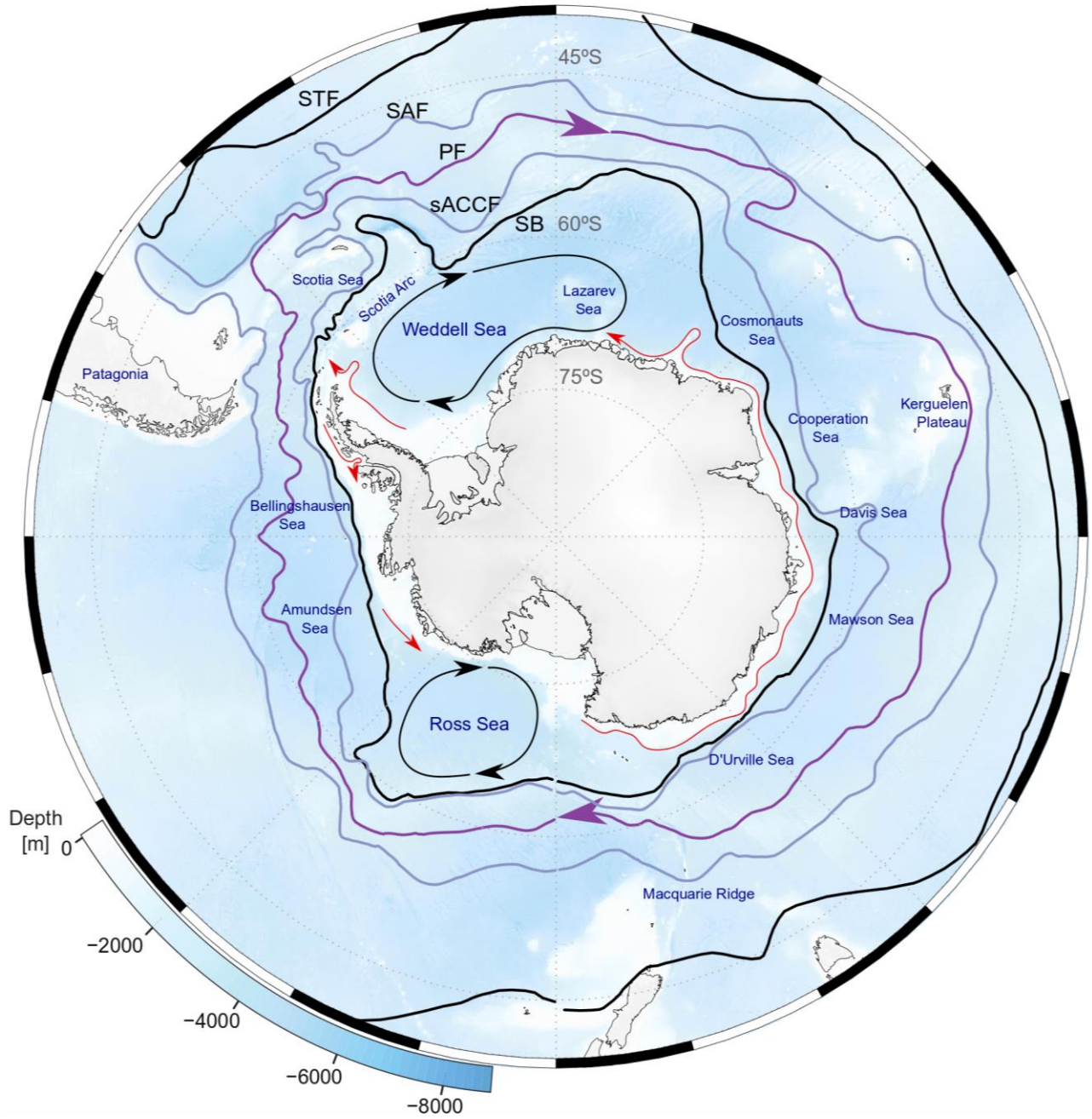



**Fig. 1.3.** Antarctica and the Southern Ocean with major current systems. The major fronts of the Antarctic Circumpolar Current are shown (from North to South): the Sub-Antarctic Front (SAF, grey), the Antarctic Polar Front (PF, magenta), and the Southern ACC front (sACCF, grey); as well as the Subtropical Front (STF, black) to the north and the Southern Boundary Front (SB=SBF, black) near the Antarctic continent. Further important hydrodynamic features are the Weddell and Ross Sea gyres (black) and the Antarctic Peninsula Coastal Current and Antarctic Slope Current (both in red and approximated after Moffat et al. (2008), Moffat & Meredith (2018) and Thompson et al. (2018)). Background shading (blue-white) indicates depth. Map produced using R package SOmap v0.5.0 (Maschette et al. 2019) with data from Orsi et al. (1995).

The Southern Ocean was established after the breakup of Gondwana, leading to an isolation of the Antarctic continent approximately 33 million years ago. Australia was then separated from Antarctica and South America from the Antarctic Peninsula. The latter event led to the opening of the Drake Passage, enabling the formation of the circumpolar current. The onset of the ACC some 25 million years ago initiated a cooling of the waters around Antarctica and has fundamentally changed the ecosystems of this region. Ice sheets thickened and led to a deepening of the continental shelf due to glacial isostasy (Anderson *et al.* 2002). Overall, the shelf of the Antarctic continent is comparatively deep (500 m on average) and narrow, although of a larger extent in the Weddell, Ross and Bellinghausen Sea. The remainder of the Southern Ocean is mostly between 3000 and 5000 m deep. Shallower areas apart from the continental shelf exist but are spatially separated, most prominently the Scotia Arc, the Kerguelen Plateau and the Macquarie Ridge (Fig. 1.3). Isolation through physical distance to other continents, the ACC, and the vast deep-sea areas has given rise to a unique fauna in the Southern Ocean (Clarke & Crame 2010). Fossil records demonstrate that millions of years ago the fauna of Antarctica was very different from its modern counterpart. The fish fauna for example was dominated by codfishes (Gadiformes) like the extant fish fauna in the northern hemisphere (Schwarzhans *et al.* 2017). This is not surprising, considering that paleotemperatures during the Eocene were well above the freezing point (Knox 1994). Now, millions of years later, charismatic animals like penguins, whales, seals, and icefish roam the cold, ice-laden waters and are a prominent attraction for explorers and tourists alike.

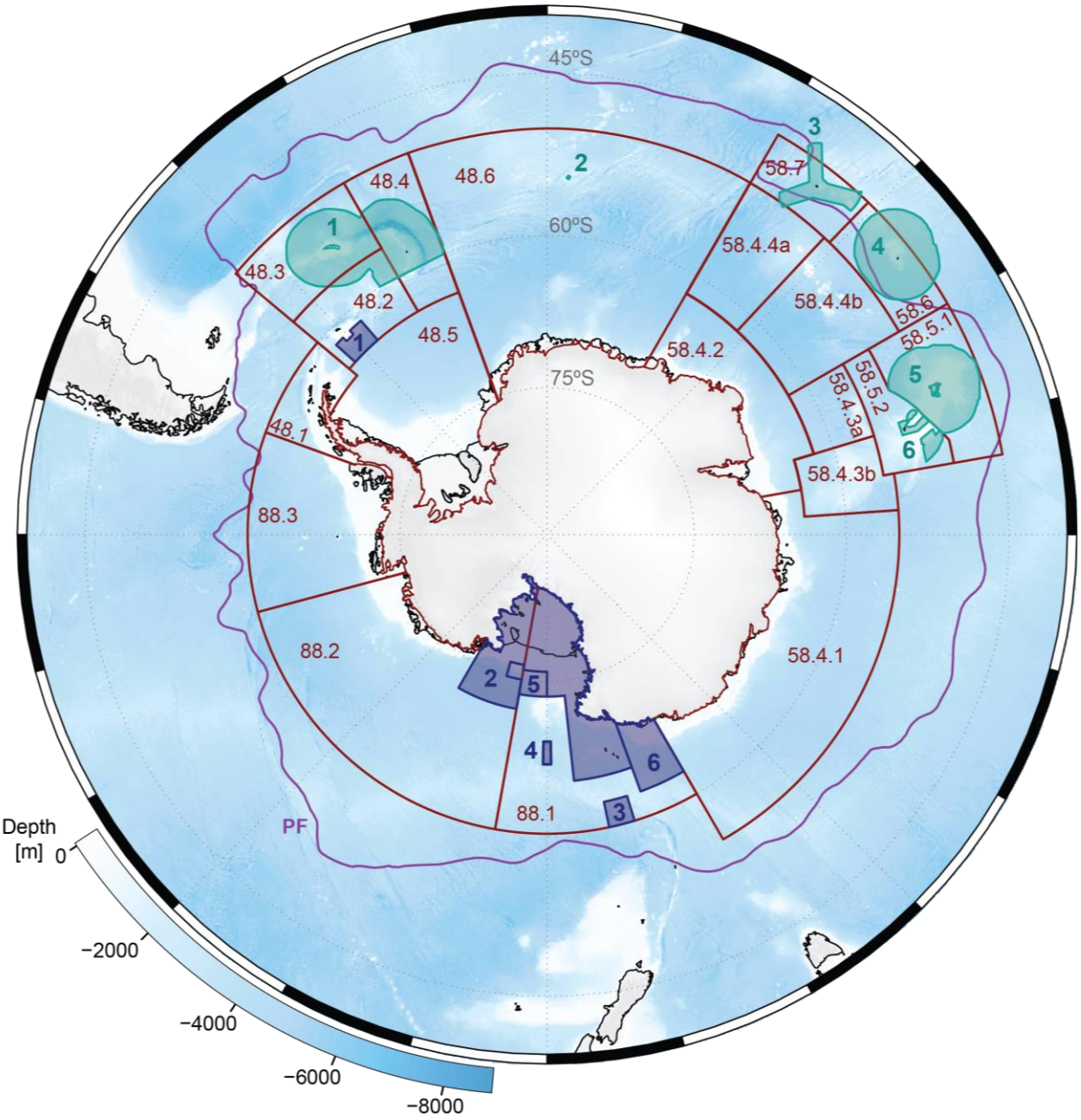



**Fig. 1.4.** Antarctica and the Southern Ocean. The convention area of the Commission for the Conservation of Antarctic Marine Living Resources (CCAMRL; red boxes) is used in this thesis as definition of the Southern Ocean. The different areas are distinct management zones of CCAMLR for the Atlantic (48.1-6), Indian (58.4-7), and Pacific (88.1-3) sectors of the Southern Ocean. Currently declared CCAMLR Marine Protected Areas (MPAs) are shown as dark blue boxes and other MPAs and Marine Reserves as dark green boxes/circles. The Antarctic Polar Front (PF) is shown as solid magenta line. Background shading (blue-white) indicates depth. Map produced using R package SOmap v0.5.0 (Maschette *et al.* 2019) with data from CCAMLR.

Exploration of the Southern Ocean and Antarctica started in the early 19th century and was at a peak during the "Heroic Age" of Antarctic exploration around 1900, when several nations competed to reach the South Pole. Direct anthropogenic impacts on the biota of the Antarctic started concomitantly with exploration. Expeditioners were using local animals to supply fresh food and other materials. Soon, extensive industries for fur seal hunting were established, later on followed by excessive whale hunting (Aronson *et al.* 2011). Shore-based sealing was a lucrative industry as seals were easy to kill and delivered highly valuable pelts. Antarctic fur seals (*Arctocephalus gazella*) were first caught in record numbers on South Georgia and as this population was nearly wiped out, sealers ventured further south (Hemmings 2015). By 1900 fur seals in the entire Southern Ocean were almost extinct and yet sealing activities fully ceased only in the 1960s (Hucke-Gaete *et al.* 2004). As sealing became difficult due to absent targets, the focus shifted successively towards baleen whales. Whaling was also conducted from the shore, but especially later increasingly from heavily motorized vessels (Hemmings 2015). These activities focused on different whale species, usually the most valuable first and when that species became rare, others were targeted. Whaling not only dramatically reduced the numbers of whales in Antarctic waters, but pushed many whale species globally to the brink of extinction (Baker & Clapham 2004). Commercial whale hunting was banned in 1986 through a moratorium imposed by the International Whaling Commission, but illegal whaling and whaling activities under disguise of scientific research continued into recent years (Aronson *et al.* 2011; Hemmings 2015).

The first account of harvesting fishes in the Southern Ocean is documented as early as 1800, where seal hunters discovered Antarctic "cod-fish" (see Gon & Heemstra 1990 and references therein). Supposedly, catching fish for subsistence was common among sealing vessels, but commercial fisheries ignored the Antarctic for a long time. In the early twentieth century first attempts to identify the viability and prospects of commercial fishing in the Southern Ocean were made by whalers on South Georgia (Hureau & Slosarczyk 1990; Kock 1992). In the 1950s exploratory fishing was conducted by Norwegian vessels and in the 1960s the first Soviet trawlers were operating off the South Orkney Islands. But it was not until the late twentieth century, more than 150 years after the first discovery of Antarctic fish as an excellent food source, that large scale industrial fishing was initiated in the Southern Ocean, mainly by trawlers from the Soviet Union (Kock 1992). Much like during the sealing and whaling periods, fisheries at their peak in the 1970s were unregulated and unsustainable by modern standards (Ainley & Blight 2009). The establishment of CCAMLR in 1982 ended at least the most severe overfishing.

However, recovery of historically overexploited animal populations has been slow in many cases, most notably blue whale (*Balaenoptera musculus*; Branch *et al.* 2007) and southern right whale (*Eubalaena australis*; Tulloch *et al.* 2018), but also fish like the marbled rockcod (*Notothenia rossii*; Marschoff *et al.* 2012). Slow growth and long lifespan are often invoked as likely explanations for the long recovery time of Antarctic species (Mintenbeck *et al.* 2012; Koubbi *et al.* 2017).

Nowadays, exploitation is not blatantly unsustainable as centuries to decades ago, but serious commercial interests in the region persist. Fisheries for Antarctic and Patagonian toothfish (*Dissostichus mawsoni and D. eleginoides*) offer high revenue and Antarctic krill (*Euphausia superba* and other spp.) is regarded as one of the most abundant living marine resources of the world. There is much debate about the sustainable nature of these fisheries (Ainley & Pauly 2013; Abrams 2014; Hanchet *et al.* 2015a; Abrams *et al.* 2016; Ainley *et al.* 2016, 2017) and whether or not they can be managed in an ecosystem-based and precautionary approach, as CCAMLR has committed itself to (Kock *et al.* 2007; Constable 2011; Brooks 2013). Other species of fish may become fishery targets (again) in the future. In addition to these direct current and potential future harvesting pressures the biota of the Southern Ocean are facing a multitude of further stressors that are more or less directly linked to anthropogenic influences (Aronson *et al.* 2011). Increasing Antarctic and Southern Ocean touristic activities are inevitably affecting the local ecosystems, although tourism in these regions is strongly regulated, albeit by the industry itself (Liggett 2016). Tourism, but also scientific activity leaves traces in the ecosystem that act together with global impacts. In addition, global processes such as atmospheric distillation and dispersal of a wide range of pollutants influence the region. Pollution and biomagnification of pollutants for example are increasingly documented in the Southern Ocean, including e.g. mercury accumulation in top predators (Yoon *et al.* 2018; Gilmour *et al.* 2019; Liu *et al.* 2019), persistent organic pollutants in fish (Corsolini & Sarà 2017; Strobel *et al.* 2018) and microplastic waste (Waller *et al.* 2017). Global climate change affects the Southern Ocean through warming sea water, increased rates of iceberg scour, ocean acidification, meltwater and sediment inflow and decreased ice cover (Turner *et al.* 2005; Constable *et al.* 2014; Rintoul *et al.* 2018). These effects are rapid in the western Antarctic Peninsula (Mulvaney *et al.* 2012), while some areas of the East Antarctic currently experience opposite trends, which also have consequences for the local fauna (Michel *et al.* 2019). Some species may benefit from these changes, but the majority are projected to suffer from future climate change (Griffiths *et al.* 2017). Instead, non-indigenous

species may invade the Southern Ocean in the near future and the isolation of Antarctica in general will decrease (Byrne *et al.* 2016; Fraser *et al.* 2018).

In a global perspective, the Southern Ocean may seem comparatively pristine (Halpern *et al.* 2008), but a closer look reveals that men[1] have not spared Antarctica in search for resources at all. Many contemporary challenges are rapidly progressing, emphasizing the importance of concerted and effective conservation action (Rintoul *et al.* 2018). These needs have been formally recognized by CCAMLR through the adoption of a framework to establish a network of Marine Protected Areas (MPAs) in the Southern Ocean (CCAMLR 2012). Nine planning domains for the development of such a MPA network were established, comprising the western Antarctic Peninsula and South Scotia Arc, the North Scotia Arc, the Weddell Sea, the Bouvet and Maud area, the Crozet – del Cano area, the Kerguelen Plateau, Eastern Antarctica, the Ross Sea, and the Amundsen – Bellingshausen Sea (CCAMLR 2011). One MPA has been established as a result of these plans. The Ross Sea MPA is to date the largest high seas protected area (Fig. 1.4), but it is also the first that is not declared in perpetuity but with an expiry date and it includes areas where fishing is allowed (Brooks *et al.* 2016). Another MPA is in force since 2009 south of the South Orkney Islands (Fig. 1.4). These CCAMLR MPAs and further additional MPAs and protected zones under national jurisdiction (that is in exclusive economic zones; Fig. 1.4) are important first steps to offer protection for the unique fauna of the Southern Ocean, including many specially adapted fishes.

#### 4. Fishes of the Southern Ocean

The ichthyofauna of the Southern Ocean has been explored for more than 200 years. Research efforts to describe fishes in and around Antarctica have been episodic at times, with many species descriptions during the early twentieth century, a steep increase in knowledge about the ecology of these fishes during the intense fisheries periods, and of course with new information rapidly being added during modern times that are facilitated by our improved ability to conduct research in such harsh conditions (Hureau & Slosarczyk 1990; Kock 1992; Eastman 1993; Jones *et al.* 2008; Griffiths *et al.* 2011; Duhamel *et al.* 2014). Taxonomy and systematics are still not fully resolved for the entire Southern Ocean ichthyofauna and new species, cryptic species, or

[1] For Antarctica's exploration history was largely written by men (Dodds 2009). Fortunately that is changing, even if slowly.

taxonomic revisions are being published on a regular basis (Bernardi & Goswami 1997; Cziko & Cheng 2006; Smith *et al.* 2008, 2011b; McMillan *et al.* 2012; Eakin *et al.* 2015; Dornburg *et al.* 2016). Notothenioidei are currently the most speciose group of fishes in these waters and are dominating the bottom fish fauna, at least close to Antarctica (Eastman 1993, 2005). Other species rich families are the myctophids (24 species in the Southern Ocean as defined by Duhamel *et al.* 2014), macrourids (10 spp.), liparids (85 spp.), and zoarcids (40 spp.) (Duhamel *et al.* 2014). Overall species richness is low compared to other oceans. In total, Duhamel *et al.* (2014) list 374 species in 47 families and another 85 species that are occasionally recorded south of the sub-tropical front. Many of the fishes living in the Southern Ocean, however, are endemic with special adaptations to thrive in this cold and relatively isolated ecosystem.

The opening of the Drake passage and formation of the ACC were key factors that now restrict the north-south and south-north migration of fauna (Barnes *et al.* 2006; Lyle *et al.* 2007), including particularly the faunal exchange among coastal, benthic and (epi-)pelagic fish species (Eastman 1993). Levels of endemism in Southern Ocean fish species are especially high in notothenioids with rates up to 97% (Eastman & McCune 2000). In addition, notothenioids have diversified ecologically and morphologically. Evolving from a benthic common ancestor, several notothenioid species have developed adaptations such as reduced ossification and lipid sacs to achieve neutral buoyancy and fill open niches in the water column (DeVries & Eastman 1978; Hagen *et al.* 2000; Near *et al.* 2003; Eastman *et al.* 2011). Feeding ecology and other life history traits such as parental investment, reproductive output and behavior vary among notothenioids (Eastman 1993; Detrich *et al.* 2005; Campbell *et al.* 2008; Rutschmann *et al.* 2011; Jones & Near 2012). These fishes are consequently viewed as a species flock that has developed through adaptive radiation (Eastman & McCune 2000; Near *et al.* 2012; Lecointre *et al.* 2013; Colombo *et al.* 2014), particularly along the benthic-pelagic axis (Klingenberg & Ekau 1996). Further adaptations to live in subzero waters include the development of antifreeze glycoproteins (AFGPs) from a trypsinogen-like protease (Chen *et al.* 1997; Nicodemus-Johnson *et al.* 2011). These AFGPs are an example of convergent evolution as they have independently evolved in codfishes (Gadidae) in the northern hemisphere, where the AFGP development occurred *de novo* (Baalsrud *et al.* 2017; Zhuang *et al.* 2019). Some notothenioids also developed traits that are arguably maladaptations, at least under future warming, that is they lost the common heat shock response of fish (Hofmann *et al.* 2000; Coppe *et al.* 2013; Huth & Place 2013) and are the only known vertebrate without hemoglobin (Ruud 1954; Sidell & Brien 2006; Xu *et al.* 2015). Notothenioids are particularly dominant in the shelf waters of the Antarctic continent, where

they account for approximately 90% of the finfish biomass (Eastman 2005) and fill most ecological niches, such as under the ice (*Pleuragramma antarctica*, *Trematomus* spp.), in the water column (*P. antarctica*, *Aethotaxis mitopteryx*, *Gvozdarus svetovidovi*, juvenile *Notothenia* spp.) and on the sea floor (various Channichthyidae, Nototheniidae, Bathydraconidae and Artedidraconidae) (Fig. 1.5). The genetic structure of notothenioids has been investigated repeatedly and by different research groups (Volckaert *et al.* 2012). Some species (e.g. *Gobionotothen gibberifrons*, *Chionodraco rastrospinosus*, *Notothenia rossii*) are characterized by high levels of gene flow (Matschiner *et al.* 2009; Damerau *et al.* 2012; Papetti *et al.* 2012; Young *et al.* 2015), while others (e.g. *Champsocephalus gunnari*, *Chaenocephalus aceratus*, *Trematomus* spp.) show signs of population differentiation (Van de Putte *et al.* 2012b; Damerau *et al.* 2014; Agostini *et al.* 2015; Young *et al.* 2015). Most of these studies focused on a limited spatial extent (but see Van de Putte *et al.* 2012) and used a handful of microsatellite loci. The driving factors of these patterns, and patterns of local adaptation are still poorly understood. It would be interesting for example to identify how common local adaptation is among notothenioids and if such patterns can be related to the demographic history of different species. If a notothenioid species was repeatedly driven into glacial refugia in times of historical climate change, this should in theory manifest in genomic signatures (Allcock & Strugnell 2012).

Towards the sub-Antarctic and deep-sea regions more and more non-notothenioid species are regularly documented. Common demersal species include Rajidae, Zoarcidae, Liparidae, Muraenolepididae, and Moridae. Some of these may represent a considerable biomass (Amsler *et al.* 2016), but their biology is still relatively poorly known (but see e.g. Meléndez & Markle 1997; Duhamel *et al.* 2010; Smith *et al.* 2011; Stein 2012; Pinkerton *et al.* 2013; Prirodina & Balushkin 2015). In the open waters of the Southern Ocean, mesopelagic and bathypelagic fish communities are present that resemble those in other areas of the world. This fauna consists predominantly of members of the Myctophidae, Bathylagidae, Gonostomatidae, Stomiidae, Paralepididae, Notosudidae, and Cetomimidae (Duhamel *et al.* 2014). The mesopelagic fauna comprises an enormous biomass that is consequently a key element of the oceanic food web with links to many top predators and energy transport pathways that connect the open ocean with the deep sea and even land (Irigoien *et al.* 2014; Saunders *et al.* 2015a; McCormack *et al.* 2019). Mesopelagic fishes prey mostly on mesozooplankton, with some differences in diet composition between species (Pakhomov *et al.* 1996; Pusch *et al.* 2004; Saunders *et al.* 2015b).

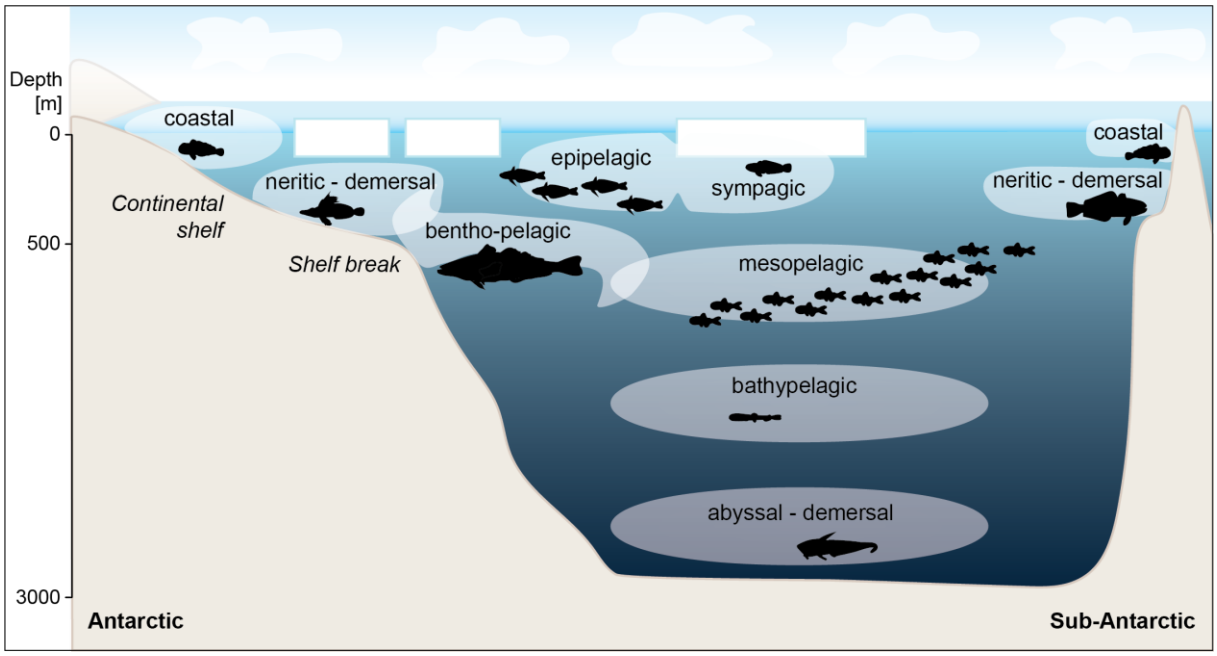



**Fig. 1.5**. Schematic depiction of important fish assemblages in the Southern Ocean. Silhouettes show common representatives of the respective assemblage (drawn by S. R. Maier), i.e. *Harpagifer antarcticus* (coastal, Antarctic), *Pagetopsis macropterus* (demersal, Antarctic), *Dissostichus mawsoni* (bentho-pelagic), *Pleuragramma antarctica* (epipelagic), *Trematomus borchgrevinki* (sympagic), *Electrona antarctica* (mesopelagic), *Bathylagus antarcticus* (bathypelagic), *Macrourus whitsoni* (abyssal demersal), *Harpagifer kerguelensis* (coastal, sub-Antarctic), and *Notothenia rossii* (demersal, sub-Antarctic). Depth indication is approximate only (note that the sea floor may be considerably deeper in parts of the Southern Ocean). Based on information found for example in Gon & Heemstra (1990); Eastman (1993, 2017); and Duhamel et al. (2014).

In turn, Antarctic foragers like seals, penguins and flying birds massively feed on mesopelagic fish (Connan *et al.* 2007; Cherel *et al.* 2009; Casaux *et al.* 2011). Most fishes in the twilight zone perform a diel vertical migration (Olivar *et al.* 2012). They ascend to shallower waters at night to feed on plankton and retreat into greater depth during daytime to evade predation (Duhamel *et al.* 2000). Bioluminescence is widespread and used for counter-illumination as well as communication (de Busserolles *et al.* 2013; de Busserolles & Marshall 2017; Gruber *et al.* 2019). The bioluminescent photophores of myctophids, the most common family of mesopelagic fishes, are a crucial morphological feature for species identification (Hulley 1990). However, myctophids are fragile and hauling onto research vessels from great depth often results in the loss of scales and other damage that leaves specimens practically unidentifiable. Otoliths and genetic tools can both help for rapid and accurate taxonomic classification in such cases (Bernal *et al.* 2014; Pappalardo *et al.* 2015). Intraspecific genetic diversity in myctophid fishes is largely unknown. Van de Putte *et al.* (2012a) documented genetic homogeneity in the Antarctic lanternfish (*Electrona antarctica*) using microsatellite markers. It remains to be investigated

how genetic diversity is geographically distributed in other myctophids species and whether all of them maintain high connectivity through their large population size.

It is difficult or likely impossible to weigh the ecological importance of specific fish assemblages. Clearly, however, mesopelagic fish match or even exceed the importance of notothenioids at least with respect to their role in the food web. More research has been conducted on notothenioids, which may be explained through, 1) the early discovery of antifreeze glycoproteins and hemoglobin loss, which both have relevance for highly applied research, 2) the evolutionary interest in adaptive radiations, 3) the fact that notothenioids are easier to sample and arguably more charismatic than small mesopelagic fish. However, mesopelagic fish are increasingly attracting global interest from fisheries for their potential use as a source of nutraceutical products or for fish meal to feed aquaculture species (St. John *et al.* 2016). Theis difference in attention from the scientific community regarding mesopelagics and notothenioids is reflected in our knowledge about genetic diversity. Many species of notothenioids have been investigated using population genetics, albeit often with microsatellites and not with the most recent genome-wide tools. Patterns of genetic local adaptation are therefore still unknown in notothenioids, although such patterns might be expected given their peculiar physiological and ecological adaptations (AFGP, buoyancy, feeding plasticity etc.). In mesopelagic fish, population genetics has very rarely been applied. In summary, both notothenioids and mesopelagic fish are very important parts of the Southern Ocean ichthyofauna and in both cases many knowledge gaps remain. Molecular methods provide much potential to address (some of) these knowledge gaps.

### 5. Research objectives

This thesis comprises four empiricial research chapters with distinct aims and hypotheses. The overall aim is to enhance our understanding of inter- and intraspecific genetic diversity and how this diversity is shaped or connected in a population context. First (**Chapter 2**, Fig. 1.6), the midwater fish community of the Southern Ocean is investigated using DNA barcoding. This important component of the Antarctic food web is still characterized by large knowledge gaps. How many mesopelagic species are commonly encountered in Antarctic waters? Can they be rapidly identified to species level using molecular methods? Are there indications of cryptic species among them? Did myctophids (the most common family of mesopelagic fish) repeatedly colonize the Southern Ocean or are they all descendants of one common ancestor that

conquered this habitat? And, what is the phylogeographic structure of the most common mesopelagic species? We hypothesize that COI barcoding is sufficient for accurate specimen identification in myctophids and that due to their abundance and entirely pelagic life style cryptic species and phylogeographic structure are absent. We use a a wide set of mesopelagic fish samples from various campaigns to test the accuracy of rapid molecular specimen identification and extend the barcode database for these fishes considerably. In addition, we use phylogenetic trees and haplotype networks in conjunction with species distribution patterns to identify potential cryptic species, phylogeographic structure and test how often myctophid species have adopted a Southern Ocean life style.

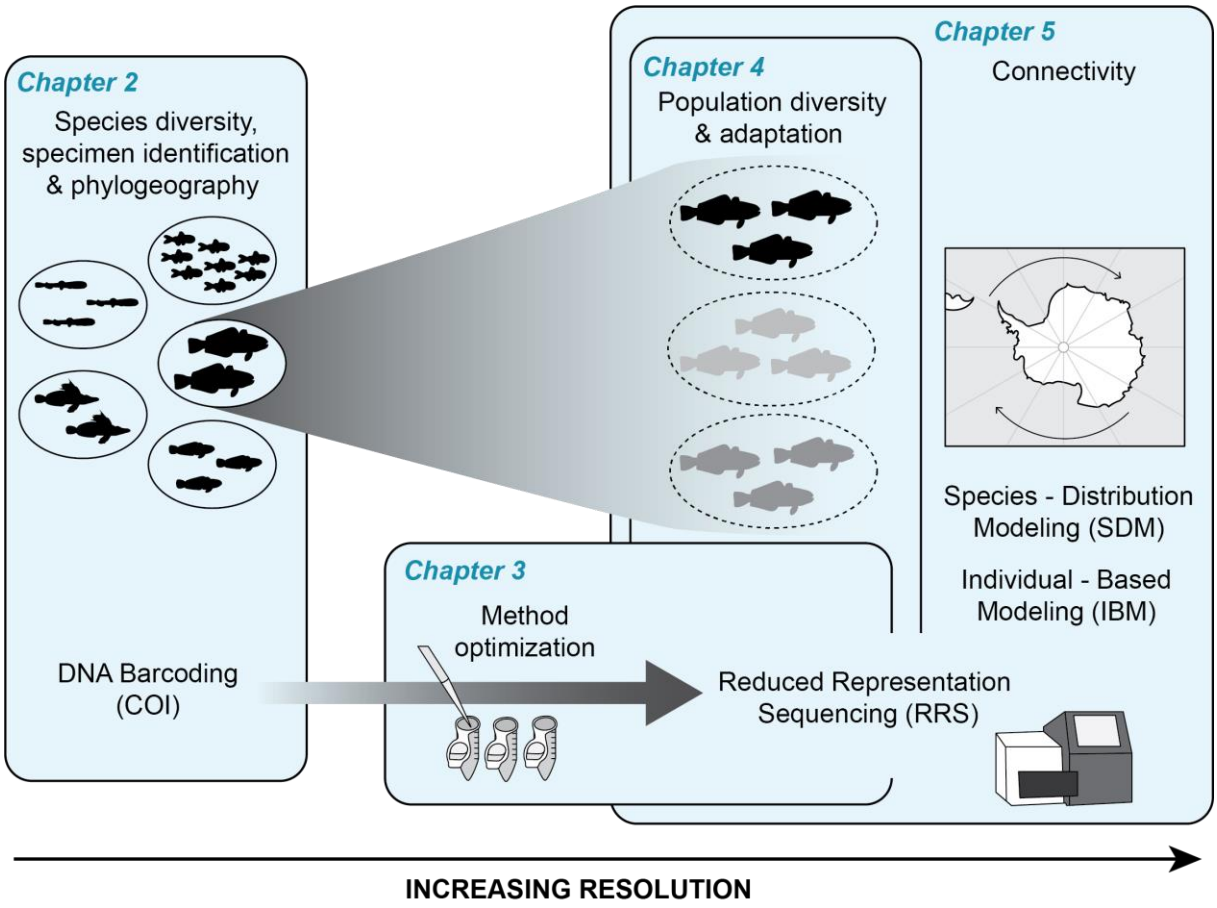



**Fig. 1.6.** Outline of the research chapters as presented in this thesis – from species to populations. Chapters 2, 4, and 5 are different case studies with increasing resolution and complexity. Chapter 3 is a technical part focused on optimization of the reduced representation sequencing approach. Chapters 3, 4, and 5 directly build on each other and include aspects of the respective preceding chapter (indicated as overlap of the light blue boxes).

The remaining empirical research chapters are designed to delve further into the complexities of intraspecific genetic diversity. Two case studies focus on two sister species of notothenioids, i.e. *Notothenia coriiceps* and *Notothenia rossii*. A prerequisite is, however, that the molecular

approach is fine-tuned for Antarctic non-model organisms. Molecular methods are rapidly developing, even in the course of a PhD many novel techniques and method refinements appear. We choose here to adapt reduced representation sequencing for Southern Ocean species as opposed to e.g. microsatellites. No published microsatellite loci are available for our target species. In addition, the application of advanced genomic methods promises to facilitate not only insights regarding genetic population structure and connectivity, but also allows us to screen for adaptive variation and correlate genomic diversity with potential drivers such as environmental variables. As outlined above, the application of RRS requires considerable expertise and optimization to avoid over- or under-sequencing (bad coverage) or the introduction of technical biases. The Belgian research projects vERSO and RECTO offered a unique opportunity to conduct RRS optimization for a wide range of Antarctic target organisms.

In **Chapter 3**, we test the applicability of several RRS protocols. Questions in this regard include, how much genomic information is available for which target organisms? Which restriction enzymes provide appropriate numbers of genomic fragments for RRS? How can the setup of RRS protocols (the number of individuals to be pooled, the size selection window and the bench protocol) be optimized for each target species? And, do test sequencing libraries provide the expected results? We use a combination of computational and empirical methods to answer these questions and thus provide a recipe to successfully apply state of the art genomic methods to most valuable samples from the remote Southern Ocean.

In **Chapter 4**, the dominant inshore fish species of much of the high-Antarctic is investigated. The bullhead notothen, *Notothenia coriiceps*, is the first Antarctic species for which a draft genome has become available (Shin *et al.* 2014). This fact and its ecological importance for coastal ecosystems renders it an ideal target species to apply RRS in a conservation genetics context. The spatial genetic structure of this species is unknown, as are levels of local adaptation. How much gene flow is present among *N. coriiceps* specimens from opposite sides of Antarctica? If there is some degree of genetic differentiation is this correlated to environmental variables? Are there signatures of recent selection pressure on the genome-wide markers? How should such patterns be incorporated into current conservation approaches? We hypothesize that populations on a large scale between East and West Antarctica are genetically differentiated, but not on a regional scale within the Scotia Sea. Due to this differentiation we expect signatures of local adaptation that correlate with environmental differences. In order to tackle these questions

we use ddRAD sequencing on 132 individuals from seven localities. Genetic structure is investigated using population clustering and individual-based statistical analyses. Genotype-environment association is tested for using redundancy analysis and genome scans for selection provide candidate loci for further investigation.

A sister species, namely *N. rossii*, provides the model for the second case study (**Chapter 5**). This species has been dramatically overfished in the 1970s. Previous population genetic and connectivity studies indicate that population differentiation is extremely low or absent (Young *et al.* 2015). These investigations, however, focused on the Scotia Sea and used microsatellite loci and hydrodynamic modelling. It remains unclear what levels of population genetic structure and connectivity *N. rossii* displays across large spatial scales. Is the species as patchily distributed as current occurrence data suggests? Are these patchy distributions well-connected through an extended planktonic egg and larval phase? Are genomic diversity and levels of effective population size reduced as a consequence of overexploitation? Here, we hypothesize again that large-scale differentiation is present, albeit high levels of gene flow. We expect reduced diversity due to the historic overfishing. In this chapter, we use additional methods to complement the population genomic approach. Species distribution modelling helps determine which localities are likely habitat for the species. A modified GBS protocol is used on more than 300 individuals to genotype thousands of SNPs. In addition, we emply individual-based modelling (IBM) based on Southern Ocean hydrodynamics to evaluate physical dispersal between the known and likely habitats. The combination of these methods provides a holistic picture of connectivity in the Southern Ocean.

Finally, in **Chapter 6**, I discuss the overall implications of the findings in this thesis. The molecular methods are critically examined and suggestions for improvements and alternative or additional future strategies are outlined. The importance of mesopelagic fish communities in the Southern Ocean and continued research to investigate their role in detail and especially under future changes is emphasized. I review the available knowledge of population genetic structure and connectivity in the Southern Ocean with a special focus on fishes. The patterns that were uncovered in previous research and during this thesis shed new light on the complex interaction between biological and physical factors in determining where and how an individual may live and reproduce in the Antarctic. It is hoped that these insights are taken into consideration in current management and conservation plans.

# CHAPTER 2: Diversity of Mesopelagic Fishes in the Southern Ocean – A Phylogeographic Perspective Using DNA Barcoding

*Note:* this manuscript has been published in *Frontiers in Ecology and Evolution*, 2018, doi: 10.3389/fevo.2018.00120.

A corrigendum has been published in *Frontiers in Ecology and Evolution*, 2018, doi: 10.3389/fevo.2018.00162. The updated article is reprinted here, with minor formatting changes.

## Diversity of Mesopelagic Fishes in the Southern Ocean – A Phylogeographic Perspective Using DNA Barcoding

Henrik Christiansen, Agnès Dettai, Franz M. Heindler, Martin A. Collins, Guy Duhamel, Mélyne Hautecoeur, Dirk Steinke, Filip A. M. Volckaert and Anton P. Van de Putte

**Abstract**

Small mesopelagic fish are ubiquitous in the ocean, representing an important trophic link between zooplankton and tertiary consumers such as larger fish, marine mammals and birds. Lanternfishes (Myctophidae) are common worldwide as well as in the Southern Ocean. However, only 17 of the approximately 250 myctophid species occur exclusively in sub-Antarctic or Antarctic waters. It is unclear whether they colonized these latitudes once and diversified from there, or whether multiple colonization events took place in which multiple ancestral phenotypes entered the Southern Ocean at various times. Phylogeographic patterns have been investigated for individual myctophid species, but so far no study has compared species across the Southern Ocean. Here, we present a dataset with previously unpublished cytochrome c oxidase I (*COI*; n = 299) and rhodopsin (*rhI*; n = 87) gene sequences from specimens collected at various locations in the Southern Ocean. Our data extend the DNA barcode library of Antarctic mesopelagic fish substantially. Combined morphological and molecular taxonomy lead to confident species level identification in 271 out of 299 cases, providing a robust reference dataset for specimen identification, independently of incomplete morphological characters. This is highly topical in light of prospective ecological metabarcoding studies. Unambiguous sequences were subsequently combined with publicly available sequences of the global DNA barcode library yielding a dataset of over 1000 individuals for phylogenetic and phylogeographic inference. Maximum likelihood trees were compared with results of recent studies and with the geographical origin of the samples. As expected for these markers, deep phylogenetic relationships remain partially unclear. However, *COI* offers unmatched sample and taxon coverage and our results at the subfamily to genus level concur to a large extent with other studies. Southern Ocean myctophids are from at least three distant subfamilies, suggesting that colonization has occurred repeatedly. Overall, spatial divergence of myctophids is rare, potentially due to their enormous abundance and the homogenizing force of ocean currents. However, we highlight potential (pseud-)cryptic or unrecognized species in *Gymnoscopelus bolini*, *Lampanyctus achirus*, and the non-myctophid genus *Bathylagus*.

## 1. Introduction

The mesopelagic fauna of the world's oceans is dominated by ubiquitous small filter feeding fish. These fishes likely represent a total biomass of up to 10 billion tons and include the perhaps most abundant vertebrate species on earth, *Cyclothone* sp. (Irigoien *et al.* 2014; Proud *et al.* 2018). Approximately 90 % of all small mesopelagic fishes belong to the bristlemouths (Gonostomatidae) and lanternfishes (Myctophidae). They form an important trophic link between primary consumers (predominantly mesozooplankton) and higher trophic levels such as large fish, squid, marine mammals, and birds (Smith *et al.* 2011a). Most small mesopelagic fish, which are generally found in the zone between approximately 200 and 1000 m, undertake a diurnal vertical migration following their prey into the epipelagic zone to feed at night (Isaacs *et al.* 1974). During daytime they retreat into the deep again, where they digest and excrete, which likely results in a substantial vertical carbon flux (Irigoien *et al.* 2014). Sonar reflections of their swim bladders cause the oceanic deep scattering layer (Barham 1966). However, despite their importance for marine food webs and organic carbon cycling, small mesopelagic fish are largely understudied.

The sub-Antarctic and Antarctic waters of the Southern Ocean are of particular importance both for global climate through ocean circulation and as a relatively pristine sanctuary for marine biodiversity. The Southern Ocean harbors considerable biodiversity (Brandt *et al.* 2007; Griffiths 2010), although species richness of fish is low compared to temperate and tropical seas with 322 currently recognized species from 19 families (Eastman 2005). Nevertheless, the Southern Ocean has been identified as an evolutionary hotspot, particularly because of the morphological and ecological diversity of species and a high degree of endemism, which amounts to 88 % in fish (Eastman 1991, 2005). It is believed that a key factor for such evolutionary uniqueness is the relative isolation of the Southern Ocean fauna, initiated approximately 24-25 mya by the formation of the Antarctic Circumpolar Current (ACC), a system of ocean currents flowing around Antarctica from West to East mostly between 50° to 60° South (Eastman 1991; Rintoul *et al.* 2001; Lyle *et al.* 2007). The ACC is the dominant hydrographic feature in the Southern Ocean (Orsi *et al.* 1995), and by providing a continuous, strong flow it forms a variable, but permanent boundary between Antarctic waters and water masses of lower latitudes (Rintoul *et al.* 2001). This greatly hampers any possible north-south (or south-north) migration of organisms. However, Saunders *et al.* (2017) recently showed that lanternfish biomass in the Scotia and Weddell Sea must be supported by mass immigration from lower latitudes. The

mesopelagic zone in temperate regions is generally strongly stratified and includes a distinct thermocline. Temperatures tend to range between 2° and 15° C with the upper layer being warmer and well mixed, followed by a sharp decrease in temperature at the thermocline (at around 50 – 400 m) and a gradual decrease of temperature with increasing depth. In contrast, the Southern Ocean is relatively well-mixed with temperatures ranging between -0.5 and 2.0° C (Ikeda 1988). Temperatures at 1000 m depth are therefore similar in temperate and Antarctic regions, whereas temperatures of upper water masses are very different.

Although less abundant than at lower latitudes, mesopelagic fish are still numerous in Antarctic and sub-Antarctic waters and represent a major part of the biomass (Eastman 1993). In terms of species richness, abundance, and biomass, the mesopelagic zone there is dominated by lanternfishes (Myctophidae) (Donnelly *et al.* 1990; Kock 1992). Myctophids are common in oceanic waters north of the Antarctic Slope Front (ASF; near the Antarctic continental shelf break), where they act as largely opportunistic mesozooplankton feeders with some interspecific dietary variation (Pakhomov *et al.* 1996; Pusch *et al.* 2004; Connan *et al.* 2010; Saunders *et al.* 2014, 2015b). Charismatic Antarctic top predators such as king penguin (Cherel *et al.* 2009), Antarctic fur seals (Casaux *et al.* 2011; Santora 2013), and seabirds (Connan *et al.* 2007) heavily rely on myctophids as a food source. Despite their small size some myctophid species (*Gymnoscopelus* spp., *Electrona carlsbergi*) were commercially exploited in the 1980s (Hulley 1990; Kock 1992). Of the approximately 240 myctophid species recognized worldwide, 68 have been recorded south of the Sub-Tropical Front, 14 of which have a sub-Antarctic (as in between sub-Antarctic Front and Polar Front), five a broadly Antarctic, and two an Antarctic distribution pattern (Duhamel *et al.* 2014). The remaining species exhibit widespread distribution patterns and only sporadically occur in the Southern Ocean.

Ecological studies are dependent on accurate biological identification to a level of taxonomic resolution appropriate for the study goal (Tautz *et al.* 2003). In myctophids, photophore patterns are mainly used to distinguish species. However, accurate identification can be impeded, because myctophids tend to lose scales during capture and are easily damaged in the net. The identification of early life stages may also be challenging. DNA barcoding is a molecular technique that uses the mitochondrial cytochrome c oxidase I gene (*COI*) as a genetic marker to provide biological identifications (Hebert *et al.* 2003). The system is now widely accepted and many taxa, including teleosts, have been successfully integrated in barcoding initiatives and data

systems (Ratnasingham & Hebert 2007, 2013; Ward *et al.* 2009). A sufficiently complete reference dataset of DNA barcodes thus enables fast and efficient verifications for morphologically identified specimens as long as *COI* exhibits levels of interspecific divergence that are higher than the intraspecific divergence of a given group. Furthermore, it can assist with the discovery of misidentified specimens, cryptic or simply not yet identified new species, help settle synonymies, or hint at intraspecific genetic structuring (Hajibabaei *et al.* 2007; see Bucklin *et al.* 2011 for an extended overview of marine barcoding applications). The latter can be used in phylogeography, a discipline concerned with phylogenetic relatedness and connectivity of species or populations with respect to geographic distribution. Genetic distance, derived from markers such as *COI*, is used to study the historical processes that may be responsible for the contemporary geographic distribution of individuals. In order to increase robustness of results derived from *COI* data, it can be useful to include an additional genetic marker, particularly nuclear and thus biparentally inherited (Cao *et al.* 2016; Thiel & Knebelsberger 2016). Rhodopsin belongs to a family of genes, the so-called G-protein-coupled receptors, that are involved in translating external information (e.g. light, molecules) into internal signals that can be processed by organisms. Rhodopsin encodes a protein that is involved in photoreception (Palczewski *et al.* 2000). It occurs on the rod cells and is extremely light sensitive enabling vision under low-light conditions (Yokoyama & Yokoyama 1996). In Actinopterygians, the rhodopsin gene generally occurs in two copies, homologous to other vertebrates. One copy, *rh1*, is an intronless retrogene that does not recombine anymore with other opsins and has proven useful for fish identification and phylogeny (Fitzgibbon *et al.* 1995; Chen *et al.* 2003; Lin *et al.* 2017; Morrow *et al.* 2017).

Less than a decade ago Grant & Linse (2009) recognized a lack of Antarctic barcoding studies. The Census of Antarctic Marine Life (CAML) set an explicit focus on DNA barcoding, resulting in many studies making significant progress in addressing this gap (Schiaparelli *et al.* 2013 and references therein). Over the past few years, Antarctic barcoding demonstrated the usefulness of *COI* sequencing e.g. to identify *Trematomus* fishes (Lautredou *et al.* 2010), and showcased the presence of cryptic species in various groups, e.g. pycnogonids (Krabbe *et al.* 2010), amphipods (Havermans *et al.* 2011), octopuses (Allcock *et al.* 2011), skates (Smith *et al.* 2008), and grenadier fishes (Smith *et al.* 2011b; McMillan *et al.* 2012). These examples clearly demonstrate that despite the fact that Antarctic biodiversity is still underexplored (Griffiths 2010; Grant *et al.* 2010), molecular techniques can enhance our understanding of contemporary diversity patterns and the processes that shaped these (Allcock & Strugnell 2012). Fish communities of the Southern

Ocean have been studied using DNA barcoding, but these studies primarily focused on benthic fish in the Scotia Sea (Rock *et al.* 2008), Dumont D'Urville Sea (Dettai *et al.* 2011), and Ross Sea (Smith *et al.* 2012). Phylogeographic patterns of myctophids have been investigated for a few species (e. g. *Electrona antarctica*, Van de Putte *et al.* 2012a), but to date no study has compared species across the Southern Ocean. Here, we present an extensive DNA barcoding approach to investigate the ecologically relevant community of Antarctic mesopelagic fish.

Our objectives were, (1) to extend the DNA barcode library of Antarctic mesopelagic fish, (2) to assess the success of specimen identification using this system, (3) to discover potential mismatches between taxonomy and genetic identification, (4) to compare our Antarctic myctophid phylogenetic data with recent myctophid phylogenies, and (5) to investigate phylogeographic patterns of common Antarctic myctophids. To achieve these objectives, we used a large-scale dataset of mesopelagic Antarctic fish, covering over 1000 specimens from a circum-Antarctic sampling range. This dataset includes 386 new samples and combines these with publicly available sequences found on the Barcode of Life Data Systems, BOLD (Ratnasingham & Hebert 2007). We focused on the analysis of *COI* but extended our results by incorporating an additional nuclear marker (*rhI*). Thus, a comprehensive picture on the inter- and intraspecific diversity of mesopelagic fishes occurring in the Southern Ocean was drawn.

## 2. Material and Methods

### 2.1 Sampling and identification

Mesopelagic fish were captured in the Southern Ocean during various expeditions. The sampling effort comprised cruise 200 with RV *James Clark Ross* (see Collins *et al.* 2012), cruises PS65 and PS69 with RV *Polarstern*, BROKE-West with RV *Aurora Australis* and additional Atlantic samples collected with RV *G.O. Sars* (BOLD project FISCO); the POKER sampling campaign 2010 off Kerguelen and additional Pacific samples from the JAMSTEC survey with RV *Hakuho Maru* (BOLD project MYCSO); cruises JR100 (Collins *et al.* 2008), JR161 and JR177 (Collins *et al.* 2012b) with RV *James Clark Ross* and few specimens from commercial vessels (BOLD project BASMF); and finally 23 myctophid specimens collected off South Africa (BOLD project DSSAU). Samples from the Atlantic and Pacific Oceans were included to provide an outgroup framework. Overall, these sampling efforts yielded a total of 386 previously unpublished specimens (Table 2.1). All specimens were identified morphologically aboard the research vessels or, in absence of

a taxonomic expert, immediately frozen or preserved whole in high-grade ethanol or formalin and identified at the respective institutions. Muscle tissue or fin biopsies were excised using sterile tools and stored in ethanol. In most cases identifications were carried out to species level and only in some instances to family or genus level (juvenile/larval or severely damaged specimens). The majority of specimens are stored at the Muséum National d'Histoire Naturelle (MNHN, Paris), KU Leuven (Belgium), the British Antarctic Survey (BAS, Cambridge) or the Natural History Museum (NHM, London), and the South African Institute for Aquatic Biodiversity (Grahamstown), respectively. Detailed collection data of all specimens are shown in Supplementary Table S2.1.

**Table 2.1.** Overview of fish specimens and species and the respective numbers of barcode compliant (non-barcode compliant in brackets) DNA sequences that were successfully obtained for the cytochrome c oxidase I (*COI*) and rhodopsin (*rh1*) gene. Project name abbreviations as used in the Barcode of Life Data Systems, BOLD (Ratnasingham & Hebert 2007). Some samples with unclear identification status were excluded from tree building.

| **Project** | **Specimens** | **Species** | ***COI* sequences** | ***Rh1* sequences** |
|---|---|---|---|---|
| FISCO | 190 | 33 | 144 | 35 |
| MYCSO | 99 | 17 | 62 | 52 |
| BASMF | 73 | 28 | 69 | 0 |
| DSSAU | 24 | 8 | 24 | 0 |
| Total (previously unpublished) | 386 | min. 57 | 299 | 87 |
| Mined from BOLD/GenBank | | | 762 | |
| Outgroup, mined from BOLD | | | 9 | 5 |
| Total (used for tree building) | | | 1073 | 90 |

### 2.2 DNA extraction, PCR and sequencing

DNA was extracted from the tissue sample using a modified standard salting-out protocol (Cruz *et al.* 2017). Extracts from the datasets FISCO, BASMF, and DSSAU were subsequently shipped to the University of Guelph, Canada, for *COI* amplification and sequencing following protocols described in Steinke & Hanner (2011). Primers used for *COI* were the cocktails C FishFlt1-C FishR1t1 as described in Ivanova *et al.* (2007). Rhodopsin gene fragments (*rh1*) were amplified using Rh193-5'CNTATGAATAYCCTCAGTACTACC3' and Rh1039r-5' TGCTTGTTCATGCAGATGTAGA3' primers (Chen *et al.* 2003). Amplification was conducted in 25 µl volume with 0.2 mM dNTP's, 2.5 mM MgO2, 20 µM primer mix, and conventional PCR buffer and Taq polymerase. PCR conditions were 2 min initial denaturation at 94 °C, followed by 30-40 cycles of 20 s at 94 °C, 30 s at 50-60 °C, 70 s at 72 °C, and a final 3 min elongation at 72 °C. The MYCSO *COI* and *rh1* dataset was generated at MNHN (France) following Dettai *et al.*

(2011) for extraction, PCR, and sequencing using standard automatic capillary sequencers. Additional *rh1* sequences were generated at KU Leuven (Belgium) for the FISCO samples following the same protocol.

### 2.3 Dataset augmentation and trimming and phylogenetic statistics

We were able to retrieve 299 *COI* and 87 *rh1* sequences of 386 specimens from 16 locations (Table 2.1, Fig. 2.1). These sequences were deposited in the BOLD datasets: "Fishes of the Scotia Sea" (FISCO), "Myctophids of the Southern Ocean" (MYCSO), "BASMF", and "DSSAU". To increase taxonomic and spatio-temporal coverage our unpublished dataset was extended with publicly available data from BOLD/GenBank including some previously published Antarctic *COI* barcode sets: Rock *et al.* 2008 (samples from the Scotia Sea, South Orkney Islands, and Elephant Island), Mabragaña *et al.* 2011 (Argentina), Smith *et al.* 2012 (Ross Sea, Heard and McDonald Islands and more). Altogether, sequences cover an unprecedented area of the Southern Ocean, although many regions remain underrepresented (Fig. 2.1). This can be attributed to the enormous logistic and financial challenges posed by Antarctic exploration. Species identity of all previously unpublished specimens from the Southern Ocean dataset was confirmed using internal tools of BOLD (Ratnasingham & Hebert 2007) using all available *COI* sequences > 500 bp with species level identification on 15$^{th}$ December 2017. If molecular and morphological identification did not match, a second morphological examination was performed and only specimens that were attributed to the species identified by BOLD were kept as such. In addition, in case of any doubt, e.g. the absence of crucial morphological characters, the specimen was excluded from further analysis.

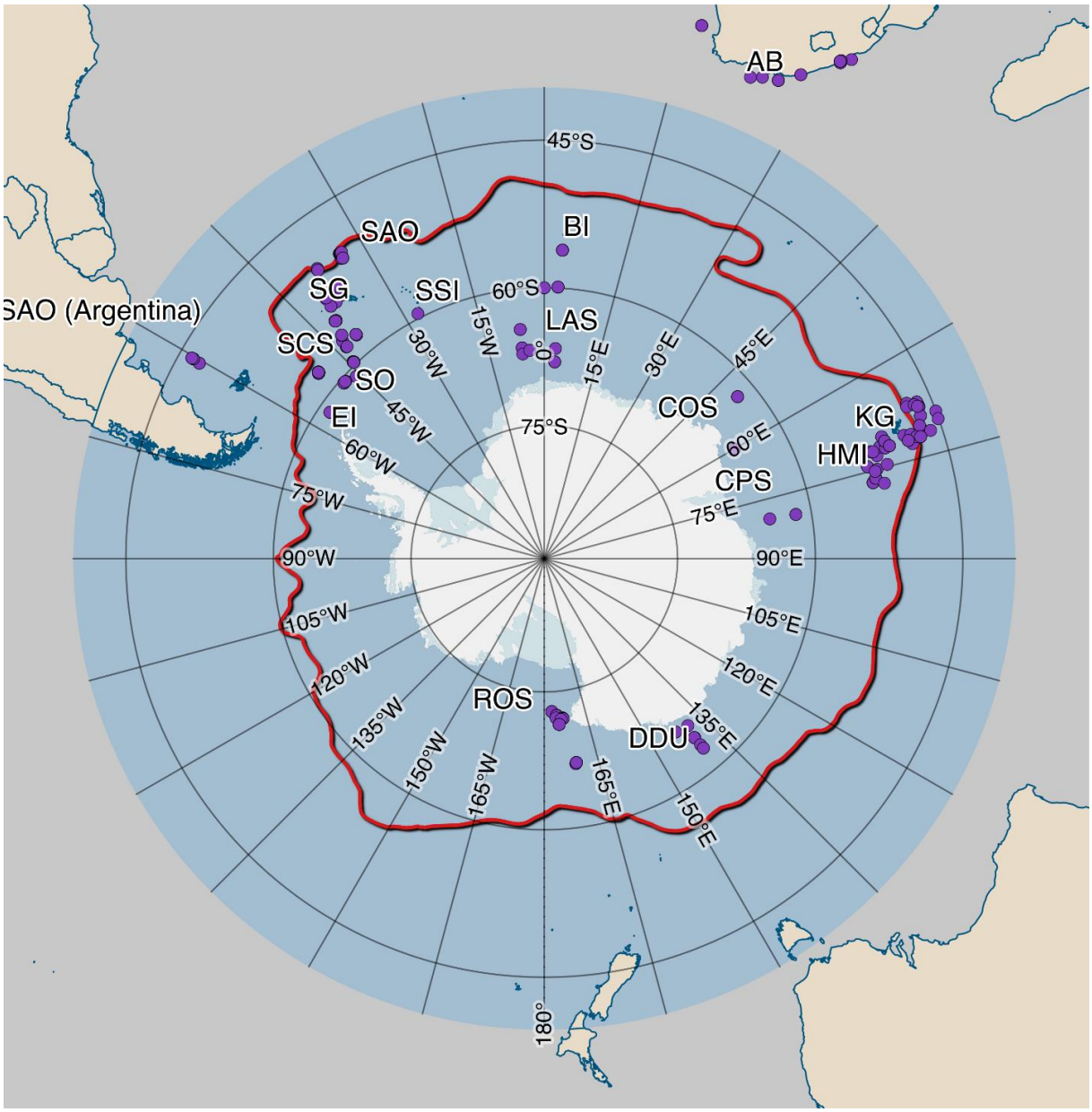



**Fig. 2.1.** Locations of mesopelagic fish sampled in the Southern Ocean and subsequently used for DNA barcoding. The bold, red line marks the approximate position of the Antarctic Polar Front. Location codes are (approx. clockwise from West to East Antarctic): SAO: South Atlantic Ocean; SCS: Scotia Sea; EI: Elephant Island; SO: South Orkney Islands; SG: South Georgia Islands; SSI: South Sandwich Islands; BI: Bouvet Island; LAS: Lazarev Sea; AB: Agulhas Bank (off South Africa); COS: Cosmonauts Sea; CPS: Cooperation Sea; KG=KI: Kerguelen Islands; HMI=HM: Heard and McDonald Islands; DDU: Dumont d'Urville Sea; ROS: Ross Sea.

Sequences from five *Synodotus binotatus* (two-spot lizardfish) specimens curated on BOLD were included for both *COI* and *rhl*. *Synodus binotatus* is an aulopiform fish, the order with closest common ancestor to myctophiform fishes (Betancur-R *et al.* 2017). These sequences were used as outgroup for phylogenetic tree rooting. Two COI sequences each of *Neoscopelus macrolepidotus* and *N. microchir* (Neoscopelidae (blackchins), the other family in

Myctophiformes, next to Myctophidae) were also included. No *rhl* sequences of Neoscopelidae were available. Three different datasets were used for phylogenetic reconstruction: (1) all available *COI* sequences (new sequences, published myctophid sequences, and outgroup; total N = 1073); (2) all available *rhl* sequences (new sequences and outgroup; N = 90); and (3) a concatenated dataset consisting of specimens from (1) and (2) for which good quality sequences of *COI* and *rhl* were available (N = 68, including outgroup). Sequences were aligned via MUSCLE (Edgar 2004) within Geneious v.8.1.5 (Biomatters Ltd) using a maximum of eight iterations and standard preset values. Tree building was performed in R v3.1.2 (R Core Team, 2016) using the packages 'ape' (Paradis *et al.* 2004; Popescu *et al.* 2012) and 'phangorn' (Schliep 2011). Kimura's two-parameter substitution model (Kimura 1980) is commonly used in DNA barcoding studies to construct genetic distance matrices, although the fit might be poor (Collins *et al.* 2012a). We decided to assess a variety of nucleotide substitution models with phangorn's 'modelTest' function. The most appropriate model for all three datasets as determined by Akaike's information criterion (AIC) was the general time reversible model with gamma distributed rate variation among sites and a proportion of invariable sites ('GTR+G+I'). This substitution model was used as initial fit and for subsequent maximum likelihood (ML) optimization using a stochastic algorithm instead of nearest-neighbor-interchange to avoid local maxima. Edge support was evaluated with 10,000 randomly seeded bootstraps. Consensus trees were created in Geneious with a support threshold of 70% (Hillis & Bull 1993) and were subsequently manually checked and annotated using MEGA7 v7.0.26 (Kumar *et al.* 2016). *COI* haplotype networks were created by median joining (Bandelt *et al.* 1999) in popART v.1.7 (Leigh & Bryant 2015).

## 3. Results

### 3.1 Extension of the DNA barcode library

Mesopelagic fish of various research expeditions were identified, catalogued, and when possible sequenced for *COI* and/or *rhl*. In some cases (not listed here) sequencing was impossible due to DNA degradation or amplification failure. Overall, 297 reliable *COI* sequences were added to BOLD after rigorous validation and exclusion of doubtful samples (two samples excluded, see below). Some of these sequences belong to larval, juvenile or incidentally caught fishes whose adult stages are generally not mesopelagic (notothenioids, grenadiers), leaving 264 Antarctic meso- or bathypelagic specimens from 35 different species with validated identification and *COI* sequences – a biogeographic assemblage that was previously almost absent in the database. The

worldwide database for myctophids was extended by 23.7 % to a total of 1021 sequences. Furthermore, 87 validated *rh1* sequences were added to BOLD. All samples and sequence IDs and associated metadata can be found in Supplementary Table S2.1.

### 3.2 Specimen identification

All 299 previously unpublished *COI* sequences were identified using BOLD data and tools (using only species level barcode records). In some instances, this revealed most likely misidentified or mislabeled sequences in BOLD. If only one *COI* sequence of a given species on BOLD was misidentified, the identification engine will declare there was no species level match. For instance, at time of study BOLD contained 68 sequences with the Barcode Index Number (BIN; Ratnasingham & Hebert 2013) corresponding to *Electrona antarctica* (Antarctic lanternfish; BOLD: AAB3737), all with low pairwise distance (average: 0.08 %, maximum: 0.78 %). The nearest neighbor of this BIN is *Symbolophorus veranyi* (large-scale lanternfish; BOLD: AAC4870; pairwise distance: 2.39 %). Yet, one of the specimens in AAB3737 (*E. antarctica*) is labelled *Nannobrachium achirus* – a more distant species well represented by 27 other sequences in the clearly distinct BIN BOLD: AAB3778. In these cases, we highlighted the likely misidentified or mislabeled sequences present in the database (Table S2.2). Five such sequences were accessible to us and are now flagged in BOLD to avoid future misidentifications when using the database.

Using this procedure, morphological and molecular identification showed high levels of congruence (97.65 %). In eleven cases the morphological identification could be confidently improved using the BOLD identification engine, as specimens were attributed at least to genus level, which was confirmed through the placement of the specimen in the phylogenetic trees. In seven further cases a mismatch between morphological and molecular identification was detected. After detailed inspection, the identification was revised (Table 2.2). Five specimens with unclear species or genus level identification could not be matched to any available *COI* sequence in BOLD. However, they could be attributed to genus level based on the phylogenetic tree (highlighted in bold italics in Fig. 2.2). Two specimens were excluded from further analysis, because *COI* and *rh1* gave conflicting results, likely indicating contamination or similar error in the laboratory. Lastly, nine specimens had matching morphological and molecular identification, although only uni-directional sequences were obtained. These were included in phylogenetic analyses, but flagged as non-barcode compliant on BOLD.

**Table 2.2.** Antarctic mesopelagic fish specimens where mismatch between morphological and molecular identification led to re-identification after detailed inspection.

| Specimen ID | Sequence ID | Initial identification | Molecular identification | Final identification |
|---|---|---|---|---|
| GYP#1 | BASMC030-09 | *Gymnoscopelus piabilis* | *G. nicholsi* | *G. nicholsi* |
| GYF#4 | BASMC039-09 | *G. fraseri* | *G. bolini* | *G. bolini* |
| PRL#1 | BASMC057-09 | *Porotomyctophum luciferum* | *P. bolini* | *P. bolini* |
| KUL_Gym_bra_PS69_FF_1524 | FISCO075-10 | *G. braueri* | *G. opisthopterus* | *G. opisthopterus* |
| KUL_Hyg_hyg_27677 | FISCO108-10 | *Hygophum* sp. | *Benthosema glaciale* | *B. glaciale* |
| KUL_Lam_mac_27733 | FISCO112-10 | *Lampanyctus* sp. | *Lobianchia dofleini* | *L. dofleini* |
| KUL_Lob_dof_27761 | FISCO114-10 | *Lobianchia dofleini* | *Lobianchia gemellarii* | *L. gemellarii* |

### 3.3 Phylogeny

The curated datasets were used to produce three phylogenetic ML consensus trees: one for *COI*, one for *rh1*, and one for both combined. Sequence alignment was not problematic, as these are coding sequences without gaps. In each case 'GTR+G+I' was identified as the most appropriate nucleotide substitution model. Clades with bootstrap support below 70 % after consensus tree building were collapsed, i.e., these splits were not retained or displayed in the figures.

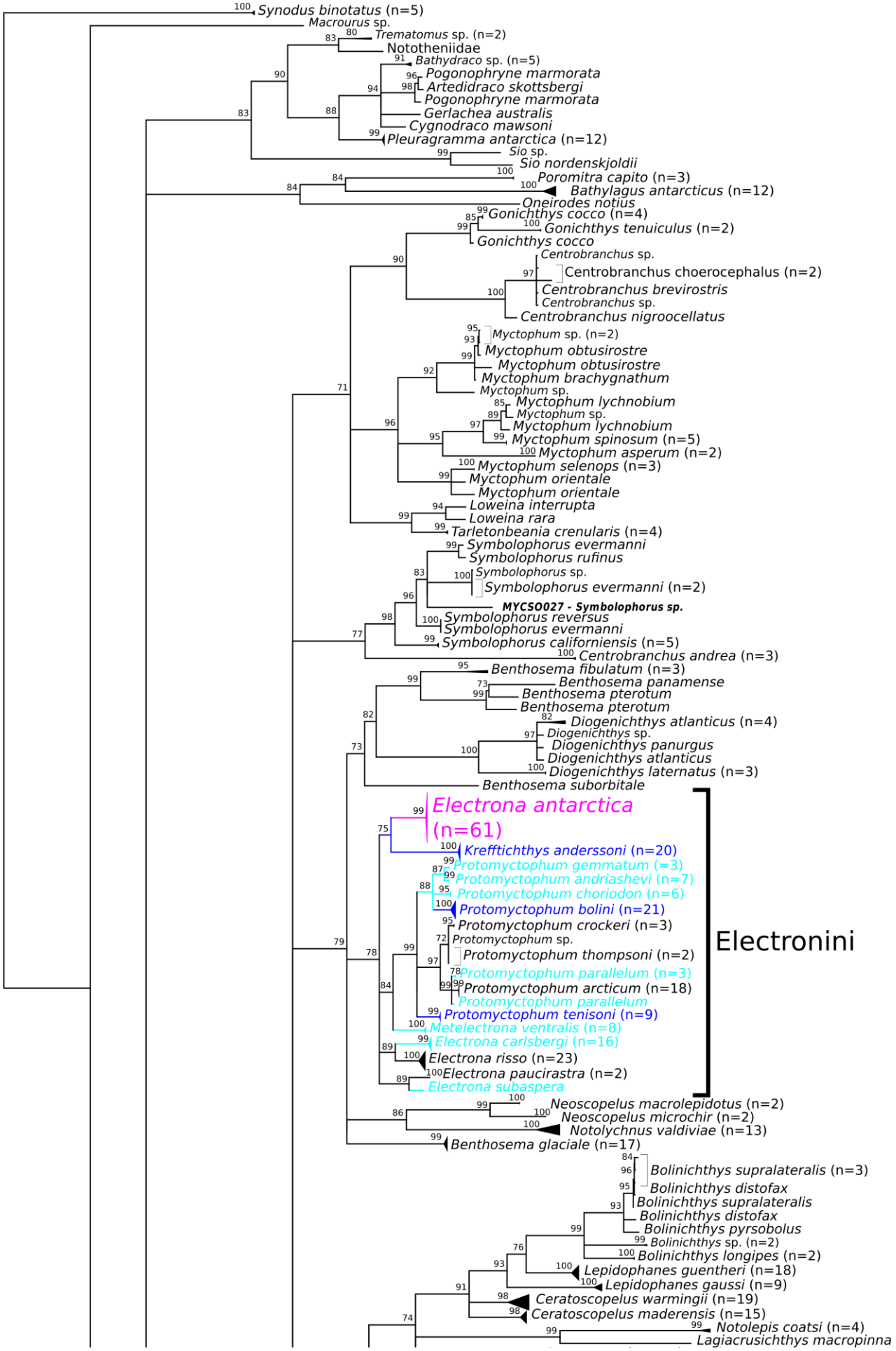
#### 3.3.1 The cytochrome c oxidase I gene

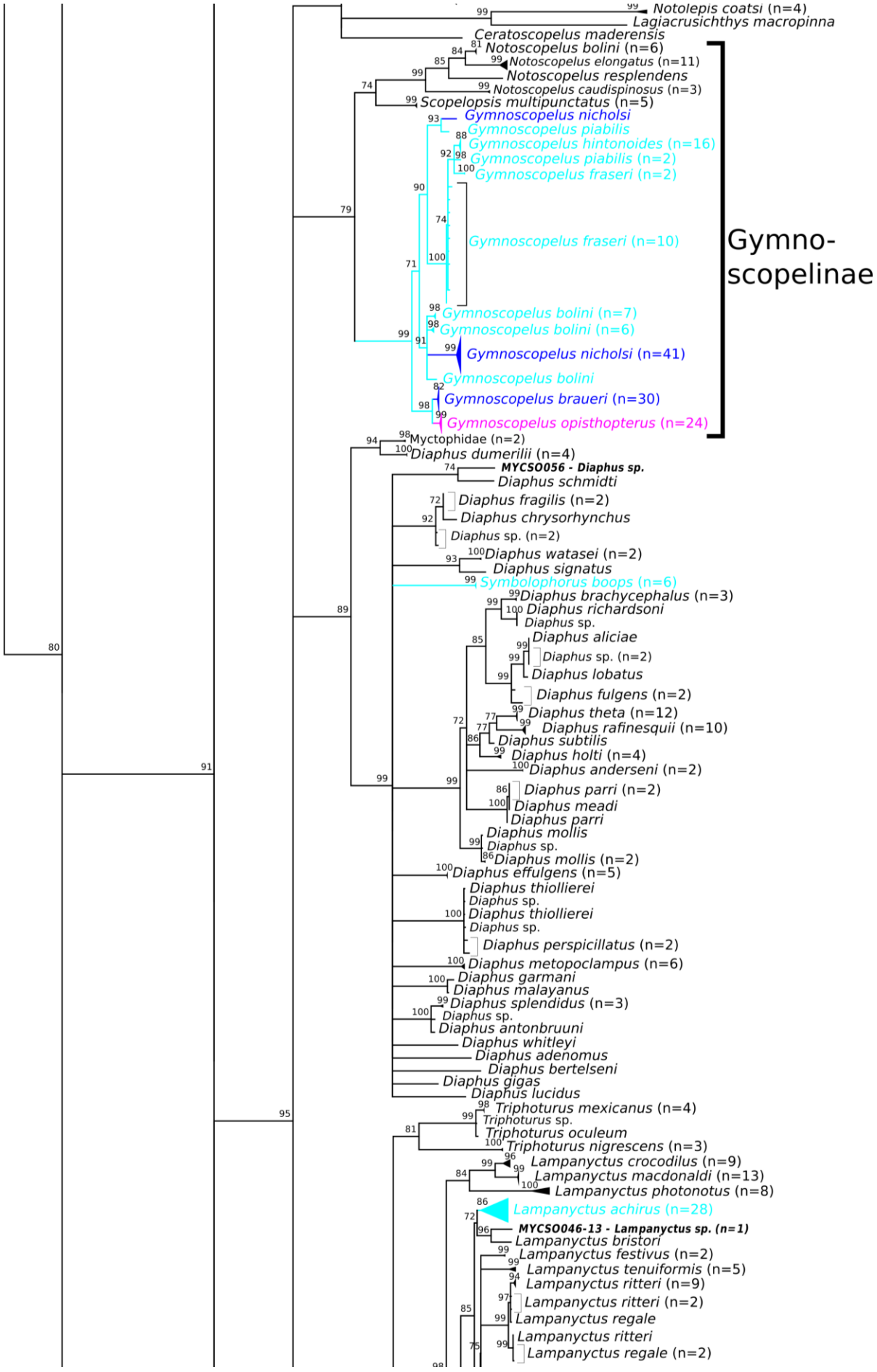
The dataset includes 1073 sequences of 539 bp length with 337 variable sites. Myctophidae are not resolved as monophyletic, because the aulopiform species *Notolepis coatsi* (Antarctic jonasfish) and *Lagiacrusichthys macropinna* (previously *Benthalbella macropinna*, see Davis, 2015) are placed within a Lampanyctinae clade and the neoscopelid species *Neoscopelus macrolepidotus* and *N. microchir* are placed next to *Notolychnus valdiviae* (topside lanternfish; Fig. 2.2). Other outgroup taxa are placed outside of Myctophidae, but their exact position is not well resolved (i.e. often < 70 % bootstrap support and thus displayed as polytomic). Within Myctophidae the tribe Electronini (*sensu* Paxton 1972) and subfamily Gymnoscopelinae (*sensu* Martin *et al.* 2018) are monophyletic with medium bootstrap support (BS = 78 and 79 %, respectively). Diaphinae (*sensu* Martin *et al.* 2018) is monophyletic as well (BS = 89 %), except for the inclusion of *Symbolophorus boops* (bogue lanternfish). The placement of these three subfamilies/tribes and the remaining genera within the Myctophidae is less clear, with bootstrap

support at times below the applied cut-off threshold. Myctophid species with their main distribution range in sub-Antarctic or Antarctic waters all belong to the three groups mentioned above, except for *S. boops* and *Lampanyctus achirus* (previously *Nannobrachium achirus*, see Martin *et al.* 2018), a bathypelagic species that was placed within a clade of *Lampanyctus* spp., sister group of *Parvilux ingens*. As the focus of this study is on (sub-)Antarctic mesopelagics, further description and discussion is restricted to these species and their position in the phylogenetic trees.

Within Electronini, the position of the genera *Metelectrona*, *Electrona*, *Krefftichthys*, and *Protomyctophum* is unclear. *Electrona antarctica* forms a clade with *K. anderssoni* (BS = 75 %). *Metelectrona ventralis* (flaccid lanternfish) is resolved as sister group (BS = 84 %) to *Protomyctophum*, the only monophyletic genus (BS = 99 %). *Electrona subaspera* (rough lanternfish) and *E. paucirastra* (belted lanternfish), and *E. risso* (electric lanternfish) and *E. carlsbergi* (electron subantarctic lanternfish), respectively, appear to be closely related (BS = 89 % in both cases). Within *Protomyctophum*, the split into the subgenera *Hierops* and *Protomyctophum* is supported except for *P. tenisoni*, which is placed next to these subgenera (BS = 99 %). *Hierops* contains *P. parallelum* (parallel lanternfish), *P. thompsoni* (bigeye lanternfish), *P. arcticum* (Arctic telescope), and *P. crockeri* (California flashlightfish) (BS = 97 %) and the subgenus *Protomyctophum* contains *P. bolini*, *P. choriodon*, *P. andriashevi*, and *P. gemmatum* (BS = 88 %).

In Gymnoscopelinae (*sensu* Martin *et al.* 2018), *Gymnoscopelus* is monophyletic (BS = 99 %) and sister group to a clade with medium support (BS = 74 %) containing *Scopelopsis multipunctatus* and the also monophyletic *Notoscopelus* (BS = 99 %). Within *Gymnoscopelus*, two single specimens identified as *G. piabilis* (Southern blacktip lanternfish) and *G. nicholsi* (Nichol's lanternfish) form their own clade apart from all other specimens with the same identification (BS = 93 %). Sister group to these is a clade comprised of *G. hintonoides* (false-midas lanternfish), *G. piabilis*, and *G. fraseri* (BS = 100 %), in which *G. hintonoides* and *G. piabilis* are resolved as monophyletic (BS = 88 and 98 %), but *G. fraseri* not. *Gymnoscopelus bolini* and *G. nicholsi* form a clade (BS = 91 %), in which *G. nicholsi* is placed as one group (BS = 99 %), but *G. bolini* as three. Lastly, *G. braueri* and *G. opisthopterus* are monophyletic sister group to all others (BS = 98 %) and also monophyletic within each species (BS = 82 % for *G. braueri* and 99 % for *G. opisthopterus*).


Synodus binotatus (n=5)
Macrourus sp.
Trematomus sp. (n=2)
Nototheniidae
Bathydraco sp. (n=5)
Pogonophryne marmorata
Artedidraco skottsbergi
Pogonophryne marmorata
Gerlachea australis
Cygnodraco mawsoni
Pleuragramma antarctica (n=12)
Sio sp.
Sio nordenskjoldii
Poromitra capito (n=3)
Bathylagus antarcticus (n=12)
Oneirodes notius
Gonichthys cocco (n=4)
Gonichthys tenuiculus (n=2)
Gonichthys cocco
Centrobranchus sp.
Centrobranchus choerocephalus (n=2)
Centrobranchus brevirostris
Centrobranchus sp.
Centrobranchus nigroocellatus
Myctophum sp. (n=2)
Myctophum obtusirostre
Myctophum obtusirostre
Myctophum brachygnathum
Myctophum sp.
Myctophum lychnobium
Myctophum sp.
Myctophum lychnobium
Myctophum spinosum (n=5)
Myctophum asperum (n=2)
Myctophum selenops (n=3)
Myctophum orientale
Myctophum orientale
Loweina interrupta
Loweina rara
Tarletonbeania crenularis (n=4)
Symbolophorus evermanni
Symbolophorus rufinus
Symbolophorus sp.
Symbolophorus evermanni (n=2)
MYCSO027 - Symbolophorus sp.
Symbolophorus reversus
Symbolophorus evermanni
Symbolophorus californiensis (n=5)
Centrobranchus andrea (n=3)
Benthosema fibulatum (n=3)
Benthosema panamense
Benthosema pterotum
Benthosema pterotum
Diogenichthys atlanticus (n=4)
Diogenichthys sp.
Diogenichthys panurgus
Diogenichthys atlanticus
Diogenichthys laternatus (n=3)
Benthosema suborbitale
Electrona antarctica (n=61)
Krefftichthys anderssoni (n=20)
Protomyctophum gemmatum (=3)
Protomyctophum andriashevi (n=7)
Protomyctophum choriodon (n=6)
Protomyctophum bolini (n=21)
Protomyctophum crockeri (n=3)
Protomyctophum sp.
Protomyctophum thompsoni (n=2)
Protomyctophum parallelum (n=3)
Protomyctophum arcticum (n=18)
Protomyctophum parallelum
Protomyctophum tenisoni (n=9)
Metelectrona ventralis (n=8)
Electrona carlsbergi (n=16)
Electrona risso (n=23)
Electrona paucirastra (n=2)
Electrona subaspera
Electronini
Neoscopelus macrolepidotus (n=2)
Neoscopelus microchir (n=2)
Notolychnus valdiviae (n=13)
Benthosema glaciale (n=17)
Bolinichthys supralateralis (n=3)
Bolinichthys distofax
Bolinichthys supralateralis
Bolinichthys distofax
Bolinichthys pyrsobolus
Bolinichthys sp. (n=2)
Bolinichthys longipes (n=2)
Lepidophanes guentheri (n=18)
Lepidophanes gaussi (n=9)
Ceratoscopelus warmingii (n=19)
Ceratoscopelus maderensis (n=15)
Notolepis coatsi (n=4)
Lagiacrusichthys macropinna


Notolepis coatsi (n=4)
Lagiacrusichthys macropinna
Ceratoscopelus maderensis
Notoscopelus bolini (n=6)
Notoscopelus elongatus (n=11)
Notoscopelus resplendens
Notoscopelus caudispinosus (n=3)
Scopelopsis multipunctatus (n=5)
Gymnoscopelus nicholsi
Gymnoscopelus piabilis
Gymnoscopelus hintonoides (n=16)
Gymnoscopelus piabilis (n=2)
Gymnoscopelus fraseri (n=2)
Gymnoscopelus fraseri (n=10)
Gymnoscopelus bolini (n=7)
Gymnoscopelus bolini (n=6)
Gymnoscopelus nicholsi (n=41)
Gymnoscopelus bolini
Gymnoscopelus braueri (n=30)
Gymnoscopelus opisthopterus (n=24)
Gymno-
scopelinae
Myctophidae (n=2)
Diaphus dumerilii (n=4)
MYCSO056 - Diaphus sp.
Diaphus schmidti
Diaphus fragilis (n=2)
Diaphus chrysorhynchus
Diaphus sp. (n=2)
Diaphus watasei (n=2)
Diaphus signatus
Symbolophorus boops (n=6)
Diaphus brachycephalus (n=3)
Diaphus richardsoni
Diaphus sp.
Diaphus aliciae
Diaphus sp. (n=2)
Diaphus lobatus
Diaphus fulgens (n=2)
Diaphus theta (n=12)
Diaphus rafinesquii (n=10)
Diaphus subtilis
Diaphus holti (n=4)
Diaphus anderseni (n=2)
Diaphus parri (n=2)
Diaphus meadi
Diaphus parri
Diaphus mollis
Diaphus sp.
Diaphus mollis (n=2)
Diaphus effulgens (n=5)
Diaphus thiollierei
Diaphus sp.
Diaphus thiollierei
Diaphus sp.
Diaphus perspicillatus (n=2)
Diaphus metopoclampus (n=6)
Diaphus garmani
Diaphus malayanus
Diaphus splendidus (n=3)
Diaphus sp.
Diaphus antonbruuni
Diaphus whitleyi
Diaphus adenomus
Diaphus bertelseni
Diaphus gigas
Diaphus lucidus
Triphoturus mexicanus (n=4)
Triphoturus sp.
Triphoturus oculeum
Triphoturus nigrescens (n=3)
Lampanyctus crocodilus (n=9)
Lampanyctus macdonaldi (n=13)
Lampanyctus photonotus (n=8)
Lampanyctus achirus (n=28)
MYCSO046-13 - Lampanyctus sp. (n=1)
Lampanyctus bristori
Lampanyctus festivus (n=2)
Lampanyctus tenuiformis (n=5)
Lampanyctus ritteri (n=9)
Lampanyctus ritteri (n=2)
Lampanyctus regale
Lampanyctus ritteri
Lampanyctus regale (n=2)

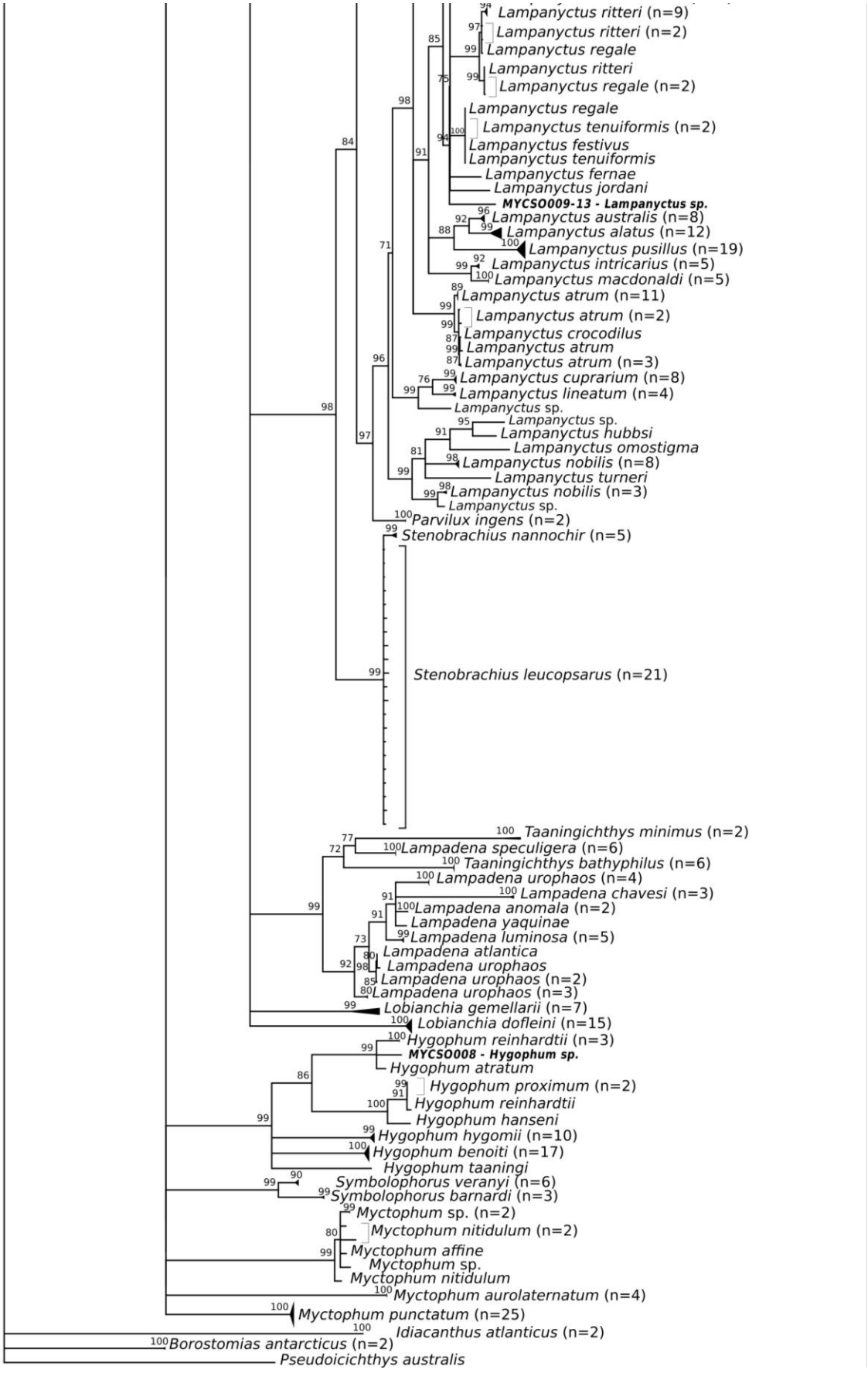

Lampanyctus ritteri (n=9)
Lampanyctus ritteri (n=2)
Lampanyctus regale
Lampanyctus ritteri
Lampanyctus regale (n=2)
Lampanyctus regale
Lampanyctus tenuiformis (n=2)
Lampanyctus festivus
Lampanyctus tenuiformis
Lampanyctus fernae
Lampanyctus jordani
MYCSO009-13 - Lampanyctus sp.
Lampanyctus australis (n=8)
Lampanyctus alatus (n=12)
Lampanyctus pusillus (n=19)
Lampanyctus intricarius (n=5)
Lampanyctus macdonaldi (n=5)
Lampanyctus atrum (n=11)
Lampanyctus atrum (n=2)
Lampanyctus crocodilus
Lampanyctus atrum
Lampanyctus atrum (n=3)
Lampanyctus cuprarium (n=8)
Lampanyctus lineatum (n=4)
Lampanyctus sp.
Lampanyctus sp.
Lampanyctus hubbsi
Lampanyctus omostigma
Lampanyctus nobilis (n=8)
Lampanyctus turneri
Lampanyctus nobilis (n=3)
Lampanyctus sp.
Parvilux ingens (n=2)
Stenobrachius nannochir (n=5)
Stenobrachius leucopsarus (n=21)
Taaningichthys minimus (n=2)
Lampadena speculigera (n=6)
Taaningichthys bathyphilus (n=6)
Lampadena urophaos (n=4)
Lampadena chavesi (n=3)
Lampadena anomala (n=2)
Lampadena yaquinae
Lampadena luminosa (n=5)
Lampadena atlantica
Lampadena urophaos
Lampadena urophaos (n=2)
Lampadena urophaos (n=3)
Lobianchia gemellarii (n=7)
Lobianchia dofleini (n=15)
Hygophum reinhardtii (n=3)
MYCSO008 - Hygophum sp.
Hygophum atratum
Hygophum proximum (n=2)
Hygophum reinhardtii
Hygophum hanseni
Hygophum hygomii (n=10)
Hygophum benoiti (n=17)
Hygophum taaningi
Symbolophorus veranyi (n=6)
Symbolophorus barnardi (n=3)
Myctophum sp. (n=2)
Myctophum nitidulum (n=2)
Myctophum affine
Myctophum sp.
Myctophum nitidulum
Myctophum aurolaternatum (n=4)
Myctophum punctatum (n=25)
Idiacanthus atlanticus (n=2)
Borostomias antarcticus (n=2)
Pseudoicichthys australis

**Fig. 2.2 (previous pages).** Phylogenetic consensus tree of myctophid fishes based on cytochrome c oxidase I (*COI*) gene variation using a maximum likelihood analysis with 10,000 bootstrap permutations. Bootstrap support > 70 is shown above the branches; branches with lower support are collapsed to polytomies; species with Antarctic, broadly Antarctic, and sub-Antarctic distribution pattern following Duhamel et al. (2014) are depicted in purple, dark blue, and light blue, respectively (color figure available online). Number of collapsed samples noted in brackets. Samples where genus level identify was added a posteriori based on position in the tree are noted in bold italics.

### 3.3.2 The rhodopsin gene

The compiled rhodopsin dataset includes 90 sequences of 820 bp length including 511 variable sites. The Myctophidae are monophyletic with 100 % bootstrap support (Fig. 2.3). Diaphinae and Gymnoscopelinae are resolved as monophyletic similarly to the *COI* tree, but Electronini are not. However, the taxonomic sampling is much smaller than for the *COI* dataset and covers 21 myctophid species, whereas Duhamel *et al.* (2014) report 66 species that are at least occasionally recorded south of the Sub-Tropical Front.

Electronini are paraphyletic with the inclusion of *Diogenichthys* sp. and *Myctophum* species. *Electrona antarctica* is placed outside the remaining Electronini and and *Myctophum* spp. as sister group to *Diogenichthys* sp. (BS = 86 %). Within the other Electronini, *Krefftichtys anderssoni* and *E. carlsbergi* diverge first from the monophyletic *Protomyctophum* (BS = 98 %), represented by *P. bolini* and *P. choriodon*.

The genus *Gymnoscopelus* (no other Gymnoscopelinae were available for *rh1*) forms a monophyletic group with high bootstrap support (100 %), with two *G. fraseri* and four *G. nicholsi* diverging first from all other specimens (BS = 99 %). Three further *G. fraseri* are resolved within the remaining clade, next to *G. bolini* and *G. hintonoides* (BS = 78 %) and another clade that comprises *G. braueri* and *G. opisthopterus* (BS = 100 %). However, the resolved topology differs from the *COI* tree, although *G. braueri* and *G. opisthopterus* are resolved as sister taxa in both analyses. *Lampanyctus achirus* – the only other myctophid common in sub-Antarctic waters in this dataset – clusters with a clade of *Lampanyctus* spp. similarly to *COI*.

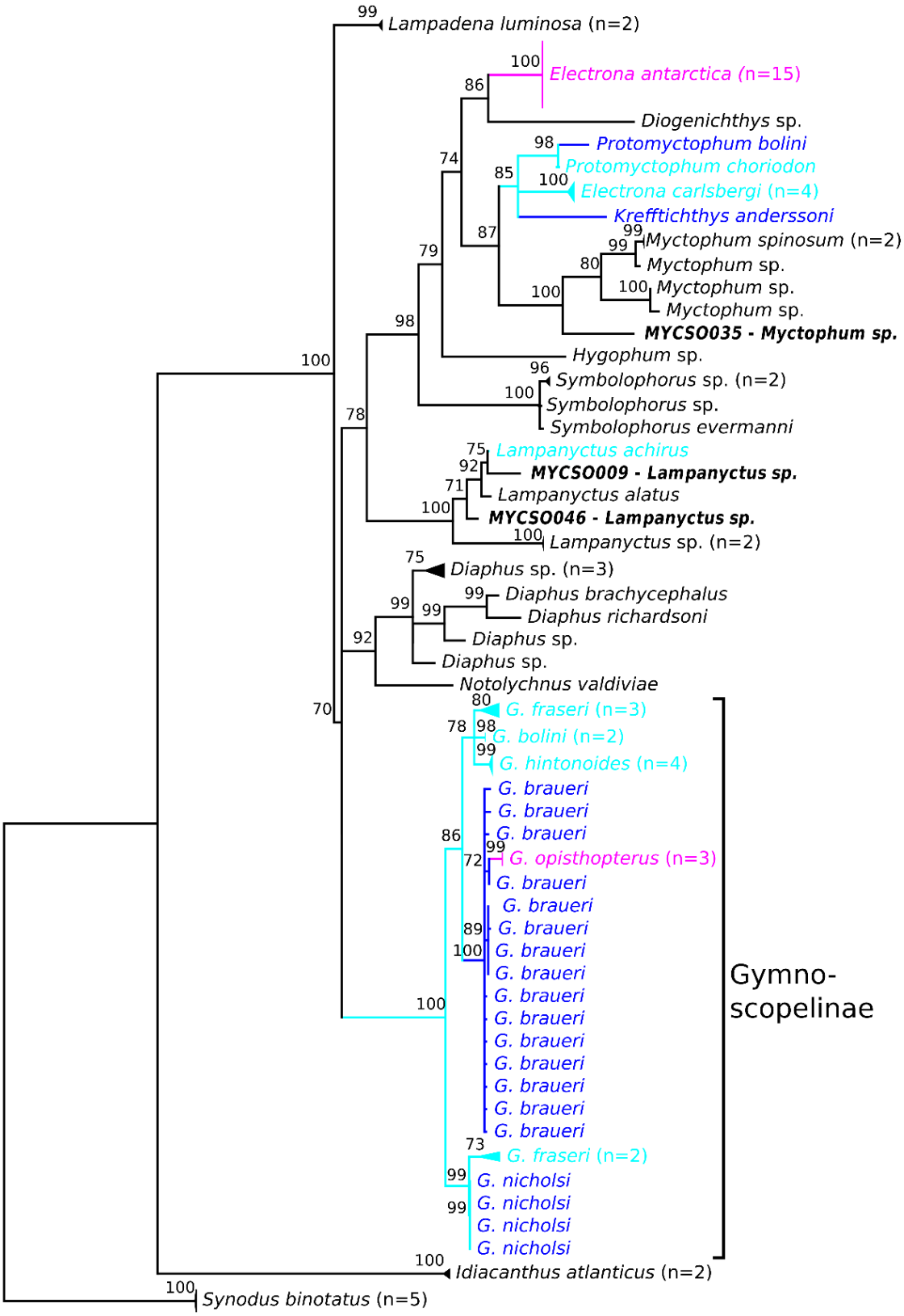



**Fig. 2.3.** Phylogenetic consensus tree of myctophid fishes based rhodopsin (*rh1*) gene variation using a maximum likelihood analysis with 10,000 bootstrap permutations. Bootstrap support > 70 is shown above the branches; branches with lower support are collapsed to polytomies; species with Antarctic, broadly Antarctic, and sub-Antarctic distribution pattern following Duhamel et al. (2014) are depicted in purple, dark blue, and light blue, respectively (color figure available online). Number of collapsed samples noted in brackets.

### 3.3.2 Both markers combined

The concatenated dataset comprises 68 specimens. In total this dataset has 795 variable sites. As expected the concatenation of both markers reduced the size of the dataset, but provided at times higher confidence in the resolved consensus topology. The Myctophidae are monophyletic with 100% bootstrap support (Fig. 2.4). The tribe Electronini is monophyletic except for the inclusion of a *Diogenichthys* sp. and placed within a clade also containing *Symbolophorus* spp., *Hygophum* spp., and *Myctophum* species. This entire clade is a sister group of a clade containing *Notolychnus valdiviae*, *Lampadena luminosa* (luminous lanternfish), a Diaphinae clade, and a clade of *Gymnoscopelus* and *Lampaynctus*. The latter are both monophyletic with 100 % bootstrap support.

Within the Electronini a clade comprising *E. antarctica* and the single *Diogenichthys* sample diverges first from the remaining samples. *Krefftichthys anderssoni* is sister group to a clade with *E. carlsbergi* and the remaining *Protomyctophum* (*P. bolini* and *P. choriodon*). Hence, *Protomyctophum* is monophyletic (BS = 100 %), but *Electrona* and *Krefftichthys* are not. *Diaphinae* are monophyletic, but only *D. richardsoni*, *D. brachycephalus* (short-headed lanternfish), and unidentified *Diaphus* spp. are included; therefore, further inferences are impossible.

Gymnoscopelinae (although only *Gymnoscopelus* is present in this dataset, neither *Scopelopsis* nor *Notoscopelus*) is resolved as monophyletic with 100 % bootstrap support. Within *Gymnoscopelus*, *G. nicholsi* diverges first from other taxa, with high support (BS = 100 %). *G. bolini* is sister group (BS = 96 %) to a clade that contains two other clades, with *G. fraseri* and *G. hintonoides* (BS = 100 %) and *G. braueri* and *G. opisthopterus* (BS = 100 %), respectively. *Lampanyctus achirus* is again placed inside a clade of *Lampanyctus* spp., here sister group of the *Gymnoscopelus* clade (BS = 80 %). Sister to these two clades are the *Diaphinae*, *Notolychnus valdivae*, and two samples of *Lampadena luminosa*.

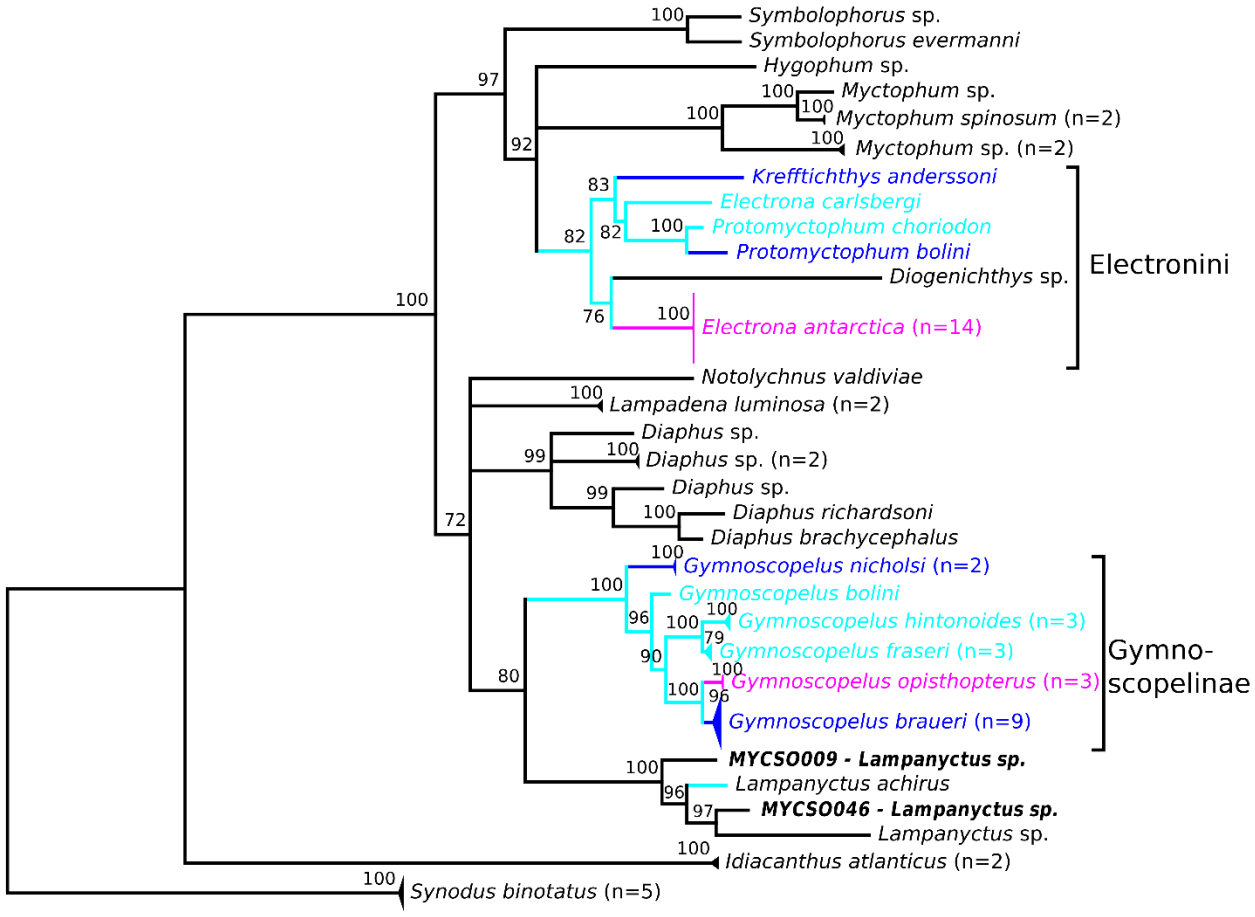


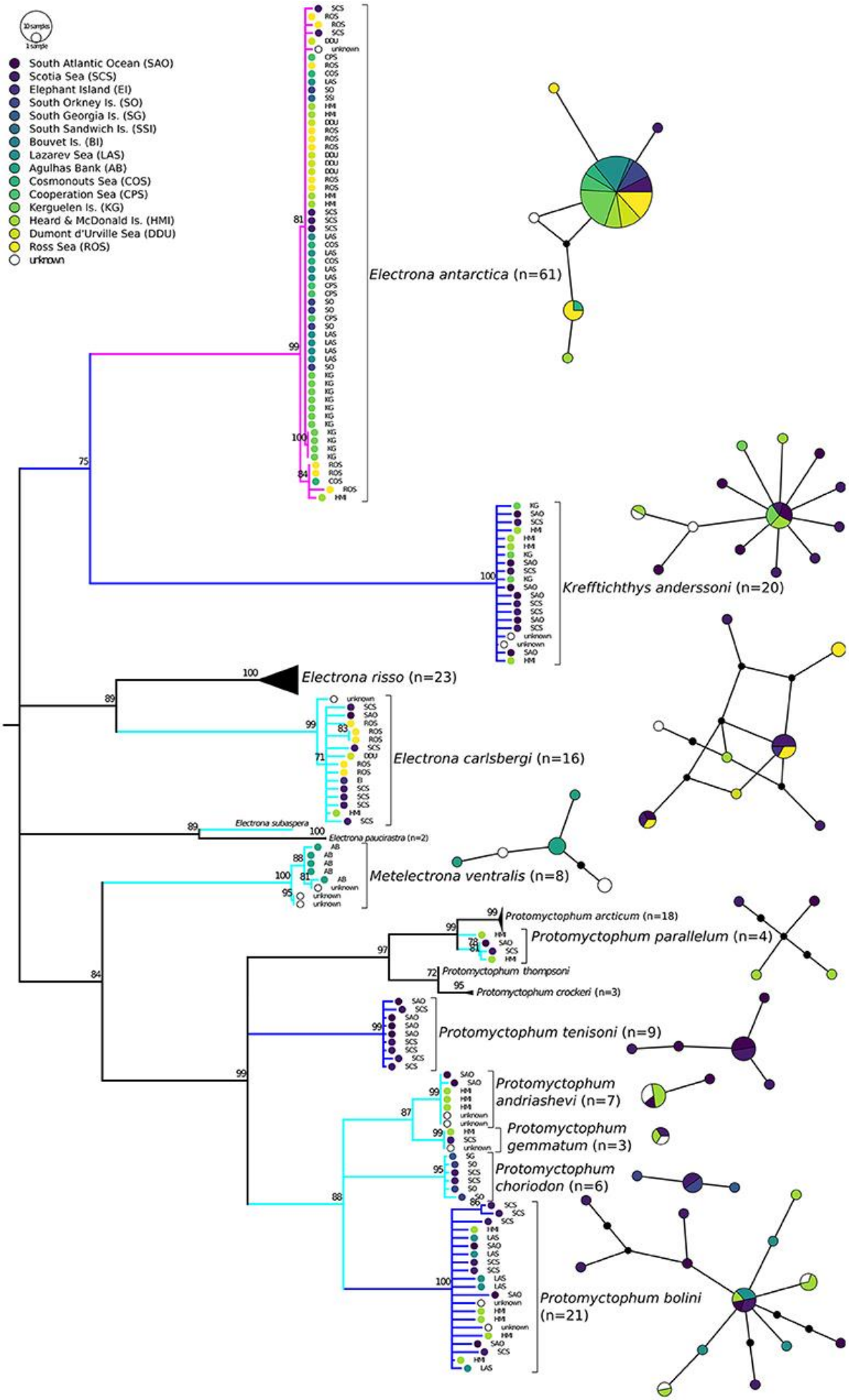

**Fig. 2.4.** Phylogenetic consensus tree of myctophid fishes based on cytochrome c oxidase I (*COI*) and rhodopsin (*rh1*) gene variation using a maximum likelihood analysis with 10,000 bootstrap permutations. Bootstrap support > 70 is shown above the branches; branches with lower support are collapsed to polytomies; species with Antarctic, broadly Antarctic, and sub-Antarctic distribution pattern following Duhamel et al. (2014) are depicted in purple, dark blue, and light blue, respectively (color figure available online). Number of collapsed samples noted in brackets.

### 3.4 Phylogeography and cryptic species

In addition to further parameterizing the BOLD database and investigating phylogenetic relationships of particularly sub-Antarctic and Antarctic myctophids, our data were used to identify phylogeographic diversity patterns of Southern Ocean myctophids. These analyses focused on the largest dataset, *COI*, and mainly on species of the tribe Electronini and the subfamily Gymnoscopelinae, comprising 203 specimens from 16 species and 167 specimens from 12 species, respectively. They were plotted as sub-trees of the *COI* tree (Fig. 2.2) with all (sub-)Antarctic species coded corresponding to sampling locality and associated haplotype networks (Fig. 2.5 & 2.6; codes as in Fig. 2.1). The geographical coverage within species varies from circum-Antarctic to only a few single sites. In general, these data cover specimens from most waters around Antarctica, as well as more northerly areas (Scotia Sea, South Atlantic Ocean north of

the Scotia Sea and off Argentina, waters around Bouvet Island, and the Kerguelen and Heard and McDonald Islands Plateaus, and off South Africa; Fig. 2.1). *Lampanyctus achirus* is not resolved as a single clade, but rather two groups – one with seven individuals caught off South Africa, and one with 21 individuals from the Ross Sea, Dumont d'Urville Sea, Scotia Sea, Kerguelen Islands, and also one individual from South Africa (Fig. S2.1). *Symbolophorus boops* is resolved as group of six individuals with low *COI* variation. Phylogeographic patterns of the other (sub-)Antarctic myctophids are discussed by tribe/subfamily below.

**Fig. 2.5 (next page).** Haplotype networks and phylogenetic consensus tree of myctophid fishes of the tribe Electronini (*sensu* Paxton, 1972) based on cytochrome c oxidase I (*COI*) gene variation using a maximum likelihood analysis with 10,000 bootstrap permutations. Bootstrap support > 70 is shown above the branches; branches with lower support are collapsed to polytomies. Branches of species with Antarctic, broadly Antarctic, and sub-Antarctic distribution pattern following Duhamel et al. (2014) are depicted in purple, dark blue, and light blue, respectively (color figure available online). Geographic origin is reflected by colored circles in the tree and networks, approx. clockwise from West (dark) to East Antarctic (light). In the haplotype networks one branch represents one mutation. Additional mutation steps between samples are indicated with small black circles.


10 samples
1 sample
South Atlantic Ocean (SAO)
Scotia Sea (SCS)
Elephant Island (EI)
South Orkney Is. (SO)
South Georgia Is. (SG)
South Sandwich Is. (SSI)
Bouvet Is. (BI)
Lazarev Sea (LAS)
Agulhas Bank (AB)
Cosmonouts Sea (COS)
Cooperation Sea (CPS)
Kerguelen Is. (KG)
Heard & McDonald Is. (HMI)
Dumont d'Urville Sea (DDU)
Ross Sea (ROS)
unknown
Electrona antarctica (n=61)
Krefftichthys anderssoni (n=20)
Electrona risso (n=23)
Electrona carlsbergi (n=16)
Electrona subaspera
Electrona paucirastra (n=2)
Metelectrona ventralis (n=8)
Protomyctophum arcticum (n=18)
Protomyctophum parallelum (n=4)
Protomyctophum thompsoni
Protomyctophum crockeri (n=3)
Protomyctophum tenisoni (n=9)
Protomyctophum andriashevi (n=7)
Protomyctophum gemmatum (n=3)
Protomyctophum choriodon (n=6)
Protomyctophum bolini (n=21)

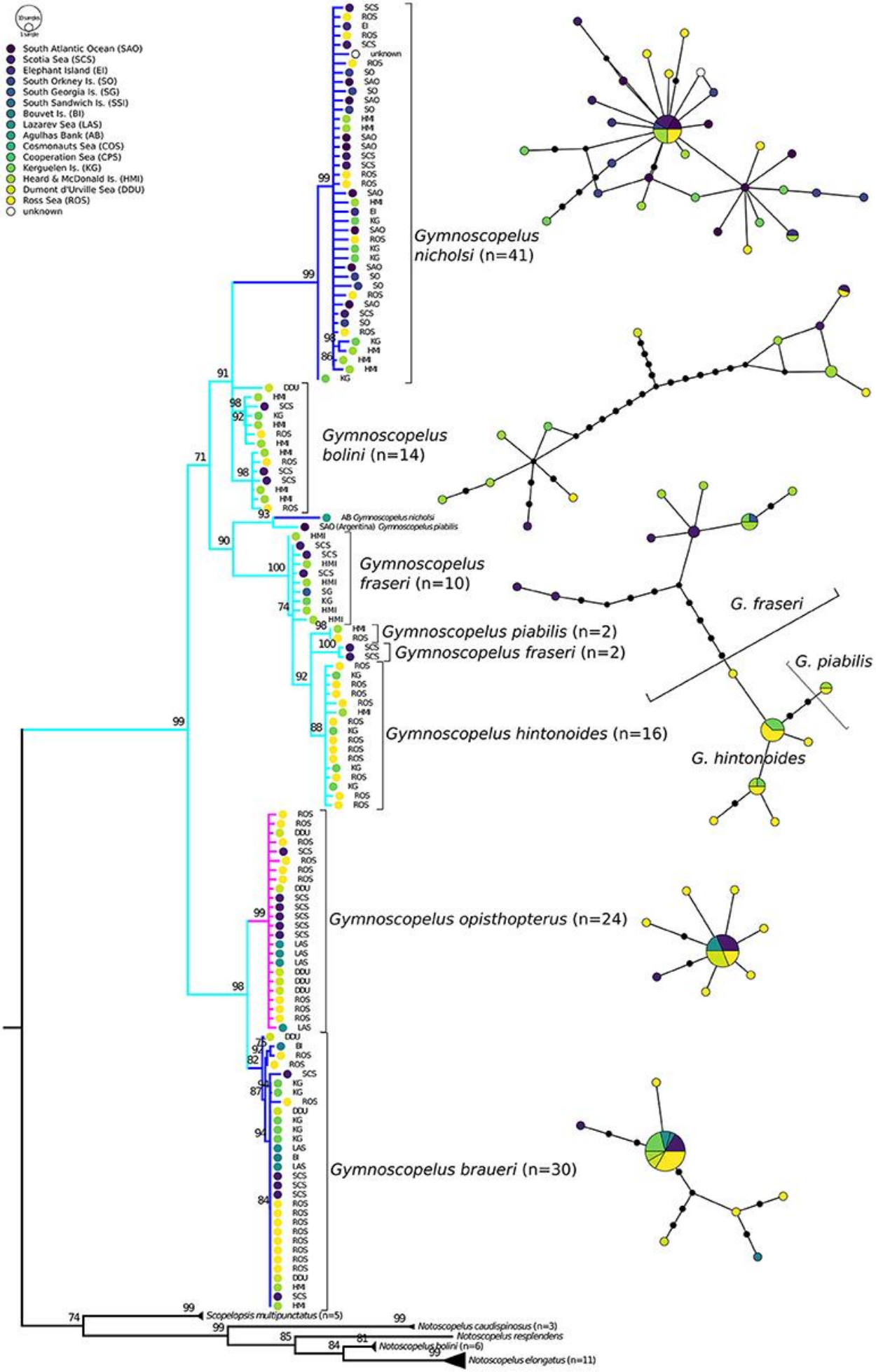

South Atlantic Ocean (SAO)
Scotia Sea (SCS)
Elephant Island (EI)
South Orkney Is. (SO)
South Georgia Is. (SG)
South Sandwich Is. (SSI)
Bouvet Is. (BI)
Lazarev Sea (LAS)
Agulhas Bank (AB)
Cosmonauts Sea (COS)
Cooperation Sea (CPS)
Kerguelen Is. (KG)
Heard & McDonald Is. (HMI)
Dumont d'Urville Sea (DDU)
Ross Sea (ROS)
unknown
Gymnoscopelus nicholsi (n=41)
Gymnoscopelus bolini (n=14)
AB Gymnoscopelus nicholsi
SAO (Argentina) Gymnoscopelus piabilis
Gymnoscopelus fraseri (n=10)
Gymnoscopelus piabilis (n=2)
Gymnoscopelus fraseri (n=2)
Gymnoscopelus hintonoides (n=16)
Gymnoscopelus opisthopterus (n=24)
Gymnoscopelus braueri (n=30)
Scopelopsis multipunctatus (n=5)
Notoscopelus caudispinosus (n=3)
Notoscopelus resplendens
Notoscopelus bolini (n=6)
Notoscopelus elongatus (n=11)
G. fraseri
G. piabilis
G. hintonoides

**Fig. 2.6 (previous page).** Haplotype networks and phylogenetic consensus tree of myctophid fishes of the subfamily Gymnoscopelinae (*sensu* Martin et al., 2018) based on cytochrome c oxidase I (*COI*) gene variation using a maximum likelihood analysis with 10,000 bootstrap permutations. Bootstrap support > 70 is shown above the branches; branches with lower support are collapsed to polytomies. Branches of species with Antarctic, broadly Antarctic, and sub-Antarctic distribution pattern following Duhamel et al. (2014) are depicted in purple, dark blue, and light blue, respectively (color figure available online). Geographic origin is reflected by colored circles in the tree and networks, approx. clockwise from West (dark) to East Antarctic (light). In the haplotype networks one branch represents one mutation. Additional mutation steps between samples are indicated with small black circles.

#### 3.4.1 Electronini

*Electrona antarctica* is the most common species in available *COI* sequences (N = 61). Nonetheless, intraspecific variation appears to be minimal, with only one small group of five individuals clustering apart with moderate support (BS = 84 %) and the vast majority of specimens showing one identical haplotype (Fig. 2.5). The group that clusters apart comprises samples from the Ross Sea, Heard and McDonald Islands, and Cosmonauts Sea, all locations also present in the other group. The intraspecific diversity is low and appears not to be related to geography. *Krefftichthys anderssoni* is present in sufficient numbers (N = 20) and with relatively broad geographical coverage, but no structure is apparent, although haplotype diversity is a little higher compared to *E. antarctica*. The sub-Antarctic *Electrona carlsbergi* features one sample of unknown origin that is separated with moderate support (BS = 71 %). The remaining samples show some haplotype diversity, but no phylogeographic pattern. *Electrona subaspera* as well as *Protomyctophum parallelum* are only present in small numbers (N = 1 and 4, respectively). *Metelectrona ventralis* is represented by only eight samples, five of which come from Agulhas Bank off South Africa (Fig. 2.5). The remaining three samples have no public locality information. However, two of these build a distinct cluster divergent from all others (BS = 95 %). All nine *P. tenisoni* are from the Scotia Sea or the South Atlantic Ocean, showing no signs of phylogeographic diversity. The remaining *Protomyctophum*, i.e., *P. bolini*, *P. choriodon*, *P. andriashevi*, and *P. gemmatum*, also show no sign of elevated intraspecific variability or clustering by location.

#### 3.4.2 Gymnoscopelinae

In *Gymnoscopelus* (Fig. 2.6), *G. nicholsi* shows moderate diversity with all but one specimen represented in one clade, but 29 different haplotypes, while *E. Antarctica* shows only six haplotypes despite larger sample size. The only outlier in *G. nicholsi* is a specimen from the Kerguelen Plateau (BS = 99 %), although the remaining samples include individuals from the

Kerguelen area as well. In contrast, *G. bolini* is split into three groups with high support: (1) one group with two samples each from the Ross Sea, Heard and McDonald Island, and Scotia Sea (BS = 98 %); (2) one group (including a sub-split) with samples from Heard and McDonald as well as Kerguelen Islands, but also Ross and Scotia Seas (BS = 92 %); and (3) one individual from Dumont d'Urville Sea (DDU). East Antarctic coastal waters seem to stand out, but the sampling density is too low for solid inferences. Two other individuals, one nominal *G. piabilis* caught off Argentina and one nominal *G. nicholsi* collected off South Africa, appear as sister group to a clade of *G. fraseri,* remaining *G. piabilis, and G. hintonoides* (BS = 93 %). The latter form a polyphyletic group with only *G. hintonoides* monophyletic in the tree (BS = 88 %), but all three species separated in the haplotype network. All 24 individuals of *G. opisthopterus* form one group with low *COI* variation. *Gymnoscopelus braueri* in turn exhibits more divergence, with one DDU individual clustering basal to a group of two Ross Sea and one Bouvet Island samples (BS = 92 %), which diverge from the 26 remaining individuals. However, the haplotype network of *G. opisthopterus* and *G. braueri* rather resembles the pattern of *E. antarctica* (Fig. 2.5) with one very common shared haplotype.

## 4. Discussion

Our study adds to the increasing knowledge and baseline data of Antarctic marine biodiversity as envisioned by the Census of Antarctic Marine Life and associated initiatives (Schiaparelli *et al.* 2013). It successfully uses the BOLD database to uncover mismatches between morphological and molecular specimen identification and highlights targets for deeper phylogenetic studies to ascertain the position of some species and specimens in the lanternfish family. Lastly, we discuss phylogeographic patterns and the evolution of the family Myctophidae in the Southern Ocean in general and hypothesize that the presence of myctophids in the high polar seas is the result of multiple colonization events.

### 4.1 Extending and using the DNA barcode library for specimen identification

The newly added sequences expand the public DNA barcoding database of Myctophidae to more than 1000 individual sequences. Of these, 263 belonged to specimens captured in the Southern Ocean, which represents a substantial increase of the barcode library of Antarctic mesopelagic fish. This will be of major importance for future ecological studies that intend to use the library for specimen identification. The value of the BOLD database is likely to increase even further

with the development of metabarcoding studies. Recent approaches include for example the detection of tropical sharks (Bakker *et al.* 2017), large-scale larval fish ecology through efficient identification of thousands of larvae (Kimmerling *et al.* 2018), as well as Antarctic studies characterizing notothenioid fish assemblages (Cowart *et al.* 2017) and toothfish diet (Yoon *et al.* 2017). All these examples are fully dependent on a high-quality reference database to match metabarcoding sequences, as the lack of identification can lead to reduced or biased results and interpretations. Good coverage of the (sub-) Antarctic teleost fauna, which now includes demersal fishes (Rock *et al.* 2008; Dettai *et al.* 2011; Smith *et al.* 2011b, 2012; Mabragaña *et al.* 2016) and large parts of the meso- and bathypelagic fish fauna (this study), will likely be highly valuable for future metabarcoding studies investigating, for example, the diet and trophic position of top predators. Species that are quickly digested can be detected with this molecular approach, although quantification remains challenging. Such studies can contribute to refine our understanding of food webs in Antarctic and sub-Antarctic waters (Cornejo-Donoso & Antezana 2008; Pinkerton & Bradford-Grieve 2014), a task of high relevance with regard to the ecosystem approach of the Commission for the Conservation of Antarctic Marine Living Resources (CCAMLR; Kock *et al.* 2007; Constable 2011).

Morphological specimen characterizations were verified with DNA barcoding. Generally, the success of specimen identification using the BOLD database (Ratnasingham & Hebert 2007) was very high. Where morphological identification is challenging because discriminating characteristics are frequently lost during sampling, such as in the Myctophidae, DNA barcoding represents a useful complement to traditional identification. In eleven cases we were able to improve the initial identification and in seven further cases, we discovered mismatches between taxonomy and genetic signature, which were attributed to initial misidentification or mislabeling. Without DNA barcoding these would likely have retained an erroneous identification, which in turn poses problems for further ecological analysis and interpretation. As other studies have shown, DNA barcoding of Antarctic fishes and skatesis a useful molecular taxonomic approach (Smith *et al.* 2008, 2011b, 2012; Rock *et al.* 2008; Duhamel *et al.* 2010; Lautredou *et al.* 2010; Rey *et al.* 2011). Furthermore, it may also serve as a starting point for phylogenetic and phylogeographic investigations (Duhamel *et al.* 2014; Mabragaña *et al.* 2016). Such help for taxonomy is highly needed in times where classical taxonomic expertise has become rare (Cao *et al.* 2016). At least some myctophids, however, may be particularly prone to DNA degradation problems. Some of the authors observed that samples from non-myctophid fishes collected during the same expeditions and processed in the same way had much higher

amplification success rates. This might be linked to (taxon-)specific degradation processes and we therefore recommend that myctophids are processed first when treating a fish catch for scientific purposes.

### 4.2 Phylogeny and phylogeography of Southern Ocean mesopelagic fishes

The topology of phylogenetic trees constructed using *COI* and *rh1* concur to a great extent with recent multi-marker phylogenies (Poulsen *et al.* 2013; Davis *et al.* 2014; Denton 2014; Martin *et al.* 2018). However, our data cannot be used to discuss the relationship between myctophid tribes. The elsewhere well-supported monophyly of Myctophidae is not resolved in our full *COI* data set with some aulopiform and neoscopelid species placed within Myctophidae. In addition, the first split of Myctophidae showed polytomy with nine branches; in other words, the placement of these clades is unclear (Fig. 2.2). The *rh1* tree spans fewer taxa and resolved *Lampadena* as sister group to all other Myctophidae, which is not supported by other studies. *COI* and *rh1* alone are clearly not sufficient to accurately resolve deeper phylogenetic relationships of Myctophidae. However, these markers offer the most comprehensive datasets, which are important for a holistic understanding of the approximately 395 currently described Myctophidae (Eschmeyer and Fong, 2018). High taxon density can positively affect tree topology through breaking down of otherwise very long branches. *COI* in particular is unmatched regarding coverage with more than 700 sequences from 149 species already previously available in BOLD for our analysis. At the subfamily to genus level our results indeed match better with recent hypotheses of myctophid intrarelationships. For example, within Lampanyctinae (*sensu* Martin *et al.* 2018) the *COI* tree shows that the genera *Lampadena* and *Taaningichthys* are related, as are *Bolinichthys*, *Lepidophanes*, and *Ceratoscopelus* and, lastly, *Stenobrachius*, *Triphoturus*, *Parvilux*, and *Lampanyctus*. This concurs with the multi-marker results of Denton (2014) and Martin *et al.* (2018). The concatenated dataset even resolves an initial split into a clade with Diaphinae, Lampanyctinae, Gymnoscopelinae, and Notolychninae and another clade with Electronini and Myctophini (Fig. 2.4). This pattern matches the analyses of Poulsen *et al.* (2013), Davis *et al.* (2014), Denton (2014), and of Martin *et al.* (2018), with the exception that the latter find Diaphinae closer related to Myctophinae. Ultimately, the deep phylogenetic hypotheses of Myctophidae still need further work. With respect to phylogeny our data can serve as starting point to highlight genera or species that are of particular interest for further analysis, such as studies analyzing entire mitogenomes or large numbers of single nucleotide polymorphisms. We restrict this discussion here to species common in the Southern Ocean but

extend it towards intraspecific genetic diversity by analyzing phylograms and haplotype networks in relation to geography. Such phylogeographic patterns are important to understand the distribution of biodiversity in the mesopelagic zone of the Southern Ocean and to plan conservation and management actions accordingly in light of climatic changes. We discuss specific phylogenetic and phylogeographic implications and recommendations of our study as compared to other recent phylogenies for: (1) Electronini, (2) Gymnoscopelinae (*sensu* Martin *et al.* 2018); (3) other Southern Ocean myctophids, and (4) other Southern Ocean mesopelagic, non-myctophids fishes.

The tribe Electronini is monophyletic within our *COI* analysis, but not in the *rhI* tree, where *Diogenichthys* appears related to *Electrona antarctica*. (Fig. 2.3). This signal is likely the reason that *Diogenichthys* is also placed within Electronini in the concatenated analysis (Fig. 2.4). Other recent studies have all corroborated the monophyly of Electronini (Poulsen *et al.* 2013; Davis *et al.* 2014; Denton 2014; Martin *et al.* 2018). It is possible that the accuracy of the *rhI* marker here is affected by a bias in base composition across taxa (Chen *et al.* 2003). Inferences from *rhI* alone as well as results from the concatenated dataset that may stem primarily from the *rhI* signal as is the case here must therefore be interpreted with caution. The relationships within Electronini are still somewhat obscure (Denton 2014). All our trees support the monophyly of *Protomyctophum*, but the placement of *Protomyctophum* and the other genera, *Metelectrona*, *Electrona*, and *Krefftichthys* remains unclear. In the *COI* dataset, *E. antarctica* and *K. anderssoni* cluster together and apart from remaining *Electrona* spp. With *rhI* and the concatenated dataset *K. anderssoni* rather appears to be a sister group of all other species, except *E. antarctica*, somewhat similar to Martin *et al.* (2018). However, Denton (2014) resolved *Krefftichthys* as sister group of only *Protomyctophum*. These contradictory results are evidence that more detailed studies are needed to clarify relationships within this tribe. Within *Protomyctophum* and using *COI*, where more than two *Protomyctophum* species were included, the split between the subgenera *Protomyctophum* and *Hierops* is supported except for *P. tenisoni* which diverges first (Fig. 2.2; Gordeeva 2013; Denton 2014). With some additional support a revision of *Protomyctophum* as suggested by Denton (2014) appears sensible.

All Electronini species are recovered as single clusters, with low to moderate (1-13 haplotypes) intraspecific levels of diversity (Fig. 2.5). They show no divergent groups that might point to undescribed or cryptic species. The striking example of this is *E. antarctica*, where the majority

of individuals belong to the same haplotype despite the distant locations. A dominant, widespread haplotype may indicate reduced genetic diversity due to, for example, a recent bottleneck. However, more variable markers show high levels of genetic diversity in this species (Van de Putte *et al.* 2012a). *Electrona antarctica* is arguably the most common myctophid in Antarctic waters with a circum-Antarctic distribution and preference for water temperatures below 2.5° C (Hulley 1990; Duhamel *et al.* 2014). Similar to Antarctic krill (Deagle *et al.* 2015), its enormous abundance may be the key factor leading to virtual panmixia, i.e. genetic homogeneity across the entire distribution range (Van de Putte *et al.* 2012a). Other Electronini species, in particular *Protomyctophum bolini* and *Electrona carlsbergi*, show more variability in their *COI* haplotypes, but no relation to geography was detectable. It seems that in addition to high effective population size, the strong flows of the Antarctic Circumpolar Current causes high connectivity, resulting in at most subtle spatial genetic structure, but clearly no pronounced phylogeographic structure of fishes in the tribe Electronini.

The subfamily Gymnoscopelinae is well supported as monophyletic group in all our analyses confirming findings of other phylogenetic studies (Poulsen *et al.* 2013; Davis *et al.* 2014; Denton 2014; Martin *et al.* 2018). We therefore adopted the taxonomic revision of Martin *et al.* (2018), who promoted the former lampanyctine tribe of Gymnoscopelini to Gymnoscopelinae. The genus *Gymnoscopelus* appears clearly monophyletic. Our concatenated dataset resolves *G. nicholsi* as sister group to all other species (BS = 100%; Fig. 2.4), as suggested by Denton (2014), the only other study that included more than two *Gymnoscopelus* species. In the individual *COI* and *rhI* datasets the placement of *G. nicholsi* was not apparent (Fig. 2.2-3), highlighting the value of an additional nuclear marker to increase confidence in phylogenetic positioning. There are examples where rhodopsin can distinguish fish species, where *COI* fails (Thiel & Knebelsberger 2016), such as in the present case. In our case all *Gymnoscopelus* spp. can be delineated using *COI*, but the exact position within the tree remains to be evaluated, except where they are corroborated by Denton's (2014) results.

With respect to phylogeography, the Gymnoscopelinae show a different pattern from the Electronini (Fig. 2.6). *G. nicholsi* features various *COI* haplotypes, but only one individual, collected off the Kerguelen Islands, stands out in the phylogenetic tree (Fig. 2.6). This is similar to the pattern of *P. bolini* within the Electronini. *G. bolini* on the other hand splits into three groups based on *COI* variation (Fig. 2.6). This pattern might hint at genetic structuring, but this

needs to be investigated further, as the single sample from East Antarctica (DDU) is divergent, while samples from Ross and Scotia Sea, and Heard and McDonald, and Kerguelen Islands group together. Other factors than circumpolar position may be at play here, for instance trophic niche partitioning, sexual selection, or simply increased levels of genetic variability. The other species show relatively low variability and no pattern related to sampling locality, with two exceptions, both samples collected in lower latitudes (off South Africa and Argentina, respectively). These two samples were identified as *G. piabilis* and *G. nicholsi*. Given their very high *COI* sequence divergence (23 mutations apart from the nearest neighbor in median joining networks, not shown in figure), we recommend re-examination of the specimens, if feasible, to investigate whether they possibly belong to different (cryptic) species or sub-species.

The available sequences identified as *Symbolophorus boops* (BOLD references DSFSE476-08 to DSFSE480-08 and DSFSG260-10) cluster apart from the two other *Symbolophorus* clades resolved in our *COI* tree (one composed of *S. californiensis*, *S. reversus*, *S. evermanni*, *Symbolophorus* sp., and *S. rufinus* and the other composed of *S. barnardi* and *S. veranyi*; Figure 2.2). Instead these sequences settle within the Diaphinae (*sensu* Martin *et al.* 2018). Unfortunately we discovered a posteriori that the COI sequences included here as *S. boops* were likely misidentified on BOLD. These sequences are probably from a *Diaphus* species (P. A. Hulley, pers. comm.) currently also not present on BOLD, but the specimens are in poor condition, preventing definite identification. The correction has been transmitted to the BOLD database. Other studies that included genetic data proposed that *Symbolophorus* is closer related to *Myctophum*, *Hygophum*, and other genera, as opposed to Diaphinae, but they all lacked specimens of *S. boops* (Poulsen *et al.* 2013; Denton 2014; Martin *et al.* 2018). Therefore, we highly recommend the collection of further samples/sequences in order to resolve the phylogenetic position of *S. boops*, and to re-identify the specimens erroneously labeled as *Symbolophorus boops*. In fact, the entire genus would benefit from a detailed systematic revision as already noted by Wisner (1976).

The genus *Nannobrachium* was recently placed into synonymy with *Lampanyctus* (Martin *et al.* 2018). Our results fully support this across all datasets and all former *Nannobrachium* were therefore labeled as *Lampanyctus*. Interestingly, *Lampanyctus achirus* is the only species of the Lampanyctinae with a sub-Antarctic distribution (*sensu* Martin *et al.* 2018). Based on *COI* the species also splits into two divergent clades with 99 and 100 % bootstrap support, respectively

(Fig. S2.1). One clade consists entirely of specimens caught off South Africa and the other clade of specimens from around Antarctica. These results, however, need to be corroborated with nuclear and morphological data. This again underscores the importance of specimen vouchers for groups with difficult morphological characters and uncertain species delineation. Nevertheless, recent studies have shown that undescribed species abound even in groups that were thought well known (Geiger *et al.* 2014). The pattern observed for *L. achirus* might be indicative of cryptic or undescribed species as found before in Myctophidae, for example in *Benthosema pterotum* (skinnycheek lanternfish; Zahuranec *et al.* 2012). Currently, there are two specimens of *B. pterotum* with *COI* sequences in BOLD (from Poulsen *et al.* 2013 and Denton 2014) and they also show a deep split. Additional *COI* sequencing of the specimens used by Zahuranec *et al.* (2012) could therefore enable fast, cost-efficient, and confident discrimination between the two cryptic species with *COI* in the future.

According to Duhamel *et al.* (2014) the most abundant non-myctophid mesopelagic fish families in the Southern Ocean are Bathylagidae (deep-sea smelts; 5 species), Gonostomatidae (bristlemouths; 4 spp.), Notosudidae (waryfishes; 2 spp.), Paralepididae (barracudinas; 4 spp.), and 5 species of Stomiidae (barbeled dragonfishes). We found *COI* sequences for only four (*Notolepis coatsi*, *Bathylagus antarcticus*, *Idiacanthus atlanticus*, and *Borostomias antarcticus*), plus five bathypelagic species (*Icichthys australis*, *Lagiacrusichthys macropinna*, *Poromitra capito*, *Sio nordensjkoldii*, and *Oneirodes notius*). Apart from Myctophidae and Nototheniidae, and apart from species only occasionally recorded south of the Sub-Tropical Convergence, Duhamel *et al.* (2014) list 51 species for the whole Southern Ocean pelagic zone. Thirty-four of those are present with *COI* in BOLD (January 2018). Many less abundant species remain to be sequenced in order to complete the reference database. Intraspecific variability is difficult to assess for the available species due to the limited number of samples. *Poromitra capito* (N = 3) showed two haplotypes, *Notolepis coatsi* (N = 4) showed three haplotypes. The genus *Bathylagus* is believed to comprise at least three species in the Southern Ocean, although morphological discrimination is very difficult. Preliminary evidence suggests that four species with distinct *COI* signature are present in the Scotia Sea (Collins et al., unpublished). In this study twelve *Bathylagus antarcticus* (Antarctic deep-sea smelt), all collected in the Lazarev Sea, were included, which had twelve unique haplotypes and showed at least two divergent clades in our *COI* tree (BS = 97 and 84 %). Dettai *et al.* (2011) also found diverging clades of *B. antarcticus* in the Dumont d'Urville Sea. We recommend a detailed integrative taxonomic investigation of all available *Bathylagus* specimens using morphology and several genetic markers to clarify the

status of this genus. The other species mentioned above were only available in low numbers (N ≤ 2), which does not permit an examination of intraspecific variation.

### 4.3 Evolution of Myctophidae in the Southern Ocean

Overall, intraspecific genetic divergences are very low with only one case where *COI* variation clearly relates to sampling locality (*Lampanyctus achirus*). This may be expected, because large abundance promotes gene flow and homogenization (Hauser & Carvalho 2008), especially in conjunction with the strong oceanographic connectivity enforced by the ACC (Orsi *et al.* 1995; Young *et al.* 2015). Another key contributing factor is the pelagic lifestyle of myctophids, characterized by seemingly free dispersal and the lack of a connection to a specific benthic habitat. Recent analyses suggest that myctophid biomass in Antarctic waters is dependent on mass immigration from lower latitudes (except for *E. antarctica* and *K. anderssoni*), which may support the idea of high connectivity (Saunders *et al.* 2017). However, examples of extended geographic structure despite high dispersal capabilities have been found in the Antarctic (Havermans *et al.* 2011; Damerau *et al.* 2014). In addition, weak genetic structuring has been observed for myctophids, although this was between the Mediterranean Sea and the Atlantic Ocean (Pappalardo *et al.* 2015). Circumpolar connectivity patterns in the sub-Antarctic and Antarctic are complex and variable, largely depending on the interplay of oceanography and life history traits (Moon *et al.* 2017 and references therein). For myctophids the combination of large abundance and a free-roaming, pelagic life style seems to cause a lack of genetic structuring. Our analysis is only a preliminary attempt to characterize such structure and is inherently biased towards common species, for which sufficient numbers of samples were available. If abundance indeed has a strong effect on population genetic or phylogeographic structure of lanternfishes, especially rare species should be investigated in detail. So far, only one study has investigated genetic structure of an Antarctic myctophid with multiple, variable markers (Van de Putte *et al.* 2012a). Insights into the genetics of myctophid populations would be useful in order to optimize current modeling efforts (Koubbi *et al.* 2011; Freer *et al.* 2018), which in turn are important for conservation planning in the Southern Ocean (Hill *et al.* 2017). Attempts to explain and forecast mesopelagic fish distribution ranges typically use oceanographic parameters, particularly temperature and salinity (Koubbi *et al.* 2011; Duhamel *et al.* 2014; Olivar *et al.* 2017). Unsurprisingly, characteristics of deep water masses better explain myctophid species occurrence than surface water properties (Olivar *et al.* 2017). For both the characterization of current distribution and for future predictions the questions arise: What temperatures can

Southern Ocean myctophids physiologically sustain? And to what level have they already adapted to colder waters and can they adapt to current rates of environmental change?

Benthic biodiversity in the Southern Ocean is comparably high, including many endemic species (Brandt *et al.* 2007). This is not the case for the mesopelagic fauna, mostly because it is not as isolated as Antarctic shallow water systems. However, it appears that only a few myctophid species adapted to permanently thrive under the prevailing environmental conditions. Hence, only 21 of 68 species ever recorded in the Southern Ocean are truly Antarctic and sub-Antarctic and endemic to these waters. This corresponds to 7.1% of all described myctophids, probably an underestimate considering the deep molecular divergences within the non-Antarctic species in the group, as e.g. in the supposedly monotypic genus *Notolychnus*. In contrast, equatorial and tropical fish communities feature high myctophid species richness (Olivar *et al.* 2017). Compared to for example pycnogonids, where 17.3% of all species are endemic to (sub-)Antarctic waters (Krabbe *et al.* 2010), myctophids seem to have diversified less in the Southern Ocean. In fact, just two species (*E. antarctica* and *G. opisthopterus*) exhibit what is described as an Antarctic distribution pattern (Duhamel *et al.* 2014). Looking at the phylogenetic trees, it becomes even clearer that adaptation to Southern Ocean conditions must have occurred repeatedly. There is no single species flock of Southern Ocean myctophids, but species from at least three subfamilies *sensu* Martin *et al.* (2018) are in fact true Southern Ocean species, although the vast majority belongs either to the Electronini or *Gymnoscopelus*. This suggests parallel evolution within similar environments based on similar genomic architecture. Denton (2018) recently showed that lanternfishes experienced elevated diversification rates initiated around the Eocene-Oligocene transition, which on the other hand could indicate that the formation of the ACC was an important evolutionary trigger for mesopelagic fish species. Southern Ocean lanternfishes are an interesting model to study evolution and speciation in the deep sea (de Busserolles *et al.* 2013; Denton 2014). The diversification of Electronini (especially *Protomyctophum*, see also Denton 2018) seems particularly intriguing, as it might represent an example of a (relatively small scale) marine adaptive radiation.

## 5. Conclusion

With this study we substantially extend the DNA barcode library of Antarctic mesopelagic fish, particularly lanternfishes. The combination of morphological and molecular identification led to confident species level identification in 281 out of 299 cases. Several misidentifications or

otherwise uncertain samples were identified in the database. Overall, DNA barcode libraries provide a robust reference dataset for specimen identification, especially to the rescue of fragile morphological characters. As expected, the mitochondrial *COI* and nuclear *rhI* genetic markers were not sufficient to resolve deep phylogenetic relationships. However, our results are largely congruent with recent phylogenetic studies of the family. Some of our findings suggest the importance of further study or re-identification, e.g., of *Symbolophorus boops*. In addition, we highlight potential (pseudo-)cryptic or unrecognized species and recommend further investigation of *Gymnoscopelus bolini*, two specific *Gymnoscopelus* specimens (nominally identified as *G. piabilis* and *G. nicholsi*), *Lampanyctus achirus* and the non-myctophid genus *Bathylagus*. That myctophid species from at least three subfamilies are Southern Ocean species suggests that colonization and adaptation to this environment has occurred repeatedly. Overall, spatial divergence of species is rare in this family, potentially due to the enormous abundance of many myctophids and the homogenizing force of ocean currents. Finally, this study provides an overview of currently available Antarctic samples and associated levels of intraspecific diversity, both of which may facilitate future ecological, phylogenetic, and evolutionary investigations of Southern Ocean myctophids, a fish family that surely warrants increased scientific attention.

## 6. Acknowledgements

We thank the officers, crew and scientists of the cruises involved in the capture of the samples. The cruises with RV James Clark Ross were supported by British Antarctic Survey (BAS) and its Discovery 2010 project. The RV Polarstern cruises were supported by the Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research (AWI). The CAML-CEAMARC cruises of RV Aurora Australis and RV Umitaka Maru (IPY project no. 53) were supported by the Australian Antarctic Division (AAD), the Japanese Science Foundation, the French polar institute IPEV, the CNRS, the MNHN, and the ANR (White project ANTFLOCKs USAR n07-BLAN-0213-01). The “Austral” cruise POKER 2 was supported by specific grants of the Ministry of Alimentation, Agriculture and Fisheries (MAAP), the Marine Reserve of TAAF and the contributions of French ship owners involved in the Kerguelen Islands fisheries, with the logistic help of SAPMER and TAAF. We thank F. Busson and the 2013 JAMSTEC RV Hakuho Maru campaign led by K. Tsukamoto. We thank the editor and reviewers for constructive comments.

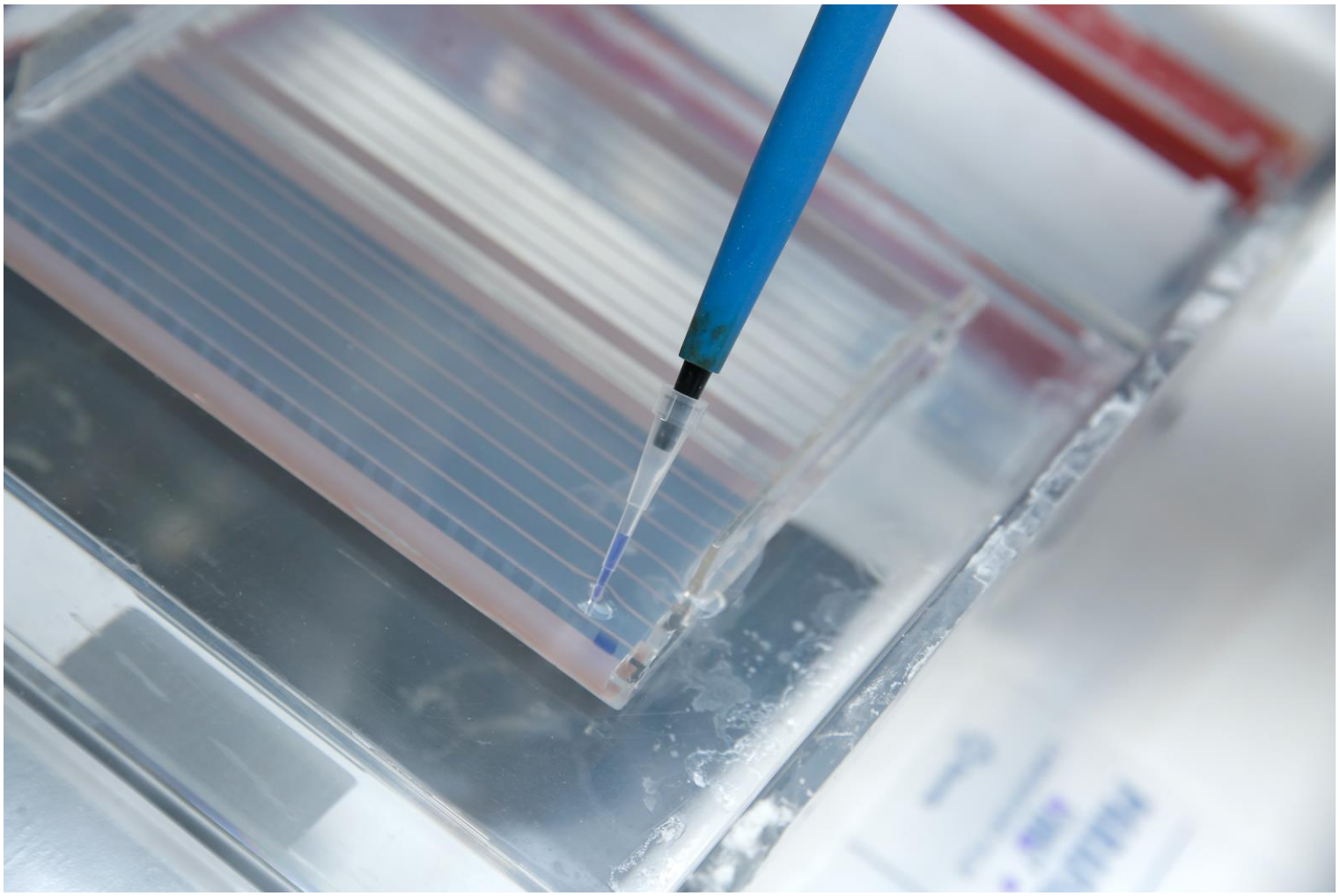
## 7. Funding

Molecular work including sequencing was supported by the governments of France and Canada. The former through the Service de Systematique Moleculaire of the Muséum National d'Histoire Naturelle (MNHN; UMS2700), supported by the network Bibliothèque du Vivant funded by the CNRS, the MNHN, the INRA, and the CEA (Centre Nacional de Séquençage). The latter through Genome Canada and the Ontario Genomics Institute for the International Barcode of Life Project. This work was further supported by the Belgian projects vERSO and RECTO (http://rectoversoprojects. be). This is contribution no. 25 to the vERSO project, funded by the Belgian Science Policy Office (BELSPO, Contract no. BR/132/A1/vERSO). The first author was funded by a grant from the former Flemish agency for Innovation by Science and Technology, now Flanders Innovation and Entrepreneurship (VLAIO, grant no. 141328). DS was supported by the Alfred P. Sloan Foundation and the Food From Thought research program funded by the Canada First Research Excellence Fund.

## 8. Data archiving statement

All samples and sequence IDs and associated metadata can be found in Supplementary Table S2.1 and on BOLD.

# CHAPTER 3: Facilitating population genomics of non-model organisms through optimized experimental design for reduced representation sequencing



*Note*: this manuscript is in preparation for publication.

**Facilitating population genomics of non-model organisms through optimized experimental design for reduced representation sequencing**

Henrik Christiansen, Franz M. Heindler, Bart Hellemans, Quentin Jossart, Francesca Pasotti, Henri Robert, Marie Verheye, Bruno Danis, Marc Kochzius, Frederik Leliaert, Camille Moreau, Tasnim Patel, Anton P. Van de Putte, Ann Vanreusel, Filip A. M. Volckaert and Isabelle M. Schön

**Abstract**

Genome-wide data are invaluable to characterize differentiation and adaptation of natural populations. In reduced representation sequencing a genome is subsampled repeatedly across many individuals. However, this approach requires careful optimization and fine-tuning to deliver high marker density while being cost-efficient. The number of genomic fragments created through restriction enzyme digestion and the sequencing library setup must be matched to achieve sufficient sequencing coverage per locus. Here, we present a workflow based on published information and computational and experimental procedures to investigate and streamline the applicability of reduced representation sequencing. In an iterative process genome size estimates, restriction enzymes and size selection windows were tested and scaled in six classes of marine Antarctic taxa (Ostracoda, Malacostraca, Bivalvia, Asteroidea, Actinopterygii, Aves). Achieving high marker density would be costly in amphipods, our malacostracan target taxon, due to their large genome size. We propose alternative approaches such as mitogenome or capture sequencing for this group. Pilot libraries were sequenced for all other target taxa. Ostracods, bivalves, sea stars, and fish showed overall good coverage and marker numbers for downstream population genomic analyses. In contrast, the bird test library produced low coverage and few polymorphic loci, likely due to degraded DNA. Prior testing and optimization are important to identify which groups are amenable for reduced representation sequencing and where alternative methods may currently offer better cost-benefit ratios. The steps outlined here are easy to follow for other non-model taxa with little genomic resources, thus stimulating efficient resource use for the many pressing research questions in molecular ecology.

**1. Introduction**

Evolutionary and ecological population genetic studies are important to understand how the diversity of life on earth is distributed, has evolved and may respond to future environmental changes (Ellegren 2014). A grand challenge has been to document this biodiversity and understand its role in maintaining ecosystem functionality, particularly in the ocean (Borja 2014) and even more so in frontier areas such as the deep-sea and polar regions (Brandt *et al.* 2007). Molecular data collection has benefitted from a revolution in sequencing technologies such that genomics, where billions of nucleotides are screened simultaneously, is now an integral part of the biological toolbox (Andrews *et al.* 2016; Kelley *et al.* 2016; Matz 2017). Genome-wide data open new avenues of ecological and evolutionary research, especially to study local adaptation (Reitzel *et al.* 2013; Savolainen *et al.* 2013). Given ever-increasing rates of anthropogenic disturbance, it is highly important to assess spatio-temporal genomic diversity, adaptation patterns and resilience of non-model organisms (Funk *et al.* 2012; Hoffmann *et al.* 2015).

Similar to previous methodology shifts in population genetics (e.g. from Amplified Fragment Length Polymorphisms [AFLP] to microsatellites), the transition to novel methods requires detailed understanding of the new technology, its potential as well as its pitfalls, and careful experimental planning. While some study systems are moving towards population-specific shallow re-sequencing of whole genomes (e.g. important commercial fish species; Fuentes-Pardo and Ruzzante, 2017; Lamichhaney et al., 2017; Therkildsen and Palumbi, 2017), many species of interest with less extensive genomic resources rely on reduced representation genome sequencing techniques. These methods are attractive because they make more frugal use of sequencing volume. They commonly apply one or several restriction endonuclease enzymes to fragment the target genome into smaller portions to be sequenced, thus reducing costs. Millions of reads from high-throughput sequencing platforms are then aligned against either a reference genome or, alternatively, a *de novo* reference catalog of loci. Subsequently, genetic variants, most commonly single nucleotide polymorphisms (SNPs), but also microsatellites, introns, copy number variants (CNV) and microhaplotypes are determined (Davey *et al.* 2011; Willis *et al.* 2017; Baetscher *et al.* 2018; Rochette *et al.* 2019).

In ecological and evolutionary research of natural populations, Restriction site-Associated DNA sequencing (RADseq) and derivatives (Baird *et al.* 2008; Peterson *et al.* 2012; Toonen *et al.* 2013),

and Genotyping by Sequencing (GBS; Elshire et al., 2011; Poland et al., 2012) are the most commonly used terms for a variety of methods that essentially employ the same principle. Unfortunately, the scientific nomenclature for such approaches lacks consistency (Campbell *et al.* 2018; Hadfield & Retief 2018). Here, we follow the reasoning of Campbell et al. (2018) and use the term reduced representation sequencing (RRS; Altshuler et al., 2000) to refer to all of these methods. RRS has provided many important insights across a wide range of taxa from different ecosystems, e.g. with respect to population structure and demography, as well as hybridization, landscape or seascape genomics, QTL mapping, phylogeography and shallow phylogenies (e.g. Andrews et al., 2016; Fang et al., 2018; Franchini et al., 2014; Gaither et al., 2018; Pante et al., 2014; Ravinet et al., 2015; Xuereb et al., 2018).

In order to be effective and cost-efficient RRS experiments need to be well designed. First, it must be established whether the species of interest is in fact one species or if cryptic species are present. This can be problematic in non-model taxa and has potentially large downstream implications for RRS such as high divergence but few shared loci (Burns et al. 2017; Paris *et al.* 2017). A useful complement is therefore DNA barcoding to screen for cryptic species (Smith *et al.* 2011; Christiansen *et al.* Chapter 2). Alternatively, RRS can be specifically employed for species delimitation purposes (Ogden *et al.* 2013; Pante *et al.* 2014; Ceballos *et al.* 2019), but this should be a deliberate choice before designing the RRS setup. For such a scenario it would be especially important to sequence many fragments thereby increasing the likelihood of capturing genetic markers that are conserved across, yet discriminatory between species. In general, the research question fundamentally determines whether RRS is in fact necessary. For example, providing evidence for significant, evolutionary neutral genetic population structure may be easier and less expensive with a good number (>10) of multi-allelic microsatellites (Langin *et al.* 2018). However, RRS may be better suited to identify loci that are putatively affected by spatially variable selection and therefore involved in local adaptation. To this end, the density of markers (SNPs across the genome) that can be realized for a given species, which depends on genome size and complexity, as well as research budget, should be considered.

With low marker density one may run the risk of accepting unreasonably high rates of false positives (outliers that are not based on biological reality) in genome scans leading to biased or erroneous inferences (Whitlock & Lotterhos 2015; Hoban *et al.* 2016). Consequently, there is debate about the usefulness of RRS (or RADseq in particular), especially in taxa with little to no

genomic resources (Lowry *et al.* 2016; Catchen *et al.* 2017). The genomic characteristics of a target species, especially the genome size and the level of linkage disequilibrium (LD), are important to design a RRS experiment. With little genomic information, *a priori* calculations may be inaccurate. In order to avoid creating sequence data unsuitable to answer a study question and consequently inefficient resource use, it is therefore crucial to assess, optimize and critically ponder the advantages and limitations of applying RRS in a given research project. A most critical point in this context is to properly strike a balance between sequencing depth (coverage) and number of fragments, which is roughly proportional to the number of genetic markers. The estimated number of fragments generated from a genome is critical for the marker density (as the number of fragments translates relatively directly into the number of SNPs), while avoiding unnecessary "over"-sequencing of the genomic fragments, i.e. loci or RADtags, to save sequencing costs. Both excessive (> 100x) and uneven or too low (< 10x) coverage is detrimental for accurate locus reconstruction and SNP calling, particularly in *de novo* approaches (Rochette & Catchen 2017). Hence, RRS experimental procedures may benefit from thorough optimization. In this context, we used the framework of a large research project ("Refugia and Ecosystem Tolerance in the Southern Ocean") to optimize RRS for a diverse set of marine taxa in parallel. The Southern Ocean hosts a unique marine fauna with high levels of endemism (Rogers 2007; Crame 2018), but is increasingly subject to external pressures, such as warming, pollution and living resource exploitation (Aronson *et al.* 2011; Nicol *et al.* 2012; Griffiths *et al.* 2017; Mangano *et al.* 2017). Population genomic approaches are needed to understand the genetic structure and connectivity of Antarctic fauna, so that appropriate management and conservation actions can be developed (e.g. Younger *et al.* 2017; Clucas *et al.* 2018: Rintoul *et al.* 2018).

In this pilot experiment we seek to investigate and optimize the applicability of RRS to a range of Antarctic non-model taxa across the animal kingdom. The target organisms are ecologically important, abundant, and widely distributed in the Southern Ocean and cover a variety of habitats – from benthos to pelagic birds. Specifically, we aim to develop economic and robust experimental setups for RRS population genomic studies in an ostracod group (*Macroscapha opaca-tensa* species complex, Brandão et al., 2010), two amphipod species (*Charcotia obesa* Chevreux, 1906 and *Eusirus* aff. *perdentatus* Chevreux, 1912), two bivalve species (*Laternula elliptica* P.P. King, 1832 and *Aequiyoldia eightsii* Jay, 1839), two sea star species (*Bathybiaster loripes* Sladen, 1889 and *Psilaster charcoti* Koehler, 1906), two fish species (*Trematomus bernacchii* Boulenger, 1902 and *Trematomus loennbergii* Regan, 1913), and the two snow petrel

sub-species (*Pagodroma nivea nivea* Forster, 1777 and *Pagodroma nivea confusa* Clancey, Brooke & Sinclair, 1981). The outlined approach should be readily adoptable for other taxa of interest. We lay out a clear and concise protocol to follow *a priori* for any RRS experiment on non-model species that will help researchers to evaluate the costs, benefits and risks of such projects.

We specifically aim to (1) collate information about the genomic properties of the target taxa; (2) assess *in silico* which restriction enzymes are likely to yield the desired number of fragments; (3) test selected restriction enzyme digestions in the laboratory; (4) optimize restriction enzyme choice, size selection window and the number of individuals to be pooled per sequencing library (based on the previous results); and (5) sequence and analyze test RRS libraries of promising experimental setups. These extensive pilot analyses – including literature research, computational analyses and laboratory work – are designed to comprehensively evaluate all information for each target species or species complex. In the workflow of optimizing the setup for each target taxon, we strive to use the same restriction enzymes (or combinations) for several taxa whenever possible in order to reduce the costs for specifically designed barcodes and adaptors. Results shall ultimately facilitate informed decisions about whether and how RRS for each taxon could be conducted. We critically discuss these considerations and suggest alternative approaches in some cases.

## 2. Material and Methods

The optimization process of RRS experimental setups for non-model species is iterative and includes many deliberate choices that must be made based on the best available knowledge (Fig. 3.1). Relatively constant variables, i.e. the number and quality of samples, the research budget and the main research question, should be considered during the entire process and flexible variables, such as restriction enzymes, size selection window and the number of individuals to be pooled, should be adjusted to reach the desired outcome.

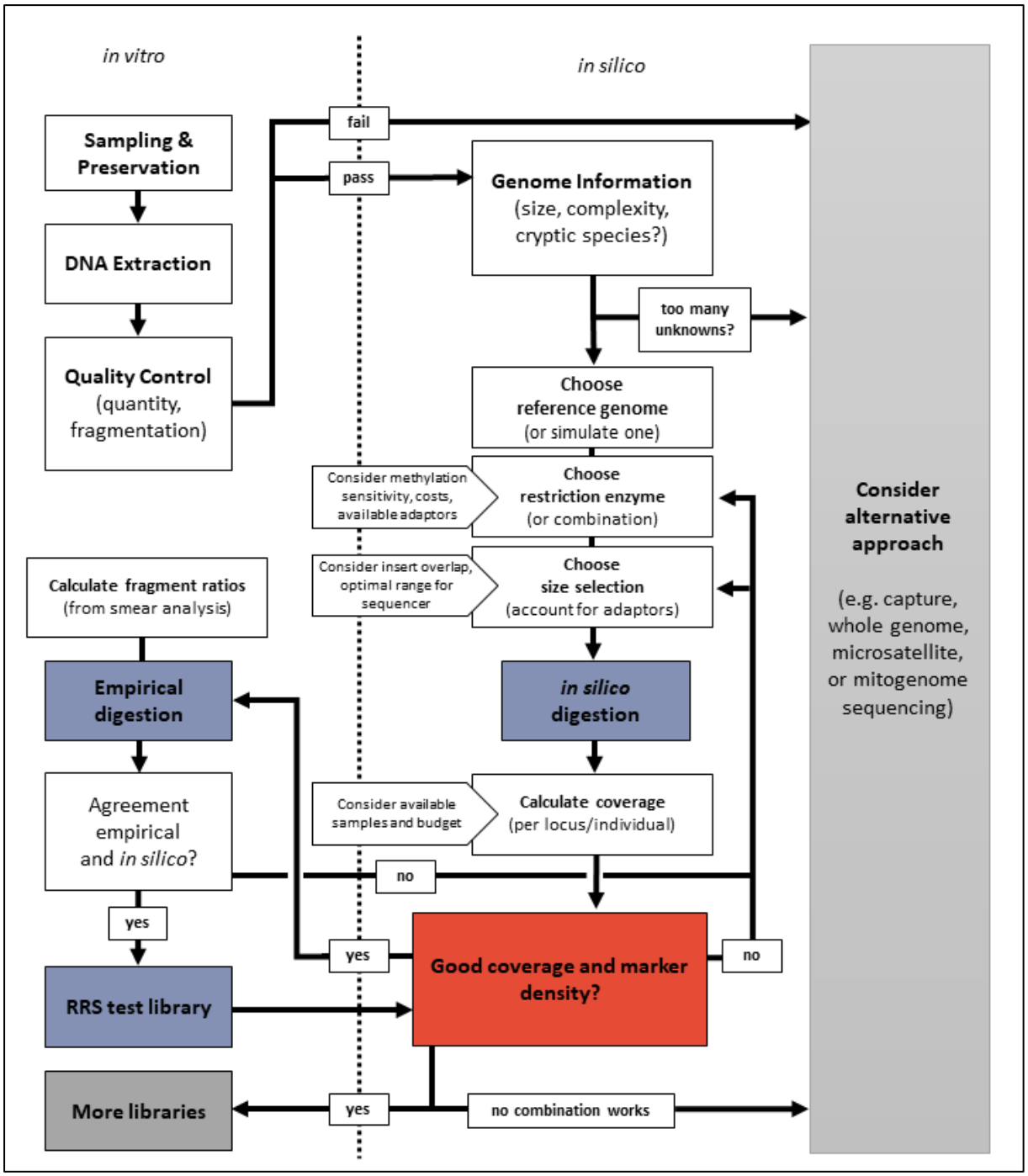



**Figure 3.1**. The iterative process of reduced representation sequencing (RRS) optimization with empirical (*in vitro*, left of the dotted line) and computational (*in silico*, right of dotted line) analyses. Core procedures to identify suitable experimental setups are *in silico* and empirical enzyme digestion and sequencing of a pilot RRS library (blue boxes). The coverage and marker density that can be achieved with a given setup needs to be repeatedly checked and fine-tuned (red box).

### 2.1 Specimen sampling

Samples of all target species were available from recent expeditions to the Southern Ocean (Supplementary Table S3.1). For ostracods, we used existing DNA extractions of Macrocyprididae from the Southern Ocean that were already taxonomically identified and described (Brandão 2010; Brandão *et al.* 2010). The amphipod target species were collected during RV *Polarstern* (AWI, 2017) expedition ANTXXIX-3 PS81. More details on *Eusirus* aff. *perdentatus* are provided in Verheye et al. (in prep.), while details of investigated *Charcotia obesa* are given in d'Udekem d'Acoz et al. (2018). The bivalves *Laternula elliptica* and *Aequiyoldia eightsii* were sampled by scuba diving in the shallow water of Potter Cove (King George Island, western Antarctic Peninsula; by F. Pasotti) and Rothera station (Adelaide Island, West Antarctic Peninsula; courtesy of the British Antarctic Survey) in 2016. Two sea star species (*Bathybiaster loripes* and *Psilaster charcoti*) were collected during international expeditions with RV *James Clark Ross* and RV *Polarstern* to the South Orkney Islands (JR15005 in 2016, PS77 in 2011), the Weddell Sea (PS81 in 2013), western Antarctic Peninsula (PS77 in 2011), and with RV *L'Astrolabe* to Adélie Land (REVOLTA 1 in 2010). Emerald rockcods (*Trematomus bernacchii*) were sampled in 2014 around James Ross Island with gill nets (Jurajda *et al.* 2016). Scaly rockcods (*Trematomus loennbergii*) were sampled in the Ross Sea as bycatch of the exploratory Antarctic toothfish (*Dissostichus mawsoni*) longline fishery. Dead birds and feathers of snow petrels (*Pagodroma* spp.) were sampled during the BELARE 2017-2018 expedition in the vicinity of the Princess Elisabeth Station, and additional samples were obtained from Signy and Adelaide Islands courtesy of the British Antarctic Survey. Samples were stored frozen, dried, or in >90% ethanol until DNA extraction.

### 2.2 Genomic resources

Prior to computational analyses, genomic information was collated for all target species or, if such information was not available, from the closest related species. Published reference genomes were collected from the literature and online resources, such as GenBank and Ensembl (Hubbard et al., 2002). In addition, genome size estimates were retrieved from genomesize.org (Gregory 2018) and other published estimates based on flow cytometry (e.g. Auvinet et al., 2018). Genome size estimates as C values were transformed to Mb for comparison (1 pg = 978 Mb; Dolezel et al., 2003).

### 2.3 *In silico* genome digestion analyses

We used SimRAD to computationally digest genomic DNA at sites matching a restriction enzyme recognition site (Lepais & Weir 2014). In total, seven restriction enzymes and combinations thereof were tested (Table 3.1). These were chosen based on what is commonly used in comparable studies and to cover a variety of enzymes ranging from very common (*MseI*, *MspI*, *ApeKI*) to medium (*EcoRI*, *SphI*, *PstI*) and rare cutters (*SbfI*). Reference genomes from related species as well as two simulated genomes per taxonomic class were used for these *in silico* digestions. Simulated genomes were modelled with GC content as in the available reference genome(s) and with two different sizes per taxonomic class to cover the approximate range of genome sizes known for this class (Table 3.2). The total number of fragments that these enzymes (or enzyme combinations for double digest setups) produced were estimated, as well as the number of fragments in size selection windows of 100 bp width (240-340 bp) and 50 bp width (210-260 bp), respectively. Approximate targets for the number of fragments in each species of interest were defined (Table 3.2) and restriction enzyme and size selection combinations that fell within the target range ± 10,000 fragments were retained for downstream testing. After narrowing down the enzyme choice and conducting empirical digestion analyses, we ran additional *in silico* digestions for a final optimization of the size window and thus number of fragments for each specific case.

### 2.4 Empirical genome digestion analyses

Laboratory experiments were conducted with promising restriction enzymes to verify results from computational analyses. For each species, DNA from three individuals was used to test two or three restriction enzymes or enzyme combinations. Genomic DNA was extracted using either the commercial DNA extraction kits NucleoSpin Tissue (Macherey-Nagel) or DNeasy Blood & Tissue (Qiagen) and following the manufacturer's guidelines, or with a standard salting out protocol (Cruz *et al.* 2017), or, for the bivalves, with a standard cetyl trimethylammonium bromide (CTAB) protocol. Subsequently, DNA quality and quantity were checked using the fluorescence assay Quant-iT PicoGreen dsDNA (Thermo Fisher Scientific Inc.), an Infinite M200 microplate reader (Tecan Group Ltd.) and 1 % agarose gel electrophoresis. Whenever possible, only high-quality DNA extractions were used.

**Table 3.1.** Restriction enzymes and combinations used for reduced representation sequencing (RRS) optimization. Recognition site, the approximate expected fragment number in a 1000 Mb genome, any special enzyme characteristic and empirical studies that recently used this enzyme (combination) are listed.

| **Restriction enzyme (combination)** | **Recognition site** | **Approximate Fragment number**† | **Special features** | **Citation** |
|---|---|---|---|---|
| *SbfI* | 5'--CCTGCA\|GG--3' | 6,000 | | e.g. Deagle et al., 2015; Hohenlohe et al., 2013; Ogden et al., 2013; Rodríguez-Ezpeleta et al., 2016 |
| *EcoRI* | 5'--G\|AATTC--3' | 320,000 | Methylation sensitive | Baird et al., 2008; Jacobsen et al., 2014 |
| *SphI* | 5'--GCATG\|C--3' | 150,000 | | |
| *PstI* | 5'--CTGCA\|G--3' | 150,000 | | Bolton et al., 2016; Herrera and Shank, 2016 |
| *ApeKI* | 5'--G\|CWGC--3' | 1,000,000 | Methylation sensitive, degenerate site | e.g. Grewe et al., 2017; Pérez-Portela et al., 2018; Puncher et al., 2018; Raeymaekers et al., 2017 |
| *MspI* | 5'--C\|CGG--3' | 1,700,000 | | |
| *MseI* | 5'--T\|TAA--3' | 7,600,000 | | Zhu et al., 2016 |
| *SbfI_SphI* | | 12,000 | | e.g. Adenyo et al., 2016; Cruz et al., 2017; Hillen et al., 2017; Maroso et al., 2018 |
| *SbfI_MspI* | | 12,000 | | Jacobsen et al., 2017; Leaché et al., 2017 |
| *PstI_MspI* | | 275,000 | | e.g. Bernatchez et al., 2016; Henning et al., 2017; Poland et al., 2012; Recknagel et al., 2013 |
| *EcoRI_SphI* | | 250,000 | | Nunziata et al., 2016 |
| *EcoRI_MspI* | | 540,000 | | e.g. Escoda et al., 2017; Lozier et al., 2016; Ng et al., 2017; Querejeta et al., 2016 |

† for a 1000 Mb genome with 40 % GC content and no size selection

**Table 3.2.** Genomic information useful for reduced representation sequencing (RRS) optimization in target species from six organisms classes. For each class approximate targets for the number of fragments (fragments target) were defined and genome size estimates from flow cytometry are listed. In species with unknown genome size, the range of published estimates from species from the same class is listed. Available genomes from related species and two simulated genomes per class were used for *in silico* digestions. The simulated genomes were modeled according to expectations regarding genome size and GC content as known from related species.

| Class | Target species | Fragments Target | Genome size (C) | Related species genome with size (Mb) | Simulated genomes |
|---|---|---|---|---|---|
| *Ostracoda* | Macrocyprididae | 50,000 | 0.17±0.003† | *Cyprideis torosa*, 286 Mb (Asexual Genome Consortium, in prep.) | 100 Mb, 43.9 % GC<br>500 Mb, 43.9 % GC |
| *Malacostraca* | *Charcotia obesa*<br>*Eusirus perdentatus* | 50,000 | unknown (Amphipoda: 0.68-64.62†) | *Hyalella Azteca*, 551 Mb (Poynton *et al.* 2018)<br>*Parhyale hawaiensis*, 4,003 Mb (Kao *et al.* 2016) | 10,000 Mb, 29.6 % GC<br>30,000 Mb, 29.6 % GC |
| *Bivalvia* | *Laternula elliptica*<br>*Aequiyoldia eightsii* | 50,000 | unknown (0.65-5.40†) | *Crassostrea gigas*, 558 Mb (Wang *et al.* 2012)<br>*Pinctada imbricata*, 991 Mb (Du *et al.* 2017)<br>*Bathymodiolus platifrons*, 1,658 Mb (Sun *et al.* 2017) | 1,000 Mb, 35.3 % GC<br>5,000 Mb, 30.2 % GC |
| *Asteroidea* | *Bathybiaster loripes*<br>*Psilaster charcoti* | 20,000 | unknown (Asteroidea: 0.54-0.96†) | *Acanthaster planci*, 383 Mb (Hall *et al.* 2017)<br>*Patiria miniata*, 811 Mb (Cameron *et al.* 2015)<br>*Patiriella regularis*, 949 Mb (Long *et al.* 2016) | 1,000 Mb, 40.2 % GC<br>2,000 Mb, 41.2 % GC |
| *Actinopterygii* | *Trematomus bernacchii*<br>*T. loennbergii* | 20,000 | 1.12±0.019‡; 1.19§; 1.82¶<br>1.34‡ | *Notothenia coriiceps*, 637 Mb (Shin *et al.* 2014) | 1,000 Mb, 40.8 % GC<br>1,800 Mb, 40.8 % GC |
| *Aves* | *Pagodroma nivea nivea*<br>*P. nivea confusa* | 50,000 | unknown (0.91-2.16†) | *Fulmarus glacialis*, 1,141 Mb (Zhang *et al.* 2014) | 1,500 Mb, 41.2 % GC<br>2,000 Mb, 41.2 % GC |

† published estimates from various species of the same class (or where indicated order), as listed on genomesize.com on 9th January 2019; Ostracoda: Macrocyprididae: Jeffery et al., 2017
‡ Auvinet et al., 2018 § Hardie and Hebert, 2003 ¶ Morescalchi et al., 1996

Because of their small size, extractions from individual ostracods yielded insufficient quantities of DNA for downstream protocols, and sample numbers per locality were very low. Hence, the entire genomic DNA of ostracods was amplified using the REPLI-G kit (Qiagen) for whole genome amplification of 1 µL extracted DNA with high-fidelity polymerase Phi 20 and multiple displacement amplification following the manufacturer's protocol. For this purpose, extractions with the highest DNA concentrations from different species of Macrocyprididae, mainly of the *Macroscapha tensa-opaca* species complex, were selected (Brandão *et al.* 2010). For all target species, 100 ng genomic DNA of three biological replicates per species was digested with 10 units of a selected restriction enzyme at 37 °C (*EcoRI*, *MspI* and *PstI*) or 75 °C (*ApeKI*) for 2 h in a total volume of 10 µL. Reactions were purified with CleanPCR (GC Biotech) according to the manufacturer's protocol. Between 1 and 5 ng of the purified digested DNA was loaded on a High Sensitivity DNA chip (Agilent Technologies) and run on an Agilent 2100 Bioanalyzer System. The 2100 expert software (Agilent) was used to conduct a "smear analysis" of the empirical digestions to estimate the amount of DNA in thirteen size bins, including the 210 – 260 and 240 – 340 bp windows. Results were then exported as CSV files and processed further in spreadsheet software. Because it is difficult to standardize the number of fragments in an empirical digest without knowledge of the genome size, we calculated the ratios of fragments in the two size windows mentioned above in comparison to the total number of fragments produced from the enzyme digestions. *In silico* and laboratory digests were then compared through these ratios.

#### 2.5 RRS setup optimization

In order to choose a promising restriction enzyme and size selection combination, we calculated the sequencing coverage per fragment as follows:

$$coverage = \frac{\sum sequencing\ reads}{\sum individuals \sum genomic\ fragments}$$

We conservatively aimed at a coverage of approximately 30x for each fragment per individual, higher than other minimum recommendations (Peterson *et al.* 2012; Paris *et al.* 2017; Rochette & Catchen 2017). Given that the accuracy of our genome size estimates is unknown, we aimed for relatively high coverage, so that in a "worst case scenario", where the genome size is actually twice as large as we estimate (or any other factor leads to twice as many fragments as assumed), we would still reach a coverage of approximately 15x. The number of individuals per sequencing library was set to 96, corresponding to one PCR plate. Sequencing with a HiSeq 4000 platform (Illumina) should conservatively yield approximately 300 million reads per sequencing lane, on

a HiSeq 2500 we expect approximately 200 million reads. These coverage calculations were applied to fragment numbers from *in silico* results based on available reference genomes and extrapolated to a final, conservative estimate of genome size based on the best available knowledge (Table 3.3). This extrapolation is likely not biologically accurate, but serves merely as a conservative correction measure here. We then used *in silico* estimates again to further tweak the size window of a chosen restriction enzyme or enzyme combination in each target species to achieve the desired coverage, while considering the size range in which the two HiSeq machines work best. Finally, we estimated the number of SNPs across the genome as a measure of marker density (analogous to Lowry *et al.* 2016) for a chosen enzyme and size selection setup and sequencing machine, assuming one SNP every other 100 bp. The latter estimate is based on our own experience, predominantly from fish genomes (see also e.g. Gao et al., 2018a).

### 2.6 RRS library preparation and sequencing

The information collected so far convinced us not to pursue RRS in amphipods (see discussion); they were therefore not included in the test libraries. In addition, not enough high molecular weight DNA samples of *P. nivea confusa* (one of the snow petrel subspecies) were available. Eventually, five RRS test libraries for eight target species were constructed using 6, 8, 10, or 14 individuals and two controls per species and sequenced on one lane of a HiSeq 2500 unit (see Table 3.4 in results section). That way, we attempted to keep the fixed variables for our coverage calculations, i.e. an estimated 250 million reads spread over 94 individuals and between 55,770 and 81,608 fragments. We originally aimed at 96 individuals, but too many samples of low-quality DNA dropped out during sample preparation. In addition, the estimated number of fragments varied between target species, but the conservative estimates in all other aspects should allow for some flexibility here. The libraries were all prepared by the same person at the KU Leuven laboratory using custom protocols that are based on two main references: the original ddRAD protocol by Peterson et al. (2012) and the original GBS protocol by Elshire et al. (2011). We adjusted these protocols slightly and provide a full-length description of the laboratory procedure in supplementary material S3.4 and S3.5. In both cases, the standardized high-quality DNA was first digested with restriction enzyme(s), followed by adaptor and barcode ligation, purification, PCR, another purification, and quantification and pooling. The libraries were then sent to the KU Leuven Genomics Core (genomicscore.be), where all five libraries were individually size selected on a Pippin Prep unit (Sage Science), checked for quantity using qPCR, pooled, and paired-end sequenced on one lane of a HiSeq 2500 platform (Illumina).

**Table 3.3.** Reduced representation sequencing (RRS) setups for seven individually optimized protocols to be run on a HiSeq 2500 platform (Illumina). Choice of restriction enzyme(s) and size window was optimized in order to obtain approximately 30x coverage (or half that in a worst-case scenario) with the assumed genome size (conservatively estimated based on available information, see Table 2). Marker density (the number of bp per sequenced SNP) was estimated as a comparable measure to the metastudy by Lowry et al. (2016).

| Class | Target Species | Restriction Enzyme (Combination) | Size Window (bp) | Assumed Genome Size (Mb) | Coverage† | Marker Density† (bp per 1 SNP) |
|---|---|---|---|---|---|---|
| *Ostracoda* | Macrocyprididae | *ApeKI* | 200-350 | 250 | 31.9x | 1,533 |
| *Malacostraca* | *Charcotia obesa* | *SbfI_MspI* | 200-330 | 27,000 | 30.3x | 157,092 |
| | *Eusirus perdentatus* | *EcoRI_SphI* | 200-260 | 7,000 | 32.8x | 44,036 |
| *Bivalvia* | *Laternula elliptica*<br>*Aequiyoldia eightsii* | *ApeKI* | 200-260 | 3,000 | 30.2 – 7.4x | 17,383 – 21,517 |
| *Asteroidea* | *Bathybiaster loripes*<br>*Psilaster charcoti* | *ApeKI* | 200-300 | 500 | 27.1 – 33.9x | 2,598 – 3,253 |
| *Actinopterygii* | *Trematomus bernacchii*<br>*T. loennbergii* | *EcoRI_MspI* | 200-450 | 1,500 | 27.5x | 7,352 |
| *Aves* | *Pagodroma nivea nivea*<br>*P. nivea confusa* | *PstI* | 200-300 | 1,500 | 31.0x | 8,941 |

† assuming 200 million reads of 125 bp length spread over 96 individuals and 0.01 SNP/bp

### 2.7 Sequence analyses

Sequencing data were checked using FastQC v0.11.5 (Andrews 2010) and then demultiplexed and cleaned (options -c and -q) using the process_radtags module of Stacks v2.4 (Catchen *et al.* 2011, 2013). Because some of our multiplexing barcodes for the *PstI* library were contained in longer *ApeKI* barcodes, we demultiplexed the *ApeKI* libraries first and captured reads that were discarded in the process. These reads were subsequently used for demultiplexing of the *PstI* library. All demultiplexing runs were conducted without barcode rescue to avoid cross-contamination between libraries. The Stacks pipeline was also used for each target species independently to create a *de novo* assembly and call genotypes. Building contigs from paired-end reads is not possible with GBS data in Stacks (Rochette *et al.* 2019), because the orientation of the reads is ambiguous. In this case (libraries 1, 2, 3, 5), we concatenated the four output files

per individual of process_radtags to run the pipeline as if it was single-end data. Our size selection windows were designed to avoid overlap between the two reads of one fragment, so this approach should work well, albeit creating shorter haplotypes. We used Stacks' default value for *m*, i.e., a minimum stack depth coverage of 3, which generally produces consistent results at typical coverage rates (Paris *et al.* 2017). Choosing parameters *M* and *n* to control the formation of loci within and across individuals on the other hand is study dependent. We explored a parameter range of $n = M = [1 .. 9]$ following Rochette and Catchen (2017) to strike a balance between over- and undermerging alleles and loci. To compare results from the different parameters only loci present in 80 % of the samples (50 % in the case of ostracods) were retained. Further detailed filtering would be required for downstream population genomic analyses.

## 3. Results

### 3.1 Genome characteristics

The genomic resources and knowledge are highly variable across target taxa, with typically more genomic information available for vertebrate groups (Table 3.2). Genome size in ostracods varies with Macrocyprididae estimated at approximately 166 Mb (or 0.17 C, Jeffery *et al.* 2017). One ostracod species with a genome of comparable size was available. Amphipods show very large variability in genome size (Ritchie *et al.* 2017) with extreme cases that dramatically exceed the size estimates of all other target taxa studied here (up to 63,198 Mb or 64.62 C, Table 3.2). Two amphipod reference genomes were available. For bivalves and sea stars more reference genomes were available, but not from species closely related to our target species. In both cases, we selected three reference genomes of varying size (Table 3.2). The Antarctic fish target species have genome size estimates available as well as a reference genome from a related species (same family, Nototheniidae). In birds, no genome size estimates for our target species were available, but bird genome size appears to be relatively constrained between approximately 1 and 2 Gb and a reference genome from the same family (Procellariidae) was available. We decided to aim at 50,000 fragments as initial targets for our optimizations in all taxa, except for fish and sea stars (Table 3.2). In the latter target taxa, we aimed at 20,000 fragments initially, because we had more samples available and thus were interested in covering more individual samples from a wider geographic range at the expense of marker density. Note that these targets are highly study specific and depend on the budget, number of samples to be sequenced and, most importantly, exact research question of a given RRS project.

### 3.2 *In silico* digestions

We estimated how many RRS fragments twelve restriction enzymes and enzyme combinations (listed in Table 3.1) would produce. These estimates were conducted using various reference genomes and also simulated genomes. We estimated the fragment number in total as well as in two size selection windows of 240 - 340 bp and 210 - 260 bp, respectively (Supplementary Table S3.2). As expected, the fragment number is influenced primarily by the type of enzyme used and the size of the genome digested. The tested combinations produced fragment numbers close to our defined targets in all species, but with different restriction enzymes. We aimed at using as few different enzyme setups across species as possible. Using the same setup for several RRS experiments reduces costs as the same adaptor sets can be re-used multiple times. Therefore, we kept five initial setups that yielded promising fragment numbers: *EcoRI*, *PstI*, *ApeKI*, *MspI* and a double digest with *EcoRI* and *MspI*.

### 3.3 Empirical digestions

Based on the preliminary *in silico* results, we tested the genome digestion by these enzymes and enzyme combinations in the laboratory. High quality bird DNA was not available, hence preventing empirical digestion tests for this group. Ostracod DNA was whole genome amplified and this proved problematic for the Bioanalyzer, because the results indicated overloading even after multiple dilutions. In total, 75 empirical digestions were conducted, several of which produced unusable results even after repeating the experiment. With only small irregularities in the curves produced by the Bioanalyzer it was impossible to infer accurate estimates of the number of fragments, especially in size ranges below 500 bp. From the successful runs it appeared that, in general, in most cases, the ratio of fragments produced empirically was higher than estimated *in silico.* Bivalve genomes for example produced approximately twice as many fragments as expected (Fig. 3.2), while in sea stars and fish the increase was less pronounced (Fig. 3.3 & Supp. Fig. S3.1).

### 3.4 RRS setup

With all information gathered thus far, we proceeded to optimize the RRS experimental setup for each of our target taxa. We planned the same setup for species from the same class, when the genomic differences between those species were unknown (in Bivalvia and Asteroidea), or when they were related and therefore likely to have similar genomic properties (Actinopterygii

and Aves). In contrast, we designed two different setups in Malacostraca, because the genomes of *C. obesa* and *E. perdentatus* may have very different sizes (Table 3.2 & 3.3). Experimental setups, i.e. restriction enzymes and size selection window, were furthermore tuned to suit a sequencing experiment with HiSeq 2500 or 4000, respectively. The choice of sequencing platform can be modified based on instrument availability and budget. In the following, results for use with a HiSeq 2500 are listed (Table 3.3), the same results for a HiSeq 4000 can be found in Supplementary Table S3. The setup for optimizing results as listed here also includes the consideration that it would be cost-efficient to use the same enzyme or enzyme combinations for several species whenever possible, because adaptors can then be re-used. Therefore, when several enzyme (combinations) seemed promising according to *in silico* digestion, we attempted to choose setups that were also promising in other target species. For ostracods, we assumed a genome size of 250 Mb and 500 Mb as worst-case scenario. Using the *Cyprideis torosa* reference genome, a digest with *ApeKI* and size selection of 200 – 350 bp would yield 31.9x coverage (or half of that in the worst-case scenario). With this setup and genome size, we would achieve a marker density of approximately one SNP every 1.5 kb. In amphipods, different setups per species are required. Given the highly uncertain genome size of 27,000 Mb for *C. obesa* and 7,000 Mb for *E. perdentatus* (based on same family estimates; Rees et al., 2007), double digest RADseq experiments with *SbfI* and *MspI* and *EcoRI* and *SphI*, respectively, would yield the desired coverage. Marker density in both cases is expected to be low, due to the large genome size (Table 3.3). Because of uncertainty with respect to genome size and an anticipated low marker density, we stopped RRS optimization in amphipods and instead explored alternatives as envisioned in our workflow (Fig. 3.1). For both bivalve species, a genome digestion with *ApeKI* and size selection of 200 – 260 bp seemed promising with all three reference genomes and would yield around one SNP per 20 kb. Similarly, in sea stars we found setups with *ApeKI* and a slightly wider size selection that should yield good results, although results varied depending on the reference genome used. For the Antarctic fishes of the genus *Trematomus*, a double digest setup with *EcoRI* and *MspI* in a size window of 250 – 450 bp should yield desired coverage and marker density. Regarding the snow petrels, a setup with *PstI* and 200 – 300 bp size selection seemed appropriate, yielding one SNP every 9 kb. Overall, results indicate that with only three enzyme choices, it should be possible to achieve the desired coverage and marker density in eight species (Table 3.4).

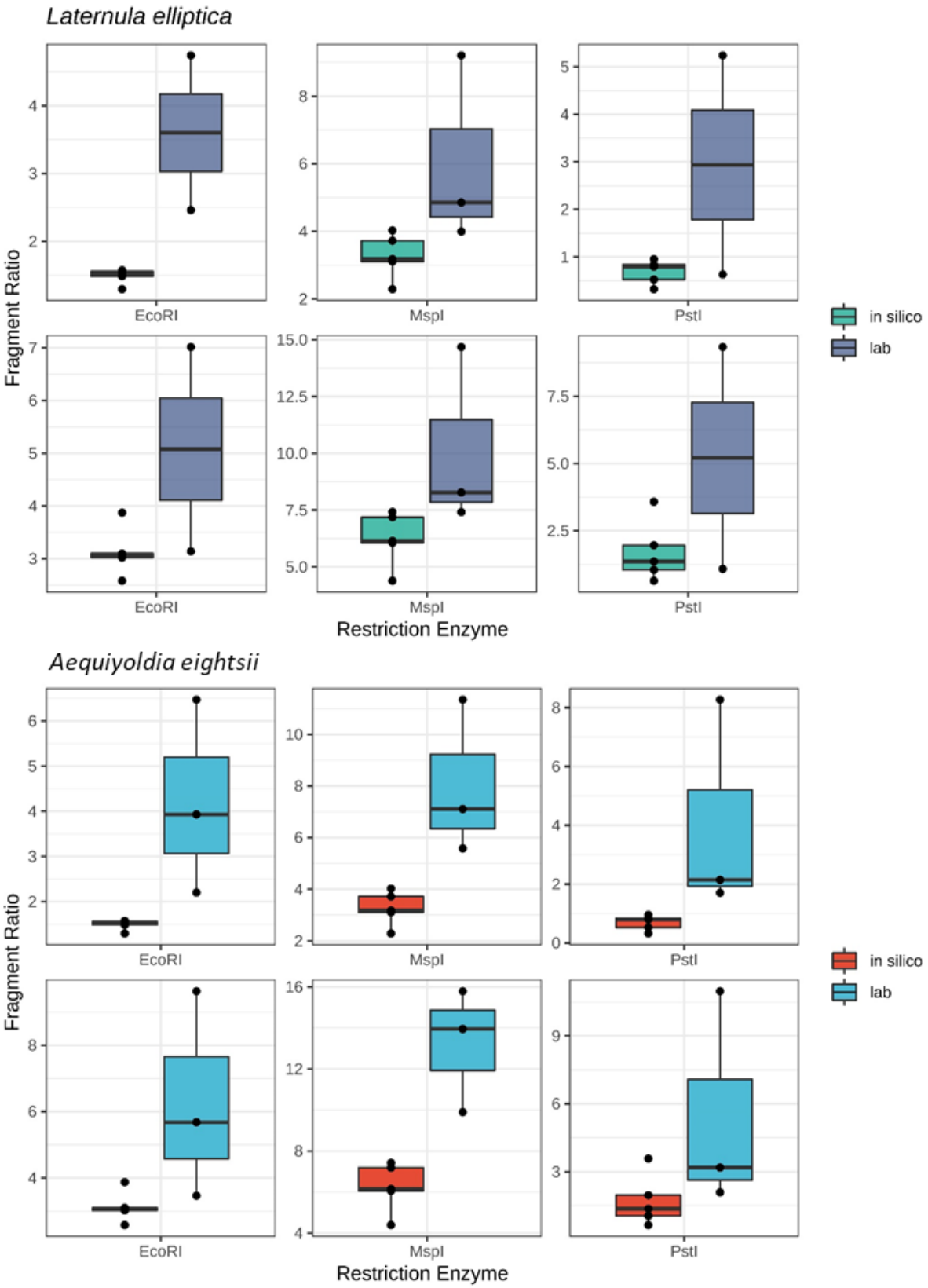



**Figure 3.2.** The ratio of fragments produced by three restriction enzymes when digesting bivalve genomes *in silico* and empirically. The ratio is calculated as the number of fragments between 210 – 260 bp (upper panel) and 240 – 340 bp (lower panel). The computational estimates include results from different reference genomes from related species (as listed in Table 3.2), as well as simulated genomes. The empirical estimates are based on three biological replicates; if fewer, the experiments failed repeatedly.

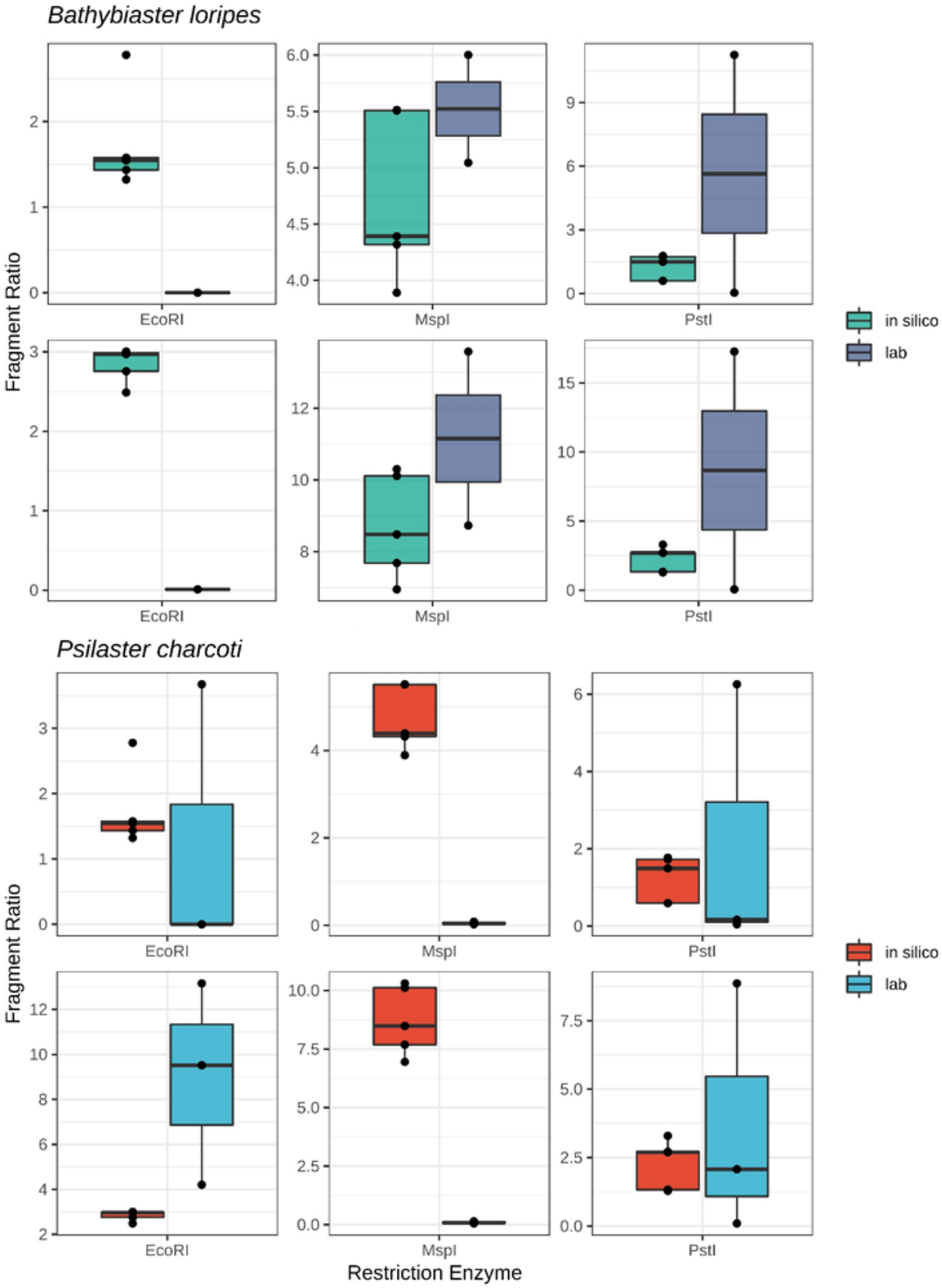



**Figure 3.3.** The ratio of fragments produced by three restriction enzymes when digesting sea star genomes *in silico* and empirically. The ratio is calculated as the number of fragments between 210 – 260 bp (upper panel) and 240 – 340 bp (lower panel). The computational estimates include results from different reference genomes from related species (as listed in Table 3.2), as well as simulated genomes. The empirical estimates are based on three biological replicates; if fewer, the experiments failed repeatedly.

### 3.5 RRS test libraries

Pilot libraries with optimized setups were sequenced, yielding a total of 531 million (M) reads. After demultiplexing and quality control, 422 M reads were retained. These reads were spread relatively evenly across libraries, species, and individuals (average and standard deviation across all taxa and libraries: 4.5 ± 2.1 M reads). All but five individuals received more than 1 M reads and most individuals received more than 3 M reads. We created *de novo* catalogs from these reads using Stacks with varying M and n parameters were created. Optimal parameters varied from 3 to 6 among taxa (Table 3.4 & Supplementary Material S3.6). Results from this parameter optimization also revealed varying levels of diversity, e.g. sea stars showed relatively high polymorphism levels, while the bird library produced many loci but few SNPs (Supplementary Material S3.6). Comparing the unfiltered numbers of loci and coverage across individuals underlined the inverse relation of these two variables (Table 3.4, Fig. 3.4). In ostracods, our target estimates were matched best. In bivalves and sea stars more loci than expected were sequenced at the expense of coverage, although coverage was still reasonable. Two individuals of *B. loripes* had low coverage due to low initial numbers of reads indicating errors during library preparation or degraded input DNA. The fish libraries contained considerably less loci than expected at high coverage, while the opposite was true for the bird library. The latter also showed very uniform low coverage at approximately 10x. Overall, these results show promise for full scale RRS libraries with sufficiently high coverage in four of five libraries.

**Table 3.4.** Setup and results of five test libraries (Nr.) for reduced representation sequencing (RRS) containing samples from seven species and one species group. RRS protocol, restriction enzyme(s) and size window (setup) is listed as well as the number of samples and controls used (N) and the Stacks parameters *M* and *n* used for bioinformatics (SP). The setup was optimized for number of fragments and coverage as in Table 3.3, these estimates are listed here again as expected values and compared to empirical results regarding average (and standard deviation of) number of loci and average (and standard deviation of) coverage based on data processing using Stacks v2.4 with optimized parameters (SP).

| Nr. | Class | Target Species | Setup | N | SP | Expected Fragments | Obtained Loci† | Expected Coverage | Obtained Coverage† |
|---|---|---|---|---|---|---|---|---|---|
| 1 | *Ostracoda* | Macrocyprididae | GBS, *ApeKI*, 200-350 | 8 + 2 | 6 | 65,245 | 69,817 (±63,114) | 31.9x | 28.2x (±5.4x) |
| 2 | *Bivalvia* | *L. elliptica* | GBS, *ApeKI*, 200-260 | 8 + 2 | 4 | 55,770–69,032 | 125,305 (±22,828) | 30.2–37.4x | 21.6x (±5.1x) |
| | | *A. eightsii* | | 10 + 2 | | | 143,551 (±28,676) | | 20.0x (±2.6x) |
| 3 | *Asteroidea* | *B. loripes* | GBS, *ApeKI* 200-300 | 10 + 2 | 5 | 61,489–76,990 | 82,945 (±43,521) | 27.1–33.9x | 21.0x (±6.8x) |
| | | *P. charcoti* | | 14 + 2 | | | 115,608 (±30,589) | | 27.6x (±4.6x) |
| 4 | *Actinopterygii* | *T. bernacchii* | ddRAD, *EcoRI_MspI*, 200-450 | 10 + 2 | 3 | 81,608 | 21,121 (±3,539) | 27.5x | 42.3x (±13.5x) |
| | | *T. loennbergii* | | 10 + 2 | | | 23,609 (±2,362) | | 49.6x (±6.2x) |
| 5 | *Aves* | *P. nivea nivea* | GBS, *PstI*, 200-300 | 6 + 2 | 3 | 67,104 | 140,972 (±26,444) | 31.0x | 10.0x (±0.4x) |

† using denovo_map.pl with m=3 and M=n as listed in column six

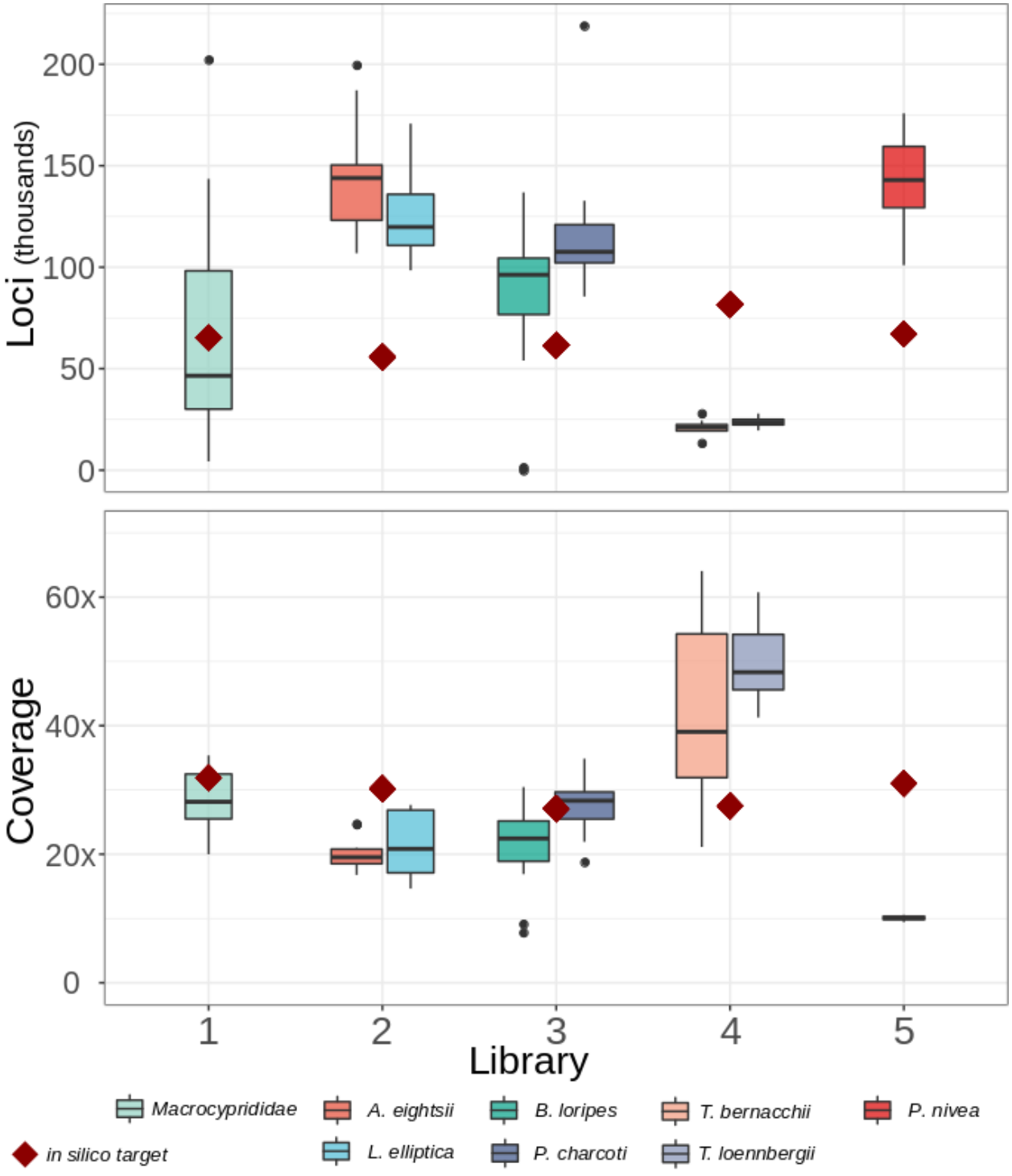



**Fig. 3.4.** The number of loci and coverage as estimated *in silico* (dark red diamonds) and realized in five test reduced representation sequencing (RRS) libraries containing DNA from eight target species. Boxplots show the median, quartiles, and outliers across individuals (N=8-14). Libraries were prepared as listed in Table 3.4.

## 4. Discussion

High-throughput sequencing methods promise new avenues of ecological and evolutionary research in non-model organisms. We provide a detailed workflow to evaluate and optimize reduced representation sequencing (RRS) techniques for any animal species of interest (Fig. 1). This approach is reproducible and ensures that researchers are well-informed about the advantages and drawbacks of RRS for their research question. Different RRS setups (i.e. various species and libraries constructed via different protocols, enzymes and size selection windows) were successfully sequenced together on one HiSeq lane. Most individuals included in this multi-library-multiplex received adequate sequencing effort, which has been problematic in other studies that pooled individuals directly after ligation (Maroso *et al.* 2018). From our experience (including this and previous studies in our laboratory; see e.g. Hillen *et al.* 2017; Raeymaekers *et al.* 2017; Maroso *et al.* 2018) it seems, that careful, repeated quantification and standardization of DNA from every individual before and after PCR are key to achieve equivalent sequencing effort across individuals. A pilot sequencing experiment can then yield valuable insights before proceeding with sequencing rounds at larger scale. Here, more loci than expected were assembled in most taxa (ostracods, bivalves, sea stars) at sufficiently high per locus coverage. This highlights the value of choosing parameters conservatively, e.g. under- rather than over-estimating the number of sequencing reads. The fish library yielded fewer loci than expected at higher coverage. Pooling more individuals, increasing the size window or changing the restriction enzyme setup altogether including new optimization are future options to further optimize this project, although the current setup also yields useful data. The bird library produced coverage that is directly at the advised limit of 10x (Rochette & Catchen 2017). This may be partly related to low quality input DNA, which in this case was mostly extracted from feathers. Alternative sampling and/or DNA extraction protocols and further testing are needed before sequencing full scale libraries for snow petrels. Overall, a few key properties determine the feasibility and cost of RRS in non-model organisms.

### 4.1 Predictability of reduced representation experiments

Planning a genome reduction through restriction enzyme digestion starts with an imperative question: how large is the target genome? Non-model species often lack information about genome size, which complicates RRS optimization (Herrera *et al.* 2015). If genome size appears relatively conserved across species within a taxonomic class (e.g. sea stars), it can be assumed that the species of interest from this class has similar genome size. Some imprecision regarding

the exact size has then only limited effects on overall accuracy. Alternatively, in other groups, such as amphipods, genome size is highly variable, spanning two orders of magnitudes (Krapp *et al.* 2006; Rees *et al.* 2007; Ritchie *et al.* 2017). In this case using an inaccurate genome size estimate has the potential to dramatically impact the parameters one aims to optimize. In addition, very large genomes are often highly repetitive, which significantly hampers downstream bioinformatics and population genomics (Deagle *et al.* 2015; McCartney-Melstad *et al.* 2016; Star *et al.* 2016). Therefore, with the current state of knowledge, we opted to exclude amphipods from our trial RRS libraries. Previously conducted ddRAD sequencing of *C. obesa* provided indeed very few useful genomic markers (Havermans et al., unpublished data). Estimating genome size with flow cytometry or conducting a series of test libraries could be alternative ways forward.

For ostracods, bivalves, sea stars, and birds more loci were found than expected. This could indicate that genome sizes were consistently larger than expected. Another, likely explanation in this case is that the used enzymes (*ApeKI*, *PstI*) produce more fragments than *in silico* digestions predict. For example, the number of fragments resulting from four base cutters may be more difficult to predict as they can produce so many fragments that effectively the entire genome would be sequenced (Herrera *et al.* 2015). The five-base recognition site of *ApeKI* features a degenerate base, which may have a similar effect. The methylation sensitivity of *ApeKI* may also provide more genomic markers in genic regions (Pootakham *et al.* 2016). It is unclear, however, how general this prediction holds across metazoans. Finally, some of the excess loci recovered may be artefacts from library preparation, PCR duplicates or incorrect locus assembly (Rochette *et al.* 2019). Rigorous downstream filtering and/or comparison of several, differently filtered data sets aids in the determination of the true biological signal. Whatever the reason, the higher than expected number of loci still led to sufficient coverage, except in the bird library. The latter is likely related to low quality/quantity of input DNA. Few bird samples were available to us, some of which only as feathers, which yielded very little DNA. Whole genome amplification (WGA) could be an option to increase yield for RRS as successfully applied in ostrocods (this study) and insects (de Medeiros & Farrell 2018).

Finally, even with reliable genome size estimates and well-tested enzymes, the empirical results may differ from *in silico* expectations. In *Trematomus* fishes, approximately half of the expected sites were found, despite well-known genome size (Morescalchi *et al.* 1996; Hardie & Hebert

2003; Auvinet *et al.* 2018). In this case, genomic architecture may play an important role in affecting the number of cut sites per restriction enzyme. We used the draft genome of a related species from the same family to estimate the number of fragments. The endemic Antarctic notothenioid fishes, however, are characterized by frequent chromosomal rearrangements and large numbers of transposable elements (Ghigliotti *et al.* 2015; Auvinet *et al.* 2018; Kim *et al.* 2019b). The genus *Trematomus* constitutes an example of a relatively recent marine adaptive radiation (Near *et al.* 2012; Lautrédou *et al.* 2012). Therefore, in this particular case, the genome of a closely related species may provide relatively poor accuracy for cut site estimations.

We have tested various enzymes and enzyme combinations that have been successfully used in RRS studies (Table 1). Yet, many previous studies achieved overall relatively little marker density, which is problematic if looking for genome-wide adaptation patterns (Lowry *et al.* 2016). With increasing output of sequencers, aiming at higher marker density is not an unachievable goal. Genome size, restriction enzyme characteristics and genomic complexity can all have effects on predictability. Altogether, our results highlight the importance of conducting test libraries before embarking on larger, multi-library sequencing projects. In our case, *ApeKI* together with a narrow size window seems robust and powerful to create many genomic fragments (and thus sufficiently high marker density) across taxa with small to medium genome size. Using the same restriction enzyme for several projects drastically reduces cost as the same custom-made barcodes and adaptors can be used.

#### 4.2 Decision making for population genomics

As we illustrated here, there are many experimental choices that may lead to inefficient or "broken" (Lowry *et al.* 2016) RRS experiments. Given the publication bias towards successful applications (Sánchez-Tójar *et al.* 2018), it is likely that a large number of unsuccessful applications of this technology to non-model species exist. It is crucially important that researchers actively engage in the decision-making process when choosing restriction enzymes, size selection windows, and the number of individuals to be pooled per sequencing lane. Furthermore, the research objectives and budget should be critically evaluated and matched. In other words, investigating genome-wide polygenic adaptation patterns in a non-model species with large, complex genomes may simply be not feasible on a small budget. The number of individuals to be included is another aspect that weighs in on these considerations. In situations where sampling is not restricted, inferences of spatial genetic structure for example may benefit

more from wider geographic sampling coverage than from higher marker density. If sampling more localities is unfeasible as may be the case in the Antarctic realm, it can be beneficial to instead invest in high density sequencing (as in several markers per block of linkage disequilibrium). With sufficient genome coverage even advanced coalescent modeling is possible using RRS data (Liu & Hansen 2017).

We recommend following a few guiding principles when planning RRS for population genomics. First, clear targets with respect to the number of individuals to be screened in a project (and/or in follow-up projects) and the marker density needed for the research objective should be defined. Second, *in silico* estimations of how these targets can be reached and approximations of the associated costs should be obtained. The number of markers and individuals must be matched to reach a certain coverage (e.g. an average target of 30x). Subsequently, it is useful to briefly evaluate the trade-offs and benefits of RRS and other methods. If a promising combination of RRS method, enzyme, size selection, and sequencing effort is found, it is often worthwhile to conduct a pilot experiment before running the full sequencing experiment. However, it is also advisable to stick to one approach afterwards and not change for example the sequencing platform, the size window or other properties of the setup that will otherwise reduce comparability between data sets.

### 4.3 Alternative approaches

In some cases, RRS might not be the right choice for molecular ecological research. A plethora of other genomic or genetic methods exists, which may offer more appropriate cost-benefit ratios. SNP genotyping arrays are a common and highly reproducible alternative, but usually only for species with more genomic resources (possibly also in the Antarctic; see Humble et al., 2018). Similarly, whole genome resequencing is providing the most extensive data sets which can be used for a wide range of analyses (Barrio *et al.* 2016; Fuentes-Pardo & Ruzzante 2017; Therkildsen & Palumbi 2017). However, this is still too costly for many research projects, especially if information across many individuals and/or localities is needed. Another option is to focus on the expressed part of the genome and use a form of sequence capture enrichment (e.g. Hoffberg et al., 2016; McCartney-Melstad et al., 2016; Puritz and Lotterhos, 2018) or RNAseq (De Wit *et al.* 2015), or both (Schmid *et al.* 2016; Linck *et al.* 2017). These approaches are versatile and can provide valuable information, even for museum samples (Bi *et al.* 2013; Li *et al.* 2015). However, substantial expertise and prior investment in custom method

development is necessary for species that have not been analyzed yet. With a limited budget and research objectives that do not depend on whole genome scans for selection, more classical molecular approaches are sometimes a good alternative. Nuclear microsatellite markers remain powerful to describe population structure and can now be multiplexed and screened in large numbers. These markers can also benefit from high-throughput sequencing (Vartia *et al.* 2016; Bradbury *et al.* 2018). Mitogenome sequencing and assembly using long-range PCR is another useful approach, particularly for phylogeographic applications (Teacher *et al.* 2012; Lait *et al.* 2018). The amphipod and bird species evaluated here may currently be more amenable to such methods instead of RRS.

## 5. Conclusion

An extensive evaluation and optimization protocol allowed us to identify whether RRS is a suitable option for population genomics in a range of Antarctic animals. We have achieved promising results in some classes (ostracods, bivalves, sea stars, and fishes) that will be further developed in the near future. In other cases (amphipods and birds/degraded samples) alternative strategies such as mitogenome, capture sequencing or microsatellites seem more appropriate. The detailed considerations outlined here are a guideline for researchers to take informed decisions about the use of RRS or alternative methods. This is particularly important for species where genomic information remains scarce.

## 6. Acknowledgements

We are very grateful to everyone that contributed to field work and provided samples. In particular, we thank S. Brandão for ostracod DNA and H. Griffiths and K. Linse (BAS), C. Held (AWI), S. Mills and D. Macpherson (NIWA), H. Baird and G. Johnstone (AAD), and the MNHN in Paris for amphipod samples. We thank M. Clark (BAS) and the scientific staff involved for the generous collection of bivalve samples from Rothera station. We thank the captains, crew and science leaders who contributed to the collection of sea stars from RRS James Clark Ross (JR15005) and RV Polarstern (PS77, PS81). We thank the scientific team of IPEV (Institut Polaire Français Paul-Emile Victor) program 1124 REVOLTA (Radiation EVOLutives en Terre Adélie; chief scientists M. Eléaume, N. Améziane and G. Lecointre). For fishes, we thank K. Roche and P. Jurajda that provided fin clips from James Ross Island. The toothfish fishery observers and S.

Parker (NIWA) are acknowledged for fish samples from the Ross Sea. We thank G.E. Maes and K. Herten from the KU Leuven Genomics Core for technical support and advice.

## 7. Funding

This is contribution no. 35 to the vERSO project, funded by the Belgian Science Policy Office (BELSPO, Contract no. BR/132/A1/vERSO) and contribution no. 8 to the RECTO project (BELSPO, Contract no. BR/154/A1/RECTO). TP thanks BELSPO for funding JPIO Mining Impact. Research was also supported by the Scientific Research Network "Eco-evolutionary dynamics in natural and anthropogenic communities" (grant W0.037.10N) and the European Marine Biological Resource Centre (EMBRC) Belgium, both funded by the Research Foundation – Flanders (FWO). The first author was supported by an individual grant from the former Flemish Agency for Innovation by Science and Technology, now managed through Flanders Innovation & Entrepreneurship (VLAIO, grant no. 141328).

## 8. Data archiving statement

Demultiplexed, but otherwise raw sequencing data are deposited on NCBI's Sequence Read Archive (SRA). This data will also be cross-linked and thus discoverable in GBIF and OBIS through biodiversity.aq. In addition, spreadsheet tables with the data and calculations regarding fragment ratios, coverage and marker density and R scripts for in silico analyses and plotting are available at https://doi.org/10.5281/zenodo.3267164.

# CHAPTER 4: Genomic differentiation and local adaptation associated with contrasting climate change in an Antarctic fish

*Note:* this manuscript is in preparation for resubmission.

## Genomic differentiation and local adaptation associated with contrasting climate change in an Antarctic fish

Henrik Christiansen, Franz. M. Heindler, Anton P. Van de Putte and Filip A. M. Volckaert

**Abstract**

Genetic differentiation depends on levels of gene flow and local adaptation and provides crucial information for conservation. The Antarctic environment is host to a unique fish fauna, but these species face rapid and locally opposing environmental change. West Antarctic fish populations are exposed to warming and sea ice reduction, while the sea ice cover recently increased in parts of East Antarctica. Here, we apply double-digest restriction-site associated DNA sequencing to investigate the genomic diversity and associations with environmental variables of the first Antarctic vertebrate with a sequenced draft genome. We genotype loci using both a *de novo* and reference genome mapping approach and compare these methods. Our data reveal that *Notothenia coriiceps*, an ecologically important and widely studied fish species of the Southern Ocean, is composed of at least two subtly ($F_{ST}$ up to 0.0239) differentiated sub-populations. While high connectivity across vast distances is realized through dispersal via ocean currents, we also observe genotype-environment associations that indicate some degree of local adaptation. In addition, genome scans for selection show that some loci are likely subject to recent, spatially-variable selection. One putatively adaptive SNP may be linked to a gene important for methionine biosynthesis and is correlated with ice cover. We also detect selection signatures in SNPs near anti-freeze glycoprotein genes. These loci should be investigated further, but already show that important genomic variability is distributed unevenly across the Southern Ocean. Conservation measures, such as marine protected areas, should consider these results to support population resilience and adaptation potential. Furthermore, we suggest that using *de novo* SNP calling, even when a reference genome is available, can provide additional insights, for example through loci that are otherwise not recognized due to spatial ascertainment bias.

## 1. Introduction

Natural populations are genetically heterogeneous. This manifests in sometimes pronounced differences in the genetic make-up of members of the same species that may result from random evolutionary processes, i.e., genetic drift, or more directed evolutionary action, such as natural selection creating local adaptation (Savolainen *et al.* 2013). Gene flow on the other hand promotes genetic homogeneity and is enhanced through increased larval dispersal and adult mobility (Bohonak 1999; Damerau *et al.* 2012; Tigano & Friesen 2016). Consequently, in marine fishes genetic differentiation was long thought to be absent or insignificant due to few geographic barriers in the ocean (Hauser & Carvalho 2008). Recently, the importance and prevalence of spatially variable adaptation patterns in the marine environment is increasingly recognized, concomitantly with advances in sequencing technology and a shift towards assessing adaptive genetic variation (Nielsen *et al.* 2009; Pardo-Diaz *et al.* 2015; Pearse 2016). Genetic divergence occurring at only few adaptive loci was revealed in various marine fish species, often in correlation with environmental variables such as ambient salinity (Lamichhaney *et al.* 2012; Teacher *et al.* 2013; Milano *et al.* 2014; Vandamme *et al.* 2014; Berg *et al.* 2015). Overall, levels of genetic diversity are reduced through harvesting, such as fisheries (Pinsky & Palumbi 2014), which also imposes artificial selection pressures and thus changes phenotypes (Olsen *et al.* 2004; Swain *et al.* 2007). Both neutral and adaptive diversity are part of a species' standing genetic variation, which is imperative to potentially adapt to an environment that changes at increasing rates (Pauls *et al.* 2013; Bernatchez 2016). Characterizing levels of gene flow and genetic differentiation is hence an important avenue of marine evolutionary biology and key for sustainable long-term management and conservation goals (Reiss *et al.* 2009; Funk *et al.* 2012).

While man-induced environmental changes are occurring worldwide, polar regions in particular are experiencing rapid warming (Vaughan *et al.* 2003; Parkinson 2004; Bromwich *et al.* 2013). In the Antarctic the link between climate and sea ice is complex, with regionally opposing effects. Despite an absence of warming in the first years of the twenty-first century (Turner *et al.* 2016) the Antarctic Peninsula region has experienced dramatic temperature increases with unprecedented break-up of ice shelves in the last century (Domack *et al.* 2005; Mulvaney *et al.* 2012; Hogg & Gudmundsson 2017). Changing marine ecosystems are the consequence (Gutt *et al.* 2014; Constable *et al.* 2014; Suprenand & Ainsworth 2017). On the other hand, parts of East Antarctica show increases in sea ice cover over the past decades, with effects on e.g. the benthic

food web (McMullin et al. 2017). Contrasting environmental changes around Antarctica may strongly affect future distributions of Antarctic marine animals as well as intraspecific diversity patterns of these taxa. It is important to establish a baseline understanding of diversity distributions to facilitate predictive modeling and future comparisons with the current state (Gutt *et al.* 2012).

The Antarctic fish fauna is dominated by teleosts, in particular members of the Notothenioidei (Eastman 1993). In a rare example of marine adaptive radiation, these fishes have diversified into more than 150 extant species (Eastman 1991; Near *et al.* 2012; Eschmeyer *et al.* 2018). Population genetic investigations of some notothenioids started several decades ago and are following technological developments in the field – from single marker to genome-wide approaches (Volckaert *et al.* 2012), although the latter are only currently being developed. Studies using microsatellite markers have revealed high gene flow in some Antarctic fish (*Gobionotothen gibberifrons* - Matschiner et al. 2009; various - Damerau et al. 2012; *Chionodraco rastrospinosus* - Papetti et al. 2012; *Notothenia rossii* - Young et al. 2015) and more pronounced genetic structure in others (*Trematomus* spp. - Van de Putte et al. 2012; *Champsocephalus gunnari* & *Chaenocepalus aceratus* - Damerau et al. 2014; Young et al. 2015, *Pleuragramma antarctica* - Agostini et al. 2015). Questions remain as to the relative contribution of life history traits and physical oceanographic forcing on spatial genetic structure (Damerau *et al.* 2014; Young *et al.* 2015; Moon *et al.* 2017). In addition, adaptive differentiation patterns and genotype-associations with environmental variables and ongoing climate change are largely unknown in Antarctic fishes.

Enormous advances in sequencing technologies have triggered an explosion of ecological, evolutionary, conservation, and fisheries studies applying a reduced representation sequencing technique such as RADseq, ddRAD, GBS, or similar, on non-model species (Baird *et al.* 2008; Peterson *et al.* 2012; Andrews *et al.* 2016; Crawford & Oleksiak 2016; Bernatchez *et al.* 2017). Initially, most of these projects relied on building a *de novo* catalog against which obtained sequences are mapped to identify variants. More recently, however, more and more (draft) reference genomes of species across the tree of life have become available and the rate of this increase in available genomic resources is likely to continue to rise (Ellegren 2014; Matz 2017; Lewin *et al.* 2018). While it is recommended to use reference mapping instead of *de novo* catalog building (Shafer *et al.* 2017), this may also lead to a loss of information due to ascertainment

bias. SNP chips are prone to ascertainment bias, but the same principle applies to reference-mapped reduced representation sequencing data, where rare or population-specific SNPs may be lost during the mapping. Especially when scanning for loci related to spatial selection, it is therefore worthwhile to explore whether exclusively mapping against a reference by default is the only way forward.

The bullhead notothen *Notothenia coriiceps* (Richardson, 1844) (Nototheniidae, Perciformes) is one of the most abundant shallow water notothenioids and occurs likely circum-Antarctic on the continental shelf as well as around sub-Antarctic islands, mostly in depths of less than 200 m (DeWitt *et al.* 1990). Many parts of the shallow high Antarctic shelves are rarely accessible and thus large knowledge gaps regarding occurrence persist (Duhamel *et al.* 2014). The species can reach 62 cm total length, though most commonly staying below 50 cm, and reaches maturity at approximately 20-40 cm total length (DeWitt *et al.* 1990; Sapota 1999; Calì *et al.* 2017). Adult fish are comparatively sedentary (North 1996). Females spawn pelagic eggs of ≥ 4 mm diameter, which develop for 100-200 days (depending on locality) until approx. 12 mm long larvae hatch in spring (Dewitt *et al.* 1990 and references therein; Sapota 1999). The pelagic larvae and young juveniles (metamorphosis occurs likely after approximately 6 months) stay in the water column for as long as one year (Sapota 1999) or even longer (Dewitt *et al.* 1990; Kellermann 1991). The behaviour of juveniles and the exact time of settling to a benthic life style seems variable (Kellermann 1991), but direct observations are lacking. Taken together, the characteristics of the bulhead notothen's life cycle suggest that connectivity between sub-populations may be high due to eggs, larvae, or juveniles covering relatively large distances during development. This may indicate low spatial genetic differentiation similar to its congener *Notothenia rossii* (Young *et al.* 2015). On the other hand, *N. coriiceps* remains much closer to shore than *N. rossii*, which could imply higher local retention of eggs and larvae. In contrast to *N. rossii*, *N. coriiceps* has been spared from commercial extinction and is no target species of past and contemporary Antarctic fisheries (Kock 1992; Kock & Jones 2005). However, it was found to be more susceptible to toxic effects of pollutants (Rodrigues *et al.* 2015). The species may also be more threatened by large scale temperature rises, as its distribution range covers large parts of the high Antarctic shelves, where a distribution shift in response to warming is not feasible.

*Notothenia coriiceps* is furthermore the first Antarctic vertebrate where a draft genome was sequenced (Shin *et al.* 2014), which should be helpful to call genotypes with RADseq. Intraspecific genomic variability of *N. coriiceps* has not been investigated on either a temporal or a spatial scale except for one preliminary single marker result (Jones *et al.* 2008). Accounting for natural diversity, however, is necessary for subsequent comparative genomic studies as well as spatial conservation planning. We here apply a reduced representation sequencing approach, more specifically modified double-digest restriction-site associated DNA sequencing (ddRAD, Peterson et al. 2012), to assess the genomic diversity of an emerging model species for cold adaptation, evolution, and development (Hsiao *et al.* 1990; Braasch *et al.* 2014; Postlethwait *et al.* 2016; Amores *et al.* 2017). Genotype-environment-associations and possible signatures of recent selection are investigated and interpreted with reference to changing environmental conditions. Genotypes are called using a *de novo* catalog assembly and mapping against the available reference genome. We expect the latter to perform better and hypothesize that populations between West and East Antarctica are differentiated, but populations on a local scale in the western Antarctic Peninsula region are genetically homogeneous.

## 2. Material and methods

### 2.1 Sampling and identification

In total 132 specimens of *Notothenia coriiceps* from six sampling localities were included (Fig. 4.1, Table 4.1). Sampling sites are distributed along the western Antarctic Peninsula and, more remotely, off Adélie Land in the Eastern Antarctic (Koubbi 2003). The wide and clustered geographic distribution of these samples provides the opportunity to compare large- and small-scale differentiation patterns. The East Antarctic samples include fish caught at several sites in the vicinity of the Dumont d'Urville Station and these were split into a "North" and a "South" sample group. Fish were caught by trammel nets, otter trawls, or hand line, identified on site, and fin or muscle biopsies were stored in ethanol until further processing. All included specimens were adults of > 25 cm total length. While no maturity data was available, based on the fish size we assume that these individuals were likely sexually mature and surely well beyond the pelagic phase, thus representing local residents. Maps of the sampling sites were created using marmap v0.9.6 (Pante & Simon-Bouhet 2013) in R v3.4.4 (R Core Team 2019).

**Table 4.1.** Sampling details of *Notothenia coriiceps* specimens used for population genomic analyses. Geographical coordinates are given in decimal degrees. At Deception Island and Adélie Land, South more than one locality in close proximity were sampled.

| Location | Location code | Latitude | Longitude | Sample size | Sampling time |
|---|---|---|---|---|---|
| South Orkney Islands | SO | -60.70 | -45.57 | 20 | March 2006 |
| King George Island, Potter Cove | PKGI | -62.23 | -58.68 | 26 | March 2006 |
| King George Island, North | NKGI | -61.97 | -57.24 | 6 | March 2006 |
| Elephant Island | EI | -61.24 | -55.62 | 20 | March 2006 |
| Deception Island | DI | -62.95; -62.97 | -60.65; -60.61 | 20 | March 2006 |
| Adélie Land, North | NTA | -66.66 | 140.02 | 20 | 2004-2006 |
| Adélie Land, South | STA | -66.67; -66.66 | 140.00; 139.99 | 20 | 2004-2006 |

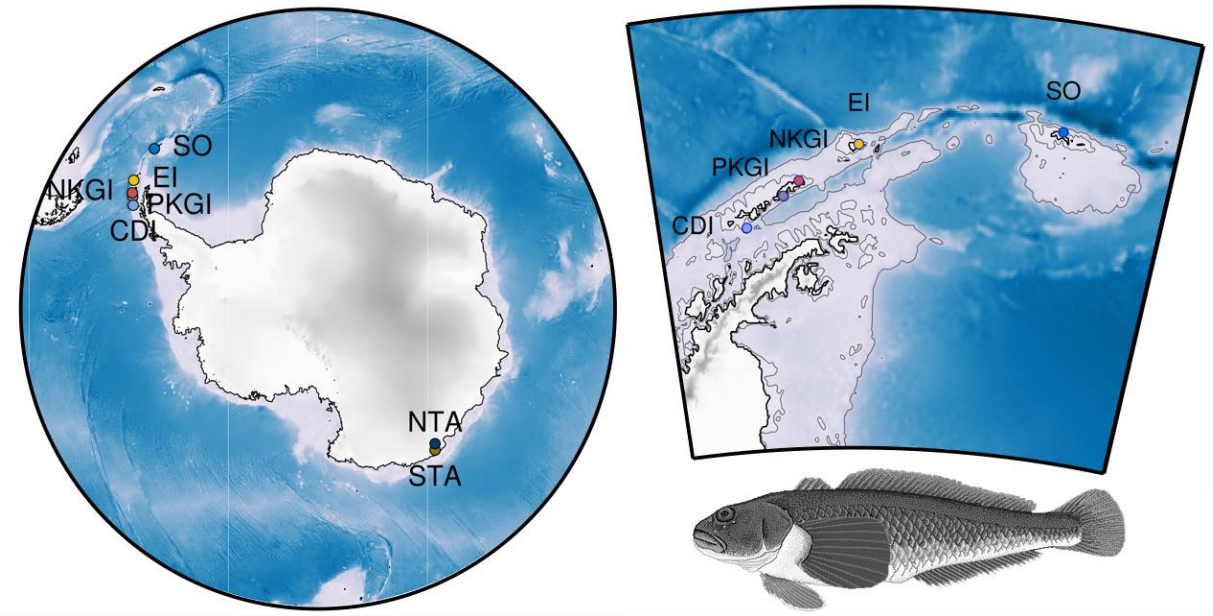



**Fig. 4.1.** Sites where *Notothenia coriiceps* (artwork by R. C. Cada for FishBase) was sampled in the Southern Ocean between 2004 and 2006. Left: all sampling localities. Right: detailed map of the tip of the Antarctic Peninsula with 200 and 1000 m depth contours. Depth indicated by color gradient from grey to dark blue (color figure available online). Sampling sites abbreviated as in Table 4.1: NTA: Adélie Land, North; STA: Adélie Land, South; CDI: Deception Island; PKGI: King George Island, Potter Cove; NKGI: King George Island; North; EI: Elephant Island; SO: South Orkney Islands.

### 2.2 Molecular analyses

Genomic DNA was extracted using a salting out protocol without vortexing to avoid shearing of the DNA, largely as described in Cruz et al. (2016). DNA quantity was assessed fluorometrically using the Quant-iT PicoGreen dsDNA kit (Thermo Fisher Scientific Inc.) and an Infinite M200 microplate reader (Tecan Group Ltd.) according to the manufacturer's instructions. DNA was then standardized to an estimated concentration of 7 ng $\mu l^{-1}$ and quality was checked via 1% agarose gel electrophoresis. Only samples of high quality were included (and listed in Table 4.1) and quantified and standardized a second time. All DNA extractions were used for DNA barcoding by amplifying an approximately 655 bp fragment of the mitochondrial DNA gene cytochrome oxidase subunit I (COI) using primers COI-fish-F1 and -R1 (Ward *et al.* 2005). Forward and reverse COI fragments were sequenced on an ABI 3130 capillary DNA sequencer (Applied Biosystems Inc.) and analyzed using GeneMapper v4.0 (Applied Biosystems Inc.) and Geneious v9.0 (Biomatters Ltd.). The Barcode of Life Data System (Ratnasingham & Hebert 2007) was used to verify morphological specimen identification with DNA barcoding data. One specimen was found to be *Notothenia rossii* and excluded from downstream bioinformatics and statistical analysis. DNA extracts were then used to construct a double-digest restriction-site associated DNA sequencing (ddRAD) library, based on the method described by Peterson et al. (2012) but with modifications. The library preparation protocol essentially follows the description in Palaiokostas et al. (2015). In brief, each sample was digested at 37° C for 45 min with restriction enzymes *SbfI* and *SphI*. Adapter combinations with unique 5 or 7 bp barcodes were then ligated to the digested DNA at 22° C for 150 min. After heat inactivation for 20 min at 65° C all samples were pooled and purified using MinElute PCR Purification Kits (Qiagen). The pooled library was size selected on agarose gel for a target range of 320 - 590 bp and subsequently purified (MinElute Gel Extraction Kit, Qiagen) and PCR amplified. Finally, after purification using MinElute PCR Purification Kits (Qiagen) the library was checked once more for quality and quantity and then frozen until sequencing.

### 2.3 Sequencing and bioinformatics

Sequencing took place at the Center for Human Genetics of the KU Leuven (GenomicsCore) using one paired-end run on one lane of a HiSeq 2500 platform with v3 chemistry (Illumina Inc.) creating some 342 million 100 bp reads. Raw sequencing data was checked for coarse sequencing error using FastQC (Andrews 2010), which revealed generally good quality (overall mean Phred score of 35.4) with lower quality in the ten last bp of the reads, which was expected

and trimmed in the process of demultiplexing. Bioinformatics were run locally, first using the module process_radtags from Stacks v1.37 (Catchen *et al.* 2011, 2013) for demultiplexing and trimming reads with low Phred scores (<20), removing reads with an uncalled base, and rescuing barcodes and RAD-tags (options: -t 90, -c, -q, -r, -s 20). From a total of approximately 342 million (M) reads, 36M reads contained ambiguous barcodes, 39M were discarded due to low quality, and 7M due to ambiguous RAD-tags, leaving 260M retained reads. Further steps followed dDocent v2.12 (Puritz *et al.* 2014a), a pipeline designed to work with reduced representation sequencing data of marine non-model species. Being essentially a series of wrapper shell scripts, dDocent relies on other free software with, in v2.12, CD-Hit (Li & Godzik 2006; Fu *et al.* 2012) at the core of the *de novo* assembly and the Burrows-Wheeler Alignment tool BWA (Li & Durbin 2009) as mapper, and FreeBayes (Garrison & Marth 2012) as variant caller. The pipeline was employed twice using largely default settings, with and without use of the *Notothenia coriiceps* reference genome (Shin *et al.* 2014). These two data sets are hereafter referred to as *de novo* and reference, and yielded initial sets of respectively 16,787 (*de novo*) and 40,343 loci (reference). Raw vcf file output of dDocent was subsequently filtered using vcftools (Danecek *et al.* 2011) requiring loci to be genotyped in 80 % of all individuals, having a minimum minor allele count of 3, and a minimum Phred score of 30. Minimum depth for a genotype call was required to be 3 and minimum mean depth had to be 3, too. Individuals with less than 80 % genotyped loci were excluded, leaving 114 individuals. A minimum minor allele frequency of 0.005 was set and loci had to be genotyped in 80 % of the individuals and at a minimum mean depth of 20. Finally, sites with allele balance not between 0.25 – 0.75 and sites with abnormally high depth (see e.g. Li 2014) were filtered and indels excluded.

### 2.4 Population genomics

Statistical analyses were conducted largely using R v3.6.0 (R Core Team 2019), which now provides a large set of packages to conduct population genetic and genomic analyses (Paradis *et al.* 2016). The R code used for our analyses is freely available under http://doi.org/10.5281/zenodo.2650260, which should make all statistical analyses and results reproducible. Output files from vcftools were converted to genind objects in R using 'vcfR' v1.8.0 (Knaus & Grünwald 2017) and 'adegenet' v2.1.1 (Jombart 2008; Jombart & Ahmed 2011) and some last filtering steps were conducted using 'adegenet' and 'poppr' v2.8.2 (Kamvar *et al.* 2014). Only bi-allelic SNPs with expected heterozygosity ($H_{exp}$) >0.5 and observed heterozygosity ($H_{obs}$) >0.5 were kept in order to remove loci that are likely results of genotyping error or paralogy. In

addition, individuals with <80% genotypes and loci with <80% call rate over individuals were filtered again to reduce the amount of missing data. After these steps, the data sets were filtered again for minor allele frequency, this time setting the minimum to 0.02. Each locus was tested in each population for agreement with Hardy-Weinberg proportions (HWP) using 'pegas' v0.11 (Paradis 2010) and loci that deviated from HWP after correcting for multiple testing using the 'qvalue' package v2.15.0 (Storey *et al.* 2019) were removed.

Observed and expected heterozygosity per sampling locality were calculated using adegenet. Overall levels of genetic differentiation as well as pairwise differentiation was assessed using R package 'hierfstat' v0.04-30 (Goudet & Jombart 2015). In addition, alternative measures of differentiation, which may be more appropriate when diversity is high (Jost 2008), were calculated using package 'mmod' v1.3.3 (Winter 2012). Pairwise differentiation was visualized with non-metric multidimensional scaling plots produced using 'MASS' v7.3-51.1 (Venables & Ripley 2002). A hierarchical analysis of molecular variance (AMOVA) following Excoffier *et al.* (1992) was performed using 'pegas'. In order to provide a genuine picture of individual based genetic differentiation, a principal component analysis (PCA) was conducted with 'adegenet', and packages 'factoextra' v1.0.5 (Kassambara & Mundt 2017) and 'ggsci' v2.9 (Xiao 2017) aiding plotting of the results. In addition, the Bayesian clustering software Structure v2.3.4 (Pritchard *et al.* 2000) was called from within R using 'ParallelStructure' v1.0 (Besnier & Glover 2013) and a function from Clark (2017) to prepare input files. Structure was run on both data sets without location priors, using the admixture model and correlated allele frequencies (Falush *et al.* 2003; Hubisz *et al.* 2009) with K ranging from two to eight, with five replicates of each run and always using 10,000 repetitions as burn-in and 100,000 subsequent iterations. Results from the Structure analysis were analysed as suggested by Pritchard & Wen (2004) and Evanno *et al.* (2005) and diagnosis plots were created using Structure Harvester (Earl & vonHoldt 2012).

#### 2.5 Genotype-environment association

Redundancy Analysis (RDA) is a powerful tool widely employed in community ecology. Here, we applied this multivariate linear regression method to detect multilocus associations between genetic and environmental data (Forester *et al.* 2018). Recent landscape or seascape genomics analyses have successfully used this technique in a variety of systems (Vandamme *et al.* 2014; Benestan *et al.* 2016; Laporte *et al.* 2016; Raeymaekers *et al.* 2017; Xuereb *et al.* 2018); and it is arguably more accurate than, for example, Mantel tests (Meirmans 2015). All RDA analyses were

conducted in R using 'vegan' v2.5-4 (Oksanen *et al.* 2018). Allele frequencies were calculated from the genotypes and Hellinger transformed (Legendre & Gallagher 2001) and subsequently summarized in 64 (*de novo* data set) or 63 (reference data set) principal components (PCs), representing >75% of the cumulative variance. Sampling coordinates were transformed to Euclidean distances and then decomposed into principal coordinates of neighborhood matrices (PCNMs, also referred to as Moran's Eigenvector Maps). A set of 58 environmental descriptors targeted towards the analysis of benthic Antarctic animals was used to describe the habitat of the sampled specimens (Guillaumot *et al.* 2018 and references therein). Most environmental variables were subsequently excluded, based on the presence of missing data, non-continuous variables, and autocorrelated variables. The final RDAs were constructed with four PCNMs and twenty key environmental descriptors, i.e., depth, mixed layer depth (MLD), sea surface current speed (SSCS), sea floor current speed (SFCS), distance to continental Antarctica (dist), ice cover range (ICR), mean ice thickness (ITM), ice thickness range (ITR), mean annual chlorophyll *a* concentration (chl_a_mean), amplitude of annual chlorophyll *a* concentration (chl_a_ampli), mean sea floor salinity (SFS), amplitude of mean sea floor salinity (SFS_ampli), mean sea floor temperature (SFT), and several variables relating to extreme events (EE) and chlorophyll *a* concentration and temperature (EE_max_chl_a, EE_max_chl_a_ampli, EE_min_chl_a, EE_min_chl_a_ampli, EE_max_temp, EE_min_temp). The ice descriptors cover the time period 1957-2017, while the other retained temporally variable descriptors (all chlorophyll, temperature, salinity and extreme event descriptors) are from 2005-2012. Significance of the results was tested with ANOVA-like permutation tests. After analyzing RDAs with all variables, a stepwise (forward) model selection was used to identify variables that explained the genotype variance best.

### 2.6 Genome scans

Finally, both *de novo* and reference SNP sets were subjected to genome scans for selection, in order to identify loci potentially relevant for local adaptation processes ("outlier loci"). Two complementary methods were used, the R package 'pcadapt' v4.1.0 (Luu *et al.* 2016) and BayeScan v2.1 (Foll & Gaggiotti 2008). The former detects outliers based on principal component analysis using a Bayesian factor model, while the latter is based on $F_{ST}$ values and employs a Bayesian framework with Monte Carlo Markov Chains to explore possible selection signatures at loci that depart from neutrality (Beaumont & Balding 2004). BayeScan was run on all putative populations (as in sampling localities, Table 3.1) with 10 prior odds, 50 pilot runs of

5,000 iterations, a burn-in of 50,000, and finally 100,000 chain iterations, and otherwise default settings. Convergence of the Markov chains was checked using Geweke's (1992) and Heidelberger and Welch's (1981) diagnostics with R package 'coda' v0.19-2 (Plummer *et al.* 2006). For pcadapt two principal components reflecting large scale population structure were retained and Mahalanobis distance of SNPs was calculated. The false discovery rate was controlled and significant loci at a q-value threshold of 0.1 were retained as candidate loci.

All candidate loci from both *de novo* and reference SNP sets and both genome detection methods were scanned for proximity to known genes using the Basic Local Alignment Search Tool (BLAST) through the NCBI website (Johnson *et al.* 2008). Where possible, the most likely gene function was assessed by identifying the function of the same gene in other species (mostly *Homo sapiens* and *Rattus norvegicus*) through UniProt (Apweiler *et al.* 2004). The RAD fragments that contained outlier loci of the *de novo* SNP set were aligned in Geneious v11.0.2 (Biomatters Ltd) to the reference genome (Shin *et al.* 2014) to test for overlap of the outliers detected between *de novo* and reference data sets. Outlier loci detected with both genome scan methods were tested for correlation to selected environmental variables using generalized linear models (GLMs) with a binomial error.

## 3. Results

In order to elucidate genomic variability within the Antarctic bullhead notothen we sequenced one ddRAD library of 132 specimens from West and East Antarctic sites. More than 300 million paired end reads of high quality (average phred score 36.6) were obtained and used for extensive bioinformatics in order to make confident genotype calls. These reads were, however, spread unevenly among the individuals with some individuals having a share of only a few thousand reads and others up to seven million reads (total average 950,413 +- 1,087,995). All individuals were kept nevertheless for bioinformatics analyses, which were generally designed to if in doubt rather keep data and instead postpone strict filtering processes for after SNP calling (see below).

### 3.1 Variant detection

The dDocent pipeline was run twice, assembling loci *de novo* and mapping against the available reference genome. These analyses yielded 16,787 and 40,343 loci, respectively, which was

reduced to 3,422 (*de novo*) and 2,783 (reference) SNPs after extensive filtering using vcftools. During the filtering with vcftools, 18 individuals were excluded, as these had <80% of loci genotyped. Further filtering steps were conducted in R, excluding loci that were likely to be genotyping error or paralogs (939 and 726 loci removed, respectively). The amount of missing data was reduced by then again removing 12 (*de novo*) and 11 (reference) individuals with > 20% missing genotypes. The individuals that were excluded (both with vcftools and R) largely corresponded to individuals that had the lowest numbers of reads in the first place. Filtering on minor allele frequency of 0.02 removed 511 (*de novo*) and 467 (reference) further loci. Twenty-four and 18 loci, respectively, were filtered due to population-specific departures from HWP. After these final filtering steps the data sets comprised genotypes from 102 individuals at 1,948 loci with 4.01% missing values (*de novo*) and 103 individuals at 1,572 loci with 3.96% missing values (reference). These are the final data sets that were used for all following analyses.

### 3.2 (Spatial) genetic variation

Overall, the data sets showed low genetic differentiation with $F_{ST}$ (Nei 1973) of 0.0097 and 0.0099, $F_{IT}$ of 0.16241 and 0.1502, and $F_{IS}$ of 0.1542 and 0.1417 (always *de novo* first and reference data set second). Mean observed heterozygosity per sampling locality ranged from 0.1534 to 0.1719 (*de novo*) and 0.1545 to 0.1767 (reference; Table 4.2). Mean expected heterozygosity per sampling locality was consistently higher than observed heterozygosity with values up to 0.1970 (*de novo*) and 0.2025 (reference) (Table 4.2; Supplementary Fig. S4.1 & S4.2).

**Table 4.2.** Mean observed ($H_O$) and expected heterozygosity ($H_E$) per sampling locality (see Table 4.1 for codes) of *Notothenia coriiceps* based on 1972 SNP genotypes derived from mapping against a *de novo* assembly and based on 1590 SNP genotypes derived from mapping against the reference genome of Shin et al. (2014). The respective standard error is shown after the value.

| | $H_O$ *de novo* | $H_E$ *de novo* | $H_O$ reference | $H_E$ reference |
|---|---|---|---|---|
| **SO** | 0.1592 ± 0.0090 | 0.1915 ± 0.0105 | 0.1576 ± 0.0087 | 0.1847 ± 0.0102 |
| **EI** | 0.1534 ± 0.0061 | 0.1836 ± 0.0070 | 0.1710 ± 0.0074 | 0.2018 ± 0.0083 |
| **PKGI** | 0.1574 ± 0.0058 | 0.1884 ± 0.0067 | 0.1702 ± 0.0069 | 0.1945 ± 0.0077 |
| **NKGI** | 0.1719 ± 0.0130 | 0.1931 ± 0.0137 | 0.1767 ± 0.0149 | 0.2025 ± 0.0162 |
| **CDI** | 0.1587 ± 0.0084 | 0.1905 ± 0.0095 | 0.1683 ± 0.0100 | 0.2022 ± 0.0116 |
| **NTA** | 0.1659 ± 0.0096 | 0.1889 ± 0.0102 | 0.1545 ± 0.0100 | 0.1862 ± 0.0118 |
| **STA** | 0.1665 ± 0.0065 | 0.1970 ± 0.0078 | 0.1620 ± 0.0076 | 0.1867 ± 0.0084 |

Pairwise genetic differentiation estimated as $F_{ST}$ (or $G_{ST}$) following Weir and Cockerham (1984) was low but significant (i.e. with 0.025 and 0.975 confidence interval quantiles > 0) in all pairwise comparisons between East Antarctic and West Antarctic samples (Table 4.3 & 4.4). In total, values ranged from -0.0021 to 0.0225 and from -0.0044 to 0.0239 in *de novo* and reference data sets, respectively. Highest values in both cases were observed between Adélie Land, South and the South Orkney Islands. In addition to the east-west differentiation two pairwise comparisons were significant as defined above: South Orkney Islands vs. Elephant Island and King George Island, North in the *de novo* set and South Orkney Islands vs. Elephant Island and Deception Island in the reference data set. These patterns were visualized using non-metric multidimensional scaling plots, which highlight that not only Adélie Land fish, but also the South Orkney Island samples diverge from the remaining localities (Supplementary Fig. S4.3 & S4.4). Alternative differentiation metrics provide similar results (Supplementary Tables S4.6 & S4.7). An AMOVA indicated that in both cases there was between group differentiation (East vs. West Antarctica, p = 0.0460, *de novo* data set, and 0.0450, reference data set), but within-group differentiation was present only in the reference data set (p = 0.0929, *de novo* and 0.0350, reference).

**Table 4.3.** Pairwise genetic differentiation of *Notothenia coriiceps* per sampling locality (see Table 4.1 for codes) based on 1948 SNP genotypes derived from mapping against a *de novo* assembly. $F_{ST}$ following Weir and Cockerham (1984) (alternatively referred to as $G_{ST}$) below the diagonal (negative values set to zero) and confidence intervals after 1000 bootstraps above the diagonal. $F_{ST}$ values where confidence intervals do not span zero are marked in bold.

| | **SO** | **EI** | **PKGI** | **NKGI** | **CDI** | **NTA** | **STA** |
|---|---|---|---|---|---|---|---|
| **SO** | | 0.0018 – 0.0093 | -0.0021 – 0.0065 | 0.0002 – 0.0159 | -0.0023 – 0.0083 | 0.0130 – 0.0268 | 0.0173 – 0.0282 |
| **EI** | **0.0053** | | -0.0009 – 0.0040 | -0.0035 – 0.0076 | -0.0041 – 0.0026 | 0.0137 – 0.0236 | 0.0127 – 0.0206 |
| **PKGI** | 0.0021 | 0.0017 | | -0.0069 – 0.0031 | -0.0039 – 0.0023 | 0.0103 – 0.0185 | 0.0134 – 0.0208 |
| **NKGI** | **0.0079** | 0.0015 | -0.0021 | | -0.0065 – 0.0060 | 0.0140 – 0.0315 | 0.0146 – 0.0278 |
| **CDI** | 0.0029 | -0.0007 | -0.0010 | -0.0002 | | 0.0134 – 0.0243 | 0.0132 – 0.0221 |
| **NTA** | **0.0199** | **0.0186** | **0.0144** | **0.0223** | **0.0191** | | -0.0047 – 0.0020 |
| **STA** | **0.0225** | **0.0165** | **0.0170** | **0.0211** | **0.0175** | -0.0013 | |

**Table 4.4.** Pairwise genetic differentiation of *Notothenia coriiceps* per sampling locality (see Table 4.1 for codes) based on 1572 SNP genotypes derived from mapping against the reference genome of Shin et al. (2014). $F_{ST}$ following Weir and Cockerham (1984) below the diagonal (negative values set to zero) and confidence intervals after 1000 bootstraps above the diagonal. $F_{ST}$ values where confidence intervals do not span zero are marked in bold.

| | **SO** | **EI** | **PKGI** | **NKGI** | **CDI** | **NTA** | **STA** |
|---|---|---|---|---|---|---|---|
| | | 0.0018 – | -0.0033 – | -0.0053 – | 0.0022 – | 0.0101 – | 0.0180 – |
| **SO** | | 0.0101 | 0.0048 | 0.0096 | 0.0146 | 0.0238 | 0.0298 |
| | | | -0.0010 – | -0.0061 – | -0.0027 – | 0.0104 – | 0.0148 – |
| **EI** | **0.0058** | | 0.0041 | 0.0057 | 0.0052 | 0.0198 | 0.0240 |
| | | | | -0.0092 – | -0.0024 – | 0.0072 – | 0.0137 – |
| **PKGI** | 0.0006 | 0.0015 | | 0.0004 | 0.0047 | 0.0169 | 0.0214 |
| | | | | | -0.0054 – | 0.0074 – | 0.0142 – |
| **NKGI** | 0.0020 | -0.0001 | -0.0044 | | 0.0083 | 0.0259 | 0.0290 |
| | | | | | | 0.0126 – | 0.0152 – |
| **CDI** | **0.0080** | 0.0013 | 0.0011 | 0.0014 | | 0.0246 | 0.0250 |
| | | | | | | | -0.0048 – |
| **NTA** | **0.0167** | **0.0150** | **0.0119** | **0.0164** | **0.0185** | | 0.0028 |
| **STA** | **0.0239** | **0.0191** | **0.0174** | **0.0213** | **0.0201** | -0.0012 | |

In addition to F statistics and related measures, individual-based analyses were applied to further elucidate genomic differentiation patterns. Principal component analysis revealed clear differentiation between East Antarctic (NTA and STA) and the Antarctic Peninsula region samples along principal component 1 (Fig. 4.2). Structure analyses indicated that K = 2 is the most likely number of genetic clusters detectable in the data (*de novo* data set: mean $L(K)$ = - 121849 and $\Delta K$ = 154; Evanno *et al.* 2005), also supported by the fact that, when plotting ancestry barplots with K > 3, no additional patterns arise. Individuals of both inferred clusters show little mixed ancestry signals (Fig. 4.2). These patterns were consistent among *de novo* and reference (reference data set: mean $L(K)$ = -100980 and $\Delta K$ = 59; Fig. S4.5) data sets. Taken together these results indicate that the East-West differentiation is the dominant pattern present in the global data set, while small-scale differentiation with respect to the South Orkney Islands is comparatively minor.

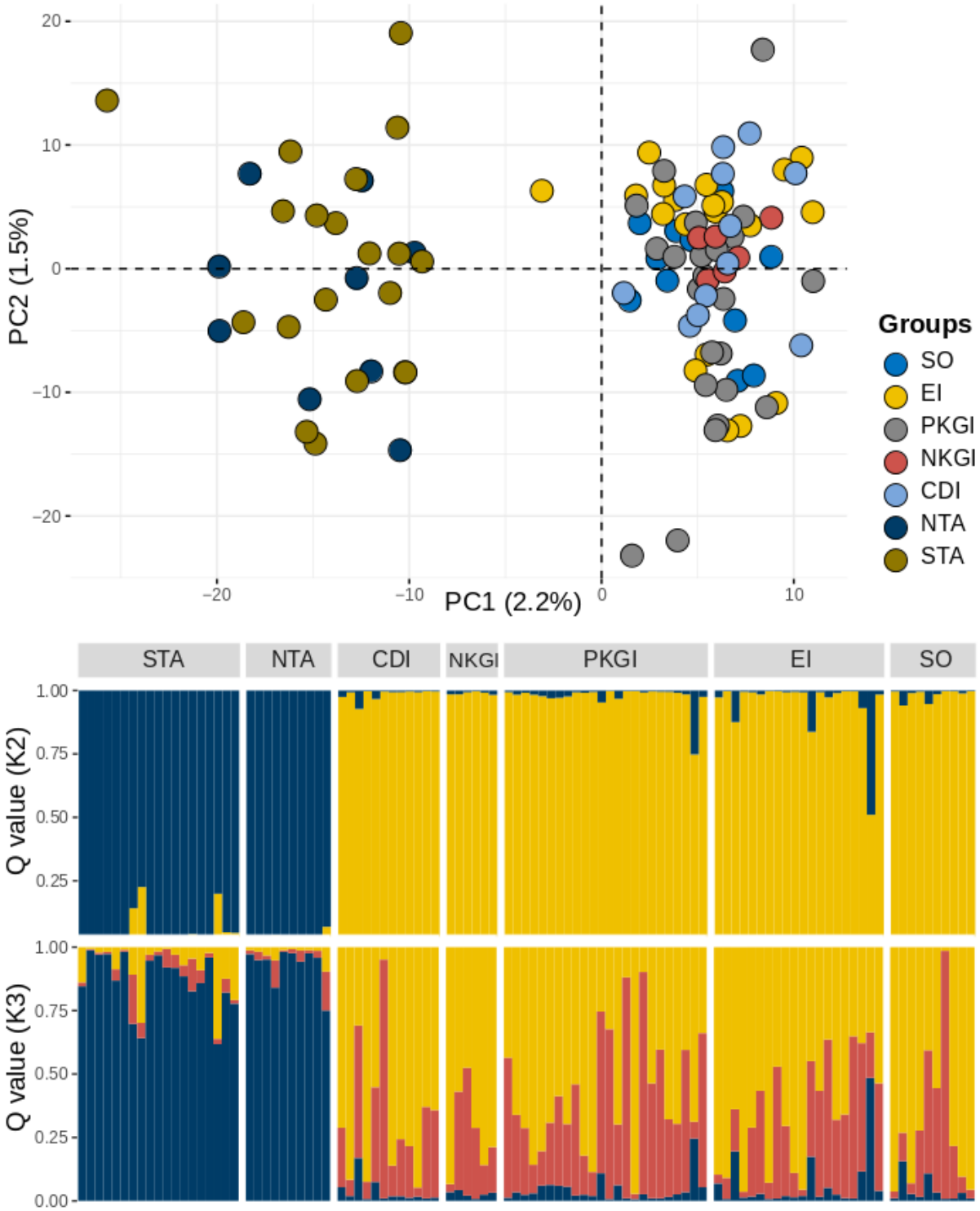



**Fig. 4.2.** Genetic differentiation of *Notothenia coriiceps* in the Southern Ocean based on 1948 SNP genotypes derived from mapping against a *de novo* assembly. East Antarctic samples (NTA, STA) are differentiated from West Antarctic samples after principal component (PC; above) and Structure analysis (below), with both two and three modeled genetic clusters (K). Sample codes as in Fig. 4.1 and Table 4.1.

### 3.3 Genotype-environment associations

The two data sets were used to run redundancy analysis (RDA) with a core set of twenty environmental and four spatial (PCNM) descriptors (Fig. 4.3). The RDA run with the *de novo* data set explained only little of the variation ($r^2$ = 0.0418; adjusted $r^2$ = 0.0125). Using the reference data set provided very similar results ($r^2$ = 0.0314; adjusted $r^2$ = 0.0120). Both RDAs were highly significant after 1000 permutations ($p \leq 0.001$; F = 1.2614 and 1.2234, respectively, always *de novo* first and reference data set second), but only the first axis was significant ($p \leq 0.001$; F = 2.8441 and 2.9812, respectively). In both data sets samples from East Antarctica (Adélie Land) were correlated with higher current speed (SSCS, SFCS) and mixed layer depth (MLD), while samples from the western Antarctic Peninsula were correlated with higher ice cover variability (ICR) and larger distance to Antarctica (dist). Spatial variability was represented by PCNM2 and PCNM3, which correlated differently to genotype PCs from the *de novo* and reference data though (Fig. 4.3a, c). After model selection only three (*de novo*) or two (reference) descriptor variables were retained (Fig. 4.3b, d). These RDAs showed similar levels of variation ($r^2$ = 0.0505; adjusted $r^2$ = 0.0214; and $r^2$ = 0.0411; adjusted $r^2$ = 0.0219) and were still significant after 1000 permutations ($p \leq 0.001$; F = 1.7360 and 2.1426, respectively), with again only the first axis being significant ($p \leq 0.001$; F = 2.8262 and 2.9733, respectively). The model selection retained PNCM1, which correlated well with the observed genetic structure in both data sets. For the *de novo* data also mean ice thickness (ITM) and maximum chlorophyll *a* concentration during extreme events (EE_max_chl_a) were retained. ITM variation was represented along the same axis as PCNM1, while the chlorophyll descriptor varied between West Antarctic samples, with fish from King George Island and Elephant Island correlated to higher values. Using the reference data only sea floor current speed (SFCS) was retained and perfectly correlated with space (PCNM1).

### 3.4 Evidence of local adaptation

*De novo* and reference data sets were subjected to two methods of genome scans for outlier detection. BayeScan run on seven putative populations identified 6 outlier loci at a q-value threshold of 0.05 and 10 loci at a threshold of 0.1 for the *de novo* set and 3 ($q < 0.05$) or 4 ($q < 0.1$) loci for the reference set. Given the generally rather low levels of $F_{ST}$ in marine fishes and low numbers of loci identified by BayeScan, we chose to retain candidate loci at the 10% false positives threshold for downstream investigations. We emphasize that these are mere candidates for further study and independent verification at this stage. Recently admixed

individuals may further complicate outlier loci detection based on *a priori* defined populations (Luu *et al.* 2016). Therefore, the individual-based pcadapt was used in addition to BayeScan. This selection scan using principal components identified 23 and 17 candidate loci at q <0.1 (12 and 8 at q <0.05) for the *de novo* and reference set, respectively. Three loci in the *de novo* set and one locus in the reference data set were identified by both BayeScan and pcadapt (Fig. 4.4).

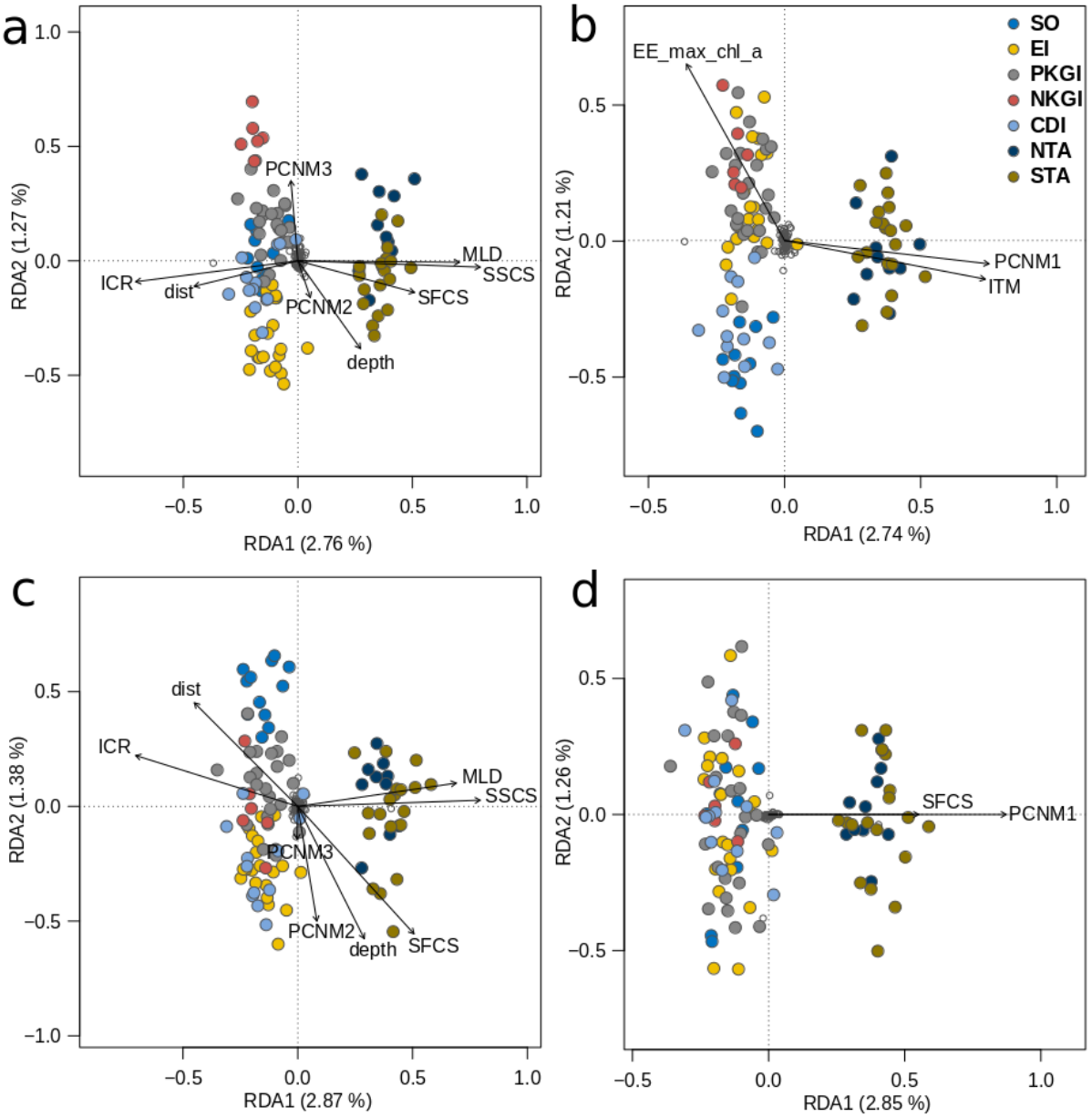



**Fig. 4.3.** Genotype-environment association of *Notothenia coriiceps* in the Southern Ocean as visualized from redundancy analysis (RDA) based on 1948 SNP genotypes derived from a de novo assembly (a, b) or from mapping against the reference genome of Shin et al. (2014) (c, d). Results from RDA with 20 environmental variables (see text for abbreviations) and four principal coordinate neighboring matrices (PCNM) representing space are shown (a, c), as well as results after forward model selection (b, d). Sample codes as in Fig. 4.1 and Table 4.1 (color figure available online).

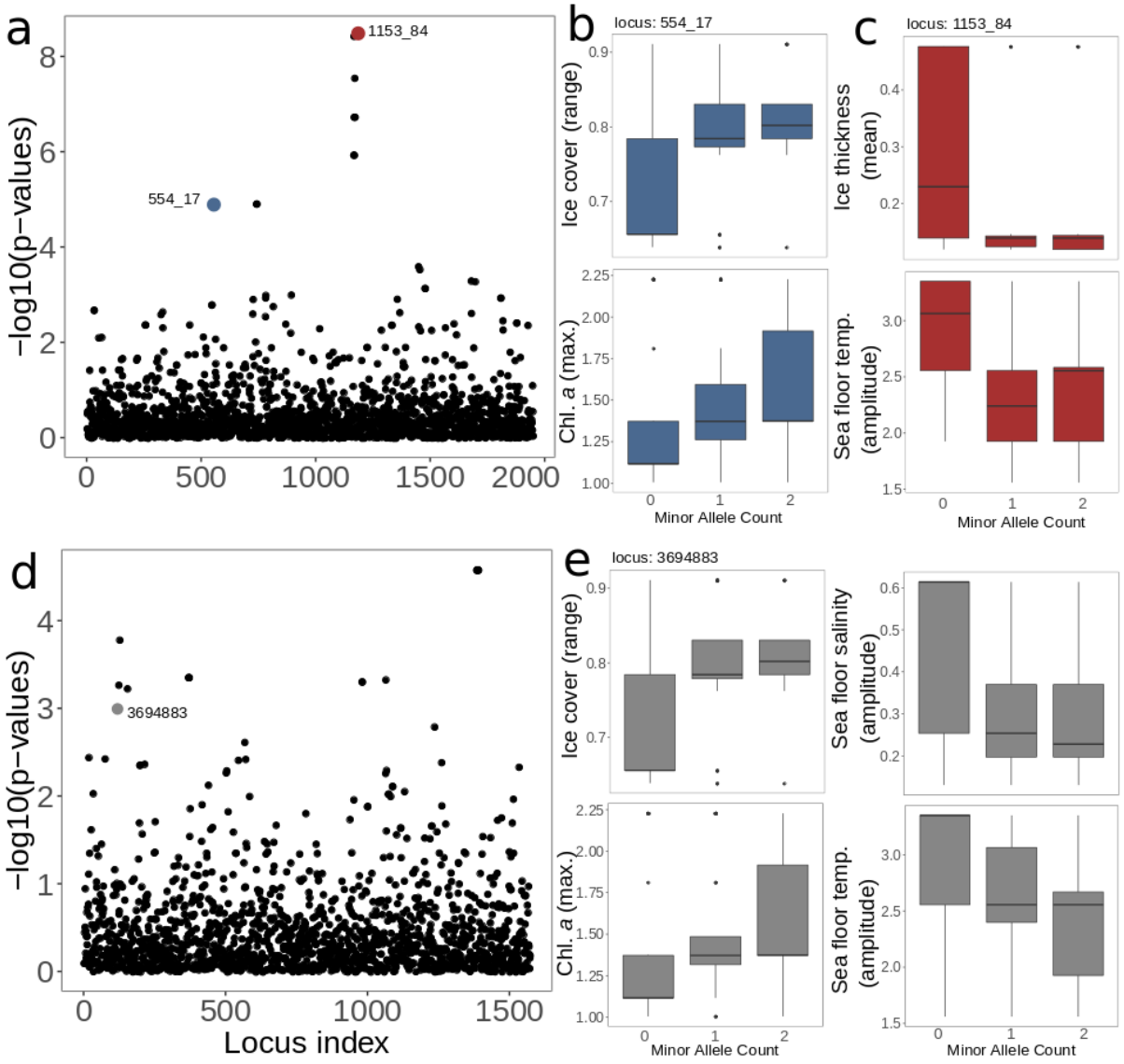



**Fig. 4.4.** Putative signatures of selection in *Notothenia coriiceps* in the Southern Ocean based on 1948 SNP genotypes derived from a de novo assembly (a, b, c) or from mapping against the reference genome of Shin et al. (2014) (d, e). Manhattan plots (a, d) show negative log10 of p-values based on pcadapt analyses to detect outlier loci. Three SNP loci and their correlation to selected environmental variables are highlighted: a SNP on contig 554 at position 17 (a, b), a SNP on contig 1153 at position 84 (a, c) and a SNP at position gi_734200844_ref_NW_011369947_1_3694883 of the *N. coriiceps* reference genome.

In the *de novo* data set 13 of 23 candidate loci detected by 'pcadapt' resulted in hits when using BLAST (Table 4.5). Five potentially linked genes were identified, two of these had two SNPs in their proximity and one gene was within 192 bp of seven candidate SNPs. Two of the latter loci were also identified by BayeScan. These seven loci matched to a methionine synthase reductase-like mRNA gene of *N. coriiceps*. Nine SNPs identified by pcadapt in the *de novo* data set could not be matched to any gene. BayeScan identified three further genes near candidate loci. Five

SNPs identified by BayeScan in this data set could not be matched to any gene. The identified genes had various functions including immune system and gene regulation and transcription processes. Regarding the reference data set, nine candidate loci detected by 'pcadapt' and four loci detected by BayeScan all showed hits when searching around the position of the SNPs for up to 1190 bp (Table 4.5). The loci detected by 'pcadapt' could be potentially linked to five genes, one of which was the same as a gene identified in the *de novo* data set, a serine/threonine-protein phosphatase subunit. Two SNPs were in the proximity of an anti-freeze glycoprotein (AFGP) encoding gene identified in *Dissostichus mawsoni* (Antarctic toothfish). One of these SNPs was identified by BayeScan as well. Three other SNPs identified by BayeScan were matched to genes not identified in other candidate loci.

In total, most BLAST matches were predicted *Notothenia coriiceps* nucleotide sequences, based on similarity to genes or proteins to closely related isolates and/or species. In two cases BLAST hits indicated matches to *D. mawsoni*, in one case each to *Larimichthys crocea* (large yellow croaker), *Scophthalmus maximus* (turbot), *Anabas testudineus* (climbing perch), *Chionodraco hamatus* (a crocodile icefish), and to *Boreogadus saida* (polar cod). All in all, 9 of 22 unique outliers detected by either pcadapt or BayeScan in the *de novo* data set could be matched to the *N. coriiceps* reference genome. Only one of these loci corresponded to the outliers detected in the reference data set. This outlier was identified as *Scophthalmus maximus* chromosome 12; Sequence ID: CP026254.1 with BLAST and no specific gene function could be identified. Furthermore, there was little overlap between the outliers detected by the two different methods (pcadapt, BayeScan). The SNP on contig 554 at position 17 from the *de novo* assembly was identified by both methods at $q < 0.05$, but no gene in the vicinity could be identified. However, this locus was significantly (GLM, $p<0.001$) correlated with ice cover range and maximum annual chlorophyll *a* concentration (Fig. 3.4b). The minor allele of two linked SNPs (at contig 1153) near a *N. coriiceps* methionine synthase reductase-like mRNA gene (LOC104943965) correlated ($p<0.001$) with lower mean ice thickness and lower amplitude in sea floor temperature (Fig. 4.4c). One SNP possibly near an orthologue of the *Dissostichus mawsoni* haplotype 1 AFGP/TLP gene locus (ID: HQ447059.1) was detected by both methods from the reference data. This locus was correlated ($p<0.001$) with various environmental descriptors, including ice cover range, maximum chlorophyll *a* concentration and the amplitude of both sea floor salinity and temperature (Fig. 4.4e).

**Table 4.5.** Candidate single nucleotide polymorphism (SNP) loci of *Notothenia coriiceps* detected to be putatively under selection and potentially linked to genes with different functions. Maximum distance to gene refers to the size window employed in the BLAST search; data set refers to the SNP calling method, either from mapping against the *N. coriiceps* reference genome or a *de novo* catalog; detection method is the genome scan method, either pcadapt (P) or BayeScan (B), and detection threshold is the q-value cut-off at which the respective SNP was detected.

| Number of SNPs | SNP position | Relevant gene in proximity | Max. distance to gene (bp) | Gene function | Data set | Detection method | Detection threshold |
|---|---|---|---|---|---|---|---|
| 1 | Contig_690_73 | PREDICTED: *Notothenia coriiceps* mitogen-activated protein kinase-binding protein 1-like (LOC104956467), mRNA | 192 | immune system | *de novo* | P | 0.1 |
| 2 | Contig_727_91&94 | PREDICTED: *Anabas testudineus* anthrax toxin receptor 1-like (LOC113159039), transcript variant X2, mRNA | 192 | cell attachment and migration | *de novo* | P | 0.1 |
| 1 | Contig_865_81 | PREDICTED: *Notothenia coriiceps* histone-lysine N-methyltransferase 2C-like (LOC104961987), mRNA | 192 | gene transcription, transcription regulation | *de novo* | P | 0.1 |
| 7 | Contig_1153_12&84&86&114 &132&186&188 | PREDICTED: *Notothenia coriiceps* methionine synthase reductase-like (LOC104943965), mRNA | 192 | methionine biosynthesis | *de novo* | P & B* | 0.05 |
| 4 | Contig_2056_79&177, gi_734298258_ref_NW_011348685_1_31128, | PREDICTED: *Notothenia coriiceps* serine/threonine-protein phosphatase 1 | 192, 192, 210, | control of chromatin | 2x *de novo* & 2x reference | P | P: 0.01 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | gi_734298258_ref_NW_0113 48685_1_31349 | regulatory subunit 10-like (LOC104949365), mRNA | 70 | structure during the cell cycle | | | |
| 5 | Contig_1316_175, 1463_65&172, 1753_114, 1769_140 | no match | 192 | - | *de novo* | P | 0.1 |
| 4 | Contig_695_160&161, 1441_15&108 | no match | 192 | - | *de novo* | P | 0.05 |
| 1 | Contig_554_17 | no match | 192 | - | *de novo* | P & B | 0.05 |
| 1 | Contig_1117_103 | PREDICTED: *Notothenia coriiceps* uncharacterized 1LOC104945202 (LOC104945202), mRNA | 192 | unknown | *de novo* | B | 0.1 |
| 1 | Contig_1244_104 | PREDICTED: *Notothenia coriiceps* interleukin-12 receptor subunit beta-2-like (LOC104952077), mRNA | 192 | immune system | *de novo* | B | 0.1 |
| 1 | Contig_1507_145 | PREDICTED: *Notothenia coriiceps* uncharacterized LOC104943630 (LOC104943630), mRNA | 192 | unknown | *de novo* | B | 0.05 |
| 2 | Contig_167_105, 397_147 | no match | 192 | - | *de novo* | B | 0.1 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 2 | Contig_908_12, 2363_102 | no match, but corresponds to reference locus gi_734300000_ref_NW_011348291_1_225190 | 192 | - | *de novo* | B | 0.05 |
| 2 | gi_734200844_ref_NW_011369947_1_3694883, gi_734222479_ref_NW_011365181_1_1452 | *Dissostichus mawsoni* haplotype 1 AFGP/TLP gene locus, partial sequence; Sequence ID: HQ447059.1 | 210, 210, 192 | antifreeze glycoprotein | reference | P & B,<br>P | 0.1 & 0.05<br>0.1 |
| 2 | gi_734222479_ref_NW_011365181_1_1197, gi_734222479_ref_NW_011365181_1_1228 | PREDICTED: *Notothenia coriiceps* uncharacterized LOC104966052 (LOC104966052), ncRNA; Sequence ID: XR_800615.1 | 490 | unknown | reference | P | 0.1 |
| 1 | gi_734200844_ref_NW_011369947_1_15486379 | PREDICTED: *Notothenia coriiceps* lumican-like (LOC104959963), mRNA | 350 | collagen binding | reference | P | 0.1 |
| 1 | gi_734200844_ref_NW_011369947_1_15486442 | *Chionodraco hamatus* partial MT gene for metallothionein-I, promoter sequence | 350 | regulation of metal ion presence | reference | P | 0.1 |
| 9 | gi_734200936_ref_NW_011369945_1_2184502, gi_734323561_ref_NW_011344899_1_50880, gi_734399223_ref_NW_011335323_1_67672 & 67688 & | *Boreogadus saida* MAK14 and RAB14 genes and AFGP1 (AFGP1) pseudogene, complete sequence; antifreeze glycoprotein polypeptide 2 (AFGP2), antifreeze glycoprotein polypeptide 3 (AFGP3), and antifreeze glycoprotein polypeptide 4 | 1845 | antifreeze glycoprotein | reference | P | 0.1 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | 67786 &67802 & 67810 & 67817 & 67848 | (AFGP4) genes, complete cds; and antifreeze glycoprotein polypeptide 5 (AFGP5) gene AFGP6 and AFGP7 (AFGP7) pseudogenes, complete sequence. | | | | | |
| 1 | gi_734211745_ref_NW_011367634_1_304678 | PREDICTED: *Notothenia coriiceps* X-linked retinitis pigmentosa GTPase regulator-like (LOC104949670), mRNA; Sequence ID: XM_010776061.1 | 1190 | cell repair/structure maintenance | reference | B | 0.05 |
| 1 | gi_734218697_ref_NW_011366068_1_13258 | *Larimichthys crocea* genome assembly, chromosome: VII; Sequence ID: LT972196.1 | 630 | unknown | reference | B | 0.1 |
| 1 | gi_734300000_ref_NW_011348291_1_225190 | *Scophthalmus maximus* chromosome 12; Sequence ID: CP026254.1 | 910 | unknown | reference | B | 0.05 |

* only SNPs at position Contig_1153_84 and Contig_1158_186 were detected by BayeScan; all were detected by pcadapt.

## 4. Discussion

### 4.1 Connectivity and evolution of *Notothenia coriiceps*

The bullhead notothen *Notothenia coriiceps* is not composed of a single, panmictic, circum-Antarctic population, but rather of at least two genetically differentiated populations in the Eastern Antarctic (Adélie Land) and the western Antarctic Peninsula. Taxonomists have sometimes considered *Notothenia neglecta* (Nybelin, 1951) as a species distinct from *Notothenia coriiceps* (Balushkin 2000), but Dewitt et al. (1990) list both species as synonymous. Without detailed morphological data it is impossible to assess whether the populations identified here also exhibit morphological variation as noted by the proponents of *N. neglecta* as a separate species. Other data types do not indicate that the species should be split. For example, prominent chromosome fusions are present in *N. coriiceps*, but no indications of karyotypic variation across space were detected (Ozouf-Costaz *et al.* 1991; Ghigliotti *et al.* 2015; Amores *et al.* 2017). Furthermore, life history trait data are available for *N. coriiceps* from many years of study, but overwhelmingly from specimens collected in the western Antarctic Peninsula (e.g. Linkowski & Zukowski 1980; Kellermann *et al.* 2002; Zamzow *et al.* 2010; Ferreira *et al.* 2017). It would be interesting to compare such ecological data across Southern Ocean bioregions, as has been done for example in *N. rossii* (Calì *et al.* 2017). The overall low levels of genetic differentiation observed here suggest regular, large scale gene flow between opposite sides of Antarctica. Therefore, we recommend to maintain *N. coriiceps* as one, likely circum-Antarctic species, but to recognize that intraspecific variations with local populations are present. Because our sampling covers only parts of *N. coriiceps*' distribution range (Duhamel *et al.* 2014), it is likely that more than two sub-populations are present. Likewise, existing knowledge regarding other notothenioid fishes is spatially fragmented (Volckaert *et al.* 2012; Damerau *et al.* 2014; Agostini *et al.* 2015; Young *et al.* 2015), primarily because of country-specific (i.e. regionally focused) expeditions and general logistic difficulties in the Southern Ocean.

There are examples of marine Antarctic organisms that exhibit no genetic population structure in the entire Southern Ocean and seem to prevail in conditions near panmixia, such as benthic invertebrates (Thornhill *et al.* 2008; Díaz *et al.* 2018), pelagic krill (Deagle *et al.* 2015), penguins (Clucas *et al.* 2016; Mura-Jornet *et al.* 2018), or even kelp (Fraser *et al.* 2009). Genetic homogeneity across large distances was also documented in the Antarctic notothenioid fishes *Gobionotothen gibberifrons* (Matschiner *et al.* 2009), *Chionodraco rastrospinosus* (Papetti *et al.* 2012) and *N. rossii* (Young *et al.* 2015). Contrastingly, and more congruent with our results, a

scenario of at least two sub-populations was uncovered in many other Antarctic species, including benthic invertebrates (Arango *et al.* 2011; Hauquier *et al.* 2017; Strugnell *et al.* 2017), seals (Bonin *et al.* 2013; Corrigan *et al.* 2016), and notothenioid fish species. Among the latter are examples of differentiated sub-populations at comparatively isolated sub-Antarctic islands, such as Bouvet Island (Damerau *et al.* 2014) and South Georgia (Van de Putte *et al.* 2012b; Damerau *et al.* 2014; Young *et al.* 2015), but also along the Antarctic Peninsula region (Agostini *et al.* 2015) and between East and West Antarctica (Van de Putte *et al.* 2012b). We observed genetic differentiation of *N. coriiceps* between the Dumont d'Urville Sea (East Antarctica) and samples from the western Antarctic Peninsula, which has likewise been documented in three *Trematomus* species (Van de Putte *et al.* 2012b). Our results also demonstrate a slight differentiation of specimens from the South Orkney Islands relative to all other regions. Similar patterns were observed in *Champsocephalus gunnari*, *Chaenocephalus aceratus*, and even in (the otherwise genetically comparatively uniform) *N. rossii* (Papetti *et al.* 2009; Damerau *et al.* 2014; Young *et al.* 2015). Overall, the extent of genetic differentiation in Antarctic fishes appears variable, from virtual panmixa, to the case of *Lepidonotothen nudifrons* (or *Nototheniops nudifrons*, see Near et al. 2018), where genetic differentiation between South Georgia and the Antarctic Peninsula was so distinct that these groups of fish should be considered two different species (Dornburg *et al.* 2016). And yet, spatial patterns seem more general and applicable across some species at least as evidenced with the South Orkney differentiation, even if the degree of divergence varies (which might also be related to methodological differences). What is the driving mechanism of observed genetic differences then?

Generally, three different scenarios are often considered as causes for spatial genetic population structure in the ocean: (1) contemporary remnants of historical constraints, such as for example separation into glacial refugia (Hewitt 2000; Hauser & Carvalho 2008), (2) ocean currents, that determine connectivity between populations (White *et al.* 2010), and (3) environmental segregation, where local conditions impose selection on incoming phenotypes (Savolainen *et al.* 2013). The relative importance of all three processes is determined to a large extent by the life history strategy of a given species (Janko *et al.* 2007; Van de Putte *et al.* 2012b; Kašparová *et al.* 2015; Moreau *et al.* 2017). Historical signatures seem especially important in high latitudes (Allcock & Strugnell 2012; Bowen *et al.* 2016). Several studies demonstrated effects of past glaciation cycles on Antarctic fishes (Janko *et al.* 2007; Near *et al.* 2012; Damerau *et al.* 2014). *Notothenia* species are between 5 and 10 mya old (Near *et al.* 2012), which coincides with the re-establishment of permanent continental ice sheets in East (10 mya) and West (5 mya) Antarctica

(Zachos *et al.* 2001). *Notothenia coriiceps* adults in particular are most commonly found in nearshore areas and are thus likely to directly experience glacial extension and retraction cycles. As a mobile species, however, they can easily react to such cycles (at least in time periods of thousands of years) as long as some refugia such as South Georgia remain viable habitat. More sedentary species may be more prone to characteristic extinction-recolonization signatures (Clarke & Crame 2010). Oceanographic connectivity may be particularly high in *N. coriiceps* due to its prolonged pelagic life stage. The Antarctic Circumpolar Current (ACC) is likely a major force that connects populations via their pelagic offspring around Antarctica and thus renders neutral spatial genetic population structure subtle at most. Other Antarctic fish species with pelagic larvae/juveniles or adults similarly show high levels of connectivity (Young *et al.* 2015; Caccavo *et al.* 2018). Life history and oceanic forcing act together in these examples to support widely distributed and well-connected populations.

### 4.2 Local adaptation of *Notothenia coriiceps* shaped by the environment?

Environmental selection as a constraint on incoming propagules (i.e. settling juveniles in this case) has received comparatively little attention in Antarctic research. Marine fish in particular with their wide distribution ranges and large effective population sizes can be subject to strong environmental selection, even if this is difficult to detect with few genetic markers (Teacher *et al.* 2013; Vandamme *et al.* 2014). We found indications for genotype-environment-association that may suggest some degree of local adaptation in *N. coriiceps*. Although it must be noted because of our wide and clustered sampling design, it is challenging to account for spatial effects. Nevertheless, it is possible that different environmental settings are constraining fingerlings of the bullhead notothen when searching for a suitable habitat. Much of the observed genotype-environment association also correlates with geography, however, so that more contiguous sampling campaigns are needed to clearly determine which abiotic factors are most important. Standardized sampling along environmental gradients are most powerful to explore genotype-environment associations. This is not an easy task but highly relevant given the rapid changes occurring at the western Antarctic Peninsula (Vaughan *et al.* 2003; Turner *et al.* 2016).

Overall, the present genetic differentiation observed here can be traced to a few loci that are strikingly divergent between sub-populations from East and West Antarctica. These loci are assorted in groups that likely diverge together with the causative genetic variation through physical linkage. This hitchhiking effect is now widely recognized (Manel *et al.* 2016). Some of

the candidate loci here may be false positives and need independent verification, but we present those as targets for further study. Three cases strike us as likely important genetic variation. First, the SNP on *de novo* contig 554 is strongly correlated to higher variability in ice cover and could be a target of very recent selection pressure. This genetic variant was clearly differentiated between East and West Antarctic fish. The locus could not be matched to any gene, but with rising genomic resources this may be possible in the future. Another instance are the seven SNPs linked to a *N. coriiceps* gene likely involved in methionine biosynthesis. Methionine is over-represented in notothenioids, which may be an adaptation to cold, highly oxygenated waters (Berthelot *et al.* 2018). Furthermore, metabolic profiling showed that *N. rossii* has higher levels of free methionine in gill tissue than *N. coriiceps*, possibly reflecting the higher base metabolism of *N. rossii* (Rebelein *et al.* 2018). It seems plausible that the signal documented here reflects contemporary adaptation to a more active life style under warming also at the intraspecific level. *Notothenia coriiceps* experiences winter metabolic suppression (Campbell *et al.* 2008), which may be more pronounced in higher latitude areas with longer sea ice cover such as Adélie Land. In contrast, temperatures in the western Antarctic Peninsula have increased over the past 70 years by 2.5 °C and are predicted to continue to rise (Turner *et al.* 2005; IPCC 2019). Moving towards a more active life style may be a recent response of local fish populations to this climate change. The mRNA LOC104943965 we detected here as putatively under selection is a starting point in the investigation of genome-wide cold adaptation as proposed by Berthelot *et al.* (2018).

Finally, two SNPs putatively under selection are in the vicinity of an anti-freeze glycoprotein (AFGP) gene. The development of AFGP was a key innovation enabling notothenioids to thrive in subzero waters (Chen *et al.* 1997; Near *et al.* 2012). However, non-Antarctic notothenioids have lost the antifreeze function (Miya *et al.* 2016a) indicating that selection may act upon this gene in warmer environments. Although the functional link remains to be validated, the signal we find here may be an indication of relaxed constraints to keep the AFGP gene intact in warming West Antarctic waters. Further, higher density, population genomic studies of *N. coriiceps* are likely to reveal additional candidates responsible for local adaptation.

### 4.3 Implications for conservation and management

Fisheries as well as protection activities in the Southern Ocean are governed by the Commission for the Conservation of Antarctic Marine Living Resources (CCAMLR). Currently, no location assessed in this study falls within a marine protected area (MPA). The shelf and high sea area

south of the South Orkney Islands and several areas in the Ross Sea are the first Southern Ocean MPAs (CCAMLR 2009, 2016). Several initiatives have proposed a network of Antarctic MPAs around the entire continent. However, the proposed establishment of the "East Antarctic Marine Protected Area", which would include the coast of Adélie Land, was rejected again in 2018, despite its significant reduction in proposed size. Our results indicate that *N. coriiceps* in East Antarctica harbor genomic diversity that is differentiated from fishes from the western Antarctic Peninsula, which should be considered in conservation planning.

Fisheries currently do not target *N. coriiceps*, but e.g. juveniles of the species may be caught as by-catch during krill trawl fisheries (Rembiszewski *et al.* 1978). Furthermore, the commercial interest in this species may increase in the future due to declines in other commercially exploited species (Ainley & Pauly 2013). Fisheries mortality in combination with rapid environmental change may decrease the overall fitness of the *N. coriiceps* metapopulation and could potentially lead to local extinctions in certain areas. Scientific sampling alone was found to have profound effects on a local *N. coriiceps* population (Casaux & Barrera-Oro 2002). In addition, strong site fidelity has been documented in adult *N. coriiceps* at different localities and spatial scales (Barrera-Oro & Casaux 1996; North 1996). Consequently, individual populations may not easily be repopulated through larvae from other source areas. Even if recolonization occurs, temporal local extinction can have devastating effects on the genetic diversity of a species (McCauley 1997). Potential for rapid adaptation is primarily rooted in the standing genetic variation of a species (Bernatchez 2016). Through local extinction genetic diversity and therefore adaptive potential may be lost, which decreases the overall likelihood of a species to cope with changing environmental pressures.

Carefully designed MPA networks that cover the entire range of the Southern Ocean are indispensable to protect genetically differentiated species such as *Notothenia coriiceps*. The distances between MPAs are ideally determined by the realized dispersal, which may change under future environmental change (see e.g. Young *et al.* 2018). Most Antarctic species will struggle with future climate change (Griffiths *et al.* 2017) and therefore supporting and protecting population resilience in the shape of spatially variable genomic diversity is highly relevant. These processes are ideally guided by regular ecological and genetic monitoring to enable long term benefits of planned and realized MPAs.

#### 4.4 Creating population genomics data for non-model species

Lastly, our results provide guidance for molecular ecologists applying RADseq and related approaches on non-model species. We compared results from *de novo* and reference mapped data sets to investigate whether exclusively relying on reference mapping to call SNPs is advisable in species with an available genome. Many of the observed patterns were similar in both cases. Large scale genetic structure for instance as identified through pairwise $F_{ST}$, PCA, and Bayesian clustering approaches was nearly identical, albeit with differing numbers of SNPs. Interestingly, we obtained more SNPs using the *de novo* catalog than the reference genome (19.37 % or 382 loci) and detected signals of selection on different candidate loci (30 in *de novo* and 20 in reference, only one in common). Other studies obtained more SNP genotypes using reference genomes (Torkamaneh *et al.* 2016; Shafer *et al.* 2017). Using a reference genome is also advocated as this facilitates easier mapping of potentially relevant loci, at least in closely related species (DiBattista *et al.* 2017). Yet, in our case the exact position of SNPs mapped to the reference genome often remains unknown, while the approximate position of SNPs from the *de novo* data set could be inferred in many instances through BLAST searches as well.

Several possibilities may explain the different SNP numbers and selection signals we found between *de novo* and reference data sets. Firstly, reference genomes are rarely complete, but instead contain gaps to various extents. The reference or draft genome of *N. coriiceps* is 637 Mb, although Detrich et al. (2010) cytometrically estimated the genome of *N. coriiceps* to be approximately 1105 Mb long (estimated at C = 1.13 ± 0.21). Another, very recent study suggests the genome is even larger (C = 1.36 ± 0.12; Auvinet et al. 2018). Results from flow cytometry and the reliably known gaps (13.1 Mb, Shin et al. 2014) clearly illustrate that the genome we used here is not complete. Therefore, sequences that we obtained from regions not present in the reference are likely discarded in the process of reference mapping. Secondly, the reference genome can also introduce spatial ascertainment bias. Shin et al. (2014) sequenced DNA from an individual from King George Island in the western Antarctic Peninsula. Our samples include specimens from remote Adélie Land in East Antarctica, which may lead to dropout of, for example, SNPs that are not present in the reference individual due to structural changes in that genomic region. High quality reference resources from different spatial scales present the next step in the genomic era of non-model species (Lewin *et al.* 2018) potentially resolving spatial ascertainment bias. Thirdly, repetitive regions or paralogues in a genome may erroneously inflate the number of SNP calls particularly in *de novo* approaches. The genome of *N. coriiceps*

is predicted to contain between 13.4 % (Detrich *et al.* 2010) and 18.15 % (Shin *et al.* 2014) of repetitive regions, hence some inaccurate SNP calls may result from repeated stretches of DNA. More inaccurate SNPs have been found with *de novo* SNP calling (Torkamaneh *et al.* 2016), but these are not necessarily derived exclusively from repetitive regions or paralogues. Ultimately, we cannot clarify this last point without whole genome resequencing and/or more detailed genomic resources of *N. coriiceps*. While it needs to be stressed that the initial sets of "outlier loci" are merely candidates for further study, combining *de novo* and reference SNP calling methods might provide a broader basis for such investigations. At least two different *de novo* loci possibly linked to different genes show strong indications of ongoing selection. Even if all other candidates were false positives, this demonstrates how reference mapping alone may miss relevant signals present in the data. As long as reference genomes from non-model species predominantly depict a single snapshot of genomic variability in space and time, we recommend that population scale genomic investigations explore both *de novo* and reference SNP calling approaches, especially when scanning for potentially adaptive loci. Such comparative analyses can also increase confidence when similar large-scale results (e.g. population structure) are obtained or else point to methodological problems during SNP calling that are common among pipelines (Torkamaneh *et al.* 2016; Shafer *et al.* 2017).

## 5. Conclusion

Population scale high throughput sequencing using *de novo* and reference-based SNP calling in parallel revealed that *Notothenia coriiceps* is likely subject to ongoing selection processes that differentiate populations from East and West Antarctica despite high levels of connectivity. It is crucial to monitor and account for such variability in present conservation plans to enable this species to persist under global climate change. Together with few other studies (Galaska *et al.* 2017; Humble *et al.* 2018), these results mark the advent of high-throughput population genomics in the Southern Ocean.

**6. Acknowledgements**

We thank everyone that helped to collect and share samples, in particular J. Rock, the captain, crew and scientific participants of cruise JR 145 aboard RV James Clark Ross, British Antarctic Survey, and P. Koubbi and all participants of the ICOTA (Ichthyologie côtière en Terre Adélie) programme. B. Hellemans, J. Hillen and T. Smit were of invaluable help during molecular lab work. Thanks also to I. Coscia, S.R. Maier, G. Maes, and C. André for helpful discussions.

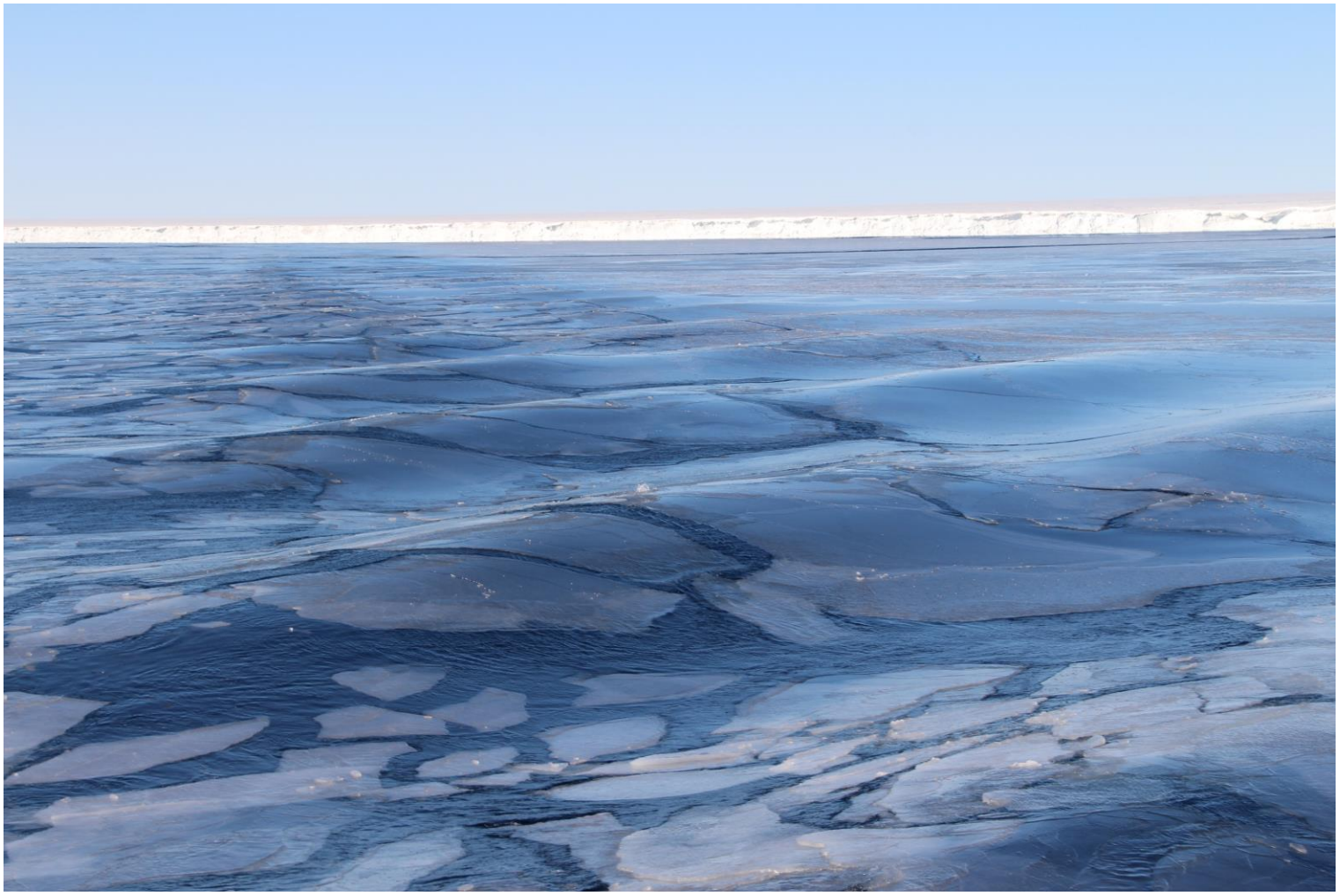
**7. Funding**

This work is contribution no. 32 to the vERSO project, funded by the Belgian Science Policy Office (BELSPO, Contract no. BR/132/A1/vERSO) and contribution no. 7 to the RECTO project (BELSPO, Contract no. BR/154/A1/RECTO). HC was supported by an individual PhD grant from the former Flemish Agency for Innovation by Science and Technology (IWT), now managed by Flanders Innovation & Entrepreneurship (VLAIO, grant no. 141328).

**8. Data archiving statement**

Demultiplexed, but otherwise raw sequencing data has been deposited on NCBI's Sequence Read Archive (SRA) under accession PRJNA533009 (reviewer link: https://dataview.ncbi.nlm.nih.gov/object/PRJNA533009?reviewer=k7oq1d2fogb9o0q4nfh3cqk6an). Metadata for each individual using the same identification codes is available and cross-linked at GeOMe (Deck *et al.* 2017) under https://n2t.net/ark:/21547/ClX2. In addition, vcf files, R scripts and additional input files for analysis are available at http://doi.org/10.5281/zenodo.2650260.

# CHAPTER 5: Integrated assessment of large-scale connectivity in a historically overexploited fish population in the Southern Ocean



*Note*: this manuscript is in preparation for publication.

**Integrated assessment of large-scale connectivity in a historically overexploited fish population in the Southern Ocean**

Henrik Christiansen, Anton P. Van de Putte, Charlène Guillaumot, Esteban Barrera-Oro, Filip A. M. Volckaert and Emma F. Young

**Abstract**

Natural populations are often heterogeneously distributed in space, potentially leading to spatial genetic structure and patterns of local adaptation. Ecological traits, the environment, and evolutionary forces determine the connectivity between population patches. In the Southern Ocean, demersal fish disperse widely between the continental shelf, and oceanic island plateaus and seamounts. The marbled rockcod *Notothenia rossii* Richardson, 1844 was historically overharvested and only recently shows signs of recovery. We applied an integrated multidisciplinary approach to determine connectivity of this ecologically important species over contemporary and evolutionary time scales. Thousands of population genomic markers reveal high levels of gene flow and a lack of genetic differentiation over vast distances. Individual-based modelling, however, suggests that large-scale connectivity can only be achieved via stepping-stone transport. In conjunction with species distribution modelling, these results highlight how genetic data alone may overestimate the extent of connectivity. Limited ecological connectivity, that is reduced exchange of larvae or juveniles within one season, and reduced effective population size may have contributed to the long recovery time of the marbled rockcod. Current conservation plans, that aim to create a network of marine protected areas in the Southern Ocean, can benefit from multi-method assessments as presented here, especially in view of global change.

## 1. Introduction

Connectivity and spatial genetic structure in marine organisms is determined through the interplay of ecological traits, such as dispersal mode, duration and behavior, and the physical setting, that is environmental conditions including hydrodynamics, and evolutionary forces, such as selection and genetic drift (Hidalgo *et al.* 2017; Hoey & Pinsky 2018; Xuereb *et al.* 2018). This multitude of factors can lead to complex patterns and makes the relative importance of each factor difficult to discriminate (Moon *et al.* 2017; Miller *et al.* 2018; Milligan *et al.* 2018). However, spatial population structure and its temporal dynamics are crucial information for sound biodiversity management and protection (Funk *et al.* 2012; Momigliano *et al.* 2019). Management of marine organisms aims to protect the biodiversity of species, populations and ecosystems under competing influences of various anthropogenic disturbances (Everson 2015; Ropert-Coudert *et al.* 2019).

Taxonomic ranks at the species level are readily available, at least for macro-organisms, but recent research has shown that it is imperative to also consider intraspecific variation (Mee *et al.* 2015; Carvalho *et al.* 2017; Des Roches *et al.* 2018; Paz-Vinas *et al.* 2018). In order to assess intraspecific variation in the ocean, putative subpopulations are characterized with respect to their ecology and evolution. These two aspects are intertwined, leading to eco-evolutionary dynamics that determine the long-term fate of a species. Similarly, a species may persist in the form of a metapopulation comprised of subpopulations that are linked through ecological and evolutionary connectivity (Cowen & Sponaugle 2009; Pinsky *et al.* 2017). Ecological connectivity is the contemporary exchange of individuals between fragmented habitats. Evolutionary connectivity additionally considers the long-term degree of connection between a given number of separate (sub)populations through genetic exchange and drift (Waples & Gaggiotti 2006). All these aspects gain relevance in an exploitation context. Fisheries in particular can have drastic and immediate effects on reproductive output and thus ecological connectivity. In addition, fisheries can reduce genetic diversity and may have evolutionary consequences by imposing artificial selection on a species (Pinsky & Palumbi 2014; Heino *et al.* 2015). Recognizing and mitigating these consequences is a global challenge for fisheries and ocean management, especially in areas beyond national jurisdiction (Ortuño Crespo *et al.* 2019).

The Southern Ocean provides an example of ecosystem- and consensus-based fisheries management in an area that is not governed by a single nation (Kock *et al.* 2007; Constable 2011;

Everson 2015). As such the Commission for the Conservation of Antarctic Marine Living Resources (CCAMLR), can be considered progressive (Constable *et al.* 2000; Nilsson *et al.* 2016; Hofman 2019), albeit constantly facing challenges (Ainley & Pauly 2013; Brooks 2013; Abrams *et al.* 2016). Before CCAMLR came into force in 1982, a number of fish populations were severely overfished. Most strikingly, the endemic marbled rockcod *Notothenia rossii* Richardson, 1844 (Nototheniidae, Perciformes) was an early target species with a cumulative reported catch of 501,262 t in the first two fishing seasons (1969/70 and 70/71) around South Georgia (Kock 1992) and near 150,000 t around Kerguelen in 1971 (Duhamel 1982). Thereafter, large trawlers were still active throughout the Southern Ocean, but with considerably lower catches, until the fishery was closed by CCAMLR in 1986/87 (Kock 1992). Since the inception of CCAMLR, conservation measures have been adopted progressively in order to assist the recovery of several notothenioid species by banning directed fisheries and establishing stringent by-catch limits in many Antarctic zones (CCAMLR 2019a). Recovery of the *N. rossii* stocks took more than 35 years with the species only in the past decade showing clear signs of increasing abundance (Barrera-Oro & Marschoff 2007; Marschoff *et al.* 2012; Barrera-Oro *et al.* 2017; Duhamel *et al.* 2017).

At present, new initiatives are underway to enhance biodiversity protection in the Southern Ocean through a network of Marine Protected Areas (MPAs). The South Orkney Islands Southern Shelf (SOISS) and large parts of the Ross Sea are designated MPAs since 2009 and 2016, respectively (Fig. 5.1). The waters around South Georgia and the South Sandwich Islands are also widely protected against overexploitation (Trathan *et al.* 2014). Additional MPAs were proposed in the Atlantic, Indian and Pacific sectors of the Southern Ocean to create a MPA network (CCAMLR 2019b). Much research has been devoted to provide the scientific information needed for the establishment of an appropriate MPA network (Teschke *et al.* 2015, 2019; Constable *et al.* 2016; Hill *et al.* 2017; Brasier *et al.* 2019; Parker *et al.* 2019). In addition, several large-scale research initiatives have increased our knowledge of biodiversity and bioregions in the Southern Ocean over the past decades (Schiaparelli *et al.* 2013; De Broyer *et al.* 2014). Nevertheless, Antarctica and the Southern Ocean remain a data-poor environment in the global context. Data availability is biased, due to high sampling effort around research bases (Griffiths 2010) and difficult accessibility of the remote and often ice-covered Southern Ocean (Convey & Peck 2019). A combined methodological approach, including genetic analyses, environmental measurements and modelling techniques, could help eliminate knowledge gaps concerning population structure and connectivity of key species in the Southern Ocean ecosystems (Gutt *et al.* 2017).

The genomic revolution has enabled the relatively fast and economic characterization of thousands of genetic markers in non-model organisms (Elshire *et al.* 2011; Andrews *et al.* 2016; Christiansen *et al.* Chapter 3). High resolution genetic data promises to yield new insights into speciation, differentiation and adaptation patterns and is thus a valuable tool to fully utilize the potential of precious Antarctic samples (Christiansen *et al.* Chapter 4). Species distribution modelling (SDM), sometimes referred to as ecological niche modelling, is a powerful technique to correlate environmental and occurrence data and subsequently predict the occurrence probability of a given species in other habitats or under changing environmental conditions (Elith & Leathwick 2009). Such techniques can be useful in data-limited situations, albeit care must be taken to ensure appropriate parameterization (Guillaumot *et al.* 2018a, 2019). Lastly, individual-based modelling (IBM) can be used to simulate dispersal and thus obtain a spatially explicit prediction of connectivity between habitats (Cowen & Sponaugle 2009). Combining genomic data and modelled connectivity estimates is most useful in the marine realm, where direct observations of migration and dispersal are often virtually impossible (Pinsky *et al.* 2017; Xuereb *et al.* 2018). We use the methods mentioned above complimentarily to advance our understanding of large-scale connectivity in *N. rossii*, a fish that is both valuable as a living resource and vulnerable to overfishing.

The marbled rockcod grows to more than 50 cm in length, can form dense shoals in sub-Antarctic and Antarctic fjords and shelf waters, and occurs widely in the Southern Ocean (DeWitt *et al.* 1990; Duhamel *et al.* 2014). Its life cycle has been well described for the population at South Georgia by Olsen (1954) and Burchett (1983) and for the population at Kerguelen Islands by Duhamel (1982). Spawning takes place between April and June on the bottom of continental shelf areas at about 200-360 m depth, where ripe adults migrate during fall. Hatching occurs between September and October in the water column, where larval and young pelagic blue-phase fingerling stages remain before they migrate inshore approximately in January-February (DeWitt *et al.* 1990, Kock & Kellermann 1991; Kock *et al.* 2000; North 2001). The fingerlings then change morphologically to the brown-phase fingerling stage and become demersal, settling in the algae beds. At about 5-7 years of age and 41-45 cm of length, *N. rossii* reaches sexual maturity and migrates to the offshore shelf feeding area joining the adult population. These offshore-inshore phases in the life cycle of the marbled rockcod are assumed to be similar in the geographical areas of its range (Kock 1992). The ecological habits of *N. rossii* as a benthic-benthopelagic species constitute an important trophic link between lower trophic

levels (macroalgae, benthic invertebrates, small fish) and Antarctic top predators, such as seals and birds (Barrera-Oro 2002; McInnes *et al.* 2017; Bertolin & Casaux 2018).

Previous studies have reasoned that this extended pelagic period contributes to the widespread distribution and low or absent genetic structure of *N. rossii* (DeWitt *et al.* 1990; Young *et al.* 2015). The Antarctic Circumpolar Current (ACC), the world's largest ocean current system, is a prime candidate to facilitate eastward advection and thus connectivity (Orsi *et al.* 1995; Matschiner *et al.* 2009; Christiansen *et al.* Chapter 2). The ACC is comprised of a series of approximately zonal fronts, where there are rapid changes in water mass properties and associated geostrophic currents. These current jets are not fixed in time and space, rather they show a high degree of mesoscale variability, with frequent splitting and merging. However, long-term trends in frontal positions show very little variability (Chapman 2017), in part due to the strong steering of the ACC by the seafloor topography. Broadly speaking, the ACC has a pervasive eastward flow. However, within the Scotia Sea this flow assumes a more northward component as the current system recovers from its most southerly excursion in Drake Passage. Population genetic investigations of *N. rossii* have focused on this area, using microsatellite loci and connectivity models (Young *et al.* 2015). However, the connection between *N. rossii* in the Scotia Sea and the Kerguelen Plateau, initially described as two subspecies (DeWitt *et al.* 1990), has not been investigated with modern genomic methods or numerical modelling.

Here, we hypothesize that despite the fragmented distribution only subtle genetic structure and adaptive divergence are present in *N. rossii* due to the long pelagic phase. Species distribution modelling is used to determine which localities are likely important habitats for *N. rossii*, while genome-wide polymorphisms are used to test for genetic population structure. Individual-based modelling is employed to quantify dispersal between sites. Combining population genomics with distribution and dispersal modelling enables us to infer gene flow and environmental specialization in a spatially explicit framework. Finally, we evaluate large-scale distribution and spatial connectivity patterns in the Southern Ocean in light of current fisheries management and conservation actions.

## 2. Material and methods

### 2.1 Species occurrence and sampling

Publicly available occurrence data for marbled rockcod *Notothenia rossii* were mined from the Ocean Biogeographic Information System (OBIS) and the Global Biodiversity Information Facility (GBIF) (both accessed September 2019) using R (R Core Team 2019) with the packages 'robis' v2.1.8 (Provoost & Bosch 2019), 'rgbif' v1.3.0 (Chamberlain *et al.* 2019) and 'SOmap' for plotting (Maschette *et al.* 2019). Duplicate entries (identical coordinates) were removed. A total of six occurrences that appear likely to be misidentified (based on these occurrences being drastically outside the generally accepted species distribution) were removed, in an attempt to provide exclusively highly reliable input data for species distribution modelling. For genetic analyses, adult fish were caught during many expeditions throughout the Southern Ocean (Fig. 5.1, Table 5.1). Fin, muscle, or liver biopsies were taken and stored in 90% ethanol or frozen until further processing. The samples from South Georgia, the South Orkney Islands, Elephant Island, and King George Island/Isla 25 de Mayo (collected in 2006) were previously analyzed using microsatellites (Young *et al.* 2015). In addition, samples from trammel nets were taken in 2016 on King George Island/Isla 25 de Mayo (Barrera-Oro *et al.* 2019) and from research trawling in 2016 on Skiff Bank (Leclaire Rise) and on the Northeast part of the Kerguelen Islands shelf. The occurrence data from these samples were added to the OBIS/GBIF occurrence data set for species distribution modelling if not already included. All available metadata per sampled individual can be found on data.biodiversity.aq.

**Table 5.1.** Sampling details (location, location code, latitude (Lat) and longitude (Lon), sample size (N) and year) and genetic diversity of *Notothenia rossii* from the Southern Ocean. Geographical coordinates are listed in decimal degrees; note that values are approximate for most locality samples; all available metadata per individual can be found on data.biodiversity.aq. Expected ($H_E$) and observed heterozygosity ($H_O$) was calculated for filtered genotypes from *de novo* and reference-based bioinformatics.

| Location | Code | Lat | Lon | N | Year | $H_O$ *de novo* | $H_O$ reference | $H_E$ *de novo* | $H_E$ reference |
|---|---|---|---|---|---|---|---|---|---|
| South Shetlands, Deception Island | SSD-06 | -62.95 | -60.65 | 34 | 2006 | 0.246 ± 0.004 | 0.212 ± 0.007 | 0.253 ± 0.005 | 0.219 ± 0.006 |
| South Shetlands, King George Island | SSK-06 | -62.23 | -58.68 | 35 | 2006 | 0.230 ± 0.004 | 0.217 ± 0.006 | 0.235 ± 0.004 | 0.227 ± 0.006 |
| South Shetlands, King George Island | SSK-15-16 | -62.23 | -58.68 | 40 | 2015/16 | 0.244 ± 0.004 | 0.213 ± 0.006 | 0.249 ± 0.004 | 0.223 ± 0.006 |
| Elephant Island | EI-02 | -61.24 | -55.62 | 33 | 2002 | 0.255 ± 0.004 | 0.215 ± 0.006 | 0.261 ± 0.004 | 0.226 ± 0.006 |
| Elephant Island | EI-06-07 | -61.24 | -55.62 | 31 | 2006/7 | 0.253 ± 0.005 | 0.211 ± 0.006 | 0.260 ± 0.005 | 0.219 ± 0.006 |
| South Orkney Islands | SO-06 | -60.70 | -45.57 | 22 | 2006 | 0.251 ± 0.005 | 0.224 ± 0.008 | 0.258 ± 0.005 | 0.236 ± 0.008 |
| South Georgia | SG-02-03 | -55.24; -53.70 | -35.6; -37.51 | 35 | 2002/3 | 0.247 ± 0.004 | 0.216 ± 0.006 | 0.252 ± 0.004 | 0.224 ± 0.006 |
| South Georgia | SG-05 | -53.70 | -37.51 | 45 | 2005 | 0.244 ± 0.004 | 0.210 ± 0.005 | 0.250 ± 0.004 | 0.220 ± 0.005 |
| Kerguelen Islands Shelf | KI-15 | -47.41 - – 48.67 | 69.7 – 70.98 | 39 | 2016 | 0.238 ± 0.004 | 0.204 ± 0.007 | 0.243 ± 0.004 | 0.213 ± 0.007 |
| Skiff Bank, Kerguelen Islands | SB-15 | -49.8 – -50.01 | 64.8 – 65.64 | 40 | 2016 | 0.236 ± 0.004 | 0.217 ± 0.006 | 0.242 ± 0.004 | 0.228 ± 0.006 |

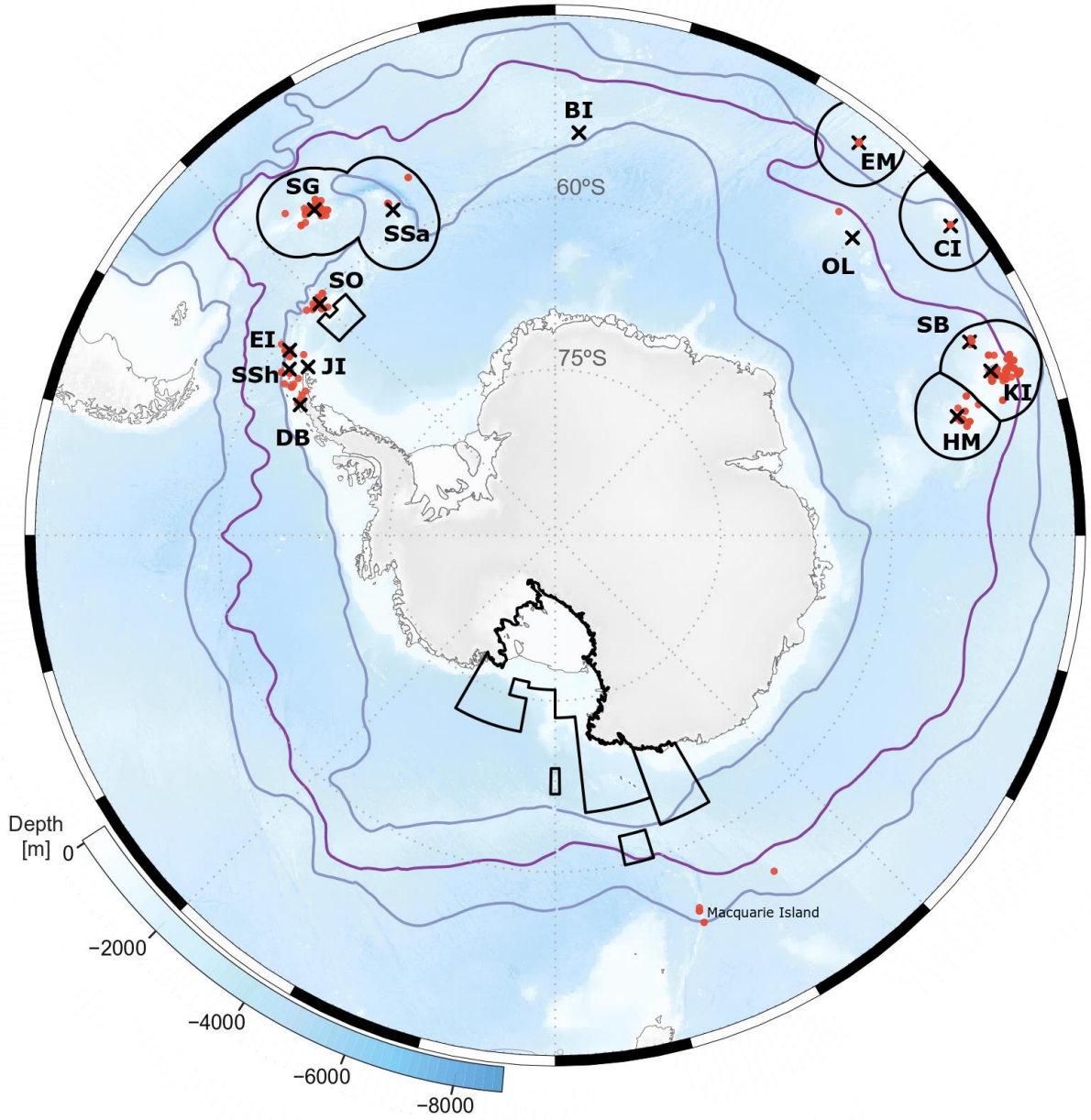



**Fig. 5.1.** Species occurrence of *Notothenia rossii* in the Southern Ocean (red dots) and localities used for individual-based hydrodynamic connectivity modelling (black crosses). Ocean fronts after Orsi et al. (1995) indicated from north to south: sub-Antarctic Front, Polar Front, Southern Boundary of the Antarctic Circumpolar Current. Current marine protected areas (MPAs), i.e. the South Orkney Islands Southern Shelf and Ross Sea MPAs, are shown as black rectangles and exclusive economic zones (EEZs) as black circles. Background shading (white-blue) reflects ocean depth. Modelling sites from west to east: Dallman Bay (DB), South Shetland Islands (SSh), Joinville Island (JI), Elephant Island (EI), South Orkney Islands (SO), South Georgia (SG), South Sandwich Islands (SSa), Bouvet Island (BI), Edward and Marion Islands (EM), Ob and Lena Banks (OL), Crozet Island (CI), Skiff Bank (SB), Kerguelen Islands (KI), Heard and McDonald Islands (HM). Samples for genetics were available from SSh, EI, SO, SG, SB, and KI (see Table 5.1).

### 2.2 Species distribution modelling

The assembled occurrence data (Fig. 5.1) and a subset of environmental variables describing the habitat of benthic Antarctic organisms at 0.1° resolution (Guillaumot *et al.* 2018b) were used for predictive species distribution modelling (SDM) with boosted regression trees (BRT, Elith *et al.* 2008). Twelve environmental variables were included after selecting the most informative variables based on biological knowledge, and the Variation Inflation Factor (VIF) stepwise procedure was used to prune highly correlated descriptors (Naimi *et al.* 2014). Retained variables were: depth, geomorphology, sediments, slope, seafloor current speed, maximum ice cover, maximum ice thickness, mixed layer depth, minimum and maximum particulate organic carbon (POC), and minimum seafloor salinity and temperature. POC, temperature and salinity are an average of interpolated values of the time period 2005-2012. Models were run using only unique occurrences per 0.1° grid-cell and environmental layers were masked to areas shallower than 1000 m depth, which encompasses the observed depth range of adult *N. rossii* (the upper 550 m of the water column (DeWitt *et al.* 1990)). The models were spatially limited to -100 and 100° longitude and -45 and -70° latitude. This longitudinal selection excludes most of the Pacific, and parts of the Indian Southern Ocean sectors. *Notothenia rossii* is documented to occur around Macquarie Island (Fig. 5.1). However, no samples from this comparatively isolated site (only known occurrence in the Pacific sector) were available. Hence, modelling focused on the area from the western Antarctic Peninsula to the Kerguelen Plateau. The latitudinal selection covers all assembled occurrences and therefore the area in which the species can likely biologically occur. Optimal BRT parameters, which in combination reduce modelling error while avoiding overfitting to the occurrence data set, were calibrated following Elith *et al.* (2008). The selected combination used for the final SDMs were: tree complexity 4, bag fraction 0.8 and learning rate 0.02. The number of background data used to characterize the environmental conditions was set at 500 and a four-fold 'CLOCK' method was applied to spatially segregate the proportion of occurrence records used to train the model (75%) and test the model (25%) (Guillaumot *et al.* 2019). In addition, a kernel density estimation layer (Phillips *et al.* 2009) and a multivariate environmental similarity surface index (Elith *et al.* 2010) were estimated and applied as described in detail in Guillaumot *et al.* (2019). These corrections were applied to correct for the influence of autocorrelation within occurrence records and model extrapolation, respectively. A final total of 240 replicate SDMs were run and the mean probabilities of occurrence were used for plotting. Model performance was assessed by measuring the Area Under the receiver operating Curve (AUC, Fielding & Bell 1997), and assessing the number of presence test data correctly classified as suitable areas by the model predictions. Analyses were conducted in R

using the packages 'ncdf4' v1.16.1 (Pierce 2019), 'raster' v3.0-2 (Hijmans *et al.* 2019), 'usdm' v1.1-18 (Naimi & Araújo 2016), 'dismo' v1.1-4 (Hijmans *et al.* 2017), 'MASS' v7.3-51.4 (Venables & Ripley 2002), and 'gbm' v2.15 (Greenwell *et al.* 2019). Input data and R scripts are available at https://doi.org/10.5281/zenodo.3552609.

### 2.3 Reduced representation sequencing

Large numbers of single nucleotide polymorphism (SNP) loci were sourced with reduced representation sequencing, a methodology that reproducibly samples the full genome (Andrews *et al.* 2016). Genomic DNA was extracted using a standard salting out protocol to avoid shearing during column purification. DNA concentration was determined using the Quant-iT PicoGreen dsDNA kit (Thermo Fisher Scientific Inc.) and an Infinite M200 microplate reader (Tecan Group Ltd.) according to the manufacturer's instructions. Extractions were then standardized, checked for signs of degradation on agarose gels and quantified and standardized again to approximately 10 ng $\mu L^{-1}$. Four reduced representation sequencing libraries containing 96 individuals each were constructed (Table 5.1). Thirty of these were within- and between-library controls (i.e. DNA replicates from identical individuals). A modified GBS library preparation protocol based on Elshire et al. (2011) but with size selection was used, as described in full detail in Christiansen et al. (Chapter 3). The restriction enzyme *ApeKI* and a size selection window of 240-340 bp, using a Pippin Prep (Sage Science), were applied to achieve high marker density. The libraries were paired-end sequenced on four lanes of a HiSeq 2500 with v4 chemistry (Illumina Inc.) at the KU Leuven Center for Human Genetics (GenomicsCore).

### 2.4 Variant calling and filtering

Sequencing data was checked for general quality using FastQC v0.11.5 (Andrews 2010). The Stacks pipeline v2.4 (Rochette *et al.* 2019) was used to genotype samples both *de novo* and using the reference genome of *Notothenia coriiceps* (Shin *et al.* 2014). First, each library was demultiplexed and quality filtered using process_radtags (options: -c -q -r). Library 2 and 4 were demultiplexed with quality control disabled (without -q), because low Phred scores at the cutsite otherwise lead to discarding of all reverse reads. For the *de novo* approach the sequences were trimmed to 119 bp, to fulfill Stacks' requirement of uniform read length. In addition, forward, reverse and remainder reads were concatenated after demultiplexing to treat both reads as individual loci, because GBS reads cannot be oriented preventing Stacks from building paired-

end contigs (Rochette *et al.* 2019). A parameter test series using a subset of 24 individuals (four from each locality) was conducted as described in Rochette & Catchen (2017). The retained parameters were m=3, M=4, n=4 (see Supplementary Material S5.1) and subsequently applied to the entire data set. A catalog was built using 10 individuals per locality (cstacks module); all individuals were matched against the catalog (sstacks) and data was transposed (tsv2bam). For the reference approach, forward and reverse reads were aligned to the *N. coriiceps* genome after demultiplexing using BWA v0.7.17 and SAMtools v1.9 and default parameters (Li & Durbin 2009; Li *et al.* 2009). In both cases (reference-based and *de novo*) genotyping of SNPs was conducted using gstacks with default parameters, that is under a Bayesian low coverage framework from Maruki & Lynch (2017).

Stringent quality control and filtering is necessary before downstream processing of GBS data, because high-throughput sequencing data has comparably high error rates (Shafer *et al.* 2017; O'Leary *et al.* 2018). In a first filtering step, genotypes of the reference-based and *de novo* data set were pruned using the populations module of Stacks, requiring loci to be present in at least 80 % of the individuals of each population, to have a minor allele frequency > 0.05 and heterozygosity < 0.7 (Rochette & Catchen, 2017). Subsequently, genepop files were imported and filtered extensively in R (R Core Team, 2019) using the 'radiator' package v1.1.1 (Gosselin, 2019). This filtering approach was conducted without and with technical replicates to assess the genotyping error rate before and after filtering. In brief, data was filtered on missing values, heterozygosity, minor allele count, coverage, SNP position, linkage disequilibrium (LD) and departures from Hardy-Weinberg proportions (Supplementary Material S5.2).

### 2.5 Population genomics

Analyses of the filtered genomic data sets were conducted mostly in R, with code and input data available under https://doi.org/10.5281/zenodo.3552609. Overall and pairwise differentiation measures ($F_{ST}$, $G_{ST}$, D), expected and observed heterozygosity and hierarchical analyses of molecular variance (AMOVA) were calculated using 'adegenet' v2.1.1 (Jombart 2008; Jombart & Ahmed 2011), 'hierfstat' v0.04-30 (Goudet & Jombart 2015), 'mmod' v1.3.3 (Winter 2012), and 'pegas' v0.11-12 R packages (Paradis 2010). Additional data filtering was conducted using 'poppr' v2.8.3 (Kamvar *et al.* 2014) (see Supplementary Material S5.2). Principal component analysis (PCA), non-metric multidimensional scaling and discriminant analyses of principal components (DAPC) were conducted using 'adegenet', 'vegan' v2.5-6 (Oksanen *et al.* 2018), 'MASS' v7.3-51.4

(Venables & Ripley 2002), and 'factoextra' v1.0.5 (Kassambara & Mundt 2017) and 'ggsci' v2.9 (Xiao 2017) aiding plotting. Following cross-validation to avoid overfitting, 30 (*de novo* data) and 50 (reference data) principal components were used for the DAPC. Migration was estimated and visualized using the divMigrate function (Sundqvist *et al.* 2016) from the 'diveRsity' package v1.9.90 (Keenan *et al.* 2013) and the 'qgraph' package v1.6.3 (Epskamp *et al.* 2012).

In addition, the Bayesian clustering software Structure v2.3.4 (Pritchard *et al.* 2000) was called from within R using 'ParallelStructure' v1.0 (Besnier & Glover 2013) and a function from Clark (2017) to prepare input files. Structure was run on both data sets with K ranging from one to ten, with five replicates of each run and always using 10,000 repetitions as burn-in and 100,000 subsequent iterations. The likely number of genetic clusters was inferred using ΔK (Evanno *et al.* 2005) and Structure Harvester (Earl & VonHoldt 2012). The contemporary effective population size ($N_e$) of each sampling location and year was estimated using the LD method (Waples & Do 2008) under a random mating model using a MAF cutoff of 0.05 with bias correction (Waples 2006) and updates for dealing with missing data (Peel *et al.* 2013), as implemented in 'NeEstimator' v2.1 (Do *et al.* 2014). Finally, a simple genome scan for signs of selection was conducted using R packages 'pcadapt' v4.1.0 (Luu *et al.* 2016) and 'qvalue' v2.16.0 (Storey *et al.* 2019). Loci with $q < 0.05$ were retained as candidate loci and the contigs that contained these SNPs were matched against the nucleotide (nt) collection of the NCBI database using BLASTN 2.10.0+ (Altschul *et al.* 1997). Only top hits with an *E*-value $\leq 1 \times 10^{-6}$ and at least 70 % similarity were retained (Benestan *et al.* 2017) and then further investigated using BLASTX 2.9.0+ and the UniProtKB vertebrate database for their putative function (Apweiler *et al.* 2004).

### 2.6 Individual-based connectivity modelling

Five-day mean flow fields for the Southern Ocean region from a state-of-the-art oceanographic modelling framework, Nucleus for European Modelling of the Ocean (NEMO), underpinned the numerical modelling simulations. Simulated flows for the period 1996–2001 were provided by the National Oceanography Centre, Southampton (UK), from a global application of NEMO with an eddy-permitting nominal horizontal resolution of 1/4°, and a partial step z-coordinate with 64 levels in the vertical. Full details of the ocean model are available at http://www.nemo-ocean.eu/About-NEMO. NEMO has been widely used over a range of spatial scales and resolutions and has been shown to provide a good representation of the dominant oceanography of the southern Atlantic Ocean region (Renner *et al.* 2009, 2012).

Mean flows from the circulation model were used to advect Lagrangian particles representing the early life stages of *Notothenia rossii*. The Lagrangian model has been applied previously to the simulation of the dispersal of the eggs and larvae of *N. rossii* around South Georgia (Young *et al.* 2012) and in the Scotia Sea region (Young *et al.* 2015, 2018). In summary, particles were advected at each model time step (5 min) according to the imposed three-dimensional velocity field, using a second-order Runge-Kutta method. Additional horizontal and vertical diffusions were included using a random-walk approach (Dyke 2001), to account for unresolved turbulent motion in the ocean model flow fields. Particles representing the early life stages of *N. rossii* were released at known spawning populations (Fig. 5.1) (DeWitt *et al.* 1990; Barrera-Oro & Casaux 1992; Duhamel *et al.* 1995; Kock *et al.* 2004). Appropriate spawning areas at each location were identified by a comparison of local model depth with the known spawning depth range, 200-360 m (Kock *et al.* 2004). Modelled particles were released randomly within appropriate grid cells, with one thousand particles released per day at each site for the duration of the observed spawning periods. Dispersal of eggs was simulated for four months (Atlantic Ocean) or three months (Indian Ocean), with subsequent larval dispersal simulated for three months. There are no data to suggest that *N. rossii* larvae perform diel vertical migration and like all notothenioids the species does not possess a swimbladder; thus model eggs and larvae were allowed to move randomly within observed depth ranges: upper 100 m for eggs and upper 50 m for larvae (A. W. North, personal communication).

The potential for successful dispersal between isolated populations was assessed by comparing the position of each model larva with recruitment boxes encompassing each known population location (Fig. 5.1) over a 4-week period centered on the end of the planktonic phase. If a larva was within a recruitment box at any point during the 4-week period, it was considered to potentially recruit successfully to a nursery ground at this site. The percentage of larvae from each spawning site successfully reaching each recruitment box was calculated, and the results were combined into a single connectivity matrix describing the proportion of individuals arriving in a destination population (rows) from a given source population (columns). Such matrices describe potential connectivity; they do not include mortality or inter-annual variability in biological processes such as spawning and development rates. The effect of inter-annual variability in the underlying flow fields on predicted dispersal and connectivity was assessed by repeating the simulations for a five-year period (1996-2000). The results were combined to give a single mean connectivity matrix, and a matrix showing the number of years in which non-zero connectivity occurred (i.e. persistence).

## 3. Results

### 3.1 Species distribution probability

Species distribution modelling results showed high modelling relevance with an AUC score of 0.975 and > 90 % of test data correctly classified (Table 5.2). Generally, model predictions are confined to areas of the Southern Ocean that are depth-wise potentially relevant for *Notothenia rossii*; large deep-sea areas are not included in the model here. The largest contribution to the modelling results was from mixed layer depth (51.6 ± 7.4%), with moderate contribution from maximum ice thickness (15.4 ± 3.7%), geomorphology (6.9 ± 4.1%), and depth (5.3 ± 2.5%), and little contribution from the remaining variables (< 5 % each). Areas with high occurrence probability overlap well with documented occurrence records. South Georgia and the Kerguelen Plateau are the largest areas of high occurrence in the part of the Southern Ocean that was evaluated (Fig. 5.2). In addition, suitable habitat for *N. rossii* is predicted in areas that are not documented in the occurrence data set: around Patagonia (low probability), around Bouvet Island and some seamounts north of that (high to medium probability), and west and east of Prydz Bay off the Antarctic continent (medium to low probability). Interestingly, the occurrence probability at the Ob and Lena banks is comparatively high. In contrast, predicted occurrence is low around the Edward and Marion Islands and Crozet Island.

**Table 5.2.** Model statistics describing the outcome of species distribution modelling to predict occurrence probability of *Notothenia rossii* in the Atlantic and Indian sectors of the Southern Ocean (mean ± standard deviation). AUC: Area Under the Receiver Operating Curve; COR: Point Biserial Correlation; TSS: True Skill Statistic; maxSSS: maximum sensitivity plus specificity threshold; Correctly classified test data (%): percentage of presence-test and background-test records falling on predicted suitable areas (prediction > maxSSS).

| Model statistic | Mean and standard deviation |
|---|---|
| AUC | 0.975 ± 0.016 |
| COR | 0.831 ± 0.069 |
| TSS | 0.768 ± 0.117 |
| maxSSS | 0.469 ± 0.230 |
| Correctly classified test data (%) | 92.3 ± 3.0 % |

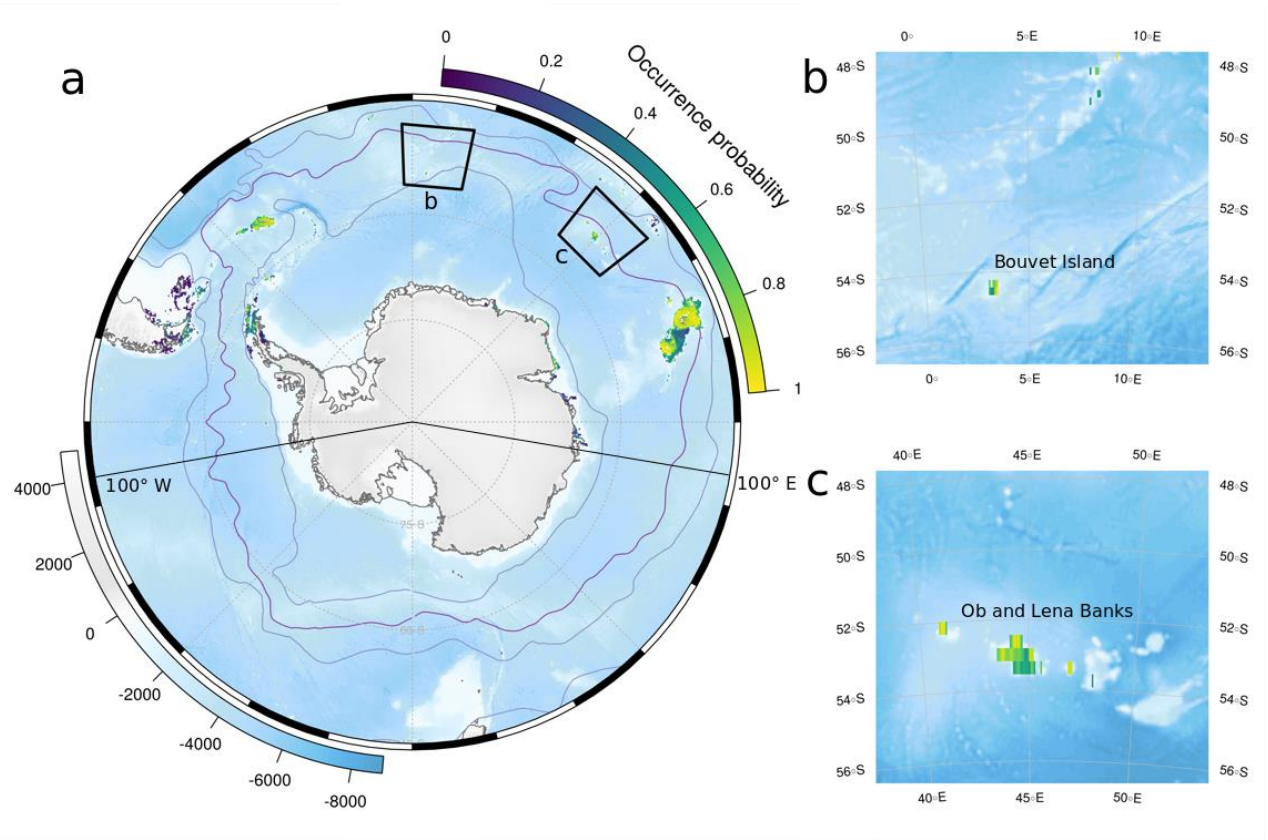



**Fig. 5.2.** Predicted species occurrence probability for *Notothenia rossii* in the Atlantic and Indian sector of the Southern Ocean (prediction only between -100 and 100° longitude) based on mean prediction values from 240 model replicates using boosted regression trees (a). Insets show predictions around Bouvet Island (b) and the Ob and Lena Banks (c). Ocean fronts after Orsi et al. (1995) indicated from north to south: sub-Antarctic Front, Polar Front, Southern Boundary of the Antarctic Circumpolar Current. Background shading (white-blue) reflects ocean depth. Only predictions in areas where the model does not extrapolate are shown.

### 3.1 Sequencing data

No sequencing problems were indicated by FastQC reports. On average each individual received 4.86 ± 2.29 million (M) reads. Four low coverage (< 1 M reads), as well as four high coverage (> 13 M reads) individuals were excluded before bioinformatics. After genotyping using Stacks, global coverage (as the average number of reads per locus per individuals) was at 10.64 ± 11.84. This coverage is at the minimum required for calling heterozygotes reliably (Rochette & Catchen 2017). Therefore, extensive downstream filtering of SNP data sets was conducted with relatively high thresholds (Supplementary Material S5.2). After pre-filtering using the populations module of Stacks, 73,554 and 85,980 SNPs were present in the reference-based and *de novo* based data sets, respectively. Global genotyping error rates of these data sets (calculated from technical replicates) were between 1.48 % and 5.30 % in the reference-based data and 0.61 % and 8.80 % in the *de novo* data. Such genotyping error rates are not ideal, but also not uncommon, and likely related to the comparably low coverage (Mastretta-Yanes *et al.* 2014; Fountain *et al.* 2016). We circumvented negative impacts as far as possible by applying very extensive downstream filtering in R (Supplementary Material S5.2). In brief, individual genotypes with high amounts of missing data, abnormal heterozygosity patterns or signs of duplicate genomes were filtered. Loci were filtered based on missing data, minor allele count, minor allele frequency, coverage, linkage, SNP position, and departure form Hardy-Weinberg proportions. Even after these filtering steps, a bias related to sequencing library remained detectable in the data, as evidenced by PCA and AMOVA. Loci that contributed to this bias were also excluded. All these steps are described in detail in Supplementary Material S5.2. Eventually, 272 individuals and 3,503 SNPs in the reference data set and 261 individuals and 7,501 SNPs in the *de novo* data set remained and were used for all subsequent analyses.

### 3.2 Genomic variability

While the number of SNPs is more than twice as high in the final filtered *de novo* data set compared to the reference-based data set, patterns of genetic diversity are largely congruent between the data sets. Observed heterozygosity was minimally lower than expected heterozygosity in both data sets (Table 5.1, Fig. 5.3a) and on average heterozygosity was slightly higher in the *de novo* data compared to the reference data (Table 5.1). Pairwise population estimates of differentiation were generally low with $F_{ST}$ values (Weir & Cockerham 1984) reaching 0.0027 and 0.0018 in the *de novo* and reference-based data sets, respectively (Table 5.3 and 5.4). Few pairwise comparisons were significantly different from zero as indicated from

confidence intervals (14 in the *de novo* data, six in the reference data). Similar results were obtained using alternative differentiation metrics (Supplementary Material S5.3). The sample from the South Orkney Islands seems slightly differentiated with five (*de novo*) or three (reference) significant pairwise estimates of $F_{ST}$. This pattern is also discernible in the NMDS plots based on Hedrick's $G_{ST}$ (Fig. 5.3c). Individual-based clustering approaches including PCA, DAPC (functions find.clusters and snapclust.choose.k) failed to identify meaningful genetic clusters (Fig. 5.3b and 5.4a). Results from STRUCTURE indicated the number of clusters based on maximum ΔK as six (*de novo* data) or five (reference data), but the log likelihood for K did not increase significantly suggesting an absence of genetic structure. Using DAPC with sampling location as priors offsets the population centroids of the South Orkneys, South Georgia and Skiff Bank slightly from the remaining locations (Fig. 5.4a). Genomic data suggests that the effective population size ($N_e$) of *N. rossii* is relatively large at most localities, with values ranging from 1,449 to 42,299, but often with upper limits of the 95 % confidence intervals as infinite (Table 5.5). The migration analysis revealed very high levels and no asymmetric pattern of gene flow, corresponding to the observed absence of distinct genetic structure (Fig. 5.4b). Genome scans for signs of selection identified 12 (*de novo*) and 37 (reference) candidate loci for further investigation (Fig 5.4c). Matching the flanking regions of these loci against the nucleotide database showed that some of them are highly similar to genomic DNA from other perciform fishes, such as *Chionodraco hamatus*, *Gymnodraco acuticeps*, *Notothenia coriiceps*, *Cottoperca gobio*, and *Dissostichus mawsoni* (Supplementary Material S5.4). The DNA sequence of *D. mawsoni* is part of an antifreeze glycoprotein locus (Nicodemus-Johnson *et al.* 2011). Several contigs matched to predicted mRNA sequences of *N. coriiceps*, e.g. with putative functions in metal ion binding (Supplementary Material S5.4). The PCA, DAPC, migration and outlier plots are only shown for the *de novo* data set (Fig. 5.3 and 5.4); the reference-based data yielded similar results (Supplementary Material S5.5).

**Table 5.3.** Pairwise genetic differentiation of *Notothenia rossii* per sampling locality (see Table 5.1 for codes) based on 7,501 SNP genotypes derived from mapping against a *de novo* assembly. $F_{ST}$ following Weir & Cockerham (1984) (also referred to as $G_{ST}$) below the diagonal (negative values set to zero) and confidence intervals after 1000 bootstraps above the diagonal. $F_{ST}$ values where confidence intervals do not span zero are marked in bold.

| | **SSD-06** | **SSK-06** | **SSK-15-16** | **EI-06-07** | **EI-02** | **SO-06** | **SG-02-03** | **SG-05** | **SB-15** | **KI-15** |
|---|---|---|---|---|---|---|---|---|---|---|
| **SSD-06** | | -0.0002 – 0.0003 | -0.0001 – 0.0003 | -0.0002 – 0.0002 | -0.0003 – 0.0001 | 0.0001 – 0.0006 | 0.0001 – 0.0005 | -0.0001 – 0.0003 | 0.0000 – 0.0004 | -0.0004 – 0.0001 |
| **SSK-06** | 0.0000 | | 0.0000 – 0.0004 | -0.0002 – 0.0002 | -0.0002 – 0.0002 | 0.0001 – 0.0006 | 0.0001 – 0.0005 | -0.0003 – 0.0000 | -0.0001 – 0.0002 | -0.0002 – 0.0003 |
| **SSK-15-16** | 0.0005 | **0.0004** | | -0.0005 – -0.0001 | -0.0001 – 0.0002 | -0.0002 – 0.0003 | -0.0003 – 0.0001 | -0.0002 – 0.0001 | -0.0001 – 0.0003 | -0.0002 – 0.0001 |
| **EI-06-07** | 0.0003 | 0.0007 | 0.0000 | | -0.0003 – 0.0001 | 0.0000 – 0.0004 | -0.0001 – 0.0003 | -0.0002 – 0.0002 | -0.0003 – 0.0001 | -0.0002 – 0.0002 |
| **EI-02** | 0.0000 | 0.0000 | 0.0000 | 0.0000 | | -0.0002 – 0.0002 | 0.0000 – 0.0004 | -0.0002 – 0.0001 | -0.0003 – 0.0001 | 0.0000 – 0.0004 |
| **SO-06** | **0.0018** | **0.0011** | 0.0002 | **0.0012** | 0.0003 | | 0.0001 – 0.0005 | -0.0004 – 0.0000 | 0.0002– 0.0006 | -0.0001 – 0.0004 |
| **SG-02-03** | **0.0005** | **0.0006** | 0.0000 | 0.0000 | **0.0000** | **0.0006** | | -0.0001 – 0.0002 | 0.0000 – 0.0004 | 0.0000 – 0.0004 |
| **SG-05** | 0.0005 | 0.0000 | 0.0002 | 0.0007 | 0.0001 | 0.0000 | 0.0000 | | -0.0003 – 0.0000 | 0.0000 – 0.0003 |
| **SB-15** | **0.0006** | 0.0002 | 0.0004 | 0.0003 | 0.0000 | **0.0007** | **0.0008** | 0.0000 | | -0.0003 – 0.0001 |
| **KI-15** | 0.0000 | 0.0003 | 0.0000 | 0.0002 | **0.0003** | 0.0004 | **0.0005** | **0.0001** | 0.0000 | |

**Table 5.4.** Pairwise genetic differentiation of *Notothenia rossii* per sampling locality (see Table 5.1 for codes) based on 3,503 SNP genotypes derived from mapping against the reference genome of *N. coriiceps* (Shin *et al.* 2014). $F_{ST}$ following Weir and Cockerham (1984) (also referred to as $G_{ST}$) below the diagonal (negative values set to zero) and confidence intervals after 1000 bootstraps above the diagonal. $F_{ST}$ values where confidence intervals do not span zero are marked in bold.

| | **SSD-06** | **SSK-06** | **SSK-15-16** | **EI-06-07** | **EI-02** | **SO-06** | **SG-02-03** | **SG-05** | **SB-15** | **KI-15** |
|---|---|---|---|---|---|---|---|---|---|---|
| **SSD-06** | | -0.0002 – 0.0003 | -0.0005 – 0.0001 | -0.0007 – 0.0000 | -0.0004 – 0.0001 | -0.0003 – 0.0004 | -0.0002 – 0.0004 | -0.0003 – 0.0002 | -0.0004 – 0.0002 | -0.0007 – -0.0001 |
| **SSK-06** | 0.0006 | | 0.0001 – 0.0007 | -0.0002 – 0.0003 | -0.0001 – 0.0004 | 0.0000 – 0.0006 | 0.0002 – 0.0008 | -0.0001 – 0.0004 | -0.0004 – 0.0001 | -0.0001 – 0.0005 |
| **SSK-15-16** | 0.0000 | **0.0017** | | 0.0000 – 0.0006 | -0.0002 – 0.0002 | -0.0005 – 0.0001 | -0.0005 – 0.0000 | -0.0002 – 0.0002 | -0.0001 – 0.0004 | -0.0001 – 0.0005 |
| **EI-06-07** | 0.0001 | 0.0013 | **0.0011** | | -0.0003 – 0.0003 | 0.0001 – 0.0007 | -0.0003 – 0.0003 | -0.0001 – 0.0005 | -0.0003– 0.0002 | -0.0005 – 0.0004 |
| **EI-02** | 0.0000 | 0.0007 | 0.0002 | 0.0002 | | -0.0002 – 0.0003 | -0.0004 – 0.0001 | -0.0003– 0.0001 | -0.0003– 0.0002 | -0.0001 – 0.0004 |
| **SO-06** | 0.0008 | **0.0027** | 0.0000 | **0.0015** | 0.0012 | | -0.0002 – 0.0005 | -0.0001 – 0.0005 | 0.0001– 0.0007 | -0.0003 – 0.0003 |
| **SG-02-03** | 0.0005 | **0.0010** | 0.0000 | 0.0005 | 0.0000 | 0.0009 | | -0.0002 – 0.0003 | -0.0002 – 0.0003 | -0.0002 – 0.0004 |
| **SG-05** | 0.0000 | 0.0003 | 0.0002 | 0.0004 | 0.0000 | 0.0007 | 0.0000 | | -0.0001 – 0.0004 | -0.0002 – 0.0003 |
| **SB-15** | 0.0000 | 0.0001 | 0.0010 | 0.0000 | 0.0003 | **0.0016** | 0.0003 | 0.0002 | | -0.0001 – 0.0004 |
| **KI-15** | 0.0000 | 0.0005 | 0.0004 | 0.0000 | 0.0007 | 0.0005 | 0.0004 | 0.0001 | 0.0000 | |

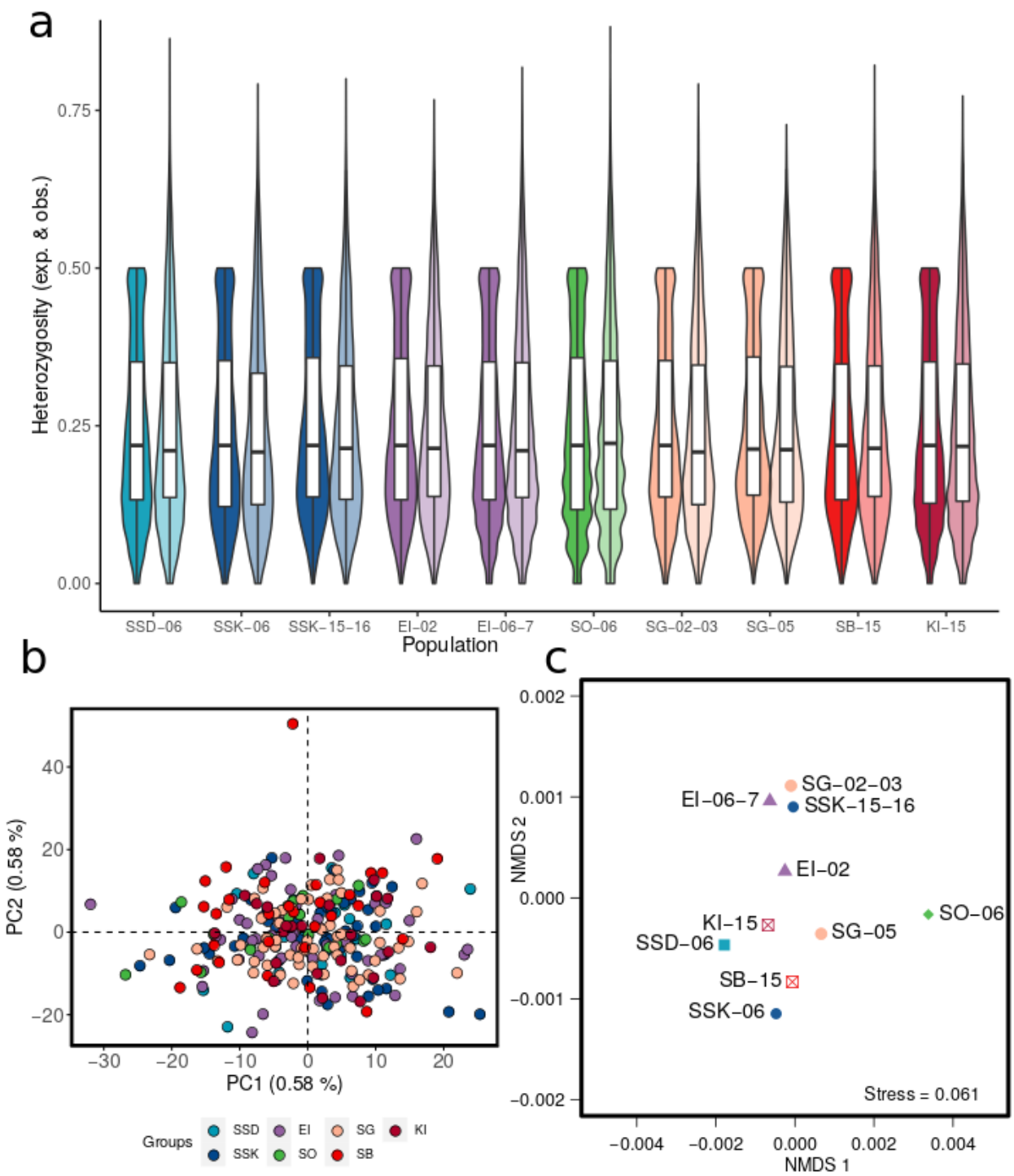



**Fig. 5.3** Genomic diversity of *Notothenia rossii* in the Southern Ocean based on 7,501 SNP loci. Expected (darker shading, left) and observed (light shading, right) heterozygosity is shown as box and violin plots for each genetically screened population (a). Principal component analysis (PCA) reveals little individual-based differentiation (b), while non-metric multidimensional scaling based on $G_{ST}$ distances shows subtle differences between population samples (c). Sample codes as in Table 5.1; samples from different years but same locality are not shown separately on the PCA.

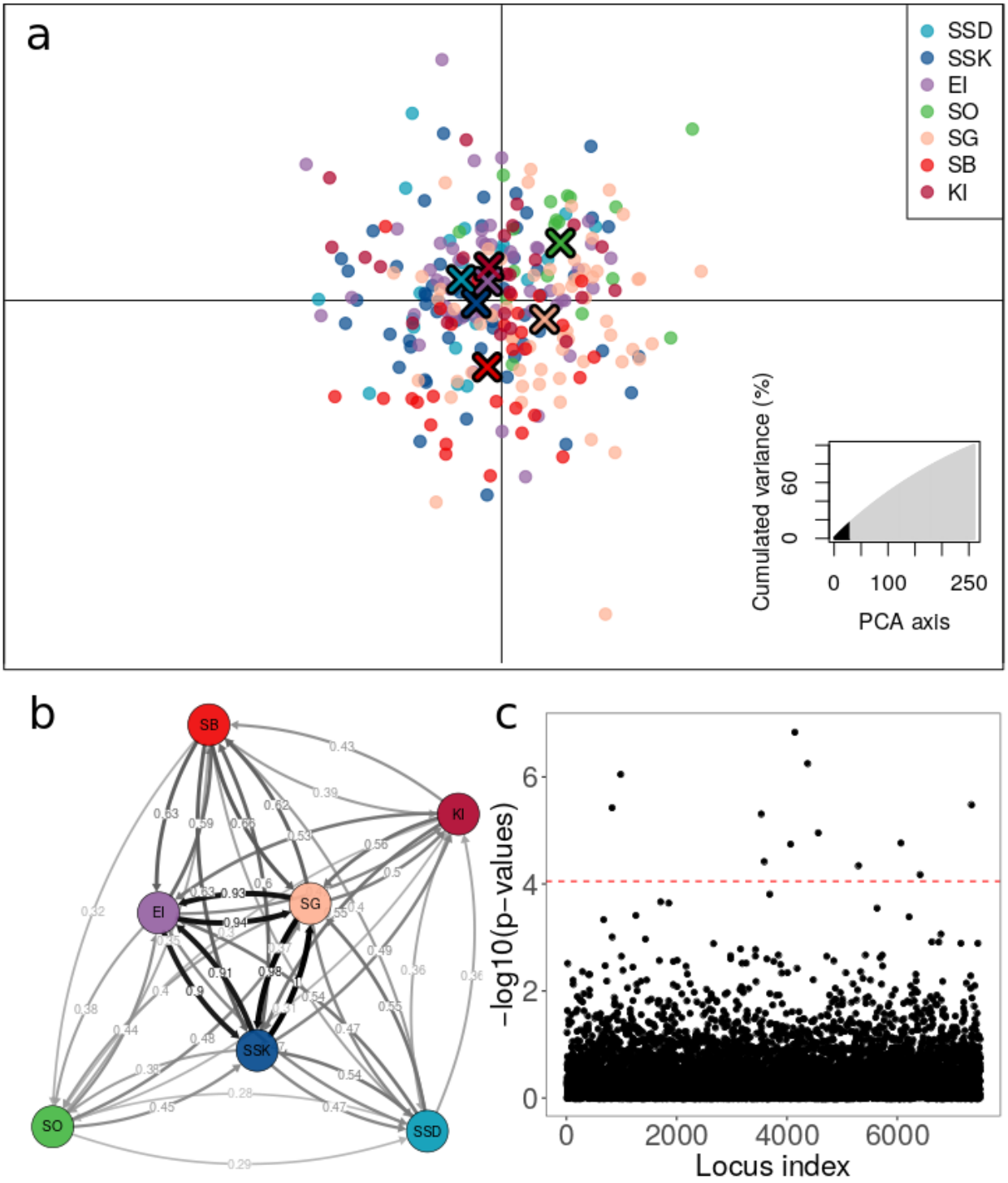



**Fig. 5.4** Genomic differentiation of *Notothenia rossii* in the Southern Ocean based on 7,501 SNP loci. Geographic clustering as attempted through discriminant analysis of principal components is shown along the first two principal components (a). Relative migration as estimated from Nei's $G_{ST}$ reveals overall high and no asymmetric gene flow (b). Genome scans for loci putatively under influence of selection detected 12 outliers at q > 0.05 (c). Sample codes as in Table 5.1; samples from different years but same locality are combined.

**Table 5.5.** Effective population size ($N_e$) of *Notothenia rossii* from various locations in the Southern Ocean. Estimates were calculated using the linkage disequilibrium method for filtered genotypes from *de novo* and reference-based bioinformatics; with 95 % confidence intervals (CI) calculated based on the jackknife method of Waples & Do (2008).

| Sample | $N_e$ *de novo* | CI *de novo* | $N_e$ reference | CI reference |
|---|---|---|---|---|
| SSD | 2,089 | 197 – Infinite | 2,677 | 319 – Infinite |
| SSK | 6,207 | 1,312 - Infinite | 6,846 | 1,302 - Infinite |
| EI | 4,777 | 1,036 - Infinite | Infinite | 1,422 - Infinite |
| SO | 6,629 | 182 - Infinite | 42,299 | 228 - Infinite |
| SG | 4,837 | 1,327 - Infinite | 21,601 | 1,864 – Infinite |
| SB | 2,327 | 385 - Infinite | 1,957 | 390 – Infinite |
| KI | 1,665 | 256 - Infinite | 1,449 | 206 - Infinite |

### 3.3 Modelled connectivity

The predicted mean connectivity matrix suggests wide dispersal of *Notothenia rossii* within the Scotia Sea, with high and persistent levels of connectivity between populations around the Antarctic Peninsula (AP) and South Georgia, and lower but persistent connectivity with the South Orkney and South Sandwich Islands (Fig 5.5). There is low but persistent connectivity from the South Sandwich Islands to Bouvet Island, and from Bouvet to populations in the Indian Ocean, in particular Edward and Marion Islands and Crozet Island. Although persistent connectivity pathways are predicted in the Indian Ocean, for example from Ob and Lena Banks to Crozet Island, Skiff Bank and Kerguelen Islands, the magnitude of connectivity is generally weaker than in the Scotia Sea, with the exception of the Kerguelen Plateau region. The patterns of connectivity suggest highly asymmetric dispersal, with a greater occurrence of non-zero values below the diagonal of the connectivity matrix, indicating unidirectional transport to the northeast across the Scotia Sea, and eastward towards and within the Indian Ocean in accordance with the dominant flows of the Antarctic Circumpolar Current. Bidirectional transport is predicted between proximate sites with complex local oceanography, in particular around the Antarctic Peninsula and the Kerguelen Plateau region. The pattern of connectivity suggests that gene flow includes an element of stepping-stone transport. *Notothenia rossii* is widely dispersed within the Scotia Sea, but there is no direct connectivity between sites in the Scotia Sea and those in the Indian Ocean. Gene flow over this larger geographic scale is achieved through stepping-stone transport via Bouvet Island.

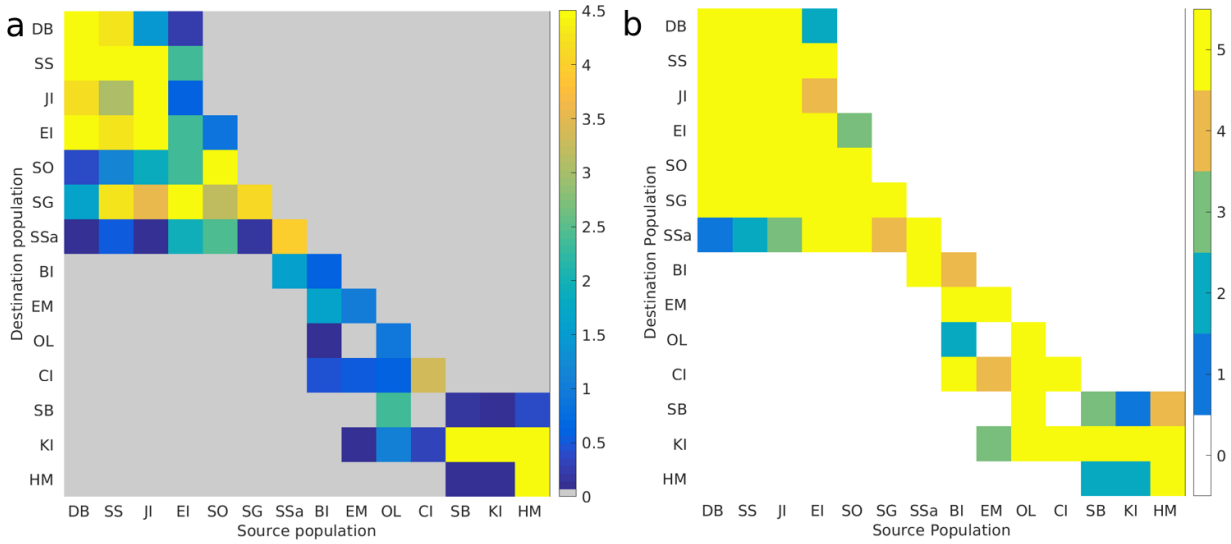



**Fig. 5.5.** Simulated dispersal connectivity of *Notothenia rossii* throughout most of the Southern Ocean: (a) mean connectivity as a percentage of particles from source populations (columns) successfully reaching destination populations (rows), on a transformed log scale [log((10x)+1)], (b) frequency of non-zero connectivity.

## 4. Discussion

### 4.1 Distribution, genomic diversity and connectivity of *Notothenia rossii*

Contrary to our working hypothesis, even thousands of genomic markers screened across many individuals from locations > 5,000 km apart reveal no evidence of genetic differentiation. The historically overexploited Antarctic fish *Notothenia rossii* therefore exhibits no contemporary spatial genetic population structure, although connectivity modelling suggests that exchange between the Scotia Sea and the Kerguelen Plateau regions occurs via stepping-stone transport. This apparent discrepancy may be due to the different temporal scales resolved by the observational and modelling analyses, i.e. ecological vs. evolutionary time. As we will further detail, our results show 1) the detailed distribution of the species in the Atlantic and Indian Ocean sectors of the Southern Ocean, 2) the contemporary genome-wide levels of diversity and 3) the direction and magnitude of dispersal connectivity.

As a shelf dwelling fish that prefers shallower waters to feed and reproduce, the occurrence of *N. rossii* is relatively localized in the Southern Ocean. Juveniles develop in algae beds (Duhamel *et al.* 1982; Barrera-Oro & Winter 2008), which are only found close to the coast (Wiencke *et al.* 2014). The cold temperatures of the high-Antarctic are, however, likely detrimental for the species as its blood equilibrium freezing point is comparatively high for a notothenioid (Bilyk & DeVries 2010; Miya *et al.* 2013). Therefore, the habitat requirements alone induce a fragmented

distribution in this species. Species distribution modelling predictions confirm this and highlight many well-documented population hotspots as highly probable habitats (Fig. 5.2). The South Shetland Islands, South Georgia and the Kerguelen Plateau are localities where *N. rossii* is most often caught (DeWitt *et al.* 1990; Duhamel *et al.* 2014). However, other areas, some of which are less well-studied, may also be relevant habitats for *N. rossii*. For example, SDM data show high occurrence probability at Bouvet Island, even though the shelf around this island is narrow and the species was never caught there (DeWitt *et al.* 1990; Jones *et al.* 2008; Padilla *et al.* 2015). However, incidental data (chinstrap penguin stomach content) indicate that *N. rossii* is present at Bouvet Island (Niemandt *et al.* 2015). This would corroborate our SDM prediction, although the record is based on a single otolith, which could also be misidentified. Ultimately, only more extensive sampling, particularly of the near shore fish fauna, will be able to resolve the matter. It also remains to be investigated whether other new occurrence localities predicted by the SDM, that is around Prydz Bay and Patagonia, are realistic. At least the latter seems very unlikely as the ecological niche in Patagonian waters is filled by different species such as *Paranotothenia magellanica* or *Patagonothen* spp. (Cousseau *et al.* 2019). *Notothenia rossii* has not been recorded off Prydz Bay so far either despite regular surveys (Hoddell *et al.* 2000; Van de Putte *et al.* 2010).

The genetic diversity of *N. rossii* in terms of heterozygosity is similar to that observed in other fishes with SNPs (Fig. 5.3a, Christiansen *et al.* Chapter 4, Berg *et al.* 2016; Pérez-Portela *et al.* 2018). In addition, there are no signs of spatial variation in heterozygosity, despite spatially heterogeneous fishing pressure (Duhamel 1982; Kock 1992). These results are in contrast to the expectation that overharvesting reduces genetic diversity (Pinsky & Palumbi 2014). However, objective comparisons between studies are challenging due to the wide variety of settings employed to generate a "final" SNP data set (Shafer *et al.* 2017). Fifty years post-exploitation the genetic diversity of *N. rossii* does not seem dramatically reduced across thousands of markers. As there is no baseline for pre-exploitation diversity levels, it remains elusive at this point whether these levels of heterozygosity are representative of the unperturbed state. In fact, even if overall average diversity is high, rare alleles, potentially important for rapid adaptation, may be lost (Pinsky & Palumbi 2014). The genetic data furthermore demonstrate a striking lack of spatial structure with very low $F_{ST}$ values and no genetic clusters discernable (Table 5.3 & 5.4, Fig. 5.3b). Consequently, this implies regular gene flow, at least via stepping stones between all established population patches of *N. rossii* (Fig. 5.4b). This is an important implication to consider in the context of the results of the connectivity modelling exercise. Finally, several

candidate loci show indications of putative recent selective pressure, despite the lack of overall population structure. This is not unexpected, given that selection and adaptation can occur in the presence of high gene flow (Tigano & Friesen 2016; Hoey & Pinsky 2018). Nevertheless, we emphasize that these candidate loci might be false positives or even related to genotyping error. High coverage studies or whole genome sequencing approaches are needed for a more detailed understanding of local or global adaptation in *N. rossii* (Booker *et al.* 2019).

Dispersal of the early life stages of *N. rossii* is generally high, but the large-scale unidirectional connectivity predicted by the modelling suggests that inter-ocean connectivity is achieved through stepping-stone transport (Fig. 5.5). The South Sandwich Islands and Bouvet Island in particular are predicted to be key links between the abundant *N. rossii* aggregations at the Antarctic Peninsula and Scotia Sea, and the Kerguelen Plateau. The relatively large effective population sizes indicated by the genomic data may secure successful large-scale connectivity over evolutionary time scales. Thus, for example, if the exchange of individuals between the Scotia Sea and the Kerguelen Plateau were to fail in some years, successful recruitment in other years may suffice to maintain gene flow.

#### 4.2 A unifying framework to explain contemporary patterns

Here, we suggest that our results can be collectively explained through a scenario that incorporates the life history, physical setting and exploitation history of *Notothenia rossii*. Three main aspects help to unify the patterns observed through SDM, genomics and dispersal modelling. First, important stepping stones such as the South Sandwich Islands and potentially Bouvet Island may also act as temporal refuges for juvenile fish. Even if small, the local benthic ecosystem at Bouvet Island provides suitable conditions including some macroalgae and a variety of invertebrate prey items for *N. rossii* fingerlings (Arntz *et al.* 2006; Jacob *et al.* 2006). The model simulations stopped after seven months, at which point observational data suggest larvae develop into brown -phase fingerlings and recruit to kelp beds (North 2001). However, the recruitment behavior of early juveniles is not well known. For example, blue-phase fingerlings may be able to continue a pelagic life style for an extended period until a suitable recruitment site is reached. Such behavior would increase the dispersal range of the early life stages, and potentially reduce the dependence on small, isolated stepping stones for inter-ocean connectivity. In addition, it is unknown so far, but not inconceivable that, for example, juveniles that reached a stepping stone may later proceed to migrate further toward habitats in the

Kerguelen Plateau where they continue to grow, mature and eventually reproduce. In fact, Shcherbich (1975) for South Georgia and Barrera-Oro *et al.* (2014) for the South Shetland Islands indicated that some juveniles spend at least a full year as blue-phase fingerlings in the water column before settling to a benthic life style. The adults also undertake at least short distance migrations, such as from coastal kelp belts to the outer archipelago (on the Kerguelen Plateau) or to their spawning ground (Duhamel 1982; DeWitt *et al.* 1990). Tagging studies have been used so far to validate age determination (Moreira *et al.* 2013), but could be used in the future to verify whether juveniles or adults of *N. rossii* are capable of more extensive migrations. Such behavior is documented in Antarctic toothfish, that undertake long-distance migrations at least occasionally (Hanchet *et al.* 2010). As the genomic data suggests circumpolar transport, it is possible that other, potentially small stepping-stone population patches exist. An important area for circumpolar connectivity could be the documented occurrence off Macquarie Island in the Pacific sector of the Southern Ocean. The lack of samples and data precluded us from investigating this further. Model simulations, however, suggested the potential for transport from the Kerguelen Plateau to Macquarie Island with an extended dispersal period of a year, although successful dispersal from Macquarie Island to the Antarctic Peninsula was not predicted within this time frame. Therefore, to achieve circumpolar connectivity, the modelling would suggest a longer pelagic phase or the existence of undocumented population patches, or both.

The slow recovery of *N. rossii* following severe overfishing may be the result of historically diminished effective population size in conjunction with the stepping-stone nature of large-scale connectivity. Large stocks around South Georgia and Kerguelen were heavily exploited in the 1970s, likely leading to considerably reduced effective population size at these localities. Possibly, the South Georgia stock would have been resupplied through dispersal from the western Antarctic Peninsula in the years following its overexploitation. However, the spawning stock at the Antarctic Peninsula was also largely removed through fishing in 1979/80 (Kock 1992). In turn, the Kerguelen stock was not supplied sufficiently, because the influx of larvae or fingerlings from South Georgia via stepping stones was interrupted. Thus, the already low levels of long-distance ecological connectivity that we estimated here may be an explanation for the long recovery time. This could mean that the original, unexploited population went through a genetic bottleneck leading to large genetic homogeneity of the extant population. In addition, other species, for example the opportunistically feeding grey notothen *Lepidonotothen squamifrons* at South Georgia (Gregory *et al.* 2014), may have filled vacant ecological niches in

the meantime and further hampered the re-establishment of highly abundant *N. rossii* stocks. Kock *et al.* (2004) rightfully pointed out that the biomass estimation is particularly difficult in *N. rossii* due to its patchy distribution as adults. Yet, the most recent surveys were able to document large catches of *N. rossii* once again, at least in the Kerguelen Plateau (Duhamel *et al.* 2017). This trend is corroborated in the South Shetland Islands (Barrera-Oro *et al.* 2017), suggesting that the species is indeed slowly recovering.

Lastly, a comparison with other Southern Ocean fish indicates that relatively high connectivity may be the most common scenario among sub-Antarctic fish. Matschiner *et al.* (2009) summarized population genetic studies of notothenioids over a period of 15 years and noted that significant differentiation over small scales (< 100 km) was only exceptionally documented in three out of 27 cases. More recently, several studies have uncovered previously unknown genetic differentiation, for example in *Trematomus* spp. (Van de Putte *et al.* 2012b), *Champsocephalus gunnari* and *Chaenocephalus aceratus* (Damerau *et al.* 2014; Young *et al.* 2015), *Pleuragramma antarctica* (Agostini *et al.* 2015; Caccavo *et al.* 2019), and *Notothenia coriiceps* (Christiansen *et al.* Chapter 4). However, except for *C. gunnari* and *C. aceratus* these results relate to long-distance differentiation in high-Antarctic species. On the one hand, habitats for non-deep-sea species near the Antarctic continent may be physically closer to each other than offshore habitats of the sub-Antarctic, where vast deep sea basins cause habitat fragmentation. On the other hand, ice cover, iceberg scouring and the advance and retreat of ice during glacial cycles may have driven high-Antarctic species into local refugia, leading to genetic dissimilarities that are still traceable in the genome (Allcock & Strugnell 2012). In the sub-Antarctic the habitat is less affected by ice, but discontinuous for benthic species through the sheer geographical setting, while it is comparatively barrier-free for pelagic species. In order to persist in this habitat and maintain vast distribution ranges, species may have adapted their dispersal capabilities to achieve persistent (even if low) long-distance connectivity as in the case of *N. rossii*, but also kelp, toothfish and crustaceans, for example (see Moon *et al.* 2017 and references therein).

### 4.3 Implications for MPA design and fisheries management

The fisheries data shows clearly that *Notothenia rossii* experienced a dramatic overexploitation in the 1970s (Kock 1992; CCAMLR 2019a). The recovery is more difficult to assess due to fewer systematic stock assessment methods (trawling, acoustics) compared to less remote oceans. Nevertheless, several recent ecological and fisheries surveys indicate an ongoing recovery

(Marschoff *et al.* 2012; Barrera-Oro *et al.* 2017; Duhamel *et al.* 2017). If a slow recovery process was indeed the result of long generation time, slow growth, reduced effective population size and stepping stone connectivity, as we suggest here, then future management plans should remain very precautionary, which is in accordance with CCAMLR. Importantly, a precautionary approach should not only regard each management area separately, but consider the interconnectivity between these areas. Successful ecological connectivity in at least some years may be an important prerequisite for a stable circum-Antarctic population. In turn, this suggests that it is important to protect areas that act as key stepping stone habitats. The waters around Bouvet Island have been fished in the past (Arntz *et al.* 2006), but are currently a designated marine reserve to 12 nautical miles from the coast. Controversially though, some krill fisheries permits are granted in this area as well. Our results demonstrate that the Bouvet Island marine ecosystem may be a unique stepping stone of large ecological importance. Endemism levels at Bouvet are very low, further supporting the premise that many species are in fact transported here by advection (Arntz *et al.* 2006; Gutt *et al.* 2006; Moreau *et al.* 2017). Therefore, not only *N. rossii* but also other invertebrate and vertebrate species, particularly those present in both the Scotia Sea and Kerguelen Plateau regions, may rely on this comparatively tiny ecological hotspot, which should receive adequate protection. Long-distance connectivity has clear benefits for the effectiveness of MPA networks, although it has previously rarely been quantified (Manel *et al.* 2019).

A second important conservation implication concerns the Antarctic Peninsula. A recent MPA proposal that was presented to CCAMLR concerns this area and was put forward by Argentina and Chile during 2018, but so far not adopted (CCAMLR 2019b). The modelled oceanographic connectivity strongly suggests that *N. rossii* populations at the Antarctic Peninsula are sources for the fish assemblages off South Georgia and, to a lesser extent, the South Sandwich and South Orkney Islands. They are therefore important for re-establishing and maintaining a large population throughout the Scotia Sea. In addition, if the population of *N. rossii* is to be managed throughout its range in a precautionary approach in the future, conservation and monitoring of the northwest region of the Antarctic Peninsula marine ecosystem, including Bransfield Strait, will be important. This region experiences drastic climate change effects, including increases in temperature and reduction of ice cover, with consequences for the entire ecosystem (McClintock *et al.* 2008; Ducklow *et al.* 2013). It can therefore be a natural laboratory to detect the effects of global warming, for example, on the high-Antarctic and sub-Antarctic *Notothenia* species (*N. coriiceps* and *N. rossii*) that occur here in sympatry. Adaptive genetic variance in these species

may bear the potential to mitigate climate change effects as shown theoretically in terrestrial species (Razgour *et al.* 2019). Continued multidisciplinary investigations as presented here could help achieve adequate monitoring and prevent unsustainable loss of biodiversity.

### 4.4 Methodological considerations and future research perspectives

Applying species distribution models at large scales and in data-poor environments is challenging. Particular problems include spatially aggregated data, presence-only data, and extensive gaps in environmental data, which can be partly circumvented with appropriate calibration and validation methods (Guillaumot *et al.* 2018a, 2019). In conjunction with other methods, SDM data can be used successfully to fill specific knowledge gaps, such as the case of Bouvet Island, presented here. Sometimes, interpolating from data collected elsewhere is a valuable alternative to costly or near impossible direct observation (Gutt *et al.* 2012). In addition to collecting more data, a future improvement of SDM approaches would be the separation of the model by life stages. For *Notothenia rossii* in particular it could be highly informative to generate a robust SDM for larval stages, provided that sufficient ecological information can be gathered.

The genomic data created here have limitations especially with regard to sequencing coverage, which causes further downstream issues such as potential genotyping error and low quality genotypes that necessitated very extensive filtering (Mastretta-Yanes *et al.* 2014). The available data were thus dramatically reduced, but the overall pattern of little genetic structure and comparatively normal diversity levels is likely accurate. In contrast, determining with certainty which loci or genes are important for local adaptation would require further efforts. The reason for our low coverage is likely an underestimation of the true genome size of *N. rossii* or of the number of fragments that the restriction enzyme *ApeKI* produces. More extensive *a priori* testing could help alleviate such issues (Christiansen *et al.* Chapter 3) as well as a high quality reference genome (Fountain *et al.* 2016; Shafer *et al.* 2017).

The oceanographic model used in this study has a relatively coarse resolution, due to the large spatial scale at which it is applied and resultant computation demands. The development of high-resolution oceanographic models that better resolve fine-scale circulation features has the potential to reveal further details on the connectivity of *N. rossii*, for example regarding the

extent of local retention. In addition, the skill of the IBM is highly dependent on the accuracy of its biological parameterization. We have used the best available biological knowledge gathered over decades of research, but some uncertainty remains. In particular, there is uncertainty over the total permissible length of the pelagic phase, and the active behavior of larvae, fingerlings, and juveniles. Further knowledge of the behavior of fingerlings and juveniles, for example, would allow the incorporation of these additional life stages into the IBM. Observational evidence for feeding behavior, active swimming and diel vertical migration, would allow further refinement of the IBM, improving its predictive skill. In addition, both SDM and IBM models could be refined by further integrating spatially variable biological traits, when additional such information becomes available. Even in the absence of genetic differentiation, some spatial differences in life history parameters between *N. rossii* assemblages have been reported over large distances, but recent ecological comparisons are scarce (DeWitt *et al.* 1990; but see also Calì *et al.* 2017). Considering intraspecific diversity is difficult but the necessary next step for most accurate biological models (in SDM, IBM and population genomics) with clear conservation benefits (Mee *et al.* 2015; Marcer *et al.* 2016; Des Roches *et al.* 2018; Paz-Vinas *et al.* 2018).

## 5. Conclusion

Multidisciplinary approaches to assess connectivity are extremely useful in data-limited situations as is the case in the vast and remote Southern Ocean. The integration of data from three different sources allowed us to identify areas important for conservation and provide a hypothesis that explains the slow recovery of *Notothenia rossii* stocks. These results are relevant for the ongoing effort to establish a network of MPAs and implement ecosystem based management for the region. Further challenges lay ahead, with climate change potentially altering the suitable habitat and connectivity, which demands continued research and monitoring, and flexible, adaptive management.

**6. Acknowledgments**

We gratefully acknowledge the contribution of those colleagues providing samples for this study. In particular we thank the British Antarctic Survey, the crew and captain of RV *James Clark Ross*, the Alfred Wegener Institute, captain and crew of RV *Polarstern*, the National Natural History Museum in Paris, France and the help of G. Carvalho, G. Duhamel, M. Hautecoeur, E. Moreira, J. Rock, A. North, L. Zane, M. Belchier, N. Tysklind, and M. De Bruyn. Simulated currents from NEMO were provided by the National Oceanography Centre, Southampton. We thank F. Heindler, B. Hellemans, I. Coscia, A. McDevitt, S. R. Maier, F. Calboli and C. Moreau for laboratory help and/or advice.

**7. Funding**

This work is contribution no. 38 to the vERSO project, funded by the Belgian Science Policy Office (BELSPO, Contract no. BR/132/A1/vERSO) and contribution no. 11 to the RECTO project (BELSPO, Contract no. BR/154/A1/RECTO).The modelling work was supported by Natural Environment Research Council (NERC, UK) grant NE/H0230238/1. HC was supported by an individual PhD grant from Flanders Innovation & Entrepreneurship (VLAIO, grant no. 141328).

**8. Data archiving statement**

Demultiplexed, but otherwise raw sequencing data has been deposited on NCBI's Sequence Read Archive (SRA). Metadata for each individual using the same identification codes is available and cross-linked at GeOMe (Deck *et al.* 2017) and data.biodiversity.aq. In addition, vcf and genepop files, R scripts and additional input files for analysis are available at https://doi.org/10.5281/zenodo.3552609.

# CHAPTER 6: Current progress in population genomics, connectivity and diversity assessments of Southern Ocean fish

# Current progress in population genomics, connectivity and diversity assessments of Southern Ocean fish

Henrik Christiansen

## Overview

The ichthyofauna of the Southern Ocean represents a unique assemblage with special adaptations to freezing waters. Mesopelagic species such as lanternfish (Myctophidae) are increasingly recognized for their importance in the Antarctic food web (McCormack *et al.* 2019), while notothenioids are prime models to study adaptive radiation (Near *et al.* 2012), ecophysiology (Papetti *et al.* 2016b), and the response to global change (Mintenbeck *et al.* 2012). These features generated considerable scientific attention. However, the diversity among and within Southern Ocean fishes is still not fully understood. Marine Antarctic ecosystems are threatened by rapid environmental change and increasing anthropogenic interference. It is therefore important to continue to describe and classify Antarctic biodiversity, ideally before baseline information is irrevocably lost. Molecular approaches, especially genome-wide sequencing, modern genetic databases, and statistics can help with the identification, characterization and ultimately conservation of biodiversity (Gutt *et al.* 2017). The data, analyses and discussions presented in the preceding chapters contribute to our understanding of inter- and intraspecific diversity of the fishes of the Southern Ocean. Specifically, Chapter 2 contributed new genetic barcodes to the online database that facilitates rapid taxonomic identification of fish, stomach, or environmental DNA samples for ecological or evolutionary research. The DNA barcoding data was also used to investigate phylogeographic patterns of midwater fishes and highlight taxa that need further investigation. We then conducted a large multi-taxon pilot experiment to optimize the reduced representation sequencing technology for Antarctic animals (Chapter 3). Adaptations of this method were used to investigate the population genetic diversity and connectivity of *Notothenia coriiceps* (Chapter 4) and *N. rossii* (Chapter 5). The data reveal subtle genetic structure and putative signs of adaptation in the coastal *N. coriiceps*, while the population of the largely sub-Antarctic, bentho-pelagic congener *N. rossii* is characterized by low, but persistent long-distance connectivity leading to absent spatial genetic divergence. Here, I put the results of these empirical research chapters into perspective, starting with a critical examination of current genomic techniques. I review the state of knowledge regarding both within and between species diversity in Southern Ocean

fishes and provide a detailed synthesis and outlook regarding population genetics of notothenioid fishes. Lastly, I discuss the conservation implications of these results.

### 1. A critical evaluation of recent molecular approaches for ecological and evolutionary research

Molecular methods are of unabated importance for evolutionary biology. In addition, many recent developments have broadened the scope of molecular tools towards fundamental questions in ecological research (Hofmann & Place 2007; Thomsen & Willerslev 2015; Creer *et al.* 2016). The spectrum of studies that employ molecular technologies is currently larger than ever before. With such widespread use, however, come certain risks (Box 6.1). In the following, I identify what such risks might be, examine why they are problematic and how they could be alleviated.

One problem that arises with increasing variety and complexity in molecular methods is the dilution of knowledge. Depending on the research question and the technique employed to answer the question, several laboratory and analytical steps may require expert knowledge, that is not always available. In such cases it is not uncommon that the technique is applied without detailed knowledge, because seemingly easy instructions or commercial kits (e.g. for DNA purification) exist. Such an approach may yield satisfactory results in many cases (Graham *et al.* 2015). At times, however, this is the first step where poor results are obtained, but may remain undetected due to the "black box" nature of both molecular laboratory and bioinformatics work. At the end of a long workflow (see Fig. 1.5, p. 22), it is often impossible to trace back where errors were introduced. Examples are the reduced representation sequencing (RRS) method, DNA metabarcoding and RNAseq. Errors that could be avoided or corrected for can then possibly not be distinguished anymore from systematic errors inherent to the technique such as allele dropout or PCR duplicates. This problem is exacerbated when project work is divided among several researchers, research teams or even laboratories. For instance, during sampling erroneous preservation can lead to degraded material, or, worse, a mixture between (DNA of) individual samples. Similar problems can occur during laboratory work. In addition, due to the complexity of bioinformatics alone, this task, for example, is regularly conducted by a different researcher than the laboratory work. However, the two processes are intimately linked and require very clear and close consultation (DaCosta & Sorenson 2014; Leigh *et al.* 2018). Therefore, in order to avoid downstream issues, it is strongly recommended that one main responsible researcher plans and executes the entire operation at least once, especially in RRS,

before creating larger data sets possibly in collaboration with several researchers. Naturally, such a holistic undertaking will consume considerable amounts of time and effort, but will likely be worthwhile in the long run. Gaining a deep understanding of the intricacies of RRS for example, enables researchers furthermore to successfully adapt a working method for new projects. In this thesis, Chapter 4 is an example where the method was largely inherited from existing projects. The molecular data is satisfactory and more than sufficient to characterize neutral genetic population structure, but the inferences about adaptation remain preliminary due to comparatively little marker density. Furthermore, the coverage of the fragments is in fact much higher than required. While this fact causes most likely no harm, it indicates that the available resources have not been used as efficiently as possible. Following the experiences made during this project, the RRS method was adjusted for Chapter 5, leading to higher marker density, but at the expense of coverage (Table 6.1). Therefore, the opposite pattern was demonstrated in this case, i.e. too many genomic fragments can lead to too little coverage, which eventually results in low marker density due to many genotypes that cannot be called with high confidence. Chapter 3 outlines in detail how one can aim at a balance between these two scenarios. It seems clear that this not trivial, particularly in non-model species with little genomic information.

**Box 6.1.** Common problems associated with applying modern molecular tools in ecological and evolutionary research and possible remedies (note that many of these aspects are intertwined).

| | | |
|---|---|---|
| **Dilution of knowledge** | Many different techniques are applied by different persons instead of relying on experts for the entire workflow. | *Remedy:* Developing and cultivating expertise in one method from start to end. |
| **Inefficient use** | Suboptimal experimental setup leading to a waste of resources. | *Remedy: A priori* testing and optimization. |
| **Erroneous use** | Introducing errors through inappropriate use of laboratory or bioinformatics procedures. | *Remedy:* Extensive technical replication within and/or between methods. |
| **Fast technological development** | Extremely rapid development in sequencing technology outruns the skill development of individual researchers and the development of appropriate statistics and replicated technological assessment studies. | *Remedy:* Choosing only well-established methods or planning considerable time and resources for optimization and verification. |
| **False incentives** | Using genetics/genomics for the sake of itself or for the perceived advantage of applying modern methods. | *Remedy:* Critical evaluation of the research question and which methodology is most appropriate to answer it. |

Methodological setups that are not fully optimized can lead to inefficient, or worse, erroneous use of a molecular method. As outlined above, too high coverage in a RRS setup e.g., translates eventually into wasting precious research funds and energy. However, the degree of optimization versus the amount of inefficiency that is tolerable must be balanced by each group individually. This trade-off should be carefully evaluated, as optimization itself also consumes time and resources. Here, a parallel optimization strategy for several taxa/projects at the same time can improve overall efficiency. Chapter 3 is an example of how larger research consortia may work together to achieve reasonable optimization without excessive amounts of time and funding. Similar endeavors have been conducted elsewhere (DaCosta & Sorenson 2014; Henri *et al.* 2015; Herrera *et al.* 2015; Burford Reiskind *et al.* 2016; McCartney-Melstad *et al.* 2016), highlighting the utility of this approach independent of the ecosystem or species in focus.

Erroneous inferences are arguably the most serious problem associated with recent molecular approaches. There are many possibilities during most advanced methods such as RRS or large scale metabarcoding where errors or biases can be introduced and lead to false outcomes. For example, the process of DNA extraction itself can introduce bias between samples when libraries with DNA from many individuals are constructed (Graham *et al.* 2015). Sex specific differences between individuals and contamination issues must also be considered (Benestan *et al.* 2017). Furthermore, high-throughput sequencing data generally contain many more genotyping errors than traditional Sanger sequencing data (Shendure & Ji 2008). Different types of sequencing platforms can cause severe compatibility issues and should be avoided within projects (Leigh *et al.* 2018). The same applies to different RRS protocols (Flanagan & Jones 2018), which may be partly due to issues such as allele dropout or PCR duplicates (Gautier *et al.* 2013). When choosing an RRS method it should be kept in mind whether the method allows for technical exclusion of PCR duplicates (such as the original RADseq; Baird *et al.* 2008) or if a specific laboratory step can be added to achieve this (e.g. Schweyen *et al.* 2014). Recent investigations indicate that PCR duplicates are likely no substantial source of bias for RRS (Euclide *et al.* 2019). Nevertheless, as a consequence of the differences between platforms and protocols, projects need to define the methodological details to be used early on and should stick to these whenever possible. In addition, very stringent filtering and technical replication is necessary (Mastretta-Yanes *et al.* 2014; Fountain *et al.* 2016; O'Leary *et al.* 2018). However, the data processing, i.e. in high-throughput data typically bioinformatics procedures, is error-prone as well and highly depends on the software, pipeline and parameters employed (Paris *et al.* 2017; Shafer *et al.* 2017; Díaz-Arce & Rodríguez-Ezpeleta 2019). Within Chapter 3, these considerations were recognized and

mitigated as well as possible, although the RRS protocols used there do not enable the identification of PCR duplicates. Addition of random oligo-nucleotides into the barcodes of these protocols could be future improvements. However, such an approach would also increase costs significantly as the barcode adapters would have to be designed and ordered anew. The uncertainty surrounding recent methods, however, indicates that these are still far from being established and no consensus on the most appropriate, statistically robust and replicable approach exists yet (Andrews *et al.* 2014; Puritz *et al.* 2014b).

The difficulties regarding very recent molecular tools is also related to their fast technological development. Output of sequencing technologies has accelerated dramatically in the early 21st century with the development of high-throughput machines that perform millions of sequencing reactions in parallel (Heather & Chain 2016). The first platforms for massively parallel sequencing were the 454 Life Sciences machines, which, however, were already discontinued in 2013, because they became noncompetitive. Instead, the Illumina HiSeq platform had established itself as the *de facto* standard for the so-called next generation sequencing. By now, HiSeq machines are already being replaced by NextSeq platforms (e.g. genomicscore.be). The sequencing methods and platforms listed here are merely a superficial look at the most well-known approaches. Detailed reviews highlight the many varieties and technologies that have rapidly been developed (and often superseded) in the past couple of years (Shendure & Ji 2008; Metzker 2010; Nielsen *et al.* 2011; Moriarty Lemmon & Lemmon 2013; Goodwin *et al.* 2016; Heather & Chain 2016). This wealth of sequencing methods illustrates how fast and at times confusing the technological development is. Even experts are therefore challenged in order to stay informed and must constantly update their knowledge and skills, possibly faster than in many other biological disciplines. The fact remains and cannot be overstated that one should always critically assess whether a particular molecular method is really suitable or necessary for a project and if the required expertise is available. In addition, different sequencing methods and their statistics may be difficult to compare leading to reproducibility problems. A typical problem of population genetics is that studies from different time periods are often incomparable, because it is impossible to distinguish differences caused through true biological changes from methodological artifacts. Many studies have, for example, attempted to compare microsatellites and SNPs (Bradbury *et al.* 2015; Jeffries *et al.* 2016; Fischer *et al.* 2017; Bohling *et al.* 2019), but even comparisons between these studies is difficult due to the large variety of methodological flavours used within them.

**Table 6.1.** Overview of genomic data created within this thesis and associated projects with details on the target species and estimated/assumed genome size (GS in Gbp; for details on this see respective chapter), reduced representation sequencing method (RRS), restriction enzyme (RE), size selection window (SSW in bp), sequencing platform, and the coverage, number of loci (before filtering) and coverage obtained.

| PhD Chapter | Species | GS | RRS | RE | SSW | Platform | Loci | Coverage |
|---|---|---|---|---|---|---|---|---|
| Christiansen, Ch. 3 | Macrocyprididae | 0.29 | GBS | *ApeKI* | 200-350 | HiSeq 2500 | 69,817 | 28.2x |
| Christiansen, Ch. 3 | *Laternula elliptica, Aequiyoldia eightsii* | 3.00 | GBS | *ApeKI* | 200-260 | HiSeq 2500 | 125,305, 143,551 | 21.6x, 20.0x |
| Christiansen, Ch. 3 | *Bathybiaster loripes, Psilaster charcoti* | 0.50 | GBS | *ApeKI* | 200-300 | HiSeq 2500 | 82,945, 115,608 | 21.0x, 27.6x |
| Christiansen, Ch. 3 | *Pagodroma nivea* | 1.14 | GBS | *PstI* | 200-300 | HiSeq 2500 | 140,972 | 10.0x |
| Christiansen, Ch. 4 | *Notothenia coriiceps* | 0.64 | ddRAD | *SbfI* & *SphI* | 320-590 | HiSeq 2500 | 40,342‡ | 58.9x |
| Christiansen, Ch. 5 | *Notothenia rossii* | 0.64 | GBS | *ApeKI* | 240-340 | HiSeq 2500 | 73,554‡ | 10.6x |
| Delerue-Ricard, Ch. 3* | *Solea solea* | 0.70 | ddRAD | *SbfI* & *SphI* | 300-600 | HiSeq 2500 & 4000 | 18,442‡ | 214.1x |
| Heindler, Ch. 5† & Christiansen, Ch. 3 | *Trematomus* spp. | 1.50 | ddRAD | *EcoRI* & *SbfI* | 200-450 | HiSeq 2500 | 23,609 | 49.6x |

* Delerue-Ricard, S., Christiansen, H., Maes, G.E., Manchado, M., Hellemans, B., Coscia, I. & Volckaert, F.A.M. (2019) Small-scale population genomic patterns of various life stages of the flatfish sole in the Northeast Atlantic Ocean. *In:* Connectivity of larval and juvenile common sole at a small and large spatial scale. PhD Thesis. 2019, KU Leuven.

† Heindler, F.M., Christiansen, H., Hellemans, B., Maes, G. E., Van de Putte, A.P. & Volckaert, F.A.M. (2019) Diffuse species separation and habitat associated differentiation in a marine adaptive radiation. *In:* Ecology and Evolution of Fishes of the Southern Ocean. PhD Thesis. 2019, KU Leuven.

‡ reference-based analysis.

Despite the difficulties relating to its pace, the technological development enables vast opportunities for genomics in, for example, conservation and fisheries biology, and has consequently led to repeated calls to "harness" the power of these modern methods for specific research fields (Funk *et al.* 2012; Ovenden *et al.* 2015; Andrews *et al.* 2016; Bernatchez *et al.* 2017; Obregón *et al.* 2018). Of course, such claims are justified and indeed many successful applications are documented (see e.g. McKinney *et al.* 2016 and Catchen *et al.* 2017 for a list of positive RADseq examples). Nevertheless, one might point out that there are cases where other non-molecular methods are equally well or better suited to answer a question. Traditional taxonomy and detailed morphological analyses for example are lagging compared to genetics/genomics, which are now very commonly applied in biological research (Godfray 2002; Agnarsson & Kuntner 2007). Sometimes one has the impression that genetics/genomics are used for the sake of genomics and not for the specific research question. The "profusion of confusion" (Hadfield & Retief 2018) in sequencing methods is possibly contributing to this problem, alongside incentives that drive researchers to always strive for the newest, most modern, most "state of the art" approach. It is argued that much of the research presented in this thesis in fact depends on molecular approaches and that these have provided novel insights that contribute to a better understanding of the evolution of Antarctic fishes. It must not be forgotten though, that one should always critically evaluate whether a genetic method in general and which one in particular is required for the task at hand (McMahon *et al.* 2014). Taxonomic expertise for myctophid fishes (Chapter 2), for example, is scarce and difficult to train. More expertise in this specific field would benefit the global research community. Regarding population connectivity studies, genomic methods are certainly powerful and provide useful data, but other approaches could be used as well. Tagging, direct observations, otolith microchemistry analysis, parasitology, or dispersal modelling are all used successfully for connectivity assessments (Ashford *et al.* 2017; Pinsky *et al.* 2017; Paris *et al.* 2018; Wright *et al.* 2018). In Chapter 5, dispersal modelling was used in conjunction with genomics to facilitate a better, more complete understanding of biological reality. Using several methodologies in combination is certainly a promising approach to evaluate connectivity between populations and is encouraged for future studies (Manel *et al.* 2019).

### 1.1 The case of reduced representation sequencing – beyond RAD?

The case of RADseq or, more generally, RRS, features all above-mentioned problems. As with other methods, laboratory work and bioinformatics are regularly implemented by different

persons, which is not ideal as the two processes strongly affect each other. For example, the choice of RRS protocol determines whether it is possible to computationally identify and exclude PCR duplicates (Andrews *et al.* 2014). We also noticed that sequencing reads are more evenly distributed across individuals when pooling only after the second PCR (Chapter 3). Other properties of the laboratory work, that need consideration when processing the data, are the DNA quality and quantity, the choice of barcodes, the number of PCR cycles, and the size selection (Fig. 6.1). These and the experimental setup (most importantly the restriction enzyme, number of individuals, libraries and the sequencing platform) determine the coverage. Efficiency and accuracy both depend on coverage. Inefficient use can be mitigated using pilot experiments and prior optimization of several taxa/projects in parallel, as conducted in Chapter 3. One of the most pervasive issues is possible erroneous inferences with adverse consequences especially in situations with practical conservation implications. In Chapter 5, a global genotyping error rate (importantly this includes possible errors introduced through PCR duplicates) was calculated as ranging between 1.48 % and 5.30 %. Arguably, such error rates are dramatic and would not be accepted in other disciplines. However, population genomics belongs to what is now often called the "big data" community and the sheer amount of information would suggest that average results across these large amounts of data are still more accurate than information from a handful of markers. Microsatellite genotyping for example can also contain errors, due to the subjective nature of genotype calls, null alleles, stutter peaks or incorrect statistic assumptions (Chistiakov *et al.* 2006; Putman & Carbone 2014). Indeed, quantitative comparisons between microsatellites and/or mitochondrial DNA markers and RRS data overwhelmingly suggest that the latter perform at least as well or better than the former (Jeffries *et al.* 2016; Hodel *et al.* 2017; Langin *et al.* 2018; Bohling *et al.* 2019). The original description of RADseq dates back to 2008 when a single enzyme restriction-associated DNA sequencing protocol with random shearing was developed (Baird *et al.* 2008). In the meantime, at least a dozen alternatives have been published (Elshire *et al.* 2011; Peterson *et al.* 2012; Poland *et al.* 2012; Toonen *et al.* 2013; Heffelfinger *et al.* 2014; Ali *et al.* 2016; Suchan *et al.* 2016; Franchini *et al.* 2017; Campbell *et al.* 2018; Bayona-Vásquez *et al.* 2019), making it complicated to choose the appropriate protocol (see also literature discussions about this issue: Andrews *et al.* 2014; Puritz *et al.* 2014). Ultimately it depends on the research objectives as well as the facilities and expertise available for the project to choose a method that is well understood, including its potential pitfalls (Puritz *et al.* 2014b). Lastly, there might be a certain "hype" around RRS methodologies, and it is therefore important to realize that other options exist that are versatile and sometimes better alternatives.

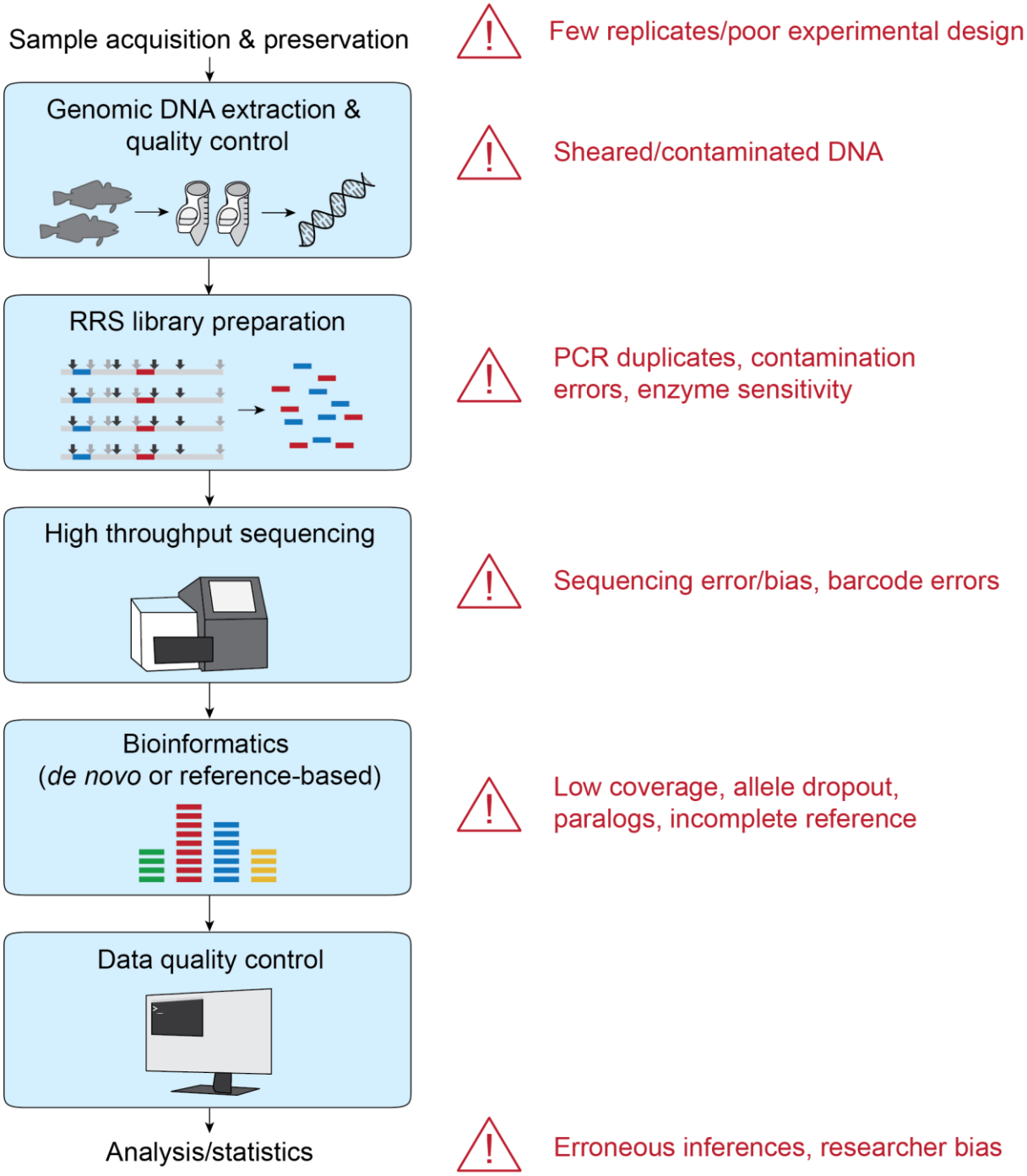



**Fig. 6.1.** Typical problems associated with reduced representation sequencing that can lead to serious downstream issues. The standard procedural steps are indicated (see also Fig. 1.5, p. 22) in boxes and the most common problems are listed in red (note that this list is not exhaustive).

What could thus be ways forward, beyond reduced representation sequencing? One logical advance is the progression towards shallow or deep re-sequencing of entire genomes across many individuals (Fuentes-Pardo & Ruzzante 2017; Therkildsen & Palumbi 2017). Whole genome sequencing is providing the most complete information about the genomic make-up of wild populations. However, it is also substantially more expensive than for example RADseq, although recent advances have decreased the per-sample cost for shallow resequencing dramatically (Therkildsen & Palumbi 2017). Not long ago, it was argued that whole genome sequencing is not necessary for most population genomic research questions (Allendorf *et al.* 2010). Recently, however, evidence is accumulating that large structural genomic changes are most important in shaping the adaptation and divergence of fishes (Barrio *et al.* 2016; Berg *et al.* 2016; Barth *et al.* 2019; Catanach *et al.* 2019). Chromosomal inversions in particular seem important for a wide variety of ecological and evolutionary processes (Wellenreuther & Bernatchez 2018). In Antarctic fishes, recent research shows that transposable elements, for example, are related to chromosome variability and rapid speciation (Auvinet *et al.* 2018). The number of chromosomes among notothenioids is highly variable, ranging from 22 to 58 (2n) (Ozouf-Costaz *et al.* 1991; Detrich *et al.* 2010; Amores *et al.* 2017). Therefore, moving beyond RADseq will ultimately be indispensable to fully appreciate the complex interplay between genomes, ecology, connectivity, and speciation. As we will see later (2.3), draft genomes of Antarctic fishes are increasingly becoming available. Shallow whole genome re-sequencing thus seems feasible already now, at least for some species. Higher mapping precision and improved statistical inferences to detect, quantify and understand genomic adaptation can be expected.

### 2. The status of intra- and interspecific fish diversity in the Southern Ocean

The fishes of the Southern Ocean, particularly the notothenioids, have received a lot of research attention over the past decades. They are not only fascinating and exotic, but also provide good models to study adaptation to climate extremes, strong seasonality, and developmental peculiarities (Detrich & Amemiya 2010; Braasch *et al.* 2014). Here, I present a brief overview of the current state of knowledge, including the progress presented in the previous chapters, on 1) the status of interspecific fish diversity, i.e. the number of currently recognized species and how much genetic information is available for different families of Southern Ocean fish, 2) the status of intraspecific diversity, i.e. results from recent population genetic and comparative ecological studies with a focus on notothenioids, and 3), recent advances and perspectives in Southern Ocean fish genomics.

### 2.1 The number of Southern Ocean fish species and available data for non-notothenioids

In 2014, one of the most comprehensive resources on the biota of the Southern Ocean, with special reference to their distribution, was published (De Broyer *et al.* 2014). A total of 451 fish species are listed, 82 of which are only occasionally recorded south of the Subtropical Convergence (STC) (Duhamel *et al.* 2014). The authors separated the fish fauna into neritic, deep sea and pelagic compartments, respectively. Since then, one may add 9 additional species (Balushkin 2013; Balushkin & Spodareva 2013a; b, 2015; Spodareva & Balushkin 2014; Eakin *et al.* 2015; Balushkin & Zhukov 2016; Balushkin & Moganova 2017, 2018), resulting in a total of 460 species (Table 6.2). The total fish species diversity in the Southern Ocean is still expected to increase, with several recent new species descriptions, and more to be expected, particularly among snailfishes (Liparidae) and plunderfishes (Artedidraconidae) (Eastman 2005). Nevertheless, in comparison with other oceans the fish diversity here is clearly impoverished (Duhamel *et al.* 2014). This is believed to be the result of the extreme conditions and repeated glacial cycles that only few species could adapt to (Eastman & McCune 2000; Eastman 2005). As we have seen, the fish that do prevail in this icy habitat are shaped by their environment and often unique, i.e. endemic to this region. The adaptive radiation of the Notothenioidei is of special interest to many researchers spanning various disciplines and has thus received considerable research attention (some recent examples: Bogan & Place 2019; Desvignes *et al.* 2019; Ferrando *et al.* 2019; Kim *et al.* 2019a; Roche *et al.* 2019). Much remains to be discovered though, as neither the phylogenetic status of all species of Notothenioidei is resolved (Near & Cheng 2008; Lautrédou *et al.* 2012; Dettai *et al.* 2012; Near *et al.* 2018), nor are all of them discovered yet (at least that is very unlikely; Eastman 2005).

Regarding other, non-notothenioid fish species of the Southern Ocean knowledge is scarce (but see e.g. Duhamel *et al.* 2010; Horn & Sutton 2015; Amsler *et al.* 2016; Vedishcheva *et al.* 2019). Only lanternfishes (Myctophidae) have recently received increased attention, in part because they are now appreciated as an important, alternative energy pathway in the Southern Ocean food web next to Antarctic krill (McCormack *et al.* 2019; Saunders *et al.* 2019). It is still difficult to accurately estimate the biomass of these fishes, even using acoustic methods, because their backscatter properties change in high latitudes (Dornan *et al.* 2019). In addition, the available biological information about myctophids is heavily spatially biased, with most data pertaining to the Scotia Sea region (Saunders *et al.*, 2014, 2015c, b; a; Lourenço *et al.* 2016; Belcher *et al.* 2019). Trawl data suggests that the lanternfish assemblage in this region depends on mass

immigration from lower latitudes (Saunders *et al.* 2017). Further studies investigating the ecological role and trophic importance of myctophids in other regions are in progress (e.g. Cipro *et al.* 2017; Moteki *et al.* 2017; Clarke *et al.* 2018). A synthesis to determine their overall status in the Southern Ocean and especially their response to future global change is needed to develop suitable conservation plans (Freer *et al.* 2018, 2019). All of these studies depend on accurate specimen identification, which can be rapidly delivered using the DNA barcode database, to which we contributed in Chapter 2. Given the paucity of taxonomic expertise, it will be important to further update this database with hitherto unrepresented taxa. Such a task extends beyond the myctophids, as especially bathypelagic and benthic deep-sea species are vastly understudied and for many species no DNA barcodes are available yet (Table 6.2). In addition, it will be important to understand the geographic variation in diversity of non-notothenioid fishes, both in terms of species richness, as well as with regards to intraspecific variability. The latter is almost completely unknown for non-notothenioid Southern Ocean fish. Only the most common Antarctic myctophid was investigated for population structure using several genetic markers. This species, *Electrona antarctica*, is present in large numbers, which in conjunction with high oceanographic connectivity causes genetic homogeneity (Van de Putte *et al.* 2012a). As our preliminary phylogeographic analyses suggest, such a pattern might be common among many Southern Ocean myctophids (Chapter 2). Even fewer studies have looked at genetic variability in Southern Ocean fishes that are neither notothenioids nor myctophids. DNA barcoding was successfully used to highlight and subsequently describe a new species of grenadier (Macrouridae) from the Southern Ocean (Smith *et al.* 2011b; McMillan *et al.* 2012). Similar insights can be expected among other, understudied fish families of the Southern Ocean. For example, the lanternfish *Gymnoscopelus bolini* should be studied in more detail for intraspecific diversity and the deep-sea smelts *Bathylagus* spp. for the presence of cryptic species (Chapter 2). The bottleneck to such investigations is no longer sequencing capability, but taxonomic expertise and sampling in remote areas.

**Table 6.2.** Overview of fish species in the Southern Ocean, updated (updates underlined) after Duhamel *et al.* (2014). Families in bold denote fish families that have been studied in this thesis (Chapter 3, 4, and 5: only Nototheniidae; all others studied in Chapter 2). Numbers in brackets denote species that were already listed in the previous categories (neritic and/or deep sea; thus not contributing to the total number). Species with barcodes available in the barcode of life data system (BOLD; Ratnasingham & Hebert 2007) are listed in the last column.

| Family | Neritic | Deep sea | Pelagic | Total | BOLD |
|---|---|---|---|---|---|
| Myxinidae* | 1* | | | **1*** | 1* |
| Petromyzonidae | 1 | | | **1** | 1 |
| Lamnidae | 1 | | 1 | **1** | 1 |
| Squalidae | 1* | 3 | (1*) | **4** | 2 |
| Somniosidae | 1 | (1) | | **1** | 0 |
| Rajidae | 7 | 3 (5) | | **10** | 10 |
| Halosauridae | | 2 | | **2** | 1 |
| Notacanthidae | | 2 | | **2** | 2 |
| Synaphobranchidae | | 3 | | **3** | 3 |
| Nemichthyidae* | | | 4* | **4*** | 3* |
| Serrivomeridae* | | | 1* | **1*** | 1* |
| Microstomatidae | | | 1 | **1** | 1 |
| **Bathylagidae** | | | 5 | **5** | 3 |
| Alepocephalidae | | 4 | | **4** | 4 |
| Platytroctidae* | | | 2* | **2*** | 2* |
| Gonostomatidae | | | 3+3* | **3+3*** | 2+3* |
| Sternoptychidae | | | 9* | **9*** | 8* |
| **Stomiidae** | | | 4+3* | **4+3*** | 3+3* |
| **Scopelarchidae** | | | 2 | **2** | 2 |
| Notosudidae | | | 2 | **2** | 2 |
| **Paralepididae** | | | 4+1* | **4+1*** | 3+0* |
| Anotopteridae | | | 1 | **1** | 0 |
| Alepisauridae | | | 1 | **1** | 1 |
| Ipnopidae | | 1 | | **1** | 1 |
| **Myctophidae** | | | 24+44* | **24+44*** | 24+37* |
| Muraenolepididae | 8 | | | **8** | 1 |
| **Macrouridae** | | 10+1* | (1) | **10+1*** | 9+1* |
| Moridae | | 6 | | **6** | 6 |
| Melanonidae | | | 1 | **1** | 1 |
| Gadidae* | | | 1* | **1*** | 1* |
| Merlucciidae* | | | 1* | **1*** | 1* |

| | | | | | |
|---|---|---|---|---|---|
| Ophidiidae | | 4 | | **4** | 1 |
| Bythitidae | | 1 | | **1** | 1 |
| Carapidae | | 1 | | **1** | 1 |
| Ceratiidae | | | 1 | **1** | 1 |
| **Oneirodidae** | | | 1 | **1** | 1 |
| Melanocetidae | | | 2 | **2** | 0 |
| Gigantactinidae | | | 1 | **1** | 0 |
| Lampridae | | | 2 | **2** | 2 |
| **Melamphaidae** | | | 3 | **3** | 3 |
| Cetomimidae | | | 3 | **3** | 1 |
| Oreosomatidae | | 1+1* | | **1+1*** | 1+1* |
| Congiopodidae | 2 | | | **2** | 1 |
| Psychrolutidae | | 2 | | **2** | 1 |
| Liparidae | 1 | 85+7* | (2) | **85+7*** | 16+0* |
| Bathylutichthyidae | | | 1 | **1** | 0 |
| Zoarcidae | 1 | 38 | 3 (4) | **42** | 14 |
| Bovichtidae | 1 | | | **1** | 0 |
| **Nototheniidae** | 90+1* | 22 (5) | 9 | **121+1*** | 84+0 |
| Tripterygiidae | | 1 | | **1** | 0 |
| Chiasmodontidae | | | 3 | **3** | 2 |
| Gempylidae | | 1 | (1) | **1** | 1 |
| Scombridae* | | | 2* | **2*** | 2* |
| **Centrolophidae** | | 1 | (1) | **1** | 1 |
| Achiropsettidae | 4 | (2) | | **4** | 2 |
| Total | 117+3* | 205+9* | 86+72* | **378+82*** | 217+64* |

* only occasionally recorded south of the Subtropical Convergence

### 2.2 Insights from 25 years of population genetic research on notothenioids

In contrast to most other Southern Ocean fish families, the notothenioid families or subfamilies (*sensu* Duhamel *et al.* 2014) have been studied extensively. Population genetics of notothenioid fishes has been an active research field since the 1990s with now approximately 38 published studies (Table 6.3). And yet it seems that no consensus has been reached regarding the prevalence and extent of population differentiation. Generally, genetic population differentiation emerges when two or more (sub-) populations of the same species experience limited gene flow due to physical (allopatry) or biological barriers (sympatry), such as

reproductive isolation. In the marine realm, genetic differentiation was believed to be largely absent, because few physical barriers exist compared to terrestrial systems that are often interrupted by mountains, valleys, rivers, etc. (Hauser & Carvalho 2008). A closer look, however, reveals that marine ecosystems are geographically structured as well, predominantly through the characteristics and movement of different water masses (White *et al.* 2010; Selkoe *et al.* 2010). Examples of biological barriers among fish populations are well-described, too, such as temporally separated spawning (Barrio *et al.* 2016; Lamichhaney *et al.* 2017). In addition, partial isolation (parapatry) may be one of the most common modes of speciation among marine fishes (Rocha & Bowen 2008).

Regarding the notothenioids, several studies have attempted to elucidate the status of genetic differentiation and which factors might determine it. Matschiner *et al.* (2009) provided an extensive overview in a similar fashion as listed here (Table 6.3). These authors concluded that gene flow facilitated by long pelagic larval duration is the most common mode among notothenioids causing genetic homogeneity over large distances (Matschiner *et al.* 2009). In the meantime, however, it has been unveiled that several notable exceptions to this rule exist. Some benthic *Trematomus* species (Janko *et al.* 2007; Van de Putte *et al.* 2012b), the icefishes *Chaenocephalus aceratus* and *Champsocephalus gunnari* (Kuhn & Gaffney 2006; Papetti *et al.* 2009; Damerau *et al.* 2012, 2014; Young *et al.* 2015), and *Notothenia coriiceps* (Chapter 3) all show some spatial genetic structure. Clearly the genetic structure cannot simply be predicted from the phylogenetic grouping, as several closely related species show contrasting population genetic signals (Van de Putte *et al.* 2012b). An interesting case is the *Lepidonotothen* genus. Here, several species show virtually not any or very little population differentiation (*L. larseni*, *L. squamifrons*). Initially, *Lepidonotothen squamifrons* and *L. kempi* were even described as geographically separated, distinct species (DeWitt *et al.* 1990). Genetic analyses, however, clearly indicate large amounts of gene flow that do not substantiate the species split or a pronounced genetic structure (Miya *et al.* 2016b), albeit some differentiation over large distances (Schneppenheim *et al.* 1994; Jones *et al.* 2008). No differentiation between various population samples was found in *L. larseni* (Jones *et al.* 2008; Miya *et al.* 2016b; Deli Antoni *et al.* 2019). In contrast, *L. nudifrons* was recently split into two, morphologically cryptic species that are distributed at the western Antarctic Peninsula and the South Georgia and South Sandwich Islands, respectively (Dornburg *et al.* 2016). The species furthermore also shows genetic differentiation between the Antarctic Peninsula and the South Shetland Islands (Deli Antoni *et al.* 2019). *Lepidonotothen mizops* is another close relative of *L. nudifrons* and *L. cf. nudifrons*,

that inhabits the Kerguelen Plateau (DeWitt *et al.* 1990). Phylogenomic reconstructions reveal that *L. squamifrons* is a sister lineage to *Patagonotothen*; the remaining *Lepidonotothen* species are therefore recommended to be considered as *Notothenio*ps (Near *et al.* 2018). Irrespective of taxonomic revision, the example of *Lepidonotothen*/*Notothenio*ps may serve as valuable illustration of the factors that are of primary importance in determining the genetic differentiation of notothenioids across space, i.e. demographic history, the environment, and the species' autecology (Box 6.2).

**Box 6.2.** Biological (B) and physical (P) factors that influence the genetic population structure of Southern Ocean fish.

| | |
|---|---|
| **Oceanographic currents (P)** | Current strength, direction and seasonality strongly affect the dispersal of eggs and or larvae. |
| **Physical barriers (P)** | Geographic features such as water depth, fjords, deep sea trenches, canyons, seamounts or land masses influence the available habitat. |
| **Environmental setting (P)** | Water characteristics like temperature, salinity and nutrient concentration determine which habitat is suitable and thus affect the amount of habitat continuity or fragmentation. |
| **Adult life style (B)** | Migration and generally differences along the bentho-pelagic axis contribute to increased or decreased genetic differentiation. |
| **Pelagic larval duration (B)** | Long pelagic larval durations, as common in notothenioids, promote gene flow and thus genetic homogeneity. |
| **Parental care and offspring size (B)** | Fewer, larger eggs that are actively guarded may restrict gene flow in contrast to broadcasting spawners that release many eggs into the water column (see also fecundity). |
| **Fecundity (B)** | Number of eggs per female released affects the chances of successful dispersal. |
| **Population size (B)** | Abundance and especially effective population size (*Ne*) influence how many migrants per generation may connect spatially separate occurrences. |
| **Competition (B)** | Predation may counteract high population size and fecundity. |

The extant ichthyofauna of the Southern Ocean is a result of the geological history of Antarctica and its surrounding ocean (Eastman 1993). Consequently, processes such as glacial advances and retreats during the Pleistocene have likely left an imprint on notothenioid fish populations. It has been hypothesized that climate fluctuations may have impacted benthic and pelagic fish species differently leading to distinct demographic histories and population genetic signatures (Janko *et al.* 2007). Indeed, benthic species have experienced more severe population

bottlenecks and expansions than pelagic taxa (Kašparová *et al.* 2015). While it is certainly important to consider the demographic history over large time scales, biological specializations along the bentho-pelagic axis also affect contemporary population structure (Van de Putte *et al.* 2012b; Young *et al.* 2015). The coastal, inshore notothenioid *L. nudifrons* is a typical example of a species that was subject to strong population fluctuations, because its global metapopulation was repeatedly driven into refugia when glacial ice cover advanced in the past (Kašparová *et al.* 2015). The species has the most pronounced spatial genetic structure among notothenioids investigated to date (Kašparová *et al.* 2015; Dornburg *et al.* 2016; Deli Antoni *et al.* 2019). *Notothenia coriiceps* resembles *L. nudifrons* in its coastal occurrence, but is likely more abundant and with longer pelagic larval duration. These features may explain the low genetic differentiation in *N. coriiceps* (Chapter 4).

In addition, the physical setting, in particular the oceanographic features of the Southern Ocean profoundly affect its biota (Young *et al.* 2012, 2015). Wind and ocean currents can facilitate large scale dispersal of Antarctic organisms (Barnes *et al.* 2006; Fraser *et al.* 2018). The considerable extent and velocity of the Antarctic Circumpolar Current is a major force that connects populations on trans-oceanic scales. *Lepidonotothen larseni* exhibits a very long pelagic larval phase of likely more than 12 months, among the longest of all notothenioids. This trait exposes the offspring of *L. larseni* for a long time to the Southern Ocean current systems and thus enables regular long-distance connectivity. It is therefore not surprising that *L. larseni* shows signs of high gene flow and absent spatial genetic divergence (Jones *et al.* 2008; Miya *et al.* 2016b; Deli Antoni *et al.* 2019). This connectivity mechanism is analogous to the findings regarding *N. rossii* as presented in Chapter 5, although our detailed dispersal modelling data suggests that large-distance gene flow happens only via stepping stones. Even indirect trans-oceanic ecological connectivity may be sufficient to maintain long term evolutionary connectivity.

Lastly, the ecology of each species cannot be neglected when describing and interpreting the presence or absence of genetic differentiation. In this context, it is important to note that the characterization of species as benthic, pelagic or intermediate may be a useful proxy for studying life history effects (Janko *et al.* 2007; Van de Putte *et al.* 2012b; Kašparová *et al.* 2015), but it is also a strong simplification. The classification of notothenioids according to life style is furthermore sometimes contradictory or ambiguous. The aforementioned *L. larseni* for instance is sometimes described as benthic (DeWitt *et al.* 1990), and elsewhere as semi-pelagic (Eastman

1993). Given that the species feeds mainly on krill (DeWitt *et al.* 1990; Barrera-Oro 2002; Kock *et al.* 2013), it seems likely that it forages at least occasionally in the water column, which would contradict with a strictly benthic life style. Other biological traits that likely have a strong influence on the dispersal potential and therefore the genetic diversity in space are the overall population size, particularly the effective population size (*Ne*), the pelagic larval duration as mentioned above, but also the egg size, the type of eggs (demersally attached or pelagic), parental care, the size at hatching and the fecundity (Box 6.2). *Lepidonotothen squamifrons* is arguably a more benthic species that *L. larseni*, but also shows high levels of gene flow (Schneppenheim *et al.* 1994; Jones *et al.* 2008; Miya *et al.* 2016b). This species has the highest fecundity of all notothenioids that were investigated with population genetic methods (according to the information available in DeWitt *et al.* 1990). Such strong reproductive output in conjunction with an again long pelagic larval phase and a widespread distribution may together explain the largely absent genetic structure.

**Table 6.3.** Overview of population genetic/genomic investigations of notothenioid fishes in the Southern Ocean, updated after Matschiner *et al.* (2009). The genetic method used is listed (A, allozyme electrophoresis; RFLP, restriction fragment length polymorphism; mtDNA, mitochondrial DNA locus/loci; nDNA, nuclear DNA locus; STR, microsatellites; RAPD, randomly amplified polymorphic DNA; RRS, reduced representation sequencing), as well as the shortest distance over which significant differentiation was found (ds) and the longest distance over which no significant differentiation could be detected (dns). The distances (taken from Matschiner *et al.* (2009) or approximated using Google Maps) are accompanied with location codes denoting the localities between which such (non) differentiation was found. Location codes are as follows (approximately clockwise): western Antarctic Peninsula (AP), South Shetland Islands (SSh), Joinville Island (JI), Elephant Island (EI), North Scotia Ridge (NSR), Falkland Islands (FI), Weddell Sea (WS), South Orkney Islands (SO), Shag Rocks (SR), South Georgia (SG), South Sandwich Islands (SSa), Bouvet Island (BI), Edward and Marion Islands (EM), Ob and Lena Banks (OL), Crozet Islands (CI), Cosmonauts Sea (COS), Cooperation Sea (CPS), Kerguelen Islands (KI), Heard and MacDonald Islands (HM), Mawson Sea (MAS), Dumont D'Urville Sea (DDU), Macquarie Island (MQ), Ross Sea (ROS).

| **Species** | **Method** | **ds (km)** | **dns (km)** | **Reference** |
|---|---|---|---|---|
| *Champsocephalus gunnari* | RFLP | - | 400 KI-HM | Williams *et al.* 1994 |
| *Lepidonotothen squamifrons* | A | 6400 | 1300 | Schneppenheim *et al.* 1994 |
| *Champsocephalus gunnari* | A | - | 6400 SO-KI | Duhamel *et al.* 1995 |
| *Notothenia rossii* | A | - | 400 SB-KI | Duhamel *et al.* 1995 |
| *Chionodraco myersi* | A | 16 WS-WS | < 1000* WS | Clément *et al.* 1998 |
| *Neopagetopsis ionah* | A | 4600 WS-CPS | < 1000* WS | Clément *et al.* 1998 |
| *Dissostichus eleginoides* | STR | 60 MQ-MQ | - | Reilly & Ward 1999 |
| *Dissostichus eleginoides* | A & STR | 2000 MQ-ROS | 6000 HM-MQ | Smith & McVeagh 2000 |
| *Dissostichus eleginoides* | A & mtDNA & STR | ~500 | 8300 | Smith & Gaffney 2000 |

| | | | | |
|---|---|---|---|---|
| *Dissostichus eleginoides* | RFLP & STR | 5200 MQ-HM | - | Appleyard *et al.* 2002 |
| *Dissostichus mawsoni* | RAPD | 4700 ROS-AP[$] | - | Parker *et al.* 2002 |
| *Chionodraco hamatus* | mtDNA | 1000 DDU-RS | 9300 WS-DDU | Patarnello *et al.* 2003 |
| *Dissostichus eleginoides* | RFLP & STR | - | 2600 EM-HM | Appleyard *et al.* 2004 |
| *Dissostichus eleginoides* | RFLP & STR | 500 NSR-SR | 1300 Patagonian Shelf-NSR | Shaw *et al.* 2004 |
| *Dissostichus mawsoni* | mtDNA & STR | - | 5000 ROS-SSh[$] | Smith & Gaffney 2005 |
| *Champsocephalus gunnari* | mtDNA & nDNA | 1200 SSh-SR | 4400 BI-HM | Kuhn & Gaffney 2006 |
| *Dissostichus eleginoides* | mtDNA & STR | 1200 FI-SG | 5100 SG-OL | Rogers *et al.* 2006 |
| *Pleuragramma antarctica* | mtDNA | 0[†] | 7000 WS-ROS[$] | Zane *et al.* 2006 |
| *Trematomus bernacchii* | mtDNA & nDNA | 4900 ROS-SSh | 1600 DDU-ROS | Janko *et al.* 2007 |
| *Trematomus newnesi* | mtDNA & nDNA | - | 4900 ROS-SSh | Janko *et al.* 2007 |
| *Chaenocephalus aceratus* | STR | - | 300 SSh-EI | Papetti *et al.* 2007 |
| *Dissostichus mawsoni* | mtDNA & nDNA | 1000 COS-CPS | 6200 SSh-CPS | Kuhn & Gaffney 2008 |
| *Chaenocephalus aceratus* | mtDNA | - | 3900 AP-BI | Jones *et al.* 2008 |
| *Lepidonotothen squamifrons* | mtDNA | 3400 EI-BI | 1800 SSa-BI | Jones *et al.* 2008 |
| *Notothenia coriiceps* | mtDNA | - | 3900 SSh-BI | Jones *et al.* 2008 |
| *Lepidonotothen larseni* | mtDNA | - | 3400 EI-BI | Jones *et al.* 2008 |
| *Gobionotothen gibberifrons* | mtDNA & STR | - | 1900 SSh-SSa | Matschiner *et al.* 2009 |
| *Chaenocephalus aceratus* | STR | 500 EI-SO | 300 SSh-EI | Papetti *et al.* 2009 |
| *Trematomus nicolai* | mtDNA | - | 5000 ROS-AP | Kuhn *et al.* 2009 |
| *Eleginops maclovinus* | mtDNA | ?[¶] | ?[¶] South American shelf | Ceballos *et al.* 2012 |
| *Trematomus newnesi* | mtDNA & STR | 10000 SO-DDU | 800 SO-SSh | Van de Putte *et al.* 2012b |
| *Trematomus hansoni* | mtDNA & STR | 800 SO-SG | 600 JI-SO | Van de Putte *et al.* 2012b |
| *Trematomus bernacchii* | mtDNA & STR | 1500 MAS-DDU | 2100 DDU-ROS | Van de Putte *et al.* 2012b |
| *Gobionotothen gibberifrons* | mtDNA & STR | - | 500 EI-SO | Damerau *et al.* 2012 |
| *Lepidonotothen squamifrons* | mtDNA & STR | - | 500 EI-SO | Damerau *et al.* 2012 |
| *Trematomus eulepidotus* | mtDNA & STR | - | 500 EI-SO | Damerau *et al.* 2012 |
| *Trematomus newnesi* | mtDNA & STR | - | 500 EI-SO | Damerau *et al.* 2012 |
| *Chaenocephalus aceratus* | mtDNA & STR | 500 EI-SO | - | Damerau *et al.* 2012 |

| | | | | |
|---|---|---|---|---|
| *Champsocephalus gunnari* | mtDNA & STR | 500 EI-SO | - | Damerau *et al.* 2012 |
| *Chionodraco rastrospinosus* | mtDNA & STR | - | 500 EI-SO | Damerau *et al.* 2012 |
| *Chionodraco rastrospinosus* | STR | - | 800 SSh-SO | Papetti *et al.* 2012 |
| *Chionodraco hamatus* | STR | - | 10000 WS-ROS | Agostini *et al.* 2013 |
| *Chionodraco myersi* | STR | - | 10000 WS-ROS | Agostini *et al.* 2013 |
| *Chaenocephalus aceratus* | mtDNA & STR | 500 EI-SO | - | Damerau *et al.* 2014 |
| *Champsocephalus gunnari* | mtDNA & STR | 500 EI-SO | - | Damerau *et al.* 2014 |
| *Dissostichus mawsoni* | nDNA | - | 10000 BI-ROS | Mugue *et al.* 2014 |
| *Notothenia rossii* | STR | - | 1500 SSh-SG | Young *et al.* 2015 |
| *Champsocephalus gunnari* | STR | 350 SR-SG | 500 AP-EI | Young *et al.* 2015 |
| *Pleuragramma antarctica* | STR | 1000 AP-JI | 1500 AP-WS | Agostini *et al.* 2015 |
| *Dissostichus eleginoides* | mtDNA & nDNA | 800 SG-SSa (mtDNA) 5000 SSa-CI (nDNA) | 400 KI-HM (mtDNA & nDNA) | Toomey *et al.* 2016 |
| *Lepidonotothen squamifrons* | mtDNA & nDNA | - | 10000 BI-ROS | Miya *et al.* 2016 |
| *Lepidonotothen larseni* | mtDNA & nDNA | - | 6000 AP-EM | Miya *et al.* 2016 |
| *Lepidonotothen nudifrons* | mtDNA & nDNA | 800 SG-SSa‡ | 800 SSh-SO | Dornburg *et al.* 2016 |
| *Pleuragramma antarctica* | STR | 1500 AP-SO | 10000 WS-ROS | Caccavo *et al.* 2018 |
| *Dissostichus eleginoides* | STR | 1200 FI-SG | 6500 South American shelf | Canales-Aguirre *et al.* 2018 |
| *Lepidonotothen larseni* | mtDNA | - | 200 AP-SSh | Deli Antoni *et al.* 2019 |
| *Lepidonotothen nudifrons* | mtDNA | 200 AP-SSh | - | Deli Antoni *et al.* 2019 |
| *Notothenia coriiceps* | RRS | 500 EI-SO | 800 SSh-SO | Christiansen *et al.*, Ch. 4 |
| *Notothenia rossii* | RRS | - | 7000 SSh-KI | Christiansen *et al.*, Ch. 5 |

* Non-significant differentiation within the Weddell Sea.
$ Counter-clockwise distance.
† Temporal differentiation.
¶ Details not provided in publication.
‡ *L. nudifrons* and *L.* cf. *nudifrons* from AP vs. SG and SSa were determined as different, cryptic species.

### 2.3 Synthesis and perspectives

It appears that no straightforward relationship exists between genetic structure and a single explanatory factor in Southern Ocean notothenioids. Rather the combination of several life history parameters, the demographic history and the contemporary environmental setting, especially in terms of oceanographic circulation, are determining the level of gene flow in a complex interplay (Fig. 6.2).

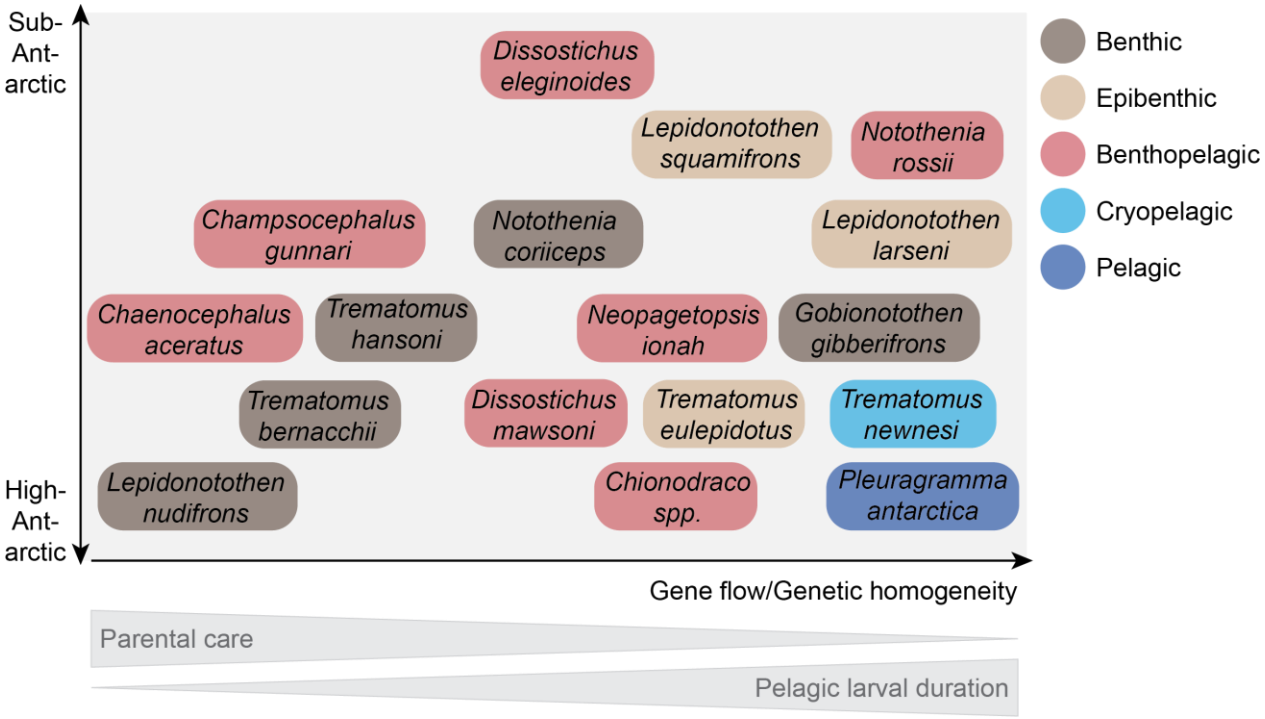



**Fig. 6.2.** A synthesis of the current state of knowledge of notothenioid population genetics. Investigated species are shown with respect to their degree of genetic differentiation, latitudinal distribution and life history traits such as habitat, parental care and pelagic larval duration (see also Table 6.3). Genetic differentiation cannot be explained solely by life history strategy, instead other factors such as demographic history and oceanographic setting (both not shown here) are influencing the amount of gene flow. Note that the arrangement of species within this continuum of population divergence is tentative and not meant as an absolute classification.

As notothenioids are relevant models for various research questions, there are strong incentives to decipher their genomes (Bargelloni *et al.* 2019; Chen *et al.* 2019). Recently, several notothenioid draft genomes and transcriptomes have become available (Table 6.4). These genomic resources will be invaluable to further our understanding of population divergence of Southern Ocean organisms in an ecological, evolutionary and conservation context. As Dornburg *et al.* (2016) put it, "this is an exciting time for polar biology". Future studies should aim at increasing the species diversity of population genetic studies – the studies listed in Table 6.3 have often focused on the same species repeatedly – and large-scale sampling. The latter is a

typical problem of Antarctic research (Griffiths *et al.* 2011; Guillaumot *et al.* 2018a). Many population genetic studies to date (including Chapters 4 & 5) have limited spatial coverage with respect to the target species' distribution. This is unfortunate as it prevents fully resolved and comparable results. Latest advances in database management, international collaboration and openly available data should help alleviate this issue (e.g. Deck *et al.* 2017). In addition, available draft genomes now facilitate the advanced study of adaptation in Southern Ocean fishes, which will likely be a major focus of future research (Convey & Peck 2019). Good reference genomes will also increase the opportunity to use old samples for historic comparisons (ancient DNA; see de Bruyn *et al.* 2014; Schmid *et al.* 2016; Star *et al.* 2017; Heindler *et al.* 2018). Managing and conserving the Antarctic fauna in times of rapid environmental change and increasing anthropogenic disturbance and pollution is of utmost importance. Ultimately, active interference with Southern Ocean animals (e.g. translocation, genetic rescue, targeted breeding, etc.) in order to save (some) biodiversity from extinction might become a subject for debate (Whiteley *et al.* 2014; Hoffmann *et al.* 2017; Razgour *et al.* 2018). Such endeavors would depend upon good knowledge regarding adaptation and historic and contemporary baselines of diversity. The data presented in this thesis, especially Chapters 4 & 5, represent a first step in that direction.

**Table 6.4.** Overview of sequenced (draft) genomes and transcriptomes of Southern Ocean fish.

| Species | Genome | BUSCO completeness | N50 contig length | Linkage map | Transcriptomic data | Reference |
|---|---|---|---|---|---|---|
| *Notothenia coriiceps* | 0.64 Gbp | 80.6 %* | 11,600 | 11 LG$ | Muscle | Shin *et al.* 2014 |
| *Parachaenichthys charcoti* | 0.81 Gbp | 86.3 % | 6,145 | | Muscle | Ahn *et al.* 2017 |
| *Chaenocephalus aceratus* | 1.1 Gbp | 87.0 %* | 1,500,626 | 24 LG | Muscle & other tissues | Kim *et al.* 2019 |
| *Eleginops maclovinus* | 0.73 Gbp | 95.0 %* | 10,900 | | Muscle & other tissues | Chen *et al.* 2019 |
| *Dissostichus mawsoni* | 0.84 Gbp | 97.2 %* | 23,100 | | Muscle | Chen *et al.* 2019 |
| *Chionodraco hamatus* | 0.83 Gbp | 66.8 %* | 2,706 | | - | Bargelloni *et al.* 2019 |
| *Chionodraco myersi* | 1.12 Gbp | 86.8 %* | 45,409 | | Muscle & other tissues | Bargelloni *et al.* 2019 |

* based on the analysis presented in Bargelloni *et al.* 2019
$ Amores *et al.* 2017

### 3. Conservation prospects and research perspectives

The Southern Ocean is currently in a transition phase regarding its resource management. While some 50 years ago many fish species were heavily exploited, the implementation of CCAMLR in 1982 has ended severe overfishing and introduced fishing bans, conservation measures (CMs) and regular stock assessments (Kock 1992). Fishing activities in the Southern Ocean have shifted in response to the CCAMLR measures, the collapse of the Soviet Union, and changed global demands. In the past 20-30 years the fishery focus has been on toothfish and krill. CCAMLR continued to manage the fisheries in an ecosystem based approach (Kock *et al.* 2007), but illegal fishing activities and uncertainties about ecosystem impacts of fishing have sparked concern among conservation-oriented scientists and non-governmental organizations (NGOs) (Stokstad 2010; Österblom & Sumaila 2011; Abrams 2014; Abrams *et al.* 2016; Brooks *et al.* 2016). Illegal, unreported and unregulated fisheries of toothfish could be minimized in the twenty-first century, not least due to collaboration between different stakeholders, including the legal fishing operators (Österblom & Bodin 2012). In addition, CCAMLR has recently committed itself to the establishment of a representative network of marine protected areas (MPAs) throughout the Southern Ocean (CCAMLR 2012). The idea of MPA networks, that increase the ecological benefits of individual MPAs through connectivity and good representation of biodiversity, is not unique to the Southern Ocean, but has been formulated on a global level (Costello *et al.* 2010). Effective MPA networks consider ecological criteria, but also the socio-economic background, and foresee integrated and adaptive management practices. However, progress towards the realization of such plans is slow, both in CCAMLR and elsewhere. Limited knowledge on connectivity and spatial genetic diversity and adaptation patterns is a contributing factor that hinders progress in implementing MPA networks.

#### 3.1 Current MPAs and their relevance for fishes

The South Orkney Islands Southern Shelf (SOISS) MPA, proposed by the United Kingdom, has been in effect since 2009 and presents a landmark in marine conservation as the first MPA established in international waters (Trathan & Grant 2020). And yet, the current MPA extent fails to protect important biodiversity hotspots, penguin foraging grounds and candidate fish spawning areas. However, existing CMs guarantee a certain amount of protection (no bottom trawling; long line fishing only in waters deeper than 550 m) for dermersal fish around the South Orkney Islands. The adult sub-population of *Notothenia coriiceps* that occurs here should be relatively well protected through these measures. As these fish are slightly differentiated from

(some) populations at the tip of the Antarctic Peninsula (Chapter 4), they may harbour unique adaptive genetic variation that is likely linked to contrasting climate change effects. While indicative at the moment, these genetic variants should be further studied, as they might represent crucial diversity that can help the species cope with climate change. It is vital to conserve such variation, especially considering the rapid warming of other *N. coriiceps* habitats such as along the Antarctic Peninsula (IPCC 2019). In that case, the respective population diversity must be protected to reach CCAMLR's ecosystem conservation objective. Standing genetic variation is in fact the most important form of genetic diversity, that enables species to adapt to changing conditions (Bernatchez 2016). Therefore, preserving both adaptive and neutral genetic diversity is likely important for long-term conservation goals

A similar argument is valid for *Notothenia rossii*, although we have currently little evidence for the presence of adaptive variation (Chapter 5). However, our results demonstrate that for *N. rossii* the South Orkney Islands are an important stepping stone to facilitate connectivity between the Antarctic Peninsula and the South Sandwich Islands, and in turn for the Atlantic and Indian Ocean sector connections (Chapter 5). Currently, the *N. rossii* population appears well connected. Targeting the South Orkney stepping stone by fisheries, like in the 1980s (Kock *et al.* 1992), may disrupt the present-day connectivity, with unknown effects for the entire populations. Possibly, neutral genetic differentiation between *N. rossii* of the Scotia Arc and the Kerguelen Plateau would increase over time. Such a scenario does not necessarily entail negative consequences. However, as the *N. rossii* stock at South Orkney likely has always been comparatively small (< 20.000 t; Kock *et al.* 1992), but is an important link, it should be managed with great precaution. In addition, it is unclear if the historically exploited spawning biomass has fully recovered. Biomass estimates have increased from 412 t in 1991 to 3278 t in 1999 (Kock & Jones 2005), but a recent survery estimated the biomass of *N. rossii* around the South Orkneys at only 161 t in 2018 (Arana *et al.* 2020). Therefore, in a truly precautionary approach fishing bans should remain in place, at least until the stock structure, recovery process and future connectivity prognoses are better understood. Forecast connectivity modelling using climate projections can be used to this end (Young *et al.* 2018) and should be combined with forecast SDM and genetic projection models.

In addition, it is important to not only consider protection of the habitat for adult fish, but also for early life stages. The distribution, behavior and general ecology of larvae and pelagic blue

phase fingerlings of both *N. coriiceps* and *N. rossii* remain poorly understood. Fish larvae and juveniles are taken as bycatch in the krill fishery (e.g. Everson *et al.* 1992), but it is unclear whether or how much *Notothenia* spp. are affected, also because fish larvae from bycatch are often not identified to species level. DNA barcoding of individual bycatch specimens or DNA metabarcoding of bulk samples from krill fishing vessels could be used in the future to improve our understanding of fish bycatch in the krill exploitation. Finally, krill fisheries may also exert pressure on *Notothenia* stocks through the removal of one of their preferred food sources (see Kock & Jones 2005 and Heindler *et al.* 2019 for an overview of the species' diet). Such indirect ecosystem effects have not been considered much for fish, in contrast to Antarctic top predators, such as whales, seals and seabirds, where the potential interaction with krill fisheries is a topic of active research and debate (Barbraud *et al.* 2012; Hinke *et al.* 2017; Watters *et al.* 2020). A combination of traditional and molecular trophic ecology and food web modelling should be helpful to illuminate possible ecosystem effects of krill fisheries.

Next to the SOISS MPA, one other MPA has been declared by CCAMLR in the Ross Sea following years of negotiations (Brooks *et al.* 2019). The Ross Sea MPA will bethe largest international MPA on Earth and shall protect wide areas of the Ross Sea shelf, although zones for exploratory fishing and krill research are included as well. As CCAMLR parties currently have not been able to reach consensus on a research and monitoring plan, the MPA has not yet been fully implemented. In addition, the Ross Sea MPA is declared for limited time, i.e. must be renegotiated after 2052. This fact is a precedential case in marine conservation and a concession made towards the countries whose primary concern is not to seriously restrict fishing. Similarly, the rapid success of the SOISS MPA establishment was likely due to the absence of conflicting objectives, as the currently designated SOISS MPA area does not overlap with areas in use by (krill) fisheries (Trathan & Grant 2020). As no samples from the Ross Sea were available for Chapter 4 and 5, no inferences about the importance of the Ross Sea MPA for *N. coriiceps* and *N. rossii* can be made. Regular biological monitoring inside and outside the MPA areas should help determine the relevance and impact of this conservation measure.

### 3.2 Research in support of future MPAs

The scientific basis for the development of Southern Ocean MPAs is developed through bioregionalisation approaches that aim to identify the spatial distribution of marine ecosystems (Hill *et al.* 2017). Challenges in such approaches are the varied nature of benthic and pelagic

ecosystems and the large number of Antarctic top predators that transverse management boundaries. Systematic conservation planning is endorsed by CCAMLR in an attempt to adequately identify areas for protection. The SOISS MPA is the direct result of systematic conservation planning (see Trathan & Grant 2020 and references therein). Several additional MPA proposals have been submitted to CCAMLR (Fig. 6.3), but none have been endorsed so far. In 2011, Australia, France and the European Union submitted an MPA proposal for East Antarctica for the first time. The European Union also introduced a proposal for a Weddell Sea MPA in 2016 and submitted a revised version in 2018. Within the western Antarctic Peninsula and Scotia Sea an MPA proposal has been submitted by Argentina and Chile in 2018. All these MPA proposals are still being debated, refined, resubmitted, and evaluated and have not been implemented so far. The debate about protection or rational use (of fish and krill resources) has been raging in CCAMLR for the past decade and is likely to continue in the near future (Jacquet *et al.* 2016; Liu & Brooks 2018; Brooks *et al.* 2019; Hofman 2019; Sylvester & Brooks 2019). It is a debate with large political implications and also reflects a cultural divide. CCAMLR itself is in a period of institutional growth and learning to meet the challenges of juggling opposing objectives in times of rapid environmental change and increasing pressures through tourism, research and fisheries. To overcome divisions between CCAMLR members it has been argued that MPAs should be more clearly linked to counter measurable threats instead of being declared on a precautionary basis. However, from a biological standpoint an element of precaution is always advisable when threats are unknown. Thus, defining the objectives of an MPA is an important research area in itself. The proposal for a western Antarctic MPA would certainly be beneficial to counter threats of overfishing for *N. rossii* – a species that clearly has been dramatically affected by historic overharvest (Kock 1992; Marschoff *et al.* 2012; Barrera-Oro *et al.* 2017).

Multidisciplinary studies as presented here (especially Chapter 5) help to identify research and conservation priorities. After all, the Southern Ocean is still a severely data-limited study system. Long-term programs, open data, and collaboration can provide useful large data sets to synthesize spatial information as recently shown with top predator tracking data (Hindell *et al.* 2020). Several important areas of ecological significance (AES) could be determined across the entire Southern Ocean, that match well with planned MPAs (39 % coverage; Hindell *et al.* 2020). This study also highlights the value of extrapolation through advanced modelling techniques. Much like the global deep sea, the Southern Ocean is too vast to sample all areas intensively. Instead, it is worthwhile to focus on available data and samples and use these to determine what

should be investigated in more detail. The research presented in this thesis can be understood in that context. The mesopelagic fish communities are a central element in the Antarctic food web next to krill (McCormack *et al.* 2019). The samples used in Chapter 2 show that the occurrence of myctophids corresponds well to the sub-Antarctic AES identified by Hindell *et al.* (2020). Future research should focus on investigating the community dynamics in these particular areas and how they may react to future change. In addition, phylogenomic approaches should be used on *Gymnoscopelus bolini*, *Lampanyctus achirus* and *Bathylagus* spp. as these were identified to possibly comprise cryptic variation (Chapter 2). The method optimization in Chapter 3 provides the tools to tackle species-specific questions about population genetics and adaptation in a wide array of Southern Ocean taxa. Multi-taxon comparisons of drivers of connectivity and genetic structure are of primary interest for the ongoing MPA planning process. Whenever possible, population genomics should be complemented with modelling techniques and used on samples from proposed MPA areas and/or AES. Our knowledge of at least two genetic subpopulation of *N. coriiceps* should be extended in the same spirit as soon as more samples become available. Furthermore, targeted future sequencing projects may focus on genomic regions that are possibly relevant for adaptation to a changing climate as indicated in Chapter 4. Finally, in Chapter 5 some very specific future research directions are apparent. The fish fauna around Bouvet Island, including adult demersal species and ichthyoplankton assemblages should be targeted with standardized and ideally seasonally repeated research sampling. This information is critical to make predictions about the impact of krill fisheries on the fish species in this area. Additional insights about population dynamics and life history traits of *N. rossii* across its distribution range and throughout the year are needed to better understand the long recovery process of this species and avoid similar situations in the future.

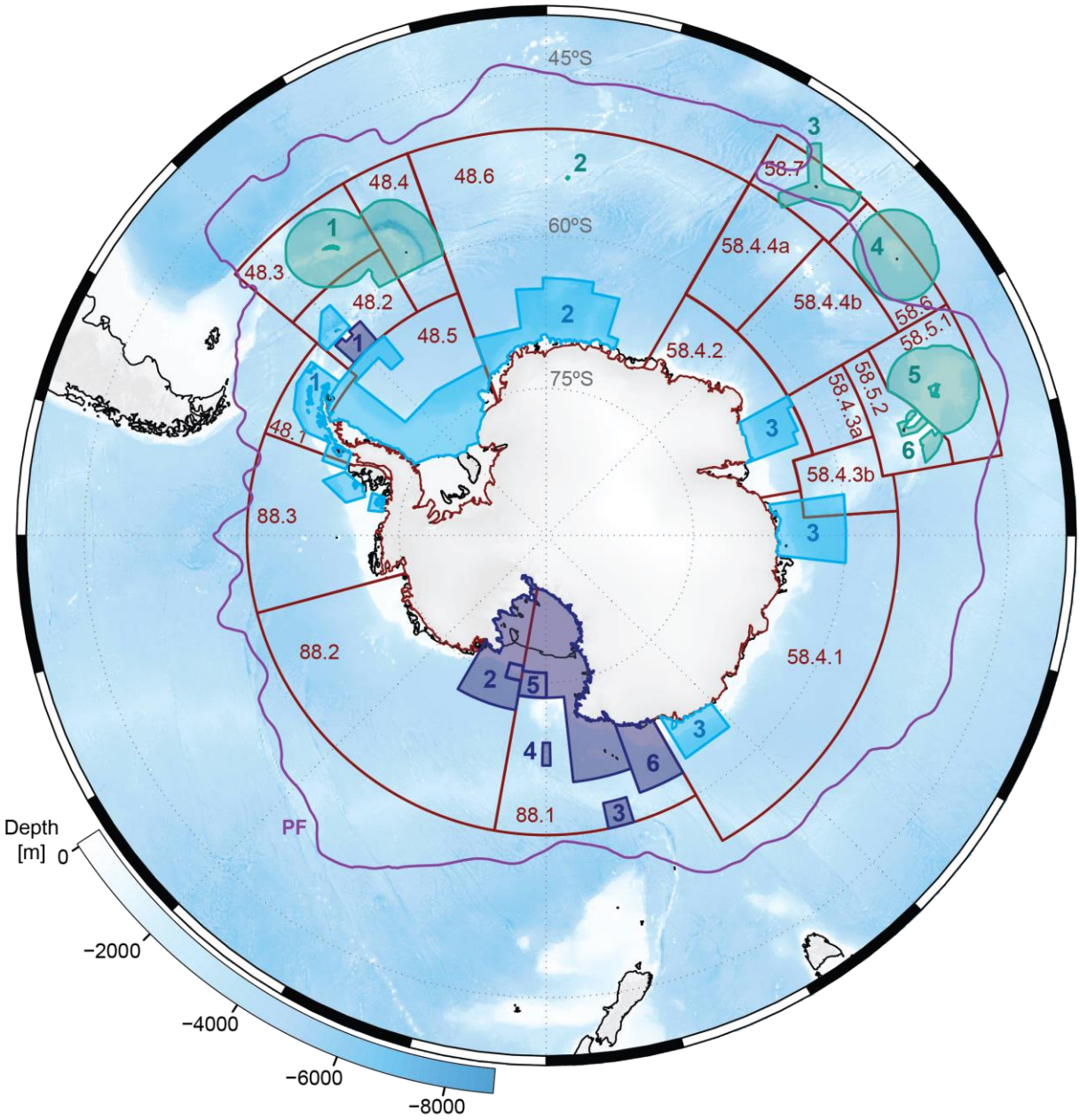



**Fig. 6.3.** Current and proposed Marine Protected Areas (MPAs) in the Southern Ocean. The convention area of the Commission for the Conservation of Antarctic Marine Living Resources (CCAMRL; red boxes) uses distinct management zones for the Atlantic (48.1-6), Indian (58.4-7), and Pacific (88.1-3) sectors of the Southern Ocean. Currently declared MPAs are shown in dark blue, proposed MPAs in light blue, and other protection areas in dark green. The Antarctic Polar Front (PF) is shown in magenta. Background shading (blue-white) indicates depth. Map produced using R package SOmap v0.5.0 (Maschette *et al.* 2019) with data from CCAMLR.

The holy grail of fisheries and conservation in the Southern Ocean is currently the Antarctic toothfish (*Dissostichus mawsoni*), where opinions on if and how this fishery should be allowed strongly diverge (Abrams 2014; Hanchet *et al.* 2015b; Abrams *et al.* 2016; Hofman 2019). An in-depth review and examination of this issue is beyond the scope of this thesis, but it should be noted that genetic diversity in terms of evolutionary potential and adaptive capacity is clearly understudied in *D. mawsoni* at the moment (see also Table 5.3). With this in mind, it seems that CCAMLR is neglecting its responsibility to employ a precautionary approach also with regard to Antarctic climate change. Given the rapid rates of warming in parts of the Southern Ocean it seems questionable whether Antarctic toothfish can adapt to changing environmental conditions, if they are concomitantly facing considerable fishery pressure. Toothfish are central parts of the food web and therefore ecologically important, especially in the Ross Sea (Ainley & Siniff 2009; Pinkerton *et al.* 2016; Salas *et al.* 2016). As with the here studied *Notothenia* species it is unclear how well the global toothfish population will cope with climate change. An integrative, multi-method approach as applied in Chapter 5 would be beneficial to at least resolve the circumpolar connectivity patterns among toothfish habitats. Notwithstanding, toothfish adults can perform considerable migrations, which complicates the assessment of connectivity in this species (Hanchet *et al.* 2015a). Examples from other marine fishes clearly show that fisheries can seriously undermine the genetic diversity of fish populations that is needed to adapt to rapidly changing environmental conditions (Pinsky & Palumbi 2014; Walsworth *et al.* 2019; Rogers *et al.* 2020). A truly precautionary approach would therefore take this information into consideration and protect Antarctic fish diversity before consequences such as further population declines become evident

In conclusion, the unique Southern Ocean remains threatened through increasing impacts from scientists, tourists, global pollution, expanding fisheries, and climate change. While many of these problems are not exclusive to the Southern Ocean and must be tackled through global actions, it is imperative to protect representative areas of the Southern Ocean, not least to study the impacts of these stressors (Trathan & Grant 2020).

#### 4. Conclusion

This thesis increased our scientific understanding of fish diversity in the Southern Ocean in space and time and from species to populations. Alongside this endeavor a foray into current genomic methods applied to non-model species was undertaken. The sequencing revolution

offers unprecedented opportunities to create vast amounts of data and potentially answer long-standing research questions, while also raising a manifold of new questions. The results presented in Chapters 3, 4, and 5 offer an insight into the complexity of adapting modern molecular methods, in particular reduced representation sequencing, for non-model species with limited genomic background knowledge. These examples may provide guidance for future applications, that will also benefit from the increasing availability of reference genomes for Antarctic fishes. In addition, the database of genetic information in terms of barcodes and georeferenced RRS data was substantially extended in Chapters 2, 4 and 5. The COI barcode library of Southern Ocean fishes still needs further addition, in particular of rare, non-notothenioid species. Such a reference database is critical to fully exploit the large potential of metabarcoding studies for biodiversity assessments and trophic ecology. The RRS data created for *Notothenia coriiceps* and its congener *N. rossii* in Chapters 4 and 5 are the first of their kind for Antarctic fishes and thus merely a start into deciphering the population genomic makeup of Southern Ocean fishes. Given the tremendous advances made in temperate species such as Atlantic cod (for instance with regard to the role of genomic architecture ), it can be expected that much will be learnt from further studies that follow this path. Ultimately, it is hoped that the findings presented here will be considered in ongoing conservation plans and contribute to better management and protection of the unique ichthyofauna of the Southern Ocean. Additional, non-traditional scientific exercises that go beyond the research itself and instead focus on science communication to raise awareness may help realize conservations goals (see Appendix III). Perhaps in the not too distant future the Southern Ocean will become indeed a biodiversity sanctuary – it would be about time.

## BIBLIOGRAPHY


Abrams PA (2014) How precautionary is the policy governing the Ross Sea Antarctic toothfish (*Dissostichus mawsoni*) fishery? *Antarctic Science*, **26**, 3–13.

Abrams PA, Ainley DG, Blight LK *et al.* (2016) Necessary elements of precautionary management: implications for the Antarctic toothfish. *Fish and Fisheries*, **17**, 1152-1174.

Addamo AM, Vertino A, Stolarski J *et al.* (2016) Merging scleractinian genera : the overwhelming genetic similarity between solitary *Desmophyllum* and colonial *Lophelia*. *BMC Evolutionary Biology*, **16**, 108.

Adenyo C, Ogden R, Kayang B *et al.* (2016) Genome-wide DNA markers to support genetic management for domestication and commercial production in a large rodent, the Ghanaian grasscutter (*Thryonomys swinderianus*). *Animal Genetics*, **48**, 113-115.

Agnarsson I, Kuntner M (2007) Taxonomy in a changing world: Seeking solutions for a science in crisis. *Systematic Biology*, **56**, 531–539.

Agostini C, Papetti C, Patarnello T *et al.* (2013) Putative selected markers in the *Chionodraco* genus detected by interspecific outlier tests. *Polar Biology*, **36**, 1509–1518.

Agostini C, Patarnello T, Ashford JR *et al.* (2015) Genetic differentiation in the ice- dependent fish *Pleuragramma antarctica* along the Antarctic Peninsula. *Journal of Biogeography*, **42**, 1103–1113.

Ahn D-H, Shin SC, Kim B-M *et al.* (2017) Draft genome of the Antarctic dragonfish, *Parachaenichthys charcoti*. *GigaScience*, **6**, 1–6.

Ainley DG, Blight LK (2009) Ecological repercussions of historical fish extraction from the Southern Ocean. *Fish and Fisheries*, **10**, 13–38.

Ainley DG, Crockett EL, Eastman JT *et al.* (2017) How overfishing a large piscine mesopredator explains growth in Ross Sea penguin populations: A framework to better understand impacts of a controversial fishery. *Ecological Modelling*, **349**, 69–75.

Ainley DG, Eastman JT, Brooks CM (2016) Comments on "The Antarctic toothfish (*Dissostichus mawsoni*): biology, ecology, and life history in the Ross Sea region," by S. Hanchet et al. *Hydrobiologia*, **771**, 1-7.

Ainley DG, Pauly D (2013) Fishing down the food web of the Antarctic continental shelf and slope. *Polar Record*, **50**, 92–107.

Ainley DG, Siniff DB (2009) The importance of Antarctic toothfish as prey of Weddell seals in the Ross Sea. *Antarctic Science*, **21**, 317-327.

Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung, Bremerhaven, Germany (2017) Polar Research and Supply Vessel POLARSTERN Operated by the Alfred-Wegener-Institute. *Journal of large scale research facilities*, **3**, A119.

Ali OA, O'Rourke SM, Amish SJ *et al.* (2016) RAD Capture (Rapture): Flexible and efficient sequence-based genotyping. *Genetics*, **202**, 389–400.

Allcock AL, Barratt I, Eléaume M *et al.* (2011) Cryptic speciation and the circumpolarity debate: A case

study on endemic Southern Ocean octopuses using the COI barcode of life. *Deep Sea Research Part II: Topical Studies in Oceanography*, **58**, 242–249.

Allcock AL, Strugnell JM (2012) Southern Ocean diversity: new paradigms from molecular ecology. *Trends in Ecology & Evolution*, **27**, 520–8.

Allendorf FW, England PR, Luikart G *et al.* (2008). Genetic effects of harvest on wild animal populations. *Trends in Ecology & Evolution*, **23**, 327–337.

Allendorf FW, Hohenlohe P a, Luikart G (2010) Genomics and the future of conservation genetics. *Nature Reviews Genetics*, **11**, 697–709.

Altschul S, Madden T, Schaffer A *et al.* (1997) Gapped blast and psi-blast: a new generation of protein database search programs. *Nucleic Acids Research*, **25**, 3389–3402.

Altshuler D, Pollara VJ, Cowles CR, Lander ES (2000) An SNP map of the human genome generated by reduced representation shotgun sequencing. *Nature*, **407**, 513–516.

Amores A, Wilson CA, Allard CAH, Detrich HW, Postlethwait JH (2017) Cold Fusion: Massive Karyotype Evolution in the Antarctic Bullhead Notothen *Notothenia coriiceps*. *G3 Genes Genomes Genetics*, **7**, 2195–2207.

Amsler MO, Eastman JT, Smith KE *et al.* (2016) Zonation of demersal fishes off Anvers Island, western Antarctic Peninsula. *Antarctic Science*, **28**, 44–50.

Anderson JB, Shipp SS, Lowe AL, Wellner JS, Mosola AB (2002) The Antarctic Ice Sheet during the Last Glacial Maximum and its subsequent retreat history: a review. *Quaternary Science Reviews*, **21**, 49–70.

Andrews S (2010) FastQC: a quality control tool for high throughput sequencing data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc

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

Andrews KR, Luikart G (2014) Recent novel approaches for population genomics data analysis. *Molecular Ecology*, **23**, 1661–1667.

Andrews KR, Paul A, Miller MR *et al.* (2014) Trade-offs and utility of alternative RADseq methods : Reply to Puritz et al . 2014. *Molecular Ecology*, **23**, 5943–5946.

Appleyard SA, Ward RD, Williams R (2002) Population structure of the Patagonian toothfish around Heard, McDonald and Macquarie Islands. *Antarctic Science*, **14**, 364–373.

Appleyard SA, Williams R, Ward RD (2004) Population genetic structure of Patagonian toothfish in the West Indian Ocean sector of the Southern Ocean. *CCAMLR Science*, **11**, 21–32.

Apweiler R, Bairoch A, Wu CH *et al.* (2004) UniProt : the Universal Protein knowledgebase. *Nucleic Acids Research*, **32**.

Arana, PM, Jones CD, Alegría NA *et al.* (2020). Antarctic demersal finfish around the Elephant and the South Orkney islands: distribution , abundance and biological characteristics. *Latin American Journal of Aquatic Research*, **48**, in press.Arango CP, Soler-Membrives A, Miller KJ (2011) Genetic differentiation in the circum—Antarctic sea spider *Nymphon australe* (Pycnogonida;

Nymphonidae). *Deep Sea Research Part II: Topical Studies in Oceanography*, **58**, 212–219.

Ardoin NM, Heimlich JE (2013) Views from the field: Conservation educators and practitioners perceptions of education as a strategy for achieving conservation outcomes. *Journal of Environmental Education*, **44**, 97–115.

Arntz WE, Thatje S, Linse K *et al.* (2006) Missing link in the Southern Ocean: Sampling the marine benthic fauna of remote Bouvet Island. *Polar Biology*, **29**, 83–96.

Aronson RB, Thatje S, Mcclintock JB, Hughes KA (2011) Anthropogenic impacts on marine ecosystems in Antarctica. *Annals of the New York Academy of Sciences*, **1223**, 82–107.

Ashford J, Dinniman M, Brooks C (2017) Physical-biological interactions influencing large toothfish over the Ross Sea shelf. *Antarctic Science*, **29**, 487–494.

Auvinet J, Graça P, Belkadi L *et al.* (2018) Mobilization of retrotransposons as a cause of chromosomal diversification and rapid speciation: the case for the Antarctic teleost genus *Trematomus*. *BMC Genomics*, **19**, 339.

Baalsrud HT, Tørresen OK, Solbakken MH *et al.* (2017) De Novo Gene Evolution of Antifreeze Glycoproteins in Codfishes Revealed by Whole Genome Sequence Data. *Molecular Biology and Evolution*, **35**, 593–606.

Baetscher DS, Clemento AJ, Ng TC, Anderson EC, Garza JC (2018) Microhaplotypes provide increased power from short-read DNA sequences for relationship inference. *Molecular Ecology Resources*, **18**, 296–305.

Baeza JA, Behringer DC (2017) Integrative taxonomy of the ornamental 'peppermint' shrimp public market and population genetics of *Lysmata boggessi* , the most heavily traded species worldwide. *PeerJ*, **5**, e3786.

Baird NA, Etter PD, Atwood TS *et al.* (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One*, **3**, e3376.

Baker CS, Clapham PJ (2004) Modelling the past and future of whales and whaling. *Trends in Ecology and Evolution*, **19**, 365–371.

Bakker J, Wangensteen OS, Chapman DD *et al.* (2017) Environmental DNA reveals tropical shark diversity and abundance in contrasting levels of anthropogenic impact. *Scientific Reports*, **7**, 16886.

Balushkin A V (2000) Morphology, classification and evolution of notothenioid fishes of the Southern Ocean. *Journal of Ichthyology*, **40**, S74–S109.

Balushkin A V. (2013) A new species of *Pogonophryne* (Perciformes: Notothenioidei: Artedidraconidae) from the deep Ross Sea, Antarctica. *Proceedings of the Zoological Institute RAS*, **317**, 119–124.

Balushkin A V., Moganova M V. (2017) New species of pelagic eelpouts of the genus *Melanostigma* (Zoarcidae) from Western Antarctica. *Journal of Ichthyology*, **57**, 171–176.

Balushkin A V., Moganova M V. (2018) *Melanostigma kharini* sp. nov. (Zoarcidae), a new species of pelagic eelpout from the Ridge of Hercules (Pacific-Antarctic Rise). *Journal of Ichthyology*, **58**, 139–147.

Balushkin A V., Spodareva V V. (2013a) *Pogonophryne skorai* sp. n. (Perciformes: Artedidraconidae), a

new species of toadlike plunderfish from the Bransfield Strait and coastal waters of the south Shetland Islands, Antarctica. *Russian Journal of Marine Biology*, **39**, 190–196.

Balushkin A V, Spodareva V V (2013b) Notothenioidei; Perciformes – The deep-water species of Antarctic plunderfishes from the Ross Sea (Southern Ocean). *Proceedings of the Zoological Institute RAS*, **1937**, 275–281.

Balushkin A V., Spodareva V V. (2015) New species of the toad plunderfish of the "albipinna" group, genus *Pogonophryne* (Artedidraconidae) from the Ross Sea (Antarctica). *Journal of Ichthyology*, **55**, 757–764.

Balushkin A V., Zhukov MY (2016) Polytypy of the genus *Zanclorhynchus* (Zanclorhynchinae: Congiopodidae): Description of the new species *Z. chereshnevi* sp. n. from the Indian sector of the Southern Ocean. *Journal of Ichthyology*, **56**, 791–798.

Bandelt H-J, Forster P, Röhl A (1999) Median-Joining Networks for Inferring Intraspecific phylogenies. *Molecular Biology*, **16**, 37–48.

Bar-On YM, Phillips R, Milo R (2018) The biomass distribution on Earth. *Proceedings of the National Academy of Sciences*, **115**, 6506–6511.

Barbraud C, Rolland V, Jenouvrier S *et al.* (2012). Effects of climate change and fisheries bycatch on Southern Ocean seabirds: A review. *Marine Ecology Progress Series*, **454**, 285–307.

Bargelloni L, Babbucci M, Ferraresso S *et al.* (2019) Draft genome assembly and transcriptome data of the icefish *Chionodraco myersi* reveal the key role of mitochondria for a life without hemoglobin at subzero temperatures. *Communications Biology*, **2**, 443.

Barham EG (1966) Deep scattering layer migration and composition: observations from a diving saucer. *Science*, **151**, 1399–1403.

Barnes DKA, Hodgson DA, Convey P, Allen CS, Clarke A (2006) Incursion and excursion of Antarctic biota: Past, present and future. *Global Ecology and Biogeography*, **15**, 121–142.

Barrera-Oro E (2002) The role of fish in the Antarctic marine food web: differences between inshore and offshore waters in the southern Scotia Arc and west Antarctic Peninsula. *Antarctic Science*, **14**, 293-309.

Barrera-Oro ER, Casaux RJ (1992) Age estimation for juvenile *Notothenia rossii* from Potter Cove, South Shetland Islands. *Antarctic Science*, **4**, 131–136.

Barrera-Oro ER, Casaux RJ (1996) Validation of age determination in *Notothenia coriiceps*, by means of a tag-recapture experiment at Potter Cove, South Shetland Islands. *Archive of Fishery and Marine Research*, **43**, 205–216.

Barrera-Oro ER, Marschoff ER (2007) Information on the status of fjord *Notothenia rossii*, *Gobionotothen gibberifrons* and *Notothenia coriiceps* in the lower South Shetland Islands, derived from the 2000-2006 monitoring program at Potter Cove. *CCAMLR Science*, **14**, 83–87.

Barrera-Oro E, Marschoff E, Ainley D (2017) Changing status of three notothenioid fish at the South Shetland Islands (1983–2016) after impacts of the 1970–80s commercial fishery. *Polar Biology*, **40**, 2047–2054.

Barrera-Oro E, La Mesa M, Moreira E (2014) Early life history timings in marbled rockcod (*Notothenia rossii*) fingerlings from the South Shetland Islands as revealed by otolith microincrement. *Polar Biology*, **37**, 1099–1109.

Barrera-Oro E, Moreira E, Seefeldt MA, Valli Francione M, Quartino ML (2019) The importance of macroalgae and associated amphipods in the selective benthic feeding of sister rockcod species *Notothenia rossii* and *N. coriiceps* (Nototheniidae) in West Antarctica. *Polar Biology*, **42**, 317–334.

Barrera-Oro ER, Winter DJ (2008) Age composition and feeding ecology of early juvenile *Notothenia rossii* (Pisces, Nototheniidae) at Potter Cove, South Shetland Islands, Antarctica. *Antarctic Science*, **20**, 339–341.

Barrio AM, Lamichhaney S, Fan G, Rafati N (2016) The genetic basis for ecological adaptation of the Atlantic herring revealed by genome sequencing. *eLife*, 1–32.

Barth JMI, Villegas-Ríos D, Freitas C *et al.* (2019) Disentangling structural genomic and behavioural barriers in a sea of connectivity. *Molecular Ecology*, **28**, 1394–1411.

Bayona-Vásquez NJ, Glenn TC, Kieran TJ *et al.* (2019) Adapterama III: Quadruple-indexed, double/triple-enzyme RADseq libraries (2RAD/3RAD). *PeerJ*, **2019**, 1–25.

Beaumont MA, Balding DJ (2004) Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology*, **13**, 969–980.

Belcher A, Saunders RA, Tarling GA (2019) Respiration rates and active carbon flux of mesopelagic fishes (Family Myctophidae) in the Scotia Sea, Southern Ocean. *Marine Ecology Progress Series*, **610**, 149-162.

Benestan L, Moore J-S, Sutherland BJG *et al.* (2017) Sex matters in massive parallel sequencing: Evidence for biases in genetic parameter estimation and investigation of sex determination systems. *Molecular Ecology*, **26**, 6767-6783.

Benestan L, Quinn BK, Maaroufi H *et al.* (2016) Seascape genomics provides evidence for thermal adaptation and current-mediated population structure in American lobster (*Homarus americanus*). *Molecular Ecology*, **25**, 5073–5092.

Berg PR, Jentoft S, Star B *et al.* (2015) Adaptation to low salinity promotes genomic divergence in Atlantic cod (*Gadus morhua* L.). *Genome Biology and Evolution*, **7**, 1644–1663.

Berg PR, Star B, Pampoulie C, Sodeland M, Barth JMI (2016) Three chromosomal rearrangements promote genomic divergence between migratory and stationary ecotypes of Atlantic cod. *Scientific Reports*, **6**, 23246.

Bernal A, Viñas J, Pilar Olivar M (2014) Genetic support for the morphological identification of larvae of Myctophidae, Gonostomatidae, Sternoptychidae and Phosichthyidae (Pisces) from the western Mediterranean. *Scientia Marina*, **78**, 461–471.

Bernardi G, Goswami U (1997) Molecular evidence for cryptic species among the Antarctic fish *Trematomus bernacchii* and *Trematomus hansoni*. *Antarctic Science*, **9**, 381–385.

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

Bernatchez S, Laporte M, Perrier C, Sirois P, Bernatchez L (2016) Investigating genomic and phenotypic parallelism between piscivorous and planktivorous lake trout (*Salvelinus namaycush*) ecotypes by means of RADseq and morphometrics analyses. *Molecular Ecology*, **25**, 4773–4792.

Bernatchez L, Wellenreuther M, Araneda C *et al.* (2017) Harnessing the Power of Genomics to Secure the Future of Seafood. *Trends in Ecology & Evolution*, **32**, 665–680.

Berthelot C, Clarke J, Desvignes T *et al.* (2018) Adaptation of proteins to the cold in Antarctic fish: A role for Methionine? *Genome Biology and Evolution*.

Bertolin ML, Casaux R (2018) Diet overlap among top predators at the South Orkney Islands, Antarctica. *Polar Biology*, **42**, 371-383.

Besnier F, Glover KA (2013) ParallelStructure: a R package to distribute parallel runs of the population genetics program STRUCTURE on multi-core computers. *PLoS ONE*, **8**, 1–5.

Betancur-R R, Wiley EO, Arratia G *et al.* (2017) Phylogenetic classification of bony fishes. *BMC Evolutionary Biology*, **17**, 1–40.

Bi K, Linderoth T, Vanderpool D *et al.* (2013) Unlocking the vault: Next-generation museum population genomics. *Molecular Ecology*, **22**, 6018–6032.

Bilyk KT, DeVries AL (2010) Freezing avoidance of the Antarctic icefishes (Channichthyidae) across thermal gradients in the Southern Ocean. *Polar Biology*, **33**, 203–213.

Bjorkland RD, Pringle CM (2001) Educating our communities and ourselves about conservation of aquatic resources through environmental outreach. *BioScience*, **51**, 279.

Bogan SN, Place SP (2019) Accelerated evolution at chaperone promoters among Antarctic notothenioid fishes. *BMC Evolutionary Biology*, **51**, 279-282.

Bohling J, Small M, Von Bargen J, Louden A, DeHaan P (2019) Comparing inferences derived from microsatellite and RADseq datasets: a case study involving threatened bull trout. *Conservation Genetics*, **20**, 329–342.

Bohonak AJ (1999) Dispersal, gene flow, and population structure. *The Quarterly Review of Biology*, **74**, 21–45.

Bolton PE, West AJ, Cardilini APA *et al.* (2016) Three molecular markers show no evidence of population genetic structure in the Gouldian finch (*Erythrura gouldiae*). *PLoS ONE*, **11**, 1–19.

Bonin CA, Goebel ME, Forcada J, Burton RS, Hoffman JI (2013) Unexpected genetic differentiation between recently recolonized populations of a long-lived and highly vagile marine mammal. *Ecology and Evolution*, **3**, 3701–3712.

Booker TR, Yeaman S, Whitlock MC (2019) Global adaptation confounds the search for local adaptation. *bioRxiv*, 742247.

Borja A (2014) Grand challenges in marine ecosystems ecology. *Frontiers in Marine Science*, **1**.

Bowen BW, Gaither MR, DiBattista JD *et al.* (2016) Comparative phylogeography of the ocean planet. *Proceedings of the National Academy of Sciences of the United States of America*, **113**, 7962–7969.

Boyer TP, Antonov JI, Baranova OK *et al.* (2013) World Ocean Database 2013. NOAA Atlas NESDIS 72, National Oceanographic Data Center (US), Ocean Climate Laboratory, 209 pp. doi:

10.7289/V5NZ85MT

Braasch I, Peterson SM, Desvignes T *et al.* (2014) A new model army: Emerging fish models to study the genomics of vertebrate Evo-Devo. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, **324**, 316-341.

Bradbury IR, Hamilton LC, Dempson JB *et al.* (2015). Transatlantic secondary contact in Atlantic salmon, comparing microsatellites, a SNP array, and Restriction Associated DNA sequencing for the resolution of complex spatial structure. *Molecular Ecology*, **24**, 5130–5144.

Bradbury IR, Wringe BF, Watson B *et al.* (2018) Genotyping-by-sequencing of genome-wide microsatellite loci reveals fine-scale harvest composition in a coastal Atlantic salmon fishery. *Evolutionary Applications*, **11**, 918-930.

Branch TA, Stafford KM, Palacios DM *et al.* (2007) Past and present distribution, densities and movements of blue whales *Balaenoptera musculus* in the Southern Hemisphere and northern Indian Ocean. *Mammal Review*, **37**, 116–175.

Brandão SN (2010) Macrocyprididae (Ostracoda) from the Southern Ocean: Taxonomic revision, macroecological patterns, and biogeographical implications. *Zoological Journal of the Linnean Society*, **159**, 567–672.

Brandão SN, Sauer J, Schön I (2010) Circumantarctic distribution in Southern Ocean benthos? A genetic test using the genus Macroscapha (Crustacea, Ostracoda) as a model. *Molecular Phylogenetics and Evolution*, **55**, 1055–1069.

Brandt A, Gooday AJ, Brandão SN *et al.* (2007) First insights into the biodiversity and biogeography of the Southern Ocean deep sea. *Nature*, **447**, 307–311.

Brasier MJ, Constable A, Melbourne-Thomas J *et al.* (2019) Observations and models to support the first Marine Ecosystem Assessment for the Southern Ocean (MEASO). *Journal of Marine Systems*, **197**, 103182.

Bromwich D, Nicolas J, Monaghan A *et al.* (2013) Central west Antarctica among the most rapidly warming regions on Earth. *Nature Geoscience*, **6**, 139–145.

Brooks CM (2013) Competing values on the Antarctic high seas: CCAMLR and the challenge of marine-protected areas. *The Polar Journal*, **3**, 277–300.

Brooks CM, Crowder LB, Curran LM *et al.* (2016) Science-based management in decline in the Southern Ocean. *Science*, **354**, 185–187.

Brooks CM, Crowder LB, Österblom H, Strong AL (2019) Reaching consensus for conserving the global commons : The case of the Ross Sea , Antarctica. *Conservation Letters*, e12676.

De Broyer C, Koubbi P, Griffiths H *et al.* (2014) *Biogeographic Atlas of the Southern Ocean*. Scientific Committee on Antarctic Research, Cambridge, 510 pp., ISBN 978-0-948277-28-3.

de Bruyn M, Pinsky ML, Hall B *et al.* (2014) Rapid increase in southern elephant seal genetic diversity after a founder event. *Proceedings of the Royal Society B: Biological Sciences*, **281**, 20133078.

Bucklin A, Steinke D, Blanco-Bercial L (2011) DNA barcoding of marine metazoa. *Annual Review of Marine Science*, **3**, 471–508.

Burchett MS (1983) The life cycle of *Notothenia rossii* from South Georgia. *British Antarctic Survey Bulletin*, **61**, 71-73.

Burford Reiskind M, Coyle K, Daniels H *et al.* (2016) Development of a universal double-digest RAD sequencing approach for a group of non-model, ecologically and economically important insect and fish taxa. *Molecular Ecology Resources*, **16**, 1303-1314.

Burns M, Starrett J, Derkarabetian S *et al.* (2017). Comparative performance of double-digest RAD sequencing across divergent arachnid lineages. *Molecular Ecology Resources*, **17**, 418–430.

Butchart SHM, Walpole M, Collen B *et al.* (2010) Global biodiversity: indicators of recent declines. *Science*, **328**, 1164–8.

Byrne M, Gall M, Wolfe K, Aguera A (2016) From pole to pole: the potential for the Arctic seastar *Asterias amurensis* to invade a warming Southern Ocean. *Global Change Biology*, 22, 3874-3887.

Caccavo J, Ashford J, Ryan S *et al.* (2019) Spatial structuring and life history connectivity of Antarctic silverfish along the southern continental shelf of the Weddell Sea. *Marine Ecology Progress Series*, **624**, 195–212.

Caccavo JA, Papetti C, Wetjen M *et al.* (2018) Along-shelf connectivity and circumpolar gene flow in Antarctic silverfish (*Pleuragramma antarctica*). *Scientific Reports*, **8**, 17856.

Calì F, Riginella E, La Mesa M, Mazzoldi C (2017) Life history traits of *Notothenia rossii* and *N. coriiceps* along the southern Scotia Arc. *Polar Biology*, **40**, 1409–1423.

Cameron RA, Kudtarkar P, Gordon SM, Worley KC, Gibbs RA (2015) Do echinoderm genomes measure up? *Marine Genomics*, **22**, 1–9.

Campbell EO, Brunet BMT, Dupuis JR, Sperling FAH (2018) Would an RRS by any other name sound as RAD? *Methods in Ecology and Evolution*, **9**, 1920–1927.

Campbell HA, Fraser KPP, Bishop CM, Peck LS, Egginton S (2008) Hibernation in an Antarctic fish: On ice for winter. *PLoS ONE*, **3**, e1743.

Canales-Aguirre CB, Ferrada-Fuentes S, Galleguillos R, Oyarzun FX, Hernández CE (2018) Population genetic structure of Patagonian toothfish (*Dissostichus eleginoides*) in the Southeast Pacific and Southwest Atlantic Ocean. *PeerJ*, **6**, e4173.

Cao X, Liu J, Chen J *et al.* (2016) Rapid dissemination of taxonomic discoveries based on DNA barcoding and morphology. *Scientific Reports*, **6**, 1–13.

Carlson D (2015) Meeting Future Challenges of Antarctic Research - International and Interdisciplinary Cooperation. In: *Exploring the Last Continent* (eds Liggett D, Storey B, Cook Y, Meduna V), pp. 557–571. Springer International Publishing, Switzerland.

Carvalho SB, Velo-Antón G, Tarroso P *et al.* (2017) Spatial conservation prioritization of biodiversity spanning the evolutionary continuum. *Nature Ecology and Evolution*, **1**, 0151.

Casaux RJ, Barrera-Oro ER (2002) Effect of a shore-based sampling programme on *Notothenia coriiceps* populations. *Antarctic Science*, **14**, 221–224.

Casaux R, Bertolin ML, Carlini A (2011) Feeding habits of three seal species at the Danco Coast, Antarctica: A re-assessment. *Polar Biology*, **34**, 1615–1620.

Catanach A, Crowhurst R, Deng C *et al.* (2019) The genomic pool of standing structural variation outnumbers single nucleotide polymorphism by threefold in the marine teleost *Chrysophrys auratus*. *Molecular Ecology*, **28**, 1210–1223.

Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH (2011) Stacks: building and genotyping loci de novo from short-read sequences. *G3 Genes Genomes Genetics*, **1**, 171–82.

Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool set for population genomics. *Molecular Ecology*, **22**, 3124–3140.

Catchen JM, Hohenlohe PA, Bernatchez L *et al.* (2017) Unbroken: RADseq remains a powerful tool for understanding the genetics of adaptation in natural populations. *Molecular Ecology Resources,* **17**, 362-365.

CCAMLR (2009) *Conservation Measure 91-03*. Hobart, Australia. https://www.ccamlr.org/measure-91-03-2009

CCAMLR (2011) *Report of the thirtieth meeting of the commission*. Hobart, Australia.

CCAMLR (2012) *Conservation Measure 91-04*. General framework for the establishment of CCAMLR Marine Protected Areas. Hobart, Australia. https://www.ccamlr.org/measure-91-04-2012

CCAMLR (2016) *Conservation Measure 91-05*. Hobart, Australia. https://www.ccamlr.org/measure-91-05-2016

CCAMLR (2018) Convention area. https://www.ccamlr.org/en/organisation/convention-area

CCAMLR (2019a) *Schedule of Conservation Measures in Force 2019/20*. https://www.ccamlr.org/en/system/files/e-schedule2019-20_1.pdf

CCAMLR (2019b) *Report of the Thirty-eighth meeting of the Commission*. Hobart, Australia. https://www.ccamlr.org/en/system/files/e-cc-38-prelim-v2.pdf

Ceballos SG, Lessa EP, Victorio MF, Fernández DA (2012) Phylogeography of the sub-Antarctic notothenioid fish *Eleginops maclovinus*: evidence of population expansion. *Marine Biology*, **159**, 499–505.

Ceballos SG, Roesti M, Matschiner M *et al.* (2019). Phylogenomics of an extra-Antarctic notothenioid radiation reveals a previously unrecognized lineage and diffuse species boundaries. *BMC Evolutionary Biology*, **19**, 13.

Chamberlain S, Barve V, Mcglinn D *et al.* (2019) rgbif: Interface to the Global Biodiversity Information Facility API. https://CRAN.R-project.org/package=rgbif

Chapman CC (2017) New perspectives on frontal variability in the Southern Ocean. *Journal of Physical Oceanography*, **47**, 1151–1168.

Chapman CC, Lea MA, Meyer A *et al.* (2020). Defining Southern Ocean fronts and their influence on biological and physical processes in a changing climate. *Nature Climate Change*, **10**, 209–219.

Chen W, Bonillo C, Lecointre G (2003) Repeatability of clades as a criterion of reliability: a case study for molecular phylogeny of Acanthomorpha (Teleostei) with larger number of taxa. *Molecular Phylogenetics and Evolution*, **26**, 262–288.

Chen L, DeVries AL, Cheng C-HC (1997) Evolution of antifreeze glycoprotein gene from a trypsinogen

gene in Antarctic notothenioid fish. *Proceedings of the National Academy of Sciences*, **94**, 3811–3816.

Chen L, Lu Y, Li W *et al.* (2019) The genomic basis for colonizing the freezing Southern Ocean revealed by Antarctic toothfish and Patagonia robalo genomes. *GigaScience*, **8**, giz016.

Cherel Y, Fontaine C, Richard P, Labat J (2009) Isotopic niches and trophic levels of myctophid fishes and their predators in the Southern Ocean. *Limnology and Oceanography*, **55**, 324–332.

Cheung WWL, Watson R, Pauly D (2013) Signature of ocean warming in global fisheries catch. *Nature*, **497**, 365–369.

Chistiakov DA, Hellemans B, Volckaert FAM (2006) Microsatellites and their genomic distribution , evolution, function and applications : A review with special reference to fish genetics. *Aquaculture*, **255**, 1–29.

Christiansen H, Fournier N, Hellemans B, Volckaert FAM (2018) Seafood substitution and mislabeling in Brussels' restaurants and canteens. *Food Control*, **85**, 66–75.

Cipro CVZ, Cherel Y, Bocher P *et al.* (2017) Trace elements in invertebrates and fish from Kerguelen waters, southern Indian Ocean. *Polar Biology*, **41**, 175-191.

Clark L (2017) lvclark/R_genetics_conv: R_genetics_conv 1.1 (Version v1.1). http://doi.org/10.5281/zenodo.846816

Clarke A, Barnes DKA, Hodgson DA (2005) How isolated is Antarctica? *Trends in Ecology and Evolution*, **20**, 1–3.

Clarke A, Crame JA (2010) Evolutionary dynamics at high latitudes: speciation and extinction in polar marine faunas. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, **365**, 3655–3666.

Clarke LJ, Trebilco R, Walters A, Polanowski AM, Deagle BE (2018) DNA-based diet analysis of mesopelagic fish from the southern Kerguelen Axis. *Deep-Sea Research Part II*. In press.

Clément O, Ozouf-Costaz C, Lecointre G, Berrebi P (1998) Allozymic Polymorphism and Phylogeny of the Family Channichthyidae. In: *Fishes of Antarctica. A biological overview* (eds di Prisco G, Pisano E, Clarke A), pp. 299–309. Springer-Verlag Italia, Milano.

Clucas GV, Younger JL, Kao D *et al.* (2016) Dispersal in the sub-Antarctic: king penguins show remarkably little population genetic differentiation across their range. *BMC Evolutionary Biology*, **16**, 1–14.

Clucas GV, Younger JL, Kao D *et al.* (2018). Comparative population genomics reveals key barriers to dispersal in Southern Ocean penguins. *Molecular Ecology*, **27**, 4680–4697.

Coates DJ, Byrne M, Moritz C (2018) Genetic diversity and conservation units: Dealing with the species-population continuum in the age of genomics. *Frontiers in Ecology and Evolution*, **6**, 1–13.

Collins RA, Boykin LM, Cruickshank RH, Armstrong KF (2012a) Barcoding's next top model: An evaluation of nucleotide substitution models for specimen identification. *Methods in Ecology and Evolution*, **3**, 457–465.

Collins RA, Cruickshank RH (2013) The seven deadly sins of DNA barcoding. *Molecular Ecology Resources*, **13**, 969–975.

Collins MA, Stowasser G, Fielding S *et al.* (2012b) Latitudinal and bathymetric patterns in the distribution and abundance of mesopelagic fish in the Scotia Sea. *Deep Sea Research Part II: Topical Studies in Oceanography*, **59–60**, 189–198.

Collins MA, Xavier JC, Johnston NM *et al.* (2008) Patterns in the distribution of myctophid fish in the northern Scotia Sea ecosystem. *Polar Biology*, **31**, 837–851.

Colombo M, Damerau M, Hanel R, Salzburger W, Matschiner M (2014) Diversity and disparity through time in the adaptive radiation of Antarctic notothenioid fishes. *Journal of Evolutionary Biology*, **28**, 376–394.

Connan M, Cherel Y, Mayzaud P (2007) Lipids from stomach oil of procellariiform seabirds document the importance of myctophid fish in the Southern Ocean. *Limnology and Oceanography*, **52**, 2445–2455.

Connan M, Mayzaud P, Duhamel G, Bonnevie BT, Cherel Y (2010) Fatty acid signature analysis documents the diet of five myctophid fish from the Southern Ocean. *Marine Biology*, **157**, 2303–2316.

Conover DO, Munch SB (2002) Sustaining fisheries yields over evolutionary time scales. *Science*, **297**, 94–6.

Constable AJ (2011) Lessons from CCAMLR on the implementation of the ecosystem approach to managing fisheries. *Fish and Fisheries*, **12**, 138–151.

Constable AJ, Costa DP, Schofield O *et al.* (2016) Developing priority variables ("ecosystem Essential Ocean Variables" — eEOVs) for observing dynamics and change in Southern Ocean ecosystems. *Journal of Marine Systems*, **161**, 26-41.

Constable AJ, de la Mare WK, Agnew DJ, Everson I, Miller D (2000) Managing fisheries to conserve the Antarctic marine ecosystem: practical implementation of the Convention on the Conservation of Antarctic Marine Living Resources (CCAMLR). *ICES Journal of Marine Science*, **57**, 778–791.

Constable AJ, Melbourne-Thomas J, Corney SP *et al.* (2014) Climate change and Southern Ocean ecosystems I: how changes in physical habitats directly affect marine biota. *Global Change Biology*, **20**, 3004-3025.

Convey P, Peck LS (2019) Antarctic environmental change and biological responses. *Science Advances*, **5**, eaaz0888.

Coppe A, Agostini C, Marino IAM *et al.* (2013) Genome evolution in the cold: Antarctic icefish muscle transcriptome reveals selective duplications increasing mitochondrial function. *Genome Biology and Evolution*, **5**, 45–60.

Corander J, Majander KK, Cheng L, Merilä J (2013) High degree of cryptic population differentiation in the Baltic Sea herring *Clupea harengus*. *Molecular Ecology*, **22**, 2931–2940.

Cornejo-Donoso J, Antezana T (2008) Preliminary trophic model of the Antarctic Peninsula Ecosystem (Sub-area CCAMLR 48.1). *Ecological Modelling*, **218**, 1–17.

Corrigan LJ, Fabiani A, Chauke LF *et al.* (2016) Population differentiation in the context of Holocene climate change for a migratory marine species, the southern elephant seal. *Journal of Evolutionary*

*Biology*, **29**, 1667–1679.

Corsolini S, Sarà G (2017) The trophic transfer of persistent pollutants (HCB, DDTs, PCBs) within polar marine food webs. *Chemosphere*, **177**, 189-199.

Costello C, Rassweiler A, Siegel D *et al.* (2010) The value of spatial information in MPA network design. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 18294–18299.

Cousseau MB, Pequeño G, Mabragaña E *et al.* (2019) The Magellanic Province and its fish fauna (South America): Several provinces or one? *Journal of Biogeography*, in press.

Cowart DA, Murphy KR, Cheng C-HC (2017) Metagenomic sequencing of environmental DNA reveals marine faunal assemblages from the West Antarctic Peninsula. *Marine Genomics*, **37**, 148-160.

Cowen RK, Sponaugle S (2009) Larval dispersal and marine population connectivity. *Annual review of marine science*, **1**, 443–466.

Crain CM, Kroeker K, Halpern BS (2008) Interactive and cumulative effects of multiple human stressors in marine systems. *Ecology Letters*, **11**, 1304–1315.

Crame JA (2018) Key stages in the evolution of the Antarctic marine fauna. *Journal of Biogeography*, **45**, 986-994.

Crawford DL, Oleksiak MF (2016) Ecological population genomics in the marine environment. *Briefings in Functional Genomics*, **15**, 342–351.

Creer S, Deiner K, Frey S *et al.* (2016) The ecologist's field guide to sequence-based identification of biodiversity. *Methods in Ecology and Evolution*, **7**, 1008–1018.

Cruz VP, Vera M, Pardo BG *et al.* (2017) Identification and validation of single nucleotide polymorphisms as tools to detect hybridization and population structure in freshwater stingrays. *Molecular Ecology Resources*, **17**, 550–556.

Cuveliers EL, Volckaert FAM, Rijnsdorp AD, Larmuseau MHD, Maes GE (2011) Temporal genetic stability and high effective population size despite fisheries-induced life-history trait evolution in the North Sea sole. *Molecular Ecology*, **20**, 3555–68.

Cziko PA, Cheng CHC (2006) A new species of nototheniid (Perciformes: Notothenioidei) fish from McMurdo Sound, Antarctica. *Copeia*, **2006**, 752–759.

d'Udekem d'Acoz C, Schön I, Robert H (2018) The genus *Charcotia* Chevreux , 1906 in the Southern Ocean, with the description of a new species. *Belgian Journal of Zoology*, **148**, 31–82.

DaCosta JM, Sorenson MD (2014) Amplification biases and consistent recovery of loci in a double-digest RAD-seq protocol. *PloS One*, **9**, e106713.

Damerau M, Matschiner M, Salzburger W, Hanel R (2012) Comparative population genetics of seven notothenioid fish species reveals high levels of gene flow along ocean currents in the southern Scotia Arc, Antarctica. *Polar Biology*, **35**, 1073–1086.

Damerau M, Matschiner M, Salzburger W, Hanel R (2014) Population divergences despite long pelagic larval stages: lessons from crocodile icefishes (Channichthyidae). *Molecular Ecology*, **23**, 284–299.

Danecek P, Auton A, Abecasis G *et al.* (2011) The variant call format and VCFtools. *Bioinformatics*, **27**, 2156–2158.

Davey JW, Blaxter ML (2010) RADSeq: next-generation population genetics. *Briefings in Functional Genomics*, **9**, 416–23.

Davey JW, Cezard T, Fuentes-Utrilla P *et al.* (2013) Special features of RAD sequencing data: implications for genotyping. *Molecular Ecology*, **22**, 3151–3164.

Davey JW, Hohenlohe PA, Etter PD *et al.* (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics*, **12**, 499–510.

Davis MP, Holcroft NI, Wiley EO, Sparks JS, Leo Smith W (2014) Species-specific bioluminescence facilitates speciation in the deep sea. *Marine Biology*, **161**, 1139–1148.

de Busserolles F, Fitzpatrick JL, Paxton JR, Marshall NJ, Collin SP (2013) Eye-size variability in deep-sea lanternfishes (Myctophidae): an ecological and phylogenetic study. *PloS one*, **8**, e58519.

de Busserolles F, Marshall NJ (2017) Seeing in the deep-sea: visual adaptations in lanternfishes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **372**, 20160070.

De Deckker P (2018) On the long-ignored scientific achievements of the Belgica expedition 1897–1899. *Polar Research*, **37**, 1474695.

de Medeiros BAS, Farrell BD (2018) Whole genome amplification in double-digest RAD-seq results in adequate libraries but fewer sequenced loci. *PeerJ*, **17**, e5089.

De Queiroz K (2007) Species concepts and species delimitation. *Systematic Biology*, **56**, 879–886.

De Wit P, Pespeni MH, Palumbi SR (2015) SNP genotyping and population genomics from expressed sequences - Current advances and future possibilities. *Molecular Ecology*, **24**, 2310–2323.

Deagle BE, Faux C, Kawaguchi S, Meyer B, Jarman SN (2015) Antarctic krill population genomics: apparent panmixia, but genome complexity and large population size muddy the water. *Molecular Ecology*, **24**, 4943–4959.

Deagle BE, Jarman SN, Coissac E *et al.* (2014) DNA metabarcoding and the cytochrome c oxidase subunit I marker: not a perfect match. *Biology Letters*, **10**, 20140562.

Deck J, Gaither MR, Ewing R *et al.* (2017) The Genomic Observatories Metadatabase (GeOMe): A new repository for field and sampling event metadata associated with genetic samples. *PLoS Biology*, **15**, e2002925.

Deli Antoni MY, Delpiani SM, González-Castro M *et al.* (2019) Comparative populational study of *Lepidonotothen larseni* and *L. nudifrons* (Teleostei: Nototheniidae) from the Antarctic Peninsula and the South Shetland Islands, Antarctica. *Polar Biology*, **42**, 1537-1547.

Denton JSS (2014) Seven-locus molecular phylogeny of Myctophiformes (Teleostei; Scopelomorpha) highlights the utility of the order for studies of deep-sea evolution. *Molecular Phylogenetics and Evolution*, **76**, 270–92.

Denton JSS (2018) Diversification patterns of lanternfishes reveal multiple rate shifts in a critical mesopelagic clade targeted for human exploitation. *Current Biology*, **28**, 933-940.

Des Roches S, Post DM, Turley NE *et al.* (2018) The ecological importance of intraspecific variation. *Nature Ecology and Evolution*, **2**, 57–64.

Desvignes T, Le François NR, Goetz LC *et al.* (2019) Intergeneric hybrids inform reproductive isolating

barriers in the Antarctic icefish radiation. *Scientific Reports*, **9**, 1–10.

Detrich HW, Amemiya CT (2010) Antarctic notothenioid fishes: genomic resources and strategies for analyzing an adaptive radiation. *Integrative and Comparative Biology*, **50**, 1009–17.

Detrich HW, Jones CD, Kim S *et al.* (2005) Nesting behavior of the icefish *Chaenocephalus aceratus* at Bouvetoya Island, Southern Ocean. *Polar Biology*, **28**, 828–832.

Detrich HW, Stuart A, Schoenborn M *et al.* (2010) Genome enablement of the notothenioidei: genome size estimates from 11 species and BAC libraries from 2 representative taxa. *Journal of Experimental Zoology Part B, Molecular and Developmental Evolution*, **314**, 369–81.

Dettai A, Berkani M, Lautredou A-C *et al.* (2012) Tracking the elusive monophyly of nototheniid fishes (Teleostei) with multiple mitochondrial and nuclear markers. *Marine Genomics*, **8**, 49–58.

Dettai A, Lautredou AC, Bonillo C *et al.* (2011) The actinopterygian diversity of the CEAMARC cruises: Barcoding and molecular taxonomy as a multi-level tool for new findings. *Deep-Sea Research Part II: Topical Studies in Oceanography*, **58**, 250–263.

DeVries AL, Eastman JT (1978) Lipid sacs as a buoyancy adaptation in an Antarctic fish. *Nature*, **271**, 352–353.

DeWitt HH, Heemstra PC, Gon O (1990) Nototheniidae. In: *Fishes of the Southern Ocean* (eds Gon O, Heemstra PC), pp. 279–331. J.L.B. Smith Institute of Ichthyology, Grahamstown, South Africa.

Díaz-Arce N, Rodríguez-Ezpeleta N (2019) Selecting RAD-Seq Data Analysis Parameters for Population Genetics: The More the Better? *Frontiers in Genetics*, **10**, 1–10.

Díaz A, Gérard K, González-Wevar C *et al.* (2018) Genetic structure and demographic inference of the regular sea urchin *Sterechinus neumayeri* (Meissner, 1900 ) in the Southern Ocean: The role of the last glaciation. *PLoS One*, **13**, e0197611.

DiBattista JD, Saenz-Agudelo P, Piatek MJ *et al.* (2017) Using a butterflyfish genome as a general tool for RAD-Seq studies in specialized reef fish. *Molecular Ecology Resources*, **17**, 1330–1341.

Dirzo R, Young HS, Galetti M *et al.* (2014) Defaunation in the Anthropocene. *Science*, **345**, 401–406.

Do C, Waples RS, Peel D *et al.* (2014) NeEstimator v2: Re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Molecular Ecology Resources*, **14**, 209–214.

Dobzhansky T (1950) Mendelian populations and their evolution. *American Naturalist*, **84**, 401–418.

Dodds K (2009) Settling and unsettling Antarctica. *Signs*, **34**, 505–509.

Dolezel J, Bartos J, Voglmayr H, Greilhuber J (2003) Nuclear DNA content and genome size of trout and human. *Cytometry A*, **51**, 127–8.

Domack E, Duran D, Leventer A *et al.* (2005) Stability of the Larsen B ice shelf on the Antarctic Peninsula during the Holocene epoch. *Nature*, **436**, 681–685.

Donnelly J, Torres JJ, Hopkins TL, Lancraft TM (1990) Proximate composition of Antarctic mesopelagic fishes. *Marine Biology*, **106**, 13–23.

Donohue KA, Tracey KL, Watts DR, Chidichimo MP, Chereskin TK (2016) Mean Antarctic Circumpolar Current transport measured in Drake Passage. *Geophysical Research Letters*, **43**, 11,760-11,767.

Dornan T, Fielding S, Saunders RA, Genner MJ, Dornan T (2019) Swimbladder morphology masks Southern Ocean mesopelagic fish biomass. *Proceedings of the Royal Society B: Biological Sciences*, **286**, 20190353.

Dornburg A, Federman S, Eytan RI, Near TJ (2016) Cryptic species diversity in Sub-Antarctic islands: A case study of *Lepidonotothen*. *Molecular Phylogenetics and Evolution*, **104**, 32–43.

Dotto TS, Naveira Garabato A, Bacon S *et al.* (2018). Variability of the Ross Gyre, Southern Ocean: drivers and responses revealed by satellite altimetry. *Geophysical Research Letters*, **45**, 6195–6204

Doubleday ZA, Connell SD (2017) Publishing with objective charisma: Breaking science's paradox. *Trends in Ecology and Evolution*, **32**, 803–805.

Du X, Fan G, Jiao Y *et al.* (2017) The pearl oyster *Pinctada fucata martensii* genome and multi-omic analyses provide insights into biomineralization. *GigaScience*, **6**, 1–12.

Ducklow HW, Fraser WR, Meredith MP *et al.* (2013) West Antarctic peninsula: An ice-dependent coastal marine ecosystem in transition. *Oceanography*, **26**, 190–203.

Duhamel G (1982) Biology and Population Dynamics of *Notothenia rossii rossii* from the Kerguelen Islands (Indian Sector of Southern Ocean). *Polar Biology*, **1**, 141–151.

Duhamel G, Hautecoeur M, Dettai A *et al.* (2010) Liparids from the Eastern sector of Southern Ocean and first information from molecular studies. *Cybium*, **34**, 319–343.

Duhamel G, Hulley P-A, Causse R *et al.* (2014) Chapter 7. Biogeographic patterns of fish. In: *Biogeographic Atlas of the Southern Ocean* (eds De Broyer C, Koubbi P, Griffiths H, et al.), pp. 327–362. Scientific Committee on Antarctic Research, Cambridge.

Duhamel G, Hureau J-C, Ozouf-Costaz C (1982) Ecological Survey of the notothenioid fishes in the Southern Ocean from Bouvet to Kerguelen Islands. In: *Proceedings of the BIOMASS Colloqium in 1982* (eds Nemoto T, Matsuda T). National Institute of Polar Research, Tokyo.

Duhamel G, Koubbi P, Ravier C (2000) Day and night mesopelagic fish assemblages off the Kerguelen Islands (Southern Ocean). *Polar Biology*, **23**, 106–112.

Duhamel G, Ozouf-Costaz C, Cattaneo-Berrebi G, Berrebi P (1995) Interpopulation relationships in two species of Antarctic fish *Notothenia rossii* and *Champsocephalus gunnari* from the Kerguelen Islands: an allozyme study. *Antarctic Science*, **7**, 351–356.

Duhamel G, Péron C, Sinègre R *et al.* (2017) Important readjustments in the biomass and distribution of groundfish species in the northern part of the Kerguelen Plateau and Skiff Bank. In: *Scond Kerguelen Plateau Symposium: marine ecosystem and fisheries*, pp. 135–184.

Dyke PPG (2001) *Coastal and Shelf Sea Modelling*. Kluwer Academic, Boston, MA.

Eakin RR, Riginella E, La Mesa M (2015) A new species of *Artedidraco* (Pisces: Artedidraconidae) from the Weddell Sea, Antarctica. *Polar Biology*, **38**, 1597–1603.

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER : a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resourcesa*, **4**, 359–361.

Eastman JT (1991) Evolution and diversification of Antarctic notothenioid fishes. *American Zoologist*, **31**,

93–109.

Eastman JT (1993) *Antarctic Fish Biology - Evolution in a Unique Environment*. Academic Press, San Diego, CA, USA.

Eastman JT (2005) The nature of the diversity of Antarctic fishes. *Polar Biology*, **28**, 93–107.

Eastman JT (2017) Bathymetric distributions of notothenioid fishes. *Polar Biology*, **40**, 2077–2095.

Eastman JT, Barrera-Oro E, Moreira E (2011) Adaptive radiation at a low taxonomic level: Divergence in buoyancy of the ecologically similar Antarctic fish *Notothenia coriiceps* and *N. rossii*. *Marine Ecology Progress Series*, **438**, 195–206.

Eastman JT, McCune AR (2000) Fishes on the Antarctic continental shelf: evolution of a marine species flock? *Journal of Fish Biology*, **57**, 84–102.

Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, **5**, 113.

Elith J, Kearney M, Phillips S (2010) The art of modelling range-shifting species. *Methods in Ecology and Evolution*, **1**, 330–342.

Elith J, Leathwick JR (2009) Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 677–697.

Elith J, Leathwick JR, Hastie T (2008) A working guide to boosted regression trees. *Journal of Animal Ecology*, **77**, 802–813.

Ellegren H (2014) Genome sequencing and population genomics in non-model organisms. *Trends in Ecology & Evolution*, **29**, 51–63.

Elshire RJ, Glaubitz JC, Sun Q *et al.* (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One*, **6**, e19379.

Epskamp S, Cramer AOJ, Waldorp LJ, Schmittmann VD, Borsboom D (2012) qgraph: Network Visualizations of Relationships in Psychometric Data. *Journal of Statistical Software*, **48**.

Eriksen M, Lebreton LCM, Carson HS *et al.* (2014) Plastic pollution in the World' s Oceans: More than 5 trillion plastic pieces weighing over 250,000 tons afloat at sea. *Plos One*, **9**, e111913.

Eschmeyer WN, Fricke R, van der Laan R (2018) Catalog of Fishes. http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp

Eschmeyer WN, Fong JD. (2018). Species by Family/Subfamily. Available online at: http://researcharchive.calacademy.org/research/ichthyology/catalog/SpeciesByFamily.asp. Electronic version (Accessed 14 January, 2018)

Escoda L, González-Esteban J, Gómez A, Castresana J (2017) Using relatedness networks to infer contemporary dispersal: application to the endangered mammal *Galemys pyrenaicus*. *Molecular Ecology*, **26**, 3343-3357.

Euclide PT, McKinney GJ, Bootsma M *et al.* (2019). Attack of the PCR clones: Rates of clonality have little effect on RAD-seq genotype calls. *Molecular Ecology Resources*, **20**, 66–78.

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

Everson, I., Neyelov, A., & Permitin, Y. E. (1992). Bycatch of fish in the south atlantic krill fishery. *Antarctic Science*, **4**, 389–392.

Everson I (2015) Designation and management of large-scale MPAs drawing on the experiences of CCAMLR. *Fish and Fisheries*, **18**, 145-159.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of Molecular Variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.

Eytan RI, Evans BR, Dornburg A *et al.* (2015) Are 100 enough? Inferring acanthomorph teleost phylogeny using Anchored Hybrid Enrichment. *BMC Evolutionary Biology*, **15**, 113.

Falush D, Stephens M, Pritchard JK (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567–1587.

Fang B, Merilä J, Ribeiro F, Alexandre CM, Momigliano P (2018) Worldwide phylogeny of three-spined sticklebacks. *Molecular Phylogenetics and Evolution*, **127**, 613–625.

FAO (2002) *A Fishery Manager's Guidebook - Management Measures and Their Application* (KL Cochrane, Ed,). Food and Agriculture Organization of the United Nations, Rome.

FAO (2018) *Impacts of climate change on fisheries and aquaculture Synthesis of current knowledge, adaptation and mitigation options* (M Barange, T Bahri, MCM Beveridge, et al., Eds,). Food and Agriculture Organization of the United Nations, Rome.

Faust E, Halvorsen KT, Andersen P, Knutsen H, André C (2018) Cleaner fish escape salmon farms and hybridize with local wrasse populations. *Royal Society Open Science*, **5**, 171752.

Ferrando S, Amaroli A, Gallus L *et al.* (2019) Olfaction in the Antarctic toothfish *Dissostichus mawsoni*: clues from the morphology and histology of the olfactory rosette and bulb. *Polar Biology*, **42**, 1081-1091.

Ferreira MF, Varela ML, Lo Nostro F, Ansaldo M, Genovese G (2017) Reproductive aspects of *Notothenia rossii* and *N. coriiceps* (Perciformes, Nototheniidae) at Potter Cove, 25 de Mayo (King George) Island during austral summer. *Polar Biology*, **40**, 1–11.

Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, **24**, 38–49.

Fischer MC, Rellstab C, Leuzinger M *et al.* (2017). Estimating genomic diversity and population differentiation – an empirical comparison of microsatellite and SNP variation in *Arabidopsis halleri*. *BMC Genomics*, **18**, 69.

Fitzgibbon J, Hope A, Slobodyanyuk SJ *et al.* (1995) The rhodopsin-encoding gene of bony fish lacks introns. *Gene*, **164**, 273–277.

Flanagan SP, Jones AG (2018) Substantial differences in bias between single-digest and double-digest RAD-seq libraries: A case study. *Molecular Ecology Resources*, **18**, 264–280.

Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, **180**, 977–993.

Forester BR, Lasky JR, Wagner HH, Urban DL (2018) Comparing methods for detecting multilocus

adaptation with multivariate genotype–environment associations. *Molecular Ecology*, **27**, 2215–2233.

Fossheim M, Primicerio R, Johannesen E *et al.* (2015) Recent warming leads to a rapid borealization of fish communities in the Arctic. *Nature Climate Change*, **5**, 1–6.

Fountain ED, Pauli JN, Reid BN, Palsbøll PJ, Peery MZ (2016) Finding the right coverage: The impact of coverage and sequence quality on single nucleotide polymorphism genotyping error rates. *Molecular Ecology Resources*, **16**, 966–978.

Franchini P, Fruciano C, Spreitzer ML *et al.* (2014) Genomic architecture of ecologically divergent body shape in a pair of sympatric crater lake cichlid fishes. *Molecular Ecology*, **23**, 1828–45.

Franchini P, Monné Parera D, Kautt AF, Meyer A (2017) quaddRAD: A new high-multiplexing and PCR duplicate removal ddRAD protocol produces novel evolutionary insights in a non-radiating cichlid lineage. *Molecular Ecology*, **26**, 2783-2795.

Fraser CI, Morrison AK, Hogg AMC *et al.* (2018) Antarctica's ecological isolation will be broken by storm-driven dispersal and warming. *Nature Climate Change*, **8**, 704–708.

Fraser CI, Nikula R, Spencer HG, Waters JM (2009) Kelp genes reveal effects of subantarctic sea ice during the Last Glacial Maximum. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 3249–3253.

Free CM, Thorson JT, Pinsky ML *et al.* (2019) Impacts of historical warming on marine fisheries production. *Science*, **363**, 979–983.

Freeman NM, Lovenduski NS, Gent PR (2016). Temporal variability in the Antarctic Polar Front (2002–2014). *Journal of Geophysical Research: Oceans*, **121**, 7263–7276

Freer JJ, Partridge JC, Tarling GA, Collins MA, Genner MJ (2018) Predicting ecological responses in a changing ocean: the effects of future climate uncertainty. *Marine Biology*, **165**, 7.

Freer JJ, Tarling GA, Collins MA, Partridge JC, Genner MJ (2019) Predicting future distributions of lanternfish, a significant ecological resource within the Southern Ocean. *Diversity and Distributions*, **25**, 1259-1272.

Fricke R, Eschmeyer WN, Van der Laan R (2019) Eschmeyer's Catalog of Fishes: Genera, Species, References. http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp

Fu L, Niu B, Zhu Z, Wu S, Li W (2012) CD-HIT: Accelerated for clustering the next-generation sequencing data. *Bioinformatics*, **28**, 3150–3152.

Fuentes-Pardo AP, Ruzzante DE (2017) Whole-genome sequencing approaches for conservation biology: advantages, limitations, and practical recommendations. *Molecular Ecology*, **26**, 5369-5406.

Fullerton AH, Lindley ST, Pess GR *et al.* (2011) Human influence on the spatial structure of threatened Pacific salmon metapopulations. *Conservation Biology*, **25**, 932–944.

Funk WC, McKay JK, Hohenlohe PA, Allendorf FW (2012) Harnessing genomics for delineating conservation units. *Trends in Ecology & Evolution*, **27**, 489–96.

Gagnaire P-A, Broquet T, Aurelle D *et al.* (2015) Using neutral, selected and hitchhiker loci to assess connectivity of marine populations in the genomic era. *Evolutionary Applications*, **8**, 769–786.

Gaither MR, Gkafas GA, De Jong M *et al.* (2018) Genomics of habitat choice and adaptive evolution in a deep-sea fish. *Nature Ecology and Evolution*, **2**, 680–687.

Galaska MP, Sands CJ, Santos SR, Mahon AR, Halanych KM (2017) Geographic structure in the Southern Ocean circumpolar brittle star *Ophionotus victoriae* (Ophiuridae) revealed from mtDNA and single-nucleotide polymorphism data. *Ecology and Evolution*, **7**, 475–485.

Galtier N (2019) Delineating species in the speciation continuum: a proposal. *Evolutionary Applications*, **12**, 657-663.

Gao G, Nome T, Pearse DE *et al.* (2018) A new single nucleotide polymorphism database for rainbow trout generated through whole genome resequencing. *Frontiers in Genetics*, **9**, 147.

Garrison E, Marth G (2012) Haplotype-based variant detection from short-read sequencing. *arXiv preprint arXiv:1207.3907*, 9.

Gautier M, Gharbi K, Cezard T *et al.* (2013). The effect of RAD allele dropout on the estimation of genetic variation within and between populations. *Molecular Ecology*, **22**, 3165–3178.

Geiger MF, Herder F, Monaghan MT *et al.* (2014) Spatial heterogeneity in the mediterranean biodiversity hotspot affects barcoding accuracy of its freshwater fishes. *Molecular Ecology Resources*, **14**, 1210–1221.

Ghigliotti L, Cheng CC-H, Ozouf-Costaz C, Vacchi M, Pisano E (2015) Cytogenetic diversity of notothenioid fish from the Ross sea: historical overview and updates. *Hydrobiologia*, **761**, 373–396.

Gilmour ME, Holmes ND, Fleishman AB, Kriwoken LK (2019) Temporal and interspecific variation in feather mercury in four penguin species from Macquarie Island, Australia. *Marine Pollution Bulletin*, **142**, 282–289.

Glover AG, Wiklund H, Chen C, Dahlgren TG (2018) Managing a sustainable deep-sea 'blue economy' requires knowledge of what actually lives there. *eLife*, **7**, e41319.

Godfray HCJ (2002) Challenges for taxonomy. *Nature*, **417**, 17–19.

Gon O, Heemstra PC (1990) *Fishes of the Southern Ocean*. JLB Smith Institute of Ichthyology, Grahamstown.

Goodwin S, McPherson JD, McCombie WR (2016) Coming of age: ten years of next-generation sequencing technologies. *Nature Reviews Genetics*, **17**, 333–351.

Gordeeva N V. (2013) Genetic divergence in the tribe Electronini (Myctophidae). *Journal of Ichthyology*, **53**, 575–584.

Goudet J, Jombart T (2015) hierfstat: estimation and tests of hierarchical F-statistics. http://github.com/jgx65/hierfstat

Graham CF, Glenn TC, McArthur AG *et al.* (2015) Impacts of degraded DNA on restriction enzyme associated DNA sequencing (RADSeq). *Molecular Ecology Resources*, 15, 1304-1315.

Grant RA, Griffiths HJ, Steinke D, Wadley V, Linse K (2010) Antarctic DNA barcoding; a drop in the ocean? *Polar Biology*, **34**, 775–780.

Grant RA, Linse K (2009) Barcoding Antarctic biodiversity: Current status and the CAML initiative, a case study of Marine invertebrates. *Polar Biology*, **32**, 1629–1637.

Green SJ, Akins JL, Maljkovic A, Cote IM (2012) Invasive lionfish drive Atlantic coral reef fish declines. *PLoS ONE*, **7**, e32596.

Greenwell B, Boehmke B, Cunningham J, GBM Developers (2019) R package "gbm." https://github.com/gbm-developers/gbm

Gregory TR (2018) Animal Genome Size Database. http://www.genomesize.com

Gregory S, Brown J, Belchier M (2014) Ecology and distribution of the grey notothen, *Lepidonotothen squamifrons*, around South Georgia and Shag Rocks, Southern Ocean. *Antarctic Science*, **26**, 239–249.

Grewe F, Huang JP, Leavitt SD, Lumbsch HT (2017) Reference-based RADseq resolves robust relationships among closely related species of lichen-forming fungi using metagenomic DNA. *Scientific Reports*, **7**, 9884.

Griffiths HJ (2010) Antarctic marine biodiversity--what do we know about the distribution of life in the Southern Ocean? *PloS ONE*, **5**, e11683.

Griffiths HJ, Danis B, Clarke A (2011) Quantifying Antarctic marine biodiversity: The SCAR-MarBIN data portal. *Deep-Sea Research Part II*, **58**, 18–29.

Griffiths HJ, Meijers AJS, Bracegirdle TJ (2017) More losers than winners in a century of future Southern Ocean seafloor warming. *Nature Climate Change*, **7**, 749–754.

Gruber DF, Phillips BT, O'Brien R *et al.* (2019) Bioluminescent flashes drive nighttime schooling behavior and synchronized swimming dynamics in flashlight fish. *PloS ONE*, **14**, e0219852.

Guichoux E, Lagache L, Wagner S *et al.* (2011) Current trends in microsatellite genotyping. *Molecular Ecology Notes*, **11**, 591–611.

Guillaumot C, Artois J, Saucède T *et al.* (2019) Broad-scale species distribution models applied to data-poor areas. *Progress in Oceanography*, **175**, 198–207.

Guillaumot C, Martin A, Eléaume M, Saucède T (2018a) Methods for improving species distribution models in data-poor areas: Example of sub-Antarctic benthic species on the Kerguelen Plateau. *Marine Ecology Progress Series*, **594**, 149–164.

Guillaumot C, Raymond B, Danis B (2018b) Marine environmental data layers for Southern Ocean species distribution modelling. https://data.aad.gov.au/metadata/records/environmental_layers

Gustafson RG, Waples RS, Myers JM *et al.* (2007) Pacific salmon extinctions: Quantifying lost and remaining diversity. *Conservation Biology*, **21**, 1009–1020.

Gutt J, Bertler N, Bracegirdle TJ *et al.* (2014) The Southern Ocean ecosystem under multiple climate change stresses - an integrated circumpolar assessment. *Global Change Biology*, **21**, 1434–1453.

Gutt J, Fricke A, Teixidó N, Potthoff M, Arntz WE (2006) Mega-epibenthos at Bouvet Island (South Atlantic): A spatially isolated biodiversity hot spot on a tiny geological spot. *Polar Biology*, **29**, 97–105.

Gutt J, Isla E, Bertler AN *et al.* (2017) Cross-disciplinarity in the advance of Antarctic ecosystem research. *Marine Genomics*, **37**, 1-17.

Gutt J, Zurell D, Bracegridle TJ *et al.* (2012) Correlative and dynamic species distribution modelling for

ecological predictions in the Antarctic: A cross-disciplinary concept. *Polar Research*, **31**, 0–23.

Hadfield J, Retief J (2018) A profusion of confusion in NGS methods naming. *Nature Methods*, **15**, 7–8.

Hagen W, Kattner G, Friedrich C (2000) The lipid compositions of high-Antarctic notothenioid fish species with different life strategies. *Polar Biology*, **23**, 785–791.

Hajibabaei M, Singer GAC, Hebert PDN, Hickey DA (2007) DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *Trends in Genetics*, **23**, 167–172.

Hall MR, Kocot KM, Baughman KW *et al.* (2017) The crown-of-thorns starfish genome as a guide for biocontrol of this coral reef pest. *Nature*, **544**, 231–234.

Halpern BS, Walbridge S, Selkoe KA *et al.* (2008) A global map of human impact on marine ecosystems. *Science*, **319**, 948–52.

Hamilton PB, Cowx IG, Oleksiak MF *et al.* (2015) Population-level consequences for wild fish exposed to sublethal concentrations of chemicals - a critical review. *Fish and Fisheries*, **17**, 545-566.

Hanchet S, Dunn A, Parker S *et al.* (2015a) The Antarctic toothfish (*Dissostichus mawsoni*): biology, ecology, and life history in the Ross Sea region. *Hydrobiologia*, **761**, 397–414.

Hanchet SM, Mormede S, Dunn A (2010) Distribution and relative abundance of Antarctic toothfish (*Dissostichus mawsoni*) on the Ross Sea shelf. *CCAMLR Science*, **17**, 33–51.

Hanchet S, Sainsbury K, Butterworth D *et al.* (2015b) CCAMLR's precautionary approach to management focusing on Ross Sea toothfish fishery. *Antarctic Science*, **8**, 1–8.

Hardie DC, Hebert PD. (2003) The nucleotypic effects of cellular DNA content in cartilaginous and ray-finned fishes. *Genome*, **46**, 683–706.

Harrison S, Taylor AD (1997) Empirical Evidence for Metapopulation Dynamics. In *Metapopulation Biology*, 27-42, Academic Press.

Harrisson KA, Pavlova A, Telonis-Scott M, Sunnucks P (2014) Using genomics to characterize evolutionary potential for conservation of wild populations. *Evolutionary Applications*, **7**, 1008–1025.

Hartl DL, Clark AG (2007) *Principles of Population Genetics*. Sinauer Associates, Inc.

Hauquier F, Leliaert F, Rigaux A, Derycke S, Vanreusel A (2017) Distinct genetic differentiation and species diversification within two marine nematodes with different habitat preference in Antarctic sediments. *BMC Evolutionary Biology*, **17**, 120.

Hauser L, Adcock GJ, Smith PJ, Ramírez JHB, Carvalho GR (2002) Loss of microsatellite diversity and low effective population size in an overexploited population of New Zealand snapper (*Pagrus auratus*). *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 11742–11747.

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

Havermans C, Nagy ZT, Sonet G, De Broyer C, Martin P (2011) DNA barcoding reveals new insights into the diversity of Antarctic species of *Orchomene* sensu lato (Crustacea: Amphipoda: Lysianassoidea). *Deep Sea Research Part II: Topical Studies in Oceanography*, **58**, 230–241.

Heather JM, Chain B (2016) The sequence of sequencers: The history of sequencing DNA. *Genomics*, **107**,

1–8.

Hebert PDN, Cywinska A, Ball SL, DeWaard JR (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences*, **270**, 313–21.

Hedrick PW (2005) A standardized genetic differentiation measure. *Evolution*, **59**, 1633–1638.

Heffelfinger C, Fragoso CA, Moreno MA *et al.* (2014) Flexible and scalable genotyping-by-sequencing strategies for population studies. *BMC Genomics*, **15**, 979.

Heindler FM, Christiansen H, Frédérich B *et al.* (2018) Historical DNA metabarcoding of the prey and microbiome of trematomid fishes using museum samples. *Frontiers in Ecology and Evolution*, **6**, 1–14.

Heindler FM, Christiansen H, Maes, SMT *et al.* (2019) Combined microbiome and stomach content DNA metabarcoding demonstrate novel food web links and microbiome-host interactions along an evolutionary bentho-pelagic gradient in fishes of the Southern Ocean. In: *Ecology and Evolution of Fishes of the Southern Ocean: A multidimensional molecular approach* (PhD Thesis Heindler FM), pp. 79-132. F. M. Heindler, Blaubeuren.

Heino M, Díaz Pauli B, Dieckmann U (2015) Fisheries-induced evolution. *Annual Review of Ecology, Evolution, and Systematics*, **46**, 461–480.

Helyar SJ, Hemmer-Hansen J, Bekkevold D *et al.* (2011) Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. *Molecular Ecology Resources*, **11**, 123–136.

Hemmings AD (2015) Commercial harvest in Antarctica. In: *Exploring the Last Continent* (eds Liggett D, Storey B, Cook Y, Meduna V), pp. 413–428. Springer International Publishing, Switzerland.

Henning F, Machado-Schiaffino G, Baumgarten L, Meyer A (2017) Genetic dissection of adaptive form and function in rapidly-speciating cichlid fishes. *Evolution*, **71**, 1297-1312.

Henri H, Cariou M, Terraz G *et al.* (2015) Optimization of multiplexed RADseq libraries using low-cost adaptors. *Genetica*, **143**, 139–143.

Herrera S, Reyes-Herrera PH, Shank TM (2015) Predicting RAD-seq marker numbers across the eukaryotic tree of life. *Genome Biology and Evolution*, **7**, 3207–3225.

Herrera S, Shank TM (2016) RAD sequencing enables unprecedented phylogenetic resolution and objective species delimitation in recalcitrant divergent taxa. *Molecular Phylogenetics and Evolution*, **100**, 70–79.

Hewitt G (2000) The genetic legacy of the quaternary ice ages. *Nature*, **405**, 907–913.

Hidalgo M, Kaplan DM, Kerr LA *et al.* (2017) Advancing the link between ocean connectivity, ecological function and management challenges. *ICES Journal of Marine Science*, **74**, 1702–1707.

Hijmans RJ, Van Etten J, Sumner M *et al.* (2019) R package "raster." https://cran.r-project.org/package=raster

Hijmans RJ, Phillips S, Leathwick J, Elith J (2017) R package "dismo." http://rspatial.org/sdm/

Hill NA, Foster SD, Duhamel G *et al.* (2017) Model-based mapping of assemblages for ecology and conservation management: A case study of demersal fish on the Kerguelen Plateau. *Diversity and Distributions*, **23**, 1216–1230.

Hillen JEJ, Coscia I, Vandeputte M *et al.* (2017) Estimates of genetic variability and inbreeding in experimentally selected populations of European sea bass. *Aquaculture*, **479**, 742–749.

Hillis DM, Bull JJ (1993) An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Systematic Biology*, **42**, 182–192.

Hindshaw RS, Mariash H, Vick-Majors TJ *et al.* (2019) A decade of shaping the futures of polar early career researchers: A legacy of the International Polar Year. *Polar Record*, **54**, 312-323.

Hinke JT, Cossio AM, Goebel ME *et al.* (2017). Identifying Risk: Concurrent overlap of the antarctic krill fishery with krill-dependent predators in the scotia sea. *PLoS ONE*, **12**, 1–24.

Hoban S, Kelley JL, Lotterhos KE *et al.* (2016) Finding the genomic basis of local adaptation: pitfalls, practical solutions, and future directions. *The American Naturalist*, **188**, 379–397.

Hoddell RJ, Crossley AC, Williams R, Hosie GW (2000) The distribution of Antarctic pelagic fish and larvae (CCAMLR division 58.4.1). *Deep Sea Research Part II: Topical Studies in Oceanography*, **47**, 2519–2541.

Hodel RGJ, Chen S, Payton AC *et al.* (2017) Adding loci improves phylogeographic resolution in red mangroves despite increased missing data: Comparing microsatellites and RAD-Seq and investigating loci filtering. *Scientific Reports*, **7**, 17598.

Hoey JA, Pinsky ML (2018) Genomic signatures of environmental selection despite near-panmixia in summer flounder. *Evolutionary Applications*, **11**, 1732–1747.

Hoffberg S, Kieran T, Catchen J *et al.* (2016) RADcap: sequence capture of dual-digest RADseq libraries with identifiable duplicates and reduced missing data. *Molecular Ecology Resources*, **16**, 1264–1278.

Hoffmann A, Griffin P, Dillon S *et al.* (2015) A framework for incorporating evolutionary genomics into biodiversity conservation and management. *Climate Change Responses*, **2**, 1–23.

Hoffmann AA, Sgrò CM, Kristensen TN (2017) Revisiting adaptive potential, population size, and conservation. *Trends in Ecology and Evolution*, **32**, 506–517.

Hofman RJ (2019) Stopping overexploitation of living resources on the high seas. *Marine Policy*, **103**, 91–100.

Hofmann GE, Barry JP, Edmunds PJ *et al.* (2010) The effect of ocean acidification on calcifying organisms in marine ecosystems: An arganism-to-ecosystem perspective. *Annual Review of Ecology, Evolution, and Systematics*, **41**, 127–147.

Hofmann GE, Buckley BA, Airaksinen S, Keen JE, Somero GN (2000) Heat-shock protein expression is absent in the Antarctic fish *Trematomus bernacchii* (family Nototheniidae). *The Journal of Experimental Biology*, **203**, 2331–2339.

Hofmann GE, Place SP (2007) Genomics-enabled research in marine ecology: Challenges, risks and pay-offs. *Marine Ecology Progress Series*, **332**, 249–255.

Hogg AE, Gudmundsson GH (2017) Commentary: Impacts of the Larsen-C Ice Shelf calving event. *Nature Climate Change*, **7**, 540–542.

Hohenlohe PA, Bassham S, Currey M, Cresko WA (2012) Extensive linkage disequilibrium and parallel adaptive divergence across threespine stickleback genomes. *Philosophical Transactions of the Royal*

*Society B: Biological Sciences*, **367**, 395–408.

Hohenlohe PA, Day MD, Amish SJ *et al.* (2013) Genomic patterns of introgression in rainbow and westslope cutthroat trout illuminated by overlapping paired-end RAD sequencing. *Molecular Ecology*, **22**, 3002–3013.

Horn PL, Sutton CP (2015) An assessment of age and growth of violet cod (*Antimora rostrata*) in the Ross Sea, Antarctica. *Polar Biology*, **38**, 1553-1558.

Hsiao K-C, Cheng C-HC, Fernandes IE *et al.* (1990) An antifreeze glycopeptide gene from the antarctic cod *Notothenia coriiceps neglecta* encodes a polyprotein of high peptide copy number. *Proceedings of the National Academy of Sciences of the United States of America*, **87**, 9265–9269.

Hucke-Gaete R, Osman LP, Moreno CA, Torres D (2004) Examining natural population growth from near extinction: The case of the Antarctic fur seal at the South Shetlands, Antarctica. *Polar Biology*, **27**, 304–311.

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009). Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, **9**, 1322–1332.

Hulley P (1990) Family Myctophidae. In: *Fishes of the Southern Ocean* (eds Gon O, Heemstra P), pp. 146–178. J.L.B. Smith Institute of Ichthyology, Grahamstown.

Humble E, Dasmahapatra KK, Martinez-Barrio A *et al.* (2018) RAD sequencing and a hybrid Antarctic fur seal genome assembly reveal rapidly decaying linkage disequilibrium, global population structure and evidence for inbreeding. *G3 Genes Genomes Genetics*, **8**, 2709–2722.

Hureau J-C, Slosarczyk W (1990) Exploitation and conservation of Antarctic fishes and recent ichthyological research in the Southern Ocean. In: *Fishes of the Southern Ocean* (eds Gon O, Heemstra PC), pp. 52–63. J.L.B. Smith Institute of Ichthyology, Grahamstown, South Africa.

Huth TJ, Place SP (2013) De novo assembly and characterization of tissue specific transcriptomes in the emerald notothen, *Trematomus bernacchii*. *BMC Genomics*, **14**, 805.

Ikeda T (1988) Metabolism and chemical composition of crustaceans from the Antarctic mesopelagic zone. *Deep Sea Research Part A, Oceanographic Research Papers*, **35**, 1991–2002.

IPCC (2019) *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate* [H.-O. Pörtner, D.C. Roberts, V. Masson-Delmotte, P. Zhai, M. Tignor, E. Poloczanska, K. Mintenbeck, A. Alegría, M. Nicolai, A. Okem, J. Petzold, B. Rama, N.M. Weyer (eds.)]. In press.

Irigoien X, Klevjer TA, Røstad A *et al.* (2014) Large mesopelagic fishes biomass and trophic efficiency in the open ocean. *Nature Communications*, **5**, 3271.

Isaacs JD, Tont SA, Wick GL (1974) Deep scattering layers: vertical migration as a tactic for finding food. *Deep-Sea Research and Oceanographic Abstracts*, **21**, 651–656.

Ivanova N V., Zemlak TS, Hanner RH, Hebert PDN (2007) Universal primer cocktails for fish DNA barcoding. *Molecular Ecology Notes*, **7**, 544–548.

Jacob U, Brey T, Fetzer I *et al.* (2006) Towards the trophic structure of the Bouvet Island marine ecosystem. *Polar Biology*, **29**, 106–113.

Jacobsen MW, Christensen C, Madsen R *et al.* (2017) Single nucleotide polymorphism markers for

analysis of historical and contemporary samples of Arctic char (*Salvelinus alpinus*). *Conservation Genetics Resources*, **9**, 587-589.

Jacobsen MW, Pujolar JM, Bernatchez L *et al.* (2014) Genomic footprints of speciation in Atlantic eels (*Anguilla anguilla* and *A. rostrata*). *Molecular Ecology*, **23**, 4785–4798.

Jacquet J, Blood-Patterson E, Brooks C, Ainley D (2016) 'Rational use' in Antarctic waters. *Marine Policy*, **63**, 28–34.

Janko K, Lecointre G, DeVries A *et al.* (2007) Did glacial advances during the Pleistocene influence differently the demographic histories of benthic and pelagic Antarctic shelf fishes ? – Inferences from intraspecific mitochondrial and nuclear DNA sequence diversity. *BMC Evolutionary Biology*, **7**, 220.

Jeffery NW, Ellis EA, Oakley TH, Ryan Gregory T (2017) The genome sizes of ostracod crustaceans correlate with body size and evolutionary history, but not environment. *Journal of Heredity*, **108**, 701–706.

Jeffries DL, Copp GH, Lawson Handley L *et al.* (2016) Comparing RADseq and microsatellites to infer complex phylogeographic patterns, an empirical perspective in the Crucian carp, *Carassius carassius*, L. *Molecular Ecology*, **25**, 2997–3018.

Johnson M, Zaretskaya I, Raytselis Y *et al.* (2008) NCBI BLAST: a better web interface. *Nucleic Acids Research*, **36**, 5–9.

Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, **24**, 1403–1405.

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

Jones CD, Eric Anderson M, Balushkin A V. *et al.* (2008) Diversity, relative abundance, new locality records and population structure of Antarctic demersal fishes from the northern Scotia Arc islands and Bouvetøya. *Polar Biology*, **31**, 1481–1497.

Jones CD, Near TJ (2012) The reproductive behaviour of *Pogonophryne scotti* confirms widespread egg-guarding parental care among Antarctic notothenioids. *Journal of Fish Biology*, **80**, 2629–35.

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

Jurajda P, Roche K, Sedláček I, Všetičková L (2016) Assemblage characteristics and diet of fish in the shallow coastal waters of James Ross Island, Antarctica. *Polar Biology*, **39**, 2299-2309.

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

Kao D, Lai AG, Stamataki E *et al.* (2016) The genome of the crustacean *Parhyale hawaiensis*, a model for animal development, regeneration, immunity and lignocellulose digestion. *eLife*, **5**, 1–45.

Kašparová E, Van De Putte AP, Marshall C, Janko K (2015) Lifestyle and ice: the relationship between ecological specialization and response to pleistocene climate change. *PLoS One*, **10**, e0138766.

Kassambara A, Mundt F (2017) factoextra: extract and visualize the results of multivariate data analyses. R package version 1.0.5. https://cran.r-project.org/package=factoextra

Kearns AM, Restani M, Szabo I *et al.* (2018) Genomic evidence of speciation reversal in ravens. *Nature Communications*, **9**, 906.

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

Kellermann AK, Gauldie RW, Ruzicka JJ (2002) Otolith microincrements in the Antarctic fishes *Notothenia coriiceps* and *Pseudochaenichthys georgianus*. *Polar Biology*, **25**, 799–807.

Kelley JL, Brown AP, Therkildsen NO, Foote AD (2016) The life aquatic: advances in marine vertebrate genomics. *Nature Reviews Genetics*, **17**, 523–534.

Kenchington EL, Baillie SM, Kenchington TJ, Bentzen P (2017) Barcoding Atlantic Canada's mesopelagic and upper bathypelagic marine fishes. *PloS ONE*, **12**, e0185173.

Kerkhove TRH, Hellemans B, De Troch M *et al.* (2019) Isolation and characterisation of 14 novel microsatellite markers through next generation sequencing for the commercial Atlantic seabob shrimp *Xiphopenaeus kroyeri*. *Molecular Biology Reports*, **46**, 6565–6569.

Kerr LA, Hintzen NT, Cadrin SX *et al.* (2016) Lessons learned from practical approaches to reconcile mismatches between biological population structure and stock units of marine fish. *ICES Journal of Marine Science*, **74**, 1708-1722.

Kim B-M, Ahn D-H, Kang S *et al.* (2019a) Skin transcriptome profiling reveals the distinctive molecular effects of temperature changes on Antarctic bullhead notothen. *Molecular & Cellular Toxicology*, **15**, 163–172.

Kim BM, Amores A, Kang S *et al.* (2019b) Antarctic blackfin icefish genome reveals adaptations to extreme environments. *Nature Ecology and Evolution*, **3**, 469–478.

Kimmerling N, Zuqert O, Amitai G *et al.* (2018) Quantitative species-level ecology of reef fish larvae via metabarcoding. *Nature Ecology & Evolution*, **2**, 306–316.

Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, **16**, 111–120.

Klingenberg CP, Ekau W (1996) A combined morphometric and phylogenetic analysis of an ecomorphological trend: pelagization in Antarctic fishes (Perciformes: Nototheniidae). *Biological Journal of the Linnean Society*, **59**, 143–177.

Knaus BJ, Grünwald NJ (2017) VCFR: a package to manipulate and visualize variant call format data in R. *Molecular Ecology Resources*, **17**, 44–53.

Knox GA (1994) *The Biology of the Southern Ocean*. Cambridge University Press.

Kock K-H (1992) *Antarctic Fish and Fisheries*. Cambridge University Press, London.

Kock KH, Belchier M, Jones CD (2004) Is the attempt to estimate the biomass of antarctic fish from a multi-species survey appropriate for all targeted species? *Notothenia rossii* in the Atlantic Ocean sector - Revisited. *CCAMLR Science*, **11**, 141–153.

Kock KH, Gröger J, Jones CD (2013) Interannual variability in the feeding of ice fish (Notothenioidei, Channichthyidae) in the southern Scotia Arc and the Antarctic Peninsula region (CCAMLR

Subareas 48.1 and 48.2). *Polar Biology*, **36**, 1451–1462.

Kock K-H, Jones CD (2005) Fish stocks in the Southern Scotia Arc region—A review and prospects for future research. *Reviews in Fisheries Science*, **13**, 75–108.

Kock K-H, Jones CD, Wilhelms S (2000). Biological characteristics of Antarctic fish stocks in the southern Scotia Arc region. *CCAMLR Science*, **7**, 1–41.

Kock KH, Kellermann A (1991). Reproduction in Antarctic notothenioid fish. *Antarctic Science*, **3**, 125–150.

Kock K-H, Reid K, Croxall J, Nicol S (2007) Fisheries in the Southern Ocean: an ecosystem approach. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **362**, 2333–49.

Koubbi P (2003) ICOTA: Ichtyologie Côtière en Terre Adélie. https://doi.org/10.18142/118

Koubbi P, Moteki M, Duhamel G *et al.* (2011) Ecoregionalization of myctophid fish in the Indian sector of the Southern Ocean: Results from generalized dissimilarity models. *Deep Sea Research Part II: Topical Studies in Oceanography*, **58**, 170–180.

Koubbi P, Grant S, Ramm D *et al.* (2017) Conservation and management of Antarctic silverfish *Pleuragramma antarctica* populations and habitats. In: M. Vacchi et al. (eds), *The Antarctic Silverfish: a Keystone Species in a Changing Ecosystem*, Advances in Polar Ecology, **3**. Springer International Publishing AG.

Krabbe K, Leese F, Mayer C, Tollrian R, Held C (2010) Cryptic mitochondrial lineages in the widespread pycnogonid *Colossendeis megalonyx* Hoek, 1881 from Antarctic and Subantarctic waters. *Polar Biology*, **33**, 281–292.

Krapp T, Lang C, Libertini A, Melzer RR (2006) *Caprella scaura* Templeton, 1836 *sensu lato* (Amphipoda: Caprellidae) in the Mediterranean. *Organisms Diversity and Evolution*, **6**, 77–81.

Kritzer JP, Sale PF (2004) Metapopulation ecology in the sea: From Levins' model to marine ecology and fisheries science. *Fish and Fisheries*, **5**, 131–140.

Kuhn KL, Gaffney PM (2006) Preliminary assessment of population structure in the mackerel icefish (*Champsocephalus gunnari*). *Polar Biology*, **29**, 927–935.

Kuhn KL, Gaffney PM (2008) Population subdivision in the Antarctic toothfish (*Dissostichus mawsoni*) revealed by mitochondrial and nuclear single nucleotide polymorphisms (SNPs). *Antarctic Science*, **20**, 327–338.

Kuhn KL, Near TJ, Jones CD, Eastman JT (2009) Aspects of the biology and population genetics of the Antarctic nototheniid fish *Trematomus nicolai*. *Copeia*, **2009**, 320–327.

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, **33**, 1870-1874.

Lait LA, Marshall HD, Carr SM (2018) Phylogeographic mitogenomics of Atlantic cod *Gadus morhua*: Variation in and among trans-Atlantic, trans-Laurentian,Northern cod, and landlocked fjord populations. *Ecology and Evolution*, **8**, 6420-6437.

Lamichhaney S, Barrio AM, Rafati N *et al.* (2012) Population-scale sequencing reveals genetic differentiation due to local adaptation in Atlantic herring. *Proceedings of the National Academy of*

*Sciences of the United States of America*, **109**, 19345–50.

Lamichhaney S, Fuentes-Pardo AP, Rafati N *et al.* (2017) Parallel adaptive evolution of geographically distant herring populations on both sides of the North Atlantic Ocean. *Proceedings of the National Academy of Sciences*, **114**, E3452–E3461.

Langin KM, Aldridge CL, Fike JA *et al.* (2018) Characterizing range-wide divergence in an alpine-endemic bird : a comparison of genetic and genomic approaches. *Conservation Genetics*, **19**, 1471-1485.

Langlais C, Rintoul S, Schiller A (2011). Variability and mesoscale activity of the Southern Ocean fronts: Identification of a circumpolar coordinate system. *Ocean Modelling*, **39**, 79–96.

Laporte M, Pavey SA, Rougeux C *et al.* (2016) RAD sequencing reveals within-generation polygenic selection in response to anthropogenic organic and metal contamination in North Atlantic Eels. *Molecular Ecology*, **25**, 219–237.

Lautredou AC, Bonillo C, Denys G *et al.* (2010) Molecular taxonomy and identification within the Antarctic genus *Trematomus* (Notothenioidei, Teleostei): How valuable is barcoding with COI? *Polar Science*, **4**, 333–352.

Lautrédou A-C, Hinsinger DD, Gallut C *et al.* (2012) Phylogenetic footprints of an Antarctic radiation: the Trematominae (Notothenioidei, Teleostei). *Molecular Phylogenetics and Evolution*, **65**, 87–101.

Leaché AD, Grummer JA, Harris RB, Breckheimer I (2017) Evidence for concerted movement of nuclear and mitochondrial clines in a lizard hybrid zone. *Molecular Ecology*, **26**, 2306-2316.

Lecointre G, Améziane N, Boisselier M-C *et al.* (2013) Is the species flock concept operational? The Antarctic shelf case. *PLoS ONE*, **8**, e68787.

Legendre P, Gallagher ED (2001) Ecologically meaningful transformations for ordination of species data. *Oecologia*, **129**, 271–280.

Leigh JW, Bryant D (2015) POPART: Full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, **6**, 1110–1116.

Leigh DM, Lischer HEL, Grossen C, Keller LF (2018) Batch effects in a multi-year sequencing study: False biological trends due to changes in read lengths. *Molecular Ecology Resources*, **18**, 778–788.

Leisher C, Mangubhai S, Hess S *et al.* (2012) Measuring the benefits and costs of community education and outreach in marine protected areas. *Marine Policy*, **36**, 1005–1011.

Lepais O, Weir JT (2014) SimRAD: An R package for simulation-based prediction of the number of loci expected in RADseq and similar genotyping by sequencing approaches. *Molecular Ecology Resources*, **14**, 1314–1321.

Levins R (1969) Some demographic and genetic consequences of environmental heterogeneity for biological control. *American Entomologist*, **15**, 237–240.

Lewin HA, Robinson GE, Kress WJ *et al.* (2018) Earth BioGenome Project: Sequencing life for the future of life. *Proceedings of the National Academy of Sciences of the United States of America*, **115**, 4325–4333.

Li H (2014) Toward better understanding of artifacts in variant calling from high-coverage samples.

*Bioinformatics*, **30**, 2843–2851.

Li C, Corrigan S, Yang L *et al.* (2015) DNA capture reveals transoceanic gene flow in endangered river sharks. *Proceedings of the National Academy of Sciences of the United States of America*, **112**, 201508735.

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, **25**, 1754–1760.

Li W, Godzik A (2006) CD-HIT: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*, **22**, 1658–1659.

Li H, Handsaker B, Wysoker A *et al.* (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.

Liggett D (2016) Destination icy wilderness: tourism in Antarctica. *Exploring the Last Continent: An Introduction to Antarctica*, 379–398.

Lin JJ, Wang FY, Li WH, Wang TY (2017) The rises and falls of opsin genes in 59 ray-finned fish genomes and their implications for environmental adaptation. *Scientific Reports*, **7**, 15568.

Linck EB, Hanna Z, Sellas A, Dumbacher JP (2017) Evaluating hybridization capture with RAD probes as a tool for museum genomics with historical bird specimens. *Ecology and Evolution*, **23**, 4755-4767.

Linkowski TB, Zukowski C (1980) Observation on the growth of *Notothenia coriiceps neglecta* Nybelin and *Notothenia rossi marmorata* Fischer in Admiralty Bay (King George Island, South Shetland Islands). *Polish Polar Research*, **1**, 155–162.

Linnaeus C (1753) *Species Plantarum*. Stockholm.

Liu N, Brooks CM (2018) China's changing position towards marine protected areas in the Southern Ocean: Implications for future Antarctic governance. *Marine Policy*, **94**, 189–195.

Liu S, Hansen MM (2017) PSMC (pairwise sequentially Markovian coalescent) analysis of RAD (restriction site associated DNA) sequencing data. *Molecular Ecology Resources*, **17**, 631–641.

Liu H, Yu B, Fu J *et al.* (2019) Different circulation history of mercury in aquatic biota from King George Island of the Antarctic. *Environmental Pollution*, **250**, 892-897.

Lluch J, Servant F, Païssé S *et al.* (2015) The characterization of novel tissue microbiota using an optimized 16S metagenomic sequencing pipeline. *PLoS ONE*, **10**, e0143334.

Long KA, Nossa CW, Sewell MA, Putnam NH, Ryan JF (2016) Low coverage sequencing of three echinoderm genomes: The brittle star *Ophionereis fasciata*, the sea star *Patiriella regularis*, and the sea cucumber *Australostichopus mollis*. *GigaScience*, **5**, 1–4.

Lourenço S, Saunders RA, Collins M *et al.* (2016) Life cycle, distribution and trophodynamics of the lanternfish *Krefftichthys anderssoni* (Lönnberg, 1905) in the Scotia Sea. *Polar Biology*, **40**, 1229-1245.

Lowry DB, Hoban S, Kelley JL *et al.* (2016) Breaking RAD: An evaluation of the utility of restriction site associated DNA sequencing for genome scans of adaptation. *Molecular Ecology Resources*, **17**, 142-152.

Lozier JD, Jackson JM, Dillon ME, Strange JP (2016) Population genomics of divergence among extreme and intermediate color forms in a polymorphic insect. *Ecology and Evolution*, **6**, 1075–1091.

Luu K, Bazin E, Blum MGB (2016) pcadapt: An R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources*, **33**, 67–77.

Lyle M, Gibbs S, Moore TC, Rea DK (2007) Late Oligocene initiation of the Antarctic circumpolar current: Evidence from the South Pacific. *Geology*, **35**, 691–694.

Mabragaña E, de Astarloa JMD, Hanner R, Zhang J, Castro M (2011) DNA barcoding identifies Argentine fishes from marine and brackish waters. *PLoS ONE*, **6**, e28655.

Mabragaña E, Delpiani SM, Rosso JJ *et al.* (2016) Barcoding Antarctic fishes: Species discrimination and contribution to elucidate ontogenetic changes in Nototheniidae. In: *DNA Barcoding in Marine Perspectives* (ed Trivedi S), pp. 213–242. Springer International Publishing, Switzerland.

Majaneva S, Hamon G, Fugmann G, Lisowska M, Baeseman J (2016) Where are they now? - A case study of the impact of international travel support for early career Arctic researchers. *Polar Science*, **10**, 382-394.

Manel S, Loiseau N, Andrello M *et al.* (2019) Long-distance benefits of marine reserves: myth or reality ? *Trends in Ecology & Evolution*, **34**, 342–354.

Manel S, Perrier C, Pratlong M *et al.* (2016) Genomic resources and their influence on the detection of the signal of positive selection in genome scans. *Molecular Ecology*, **25**, 170–184.

Mangano MC, Sarà G, Corsolini S (2017) Monitoring of persistent organic pollutants in the polar regions: knowledge gaps & gluts through evidence mapping. *Chemosphere*, **172**, 37–45.

Marcer A, Méndez-Vigo B, Alonso-Blanco C, Picó FX (2016) Tackling intraspecific genetic structure in distribution models better reflects species geographical range. *Ecology and Evolution*, **6**, 2084–2097.

Maroso F, Hillen JEJ, Pardo BG *et al.* (2018) Performance and precision of double digestion RAD (ddRAD) genotyping in large multiplexed datasets of marine fish species. *Marine Genomics*, **39**, 64–72.

Marschoff ER, Barrera-Oro ER, Alescio NS, Ainley DG (2012) Slow recovery of previously depleted demersal fish at the South Shetland Islands, 1983–2010. *Fisheries Research*, **125–126**, 206–213.

Martin RP, Olson EE, Girard MG, Leo Smith W, Davis MP (2018) Light in the darkness: New perspective on lanternfish relationships and classification using genomic and morphological Data. *Molecular Phylogenetics and Evolution*, **121**, 71-85.

Maschette D, Sumner M, Raymond B (2019) SOmap: Southern Ocean maps. https://github.com/AustralianAntarcticDivision/SOmap

Mastretta-Yanes A, Arrigo N, Alvarez N *et al.* (2014) Restriction site-associated DNA sequencing, genotyping error estimation and de novo assembly optimization for population genetic inference. *Molecular Ecology Resources*, **15**, 28-41.

Matschiner M, Hanel R, Salzburger W (2009) Gene flow by larval dispersal in the Antarctic notothenioid fish *Gobionotothen gibberifrons*. *Molecular Ecology*, **18**, 2574–2587.

Matz M V. (2017) Fantastic beasts and how to sequence them: ecological genomics for obscure model organisms. *Trends in Genetics*, **34**, 121–132.

Mayr E (1942) *Systematics and the origin of species*. Columbia University Press, New York.

McCartney-Melstad E, Mount GG, Shaffer HB (2016) Exon capture optimization in amphibians with large genomes. *Molecular Ecology Resources*, **16**, 1084–1094.

McCauley DE (1997) Genetic population extinction and recolonization. *Trends in Ecology & Evolution*, **6**, 5-8.

McCauley DJ, Pinsky ML, Palumbi SR *et al.* (2015) Marine defaunation: Animal loss in the global ocean. *Science*, **347**, 1255641–1255641.

McClintock J, Ducklow H, Fraser W (2008) Ecological responses to climate change on the Antarctic Peninsula. *American Scientist*, **96**, 302–310.

McCormack SA, Melbourne-Thomas J, Trebilco R, Blanchard JL, Constable A (2019) Alternative energy pathways in Southern Ocean food webs: Insights from a balanced model of Prydz Bay, Antarctica. *Deep-Sea Research Part II: Topical Studies in Oceanography*, in press.

McInnes JC, Jarman SN, Lea M-A *et al.* (2017) DNA metabarcoding as a marine conservation and management tool: A circumpolar examination of fishery discards in the diet of threatened albatrosses. *Frontiers in Marine Science*, **4**, 1–22.

McKinney GJ, Larson WA, Seeb LW, Seeb JE (2016) Radseq provides unprecedented insights into molecular ecology and evolutionary genetics: comment on Breaking RAD by Lowry et al. (2016). *Molecular Ecology Resources*, **17**, 356-361.

McMahon BJ, Teeling EC, Höglund J (2014) How and why should we implement genomics into conservation? *Evolutionary Applications*, 999–1007.

McMillan P, Iwamoto T, Stewart A, Smith PJ (2012) A new species of grenadier, genus *Macrourus* (Teleostei, Gadiformes, Macrouridae) from the southern hemisphere and a revision of the genus. *Zootaxa*, **3165**, 1–24.

McMullin RM, Wing SR, Wing LC, Shatova OA (2017) Trophic position of Antarctic ice fishes reflects food web structure along a gradient in sea ice persistence. *Marine Ecology Progress Series*, **564**, 87–98.

Mee JA, Bernatchez L, Reist JD, Rogers SM, Taylor EB (2015) Identifying designatable units for intraspecific conservation prioritization: A hierarchical approach applied to the lake whitefish species complex (*Coregonus* spp.). *Evolutionary Applications*, **8**, 423–441.

Meirmans PG (2015) Seven common mistakes in population genetics and how to avoid them. *Molecular Ecology*, **24**, 3223–3231.

Meléndez RC, Markle DF (1997) Phylogeny and zoogeography of *Laemonema* and *Guttigadus* (Pisces; Gadiformes; Moridae). *Bulletin of Marine Science*, **61**, 593–670.

Metzker ML (2010) Sequencing technologies - the next generation. *Nature Reviews Genetics*, **11**, 31–46.

Michel L, Danis B, Dubois P *et al.* (2019) Increased sea ice cover disrupts food web structure in coastal Antarctica. *Scientific Reports*, **9**, 8062.

Milano I, Babbucci M, Cariani A *et al.* (2014) Outlier SNP markers reveal fine-scale genetic structuring across European hake populations (*Merluccius merluccius*). *Molecular Ecology*, **23**, 118–35.

Miller KJ, Baird HP, Oosterom J Van, Mondon J, King CK (2018) Complex genetic structure revealed in the circum-Antarctic broadcast spawning sea urchin *Sterechinus neumayeri*. *Marine Ecology Progress Series*, **601**, 153–166.

Milligan BG, Archer FI, Ferchaud AL *et al.* (2018) Disentangling genetic structure for genetic monitoring of complex populations. *Evolutionary Applications*, **11**, 1149–1161.

Mintenbeck K, Barrera-Oro ER, Brey T *et al.* (2012) Impact of climate change on fishes in complex Antarctic ecosystems. In: *Advances in Ecological Research*, pp. 351–426. Elsevier Ltd.

Miya T, Gon O, Mwale M, Cheng C-HC (2016a) Multiple independent reduction or loss of antifreeze trait in low Antarctic and sub-Antarctic notothenioid fishes. *Antarctic Science*, **28**, 17–28.

Miya T, Gon O, Mwale M, Christina Cheng C-H (2013) The effect of habitat temperature on serum antifreeze glycoprotein (AFGP) activity in *Notothenia rossii* (Pisces: Nototheniidae) in the Southern Ocean. *Polar Biology*, **37**, 367–373.

Miya T, Gon O, Mwale M, Poulin E (2016b) Molecular systematics and taxonomic status of three latitudinally widespread nototheniid (Perciformes: Notothenioidei) fishes from the Southern Ocean. *Zootaxa*, **4061**, 381–396.

Miya M, Pietsch TW, Orr JW *et al.* (2010) Evolutionary history of anglerfishes (Teleostei: Lophiiformes): a mitogenomic perspective. *BMC Evolutionary Biology*, **10**, 1–27.

Moffat C, Beardsley RC, Owens B, van Lipzig N (2008) A first description of the Antarctic Peninsula Coastal Current. *Deep-Sea Research Part II: Topical Studies in Oceanography*, **55**, 277–293.

Moffat C, Meredith M (2018). Shelf-ocean exchange and hydrography west of the Antarctic Peninsula: A review. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **376**, 20170164.

Momigliano P, Denys GPJ, Jokinen H, Merilä J (2018) *Platichthys solemdali* sp. nov. (Actinopterygii , Pleuronectiformes): A new flounder species from the Baltic Sea. *Frontiers in Marine Science*, **5**, 225.

Momigliano P, Jokinen H, Calboli F, Aro E, Merilä J (2019) Cryptic temporal changes in stock composition explain the decline of a flounder (*Platichthys* spp.) assemblage. *Evolutionary Applications*, **12**, 549–559.

Momigliano P, Jokinen H, Fraimout A *et al.* (2017) Extraordinarily rapid speciation in a marine fish. *Proceedings of the National Academy of Sciences*, **114**, 6074-6079.

Moon KL, Chown SL, Fraser CI (2017) Reconsidering connectivity in the sub-Antarctic. *Biological Reviews*, **92**, 2164–2181.

Moreau C, Saucède T, Jossart Q *et al.* (2017) Reproductive strategy as a piece of the biogeographic puzzle: a case study using Antarctic sea stars (Echinodermata, Asteroidea). *Journal of Biogeography*, **44**, 848–860.

Moreira E, Barrera-Oro E, La Mesa M (2013) Age validation of juvenile *Notothenia rossii* at Potter Cove, South Shetland Islands, using mark-recapture data. *Polar Biology*, **36**, 1845–1850.

Morescalchi A, Morescalchi MA, Odierna G, Stingo V, Capriglione T (1996) Karyotype and genome size of zoarcids and Notothenioids (Teleostei, perciformes) from the Ross Sea: Cytotaxonomic

implications. *Polar Biology*, **16**, 559–564.

Moriarty Lemmon E, Lemmon AR (2013) High-throughput genomic data in systematics and phylogenetics. *Annual Review of Ecology Evolution and Systematics*, **44**, 99–121.

Morrow JM, Lazic S, Dixon Fox M *et al.* (2017) A second visual rhodopsin gene, *rh1-2* , is expressed in zebrafish photoreceptors and found in other ray-finned fishes. *The Journal of Experimental Biology*, **220**, 294–303.

Moteki M, Fujii K, Amakasu K *et al.* (2017) Distributions of larval and juvenile/adult stages of the Antarctic myctophid fish, *Electrona antarctica*, off Wilkes Land in East Antarctica. *Polar Science*, **12**, 99-108.

Mugue NS, Petrov AF, Zelenina DA, Gordeev II, Sergeev AA (2014) Low genetic diversity and temporal stability in the Antarctic Toothfish (*Dissostichus mawsoni*) from near-continental seas of Antarctica. *CCAMLR Science*, **21**, 1-9.

Mulvaney R, Abram NJ, Hindmarsh RCA *et al.* (2012) Recent Antarctic Peninsula warming relative to Holocene climate and ice-shelf history. *Nature*, **489**, 141–144.

Mura-Jornet I, Pimentel C, Dantas GPM *et al.* (2018) Chinstrap penguin population genetic structure: one or more populations along the Southern Ocean ? *BMC Evolutionary Biology*, **18**, 1–12.

Naimi B, Araújo MB (2016) Sdm: A reproducible and extensible R platform for species distribution modelling. *Ecography*, **39**, 368–375.

Naimi B, Hamm NAS, Groen TA, Skidmore AK, Toxopeus AG (2014) Where is positional uncertainty a problem for species distribution modelling? *Ecography*, **37**, 191–203.

Narum SR, Buerkle CA, Davey JW, Miller MR, Hohenlohe PA (2013) Genotyping-by-sequencing in ecological and conservation genomics. *Molecular Ecology*, **22**, 2841–2847.

Near TJ, Cheng C-HC (2008) Phylogenetics of notothenioid fishes (Teleostei: Acanthomorpha): inferences from mitochondrial and nuclear gene sequences. *Molecular Phylogenetics and Evolution*, **47**, 832–40.

Near TJ, Dornburg A, Kuhn KL *et al.* (2012) Ancient climate change, antifreeze, and the evolutionary diversification of Antarctic fishes. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 3434–3439.

Near TJ, MacGuigan DJ, Parker E *et al.* (2018) Phylogenetic analysis of Antarctic notothenioids illuminates the utility of RADseq for resolving Cenozoic adaptive radiations. *Molecular Phylogenetics and Evolution*, **129**, 268–279.

Near TJ, Russo SE, Jones CD, DeVries AL (2003) Ontogenetic shift in buoyancy and habitat in the Antarctic toothfish, *Dissostichus mawsoni* (Perciformes : Nototheniidae). *Polar Biology*, **26**, 124–128.

Nei M (1973) Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the United States of America*, **70**, 3321–3323.

Ng NSR, Wilton PR, Prawiradilaga DM *et al.* (2017) The effects of Pleistocene climate change on biotic differentiation in a montane songbird clade from Wallacea. *Molecular Phylogenetics and Evolution*,

**114**, 353-366.

Nicodemus-Johnson J, Silic S, Ghigliotti L, Pisano E, Cheng C-HC (2011) Assembly of the antifreeze glycoprotein/trypsinogen-like protease genomic locus in the Antarctic toothfish *Dissostichus mawsoni* (Norman). *Genomics*, **98**, 194–201.

Nicol S, Foster J, Kawaguchi S (2012) The fishery for Antarctic krill - recent developments. *Fish and Fisheries*, **13**, 30–40.

Nielsen EE, Hemmer-Hansen J, Poulsen NA *et al.* (2009) Genomic signatures of local directional selection in a high gene flow marine organism; the Atlantic cod (*Gadus morhua*). *BMC Evolutionary Biology*, **9**, 276.

Nielsen R, Paul JS, Albrechtsen A, Song YS (2011) Genotype and SNP calling from next-generation sequencing data. *Nature Reviews Genetics*, **12**, 443–451.

Niemandt C, Kovacs KM, Lydersen C *et al.* (2015) Chinstrap and macaroni penguin diet and demography at Nyroysa, Bouvetoya. *Antarctic Science*, **28**, 91–100.

Nilsson JA, Fulton EA, Haward M, Johnson C (2016) Consensus management in Antarctica's high seas – Past success and current challenges. *Marine Policy*, **73**, 172–180.

North AW (1996) Locomotory activity and behaviour of the Antarctic teleost *Notothenia coriiceps*. *Marine Biology*, **126**, 125–132.

North AW (2001) Early life history strategies of notothenioids at South Georgia. *Journal of Fish Biology*, **58**, 496–505.

Nunziata SO, Lance SL, Scott DE, Lemmon EM, Weisrock DW (2016) Genomic data detect corresponding signatures of population size change on an ecological time scale in two salamander species. *Molecular Ecology*, **26**, 1060-1074.

O'Leary SJ, Puritz JB, Willis SC, Hollenbeck CM, Portnoy DS (2018) These aren't the loci you're looking for: Principles of effective SNP filtering for molecular ecologists. *Molecular Ecology*, 27, 3193-3206.

Obregón C, Lyndon AR, Barker J *et al.* (2018) Valuing and understanding fish populations in the Anthropocene: Key questions to address. *Journal of Fish Biology*, **61**, 828–845.

Ogden R, Gharbi K, Mugue N *et al.* (2013) Sturgeon conservation genomics: SNP discovery and validation using RAD sequencing. *Molecular Ecology*, **22**, 3112–23.

Oksanen J, Blanchet FG, Friendly M *et al.* (2018) vegan: Community Ecology Package. R package version 2.4-6. https://cran.r-project.org/package=vegan

Olivar MP, Bernal A, Molí B *et al.* (2012) Vertical distribution, diversity and assemblages of mesopelagic fishes in the western Mediterranean. *Deep-Sea Research Part I: Oceanographic Research Papers*, **62**, 53–69.

Olivar MP, Hulley PA, Castellón A *et al.* (2017) Mesopelagic fishes across the tropical and equatorial Atlantic: Biogeographical and vertical patterns. *Progress in Oceanography*, **151**, 116–137.

Olsen G (1954) South Georgian cod, Fischer. *Norsk Hvalfangstidende*, **43**, 373-382.

Olsen EM, Heino M, Lilly GR *et al.* (2004) Maturation trends indicative of rapid evolution preceded the collapse of northern cod. *Nature*, **428**, 4–7.

Orsi AH, Whitworth T, Nowlin Jr WD (1995) On the meridional extent and fronts of the Antarctic Circumpolar Current. *Deep Sea Research Part I: Oceanographic Research Papers*, **42**, 641–673.

Ortuño Crespo G, Dunn DC, Gianni M *et al.* (2019) High seas fish biodiversity is slipping through the governance net. *Nature Ecology & Evolution*, **3**, 1273-1276.

Österblom H, Bodin Ö (2012) Global cooperation among diverse organizations to reduce illegal fishing in the Southern Ocean. *Conservation Biology*, **26**, 638–648.

Österblom H, Sumaila UR (2011) Toothfish crises, actor diversity and the emergence of compliance mechanisms in the Southern Ocean. *Global Environmental Change*, **21**, 972–982.

Ovenden JR, Berry O, Welch DJ, Buckworth RC, Dichmont CM (2015) Ocean's eleven: A critical evaluation of the role of population, evolutionary and molecular genetics in the management of wild fisheries. *Fish and Fisheries*, **16**, 125–159.

Ozouf-Costaz C, Hureau JC, Beaunier M (1991) Chromosome studies on fish of the suborder Notothenioidei collected in the Weddell Sea during EPOS 3 cruise. *Cybium*, **15**, 271–289.

Padilla A, Zeller D, Pauly D (2015) The fish and fisheries of Bouvet Island. In: *Marine Fisheries Catches of SubAntarctic Islands, 1950 -2010* (eds Palomares MLD, Pauly D), pp. 21–30. Fisheries Centre, University of British Columbia, Vancouver, BC.

Pakhomov EAE, Perissinotto R, McQuaid CDC (1996) Prey composition and daily rations of myctophid fishes in the Southern Ocean. *Marine Ecology Progress Series*, **134**, 1–14.

Palaiokostas C, Bekaert M, Khan MGQ *et al.* (2015) A novel sex-determining QTL in Nile tilapia (*Oreochromis niloticus*). *BMC Genomics*, **16**, 171.

Palczewski K, Kumasaka T, Hori T *et al.* (2000) Crystal structure of rhodopsin : A G protein–coupled receptor. *Science*, **289**, 739-745.

Pante E, Abdelkrim J, Viricel A *et al.* (2014) Use of RAD sequencing for delimiting species. *Heredity*, **114**, 450-459.

Pante E, Simon-Bouhet B (2013) marmap: a package for importing, plotting and analyzing bathymetric and topographic data in R. *PLoS ONE*, **8**, 6–9.

Papetti C, Harms L, Jürgens J *et al.* (2016a) Microsatellite markers for the notothenioid fish *Lepidonotothen nudifrons* and two congeneric species. *BMC Research Notes*, **9**, 238.

Papetti C, Lucassen M, Pörtner H-O (2016b) Integrated studies of organismal plasticity through physiological and transcriptomic approaches: examples from marine polar regions. *Briefings in Functional Genomics*, **15**, 365-372.

Papetti C, Pujolar JM, Mezzavilla M *et al.* (2012) Population genetic structure and gene flow patterns between populations of the Antarctic icefish *Chionodraco rastrospinosus*. *Journal of Biogeography*, **39**, 1361–1372.

Papetti C, Susana E, La Mesa M *et al.* (2007) Microsatellite analysis reveals genetic differentiation between year-classes in the icefish *Chaenocephalus aceratus* at South Shetlands and Elephant Island. *Polar Biology*, **30**, 1605–1613.

Papetti C, Susana E, Patarnello T, Zane L (2009) Spatial and temporal boundaries to gene flow between

*Chaenocephalus aceratus* populations at South Orkney and South Shetlands. *Marine Ecology Progress Series*, **376**, 269–281.

Pappalardo AM, Cuttitta A, Sardella A *et al.* (2015) DNA barcoding and COI sequence variation in Mediterranean lanternfishes larvae. *Hydrobiologia*, **749**, 155–167.

Paradis E (2010) Pegas: An R package for population genetics with an integrated-modular approach. *Bioinformatics*, **26**, 419–420.

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

Paradis E, Gosselin T, Goudet J, Jombart T, Schliep K (2016) Linking genomics and population genetics with R. *Molecular Ecology Resources*, **17**, 54–66.

Pardo-Diaz C, Salazar C, Jiggins CD (2015) Towards the identification of the loci of adaptive evolution. *Methods in Ecology and Evolution*, **6**, 445–464.

Pardo MÁ, Jiménez E, Pérez-Villarreal B (2016) Misdescription incidents in seafood sector. *Food Control*, **62**, 277–283.

Paris JR, Sherman KD, Bell E *et al.* (2018) Understanding and managing fish populations: keeping the toolbox fit for purpose: *Journal of Fish Biology*, **92**, 727–751.

Paris JR, Stevens JR, Catchen JM (2017) Lost in parameter space: a road map for stacks. *Methods in Ecology and Evolution*, **8**, 1360–1373.

Park YH, Durand I, Kestenare E *et al.* (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.

Parker SJ, Mormede S, Hanchet SM *et al.* (2019) Monitoring Antarctic toothfish in McMurdo Sound to evaluate the Ross Sea region Marine Protected Area. *Antarctic Science*, **13**, 1–13.

Parker RW, Paige KN, DeVries AL (2002) Genetic variation among populations of the Antarctic toothfish: evolutionary insights and implications for conservation. *Polar Biology*, **25**, 256–261.

Parkinson CL (2004) Southern Ocean sea ice and its wider linkages: insights revealed from models and observations. *Antarctic Science*, **16**, 387–400.

Patarnello T, Marcato S, Zane L, Varotto V, Bargelloni L (2003) Phylogeography of the *Chionodraco* genus (Perciformes, Channichthydae) in the Southern Ocean. *Molecular Phylogenetics and Evolution*, **28**, 420–429.

Pauls SU, Nowak C, Bálint M, Pfenninger M (2013) The impact of global climate change on genetic diversity within populations and species. *Molecular Ecology*, **22**, 925–946.

Paxton JR (1972) Osteology and relationships of the lanternfishes (family Myctophidae). *Bulletin of the Natural History Museum of Los Angeles County*, **13**, 1–81.

Paz-Vinas I, Loot G, Hermoso V *et al.* (2018) Systematic conservation planning for intraspecific genetic diversity. *Proceedings of the Royal Society B: Biological Sciences*, **285**, 20172746.

Pearse DE (2016) Saving the spandrels? Adaptive genomic variation in conservation and fisheries management. *Journal of Fish Biology*, **89**, 2697-2716.

Peel D, Waples RS, Esplanade C *et al.* (2013) Accounting for missing data in the estimation of contemporary genetic effective population size (Ne). *Molecular Ecology Resources*, **13**, 243–253.

Pérez-Portela R, Bumford A, Coffman B *et al.* (2018) Genetic homogeneity of the invasive lionfish across the Northwestern Atlantic and the Gulf of Mexico based on Single Nucleotide Polymorphisms. *Scientific Reports*, **8**, 5062.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE*, **7**, e37135.

Phillips SJ, Dudik M, Elith J *et al.* (2009) Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological Applications*, **19**, 181–197.

Pierce D (2019) R package "ncdf4." http://cirrus.ucsd.edu/~pierce/ncdf

Pimentel JSM, Carmo AO, Rosse IC *et al.* (2018) High-throughput sequencing strategy for microsatellite genotyping using neotropical fish as a model. *Frontiers in Genetics*, **9**, 1–8.

Pinkerton MH, Bradford-Grieve JM (2014) Characterizing foodweb structure to identify potential ecosystem effects of fishing in the Ross Sea, Antarctica. *Ices Journal of Marine Science*, **71**, 1542–1553.

Pinkerton MH, Lyver PO, Stevens DW *et al.* (2016) Increases in Adélie penguins in the Ross Sea: Could the fishery for Antarctic toothfish be responsible? *Ecological Modelling*, **337**, 262–271.

Pinkerton MH, McMillan P, Forman J *et al.* (2013) Distribution, morphology and ecology of *Macrourus whitsoni* and *M. caml* (Gadiformes, Macrouridae) in the Ross Sea region. *CCAMLR Science*, **20**, 37–61.

Pinsky ML, Palumbi SR (2014) Meta-analysis reveals lower genetic diversity in overfished populations. *Molecular Ecology*, **23**, 29–39.

Pinsky ML, Saenz-Agudelo P, Salles OC *et al.* (2017) Marine dispersal scales are congruent over evolutionary and ecological time. *Current Biology*, **27**, 149–154.

Pinsky ML, Worm B, Fogarty MJ, Sarmiento JL, Levin SA (2013) Marine taxa track local climate velocities. *Science*, **341**, 1239–42.

Plummer M, Best N, Cowles K, Vines K (2006) CODA: convergence diagnosis and output analysis for MCMC. *R News*, **6**, 7–11.

Poland JA, Brown PJ, Sorrells ME, Jannink J-L (2012) Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS ONE*, **7**, e32253.

Pollard RT, Lucas MI, Read JF (2002) Physical controls on biogeochemical zonation in the Southern Ocean. *Deep-Sea Research Part II: Topical Studies in Oceanography*, **49**, 3289–3305.

Pompanon F, Bonin A, Bellemain E, Taberlet P (2005) Genotyping errors: causes, consequences and solutions. *Nature Reviews Genetics*, **6**, 846–847.

Pompanon F, Deagle BE, Symondson WOC *et al.* (2012) Who is eating what: Diet assessment using next generation sequencing. *Molecular Ecology*, **21**, 1931–1950.

Pootakham W, Sonthirod C, Naktang C *et al.* (2016) Effects of methylation-sensitive enzymes on the enrichment of genic SNPs and the degree of genome complexity reduction in a two-enzyme genotyping-by-sequencing (GBS) approach: a case study in oil palm (*Elaeis guineensis*). *Molecular Breeding*, **36**, 154.

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

Postlethwait JH, Yan Y, Desvignes T *et al.* (2016) Embryogenesis and early skeletogenesis in the Antarctic Bullhead Notothen, *Notothenia coriiceps*. *Developmental Dynamics*, **245**, 1066–1080.

Poulsen JY, Byrkjedal I, Willassen E *et al.* (2013) Mitogenomic sequences and evidence from unique gene rearrangements corroborate evolutionary relationships of myctophiformes (Neoteleostei). *BMC Evolutionary Biology*, **13**, 111.

Poynton HC, Hasenbein S, Benoit JB *et al.* (2018) The toxicogenome of *Hyalella azteca*: A model for sediment ecotoxicology and evolutionary toxicology. *Environmental Science and Technology*, **52**, 6009–6022.

Prirodina VP, Balushkin A V. (2015) First record of *Muraenolepis evseenkoi* in the Weddell Sea and identification key to the eel cods (Muraenolepididae) of continental seas of Antarctica. *Journal of Ichthyology*, **55**, 778–782.

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

Pritchard JK, Wen W. (2003). Documentation for STRUCTURE Software: Version 2. Chicago: University of Chicago Press. Retrived from: http://web.stanford.edu/group/pritchardlab/software/structure2_1.html.

Proud R, Cox MJ, Handegard NO, Kloser RJ, Brierley AS (2018) From siphonophores to deep scattering layers: an estimation of global mesopelagic fish biomass. *ICES Journal of Marine Science*, **76**, 718-733.

Provoost P, Bosch S (2019) robis: R Client to access data from the OBIS API. https://cran.r-project.org/package=robis

Puckett EE (2016) Variability in total project and per sample genotyping costs under varying study designs including with microsatellites or SNPs to answer conservation genetic questions. *Conservation Genetics Resources*, **9**, 289-304.

Puncher GN, Cariani A, Maes GE *et al.* (2018) Spatial dynamics and mixing of bluefin tuna in the Atlantic Ocean and Mediterranean Sea revealed using next generation sequencing. *Molecular Ecology Resources*, **18**, 620-638.

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

Puritz JB, Lotterhos KE (2018) Expressed exome capture sequencing: A method for cost-effective exome sequencing for all organisms. *Molecular Ecology Resources*, **18**, 1209–1222.

Puritz JB, Matz M V., Toonen RJ *et al.* (2014b) Demystifying the RAD fad. *Molecular Ecology*, **23**, 5937-

5942.

Pusch C, Hulley PA, Kock KH (2004) Community structure and feeding ecology of mesopelagic fishes in the slope waters of King George Island (South Shetland Islands, Antarctica). *Deep-Sea Research Part I: Oceanographic Research Papers*, **51**, 1685–1708.

Putman AI, Carbone I (2014) Challenges in analysis and interpretation of microsatellite data for population genetic studies. *Ecology and Evolution*, **4**, 4399–4428.

Querejeta M, González-Esteban J, Gómez A *et al.* (2016) Genomic diversity and geographical structure of the Pyrenean desman. *Conservation Genetics*, **17**, 1333-1344.

R Core Team (2016) R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing*, v3.3.1. http://www.r-project.org/

R Core Team (2019) R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing*, v3.6.1. http://www.r-project.org/

Raeymaekers JAM, Chaturvedi A, Hablützel PI *et al.* (2017) Adaptive and non-adaptive divergence in a common landscape. *Nature Communications*, **8**, 267.

Ratnasingham S, Hebert PDN (2007) BOLD: The Barcode of Life Data System (http://www.barcodinglife.org). *Molecular Ecology Notes*, **7**, 355–364.

Ratnasingham S, Hebert PDN (2013) A DNA-Based Registry for All Animal Species: The Barcode Index Number (BIN) System. *PLoS ONE*, **8**.

Ravinet M, Westram A, Johannesson K *et al.* (2015) Shared and nonshared genomic divergence in parallel ecotypes of *Littorina saxatilis* at a local scale. *Molecular Ecology*, 287–305.

Ray J (1686) *Historia Plantarum*. London.

Razgour O, Forester B, Taggart JB *et al.* (2019) Considering adaptive genetic variation in climate change vulnerability assessment reduces species range loss projections. *Proceedings of the National Academy of Sciences of the United States of America*, **116**, 10418–10423.

Razgour O, Taggart JB, Manel S *et al.* (2018) An integrated framework to identify wildlife populations under threat from climate change. *Molecular Ecology Resources*, **18**, 18–31.

Rebelein A, Pörtner H-O, Bock C (2018) Untargeted metabolic profiling reveals distinct patterns of thermal sensitivity in two related notothenioids. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, **217**, 43–54.

Recknagel H, Elmer KR, Meyer A (2013) A hybrid genetic linkage map of two ecologically and morphologically divergent Midas cichlid fishes (*Amphilophus* spp.) obtained by massively parallel DNA sequencing (ddRADSeq). *G3: Genes, Genomes, Genetics*, **3**, 65–74.

Rees DJ, Dufresne F, Glémet H, Belzile C (2007) Amphipod genome sizes: first estimates for Arctic species reveal genomic giants. *Genome*, **50**, 151–158.

Reid NM, Proestou DA, Clark BW *et al.* (2016) The genomic landscape of rapid repeated evolutionary adaptation to toxic pollution in wild fish. *Science*, **354**, 1305-1308.

Reilly A, Ward RD (1999) Microsatellite loci to determine population structure of the Patagonian toothfish *Dissostichus eleginoides*. *Molecular Ecology*, **8**, 1753–1768.

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

Reitzel AM, Herrera S, Layden MJ, Martindale MQ, Shank TM (2013) Going where traditional markers have not gone before: Utility of and promise for RAD sequencing in marine invertebrate phylogeography and population genomics. *Molecular Ecology*, **22**, 2953–2970.

Rellstab C, Gugerli F, Eckert AJ *et al.* (2015) A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, **24**, 4348–4370.

Rembiszewski JM, Krzeptowski M, Linkowski TB (1978) Fishes (Pisces) as by-catch in fisheries of krill *Euphausia superba* Dana (Euphausiacea, Crustacea). *Polskie Archiwum Hydrobiologii*, **25**, 677–695.

Renner AHH, Heywood KJ, Thorpe SE (2009) Validation of three global ocean models in the Weddell Sea. *Ocean Modelling*, **30**, 1–15.

Renner AHH, Thorpe SE, Heywood KJ *et al.* (2012) Advective pathways near the tip of the Antarctic Peninsula: Trends, variability and ecosystem implications. *Deep-Sea Research Part I: Oceanographic Research Papers*, **63**, 91–101.

Reusch TBH, Ehlers A, Hämmerli A, Worm B (2005) Ecosystem recovery after climatic extremes enhanced by genotypic diversity. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 2826–2831.

Rey O, Bonillo C, Gallut C *et al.* (2011) Naked dragonfishes *Gymnodraco acuticeps* and *G. victori* (Bathydraconidae, Notothenioidei) off Terre Adélie are a single species. *Cybium*, **35**, 111–119.

Riginos C, Crandall ED, Liggins L *et al.* (2016) Navigating the currents of seascape genomics: How spatial analyses can augment population genomic studies. *Current Zoology*, **62**, 581–601.

Rintoul SR, Chown SL, DeConto RM *et al.* (2018) Choosing the future of Antarctica. *Nature*, **558**, 233–241.

Rintoul SR, Hughes C, Olbers D (2001) The Antarctic Circumpolar Current System. In: *Ocean Circulation and Climate*, pp. 271–302. Academic Press, New York.

Rintoul SR, Naveira Garabato AC (2013). Dynamics of the Southern Ocean circulation. In: *Ocean Circulation and Climate* (2nd ed.), pp. 471–492. Elsevier Ltd.

Ritchie H, Jamieson AJ, Piertney SB (2017) Genome size variation in deep-sea amphipods. *Royal Society Open Science*, **4**, 170862.

Rocha LA, Bowen BW (2008) Speciation in coral-reef fishes. *Journal of Fish Biology*, **72**, 1101–1121.

Roche K, Kuta J, Sedlá I, Rostislav Č, Jurajda P (2019) Concentrations of thirteen trace metals in scales of three nototheniid fishes from Antarctica (James Ross Island , Antarctic Peninsula). *Biological Trace Element Research*, **191**, 214-223.

Rochette NC, Catchen JM (2017) Deriving genotypes from RAD-seq short-read data using Stacks. *Nature Protocols*, **12**, 2640–2659.

Rochette NC, Rivera-Colon AG, Catchen JM (2019) STACKS 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology*, **28**, 4737-4754.

Rock J, Costa FO, Walker DI *et al.* (2008) DNA barcodes of fish of the Scotia Sea, Antarctica indicate priority groups for taxonomic and systematics focus. *Antarctic Science*, **20**, 253–262.

Rodrigues E, Feijo-Oliveira M, Suda CNK *et al.* (2015) Metabolic responses of the Antarctic fishes *Notothenia rossii* and *Notothenia coriiceps* to sewage pollution. *Fish Physiology and Biochemistry*, **41**, 1205–1220.

Rodríguez-Ezpeleta N, Bradbury IR, Mendibil I *et al.* (2016) Population structure of Atlantic mackerel inferred from RAD-seq-derived SNP markers: Effects of sequence clustering parameters and hierarchical SNP selection. *Molecular Ecology Resources*, **16**, 991–1001.

Rogers AD (2007) Evolution and biodiversity of Antarctic organisms: A molecular perspective. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **362**, 2191–2214.

Rogers AD, Frinault BAV, Barnes DKA *et al.* (2020) Antarctic futures: An assessment of climate-driven changes in ecosystem structure, function, and service provisioning in the Southern Ocean. *Annual Review of Marine Science*, **12**, 87-120.

Rogers AD, 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**, 915–924.

Ropert-Coudert Y, Chiaradia A, Ainley D *et al.* (2019) Happy feet in a hostile world? The future of penguins depends on proactive management of current and expected threats. *Frontiers in Marine Science*, **6**.

Rutschmann S, Matschiner M, Damerau M *et al.* (2011) Parallel ecological diversification in Antarctic notothenioid fishes as evidence for adaptive radiation. *Molecular Ecology*, **20**, 4707–21.

Ruud JT (1954) Vertebrates without erythrocytes and blood pigment. *Nature*, **173**, 848–50.

Ryman N, Utter F, Laikre L (1995) Protection of intraspecific biodiversity of exploited fishes. *Reviews in Fish Biology and Fisheries*, **5**, 417–446.

Salas L, Nur N, Ainley D *et al.* (2016) Coping with the loss of large, energy-dense prey: A potential bottleneck for Weddell Seals in the Ross Sea. *Ecological Applications*, **27**, 10–25.

Salmon RA, Roop HA (2019) Bridging the gap between science communication practice and theory: Reflecting on a decade of practitioner experience using polar outreach case studies to develop a new framework for public engagement design. *Polar Record*, **55**, 297-310.

Sánchez-Tójar A, Nakagawa S, Sánchez-Fortún M *et al.* (2018) Meta-analysis challenges a textbook example of status signalling and demonstrates publication bias. *eLife*, **7**, 1–26.

Santora JA (2013) Dynamic intra-seasonal habitat use by Antarctic fur seals suggests migratory hotspots near the Antarctic Peninsula. *Marine Biology*, **160**, 1383–1393.

Sapota MR (1999) Gonad development and embryogenesis of *Notothenia coriiceps* from South Shetlands - Antarctica. *Polar Biology*, **22**, 164–168.

Saunders RA, Collins MA, Foster E *et al.* (2014) The trophodynamics of Southern Ocean *Electrona* (Myctophidae) in the Scotia Sea. *Polar Biology*, **37**, 789–807.

Saunders RA, Collins MA, Stowasser G, Tarling GA (2017) Southern Ocean mesopelagic fish communities in the Scotia Sea are sustained by mass immigration. *Marine Ecology Progress Series*, **569**, 173–185.

Saunders RA, Collins MA, Ward P *et al.* (2015a) Predatory impact of the myctophid fish community on zooplankton in the Scotia Sea (Southern Ocean). *Marine Ecology Progress Series*, **541**, 45–64.

Saunders RA, Collins MA, Ward P *et al.* (2015b) Distribution, population structure and trophodynamics of Southern Ocean *Gymnoscopelus* (Myctophidae) in the Scotia Sea. *Polar Biology*, **38**, 287–308.

Saunders RA, Collins MA, Ward P *et al.* (2015c) Trophodynamics of *Protomyctophum* (Myctophidae) in the Scotia Sea (Southern Ocean). *Journal of Fish Biology*, **87**, 1031–1058.

Saunders RA, Hill SL, Tarling GA, Murphy EJ (2019) Myctophid fish (Family Myctophidae) are central consumers in the food web of the Scotia Sea (Southern Ocean). *Frontiers in Marine Science*, **6**, 530.

Saunders RA, Lourenço S, Vieira RP *et al.* Age and growth of Brauer's lanternfish *Gymnoscopelus braueri* and rhombic lanternfish *Krefftichthys anderssoni* (Family Myctophidae) in the Scotia Sea, Southern Ocean. *Journal of Fish Biology*.

Savolainen O, Lascoux M, Merilä J (2013) Ecological genomics of local adaptation. *Nature Reviews Genetics*, **14**, 807–20.

Schiaparelli S, Danis B, Wadley V, Stoddart DM (2013) The census of Antarctic marine life: The first available baseline for Antarctic marine biodiversity. In: *Adaptation and Evolution in Marine Environments, Volume 2* (eds Verde C, di Prisco G), pp. 3–19. Springer-Verlag, Berlin, Heidelberg.

Schliep KP (2011) phangorn: Phylogenetic analysis in R. *Bioinformatics*, **27**, 592–593.

Schmid S, Genevest R, Gobet E *et al.* (2016) HyRAD-X, a versatile method combining exome capture and RAD sequencing to extract genomic information from ancient DNA. *Methods in Ecology and Evolution*, **8**, 1374-1388.

Schneppenheim R, Kock K-H, Duhamel G, Janssen G (1994) On the taxonomy of the *Lepidonotothen squamifrons* group (Pisces, Perciformes, Notothenioidei). *Archive of Fishery and Marine Research*, **422**, 137–148.

Schwarzhans W, Mörs T, Engelbrecht A, Reguero M, Kriwet J (2017) Before the freeze: otoliths from the Eocene of Seymour Island, Antarctica, reveal dominance of gadiform fishes (Teleostei). *Journal of Systematic Palaeontology*, **15**, 147–170.

Schweyen H, Rozenberg A, Leese F (2014). Detection and removal of PCR duplicates in population genomic ddRAD studies by addition of a degenerate base region (DBR) in sequencing adapters. *Biological Bulletin*, **227**, 146–160.

Seefeldt MA, Weigand AM, Havermans C, Moreira E, Held C (2017) Fishing for scavengers: an integrated study to amphipod (Crustacea: Lysianassoidea) diversity of Potter Cove (South Shetland Islands, Antarctica). *Marine Biodiversity*, **48**, 2081-2104.

Selkoe KA, Toonen RJ (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters*, **9**, 615–629.

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

Shafer ABA, Peart CR, Tusso S *et al.* (2017) Bioinformatic processing of RAD-seq data dramatically

impacts downstream population genetic inference. *Methods in Ecology and Evolution*, **8**, 907–917.

Shafer ABA, Wolf JBW, Alves PC *et al.* (2015) Genomics and the challenging translation into conservation practice. *Trends in Ecology and Evolution*, **30**, 78–87.

Shaw PW, Arkhipkin AI, 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**, 3293–303.

Shendure J, Ji H (2008) Next-generation DNA sequencing. *Nature Biotechnology*, **26**, 1135–1145.

Shin SC, Ahn DH, Kim SJ *et al.* (2014) The genome sequence of the Antarctic bullhead notothen reveals evolutionary adaptations to a cold environment. *Genome Biology*, **15**, 468.

Sidell BD, Brien KMO (2006) When bad things happen to good fish : the loss of hemoglobin and myoglobin expression in Antarctic icefishes. *The Journal of Experimental Biology*, **209**, 1791–1802.

Smedbol RK, McPherson A, Hansen MM, Kenchington E (2002) Myths and moderation in marine "metapopulations"? *Fish and Fisheries*, **3**, 20–35.

Smith ADM, Brown CJ, Bulman CM *et al.* (2011a) Impacts of fishing low-trophic level species on marine ecosystems. *Science*, **333**, 1147–1151.

Smith P, Gaffney P (2000) Toothfish stock structure revealed with DNA methods. *Water & Atmosphere*, **8**, 17–18.

Smith PJ, Gaffney PM (2005) Low genetic diversity in the Antarctic Toothfish (*Dissostichus mawsoni*) observed with mitochondrial and intron DNA markers. *CCAMLR Science*, **12**, 43–51.

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.

Smith PJ, Steinke D, Dettai A *et al.* (2012) DNA barcodes and species identifications in Ross Sea and Southern Ocean fishes. *Polar Biology*, **35**, 1297–1310.

Smith PJ, Steinke D, McMillan PJ *et al.* (2011b) DNA barcoding highlights a cryptic species of grenadier *Macrourus* in the Southern Ocean. *Journal of Fish Biology*, **78**, 355–365.

Smith PJ, Steinke D, Mcveagh SM *et al.* (2008) Molecular analysis of Southern Ocean skates (*Bathyraja*) reveals a new species of Antarctic skate. *Journal of Fish Biology*, **73**, 1170–1182.

Spodareva V V., Balushkin A V. (2014) Description of a new species of plunderfish of genus *Pogonophryne* (Perciformes: Artedidraconidae) from the Bransfield Strait (Antarctica) with a key for the identification of species of the group "marmorata." *Journal of Ichthyology*, **54**, 1–6.

St. John MA, Borja A, Chust G *et al.* (2016). A dark hole in our understanding of marine ecosystems and their services: perspectives from the mesopelagic community. *Frontiers in Marine Science*, **3**, 31.

Star B, Boessenkool S, Gondek AT *et al.* (2017) Ancient DNA reveals the Arctic origin of Viking Age cod from Haithabu, Germany. *Proceedings of the National Academy of Sciences*, **114**, 9152-9157.

Star B, Hansen MH, Skage M *et al.* (2016) Preferential amplification of repetitive DNA during whole genome sequencing library creation from historic samples. *STAR: Science & Technology of Archaeological Research*, **2**, 36–45.

Stein DL (2012) Snailfishes (Family Liparidae) of the Ross Sea, Antarctica, and closely adjacent waters.

*Zootaxa*, **120**, 1–120.

Steinke D, Hanner R (2011) The FISH-BOL collaborators' protocol. *Mitochondrial DNA*, **22**, 10–14.

Stokstad E (2010) Behind the eco-label, a debate over Antarctic toothfish. *Science*, **329**, 1596–1597.

Storey JD, Bass AJ, Dabney A, Robinson D (2019) qvalue: Q-value estimation for false discovery rate control. http://github.com/jdstorey/qvalue

Strobel A, Schmid P, Burkhardt-Holm P, Segner H, Zennegg M (2018) Persistent organic pollutants in red- and white-blooded High-Antarctic notothenioid fish from the remote Weddell Sea. *Chemosphere*, **193**, 213–222.

Strugnell JM, Allcock AL, Watts PC (2017) Closely related octopus species show different spatial genetic structures in response to the Antarctic seascape. *Ecology and Evolution*, **7**, 8087–8099.

Suchan T, Pitteloud C, Gerasimova NS *et al.* (2016) Hybridization capture using RAD probes (hyRAD), a new tool for performing genomic analyses on collection specimens. *PLoS ONE*, **11**, e0151651.

Sun J, Zhang Y, Xu T *et al.* (2017) Adaptation to deep-sea chemosynthetic environments as revealed by mussel genomes. *Nature Ecology and Evolution*, **1**, 1–7.

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

Suprenand PM, Ainsworth CH (2017) Trophodynamic effects of climate change-induced alterations to primary production along the western Antarctic Peninsula. *Marine Ecology Progress Series*, **569**, 37–54.

Swain DP, Sinclair AF, Mark Hanson J (2007) Evolutionary response to size-selective mortality in an exploited fish population. *Proceedings of the Royal Society B: Biological Sciences*, **274**, 1015–22.

Sykora-Bodie ST, Morrison TH (2019) Drivers of consensus-based decision-making in international environmental regimes: Lessons from the Southern Ocean. *Aquaric Conservation: Marine and Freshwater Ecosystems*, **29**, 2147–2161.

Sylvester ZT, Brooks CM (2019) Protecting Antarctica through co-production of actionable science: Lessons from the CCAMLR marine protected area process. *Marine Policy*, **111**, 103720.

Tautz D, Arctander P, Minelli A, Thomas RH, Vogler AP (2003) A plea for DNA taxonomy. *Trends in Ecology and Evolution*, **18**, 70–74.

Teacher AGF, André C, Jonsson PR, Merilä J (2013) Oceanographic connectivity and environmental correlates of genetic structuring in Atlantic herring in the Baltic Sea. *Evolutionary Applications*, **6**, 549–567.

Teacher AG, André C, Merilä J, Wheat CW (2012) Whole mitochondrial genome scan for population structure and selection in the Atlantic herring. *BMC Evolutionary Biology*, **12**, 248.

Teschke K, Beaver D, Bester MN *et al.* (2015) *Scientific background document in support of the development of a CCAMLR MPA in the Weddell Sea (Antarctica) - Version 2015 - Part A: General context of the establishment of MPAs and background information on the Weddell Sea MPA planning area*.

Teschke K, Pehlke H, Siegel V *et al.* (2019) An integrated data compilation for the development of a

marine protected area in the Weddell Sea. *Earth System Science Data Discussions*, **2012**, 1–31.

Therkildsen NO, Palumbi SR (2017) Practical low-coverage genomewide sequencing of hundreds of individually barcoded samples for population and evolutionary genomics in nonmodel species. *Molecular Ecology Resources*, **17**, 194–208.

Therkildsen NO, Wilder AP, Conover DO *et al.* (2019) Contrasting genomic shifts underlie parallel phenotypic evolution in response to fishing. *Science*, **365**, 487–490.

Thiel R, Knebelsberger T (2016) How reliably can northeast Atlantic sand lances of the genera *Ammodytes* and *Hyperoplus* be distinguished? A comparative application of morphological and molecular methods. *ZooKeys*, **2016**, 139–164.

Thompson AF (2008). The atmospheric ocean: Eddies and jets in the Antarctic Circumpolar Current. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **366**, 4529–4541.

Thompson AF, Stewart AL, Spence P, Heywood KJ (2018) The Antarctic Slope Current in a Changing Climate. *Reviews of Geophysics*, **56**, 741–770.

Thomsen PF, Willerslev E (2015) Environmental DNA - An emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation*, **183**, 4–18.

Thornhill DJ, Mahon AR, Norenburg JL, Halanych KM (2008) Open-ocean barriers to dispersal: A test case with the Antarctic Polar Front and the ribbon worm *Parborlasia corrugatus* (Nemertea: Lineidae). *Molecular Ecology*, **17**, 5104–5117.

Tigano A, Friesen VL (2016) Genomics of local adaptation with gene flow. *Molecular Ecology*, **25**, 2144–2164.

Toomey L, Welsford D, Appleyard SA *et al.* (2016) Genetic structure of Patagonian toothfish populations from otolith DNA. *Antarctic Science*, **14**, 1–14.

Toonen RJ, Puritz JB, Forsman ZH *et al.* (2013) ezRAD: a simplified method for genomic genotyping in non-model organisms. *PeerJ*, **1**, e203.

Torkamaneh D, Laroche J, Belzile F (2016) Genome-wide SNP calling from genotyping by sequencing (GBS) data: A comparison of seven pipelines and two sequencing technologies. *PLoS ONE*, **11**, 1–14.

Trathan PN, Collins MA, Grant SM *et al.* (2014) The South Georgia and the South Sandwich Islands MPA: Protecting A Biodiverse Oceanic Island Chain Situated in the Flow of the Antarctic Circumpolar Current. In: *Advances in Marine Biology*, **69**, 15-78. Elsevier Ltd.

Trathan PN, Grant SM (2020) The South Orkney Islands Southern Shelf Marine Protected Area : towards the establishment of marine spatial protection within international waters in the Southern Ocean. In: *Marine Protected Areas*, pp. 67–98. Elsevier Ltd.

Tulloch VJD, Plagányi ÉE, Matear R, Brown CJ, Richardson AJ (2018) Ecosystem modelling to quantify the impact of historical whaling on Southern Hemisphere baleen whales. *Fish and Fisheries*, **19**, 117–137.

Turner J, Colwell SR, Marshall GJ *et al.* (2005) Antarctic climate change during the last 50 years. *International Journal of Climatology*, **25**, 279–294.

Turner J, Lu H, White I *et al.* (2016) Absence of 21st century warming on Antarctic Peninsula consistent with natural variability. *Nature*, **535**, 411–415.

Tyberghein L, Verbruggen H, Pauly K *et al.* (2012) Bio-ORACLE: a global environmental dataset for marine species distribution modelling. *Global Ecology and Biogeography* **21,** 272–281.

United Nations (1992) *Convention on Biological Diversity*.

Van de Putte AP, Van Houdt JKJ, Maes GE *et al.* (2012a) High genetic diversity and connectivity in a common mesopelagic fish of the Southern Ocean: The myctophid *Electrona antarctica*. *Deep Sea Research Part II: Topical Studies in Oceanography*, **59–60**, 199–207.

Van de Putte AP, Jackson GD, Pakhomov E, Flores H, Volckaert FAM (2010) Distribution of squid and fish in the pelagic zone of the Cosmonaut Sea and Prydz Bay region during the BROKE-West campaign. *Deep Sea Research Part II: Topical Studies in Oceanography*, **57**, 956–967.

Van de Putte AP, Janko K, Kasparova E *et al.* (2012b) Comparative phylogeography of three trematomid fishes reveals contrasting genetic structure patterns in benthic and pelagic species. *Marine Genomics*, **8**, 23–34.

Van Houdt JKJ, Hellemans B, Van de Putte A, Koubbi P, Volckaert FAM (2006) Isolation and multiplex analysis of six polymorphic microsatellites in the Antarctic notothenioid fish, *Trematomus newnesi*. *Molecular Ecology Notes*, **6**, 157–159.

Vandamme SG, Maes GE, Raeymaekers JAM *et al.* (2014) Regional environmental pressure influences population differentiation in turbot (*Scophthalmus maximus*). *Molecular Ecology*, **23**, 618–36.

Vartia S, Villanueva-Cañas JL, Finarelli J *et al.* (2016) A novel method of microsatellite genotyping-by-sequencing using individual combinatorial barcoding. *Royal Society Open Science*, **3**, 150565.

Vaughan DG, Marshall GJ, Connolley WM *et al.* (2003) Recent rapid regional climate warming on the Antarctic Peninsula. *Climatic Change*, **60**, 243–274.

Vedishcheva E V, Korostelev NB, Gordeev II, Orlov AM (2019) A first attempt to evaluate the age and growth of blue hake *Antimora rostrata* (Moridae, Gadiformes, Teleostei) from the Lazarev and Weddell seas (Antarctic). *Polar Record*, **55**, 25–31.

Venables WN, Ripley BD (2002) *Modern Applied Statistics with S*. Springer, New York.

Volckaert FAM, Rock J, Van de Putte AP (2012) Connectivity and Molecular Ecology of Antarctic Fishes. In: *Adaptation and Evolution in Marine Environments, Volume 1* (eds di Prisco G, Verde C), pp. 75–96. Springer, Berlin, Heidelberg.

Waller CL, Griffiths HJ, Waluda CM *et al.* (2017) Microplastics in the Antarctic marine system: An emerging area of research. *Science of the Total Environment*, **598**, 220–227.

Walsworth TE, Schindler DE, Colton MA *et al.* (2019) Management for network diversity speeds evolutionary adaptation to climate change. *Nature Climate Change*, **9**, 632-636.

Wang J, Zhang G, Fang X *et al.* (2012) The oyster genome reveals stress adaptation and complexity of shell formation. *Nature*, **490**, 49–54.

Waples RS (2006) A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. *Conservation Genetics*, **7**, 167–184.

Waples RS, Do C (2008) LDNE: A program for estimating effective population size from data on linkage disequilibrium. *Molecular Ecology Resources*, **8**, 753–756.

Waples RS, Gaggiotti O (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology*, **15**, 1419–39.

Waples RS, Punt AE, Cope JM (2008) Integrating genetic data into management of marine resources: how can we do it better? *Fish and Fisheries*, **9**, 423–449.

Ward RD, Hanner R, Hebert PDN (2009) The campaign to DNA barcode all fishes, FISH-BOL. *Journal of Fish Biology*, **74**, 329–56.

Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN (2005) DNA barcoding Australia's fish species. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, **360**, 1847–1857.

Watters GM, Hinke JT, Reiss CS (2020). Long-term observations from Antarctica demonstrate that mismatched scales of fisheries management and predator-prey interaction lead to erroneous conclusions about precaution. *Scientific Reports*, **10**, 2314.

Watson R a, Cheung WWL, Anticamara J a *et al.* (2013) Global marine yield halved as fishing intensity redoubles. *Fish and Fisheries*, **14**, 493–503.

Weir BS, Cockerham CC (1984) Estimating F-Statistics for the Analysis of Population Structure. *Evolution*, **38**, 1358–1370.

Wellenreuther M, Bernatchez L (2018) Eco-evolutionary genomics of chromosomal inversions. *Trends in Ecology & Evolution*, **33**, 427–440.

White C, Selkoe KA, Watson J *et al.* (2010) Ocean currents help explain population genetic structure. *Proceedings of the Royal Society B*, **277**, 1685–94.

Whiteley AR, Fitzpatrick SW, Funk WC, Tallmon DA (2014) Genetic rescue to the rescue. *Trends in Ecology & Evolution*, **30**, 42–49.

Whitlock MC, Lotterhos KE (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**, S24-36.

Wiencke C, Amsler CD, Clayton MN (2014) Chapter 5.1 Macroalgae. In: *Biogeographic Atlas of the Southern Ocean* (eds De Broyer C, Koubbi P, Griffiths H, et al.), pp. 66–73. Scientific Committee on Antarctic Research, Cambridge.

Williams R, Smolenski AJ, White RWG (1994) Mitochondrial DNA variation of *Champsocephalus gunnari* Lonnberg (Pisces: Channichthyidae) stocks on the Kerguelen Plateau, southern Indian Ocean. *Antarctic Science*, **6**, 347–352.

Willis S, Hollenbeck C, Puritz JB, Gold J, Portnoy D (2017) Haplotyping RAD loci: an efficient method to filter paralogs and account for physical linkage. *Molecular Ecology Resources*, **17**, 26–39.

Winter DJ (2012) MMOD: An R library for the calculation of population differentiation statistics. *Molecular Ecology Resources*, **12**, 1158–1160.

Wisner R (1976) *The Taxonomy and Distribution of Lanternfishes (Family Myctophidae) of the Eastern*

*Pacific Ocean - NORDA-Report 3*. Bay St. Louis, MS.

Wong EHK, Hanner RH (2008) DNA barcoding detects market substitution in North American seafood. *Food Research International*, **41**, 828–837.

Wright PJ, Régnier T, Gibb FM, Augley J, Devalla S (2018) Identifying stock structuring in the sandeel, *Ammodytes marinus*, from otolith microchemistry. *Fisheries Research*, **199**, 19–25.

Xavier JC, Fugmann G, Beck I, Huffman L, Jensen E (2016) Education on biodiversity in the polar regions. In: *Biodiversity and Education for Sustainable Development. World Sustainability Series.* (eds Castro P, Azeiteiro U, Bacelar-Nicolau P, Filho WL, Azul A), pp. 43–57. Springer Publishing Group International, Switzerland.

Xiao N (2017) ggsci: scientific journal and sci-fi themed color palettes for "ggplot2". R package version 2.8. https://cran.r-project.org/package=ggsci

Xu Q, Cai C, Hu X *et al.* (2015) Evolutionary suppression of erythropoiesis via the modulation of TGF-β signalling in an Antarctic icefish. *Molecular Ecology*, **24**, 4664–4678.

Xuereb A, Benestan L, Normandeau É *et al.* (2018) Asymmetric oceanographic processes mediate connectivity and population genetic structure, as revealed by RADseq, in a highly dispersive marine invertebrate (*Parastichopus californicus*). *Molecular Ecology*, **27**, 2347–2364.

Yokoyama S, Yokoyama R (1996) Adaptive evolution of photoreceptors and visual pigments in vertebrates. *Annual Review of Ecology and Systematics*, **27**, 543–567.

Yoon M, Jo MR, Kim PH *et al.* (2018) Total and methyl mercury concentrations in Antarctic toothfish (*Dissostichus mawsoni*): Health risk assessment. *Bulletin of Environmental Contamination and Toxicology*, **100**, 748-753.

Yoon T-H, Kang H-E, Lee SR *et al.* (2017) Metabarcoding analysis of the stomach contents of the Antarctic Toothfish (*Dissostichus mawsoni*) collected in the Antarctic Ocean. *PeerJ*, **5**, e3977.

Young EF, Belchier M, Hauser L *et al.* (2015) Oceanography and life history predict contrasting genetic population structure in two Antarctic fish species. *Evolutionary Applications*, **8**, 486–509.

Young EF, Rock J, Meredith M *et al.* (2012) Physical and behavioural influences on larval fish retention: contrasting patterns in two Antarctic fishes. *Marine Ecology Progress Series*, **465**, 201–215.

Young EF, Tysklind N, Meredith MP *et al.* (2018) Stepping stones to isolation: impacts of a changing climate on the connectivity of fragmented fish populations. *Evolutionary Applications*, **11**, 978-994.

Younger JL, Clucas GV, Kao D *et al.* (2017). The challenges of detecting subtle population structure and its importance for the conservation of emperor penguins. *Molecular Ecology*, **26**, 3883–3897.

Zachos FE (2018) (New) Species concepts, species delimitation and the inherent limitations of taxonomy. *Journal of Genetics*, **97**, 811-815.

Zachos J, Pagani M, Sloan L, Thomas E, Billups K (2001) Trends, rhythms, aberrations to present in global climate 65 ma to present. *Science*, **292**, 686–693.

Zahuranec B, Karuppasamy PK, Valinassab T *et al.* (2012) Cryptic speciation in the mesopelagic environment: Molecular phylogenetics of the lanternfish genus *Benthosema*. *Marine Genomics*, **7**, 7–10.

Zamzow JP, Aumack CF, Amsler CD *et al.* (2010) Gut contents and stable isotope analyses of the Antarctic fish, *Notothenia coriiceps* (Richardson), from two macroalgal communities. *Antarctic Science*, **23**, 107–116.

Zane L, Bargelloni L, Patarnello T (2002) Strategies for microsatellite isolation: a review. *Molecular Ecology*, **11**, 1–16.

Zane L, Marcato S, Bargelloni L *et al.* (2006) Demographic history and population structure of the Antarctic silverfish *Pleuragramma antarcticum*. *Molecular Ecology*, **15**, 4499–511.

Zhang G, Li C, Li Q *et al.* (2014) Comparative genomics reveals insights into avian genome evolution and adaptation. *Science*, **346**, 1311–1320.

Zhu F, Cui QQ, Hou ZC (2016) SNP discovery and genotyping using Genotyping-by-Sequencing in Pekin ducks. *Scientific Reports*, **6**, 1–5.

Zhuang X, Yang C, Murphy KR, Cheng C-HC (2019) Molecular mechanism and history of non-sense to sense evolution of antifreeze glycoprotein gene in northern gadids. *Proceedings of the National Academy of Sciences*, **116**, 4400–4405.

## LIST OF PUBLICATIONS

Salvanes AG V., Christiansen H., Taha Y., Henseler C., Seivag ML., Kjesbu OS., Folkvord A., Utne-Palm AC., Currie B., Ekau W., Van der Plas AK., Gibbons MJ. (2018). Variation in growth, morphology and reproduction of the bearded goby (*Sufflogobius bibarbatus*) in varying oxygen environments of northern Benguela. *Journal of Marine Systems*, *188*, 81-97. doi:10.1016/j.jmarsys.2018.04.003.

Heindler FM., Christiansen H., Frederich B., Dettai A., Lepoint G., Maes GE., Van de Putte AP., Volckaert FAM. (2018). Historical DNA Metabarcoding of the Prey and Microbiome of Trematomid Fishes Using Museum Samples. *Frontiers in Ecology and Evolution*, *6*, 151. doi:10.3389/fevo.2018.00151.

Christiansen H., Dettai A., Heindler FM., Collins MA., Duhamel G., Hautecoeur M., Steinke D., Volckaert FA M., Van de Putte AP. (2018). Diversity of Mesopelagic Fishes in the Southern Ocean - A Phylogeographic Perspective Using DNA Barcoding. *Frontiers in Ecology and Evolution*, *6*, 120. doi:10.3389/fevo.2018.00120.

Christiansen H., Fournier N., Hellemans B., Volckaert FAM. (2018). Seafood substitution and mislabeling in Brussels' restaurants and canteens. *Food Control*, *85*, 66-75. doi:10.1016/j.foodcont.2017.09.005.

Obregon C., Lyndon AR., Barker J., Christiansen H., Godley BJ., Kurland S., Piccolo JJ., Potts R., Short R., Tebb A., Mariani S. (2018). Valuing and understanding fish populations in the Anthropocene: key questions to address. *Journal of Fish Biology*, *92* (3), 828-845. doi:10.1111/jfb.13536.

Faust E., Andre C., Meurling S., Kochmann J., Christiansen H., Jensen LF., Charrier G., Laugen AT., Strand A. (2017). Origin and route of establishment of the invasive Pacific oyster *Crassostrea gigas* in Scandinavia. *Marine Ecology Progress Series*, *575*, 95-105. doi:10.3354/meps12219

## AUTHOR AFFILIATIONS AND ADDRESSES

**Esteban Barrera-Oro**
Instituto Antártico Argentino
25 de Mayo 1143, San Martín
Buenos Aires, Argentina

**Martin A. Collins**
Centre for Environment, Fisheries & Aquaculture Science
Lowestoft, United Kingdom

**Bruno Danis**
Université Libre de Bruxelles
Marine Biology Laboratory
Brussels, Belgium

**Agnès Dettaï**
Muséum National d'Histoire Naturelle
Département Systématique et Évolution
MR 7205 ISYEB Paris, France

**Guy Duhamel**
Muséum National d'Histoire Naturelle
Département Milieux et Peuplements Aquatiques
UMR 7209 Paris, France

**Charlène Guillaumot**
Université Libre de Bruxelles
Marine Biology Laboratory
Brussels, Belgium
&
UMR 6282 Biogéosciences, Univ. Bourgogne Franche-Comté
CNRS, 6 bd Gabrie F-21000 Dijon, France

**Mélyne Hautecœur**
Muséum National d'Histoire Naturelle
Département Milieux et Peuplements Aquatiques
UMR 7209 Paris, France

**Franz M. Heindler**
KU Leuven
Laboratory of Biodiversity and Evolutionary Genomics
Ch. Deberiotstraat 32, box 2439
B-3000 Leuven, Belgium

**Bart Hellemans**
KU Leuven
Laboratory of Biodiversity and Evolutionary Genomics
Ch. Deberiotstraat 32, box 2439
B-3000 Leuven, Belgium

**Quentin Jossart**
Vrije Universiteit Brussel
Marine Biology Group
Brussels, Belgium

**Marc Kochzius**
Vrije Universiteit Brussel
Marine Biology Group
Brussels, Belgium

**Fredrik Leliaert**
Ghent University
Marine Biology Research Group
Ghent, Belgium

**Camille Moreau**
Université Libre de Bruxelles
Marine Biology Laboratory
Brussels, Belgium
&

UMR 6282 Biogéosciences, Univ. Bourgogne
Franche-Comté
CNRS, 6 bd Gabrie F-21000 Dijon, France

**Francesca Pasotti**
Ghent University
Marine Biology Research Group
Ghent, Belgium

**Tasnim Patel**
Royal Belgian Institute of Natural Sciences
OD Nature
Vautiersstraat 29
B-1000 Brussels, Belgium

**Henri Robert**
Royal Belgian Institute of Natural Sciences
OD Nature
Vautiersstraat 29
B-1000 Brussels, Belgium

**Isabelle M. Schön**
Royal Belgian Institute of Natural Sciences
OD Nature
Vautiersstraat 29
B-1000 Brussels, Belgium

**Dirk Steinke**
University of Guelph
Centre for Biodiversity Genomics
Guelph, ON, Canada
&
University of Guelph
Department of Integrative Biology
Guelph, ON, Canada

**Anton P. Van de Putte**
Royal Belgian Institute of Natural Sciences
OD Nature
Vautiersstraat 29
B-1000 Brussels, Belgium

**Ann Vanreusel**
Ghent University
Marine Biology Research Group
Ghent, Belgium

**Marie Verheye**
Royal Belgian Institute of Natural Sciences
OD Nature
Vautiersstraat 29
B-1000 Brussels, Belgium

**Filip Volckaert**
KU Leuven
Laboratory of Biodiversity and Evolutionary Genomics
Ch. Deberiotstraat 32, box 2439
B-3000 Leuven, Belgium

**Emma F. Young**
British Antarctic Survey
Cambridge, United Kingdom

# APPENDICES

### I. Acknowledgment of funding an in-kind support

The research described here was financially supported by an individual PhD Fellowship grant from the former Flemish Agency for Innovation by Science and Technology (IWT), now managed through Flanders Innovation & Entrepreneurship (VLAIO, grant no. 141328).

The Belgian Science Policy Office (BELSPO) fundamentally supported this work through the projects "Ecosystem Responses to global change: a multiscale approach in the Southern Ocean" (vERSO, contract no. BR/132/A1/vERSO) and "Refugia and Ecosystem Tolerance in the Southern Ocean" (RECTO, contract no. BR/154/A1/RECTO). Thanks to the KU Leuven and the Laboratory of Biodiversity and Evolutionary Genomics, where this thesis was hosted.

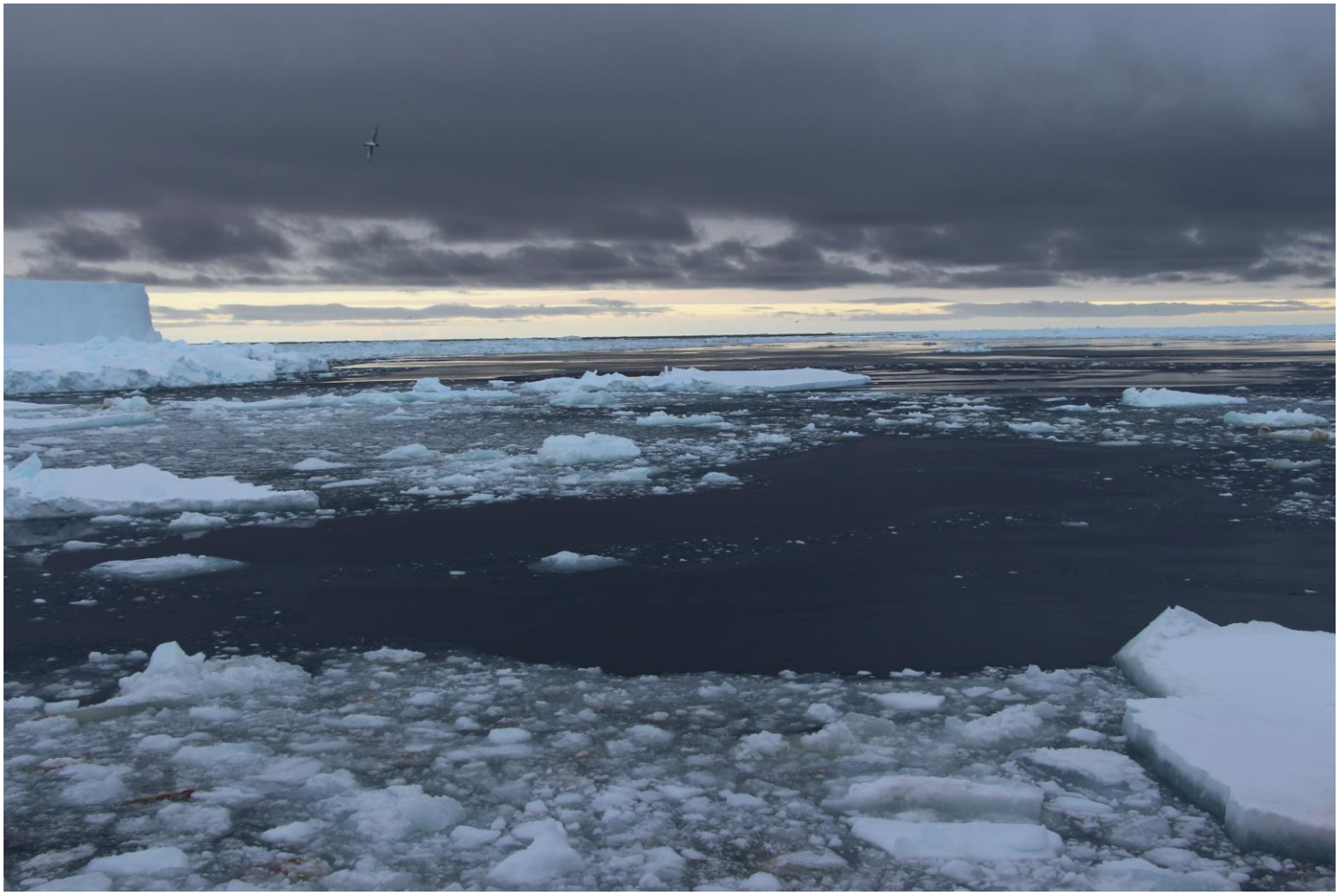
Special thanks to the institutions and personnel that supported me in the field, that is the Alfred Wegener Institute, Helmholtz Center for Polar and Marine Research (AWI), the captain and crew of RV *Polarstern* expedition PS96, all members of the fish and benthos teams during this expedition, and the captain and crew of Ocean Expeditions AV *Australis* (B. Wallis, R. Houston, K. Lucas) and the team of the Belgica 121 expedition (B. Danis, C. Guillaumot, F. Heindler, Q. Jossart, C. Moreau, F. Pasotti, H. Robert, T. Saucède). Thanks to the Research Foundation Flanders (FWO), which supported my participation in the latter expedition through a grant for long stays abroad (grant no. V407319N).

Finally, I would like to thank the institutions that have supported APECS Belgium in the past years: APECS International, BELSPO, the Belgian National Committee for Antarctic Research (BNCAR), the Royal Belgian Institute of Natural Sciences (RBINS), APECS Netherlands, the Netherlands Organization for Scientific Research (NWO), APECS Luxembourg/Polar.lu, Archives Antarctiques Belges – Belgische Antarctische Archieven (AABBAA), the Scientific Committee for Antarctic Research (SCAR), the SCAR research program AntEco, the SCAR Standing Committee on Antarctic Data Management (SCADM), KU Leuven, University of Ghent, Université de Liège, Université Libre de Bruxelles, and Vrije Universiteit Brussel.

**II. Picture credits**

Fish drawing in Fig. 3.1 by R.C. Cada for FishBase.

All other figures and images: Henrik Christiansen & Sandra Maier.



### III. Outreach and education in support of awareness and future research

The translation of fundamental and applied research into management and conservation can be challenging due to e.g. the complexity and inherent uncertainties of the research (Shafer *et al.* 2015), competing policy objectives (Brooks 2013), and insufficient or ineffective communication (Doubleday & Connell 2017). The latter is a pertinent, but frustrating issue as almost all scientists would agree that communicating their research is an integral part of science. However, the views on how that should be done differ among scientists, with a remarkable proportion of researchers communicating predominantly or even exclusively with each other via peer-reviewed articles and presentations and posters at scientific conferences. Engaging with society directly or through media, policy makers or other stakeholders is crucial for actual, real world impacts, lest science remains but an academic exercise. Importantly, by emphasizing the relevance of science communication I do not mean to contend the value of fundamental research without direct application aspect. Many discoveries have been made without any ulterior motive or direct application benefit. However, this should not prevent one from contributing to science communication. If anything, it may stimulate critical thinking and sharing the fascination one holds for a research topic. In that sense, science communication with varied audiences can be mutually beneficial for all parties involved, i.e. the scientist, science, the public, society and (in the case of biology) nature itself (Obregón *et al.* 2018). Therefore, engaging in activities that go beyond fundamental research, such as outreach, education and science communication (or "social strategies", see Ardoin & Heimlich 2013), have value and also bear the potential to positively affect the uptake of research outcomes by policy makers and other stakeholders (Bjorkland & Pringle 2001; Leisher *et al.* 2012). The polar realm and the Antarctic in particular are good examples in this regard, because of their special political background. CCAMLR as the governing body of fisheries in the Southern Ocean operates on a multilateral and consensus-based level. In turn, individual countries can only influence the policies of CCAMLR to a certain extent. Non-state actors and informal negotiations are likely to become more important in the future to support the success of CCAMLR conservation actions (Sykora-Bodie & Morrison 2019). Even within one nation, it is necessary to form clear objectives and standpoints regarding Antarctic policy. Contributing to the education of future citizens and scientists and raising awareness among the public has strong potential to ensure that policy makers act sustainably and prioritize protection over exploitation (Fig. 6.4).

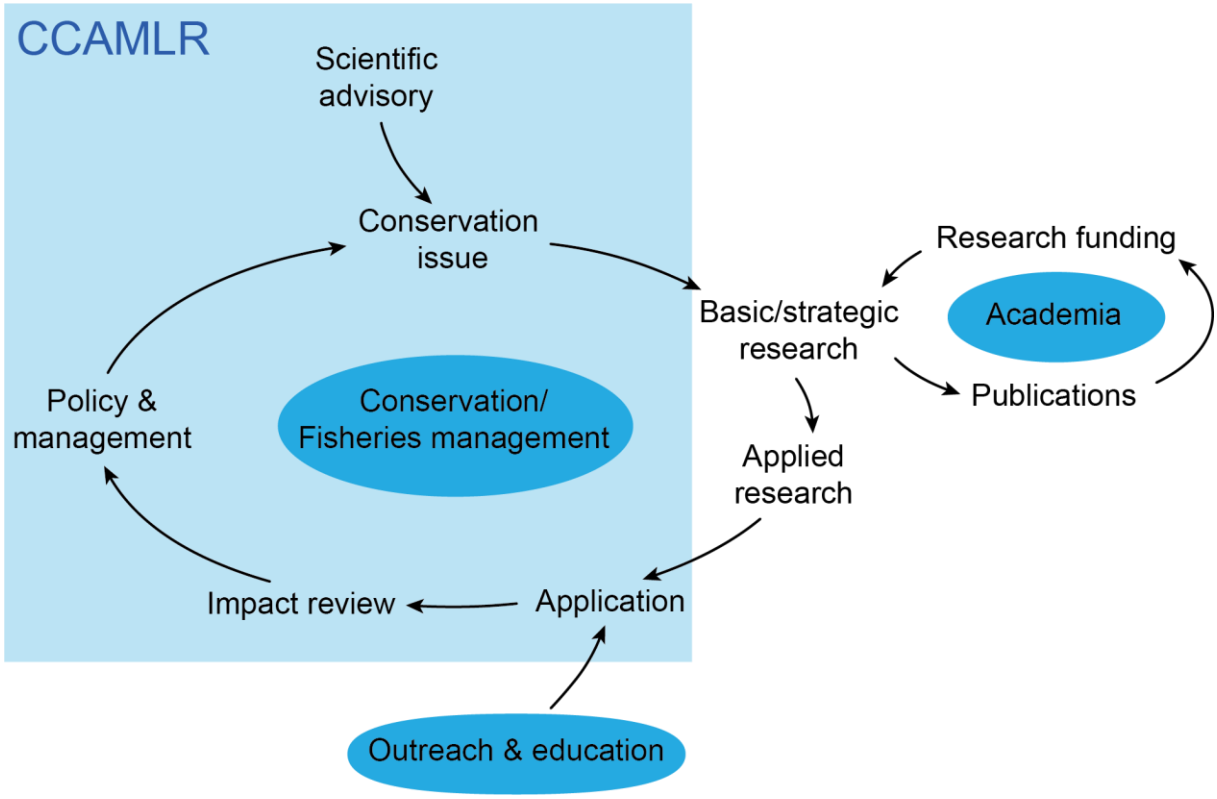



**Fig. 6.4.** The conservation and fisheries management process as implemented in the Committee for the Conservation of Antarctic Marine Living Resources (CCAMLR) and how it is affected by Academia and outreach and education activities. The latter play an important role in contributing to the successful uptake and application of research outcomes in management measures.

During the course of this thesis, I have developed an increasing appreciation of the value of public outreach and education and have initiated and contributed to a number of projects that aimed at raising awareness about polar science. In the following, I summarize a few examples of how outreach and education can be tackled within a PhD framework. The Association of Polar Early Career Scientists (APECS) provided an already well-established platform for international and national outreach activities (Hindshaw *et al.* 2019). The main goal of APECS is to "shape the future of polar research" through contributions to the training of polar researchers, as well as through outreach and education initiatives for the public at large. To achieve these goals APECS is structured as an international and interdisciplinary organization with clearly defined leadership. In addition to the global APECS structure, several national committees (NCs) of APECS for single countries (or several countries) exist and perform APECS activities on a local scale. A particular strength of APECS is its interdisciplinary nature (Carlson 2015). The uniting characteristic of all APECS members and advocates is a shared excitement and appreciation of the polar (and alpine) systems, not a focus on a single scientific discipline such as biology, physics, or social science.

The NC for Belgium, that is APECS Belgium, was founded in 2011, but became inactive some years after. The high turnover of early career researchers (ECRs), who are the main contributors to APECS, regularly leads to periods of inactivity or reduced contributions, especially in a country with comparatively small polar science communities such as Belgium. In 2016, I initiated a reinvigoration of APECS Belgium by calling on polar ECRs throughout Belgium to join for a meeting, which resulted in the formation of a leadership board. Together with six other polar ECRs we then started to reactivate the NC, planned first activities and (re-)opened internet and social media presences. By now APECS Belgium has a clearly defined structure with its own Terms of Reference, describing its purpose and organization structure, and a Letter of Agreement with APECS International. The leadership of APECS Belgium consists of an open board with a flexible number of members that are appointed for one year, from beginning of October to end of September, and led by a chair. The organization is now in its fourth year after the revival and continues to be very active and shaped by several highly motivated polar ECRs. Many different activities have been carried out by APECS Belgium during this time period (Table 6.5). For example, several workshops and networking events were conducted, largely aiming at increasing the awareness of the importance of science communication, detailing ways in which science communication, outreach and education can happen and providing a platform for polar ECRs to learn, connect and network. In addition, many activities were aimed directly at members of the public, such as school children, museum visitors, or internet users. Internet technologies are particularly suited to reach a very wide audience and complement personal interaction (Xavier *et al.* 2016).

**Table 6.5.** Non-exhaustive list of activities conducted by the Association of Polar Early Career Researchers (APECS) Belgium during the period from 2016 to 2019. In addition to the events listed below, for example, dozens of in person or online meetings of the board members and organizers of the events have taken place during this time period, as well as participation in other events/conferences. More detailed information on most activities can be found on apecsbelgium.wordpress.com. Belgian APECS activities organized by: H. Christiansen, I. Pessi, S. Berger, A. Gossart, F. Pasotti, E. Pinseel, A. Roukaerts, F. Heindler, C. Guillaumot, A. Aguera, C. Moreau, Q. Jossart, B. Durieu, V. Savaglia, C. Jacques, F. Deman, Y. M. Gan, M. Pinzone, S. Maes, D. Brode-Roger, L. De Maeyer, M. Cavitte; various other individuals and organizations supported many of these activities.

| Activity | Date | Location | Description |
|---|---|---|---|
| APECS Belgium Get to Know Meeting | 17.08.2016 | Brussels | One day meeting gathering 14 polar ECRs from throughout Belgium. |
| APECS Netherlands/Belgium Symposium | 02.11.2016 | The Hague | One day symposium with invited speakers and contributions from attendants discussing recent polar science. |
| Antarctic Photo Competition | 01.12.2016 | Online | Online picture contest, where polar enthusiasts can submit and vote for their best Antarctic picture. |
| SCAR Biology Symposium 2017 (SCARBio17): Low Fare Accommodation | 10-14.07.2017 | Leuven | Helped organize low fare accommodation for ECRs attending the conference. |
| SCARBio17: ECR Awards | 10-14.07.2017 | Leuven | Helped organize presentation and poster awards. |
| SCARBio17: Antarctic Photo Auction | 10-14.07.2017 | Leuven | Displayed and offered pictures of our photo competition in a “pen and paper auction”. |
| SCARBio17: APECS Workshop | 09.07.2017 | Leuven | Full day workshop focussed on science communication. Four invited speakers/teachers and 38 attendees. |
| SCARBio17: Social Evening | 11.07.2017 | Leuven | Organized a local pub evening for ECRs (and colleagues). |
| Polar Week Movie Screening | 27.09.2017 | Liège | Movie screening & debate. |
| Science Day | 26.11.2017 | Ghent | Science stands and exhibition material. |
| Antarctica Day 2017 | 01.12.2017 | Online | Launch of new interactive online maps, where ongoing polar research in Belgium and current polar expeditions with Belgian scientists are shown. |
| Data tutorials | 02.2018 | Online | New series of online tutorials on coding. |
| AABBAA Event | 21.03.2018 | Brussels | Help in organisation and participation in a conference evening by the “Archives Antarctiques Belges – Belgische Antarctische Archieven” (AABBAA). |

| | | | |
|---|---|---|---|
| Polar View | 03.2018 | Online | A new online blog category aiming to explain complicated polar science in simple words to an audience not familiar with it. |
| APECS World Summit | 17.06.2018 | Davos | Participation in the World Summit, where APECS international leadership and several NCs met. |
| Story Book Contest | 2018-2019 | All over Belgium | Continuation of an Antarctic story book contest with school children. Initiated by the former APECS Belgium group several years ago and now completed with the publication of the book. |
| APECS Belgium Day 2018 | 18.09.2018 | Brussels | Networking event held for all polar enthusiasts in Belgium. |
| School visits for Antarctica Day 2018 | 20-20.11.2018 | Brussels | Four visits to three schools in Brussels. Various educational activities for the school kids showcasing Antarctica, including the instalment of pinhole cameras. |
| Polar Week Q&A | 18.-23.03.2019 | Online | Question and Answer event about polar science. |
| School Visits | 06.2019 | Brussels | Revisiting the same schools as before to collect the pinhole cameras and conclude the event. |
| Summer Q&A | 07-08.2019 | Online | Question and Answer event about polar science. |
| APECS Belgium Day 2019 | 12.09.2019 | Brussels | Networking event held for all polar enthusiasts in Belgium. |
| Antarctica Day 2019 | 01.12.2019 | Brussels | Participation at the opening of the new Antarctica exhibition at the Royal Belgian Institute of Natural Sciences. |

We also attempted to link APECS Belgium (and therefore the ECRs that shape it) with higher level Belgian polar research through attendance of the biannual meetings of the Belgian National Committee on Antarctic Research (BNCAR). BNCAR has readily opened the door for APECS Belgium. Thus, this link and collaboration could be extended in the future to facilitate more direct interaction between ECRs and established researchers and (more) joint activities. The matter of consistency will still be an issue for APECS Belgium and at the moment it can only be speculated whether the NC will remain active in the future or fall into hibernation again when the current generation of ECRs progresses with their careers. Nevertheless, we have put mechanisms in place that may help to achieve some continuity. First, the NC has now a clear definition and written background documents. This is an important step towards professionalization of any outreach and education initiative (Salmon & Roop 2019). Second, there is a bank account linked to the organization and a financial backbone that can support

small running costs (partly inherited from funding from the first installment of APECS Belgium). Third, APECS Belgium has a defined leadership structure that is to be renewed on a yearly basis. The latter hopefully provides strong incentives for the respective current leadership to recruit, motivate and support new ECRs that follow along the same path. In the past four years this process has worked very well, with people that are active within the APECS Belgium board always on the lookout for new prospective members. Typically, new members join first as regular board members and can thus learn from the ones more familiar with outreach and education activities, while at the same time bringing fresh ideas to the board. The chair of the board has an important role – not necessarily as the person that puts in most work for the organization of events – but as the one that reminds other members of upcoming tasks, schedules regular meetings and prevents the board from becoming inactive. Ideally, someone already active in the board for some time can subsequently step up to take over the role of chair and then recruit a new chair before resigning from the role (current chair: V. Savaglia; previously: H. Christiansen). Rightfully, I believe, APECS is now recognized on an international level by leading experts in polar research and policy as an important actor to train new researchers, raise awareness and foster international and interdisciplinary work (Carlson 2015; Majaneva *et al.* 2016; Hindshaw *et al.* 2019). Perhaps APECS Belgium can achieve similar results on a national level, although measurable impact will of course take a long time to be built. Reaching conservation goals is even more difficult to assess and also a long-term process. APECS Belgium now at least provides a well-functioning platform that supports the personal development of the next generation of polar researchers as well as the education of the public on polar science in Belgium.

Finally, while an organization such as APECS Belgium provides an optimal platform to engage with other polar scientists and organizations and with the public, it is also possible to target the latter through entirely open, free-form ways. In the past four years, we have engaged in science communication activities through outreach in relation to the unique "Belgica 121" expedition. This research cruise was conducted 121 years after the original Belgica expedition, the first international Antarctic science cruise (De Deckker 2018). The expedition team maintained an active blog (belgica121.be) and also started a crowdfunding campaign to finance the realization of a documentary film. Under the lead of F. Heindler several company sponsors and 10,000 EUR from private supporters were collected and thus the costs for professional filming equipment and post-processing could be covered. The science crew itself operated as film crew and documentary protagonists at the same time. The result is a 66 minute long feature documentary entitled "Observations at 65° South" directed by L. Hess. At the time of writing, the film has not

been released yet (closed premiere for sponsors and collaborators took place on 10th January, 2020) and the distribution channel is still being negotiated. Depending on the distribution the film will likely have the potential to be seen by more than 10,000 people. Exposure of such an audience to the fragility of the Southern Ocean and its fauna might eventually contribute to a wider appreciation of the importance of Antarctic conservation.

## IV. Supplementary Material

All Supplementary Material reprinted here without additional formatting. The metadata, raw data, and most often also R scripts used to analyze the data are made available online. As some chapters are not published yet, the online material is sometimes not accessible yet either. However, the correct hyperlinks are in place wherever possible and should become functional in the future.

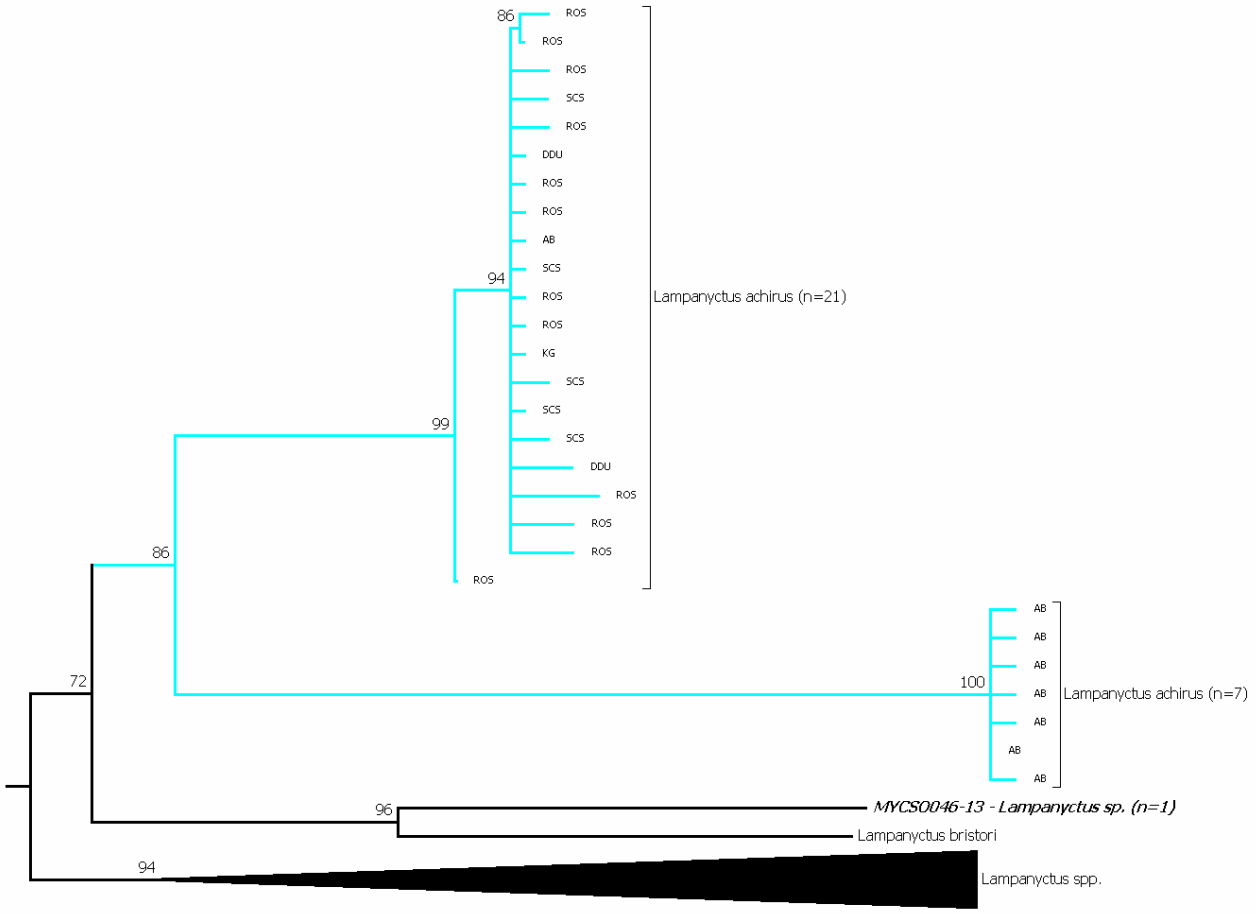
**IV. S2. Supplementary Material Chapter 2**

# *Supplementary Material*

## Diversity of Mesopelagic Fishes in the Southern Ocean - a Phylogeographic Perspective Using DNA Barcoding

**Henrik Christiansen*, Agnès Dettaï, Franz M. Heindler, Martin A. Collins, Guy Duhamel, Mélyne Hautecœur, Dirk Steinke, Filip A.M. Volckaert, Anton P. Van de Putte**

*** Correspondence:** Corresponding Author: henrik.christiansen@kuleuven.be




**Supplementary Figure S2.1.** Phylogenetic consensus tree of myctophid fishes of a *Lampanyctus* clade including the sub-Antarctic *Lampanyctus achirus* (light blue) based on cytochrome c oxidase I (*COI*) gene variation using a maximum likelihood analysis with 10,000 bootstrap permutations. Bootstrap support > 70 is shown above the branches; branches with lower support are collapsed to polytomies. Locality codes as in Fig. 2.1.

**Supplementary Table S2.2.** Mesopelagic fish used for DNA analyses with specimen and sequence ID and collection information as available in the Barcode of Life Data Systems (BOLD) as well as respective BOLD Barcode Index Number (BOLD BIN). Sampling coordinates displayed as latitude (Lat) and longitude (Lon) in decimal degrees. Cytochrome oxidase subunit I (*COI*) and/or rhodopsin (*rh1*) sequences are available for these individuals unless the respective column contains a zero. Storing institution (Inst.) abbreviations: British Antarctic Survey, Cambridge, UK (BAS), South African Institute for Aquatic Biodiversity, Grahamstown, South Africa (SAIAB), KU Leuven, Leuven, Belgium (KUL), Museum National d'Histoire Naturelle, Paris, France (MNHN). Samples from France are from the French Southern and Antarctic Lands. Further collection details for some individuals available in BOLD. Species identities highlighted in yellow have been revised or updated (see main text and table 2.2).

| BOLD dataset | Sample ID | Sequence ID | BOLD BIN | Inst. | *Species* | Collection Date | Country/Ocean | Region | Lat | Lon | *COI* | *Rhl* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| BASMF | SG09_no33_Aethotaxis | CAOII859-09 | | BAS | *Aethotaxis mitopteryx* | 19-Jan-2009 | South Georgia Islands | NW South Georgia | -53,753 | -37,466 | 658 | 0 |
| BASMF | BEM#1 | BASMC066-09 | BOLD:AAF0459 | BAS | *Lagiacrusichthys macropinna* | 30-Nov-2006 | Antarctica | Scotia Sea | -57,79 | -50,63 | 581 | 0 |
| BASMF | BOA#1 | BASMC067-09 | BOLD:AAB5140 | BAS | *Borostomias antarcticus* | 30-Oct-2006 | Antarctica | Scotia Sea | -57,76 | -50,45 | 652 | 0 |
| BASMF | BOA#2 | BASMC068-09 | BOLD:AAB5140 | BAS | *Borostomias antarcticus* | 30-Oct-2006 | Antarctica | Scotia Sea | -57,76 | -50,45 | 652 | 0 |
| BASMF | ELN#1 | BASMC006-09 | BOLD:AAB3737 | BAS | *Electrona antarctica* | 30-Oct-2006 | Antarctica | Scotia Sea | -57,79 | -50,63 | 641 | 0 |
| BASMF | ELN#2 | BASMC007-09 | BOLD:AAB3737 | BAS | *Electrona antarctica* | 01-Nov-2006 | | | -60,5 | -48,75 | 648 | 0 |
| BASMF | ELN#3 | BASMC008-09 | BOLD:AAB3737 | BAS | *Electrona antarctica* | Antarctica | Scotia Sea | Scotia Sea | -60,5 | -48,75 | 648 | 0 |
| BASMF | ELN#4 | BASMC009-09 | BOLD:AAB3737 | BAS | *Electrona antarctica* | 30-Oct-2006 | Antarctica | Scotia Sea | -57,79 | -50,63 | 648 | 0 |
| BASMF | ELN#5 | BASMC010-09 | BOLD:AAB3737 | BAS | *Electrona antarctica* | 01-Nov-2006 | Antarctica | Scotia Sea | -60,5 | -48,75 | 648 | 0 |
| BASMF | ELC#1 | BASMC011-09 | BOLD:AAB8433 | BAS | *Electrona carlsbergi* | 30-Oct-2006 | Antarctica | Scotia Sea | -57,79 | -50,63 | 648 | 0 |
| BASMF | ELC#2 | BASMC012-09 | BOLD:AAB8433 | BAS | *Electrona carlsbergi* | 29-Nov-2006 | Antarctica | Scotia Sea | -57,67 | -50,42 | 648 | 0 |
| BASMF | ELC#3 | BASMC013-09 | BOLD:AAB8433 | BAS | *Electrona carlsbergi* | 29-Nov-2006 | Antarctica | Scotia Sea | -57,67 | -50,42 | 648 | 0 |
| BASMF | ELC#4 | BASMC014-09 | BOLD:AAB8433 | BAS | *Electrona carlsbergi* | 29-Nov-2006 | Antarctica | Scotia Sea | -57,67 | -50,42 | 647 | 0 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| BASMF | ELC#5 | BASMC015-09 | BOLD:AAB8433 | BAS | *Electrona carlsbergi* | 29-Nov-2006 | Antarctica | Scotia Sea | -57,67 | -50,42 | 648 | 0 |
| BASMF | ELS#1 | BASMC042-09 | BOLD:AAD8917 | BAS | *Electrona subaspera* | 16-Jan-2008 | Antarctica | Scotia Sea | -50,55 | -34,05 | 652 | 0 |
| BASMF | GYB#1 | BASMC041-09 | BOLD:AAC7549 | BAS | *Gymnoscopelus bolini* | 16-Jan-2008 | Antarctica | Scotia Sea | -59,68 | -44,14 | 648 | 0 |
| BASMF | Gymnoscopelus_bolini_ | CAOII857-09 | | BAS | *Gymnoscopelus bolini* | 13-Jan-2009 | South Georgia and the South Sandwich Islands | NW South Georgia | -53,798 | -39,07 | 0 | 0 |
| BASMF | GYR#1 | BASMC043-09 | BOLD:AAB5232 | BAS | *Gymnoscopelus braueri* | 30-Oct-2006 | Antarctica | Scotia Sea | -57,76 | -50,45 | 648 | 0 |
| BASMF | GYR#2 | BASMC044-09 | BOLD:AAB5232 | BAS | *Gymnoscopelus braueri* | 30-Oct-2006 | Antarctica | Scotia Sea | -57,76 | -50,45 | 648 | 0 |
| BASMF | GYR#3 | BASMC045-09 | BOLD:AAB5232 | BAS | *Gymnoscopelus braueri* | 30-Oct-2006 | Antarctica | Scotia Sea | -57,79 | -50,63 | 637 | 0 |
| BASMF | GYR#4 | BASMC046-09 | BOLD:AAB5232 | BAS | *Gymnoscopelus braueri* | 30-Oct-2006 | Antarctica | Scotia Sea | -57,79 | -50,63 | 648 | 0 |
| BASMF | GYR#5 | BASMC047-09 | BOLD:AAB5232 | BAS | *Gymnoscopelus braueri* | 30-Oct-2006 | Antarctica | Scotia Sea | -57,79 | -50,63 | 648 | 0 |
| BASMF | GYF#1 | BASMC036-09 | BOLD:ACE4884 | BAS | *Gymnoscopelus fraseri* | 23-Mar-2004 | Antarctica | Scotia Sea | -53,31 | -37,98 | 648 | 0 |
| BASMF | GYF#2 | BASMC037-09 | BOLD:ACE4884 | BAS | *Gymnoscopelus fraseri* | 23-Mar-2004 | Antarctica | Scotia Sea | -53,31 | -37,98 | 648 | 0 |
| BASMF | GYF#3 | BASMC038-09 | BOLD:ABZ1021 | BAS | *Gymnoscopelus fraseri* | 28-Nov-2008 | Antarctica | Scotia Sea | -49,95 | -38,06 | 648 | 0 |
| BASMF | GYF#4 | BASMC039-09 | BOLD:AAC7549 | BAS | *Gymnoscopelus bolini* | 28-Nov-2008 | Antarctica | Scotia Sea | -49,95 | -38,06 | 648 | 0 |
| BASMF | GYF#5 | BASMC040-09 | BOLD:ABZ1021 | BAS | *Gymnoscopelus fraseri* | 28-Nov-2008 | Antarctica | Scotia Sea | -49,95 | -38,06 | 648 | 0 |
| BASMF | GYN#1 | BASMC026-09 | BOLD:AAB0899 | BAS | *Gymnoscopelus nicholsi* | 30-Oct-2006 | Antarctica | Scotia Sea | -57,79 | -40,02 | 648 | 0 |
| BASMF | GYN#2 | BASMC027-09 | BOLD:AAB0899 | BAS | *Gymnoscopelus nicholsi* | 20-Jan-2008 | Antarctica | Scotia Sea | -58,01 | -43,05 | 648 | 0 |
| BASMF | GYN#3 | BASMC028-09 | BOLD:AAB0899 | BAS | *Gymnoscopelus nicholsi* | 20-Jan-2008 | Antarctica | Scotia Sea | -58,01 | -43,05 | 648 | 0 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| BASMF | GYN#4 | BASMC029-09 | BOLD:AAB0899 | BAS | *Gymnoscopelus nicholsi* | 20-Jan-2008 | Antarctica | Scotia Sea | -58,01 | -43,05 | 648 | 0 |
| BASMF | GYO#1 | BASMC031-09 | BOLD:AAB6899 | BAS | *Gymnoscopelus opisthopterus* | 01-Nov-2008 | Antarctica | Scotia Sea | -60,47 | -48,78 | 648 | 0 |
| BASMF | GYO#2 | BASMC032-09 | BOLD:AAB6899 | BAS | *Gymnoscopelus opisthopterus* | 01-Nov-2008 | Antarctica | Scotia Sea | -60,47 | -48,78 | 648 | 0 |
| BASMF | GYO#3 | BASMC033-09 | BOLD:AAB6899 | BAS | *Gymnoscopelus opisthopterus* | 08-Nov-2008 | Antarctica | Scotia Sea | -59,63 | -44,15 | 648 | 0 |
| BASMF | GYO#4 | BASMC034-09 | BOLD:AAB6899 | BAS | *Gymnoscopelus opisthopterus* | 08-Nov-2008 | Antarctica | Scotia Sea | -59,63 | -44,15 | 648 | 0 |
| BASMF | GYO#5 | BASMC035-09 | BOLD:AAB6899 | BAS | *Gymnoscopelus opisthopterus* | 08-Nov-2008 | Antarctica | Scotia Sea | -59,63 | -44,15 | 648 | 0 |
| BASMF | GYP#1 | BASMC030-09 | BOLD:AAB0899 | BAS | *Gymnoscopelus nicholsi* | 30-Oct-2006 | Antarctica | Scotia Sea | -57,79 | -50,63 | 648 | 0 |
| BASMF | KRA #1 | BASMC001-09 | BOLD:AAC6875 | BAS | *Krefftichthys anderssoni* | 21-Nov-2006 | Antarctica | Scotia Sea | -55,26 | -41,16 | 652 | 0 |
| BASMF | KRA #2 | BASMC002-09 | BOLD:AAC6875 | BAS | *Krefftichthys anderssoni* | 21-Nov-2006 | Antarctica | Scotia Sea | -55,26 | -41,16 | 652 | 0 |
| BASMF | KRA #3 | BASMC003-09 | BOLD:AAC6875 | BAS | *Krefftichthys anderssoni* | 21-Nov-2006 | Antarctica | Scotia Sea | -55,26 | -41,16 | 652 | 0 |
| BASMF | KRA #4 | BASMC004-09 | BOLD:AAC6875 | BAS | *Krefftichthys anderssoni* | 21-Nov-2006 | Antarctica | Scotia Sea | -55,26 | -41,16 | 652 | 0 |
| BASMF | KRA #5 | BASMC005-09 | BOLD:AAC6875 | BAS | *Krefftichthys anderssoni* | 17-Nov-2006 | Antarctica | Scotia Sea | -57,3 | -42,84 | 652 | 0 |
| BASMF | LAC#1 | BASMC021-09 | BOLD:AAB3778 | BAS | *Nannobrachium achirus* | 30-Oct-2006 | Antarctica | Scotia Sea | -57,69 | -50,47 | 652 | 0 |
| BASMF | LAC#2 | BASMC022-09 | BOLD:AAB3778 | BAS | *Nannobrachium achirus* | 26-Nov-2006 | Antarctica | Scotia Sea | -57,69 | -40,02 | 652 | 0 |
| BASMF | LAC#3 | BASMC023-09 | BOLD:AAB3778 | BAS | *Nannobrachium achirus* | 26-Nov-2006 | Antarctica | Scotia Sea | -57,69 | -40,02 | 652 | 0 |
| BASMF | LAC#4 | BASMC024-09 | BOLD:AAB3778 | BAS | *Nannobrachium achirus* | 26-Nov-2006 | Antarctica | Scotia Sea | -57,69 | -40,02 | 652 | 0 |
| BASMF | LAC#5 | BASMC025-09 | BOLD:AAB3778 | BAS | *Nannobrachium achirus* | 26-Nov-2006 | Antarctica | Scotia Sea | -57,69 | -40,02 | 652 | 0 |
| BASMF | CER#1 | BASMC048-09 | BOLD:AAD9321 | BAS | *Oneirodes notius* | 23-Mar-2004 | Antarctica | Scotia Sea | -53,31 | -37,98 | 652 | 0 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| BASMF | PMC#1 | BASMC060-09 | BOLD:AAB7779 | BAS | *Poromitra capito* | 26-Nov-2006 | Antarctica | Scotia Sea | -52,85 | -40,02 | 652 | 0 |
| BASMF | PMC#2 | BASMC061-09 | BOLD:AAB7779 | BAS | *Poromitra capito* | 26-Nov-2006 | Antarctica | Scotia Sea | -52,85 | -40,02 | 652 | 0 |
| BASMF | PMC#3 | BASMC062-09 | BOLD:AAB7779 | BAS | *Poromitra capito* | 26-Nov-2006 | Antarctica | Scotia Sea | -57,76 | -50,45 | 652 | 0 |
| BASMF | PRA#1 | BASMC059-09 | | BAS | *Protomyctophum andriashevi* | 26-Nov-2006 | Antarctica | Scotia Sea | -52,9 | -40,04 | 0 | 0 |
| BASMF | PRM#1 | BASMC049-09 | BOLD:AAC7885 | BAS | *Protomyctophum bolini* | 28-Nov-2006 | Antarctica | Scotia Sea | -50,12 | -38,14 | 648 | 0 |
| BASMF | PRM#2 | BASMC050-09 | BOLD:AAC7885 | BAS | *Protomyctophum bolini* | 28-Nov-2006 | Antarctica | Scotia Sea | -50,12 | -38,14 | 648 | 0 |
| BASMF | PRM#3 | BASMC051-09 | BOLD:AAC7885 | BAS | *Protomyctophum bolini* | 05-Feb-2008 | Antarctica | Scotia Sea | -52,72 | -39,08 | 648 | 0 |
| BASMF | PRM#4 | BASMC052-09 | BOLD:AAC7885 | BAS | *Protomyctophum bolini* | 05-Feb-2008 | Antarctica | Scotia Sea | -52,72 | -39,08 | 640 | 0 |
| BASMF | PRM#5 | BASMC053-09 | BOLD:AAC7885 | BAS | *Protomyctophum bolini* | 05-Feb-2008 | Antarctica | Scotia Sea | -52,72 | -39,08 | 648 | 0 |
| BASMF | PRY#1 | BASMC054-09 | BOLD:AAD9636 | BAS | *Protomyctophum choriodon* | 15-Mar-2004 | Antarctica | Scotia Sea | -53,79 | -40,22 | 648 | 0 |
| BASMF | PRY#2 | BASMC055-09 | BOLD:AAD9636 | BAS | *Protomyctophum choriodon* | 15-Mar-2004 | Antarctica | Scotia Sea | -53,79 | -40,22 | 648 | 0 |
| BASMF | PRG#1 | BASMC056-09 | BOLD:ACE3466 | BAS | *Protomyctophum gemmatum* | 20-Nov-2006 | Antarctica | Scotia Sea | -55,21 | -41,38 | 648 | 0 |
| BASMF | PRL#1 | BASMC057-09 | BOLD:AAC7885 | BAS | *Protomyctophum bolini* | 30-Oct-2006 | Antarctica | Scotia Sea | -57,69 | -50,46 | 648 | 0 |
| BASMF | PRP#1 | BASMC058-09 | BOLD:ACF2373 | BAS | *Protomyctophum parallelum* | 30-Oct-2006 | Antarctica | Scotia Sea | -57,69 | -50,46 | 652 | 0 |
| BASMF | PRE#1 | BASMC016-09 | BOLD:AAD9622 | BAS | *Protomyctophum tenisoni* | 07-Feb-2008 | Antarctica | Scotia Sea | -50,55 | -34,05 | 648 | 0 |
| BASMF | PRE#2 | BASMC017-09 | BOLD:AAD9622 | BAS | *Protomyctophum tenisoni* | 07-Feb-2008 | Antarctica | Scotia Sea | -50,55 | -34,05 | 648 | 0 |
| BASMF | PRE#3 | BASMC018-09 | BOLD:AAD9622 | BAS | *Protomyctophum tenisoni* | 07-Feb-2008 | Antarctica | Scotia Sea | -50,55 | -34,05 | 648 | 0 |
| BASMF | PRE#4 | BASMC019-09 | BOLD:AAD9622 | BAS | *Protomyctophum tenisoni* | 07-Feb-2008 | Antarctica | Scotia Sea | -50,55 | -34,05 | 648 | 0 |
| BASMF | PRE#5 | BASMC020-09 | BOLD:AAD9622 | BAS | *Protomyctophum tenisoni* | 07-Feb-2008 | Antarctica | Scotia Sea | -50,55 | -34,05 | 648 | 0 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| BASMF | Ichychthus_australia | CAOII858-09 | | BAS | *Icichthys australis* | 13-Jan-2009 | South Georgia and the South Sandwich Islands | NW South Georgia | -53,798 | -39,07 | 0 | 0 |
| BASMF | IIC#1 | BASMC065-09 | BOLD:AAE2761 BOLD:AAA8371 | BAS | *Icichthys australis* | 17-Nov-2006 | Antarctica | Scotia Sea | -57,31 | -42,85 | 652 | 0 |
| BASMF | SG09_no67_Notothenid | CAOII860-09 | | BAS | *Trematomus vicarius* | 23-Jan-2009 | South Georgia and the South Sandwich Islands | SW South Georgia | -54,23 | -37,924 | 657 | 0 |
| BASMF | SIO#1 | BASMC063-09 | BOLD:AAJ7118 | BAS | *Sio nordenskjoldii* | 28-Nov-2006 | Antarctica | Scotia Sea | -50,03 | -38,06 | 648 | 0 |
| BASMF | MEL#1 | BASMC064-09 | BOLD:AAJ7118 | BAS | *Sio sp.* | 05-Feb-2008 | Antarctica | Scotia Sea | -52,62 | -39,12 | 652 | 0 |
| BASMF | SG09_no54_Notothenid | CAOII861-09 | | BAS | *Trematomus cf. vicarius* | 21-Jan-2009 | South Georgia and the South Sandwich Islands | SE South Georgia | -54,846 | -35,23 | 0 | 0 |
| DSSAU | ADC10_86.999 #1 | DSFSG402-10 | BOLD:AAB9650 | SAIAB | *Electrona risso* | 25-Jan-2010 | South Africa | FAO-47 | -32,317 | 16,45 | 652 | 0 |
| DSSAU | ADC10_86.999 #2 | DSFSG413-10 | BOLD:AAB9650 | SAIAB | *Electrona risso* | 25-Jan-2010 | South Africa | FAO-47 | -32,317 | 16,45 | 652 | 0 |
| DSSAU | ADC10_86.999 #3 | DSFSG417-10 | BOLD:AAB9650 | SAIAB | *Electrona risso* | 25-Jan-2010 | South Africa | FAO-47 | -32,317 | 16,45 | 652 | 0 |
| DSSAU | ADC10_86.999 #4 | DSFSG403-10 | BOLD:AAB9650 | SAIAB | *Electrona risso* | 25-Jan-2010 | South Africa | FAO-47 | -32,317 | 16,45 | 652 | 0 |
| DSSAU | ADC10_86.999 #5 | DSFSG420-10 | BOLD:AAB9650 | SAIAB | *Electrona risso* | 25-Jan-2010 | South Africa | FAO-47 | -32,317 | 16,45 | 652 | 0 |
| DSSAU | ADC09_86.53#1 | DSFSG335-10 | | SAIAB | *Gymnoscopelus nicholsi* | 19-Jan-2009 | South Africa | | | | 648 | 0 |
| DSSAU | ADC10_86.998 #1 | DSFSF421-09 | | SAIAB | *Lampadena speculigera* | 06-May-2010 | South Africa | FAO-47 | -35,167 | 23,183 | 652 | 0 |
| DSSAU | ADC09_86.97#1 | DSFSG360-10 | BOLD:AAC4451 | SAIAB | *Metelectrona ventralis* | 23-Jan-2009 | South Africa | | | | 652 | 0 |
| DSSAU | ADC09_86.97#2 | DSFSF361-09 | BOLD:AAC4451 | SAIAB | *Metelectrona ventralis* | 23-Jan-2009 | South Africa | | | | 652 | 0 |
| DSSAU | ADC09_86.97#3 | DSFSF362-09 | BOLD:AAC4451 | SAIAB | *Metelectrona ventralis* | 23-Jan-2009 | South Africa | | | | 652 | 0 |
| DSSAU | ADC09_86.97#4 | DSFSF363-09 | BOLD:AAC4451 | SAIAB | *Metelectrona ventralis* | 23-Jan-2009 | South Africa | | | | 652 | 0 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| DSSAU | ADC09_86.97#5 | DSFSF364-09 | BOLD:AAC4451 | SAIAB | *Metelectrona ventralis* | 23-Jan-2009 | South Africa | | | | 652 | 0 |
| DSSAU | ADC10_86.105 #1 | DSFSG393-10 | BOLD:AAK8894 | SAIAB | *Myctophum spinosum* | 09-May-2010 | South Africa | FAO-47 | -34,733 | 24,367 | 652 | 0 |
| DSSAU | ADC09_86.118#1 | DSFSF422-09 | BOLD:AAD1092 | SAIAB | *Scopelopsis multipunctatus* | 01-Jul-2009 | South Africa | | -33,028 | 27,929 | 652 | 0 |
| DSSAU | ADC09_86.118#2 | DSFSF423-09 | BOLD:AAD1092 | SAIAB | *Scopelopsis multipunctatus* | 01-Jul-2009 | South Africa | | -33,028 | 27,929 | 652 | 0 |
| DSSAU | ADC09_86.119 #2 | DSFSF337-09 | BOLD:AAC7164 | SAIAB | *Symbolophorus barnardi* | 20-Jan-2009 | South Africa | | | | 652 | 0 |
| DSSAU | ADC09_86.119#1 | DSFSF336-09 | BOLD:AAC7164 | SAIAB | *Symbolophorus barnardi* | 20-Jan-2009 | South Africa | | | | 652 | 0 |
| DSSAU | ADC09_86.119#3 | DSFSF343-09 | BOLD:AAC7164 | SAIAB | *Symbolophorus barnardi* | 20-Jan-2009 | South Africa | | | | 652 | 0 |
| DSSAU | ACD07_86.120 #1 | DSFSE476-08 | BOLD:AAE2682 | SAIAB | *Symbolophorus boops* | 12-Oct-2007 | South Africa | | -34,383 | 26,067 | 652 | 0 |
| DSSAU | ACD07_86.120 #2 | DSFSE477-08 | BOLD:AAE2682 | SAIAB | *Symbolophorus boops* | 12-Oct-2007 | South Africa | | -34,383 | 26,067 | 652 | 0 |
| DSSAU | ACD07_86.120 #3 | DSFSE478-08 | BOLD:AAE2682 | SAIAB | *Symbolophorus boops* | 12-Oct-2007 | South Africa | | -34,383 | 26,067 | 652 | 0 |
| DSSAU | ACD07_86.120 #4 | DSFSE479-08 | BOLD:AAE2682 | SAIAB | *Symbolophorus boops* | 12-Oct-2007 | South Africa | | -34,383 | 26,067 | 652 | 0 |
| DSSAU | ACD07_86.120 #5 | DSFSE480-08 | BOLD:AAE2682 | SAIAB | *Symbolophorus boops* | 12-Oct-2007 | South Africa | | -34,383 | 26,067 | 652 | 0 |
| DSSAU | ADC10_86.40 #1 | DSFSG260-10 | BOLD:AAE2682 | SAIAB | *Symbolophorus boops* | 01-Aug-2007 | South Africa | | -29,705 | 31,617 | 569 | 0 |
| FISCO | KUL_Art_sko_AAV3FF_00244 | FISCO151-10 | BOLD:AAB5322 | KUL | *Artedidraco skottsbergi* | 2/4/2006 | Antarctica | | -65,001 | 50,036 | 652 | 0 |
| FISCO | KUL_Bat_ant_PS65FF_13171 | FISCO189-10 | BOLD:AAB8351 | KUL | *Bathylagus antarcticus* | 04/22/04 | Antarctica | Lazarev Sea | -64,000 | -0,0038 | 647 | 0 |
| FISCO | KUL_Bat_ant_PS65FF_13190 | FISCO186-10 | BOLD:AAB8351 | KUL | *Bathylagus antarcticus* | 04/23/04 | Antarctica | Lazarev Sea | -63,000 | 0,0031 | 633 | 0 |
| FISCO | KUL_Bat_ant_PS65FF_13191 | FISCO187-10 | BOLD:AAB8351 | KUL | *Bathylagus antarcticus* | 04/23/04 | Antarctica | Lazarev Sea | -63,000 | 0,0031 | 629 | 0 |
| FISCO | KUL_Bat_ant_PS69_FF_0403 | FISCO180-10 | BOLD:AAB8351 | KUL | *Bathylagus antarcticus* | 6/6/2006 | Antarctica | Lazarev Sea | -67,963 | 3,0266 | 632 | 0 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| FISCO | KUL_Bat_ant_PS69_FF_0406 | FISCO188-10 | BOLD:AAB8351 | KUL | *Bathylagus antarcticus* | 6/6/2006 | Antarctica | Lazarev Sea | -67,963 | 3,0266 | 633 | 0 |
| FISCO | KUL_Bat_ant_PS69_FF_0407 | FISCO185-10 | BOLD:AAB8351 | KUL | *Bathylagus antarcticus* | 6/6/2006 | Antarctica | Lazarev Sea | -67,963 | 3,0266 | 644 | 0 |
| FISCO | KUL_Bat_ant_PS69_FF_0929 | FISCO058-10 | BOLD:AAB8351 | KUL | *Bathylagus antarcticus* | 7/23/2006 | Antarctica | Lazarev Sea | -61,974 | -0,026 | 652 | 0 |
| FISCO | KUL_Bat_ant_PS69_FF_1682 | FISCO059-10 | BOLD:AAB8351 | KUL | *Bathylagus antarcticus* | 8/13/2006 | Antarctica | Lazarev Sea | -59,905 | 2,8783 | 652 | 0 |
| FISCO | KUL_Bat_ant_PS69_FF_1684 | FISCO060-10 | BOLD:AAB8351 | KUL | *Bathylagus antarcticus* | 8/13/2006 | Antarctica | Lazarev Sea | -59,905 | 2,8783 | 652 | 0 |
| FISCO | KUL_Bat_ant_PS69FF_0422 | FISCO061-10 | BOLD:AAB8351 | KUL | *Bathylagus antarcticus* | 6/6/2006 | Antarctica | Lazarev Sea | -67,963 | 3,0266 | 652 | 0 |
| FISCO | KUL_Bat_ant_PS69FF_0496 | FISCO062-10 | BOLD:AAB8351 | KUL | *Bathylagus antarcticus* | 7/6/2006 | Antarctica | Lazarev Sea | -68,534 | 2,9295 | 652 | 0 |
| FISCO | KUL_Bat_ant_PS69FF_1622 | FISCO063-10 | BOLD:AAB8351 | KUL | *Bathylagus antarcticus* | 8/11/2006 | Antarctica | Lazarev Sea | -60,470 | 0,0741 | 652 | 0 |
| FISCO | KUL_Bat_sp._AAV3FF_00321 | FISCO157-10 | BOLD:AAC3458 | KUL | *Bathydraco sp.* | 2/19/2006 | Antarctica | | -65,501 | 70,001 | 619 | 0 |
| FISCO | KUL_Bat_sp._AAV3FF_00322 | FISCO158-10 | | KUL | | 2/19/2006 | Antarctica | | -65,501 | 69,997 | 0 | 0 |
| FISCO | KUL_Bat_sp._AAV3FF_00325 | FISCO161-10 | BOLD:AAC3458 | KUL | *Bathydraco sp.* | 2/19/2006 | Antarctica | | -65,501 | 69,997 | 644 | 0 |
| FISCO | KUL_Bat_sp._AAV3FF_00347 | FISCO164-10 | BOLD:AAC3458 | KUL | *Bathydraco sp.* | 2/21/2006 | Antarctica | | -67,045 | 74,968 | 650 | 0 |
| FISCO | KUL_Bat_sp._AAV3FF_00411 | FISCO175-10 | BOLD:AAC3458 | KUL | *Bathydraco sp.* | 2/26/2006 | Antarctica | | -65,002 | 79,988 | 632 | 0 |
| FISCO | KUL_Ben_gla_26311 | FISCO096-10 | | KUL | | | Antarctica | | | | 0 | 0 |
| FISCO | KUL_Ben_gla_26325 | FISCO097-10 | BOLD:AAC5632 | KUL | *Benthosema glaciale* | | Antarctica | | | | 626 | 0 |
| FISCO | KUL_Ben_gla_26339 | FISCO098-10 | BOLD:AAC5632 | KUL | *Benthosema glaciale* | | Antarctica | | | | 648 | 0 |
| FISCO | KUL_Chi_sp._AAV3FF_00283 | FISCO156-10 | | KUL | | 2/11/2006 | Antarctica | | -66,939 | 59,942 | 0 | 0 |
| FISCO | KUL_Chi_sp._AAV3FF_00340 | FISCO163-10 | | KUL | | 2/20/2006 | Antarctica | | -67,181 | 69,977 | 0 | 0 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| FISCO | KUL_Chi_sp._AAV3FF_00352 | FISCO170-10 | | KUL | | 2/21/2006 | Antarctica | | -67,045 | 74,968 | 0 | 0 |
| FISCO | KUL_Chi_sp._AAV3FF_00355 | FISCO172-10 | | KUL | | 2/21/2006 | Antarctica | | -67,045 | 74,968 | 0 | 0 |
| FISCO | KUL_Cry_ant_AAV3FF_00358 | FISCO166-10 | BOLD:ABY5228 | KUL | *Cryodraco antarcticus* | 2/23/2006 | Antarctica | | -62,792 | 75,008 | 619 | 0 |
| FISCO | KUL_Cry_ant_AAV3FF_00373 | FISCO167-10 | | KUL | | 2/24/2006 | Antarctica | | -62,23 | 80,011 | 0 | 0 |
| FISCO | KUL_Cyc_mic_PS69_FF_1695 | FISCO073-10 | | KUL | | 8/13/2006 | Antarctica | | -59,905 | 2,8783 | 0 | 0 |
| FISCO | KUL_Cyc_mic_PS69_FF_1702 | FISCO069-10 | | KUL | | 8/13/2006 | Antarctica | | -59,905 | 2,8783 | 0 | 0 |
| FISCO | KUL_Cyc_mic_PS69_FF_1703 | FISCO070-10 | | KUL | | 8/13/2006 | Antarctica | | -59,905 | 2,8783 | 0 | 0 |
| FISCO | KUL_Cyc_mic_PS69_FF_1704 | FISCO071-10 | | KUL | | 8/13/2006 | Antarctica | | -59,905 | 2,8783 | 0 | 0 |
| FISCO | KUL_Cyc_mic_PS69_FF_1705 | FISCO072-10 | | KUL | | 8/13/2006 | Antarctica | | -59,905 | 2,8783 | 0 | 0 |
| FISCO | KUL_Cyg_maw_AAV3FF_00324 | FISCO160-10 | | KUL | | 2/19/2006 | Antarctica | | -65,501 | 69,997 | 0 | 0 |
| FISCO | KUL_Cyg_maw_AAV3FF_00420 | FISCO176-10 | BOLD:AAC9735 | KUL | *Cygnodraco mawsoni* | 2/27/2006 | Antarctica | | -66,025 | 80,018 | 595 | 0 |
| FISCO | KUL_Dia_raf_27621 | FISCO099-10 | BOLD:AAC7800 | KUL | *Diaphus rafinesquii* | | Antarctica | | | | 629 | 0 |
| FISCO | KUL_Dia_raf_27635 | FISCO100-10 | BOLD:AAC7800 | KUL | *Diaphus rafinesquii* | | Antarctica | | | | 651 | 0 |
| FISCO | KUL_Dia_raf_27649 | FISCO101-10 | BOLD:AAC7800 | KUL | *Diaphus rafinesquii* | | Antarctica | | | | 645 | 0 |
| FISCO | KUL_Ele_ant_AAV3FF_00201 | FISCO088-10 | BOLD:AAB3737 | KUL | *Electrona antarctica* | 2/2/2006 | Antarctica | Cosmonaut Sea | -62,001 | 50,028 | 648 | 745 |
| FISCO | KUL_Ele_ant_AAV3FF_00202 | FISCO089-10 | BOLD:AAB3737 | KUL | *Electrona antarctica* | 2/2/2006 | Antarctica | Cosmonaut Sea | -62,001 | 50,028 | 648 | 745 |
| FISCO | KUL_Ele_ant_AAV3FF_00203 | FISCO090-10 | BOLD:AAB3737 | KUL | *Electrona antarctica* | 2/2/2006 | Antarctica | Cosmonaut Sea | -62,001 | 50,028 | 637 | 0 |
| FISCO | KUL_Ele_ant_AAV3FF_00204 | FISCO091-10 | BOLD:AAB3737 | KUL | *Electrona antarctica* | 2/2/2006 | Antarctica | Cosmonaut Sea | -62,001 | 50,028 | 648 | 745 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| FISCO | KUL_Ele_ant_AAV3FF_00278 | FISCO092-10 | BOLD:AAB3737 | KUL | *Electrona antarctica* | 2/10/2006 | Antarctica | Cooperation Sea | -65,501 | 60,029 | 648 | 745 |
| FISCO | KUL_Ele_ant_AAV3FF_00367 | FISCO093-10 | BOLD:AAB3737 | KUL | *Electrona antarctica* | 2/24/2006 | Antarctica | Cooperation Sea | -61,663 | 80,029 | 648 | 745 |
| FISCO | KUL_Ele_ant_AAV3FF_00369 | FISCO094-10 | BOLD:AAB3737 | KUL | *Electrona antarctica* | 2/24/2006 | Antarctica | Cooperation Sea | -61,663 | 80,029 | 648 | 745 |
| FISCO | KUL_Ele_ant_AAV3FF_00388 | FISCO095-10 | BOLD:AAB3737 | KUL | *Electrona antarctica* | 2/26/2006 | Antarctica | Cooperation Sea | -64,495 | 79,999 | 648 | 745 |
| FISCO | KUL_Ele_ant_PS65_F_10700 | FISCO137-10 | BOLD:AAB3737 | KUL | *Electrona antarctica* | 04/07/04 | Antarctica | Lazarev Sea | -64,337 | -6,0091 | 648 | 0 |
| FISCO | KUL_Ele_ant_PS65_F_10701 | FISCO138-10 | BOLD:AAB3737 | KUL | *Electrona antarctica* | 04/07/04 | Antarctica | Lazarev Sea | -64,337 | -6,0091 | 639 | 745 |
| FISCO | KUL_Ele_ant_PS65_F_10722 | FISCO139-10 | BOLD:AAB3737 | KUL | *Electrona antarctica* | 04/08/04 | Antarctica | Lazarev Sea | -66,337 | -5,9996 | 648 | 0 |
| FISCO | KUL_Ele_ant_PS65_F_10724 | FISCO140-10 | | KUL | *Electrona antarctica* | 04/08/04 | Antarctica | Lazarev Sea | -66,337 | -5,9996 | 0 | 0 |
| FISCO | KUL_Ele_ant_PS65_F_10744 | FISCO141-10 | BOLD:AAB3737 | KUL | *Electrona antarctica* | 04/08/04 | Antarctica | Lazarev Sea | -67,007 | -6,0001 | 648 | 745 |
| FISCO | KUL_Ele_ant_PS65_F_10850 | FISCO142-10 | | KUL | *Electrona antarctica* | 04/13/04 | Antarctica | Lazarev Sea | -66,677 | -3,9923 | 0 | 745 |
| FISCO | KUL_Ele_ant_PS69_F_0291 | FISCO064-10 | BOLD:AAB3737 | KUL | *Electrona antarctica* | 7/4/2006 | Antarctica | Lazarev Sea | -66,498 | 3,0281 | 648 | 745 |
| FISCO | KUL_Ele_ant_PS69_F_0379 | FISCO065-10 | BOLD:AAB3737 | KUL | *Electrona antarctica* | 6/6/2006 | Antarctica | Lazarev Sea | -67,963 | 3,0266 | 648 | 745 |
| FISCO | KUL_Ele_ant_PS69_F_0381 | FISCO066-10 | BOLD:AAB3737 | KUL | *Electrona antarctica* | 6/6/2006 | Antarctica | Lazarev Sea | -67,963 | 3,0266 | 648 | 745 |
| FISCO | KUL_Ele_ant_PS69_F_0387 | FISCO067-10 | BOLD:AAB3737 | KUL | *Electrona antarctica* | 6/6/2006 | Antarctica | Lazarev Sea | -67,963 | 3,0266 | 648 | 745 |
| FISCO | KUL_Ele_ant_PS69_F_0393 | FISCO068-10 | BOLD:AAB3737 | KUL | *Electrona antarctica* | 6/6/2006 | Antarctica | Lazarev Sea | -67,963 | 3,0266 | 648 | 745 |
| FISCO | KUL_Ele_ris_26353 | FISCO102-10 | BOLD:AAB9650 | KUL | *Electrona risso* | | Antarctica | | | | 607 | 0 |
| FISCO | KUL_Ele_ris_27473 | FISCO103-10 | BOLD:AAB9650 | KUL | *Electrona risso* | | Antarctica | | | | 614 | 0 |
| FISCO | KUL_Ele_ris_27565 | FISCO104-10 | BOLD:AAB9650 | KUL | *Electrona risso* | | Antarctica | | | | 614 | 0 |
| FISCO | KUL_Ele_ris_27593 | FISCO105-10 | BOLD:AAB9650 | KUL | *Electrona risso* | | Antarctica | | | | 523 | 0 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| FISCO | KUL_Ele_ris_27607 | FISCO106-10 | BOLD:AAB9650 | KUL | *Electrona risso* | | Antarctica | | | | 605 | 0 |
| FISCO | KUL_Ger_aus_AAV3FF_00260 | FISCO152-10 | BOLD:AAB3172 | KUL | *Gerlachea australis* | 2/5/2006 | Antarctica | | -66,335 | 50,17 | 651 | 0 |
| FISCO | KUL_Gym_acu_AAV3FF_00174 | FISCO149-10 | | KUL | | 1/30/2006 | Antarctica | | -67,944 | 40,06 | 0 | 0 |
| FISCO | KUL_Gym_acu_AAV3FF_00261 | FISCO153-10 | | KUL | | 2/5/2006 | Antarctica | | -66,335 | 50,17 | 0 | 0 |
| FISCO | KUL_Gym_acu_AAV3FF_00341 | FISCO168-10 | | KUL | | 2/20/2006 | Antarctica | | -67,181 | 69,977 | 0 | 0 |
| FISCO | KUL_Gym_bra_PS69_FF_1523 | FISCO074-10 | BOLD:AAB5232 | KUL | *Gymnoscopelus braueri* | 8/10/2006 | Antarctica | Lazarev Sea | -60,021 | -0,0155 | 648 | 745 |
| FISCO | KUL_Gym_bra_PS69_FF_1524 | FISCO075-10 | BOLD:AAB6899 | KUL | *Gymnoscopelus opisthopterusi* | 8/10/2006 | Antarctica | Lazarev Sea | -60,021 | -0,0155 | 648 | 0 |
| FISCO | KUL_Gym_bra_PS69_FF_1796 | FISCO076-10 | BOLD:AAB5232 | KUL | *Gymnoscopelus braueri* | 8/14/2006 | Antarctica | Bouvet Island | -56,014 | 3,3687 | 648 | 745 |
| FISCO | KUL_Gym_bra_PS69_FF_1797 | FISCO077-10 | BOLD:AAB5232 | KUL | *Gymnoscopelus braueri* | 8/14/2006 | Antarctica | Bouvet Island | -56,014 | 3,3687 | 648 | 745 |
| FISCO | KUL_Gym_sp_PS69_FF_1673 | FISCO081-10 | BOLD:AAB6899 | KUL | *Gymnoscopelus opisthopterus* | 8/13/2006 | Antarctica | Lazarev Sea | -59,905 | 2,8783 | 648 | 745 |
| FISCO | KUL_Gym_sp_PS69_FF_1675 | FISCO080-10 | BOLD:AAB5232 | KUL | *Gymnoscopelus braueri* | 8/13/2006 | Antarctica | Lazarev Sea | -59,905 | 2,8783 | 648 | 745 |
| FISCO | KUL_Gym_sp_PS69_FF_1677 | FISCO079-10 | BOLD:AAB6899 | KUL | *Gymnoscopelus opisthopterus* | 8/13/2006 | Antarctica | Lazarev Sea | -59,905 | 2,8783 | 648 | 745 |
| FISCO | KUL_Gym_sp_PS69_FF_1679 | FISCO078-10 | BOLD:AAB6899 | KUL | *Gymnoscopelus opisthopterus* | 8/13/2006 | Antarctica | Lazarev Sea | -59,905 | 2,8783 | 648 | 745 |
| FISCO | KUL_Hyg_hyg_27663 | FISCO107-10 | BOLD:AAC6495 | KUL | *Hygophum hygomii* | | Antarctica | | | | 615 | 0 |
| FISCO | KUL_Hyg_hyg_27677 | FISCO108-10 | BOLD:AAC5632 | KUL | *Benthosema glaciale* | | Antarctica | | | | 640 | 0 |
| FISCO | KUL_Hyg_hyg_27691 | FISCO109-10 | BOLD:AAC6495 | KUL | *Hygophum hygomii* | | Antarctica | | | | 648 | 0 |
| FISCO | KUL_JCR200_106_2_PRY_1 | FISCO030-10 | | KUL | | 3/25/2009 | Antarctica | | | | 0 | 0 |
| FISCO | KUL_JCR200_115_1_GYF_22 | FISCO036-10 | BOLD:ACE4884 | KUL | *Gymnoscopelus fraseri* | 3/27/2009 | Antarctica | Scotia Sea | -56,806 | -42,232 | 648 | 0 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| FISCO | KUL_JCR200_127_2_GYJ_7 | FISCO035-10 | | KUL | | 3/27/2009 | Antarctica | | -56,762 | -42,272 | 0 | 0 |
| FISCO | KUL_JCR200_128_1_ELC_1 | FISCO002-10 | BOLD:AAB8433 | KUL | *Electrona carlsbergi* | 3/28/2009 | Antarctica | Scotia Sea | -56,768 | -42,180 | 648 | 0 |
| FISCO | KUL_JCR200_128_1_ELC_2 | FISCO003-10 | | KUL | *Electrona carlsbergi* | 3/28/2009 | Antarctica | Scotia Sea | -56,768 | -42,180 | 0 | 0 |
| FISCO | KUL_JCR200_141_1_GYF_4 | FISCO037-10 | BOLD:ACE4884 | KUL | *Gymnoscopelus fraseri* | 3/30/2009 | Antarctica | Scotia Sea | -55,265 | -41,347 | 648 | 745 |
| FISCO | KUL_JCR200_17_1_PRY_41 | FISCO026-10 | BOLD:AAD9636 | KUL | *Protomyctophum choriodon* | 3/16/2009 | Antarctica | South Orkney Islands | -60,493 | -48,228 | 648 | 0 |
| FISCO | KUL_JCR200_17_2_GYN_68 | FISCO043-10 | BOLD:AAB0899 | KUL | *Gymnoscopelus nicholsi* | 3/16/2009 | Antarctica | South Orkney Islands | -60,488 | -48,266 | 648 | 0 |
| FISCO | KUL_JCR200_17_2_PRY_69 | FISCO027-10 | | KUL | | 3/16/2009 | Antarctica | | -60,488 | -48,266 | 0 | 0 |
| FISCO | KUL_JCR200_18_1_GYO_23 | FISCO032-10 | BOLD:AAB6899 | KUL | *Gymnoscopelus opisthopterus* | | Antarctica | Scotia Sea | | | 625 | 0 |
| FISCO | KUL_JCR200_185_2_GYF_36 | FISCO038-10 | BOLD:ACE4884 | KUL | *Gymnoscopelus fraseri* | 4/4/2009 | Antarctica | South Georgia | -52,822 | -39,902 | 648 | 745 |
| FISCO | KUL_JCR200_195_8_GYF_ | FISCO039-10 | BOLD:ACE4884 | KUL | *Gymnoscopelus fraseri* | | Antarctica | Scotia Sea | | | 640 | 745 |
| FISCO | KUL_JCR200_225_1_KRA_31 | FISCO006-10 | BOLD:AAC6875 | KUL | *Krefftichthys anderssoni* | 4/9/2009 | Antarctica | Southern Atlantic Ocean | -50,047 | -33,900 | 652 | 0 |
| FISCO | KUL_JCR200_225_1_KRA_32 | FISCO007-10 | BOLD:AAC6875 | KUL | *Krefftichthys anderssoni* | 4/9/2009 | Antarctica | Southern Atlantic Ocean | -50,047 | -33,900 | 652 | 0 |
| FISCO | KUL_JCR200_225_1_KRA_33 | FISCO008-10 | BOLD:AAC6875 | KUL | *Krefftichthys anderssoni* | 4/9/2009 | Antarctica | Southern Atlantic Ocean | -50,047 | -33,900 | 652 | 0 |
| FISCO | KUL_JCR200_225_1_KRA_34 | FISCO009-10 | BOLD:AAC6875 | KUL | *Krefftichthys anderssoni* | 4/9/2009 | Antarctica | Southern Atlantic Ocean | -50,047 | -33,900 | 652 | 0 |
| FISCO | KUL_JCR200_225_1_KRA_35 | FISCO010-10 | | KUL | *Krefftichthys anderssoni* | 4/9/2009 | Antarctica | Southern Atlantic Ocean | -50,047 | -33,900 | 652 | 0 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| FISCO | KUL_JCR200_225_2_PRP_11 | FISCO031-10 | BOLD:ACF2373 | KUL | *Protomyctophum parallelum* | 4/10/2009 | Antarctica | Southern Atlantic Ocean | -50,041 | -33,828 | 652 | 0 |
| FISCO | KUL_JCR200_226_1_PRA_16 | FISCO018-10 | | KUL | *Protomyctophum andriashevi* | 4/10/2009 | Antarctica | Southern Atlantic Ocean | -50,047 | -33,743 | 0 | 0 |
| FISCO | KUL_JCR200_226_1_PRM_28 | FISCO011-10 | | KUL | *Protomyctophum bolini* | 4/10/2009 | Antarctica | Southern Atlantic Ocean | -50,047 | -33,743 | 0 | 0 |
| FISCO | KUL_JCR200_226_1_PRM_29 | FISCO012-10 | | KUL | *Protomyctophum bolini* | 4/10/2009 | Antarctica | Southern Atlantic Ocean | -50,047 | -33,743 | 0 | 0 |
| FISCO | KUL_JCR200_226_1_PRM_30 | FISCO013-10 | BOLD:AAC7885 | KUL | *Protomyctophum bolini* | 4/10/2009 | Antarctica | Southern Atlantic Ocean | -50,047 | -33,743 | 648 | 0 |
| FISCO | KUL_JCR200_226_1_PRM_31 | FISCO014-10 | | KUL | *Protomyctophum bolini* | 4/10/2009 | Antarctica | Southern Atlantic Ocean | -50,047 | -33,743 | 0 | 0 |
| FISCO | KUL_JCR200_226_1_PRM_32 | FISCO015-10 | | KUL | *Protomyctophum bolini* | 4/10/2009 | Antarctica | Southern Atlantic Ocean | -50,047 | -33,743 | 0 | 0 |
| FISCO | KUL_JCR200_226_2_ELS_8 | FISCO041-10 | | KUL | | 4/10/2009 | Antarctica | | -50,044 | -33,666 | 0 | 0 |
| FISCO | KUL_JCR200_226_2_PRE_34 | FISCO021-10 | BOLD:AAD9622 | KUL | *Protomyctophum tenisoni* | 4/10/2009 | Antarctica | Southern Atlantic Ocean | -50,044 | -33,666 | 648 | 0 |
| FISCO | KUL_JCR200_226_2_PRE_35 | FISCO022-10 | BOLD:AAD9622 | KUL | *Protomyctophum tenisoni* | 4/10/2009 | Antarctica | Southern Atlantic Ocean | -50,044 | -33,666 | 648 | 0 |
| FISCO | KUL_JCR200_226_2_PRE_36 | FISCO023-10 | BOLD:AAD9622 | KUL | *Protomyctophum tenisoni* | 4/10/2009 | Antarctica | Southern Atlantic Ocean | -50,044 | -33,666 | 648 | 0 |
| FISCO | KUL_JCR200_226_2_PRE_37 | FISCO024-10 | BOLD:AAD9622 | KUL | *Protomyctophum tenisoni* | 4/10/2009 | Antarctica | Southern Atlantic Ocean | -50,044 | -33,666 | 648 | 0 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| FISCO | KUL_JCR200_226_2_PRE_38 | FISCO025-10 | BOLD:AAD9622 | KUL | *Protomyctophum tenisoni* | 4/10/2009 | Antarctica | Southern Atlantic Ocean | -50,044 | -33,666 | 652 | 0 |
| FISCO | KUL_JCR200_227_1_GYB_6 | FISCO034-10 | BOLD:AAC5123 | KUL | *Gymnoscopelus braueri* | | Antarctica | Scotia Sea | | | 648 | 745 |
| FISCO | KUL_JCR200_227_2_ELS_86 | FISCO042-10 | | KUL | | 4/10/2009 | Antarctica | | -50,050 | -33,512 | 0 | 0 |
| FISCO | KUL_JCR200_227_2_PRA_87 | FISCO019-10 | BOLD:AAC4230 | KUL | *Protomyctophum andriashevi* | 4/10/2009 | Antarctica | Southern Atlantic Ocean | -50,050 | -33,512 | 648 | 0 |
| FISCO | KUL_JCR200_227_2_PRA_88 | FISCO020-10 | BOLD:AAC4230 | KUL | *Protomyctophum andriashevi* | 4/10/2009 | Antarctica | Southern Atlantic Ocean | -50,050 | -33,512 | 648 | 0 |
| FISCO | KUL_JCR200_235_1_ELC_34 | FISCO004-10 | | KUL | *Electrona carlsbergi* | 4/10/2009 | Antarctica | Southern Atlantic Ocean | -50,600 | -33,839 | 0 | 745 |
| FISCO | KUL_JCR200_235_1_ELC_35 | FISCO005-10 | BOLD:AAB8433 | KUL | *Electrona carlsbergi* | 4/10/2009 | Antarctica | Southern Atlantic Ocean | -50,600 | -33,839 | 648 | 745 |
| FISCO | KUL_JCR200_235_1_PRM_11 | FISCO016-10 | BOLD:AAC7885 | KUL | *Protomyctophum bolini* | 4/10/2009 | Antarctica | Southern Atlantic Ocean | -50,600 | -33,839 | 648 | 0 |
| FISCO | KUL_JCR200_235_1_PRM_12 | FISCO017-10 | BOLD:AAC7885 | KUL | *Protomyctophum bolini* | 4/10/2009 | Antarctica | Southern Atlantic Ocean | -50,600 | -33,839 | 648 | 0 |
| FISCO | KUL_JCR200_236_2_GYF_49 | FISCO040-10 | | KUL | | 4/11/2009 | Antarctica | | -50,589 | -33,773 | 0 | 745 |
| FISCO | KUL_JCR200_55_1_GYN_32 | FISCO044-10 | BOLD:AAB0899 | KUL | *Gymnoscopelus nicholsi* | 3/20/2009 | Antarctica | South Orkney Islands | -59,688 | -44,067 | 648 | 745 |
| FISCO | KUL_JCR200_55_1_GYN_33 | FISCO045-10 | BOLD:AAB0899 | KUL | *Gymnoscopelus nicholsi* | 3/20/2009 | Antarctica | South Orkney Islands | -59,688 | -44,067 | 648 | 745 |
| FISCO | KUL_JCR200_55_2_PRY_133 | FISCO028-10 | BOLD:AAD9636 | KUL | *Protomyctophum choriodon* | 3/20/2009 | Antarctica | South Orkney Islands | -59,722 | -44,091 | 648 | 745 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| FISCO | KUL_JCR200_55_2_PRY_134 | FISCO029-10 | BOLD:AAD9636 | KUL | *Protomyctophum choriodon* | 3/20/2009 | Antarctica | South Orkney Islands | -59,722 | -44,091 | 648 | 0 |
| FISCO | KUL_JCR200_56_1_GYO_1 | FISCO033-10 | | KUL | | 3/20/2009 | Antarctica | | -59,698 | -44,084 | 0 | 0 |
| FISCO | KUL_JCR200_56_2_ELN_100 | FISCO050-10 | BOLD:AAB3737 | KUL | *Electrona antarctica* | 3/20/2009 | Antarctica | South Orkney Islands | -59,722 | -44,141 | 648 | 0 |
| FISCO | KUL_JCR200_56_2_ELN_101 | FISCO051-10 | BOLD:AAB3737 | KUL | *Electrona antarctica* | 3/20/2009 | Antarctica | South Orkney Islands | -59,722 | -44,141 | 648 | 0 |
| FISCO | KUL_JCR200_56_2_ELN_102 | FISCO052-10 | BOLD:AAB3737 | KUL | *Electrona antarctica* | 3/20/2009 | Antarctica | South Orkney Islands | -59,722 | -44,141 | 648 | 0 |
| FISCO | KUL_JCR200_56_2_ELN_98 | FISCO048-10 | BOLD:AAB3737 | KUL | *Electrona antarctica* | 3/20/2009 | Antarctica | South Orkney Islands | -59,722 | -44,141 | 640 | 0 |
| FISCO | KUL_JCR200_56_2_ELN_99 | FISCO049-10 | BOLD:AAB3737 | KUL | *Electrona antarctica* | 3/20/2009 | Antarctica | South Orkney Islands | -59,722 | -44,141 | 648 | 0 |
| FISCO | KUL_JCR200_56_2_GYN_3 | FISCO047-10 | BOLD:AAB0899 | KUL | *Gymnoscopelus nicholsi* | 3/20/2009 | Antarctica | South Orkney Islands | -59,722 | -44,141 | 648 | 0 |
| FISCO | KUL_JCR200_56_2_GYN_4 | FISCO046-10 | BOLD:AAB0899 | KUL | *Gymnoscopelus nicholsi* | 3/20/2009 | Antarctica | South Orkney Islands | -59,722 | -44,141 | 648 | 0 |
| FISCO | KUL_JCR200_82_2_ELC_4 | FISCO001-10 | | KUL | *Electrona carlsbergi* | 3/23/2009 | Antarctica | Scotia Sea | -58,025 | -42,917 | 0 | 745 |
| FISCO | KUL_Lam_mac_27705 | FISCO110-10 | BOLD:AAB3777 | KUL | *Lampanyctus macdonaldi* | | Antarctica | | | | 630 | 0 |
| FISCO | KUL_Lam_mac_27719 | FISCO111-10 | BOLD:AAB3777 | KUL | *Lampanyctus macdonaldi* | | Antarctica | | | | 652 | 0 |
| FISCO | KUL_Lam_mac_27733 | FISCO112-10 | BOLD:AAC1960 | KUL | *Lobianchia dofleini* | | Antarctica | | | | 624 | 0 |
| FISCO | KUL_Lob_dof_27747 | FISCO113-10 | BOLD:AAC1960 | KUL | *Lobianchia dofleini* | | Antarctica | | | | 648 | 0 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| FISCO | KUL_Lob_dof_27761 | FISCO114-10 | BOLD:AAC1960 | KUL | *Lobianchia gemellarii* | | Antarctica | | | 626 | 0 |
| FISCO | KUL_Lob_dof_27775 | FISCO115-10 | BOLD:AAC1960 | KUL | *Lobianchia dofleini* | | Antarctica | | | 609 | 0 |
| FISCO | KUL_Mac_sp._PS69F_0198 | FISCO179-10 | BOLD:AAA7123 | KUL | *Macrourus sp.* | 7/3/2006 | Antarctica | -65,508 | 3,0498 | 644 | 0 |
| FISCO | KUL_Mur_sp._PS69F_0185 | FISCO178-10 | | KUL | | 7/2/2006 | Antarctica | -65,018 | 0 | 0 | 0 |
| FISCO | KUL_Myc_pun_27873 | FISCO116-10 | BOLD:AAB5848 | KUL | *Myctophum punctatum* | | Antarctica | | | 621 | 0 |
| FISCO | KUL_Myc_pun_27887 | FISCO117-10 | | KUL | *Myctophum punctatum* | | Antarctica | | | 0 | 0 |
| FISCO | KUL_Myc_pun_27901 | FISCO118-10 | BOLD:AAB5848 | KUL | *Myctophum punctatum* | | Antarctica | | | 596 | 0 |
| FISCO | KUL_Not_coa_AAV3F_00004 | FISCO143-10 | | KUL | | 1/21/2006 | Antarctica | -65,998 | 29,983 | 0 | 0 |
| FISCO | KUL_Not_coa_AAV3F_00009 | FISCO144-10 | | KUL | | 1/23/2006 | Antarctica | -68,693 | 29,981 | 0 | 0 |
| FISCO | KUL_Not_coa_AAV3F_00135 | FISCO147-10 | | KUL | | 1/27/2006 | Antarctica | -63,319 | 40,023 | 0 | 0 |
| FISCO | KUL_Not_coa_AAV3F_00170 | FISCO148-10 | BOLD:AAC0165 | KUL | *Notolepis coatsi* | 1/29/2006 | Antarctica | -65,996 | 39,996 | 645 | 0 |
| FISCO | KUL_Not_coa_AAV3F_00277 | FISCO154-10 | BOLD:AAM8097 | KUL | *Notolepis coatsi* | 2/9/2006 | Antarctica | -64,351 | 59,937 | 652 | 0 |
| FISCO | KUL_Not_coa_AAV3F_00279 | FISCO155-10 | | KUL | | 2/10/2006 | Antarctica | -66,076 | 60,035 | 0 | 0 |
| FISCO | KUL_Not_coa_AAV3F_00377 | FISCO173-10 | BOLD:AAC0164 | KUL | *Notolepis coatsi* | 2/25/2006 | Antarctica | -63,33 | 80,02 | 633 | 0 |
| FISCO | KUL_Not_coa_AAV3F_00378 | FISCO174-10 | BOLD:AAC0164 | KUL | *Notolepis coatsi* | 2/25/2006 | Antarctica | -63,33 | 80,02 | 596 | 0 |
| FISCO | KUL_Not_kem_PS69F_0540 | FISCO182-10 | BOLD:AAA7826 | KUL | | 7/7/2006 | Antarctica | -68,968 33333 | 2,9981 | 632 | 0 |
| FISCO | KUL_Not_kro_27789 | FISCO119-10 | BOLD:AAB7073 | KUL | *Notoscopelus elongatus* | | Antarctica | | | 459 | 0 |
| FISCO | KUL_Not_kro_27803 | FISCO120-10 | BOLD:AAB7073 | KUL | *Notoscopelus elongatus* | | Antarctica | | | 570 | 0 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| FISCO | KUL_Not_kro_27817 | FISCO121-10 | BOLD:AAC5632 | KUL | *Notoscopelus elongatus* | | Antarctica | | | 522 | 0 |
| FISCO | KUL_Ple_ant_AAV3FF_00077 | FISCO085-10 | BOLD:AAB0924 | KUL | *Pleuragramma antarctica* | 1/24/2006 | Antarctica | -68,462 | 35,009 | 651 | 0 |
| FISCO | KUL_Ple_ant_AAV3FF_00078 | FISCO086-10 | BOLD:AAB0924 | KUL | *Pleuragramma antarctica* | 1/24/2006 | Antarctica | -68,462 | 35,009 | 651 | 0 |
| FISCO | KUL_Ple_ant_AAV3FF_00079 | FISCO087-10 | BOLD:AAB0924 | KUL | *Pleuragramma antarctica* | 1/24/2006 | Antarctica | -68,462 | 35,009 | 651 | 0 |
| FISCO | KUL_Ple_ant_AAV3FF_00080 | FISCO145-10 | BOLD:AAB0924 | KUL | *Pleuragramma antarctica* | 1/24/2006 | Antarctica | -68,462 | 35,009 | 651 | 0 |
| FISCO | KUL_Ple_ant_AAV3FF_00081 | FISCO146-10 | BOLD:AAB0924 | KUL | *Pleuragramma antarctica* | 1/24/2006 | Antarctica | -68,462 | 35,009 | 644 | 0 |
| FISCO | KUL_Ple_ant_AAV3FF_00336 | FISCO082-10 | BOLD:AAB0924 | KUL | *Pleuragramma antarctica* | 2/20/2006 | Antarctica | -67,181 | 69,977 | 651 | 0 |
| FISCO | KUL_Ple_ant_AAV3FF_00337 | FISCO083-10 | BOLD:AAB0924 | KUL | *Pleuragramma antarctica* | 2/20/2006 | Antarctica | -67,181 | 69,977 | 651 | 0 |
| FISCO | KUL_Ple_ant_AAV3FF_00338 | FISCO084-10 | BOLD:AAB0924 | KUL | *Pleuragramma antarctica* | 2/20/2006 | Antarctica | -67,181 | 69,977 | 651 | 0 |
| FISCO | KUL_Ple_ant_AAV3FF_00344 | FISCO169-10 | | KUL | | 2/20/2006 | Antarctica | -67,296 | 69,948 | 0 | 0 |
| FISCO | KUL_Ple_ant_AAV3FF_00353 | FISCO171-10 | BOLD:AAB0924 | KUL | *Pleuragramma antarctica* | 2/21/2006 | Antarctica | -67,045 | 74,968 | 615 | 0 |
| FISCO | KUL_Ple_ant_PS69FF_0543 | FISCO183-10 | BOLD:AAB0924 | KUL | *Pleuragramma antarctica* | 7/7/2006 | Antarctica | -69,480 | 3,1518 | 651 | 0 |
| FISCO | KUL_Ple_ant_PS69FF_0544 | FISCO184-10 | BOLD:AAB0924 | KUL | *Pleuragramma antarctica* | 7/7/2006 | Antarctica | -69,480 | 3,1518 | 623 | 0 |
| FISCO | KUL_Ple_ant_PS69FF_0553 | FISCO181-10 | BOLD:AAB0924 | KUL | *Pleuragramma antarctica* | 7/8/2006 | Antarctica | -70,006 | 3,0666 | 644 | 0 |
| FISCO | KUL_Pog_mar_AAV3FF_00323 | FISCO159-10 | BOLD:AAB5322 | KUL | *Pogonophryne marmorata* | 2/19/2006 | Antarctica | -65,501 | 69,997 | 652 | 0 |
| FISCO | KUL_Pog_mar_AAV3FF_00350 | FISCO165-10 | BOLD:AAB5322 | KUL | *Pogonophryne marmorata* | 2/21/2006 | Antarctica | -67,045 | 74,968 | 641 | 0 |
| FISCO | KUL_Pri_eva_AAV3FF_00186 | FISCO150-10 | | KUL | | 1/30/2006 | Antarctica | -68,196 | 40,018 | 0 | 0 |
| FISCO | KUL_Pro_arc_20083 | FISCO124-10 | BOLD:ACR2195 | KUL | *Protomyctophum arcticum* | | Antarctica | | | 629 | 0 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| FISCO | KUL_Pro_arc_27929 | FISCO122-10 | BOLD:ACR2195 | KUL | *Protomyctophum arcticum* | | Antarctica | | | | 637 | 0 |
| FISCO | KUL_Pro_arc_27943 | FISCO126-10 | BOLD:ACR2195 | KUL | *Protomyctophum arcticum* | | Antarctica | | | | 652 | 0 |
| FISCO | KUL_Pro_arc_27957 | FISCO132-10 | BOLD:ACR2195 | KUL | *Protomyctophum arcticum* | | Antarctica | | | | 620 | 0 |
| FISCO | KUL_Pro_arc_27971 | FISCO134-10 | BOLD:ACR2195 | KUL | *Protomyctophum arcticum* | | Antarctica | | | | 652 | 0 |
| FISCO | KUL_Pro_arc_27999 | FISCO136-10 | BOLD:ACR2195 | KUL | *Protomyctophum arcticum* | | Antarctica | | | | 620 | 745 |
| FISCO | KUL_Pro_arc_28027 | FISCO130-10 | | KUL | *Protomyctophum arcticum* | | Antarctica | | | | 0 | 0 |
| FISCO | KUL_Pro_arc_28041 | FISCO133-10 | BOLD:ACR2195 | KUL | *Protomyctophum arcticum* | | Antarctica | | | | 628 | 0 |
| FISCO | KUL_Pro_arc_28055 | FISCO123-10 | BOLD:ACR2195 | KUL | *Protomyctophum arcticum* | | Antarctica | | | | 648 | 0 |
| FISCO | KUL_Pro_arc_28097 | FISCO125-10 | | KUL | *Protomyctophum arcticum* | | Antarctica | | | | 0 | 0 |
| FISCO | KUL_Pro_arc_28111 | FISCO131-10 | BOLD:ACR2195 | KUL | *Protomyctophum arcticum* | | Antarctica | | | | 630 | 0 |
| FISCO | KUL_Pro_arc_28125 | FISCO135-10 | BOLD:ACR2195 | KUL | *Protomyctophum arcticum* | | Antarctica | | | | 635 | 0 |
| FISCO | KUL_Pro_bol_PS69_FF_1788 | FISCO053-10 | BOLD:AAC7885 | KUL | *Protomyctophum bolini* | 8/14/2006 | Antarctica | Lazarev Sea | -56,014 | 3,3687 | 648 | 0 |
| FISCO | KUL_Pro_bol_PS69_FF_1791 | FISCO054-10 | BOLD:AAC7885 | KUL | *Protomyctophum bolini* | 8/14/2006 | Antarctica | Lazarev Sea | -56,014 | 3,3687 | 648 | 0 |
| FISCO | KUL_Pro_bol_PS69_FF_1792 | FISCO055-10 | BOLD:AAC7885 | KUL | *Protomyctophum bolini* | 8/14/2006 | Antarctica | Lazarev Sea | -56,014 | 3,3687 | 648 | 745 |
| FISCO | KUL_Pro_bol_PS69_FF_1794 | FISCO056-10 | BOLD:AAC7885 | KUL | *Protomyctophum bolini* | 8/14/2006 | Antarctica | Lazarev Sea | -56,014 | 3,3687 | 648 | 0 |
| FISCO | KUL_Pro_bol_PS69_FF_1795 | FISCO057-10 | BOLD:AAC7885 | KUL | *Protomyctophum bolini* | 8/14/2006 | Antarctica | Lazarev Sea | -56,014 | 3,3687 | 648 | 0 |
| FISCO | KUL_Sym_ver_27831 | FISCO127-10 | BOLD:AAU6312 | KUL | *Symbolophorus veranyi* | | Antarctica | | | | 589 | 0 |
| FISCO | KUL_Sym_ver_27845 | FISCO128-10 | BOLD:AAC4870 | KUL | *Symbolophorus veranyi* | | Antarctica | | | | 629 | 0 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| FISCO | KUL_Sym_ver_27859 | FISCO129-10 | | KUL | *Symbolophorus veranyi* | | Antarctica | | | 0 | 0 |
| FISCO | KUL_Tre_sp._AAV3FF_00339 | FISCO162-10 | | KUL | | 2/20/2006 | Antarctica | -67,181 | 69,977 | 0 | 0 |
| FISCO | KUL_Tre_sp._PS69A_0048 | FISCO177-10 | BOLD:AAA6710 | KUL | *Trematomus sp.* | 6/28/2006 | Antarctica | -61,953 | 3,096 | 650 | 0 |
| FISCO | KUL_Unid_sp._PS65FF_644_10 | FISCO190-10 | | KUL | | 04/14/04 | Antarctica | -64,34 | -4,007 | 0 | 0 |
| MYCSO | AM2013-St51E1 | MYCSO026-13 | BOLD:ACN1843 | MNHN | *Ceratoscopelus warmingii* | 14-Feb-2013 | Pacific Ocean | -16,012 | -176,10 | 651 | 0 |
| MYCSO | AM2013-St51E2 | MYCSO027-13 | BOLD:ABV1888 | MNHN | *Symbolophorus* sp. | 14-Feb-2013 | Pacific Ocean | -16,012 | -176,10 | 651 | 814 |
| MYCSO | AM2013-St51F | MYCSO028-13 | BOLD:AAE0838 | MNHN | *Diaphus brachycephalus* | 14-Feb-2013 | Pacific Ocean | -16,012 | -176,10 | 651 | 813 |
| MYCSO | AM2013-St51G | MYCSO029-13 | BOLD:AAV7884 | MNHN | *Lampadena luminosa* | 14-Feb-2013 | Pacific Ocean | -16,012 | -176,10 | 651 | 814 |
| MYCSO | AM2013-St52D | MYCSO052-13 | BOLD:AAC6978 | MNHN | *Diaphus* | 15-Feb-2013 | Pacific Ocean | -19,998 | -176,07 | 651 | 814 |
| MYCSO | AM2013-St52H | MYCSO056-13 | BOLD:ACI7335 | MNHN | *Diaphus* | 15-Feb-2013 | Pacific Ocean | -19,998 | -176,07 | 651 | 814 |
| MYCSO | AM2013-St52J | MYCSO058-13 | | MNHN | *Lampanyctus* | 15-Feb-2013 | Pacific Ocean | -19,998 | -176,07 | 0 | 814 |
| MYCSO | AM2013-St52L | MYCSO060-13 | BOLD:AAG0288 | MNHN | *Lampanyctus* | 15-Feb-2013 | Pacific Ocean | -19,998 | -176,07 | 651 | 814 |
| MYCSO | AM2013-St52M | MYCSO016-13 | | MNHN | *Diaphus* | 15-Feb-2013 | Pacific Ocean | -19,998 | -176,07 | 0 | 814 |
| MYCSO | AM2013-St52N | MYCSO017-13 | BOLD:AAI3244 | MNHN | *Myctophum* | 15-Feb-2013 | Pacific Ocean | -19,998 | -176,07 | 651 | 814 |
| MYCSO | AM2013-St52P | MYCSO019-13 | BOLD:ACI7353 | MNHN | *Diogenichthys* | 15-Feb-2013 | Pacific Ocean | -19,998 | -176,07 | 651 | 787 |
| MYCSO | AM2013-St53B | MYCSO067-13 | | MNHN | *Symbolophorus* | 15-Feb-2013 | Pacific Ocean | -19,992 | -178,00 | 0 | 814 |
| MYCSO | AM2013-St53D | MYCSO069-13 | BOLD:AAC6978 | MNHN | *Diaphus* | 15-Feb-2013 | Pacific Ocean | -19,992 | -178,00 | 651 | 790 |
| MYCSO | AM2013-St56A | MYCSO063-13 | BOLD:ACI7772 | MNHN | *Centrobranchus* | 18-Feb-2013 | Pacific Ocean | -22,003 | -167,99 | 651 | 0 |
| MYCSO | AM2013-St56B | MYCSO064-13 | BOLD:AAD4625 | MNHN | *Symbolophorus evermanni* | 18-Feb-2013 | Pacific Ocean | -22,003 | -167,99 | 651 | 814 |
| MYCSO | AM2013-St56F | MYCSO033-13 | BOLD:AAF2697 | MNHN | *Notolychnus valdiviae* | 18-Feb-2013 | Pacific Ocean | -22,003 | -167,99 | 651 | 0 |
| MYCSO | AM2013-St56G | MYCSO034-13 | | MNHN | *Lampanyctus alatus* | 18-Feb-2013 | Pacific Ocean | -22,003 | -167,99 | 0 | 814 [5n] |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| MYCSO | AM2013-St56H | MYCSO035-13 | | MNHN | *Myctophidae* | 18-Feb-2013 | Pacific Ocean | -22,003 | -167,99 | 0 | 814 |
| MYCSO | AM2013-St56L | MYCSO039-13 | BOLD:AAF2697 | MNHN | *Notolychnus valdiviae* | 18-Feb-2013 | Pacific Ocean | -22,003 | -167,99 | 651 | 0 |
| MYCSO | AM2013-St56R | MYCSO046-13 | BOLD:ACI7621 | MNHN | *Myctophidae* | 18-Feb-2013 | Pacific Ocean | -22,003 | -167,99 | 651 | 814 |
| MYCSO | AM2013-St58A1 | MYCSO048-13 | BOLD:AAI3244 | MNHN | *Myctophum* sp. | 18-Feb-2013 | Pacific Ocean | -18,149 | -169,95 | 651 | 814 |
| MYCSO | AM2013-St58B | MYCSO062-13 | BOLD:ABW8492 | MNHN | *Myctophidae* | 18-Feb-2013 | Pacific Ocean | -18,149 | -169,95 | 651 | 814 |
| MYCSO | AM2013-St58C | MYCSO079-13 | BOLD:AAV7884 | MNHN | *Lampadena luminosa* | 18-Feb-2013 | Pacific Ocean | -18,149 | -169,95 | 651 | 814 |
| MYCSO | AM2013-St58I | MYCSO086-13 | BOLD:AAF2697 | MNHN | *Notolychnus valdiviae* | 18-Feb-2013 | Pacific Ocean | -18,149 | -169,95 | 651 | 814 |
| MYCSO | AM2013-St58J | MYCSO001-13 | BOLD:AAF2697 | MNHN | *Notolychnus valdiviae* | 18-Feb-2013 | Pacific Ocean | -18,149 | -169,95 | 651 | 0 |
| MYCSO | AM2013-St58N | MYCSO083-13 | BOLD:ACN1843 | MNHN | *Ceratoscopelus warmingii* | 18-Feb-2013 | Pacific Ocean | -18,149 | -169,95 | 651 | 0 |
| MYCSO | AM2013-St60A1 | MYCSO077-13 | BOLD:AAK8894 | MNHN | *Myctophum spinosum* | 19-Feb-2013 | Pacific Ocean | -14,002 | -169,99 | 651 | 814 |
| MYCSO | AM2013-St60A2 | MYCSO078-13 | BOLD:AAX3219 | MNHN | *Myctophum* | 19-Feb-2013 | Pacific Ocean | -14,002 | -169,99 | 651 | 814 |
| MYCSO | AM2013-St60E | MYCSO008-13 | BOLD:ACI7790 | MNHN | *Hygophum* | 19-Feb-2013 | Pacific Ocean | -14,002 | -169,99 | 651 | 814 |
| MYCSO | AM2013-St60F | MYCSO009-13 | BOLD:ACI7389 | MNHN | *Myctophidae* | 19-Feb-2013 | Pacific Ocean | -14,002 | -169,99 | 651 | 814 |
| MYCSO | AM2013-St60G | MYCSO010-13 | BOLD:AAC9267 | MNHN | *Diaphus* | 19-Feb-2013 | Pacific Ocean | -14,002 | -169,99 | 651 | 814 |
| MYCSO | AM2013-St64A | MYCSO070-13 | BOLD:AAK8894 | MNHN | *Myctophum spinosum* | 21-Feb-2013 | Pacific Ocean | -8 | -167,99 | 651 | 814 |
| MYCSO | AM2013-St64C | MYCSO072-13 | BOLD:AAE0839 | MNHN | *Diaphus richardsoni* | 21-Feb-2013 | Pacific Ocean | -8 | -167,99 | 613 | 814 |
| MYCSO | AM2013-St64F | MYCSO075-13 | | MNHN | *Symbolophorus* | 21-Feb-2013 | Pacific Ocean | -8 | -167,99 | 0 | 814 |
| MYCSO | AM2013-St65A | MYCSO061-13 | BOLD:ACI7772 | MNHN | *Centrobranchus* sp. | 23-Feb-2013 | Pacific Ocean | -12,014 | -174,04 | 651 | 0 |
| MYCSO | POKER9203 | FKCIR025-11 | | MNHN | *Lampadena sp.* | 28-Aug-2010 | France | -46,387 | 67,943 | 0 | 0 |
| MYCSO | POKER9292 | FKCIR112-11 | BOLD:AAC0459 | MNHN | *Gymnoscopelus hintonoides* | 06-Sep-2010 | France | -46,984 | 70,683 | 652 | 786 |
| MYCSO | POKER9293 | FKCIR113-11 | BOLD:AAC0459 | MNHN | *Gymnoscopelus hintonoides* | 06-Sep-2010 | France | -46,984 | 70,683 | 652 | 791 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| MYCSO | POKER9294 | FKCIR114-11 | BOLD:AAC0459 | MNHN | *Gymnoscopelus hintonoides* | 06-Sep-2010 | France | -46,984 | 70,683 | 652 | 786 |
| MYCSO | POKER9295 | FKCIR115-11 | | MNHN | *Gymnoscopelus hintonoides* | 06-Sep-2010 | France | -46,984 | 70,683 | 0 | 815 |
| MYCSO | POKER9300 | FKCIR120-11 | | MNHN | *Gymnoscopelus bolini* | 06-Sep-2010 | France | -46,984 | 70,683 | 0 | 786 |
| MYCSO | POKER9301 | FKCIR121-11 | | MNHN | *Gymnoscopelus braueri* | 06-Sep-2010 | France | -46,984 | 70,683 | 0 | 0 |
| MYCSO | POKER9302 | FKCIR122-11 | BOLD:AAB5232 | MNHN | *Gymnoscopelus braueri* | 06-Sep-2010 | France | -46,984 | 70,683 | 652 | 786 |
| MYCSO | POKER9303 | FKCIR123-11 | BOLD:AAB5232 | MNHN | *Gymnoscopelus braueri* | 06-Sep-2010 | France | -46,984 | 70,683 | 652 | 786 |
| MYCSO | POKER9304 | FKCIR124-11 | | MNHN | *Gymnoscopelus braueri* | 06-Sep-2010 | France | -46,984 | 70,683 | 0 | 786 |
| MYCSO | POKER9335 | FKCIR155-11 | BOLD:AAC6875 | MNHN | *Krefftichthys anderssoni* | 10-Sep-2010 | France | -50,939 | 71,993 | 652 | 801 |
| MYCSO | POKER9337 | FKCIR157-11 | BOLD:AAB3778 | MNHN | *Nannobrachium achirus* | 10-Sep-2010 | France | -50,939 | 71,993 | 652 | 786 |
| MYCSO | POKER9340 | FKCIR160-11 | BOLD:ACE4884 | MNHN | *Gymnoscopelus fraseri* | 11-Sep-2010 | France | -48,769 | 71,071 | 652 | 801 |
| MYCSO | POKER9347 | FKCIR167-11 | | MNHN | *Gymnoscopelus braueri* | 13-Sep-2010 | France | -50,477 | 71,885 | 0 | 801 |
| MYCSO | POKER9348 | FKCIR168-11 | | MNHN | *Gymnoscopelus microlampas* | 13-Sep-2010 | France | -50,791 | 70,945 | 0 | 0 |
| MYCSO | POKER9349 | FKCIR169-11 | BOLD:AAB5232 | MNHN | *Gymnoscopelus braueri* | 13-Sep-2010 | France | -50,791 | 70,945 | 602 | 801 |
| MYCSO | POKER9350 | FKCIR170-11 | BOLD:AAB5232 | MNHN | *Gymnoscopelus braueri* | 13-Sep-2010 | France | -50,791 | 70,945 | 652 | 786 |
| MYCSO | POKER9367 | FKCIR187-11 | BOLD:AAC8208 | MNHN | *Idiacanthus atlanticus* | 14-Sep-2010 | France | -50,699 | 70,343 | 652 | 801 |
| MYCSO | POKER9434 | FKCIR254-11 | BOLD:AAC8208 | MNHN | *Idiacanthus atlanticus* | 25-Sep-2010 | France | -49,708 | 67,489 | 652 | 786 |
| MYCSO | POKER9441 | FKCIR261-11 | | MNHN | *Electrona antarctica* | 07-Sep-2010 | France | -47,741 | 72,069 | 0 | 0 |
| MYCSO | POKER9442 | FKCIR262-11 | | MNHN | *Electrona carlsbergi* | 28-Aug-2010 | France | -46,398 | 67,617 | 0 | 0 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| MYCSO | POKER9443 | FKCIR263-11 | | MNHN | *Electrona carlsbergi* | 14-Sep-2010 | France | -50,925 | 70,153 | 0 | 0 |
| MYCSO | POKER9444 | FKCIR264-11 | | MNHN | *Electrona antarctica* | 14-Sep-2010 | France | -50,925 | 70,153 | 0 | 0 |
| MYCSO | POKER9445 | FKCIR265-11 | BOLD:AAC6875 | MNHN | *Krefftichthys anderssoni* | 21-Sep-2010 | France | -47,38 | 66,791 | 652 | 0 |
| MYCSO | POKER9446 | FKCIR266-11 | | MNHN | *Gymnoscopelus braueri* | 28-Aug-2010 | France | -46,417 | 67,177 | 0 | 815 |
| MYCSO | POKER9447 | FKCIR267-11 | | MNHN | *Gymnoscopelus braueri* | 07-Sep-2010 | France | -47,677 | 72,087 | 0 | 0 |
| MYCSO | POKER9448 | FKCIR268-11 | | MNHN | *Gymnoscopelus braueri* | 07-Sep-2010 | France | -47,677 | 72,087 | 0 | 815 |
| MYCSO | POKER9450 | FKCIR269-11 | BOLD:AAB3737 | MNHN | *Electrona antarctica* | 13-Sep-2010 | France | -50,791 | 70,945 | 652 | 0 |
| MYCSO | POKER9451 | FKCIR270-11 | BOLD:AAB3737 | MNHN | *Electrona antarctica* | 13-Sep-2010 | France | -50,791 | 70,945 | 652 | 0 |
| MYCSO | POKER9452 | FKCIR271-11 | BOLD:AAB3737 | MNHN | *Electrona antarctica* | 13-Sep-2010 | France | -50,791 | 70,945 | 603 | 0 |
| MYCSO | POKER9453 | FKCIR272-11 | BOLD:AAC6875 | MNHN | *Krefftichthys anderssoni* | 10-Sep-2010 | France | -48,214 | 72,66 | 652 | 0 |
| MYCSO | POKER9454 | FKCIR273-11 | | MNHN | *Electrona carlsbergi* | 27-Sep-2010 | France | -45,015 | 69,338 | 0 | 801 |
| MYCSO | POKER9455 | FKCIR274-11 | | MNHN | *Gymnoscopelus nicholsi* | 06-Sep-2010 | France | -47,241 | 72,038 | 0 | 0 |
| MYCSO | POKER9456 | FKCIR275-11 | | MNHN | *Gymnoscopelus braueri* | 28-Aug-2010 | France | -46,417 | 67,177 | 0 | 0 |
| MYCSO | POKER9457 | FKCIR276-11 | BOLD:AAB5232 | MNHN | *Gymnoscopelus braueri* | 28-Aug-2010 | France | -46,511 | 67,772 | 603 | 815 |
| MYCSO | POKER9458 | FKCIR277-11 | | MNHN | *Gymnoscopelus braueri* | 28-Aug-2010 | France | -46,511 | 67,772 | 0 | 801 |
| MYCSO | POKER9459 | FKCIR278-11 | | MNHN | *Gymnoscopelus braueri* | 28-Aug-2010 | France | -46,387 | 67,943 | 0 | 815 |
| MYCSO | POKER9460 | FKCIR279-11 | | MNHN | *Gymnoscopelus braueri* | 28-Aug-2010 | France | -46,387 | 67,943 | 0 | 0 |
| MYCSO | POKER9461 | FKCIR280-11 | | MNHN | *Nannobrachium achirus* | 16-Sep-2010 | France | -48,092 | 71,263 | 0 | 0 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| MYCSO | POKER9462 | FKCIR281-11 | | MNHN | *Electrona carlsbergi* | 03-Sep-2010 | France | -46,569 | 69,19 | 0 | 0 |
| MYCSO | POKER9463 | FKCIR282-11 | | MNHN | *Electrona carlsbergi* | 03-Sep-2010 | France | -46,569 | 69,19 | 0 | 0 |
| MYCSO | POKER9464 | FKCIR283-11 | | MNHN | *Electrona carlsbergi* | 03-Sep-2010 | France | -46,569 | 69,19 | 0 | 0 |
| MYCSO | POKER9465 | FKCIR284-11 | BOLD:AAB0899 | MNHN | *Gymnoscopelus nicholsi* | 06-Sep-2010 | France | -46,99 | 70,458 | 652 | 0 |
| MYCSO | POKER9466 | FKCIR285-11 | BOLD:AAB0899 | MNHN | *Gymnoscopelus nicholsi* | 06-Sep-2010 | France | -46,99 | 70,458 | 652 | 0 |
| MYCSO | POKER9467 | FKCIR286-11 | BOLD:AAB0899 | MNHN | *Gymnoscopelus nicholsi* | 06-Sep-2010 | France | -46,99 | 70,458 | 652 | 0 |
| MYCSO | POKER9468 | FKCIR287-11 | | MNHN | *Gymnoscopelus nicholsi* | 28-Aug-2010 | France | -46,511 | 67,772 | 0 | 815 |
| MYCSO | POKER9469 | FKCIR288-11 | BOLD:AAB0899 | MNHN | *Gymnoscopelus nicholsi* | 01-Sep-2010 | France | -47,507 | 67,248 | 652 | 0 |
| MYCSO | POKER9470 | FKCIR289-11 | | MNHN | *Gymnoscopelus nicholsi* | 07-Sep-2010 | France | -47,624 | 71,844 | 0 | 815 |
| MYCSO | POKER9471 | FKCIR290-11 | BOLD:AAB0899 | MNHN | *Gymnoscopelus nicholsi* | 07-Sep-2010 | France | -47,624 | 71,844 | 652 | 0 |
| MYCSO | POKER9472 | FKCIR291-11 | | MNHN | *Gymnoscopelus nicholsi* | 31-Aug-2010 | France | -46,736 | 67,422 | 0 | 0 |
| MYCSO | POKER9473 | FKCIR292-11 | | MNHN | *Gymnoscopelus nicholsi* | 31-Aug-2010 | France | -46,736 | 67,422 | 0 | 0 |
| MYCSO | POKER9474 | FKCIR293-11 | | MNHN | *Gymnoscopelus nicholsi* | 06-Sep-2010 | France | -45,052 | 70,424 | 0 | 0 |
| MYCSO | POKER9475 | FKCIR294-11 | BOLD:AAB3737 | MNHN | *Electrona antarctica* | 09-Sep-2010 | France | -48,527 | 72,013 | 603 | 0 |
| MYCSO | POKER9476 | FKCIR295-11 | BOLD:AAB3737 | MNHN | *Electrona antarctica* | 09-Sep-2010 | France | -48,527 | 72,013 | 603 | 0 |
| MYCSO | POKER9477 | FKCIR296-11 | BOLD:AAB3737 | MNHN | *Electrona antarctica* | 09-Sep-2010 | France | -48,527 | 72,013 | 652 | 0 |
| MYCSO | POKER9478 | FKCIR297-11 | BOLD:AAB3737 | MNHN | *Electrona antarctica* | 07-Sep-2010 | France | -46,121 | 71,627 | 603 | 0 |
| MYCSO | POKER9479 | FKCIR298-11 | | MNHN | *Electrona antarctica* | 28-Aug-2010 | France | -46,511 | 67,772 | 0 | 0 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| MYCSO | POKER9480 | FKCIR299-11 | BOLD:AAB3737 | MNHN | *Electrona antarctica* | 13-Sep-2010 | France | -50,477 | 71,885 | 652 | 0 |
| MYCSO | POKER9481 | FKCIR300-11 | BOLD:AAB3737 | MNHN | *Electrona antarctica* | 13-Sep-2010 | France | -50,477 | 71,885 | 652 | 0 |
| MYCSO | POKER9482 | FKCIR301-11 | BOLD:AAB3737 | MNHN | *Electrona antarctica* | 13-Sep-2010 | France | -50,477 | 71,885 | 652 | 0 |
| MYCSO | POKER9483 | FKCIR302-11 | BOLD:AAB3737 | MNHN | *Electrona antarctica* | 13-Sep-2010 | France | -50,477 | 71,885 | 652 | 0 |
| MYCSO | POKER9485 | FKCIR304-11 | BOLD:AAC0459 | MNHN | *Gymnoscopelus hintonoides* | 06-Sep-2010 | France | -47,241 | 72,038 | 603 | 0 |
| MYCSO | POKER9486 | FKCIR305-11 | | MNHN | *Gymnoscopelus bolini* | 28-Aug-2010 | France | -46,511 | 67,772 | 0 | 0 |
| MYCSO | POKER9487 | FKCIR306-11 | BOLD:AAC5123 | MNHN | *Gymnoscopelus bolini* | 06-Sep-2010 | France | -47,241 | 72,038 | 603 | 0 |

**Supplementary Table S2.2.** Cytochrome oxidase subunit I (*COI*) sequences from myctophid fishes included in the Barcode of Life Data Systems (BOLD), but based on our analyses likely mislabeled or misidentified. Except for the private/early-release sequences, these are flagged now in BOLD and thus excluded from the specimen identification engine.

| Sequence ID | Sample ID | Record ID | Source | BIN | Proclaimed species | Likely species | Reference |
|---|---|---|---|---|---|---|---|
| GBGCA11160-15 | KJ555440 | KJ555440 | Mined from GenBank, NCBI | BOLD:AAB3737 | *Lampanyctus achirus* | *Electrona antarctica* | |
| GBGCA11113-15 | KJ555403 | KJ555403 | Mined from GenBank, NCBI | BOLD:AAD9622 | *Krefftichthys anderssoni* | *Protomyctophum tenisoni* | |
| GBGCA11097-15 | KJ555393 | KJ555393 | Mined from GenBank, NCBI | BOLD:AAD8917 | *Gymnoscopelus braueri* | *Electrona paucirastra / Electrona subaspera* | |
| Private | | | | BOLD:AAD9636 | *Hygophum sp.* | *Protomyctophum choriodon* | |
| Early-Release | | | | BOLD:AAC2331 | *Loweina rara* | *Lampadena speculigera* | |
| Private | | | | BOLD:AAC5632 | *Notoscopelus elongatus kroyeri* | *Benthosema glaciale* | |
| MAECO292-06 | ME-4460 | ME-4460 | University of Bergen, Natural History Collections | BOLD:AAB3777 | *Lampanyctus achirus* | *Lampanyctus macdonaldi* | Zhang et al. unpublished |
| MAECO438-09 | ME-5649-2 | ME-5649-2 | University of Bergen, Natural History Collections | BOLD:AAB3779 | *Lampanyctus crocodilus* | *Lampanyctus atrum* | Published, Canadian Barcode of Life |

## IV. S3. Supplementary Material Chapter 3

**Supplementary Table S3.1.** Samples used for reduced representation sequencing (RRS) optimization. DNA from these samples was used for empirical restriction enzyme digestions with different enzymes (single digest *EcoRI*, *PstI*, *MspI*, or double digest *EcoRI-MspI*) and for RRS pilot libraries. Some samples were extracted twice as replicates (marked as _rep in sample ID). Three samples per species (family in the case of ostracods) were used for empirical digestions. The amphipod (*C. obesa*) samples and one *T. loennbergii* were used for empirical digestions, but not included in any RRS library.

| **Species** | **Sample ID** | **Origin** | **Empirical Digestion** | **RRS Library** |
|---|---|---|---|---|
| *Macropyxis hornei* | 280 | ANT XIX-3, St. 46-7-S | *EcoRI, MspI* | 1 |
| *Macrocyprina rocas* | 340 | Buzios | *EcoRI, MspI* | 1 |
| *Macroscapha falcis* | 176 | ANT XXII-3, St. 74-6-S | *EcoRI, MspI* | 1 |
| *Macroscapha falcis* | 187 | ANT XXII-3, St. 74-6-E | - | 1 |
| *Macroscapha solecavai* | 223 | ANT XXII-3, St. 151-7-E | - | 1 |
| *Macroscapha falcis* | 186 | ANT XXII-3, St. 74-6-E | - | 1 |
| *Macroscapha opaca* | 240 | ANT XXII-3, St. 153-7-S | - | 1 |
| *Macroscapha solecavai* | 226 | ANT XXII-3, St. 151-7-E | - | 1 |
| *Macroscapha opaca* | 240_rep | ANT XXII-3, St. 153-7-S | - | 1 |
| *Macroscapa solecavai* | 226_rep | ANT XXII-3, St. 151-7-E | - | 1 |
| *Charcotia obesa* | A77 | ANTXXIX-3 PS81, St. 162-7 | *EcoRI, MspI* | - |
| *Charcotia obesa* | A78 | ANTXXIX-3 PS81, St. 162-7 | *EcoRI, MspI* | - |
| *Charcotia obesa* | A79 | ANTXXIX-3 PS81, St. 162-7 | *EcoRI, MspI* | - |
| *Eusirus* aff. *perdentatus* | HE4 | ANTXXIX-3 PS81, St. 227-2 | - | - |
| *Eusirus* aff. *perdentatus* | HE5 | ANTXXIX-3 PS81, St. 227-2 | - | - |
| *Eusirus* aff. *perdentatus* | HE6 | ANTXXIX-3 PS81, St. 227-2 | - | - |
| *Laternula elliptica* | 4C | - | *EcoRI, PstI, MspI* | 2 |
| *Laternula elliptica* | 5C | - | *EcoRI, PstI, MspI* | 2 |
| *Laternula elliptica* | 6C | - | *EcoRI, PstI, MspI* | 2 |
| *Laternula elliptica* | KGI18 | - | - | 2 |
| *Laternula elliptica* | KGI11 | - | - | 2 |
| *Laternula elliptica* | R5 | - | - | 2 |
| *Laternula elliptica* | R6 | - | - | 2 |
| *Laternula elliptica* | R7 | - | - | 2 |
| *Laternula elliptica* | R6_rep | - | - | 2 |
| *Laternula elliptica* | R7_rep | - | - | 2 |
| *Aequiyoldia eightsii* | 1C | - | *EcoRI, PstI, MspI* | 2 |

| | | | | |
|---|---|---|---|---|
| *Aequiyoldia eightsii* | 2C | - | *EcoRI*, *PstI*, *MspI* | 2 |
| *Aequiyoldia eightsii* | 3C | - | *EcoRI*, *PstI*, *MspI* | 2 |
| *Aequiyoldia eightsii* | R10 | - | - | 2 |
| *Aequiyoldia eightsii* | R11 | - | - | 2 |
| *Aequiyoldia eightsii* | R12 | - | - | 2 |
| *Aequiyoldia eightsii* | R27 | - | - | 2 |
| *Aequiyoldia eightsii* | R28 | - | - | 2 |
| *Aequiyoldia eightsii* | KGI2 | - | - | 2 |
| *Aequiyoldia eightsii* | KGI5 | - | - | 2 |
| *Aequiyoldia eightsii* | R27_rep | - | - | 2 |
| *Aequiyoldia eightsii* | R28_rep | - | - | 2 |
| *Bathybiaster loripes* | Bat004 | Proteker II | - | 3 |
| *Bathybiaster loripes* | Bat062 | CEAMARC | - | 3 |
| *Bathybiaster loripes* | Bat076 | LASSO_ANTXXIX/3 | *EcoRI*, *PstI*, *MspI* | 3 |
| *Bathybiaster loripes* | Bat095 | ANT XXVII/3 (CAMBIO) | - | 3 |
| *Bathybiaster loripes* | Bat096 | ANT XXVII/3 (CAMBIO) | *EcoRI*, *PstI*, *MspI* | 3 |
| *Bathybiaster loripes* | Bat152 | ANT XXVII/3 (CAMBIO) | *EcoRI*, *PstI*, *MspI* | 3 |
| *Bathybiaster loripes* | Bat156 | JR230 | - | 3 |
| *Bathybiaster loripes* | Bat157 | JR275 | - | 3 |
| *Bathybiaster loripes* | Bat164 | POKER II | - | 3 |
| *Bathybiaster loripes* | Bat184 | POKER II | - | 3 |
| *Bathybiaster loripes* | Bat004_rep | Proteker II | - | 3 |
| *Bathybiaster loripes* | Bat062_rep | CEAMARC | - | 3 |
| *Psilaster charcoti* | Psi002 | JR15005 | *EcoRI*, *PstI*, *MspI* | 3 |
| *Psilaster charcoti* | Psi003 | JR15005 | *EcoRI*, *PstI*, *MspI* | 3 |
| *Psilaster charcoti* | Psi008 | JR15005 | - | 3 |
| *Psilaster charcoti* | Psi036 | REVOLTA1 | *EcoRI*, *PstI*, *MspI* | 3 |
| *Psilaster charcoti* | Psi037 | REVOLTA1 | - | 3 |
| *Psilaster charcoti* | Psi039 | JR275 | - | 3 |
| *Psilaster charcoti* | Psi040 | JR275 | - | 3 |
| *Psilaster charcoti* | Psi048 | CEAMARC | - | 3 |
| *Psilaster charcoti* | Psi063 | CEAMARC | - | 3 |
| *Psilaster charcoti* | Psi075 | CEAMARC | - | 3 |
| *Psilaster charcoti* | Psi153 | JR275 | - | 3 |
| *Psilaster charcoti* | Psi155 | JR275 | - | 3 |

| | | | | |
|---|---|---|---|---|
| *Psilaster charcoti* | Psi164 | JR15005 | - | 3 |
| *Psilaster charcoti* | Psi215 | JR308 | - | 3 |
| *Psilaster charcoti* | Psi037_rep | REVOLTA1 | - | 3 |
| *Psilaster charcoti* | Psi039_rep | JR275 | - | 3 |
| *Trematomus bernacchii* | JRI_02 | *see* Jurajda et al., 2016 | - | 4 |
| *Trematomus bernacchii* | JRI_03 | *see* Jurajda et al., 2016 | - | 4 |
| *Trematomus bernacchii* | JRI_04 | *see* Jurajda et al., 2016 | - | 4 |
| *Trematomus bernacchii* | JRI_05 | *see* Jurajda et al., 2016 | - | 4 |
| *Trematomus bernacchii* | JRI_06 | *see* Jurajda et al., 2016 | *EcoRI*, *ApeKI*, *EcoRI-MspI* | 4 |
| *Trematomus bernacchii* | JRI_07 | *see* Jurajda et al., 2016 | - | 4 |
| *Trematomus bernacchii* | JRI_08 | *see* Jurajda et al., 2016 | *EcoRI*, *ApeKI*, *EcoRI-MspI* | 4 |
| *Trematomus bernacchii* | JRI_09 | *see* Jurajda et al., 2016 | - | 4 |
| *Trematomus bernacchii* | JRI_10 | *see* Jurajda et al., 2016 | - | 4 |
| *Trematomus bernacchii* | JRI_11 | *see* Jurajda et al., 2016 | *EcoRI*, *ApeKI*, *EcoRI-MspI* | 4 |
| *Trematomus bernacchii* | JRI_03_rep | *see* Jurajda et al., 2016 | - | 4 |
| *Trematomus bernacchii* | JRI_04_rep | *see* Jurajda et al., 2016 | - | 4 |
| *Trematomus* | ROS_1352 | RSSS 2016 | - | 4 |
| *Trematomus* | ROS_1353 | RSSS 2016 | - | 4 |
| *Trematomus* | ROS_1354 | RSSS 2016 | - | 4 |
| *Trematomus* | ROS_1417 | RSSS 2016 | - | 4 |
| *Trematomus loennbergii* | ROS_1418 | RSSS 2016 | *EcoRI*, *ApeKI*, *EcoRI-MspI* | 4 |
| *Trematomus* | ROS_1419 | RSSS 2016 | - | 4 |
| *Trematomus* | ROS_1420 | RSSS 2016 | - | 4 |
| *Trematomus* | ROS_1421 | RSSS 2016 | - | 4 |
| *Trematomus* | ROS_1484 | RSSS 2016 | - | 4 |
| *Trematomus loennbergii* | ROS_1485 | RSSS 2016 | *EcoRI*, *ApeKI*, *EcoRI-MspI* | 4 |
| *Trematomus loennbergii* | ROS_1487 | RSSS 2016 | *EcoRI*, *ApeKI*, *EcoRI-MspI* | - |
| *Trematomus* | ROS_1352_rep | RSSS 2016 | - | 4 |
| *Trematomus* | ROS_1353_rep | RSSS 2016 | - | 4 |
| *Pagodroma nivea* | 1 | BAS, Rothera Point | - | 5 |

| | | | | |
|---|---|---|---|---|
| *Pagodroma nivea* | 2 | BAS, Storm Ridge | - | 5 |
| *Pagodroma nivea* | 3 | BAS, Storm Ridge | *EcoRI*, *PstI*, *MspI* | 5 |
| *Pagodroma nivea* | 4 | BAS, Signy Island | *EcoRI*, *PstI*, *MspI* | 5 |
| *Pagodroma nivea* | 5 | BAS, Signy Island | *EcoRI*, *PstI*, *MspI* | 5 |
| *Pagodroma nivea* | BEL-G05 | Ut 005 | - | 5 |
| *Pagodroma nivea* | BEL-G81 | Ta 081 | - | 5 |
| *Pagodroma nivea* | BEL-G20 | Pi 020 | - | 5 |
| *Pagodroma nivea* | BEL-G05_rep1 | Ut 005 | - | 5 |
| *Pagodroma nivea* | BEL-G05_rep2 | Ut 005 | - | 5 |

**Supplementary Table S3.2.** *In silico* estimates of the number of fragments produced through restriction enzyme digestions for reduced representation sequencing (RRS) optimized for approximately 30x coverage. The number of fragments depends on the restriction enzyme/combination, the size window, the assumed genome size and the reference genome used for *in silico* digestion. Reference genomes of related species were used as well as simulated genomes; in this case the size and GC content are listed. Only two different enzyme and size selection setups per target species are listed here (the same as in Table 3.3, Table 3.4, Supplementary Table S3.3); further estimates can be found in spreadsheets available at https://doi.org/10.5281/zenodo.3267164.

| Class | Target Species | Restriction Enzyme (Combination) | Size Window (bp) | Assumed Genome Size (Mb) | Reference genome (GC content) | Number of fragments |
|---|---|---|---|---|---|---|
| Ostracoda | Macrocyprididae | *ApeKI* | 250-500 | 250 | *C. torosa* | 88,551 |
| Ostracoda | Macrocyprididae | *ApeKI* | 250-500 | 250 | *D. pulex* | 70,618 |
| Ostracoda | Macrocyprididae | *ApeKI* | 250-500 | 250 | *A. tonsa* | 22,105 |
| Ostracoda | Macrocyprididae | *ApeKI* | 250-500 | 250 | *P. hawaiensis* | 19,078 |
| Ostracoda | Macrocyprididae | *ApeKI* | 250-500 | 250 | 100 Mb (43.9) | |
| Ostracoda | Macrocyprididae | *ApeKI* | 250-500 | 250 | 500 Mb (43.9) | |
| Ostracoda | Macrocyprididae | *ApeKI* | 200-350 | 250 | *C. torosa* | 65,245 |
| Ostracoda | Macrocyprididae | *ApeKI* | 200-350 | 250 | *D. pulex* | 53,331 |
| Ostracoda | Macrocyprididae | *ApeKI* | 200-350 | 250 | *A. tonsa* | 15,122 |
| Ostracoda | Macrocyprididae | *ApeKI* | 200-350 | 250 | *P. hawaiensis* | 14,927 |
| Ostracoda | Macrocyprididae | *ApeKI* | 200-350 | 250 | 100 Mb (43.9) | |
| Ostracoda | Macrocyprididae | *ApeKI* | 200-350 | 250 | 500 Mb (43.9) | |
| Malacostraca | *Charcotia obesa* | *SbfI_MspI* | 250-450 | 27,000 | 34.0x | 97,851 |

† assuming 200 million reads spread over 96 individuals

**Supplementary Table S3.3.** Reduced representation sequencing (RRS) setups for seven individually optimized protocols to be run on a HiSeq 4000 platform (Illumina). Choice of restriction enzyme(s) and size window was optimized in order to obtain approximately 30x coverage (or half that in a worst case scenario) with the assumed genome size. Marker density was estimated as a comparable measure to the metastudy by Lowry et al. (2016).

| Class | Target Species | Restriction Enzyme (Combination) | Size Window (bp) | Assumed Genome Size (Mb) | Coverage† | Marker Density† (bp per 1 SNP) |
|---|---|---|---|---|---|---|
| Ostracoda | Macrocyprididae | *ApeKI* | 250-500 | 250 | 35.3x | 941 |
| Malacostraca | *Charcotia obesa* | *SbfI_MspI* | 250-450 | 27,000 | 34.0x | 97,851 |
| | *Eusirus perdentatus* | *EcoRI_SphI* | 250-350 | 7,000 | 30.7x | 22,895 |
| Bivalvia | *Laternula elliptica* and *Aequiyoldia eightsii* | *ApeKI* | 250-350 | 3,000 | 29.7 – 35.5x | 9,492 – 11,357 |
| Asteroidea | *Bathybiaster loripes* and *Psilaster charcoti* | *ApeKI* | 250-400 | 500 | 31.6 – 39.1x | 1,685 – 2,052 |
| Actinopterygii | *Trematomus bernacchii* and *T. loennbergii* | *EcoRI_MspI* | 250-600 | 1,500 | 33.2x | 4,918 |
| Aves | *Pagodroma nivea nivea* and *P. nivea confusa* | *PstI* | 250-400 | 1,500 | 33.9x | 5,425 |

† assuming 300 million reads of 150 bp length spread over 96 individuals and 0.01 SNP/bp

**Supplementary Protocol S3.4.** Reduced representation sequencing (RRS) laboratory protocol based on the double digest RADseq protocol from Peterson et al. (2012). The protocol is scaled for use with 192 samples and with restriction enzymes *EcoRI* and *MspI*; the reagent volumes can be scaled down/up to suit other sample numbers; if other enzymes are used the respective reaction conditions must be adjusted.

**Step 1.** Prepare two PCR plates with 15 µL of each DNA sample at a concentration of 20 ng/µL

**Step 2.** Restriction enzyme digestion of 192 samples with *EcoRI*-HF and *MspI* (NEB, New England Biolabs)

- Prepare master mix for 220 samples:
    - Cut Smart buffer (NEB, 2 µL per sample): 440 µL
    - *EcoRI*-HF (1 µL per sample): 220 µL
    - *MspI* (1 µL per sample): 220 µL
    - Molecular grade water (1 µL per sample): 220 µL
- Vortex master mix, spin down briefly and put on ice
- Distribute 135 µL of the mix in 8 well strip and add 5 µL to each well of the sample plates with multipipet
- Total volume in each well: 20 µL
- Incubate for **3 h at 37° C**, cool down to 10° C
- Heat inactivation: **20 min at 65° C**, cool down to 10° C

**Step 3.** Ligation (T4 DNA ligase from NEB)

- Prepare master mix for 220 samples in falcon tube
    - Cut Smart buffer (NEB, 2 µl per sample): 440 µL
    - T4 ligase (1 µl per sample): 220 µL
    - *MspI* adaptor (9 µM; 2 µl per sample): 440 µl
    - rATP (1 µl per sample): 220 µl
    - Molecular grade water (12 µl per sample): 2640 µL
- Distribute in a clean plastic tray and add 18 µL to each well containing digested DNA
- Total volume in each well: 38 µL
- Add 2 µl *EcoRI* adaptor (0.6 µM) to each sample: **watch out, there are 8 different adaptors with 8 barcodes**. Put adaptor 1 in wells A1, A2, …till A12. Adaptor 2 in wells B1, B2, … till B12 and the same for the others
- Incubate **30 min at 22° C**, followed by **10 min at 65°C**

**Step 4.** Purification with CleanPCR beads (CleanNA; GC Biotech);
to reduce costs of CleanPCR beads, only 20 µL will be purified

- Add 20 µL beads to new plate
- Add 20 µL of digestion/ligation mixture
- Mix by carefully pipetting up and down 10 times to ensure proper mixing
- Incubate 5 min at room temperature
- Place plate on magnet for 5 min to separate beads from solution
- Remove 35 µL of the clear solution while the plate is still on the magnet. Discard solution. Avoid taking out any beads; leave ca. 5 µL of the solution behind.
- Add 200 µL of 70% ethanol and wait 30 s
- Remove 200 µL ethanol (beads are now attached much better to the wall)
- Add 200 µL of 70% ethanol, wait 30 s

- Remove all supernatant (230 µL of ethanol). Check whether all ethanol is removed. Take 10 µL multipipet to double check whether all wells are empty. Residual ethanol may interfere with downstream PCR
- Remove plate from magnet and add 40 µL elution buffer or pure water (Sigma)
- Mix by pipetting 10 times up and down
- Incubate 5 min
- Put plate on magnet for 5 min to separate beads from solution
- Transfer 30 µL to new plate (make sure to not transfer beads, although they are not necessarily problematic later on)

**Step 5.** PCR on individual samples

- Prepare master mix for 200 samples
    - NEB Q5 hotstart master mix (12.5 µl per sample): 2500 µl
    - Molecular grade water (8.5 µl per sample): 1700 µl
    - F-Primer (5 µM, 1 µl per sample): 200 µl
- Distribute 22 µL of the mix and add 1 µL of R-primer (5 µM): **watch out, there are 12 different primers with 12 barcodes**. Put primer 1 in wells A1, B1, … till H1. Primer 2 in wells A2, B2, … till H2 and the same for the others.
- Add 2 µl of the purified digestion-ligation mix
- Total volume: 25 µL
- Initial denaturation at **98° C for 30 s** followed by 13 cycles of **10 s at 98° C**, **30 s at 65° C** and **30 s at 72°C**. Final elongation **5 min at 72°C**.

**Step 6.** Purification with CleanPCR beads

- Purify PCR product as step 3 but add only 20 µL beads to 25 µL PCR product (0.8 ratio)
- Follow protocol in step 3
- The final elution volume is 25 µL and 20 µL is transferred to a new tube

**Step 7.** Quantification

- Use the Quant-iT PicoGreen protocol (Thermo Fisher Scientific Inc.) and a microplate reader or a similar photometric method to precisely quantify the individual, amplified ddRAD samples following the manufacturer's instructions.
- If DNA quantity at this step is too low, one may try to go back to step 5 and try with more PCR cyles. This also increases the amount of PCR duplicates of course.

**Step 8.** Pooling of the samples

- Depending on the lowest concentration take 20 ng (if possible, otherwise 10 ng or 5 ng) from each sample and transfer into one single tube.
- Quantify the pooled sample with PicoGreen again and check on gel.

**Step 9.** Export to KU Leuven Genomics Core.

The library/libraries are size selected (do not forget to add adaptor length to the chosen size window) and quantified at the Genomics Core using a Pippin Prep (Sage Science) and qPCR, respectively.
See original protocol version for further details:
Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S., Hoekstra, H.E. (2012) Double digest RADseq: an inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. PLoS One 7(5), e37135. https://doi.org/10.1371/journal.pone.0037135

**Supplementary Protocol S3.5.** Reduced representation sequencing (RRS) laboratory protocol based on the genotyping-by-sequencing protocol from Elshire et al. (2011). The protocol is scaled for use with 192 samples and with restriction enzymes *PstI* or *ApeKI*; the reagent volumes can be scaled down/up to suit other sample numbers; if other enzymes are used the respective reaction conditions must be adjusted.

**Step 1.** Prepare two PCR plates with 10 µL of each DNA sample at a concentration of 10 ng/µL

**Step 2.** Restriction enzyme digestion of 192 samples with *PstI* or *ApeKI* (NEB, New England Biolabs)

- Add 6 µL of adaptor (1/10 diluted) to each well with multipipet
- Prepare master mix for 220 samples
    - NEB buffer 3 (2 µL per sample): 440 µL
    - *PstI* or *ApeKI* (1 µL per sample): 220 µL
    - Molecular grade water (1 µL per sample): 220 µL
- Vortex master mix, spin down briefly and put on ice
- Distribute 110 µL of the mix in 8 well strip and add 4 µL to each well of the sample plates with multipipet
- Total volume in each well: 20 µL
- For *PstI*: incubate **2 h at 37° C**, cool down to 10° C
- For *ApeKI*: incubate **2 h at 75° C**, cool down to 10° C

**Step 3.** Ligation (T4 DNA ligase from NEB)

- Prepare master mix for 220 samples in falcon tube
    - 10x T4 DNA ligase buffer (5 µL per sample): 1100 µL
    - T4 ligase (1.2 µL per sample): 264 µL
    - Molecular grade water (23.8 µL per sample): 5236 µL
- Distribute in a clean plastic tray and add 30 µL to each well containing digested DNA with multipipet
- Total volume in each well: 50 µL
- Incubate **1 h at 22° C**, followed by **30 min at 65° C** (heat inactivation of enzyme)

**Step 4.** Purification with CleanPCR beads (CleanNA; GC Biotech);
to reduce costs of CleanPCR beads, only 25 µL will be purified

- Add 25 µL beads to new plate
- Add 25 µL of digestion/ligation mixture
- Mix by carefully pipetting up and down 10 times to ensure proper mixing
- Incubate 5 min at room temperature
- Place plate on magnet for 5 min to separate beads from solution
- Remove 45 µL of the clear solution while the plate is still on the magnet. Discard solution. Avoid taking out any beads; leave ca. 5 µL of the solution behind.
- Add 200 µL of 70% ethanol and wait 30 s
- Remove 200 µL ethanol (beads are now attached much better to the wall)
- Add 200 µL of 70% ethanol, wait 30 s
- Remove all supernatant (230 µL of ethanol). Check whether all ethanol is removed. Take 10 µL multipipet to double check whether all wells are empty. Residual ethanol may interfere with downstream PCR
- Remove plate from magnet and add 40 µL elution buffer (e.g. from Qiagen kit) or pure water (Sigma)
- Mix by pipetting 10 times up and down
- Incubate 5 min

- Put plate on magnet for 5 min to separate beads from solution
- Transfer 35 µL to new plate (make sure to not transfer beads, although they are not necessarily problematic later on)

**Step 5**. PCR on separate samples

- Prepare master mix for 200 samples
    - NEB Q5 hotstart master mix (12.5 µL per sample): 2500 µL
    - Molecular grade water (10.5 µL per sample): 2100 µL
    - Primer mix (contains F and R primer, each at 5 µM, 1 µL per sample): 200 µL
- Distribute 24 µL of the mix and add 1 µL of cleaned ligation product
- Total volume: 25 µL
- Initial denaturation at **98° C for 30 s**, followed by 18 cycles of **10 s at 98° C, 30 s at 65° C** and **30 s at 72° C**. Final elongation **5 min at 72° C**.

**Step 6.** Purification with CleanPCR beads

- Purify PCR product as in step 3 but add only 20 µL of beads to 25 µL PCR product (0.8 ratio)
- Follow protocol in step 3
- The final elution volume is 30 µL and 25 µL is transferred to a new tube

**Step 7.** Quantification

- Use the Quant-iT PicoGreen protocol (Thermo Fisher Scientific Inc.) and a microplate reader or a similar photometric method to precisely quantify the individual, amplified ddRAD samples following the manufacturer's instructions.
- If DNA quantity at this step is too low, one may try to go back to step 5 and try with more PCR cyles. This also increases the amount of PCR duplicates of course.

**Step 8.** Pooling of the samples

- Depending on the lowest concentration take 5 or 10 ng from each sample and transfer into one single tube.
- Quantify the pooled sample with PicoGreen again and check on gel.

**Step 9.** Export to KU Leuven Genomics Core.

The library/libraries are size selected (do not forget to add adaptor length to the chosen size window) and quantified at the Genomics Core using a Pippin Prep (Sage Science) and qPCR, respectively.

See original protocol version for further details:

Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S., Mitchell, S.E. (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One 6(5), e19379. https://doi.org/10.1371/journal.pone.0019379

**Supplementary Material S3.6.** Results from parameter optimization with Stacks v.2.4 (Rochette et al. 2019) for de novo assembly and genotyping of five reduced representation sequencing (RRS) libraries.

Eight parameter optimization series were conducted following Rochette & Catchen (2017); one for each species/species complex. The Stacks parameter m was kept constant (m = 3), while parameters M and n were varied together from 1 to 9. Subsequently, only loci present in 80 % of the samples were retained and for each M=n parameter the number of loci and polymorphic loci was plotted, as well as the proportion of these loci containing 0 to 10 or >10 SNPs. In ostracods, the library contained DNA from a species-complex, resulting in very few shared loci across 80 % of the samples. Therefore, in this case results based on loci shared by 50 % of samples are shown. Optimal M=n values were decided in all cases with this information (and reported in Table 3.4). Note, however, that it is impossible to make absolute calls regarding the ideal value.

References

Rochette, N.C. and Catchen, J.M. (2017) Deriving genotypes from RAD-seq short-read data using Stacks. Nat. Protoc. 12, 2640–2659 https://doi.org/10.1038/nprot.2017.123

Rochette, N.C., Rivera-Colón, A.G. and Catchen, J.M. (2019) STACKS 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. bioRxiv http://dx.doi.org/10.1101/615385

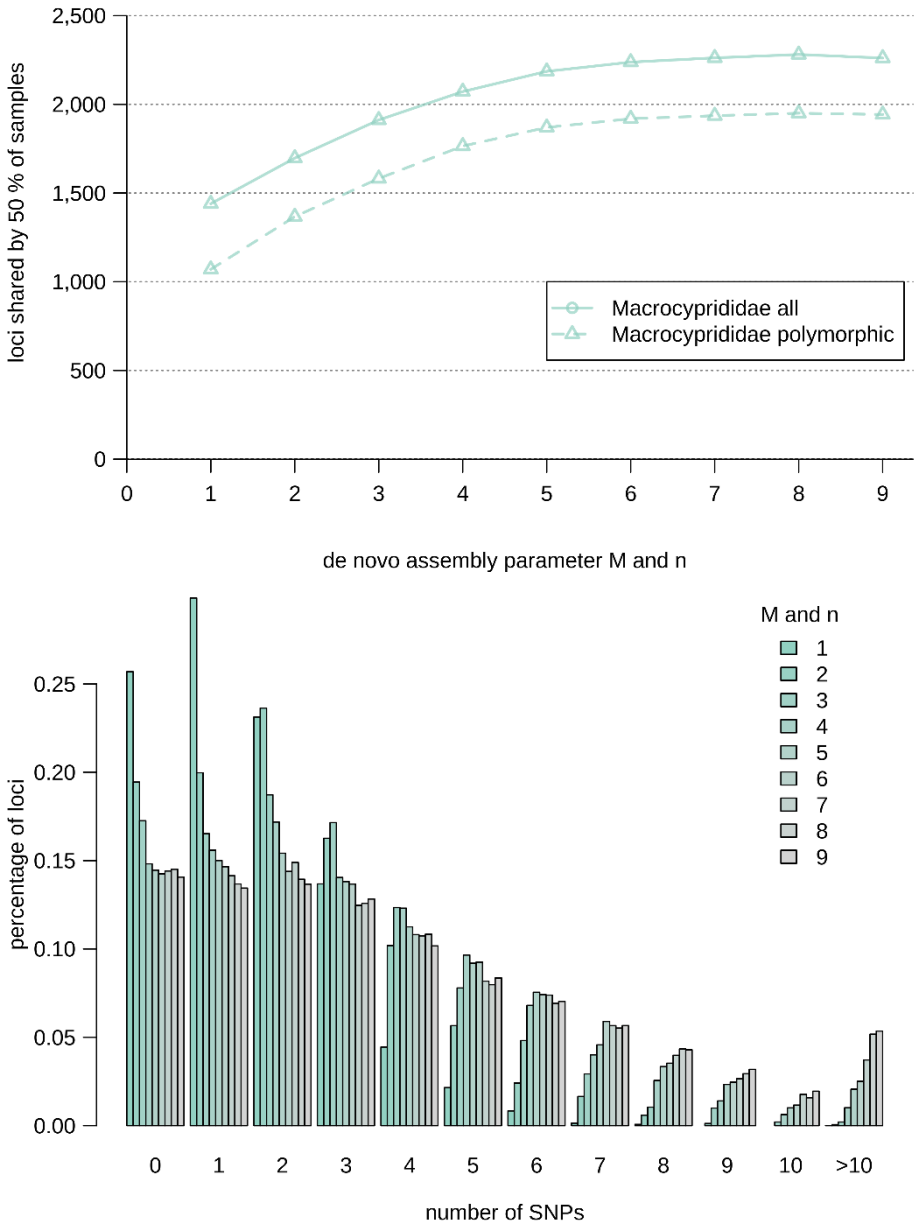



**Figure S3.6.1.** Number of loci and polymorphic loci shared by 50 % of samples from test library 1 across nine values for parameter M and n in Stacks v.2.4 (top) and number of SNPs per locus across the same parameter range (bottom). M = n = 6 was retained. Note that library 1 contained a species complex (*Macroscapha opaca-tensa* complex) with only few loci shared across many samples. Therefore results of loci shared by only 50 % of the samples are shown.

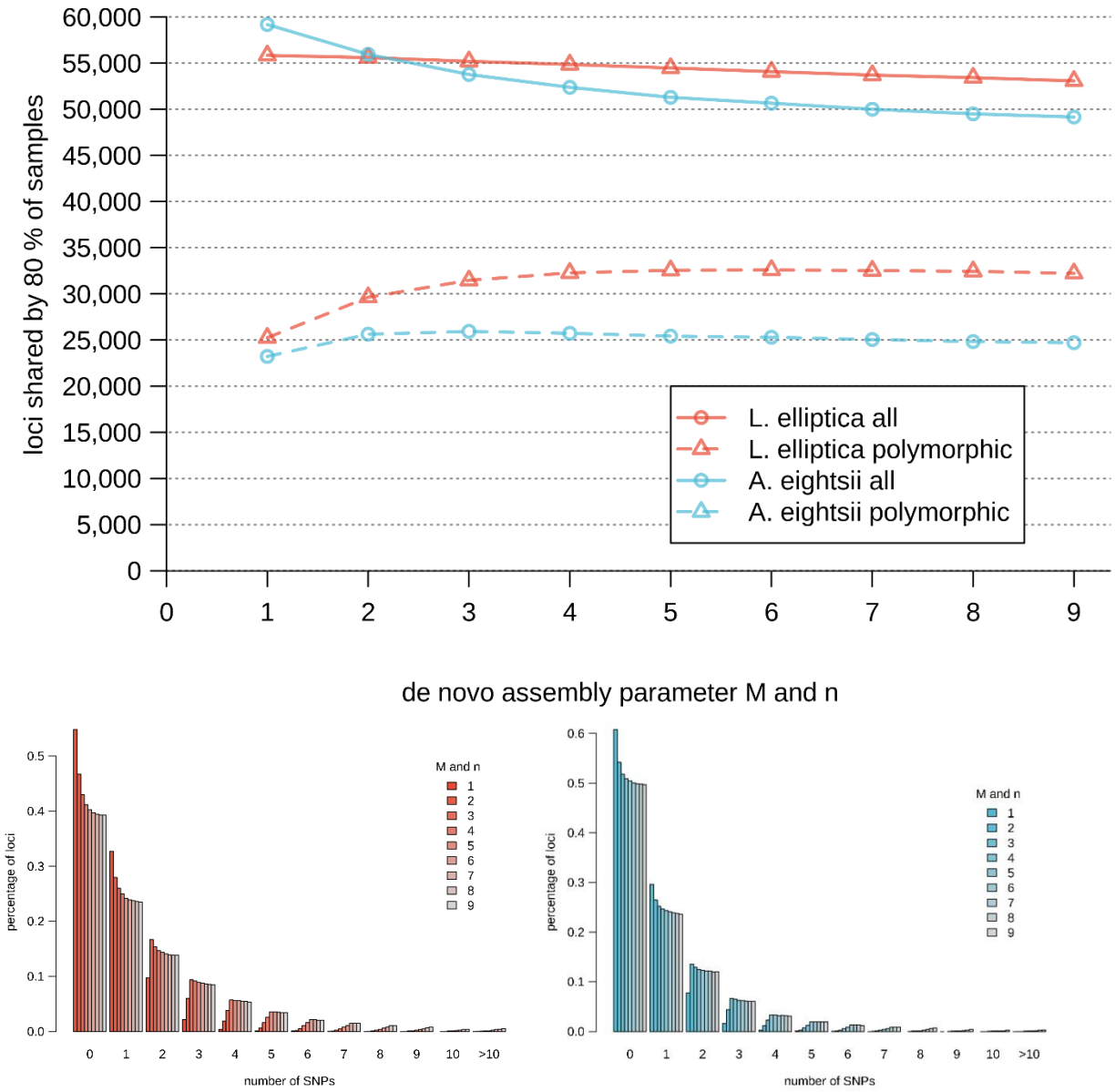



**Figure S3.6.2.** Number of loci and polymorphic loci shared by 80 % of samples from test library 2 across nine values for parameter M and n in Stacks v.2.4 (top) and number of SNPs per locus across the same parameter range for *Laternula elliptica* (bottom left) and *Aequiyoldia eightsii* (bottom right). M = n = 4 was retained.

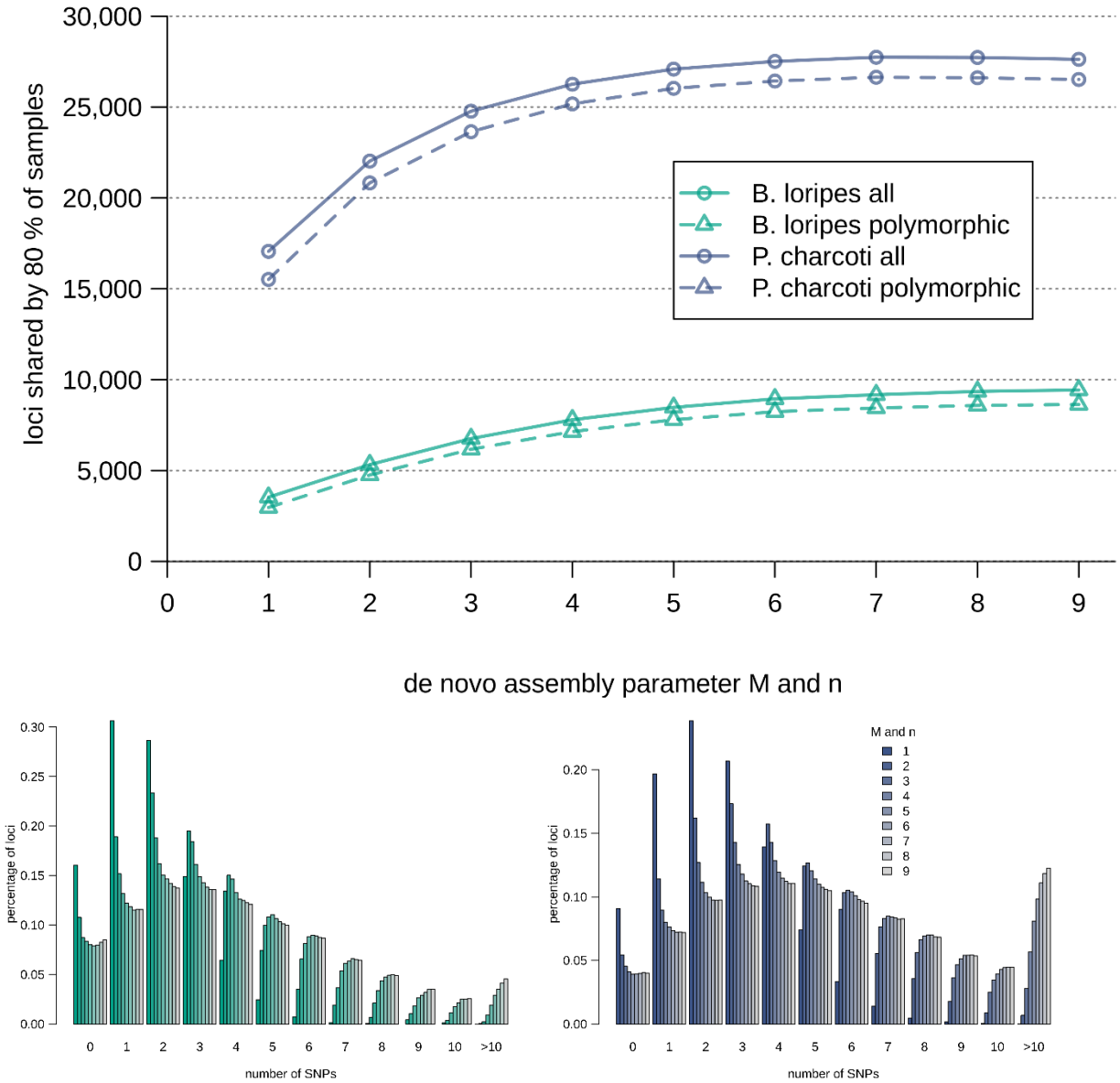



**Figure S3.6.3.** Number of loci and polymorphic loci shared by 80 % of samples from test library 3 across nine values for parameter M and n in Stacks v.2.4 (top) and number of SNPs per locus across the same parameter range for *Bathybiaster loripes* (bottom left) and *Psilaster charcoti* (bottom right). M = n = 5 was retained.

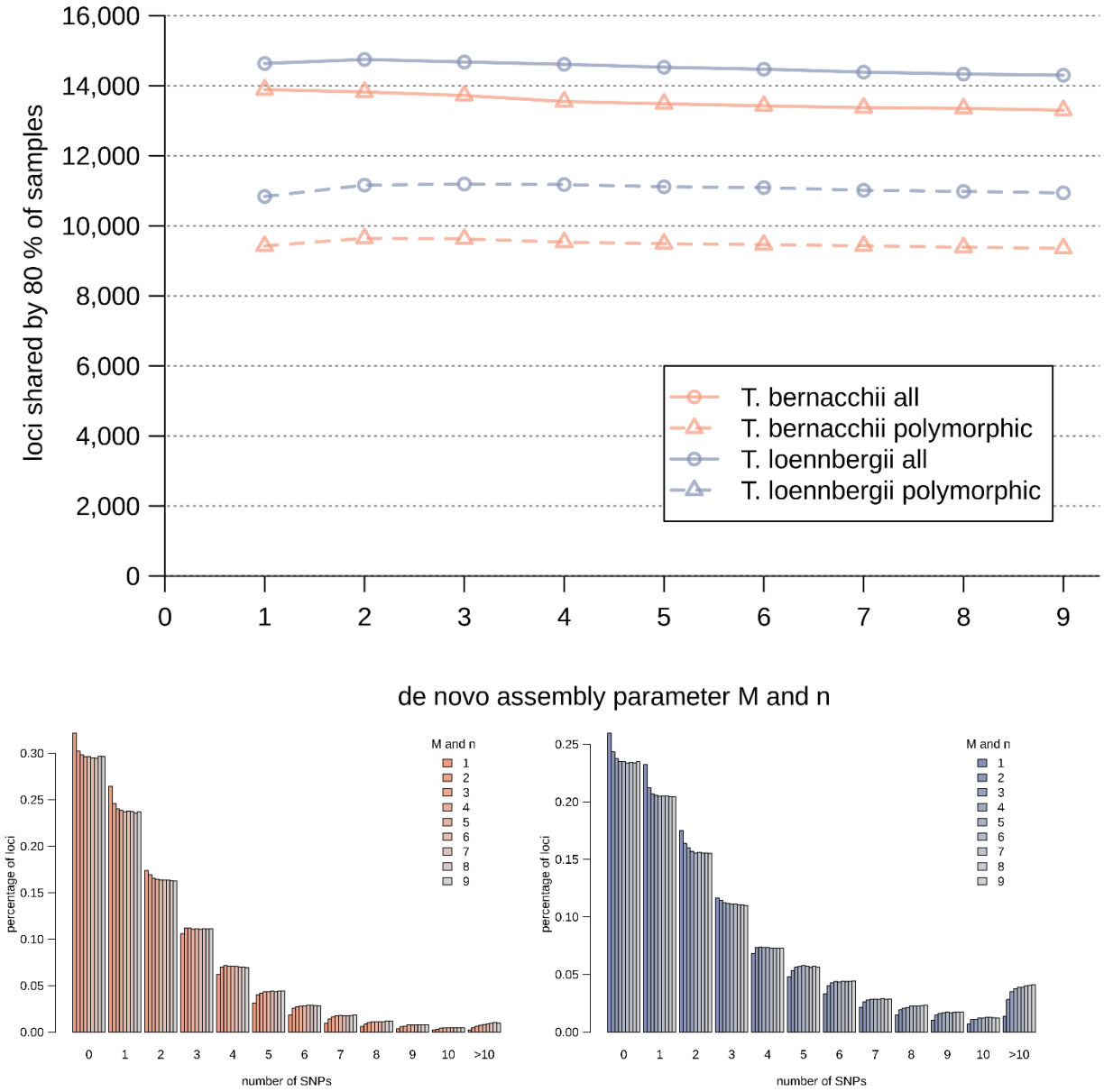



**Figure S3.6.4.** Number of loci and polymorphic loci shared by 80 % of samples from test library 4 across nine values for parameter M and n in Stacks v.2.4 (top) and number of SNPs per locus across the same parameter range for *Trematomus bernacchii* (bottom left) and *T. loennbergii* (bottom right). M = n = 3 was retained.

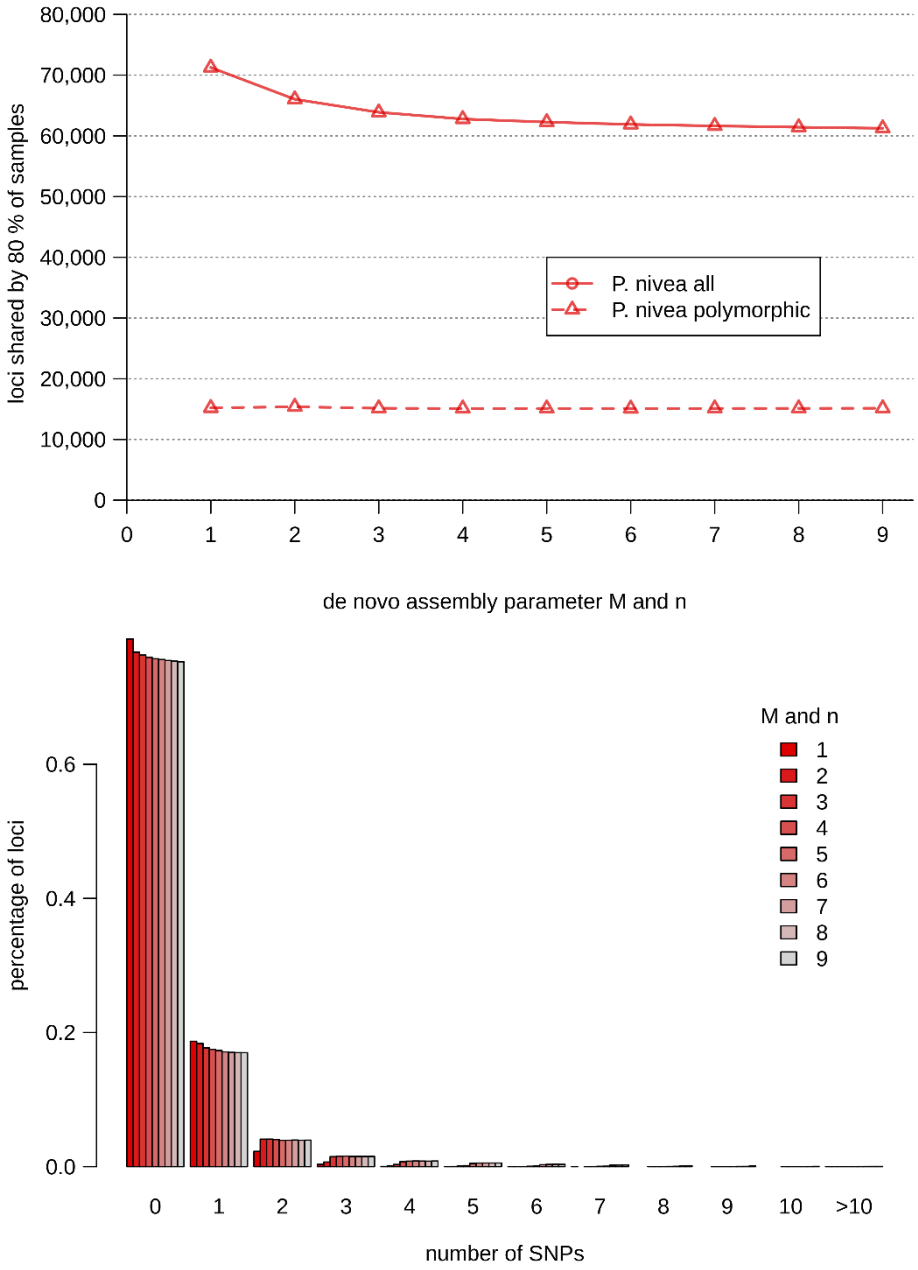



**Figure S3.6.1.** Number of loci and polymorphic loci shared by 50 % of samples from test library 5 across nine values for parameter M and n in Stacks v.2.4 (top) and number of SNPs per locus across the same parameter range (bottom). M = n = 3 was retained. Note that library 2 contained only few samples with likely high levels of degradation, possibly explaining the low amount of polymorphism detected.

## IV. S4. Supplementary Material Chapter 4

**Fig. S4.1.** Observed (Hobs) and expected (Hexp) heterozygosity and standard error of *Notothenia coriiceps* from seven localities in the Southern Ocean based on 1948 SNP genotypes derived from mapping against a de novo assembly. Sampling sites abbreviated as in Table 4.1: NTA: Adélie Land, North; STA: Adélie Land, South; CDI: Deception Island; PKGI: King George Island, Potter Cove; NKGI: King George Island; North; EI: Elephant Island; SO: South Orkney Islands.

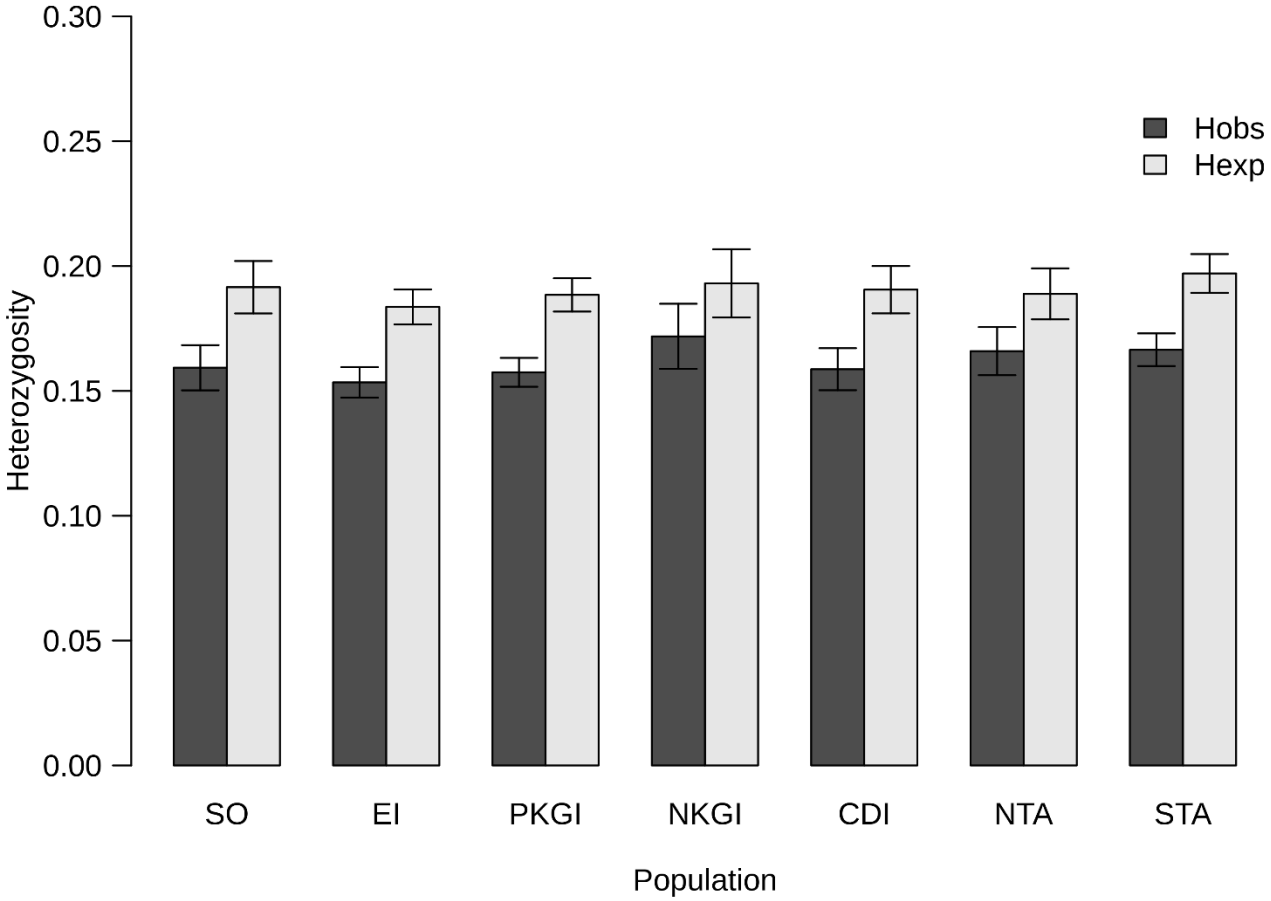


**Fig. S4.2.** Observed (Hobs) and expected (Hexp) heterozygosity and standard error of *Notothenia coriiceps* from seven localities in the Southern Ocean 1572 SNP genotypes derived from mapping against the reference genome of Shin et al. (2014). Sampling sites abbreviated as in Table 4.1: NTA: Adélie Land, North; STA: Adélie Land, South; CDI: Deception Island; PKGI: King George Island, Potter Cove; NKGI: King George Island; North; EI: Elephant Island; SO: South Orkney Islands.

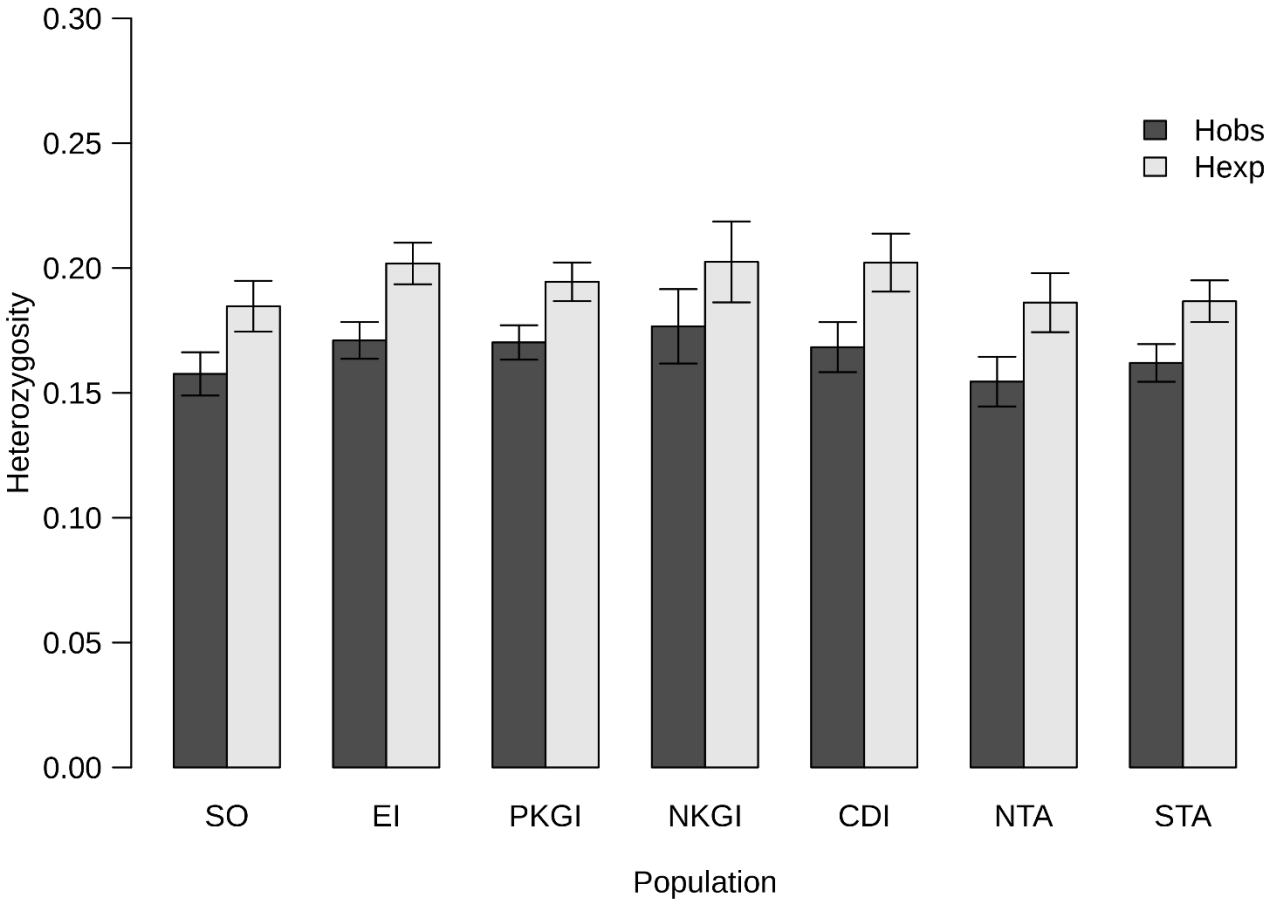


**Fig. S4.3.** Genetic differentiation of *Notothenia coriiceps* in the Southern Ocean as visualized by non-metric multidimensional scaling of pairwise values of Hedrick's $G_{ST}$ (2005) based on 1948 SNP genotypes derived from mapping against a de novo assembly. Sampling sites abbreviated as in Table 4.1: NTA: Adélie Land, North; STA: Adélie Land, South; CDI: Deception Island; PKGI: King George Island, Potter Cove; NKGI: King George Island; North; EI: Elephant Island; SO: South Orkney Islands.

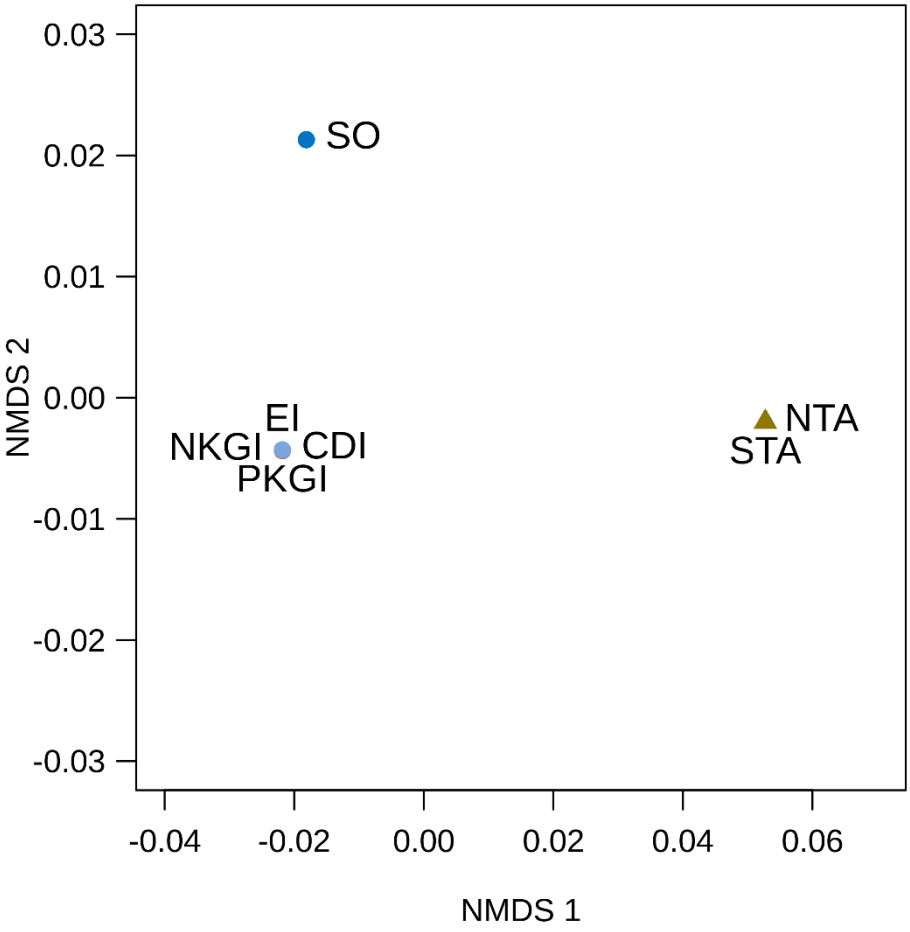


**Fig. S4.4.** Genetic differentiation of Notothenia coriiceps in the Southern Ocean as visualized by non-metric multidimensional scaling of pairwise values of Hedrick's $G_{ST}$ (2005) based on 1572 SNP genotypes derived from mapping against the reference genome of Shin et al. (2014). Sampling sites abbreviated as in Table 4.1: NTA: Adélie Land, North; STA: Adélie Land, South; CDI: Deception Island; PKGI: King George Island, Potter Cove; NKGI: King George Island; North; EI: Elephant Island; SO: South Orkney Islands.

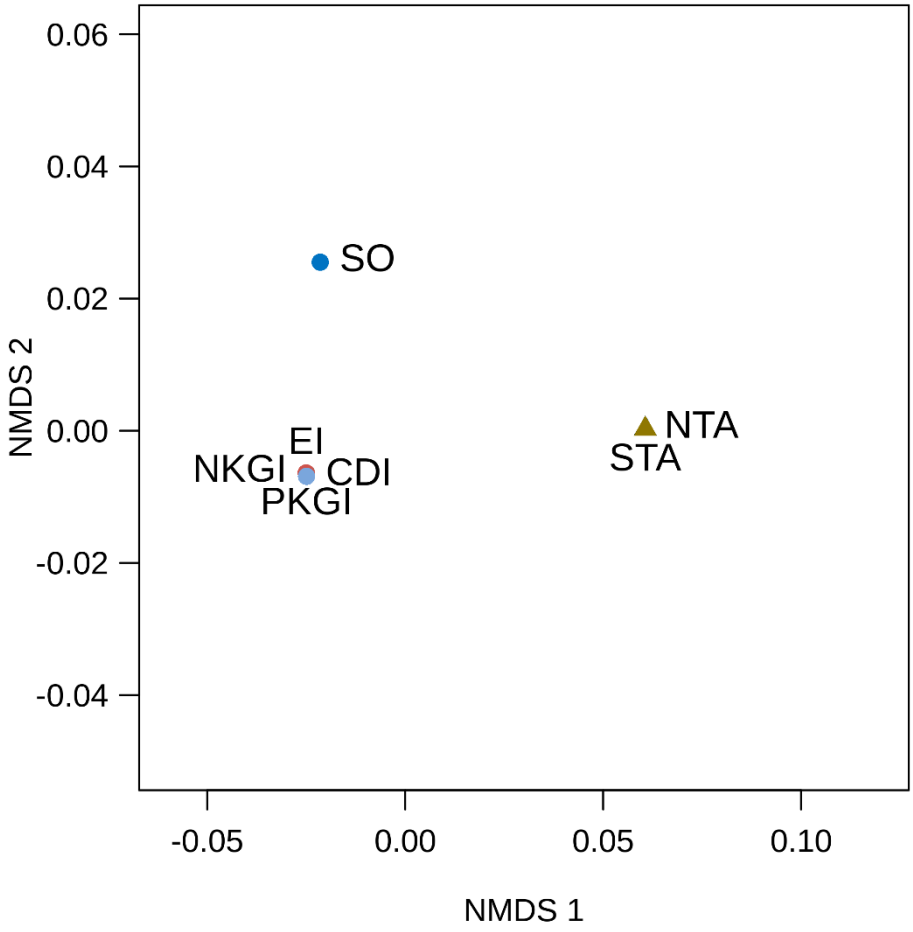


**Fig. S4.5.** Genetic differentiation of *Notothenia coriiceps* in the Southern Ocean based on 1572 SNP genotypes derived from mapping against the reference genome of Shin et al. (2014). East Antarctic samples (NTA, STA) cluster apart from West Antarctic samples after principal component (PC; above) and STRUCTURE analysis (below), with both two and three modeled genetic clusters (K). Sample codes as in Fig. 4.1 and Table 4.1.

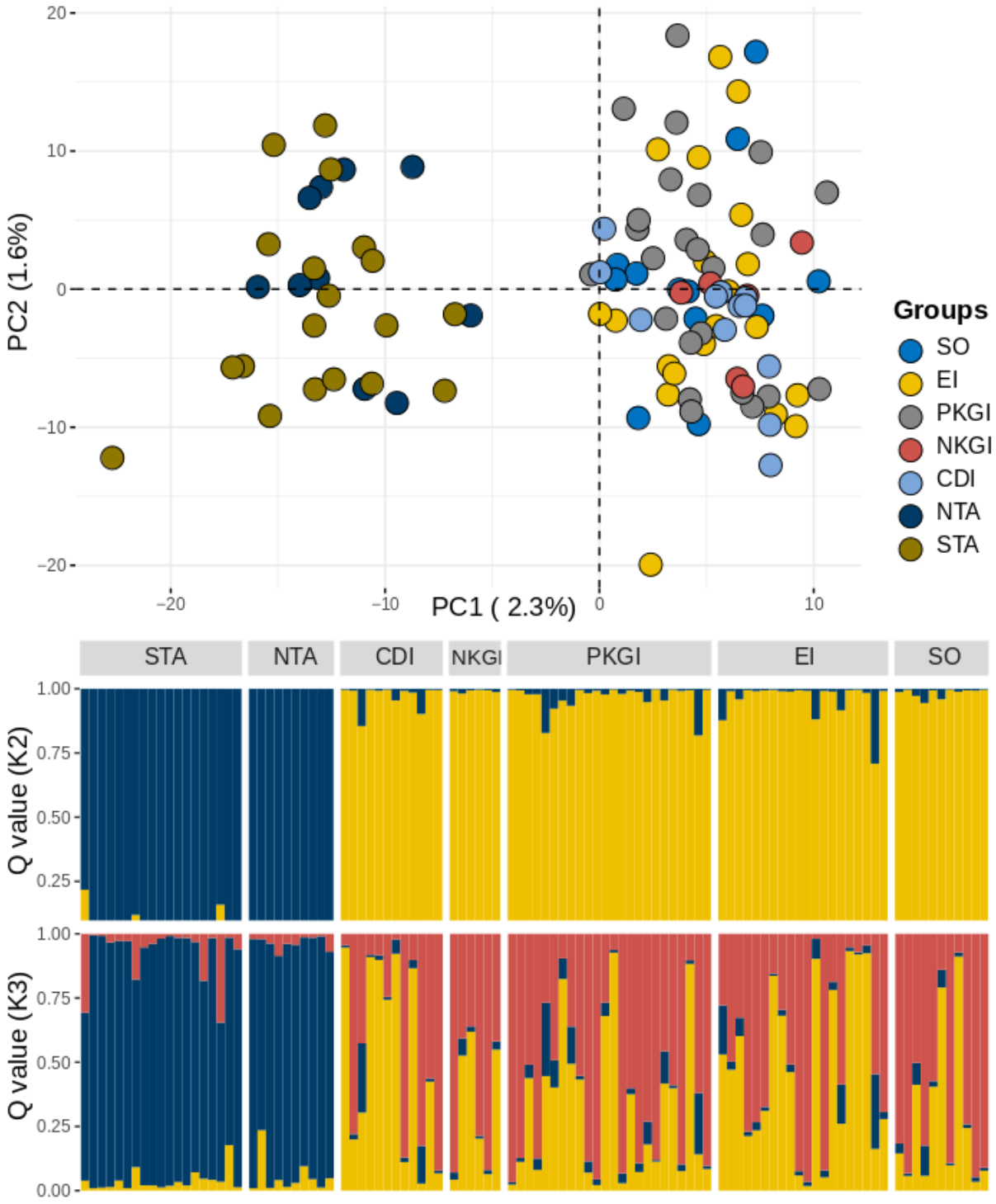


**Table S4.6.** Pairwise genetic differentiation of *Notothenia coriiceps* based on 1948 SNP genotypes derived from mapping against a de novo assembly. Jost's D (2008) below the diagonal and Hedrick's $G_{ST}$ (2005)above the diagonal, as calculated with R package mmod (Winter 2012).

| | **SO** | **EI** | **PKGI** | **NKGI** | **CDI** | **NTA** | **STA** |
|---|---|---|---|---|---|---|---|
| **SO** | | 0.0162 | 0.0125 | 0.0237 | 0.0170 | 0.0374 | 0.0356 |
| **EI** | 0.0030 | | 0.0074 | 0.0113 | 0.0069 | 0.0309 | 0.0246 |
| **PKGI** | 0.0023 | 0.0014 | | 0.0079 | 0.0068 | 0.0261 | 0.0254 |
| **NKGI** | 0.0045 | 0.0022 | 0.0015 | | 0.0119 | 0.0376 | 0.0326 |
| **CDI** | 0.0031 | 0.0013 | 0.0013 | 0.0023 | | 0.0341 | 0.0278 |
| **NTA** | 0.0071 | 0.0060 | 0.0050 | 0.0073 | 0.0066 | | 0.0046 |
| **STA** | 0.0068 | 0.0048 | 0.0049 | 0.0064 | 0.0054 | 0.0009 | |

**Table S4.7.** Pairwise genetic differentiation of *Notothenia coriiceps* based on 1572 SNP genotypes derived from mapping against the reference genome of Shin et al. (2014). Jost's D (2008) below the diagonal and Hedrick's $G_{ST}$ (2005) above the diagonal, as calculated with R package mmod (Winter 2012).

| | **SO** | **EI** | **PKGI** | **NKGI** | **CDI** | **NTA** | **STA** |
|---|---|---|---|---|---|---|---|
| **SO** | | 0.0153 | 0.0093 | 0.0164 | 0.0213 | 0.0328 | 0.0362 |
| **EI** | 0.0029 | | 0.0066 | 0.0087 | 0.0086 | 0.0259 | 0.0273 |
| **PKGI** | 0.0018 | 0.0013 | | 0.0044 | 0.0087 | 0.0225 | 0.0255 |
| **NKGI** | 0.0032 | 0.0017 | 0.0009 | | 0.0129 | 0.0315 | 0.0324 |
| **CDI** | 0.0041 | 0.0017 | 0.0017 | 0.0025 | | 0.0329 | 0.0304 |
| **NTA** | 0.0064 | 0.0051 | 0.0044 | 0.0063 | 0.0064 | | 0.0043 |
| **STA** | 0.0071 | 0.0054 | 0.0051 | 0.0065 | 0.0060 | 0.0008 | |

## IV. S5. Supplementary Material Chapter 5

**Supplementary Material S5.1**. Results from parameter optimization with Stacks v2.4 (Rochette et al. 2019) for *de novo* assembly and genotyping of four genotyping-by-sequencing (GBS) libraries of *Notothenia rossii*.

A parameter test series using a subset of 24 individuals (4 from each population) was conducted as described in Rochette & Catchen (2017). The Stacks parameter m was kept constant in two test series (m = 2 and m = 3), while parameters M and n were varied together from 1 to 9. Subsequently, only loci present in 80 % of the samples were retained and for each M=n parameter the number of loci and polymorphic loci was plotted, as well as the proportion of these loci containing 0 to 10 or >10 SNPs. Optimal values were inferred from these results as m=3 and M=n=4.

### References

Rochette, N.C. and Catchen, J.M. (2017) Deriving genotypes from RAD-seq short-read data using Stacks. Nat. Protoc. 12, 2640–2659 https://doi.org/10.1038/nprot.2017.123

Rochette, N.C., Rivera-Colón, A.G. and Catchen, J.M. (2019) STACKS 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. bioRxiv http://dx.doi.org/10.1101/615385

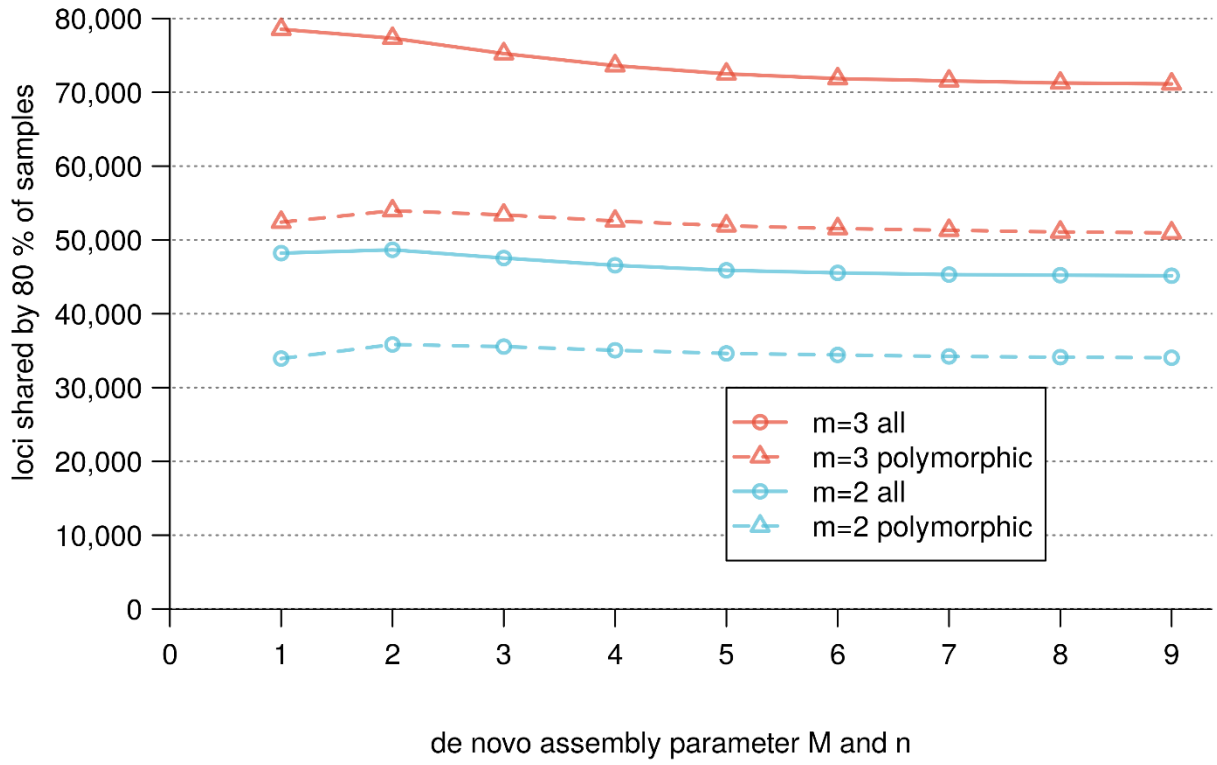



**Fig. S5.1.1.** Number of loci and polymorphic loci shared by 80 % of samples from the *N. rossii* GBS libraries across nine values for parameter M and n and two values for parameter m (2 and 3; blue vs. red) in Stacks v.2.4.

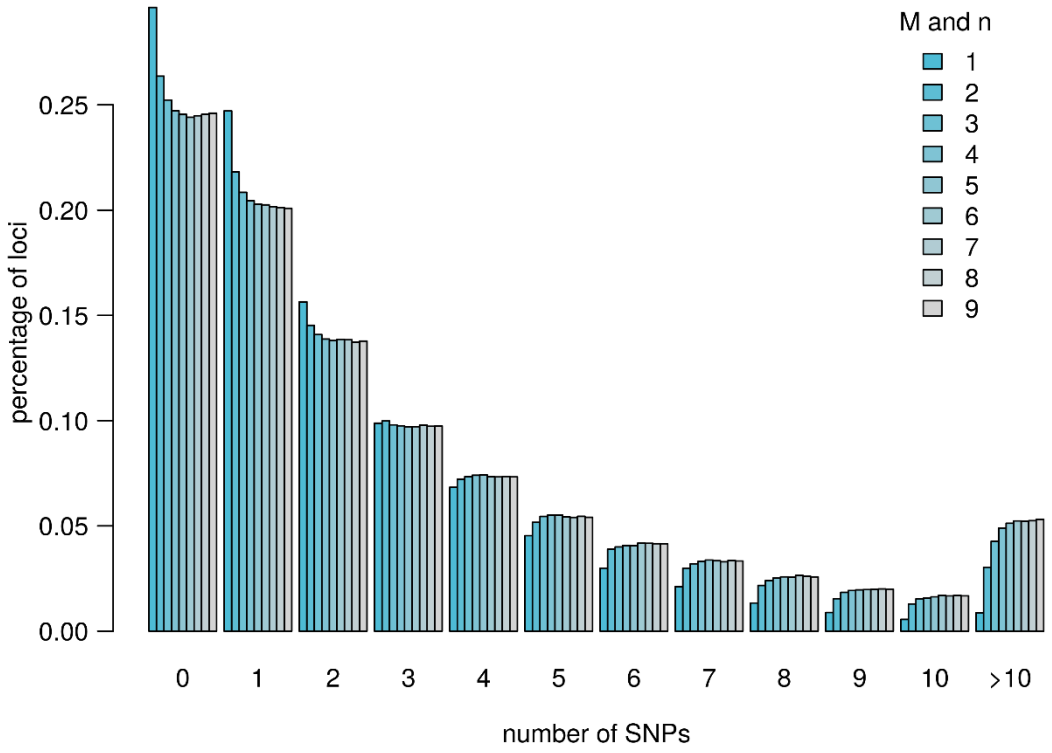



**Fig. S5.1.2.** Number of SNPs per locus shared by 80 % of samples from the *N. rossii* GBS libraries across nine parameters of M and n under constant m = 2 in Stacks v.2.4.

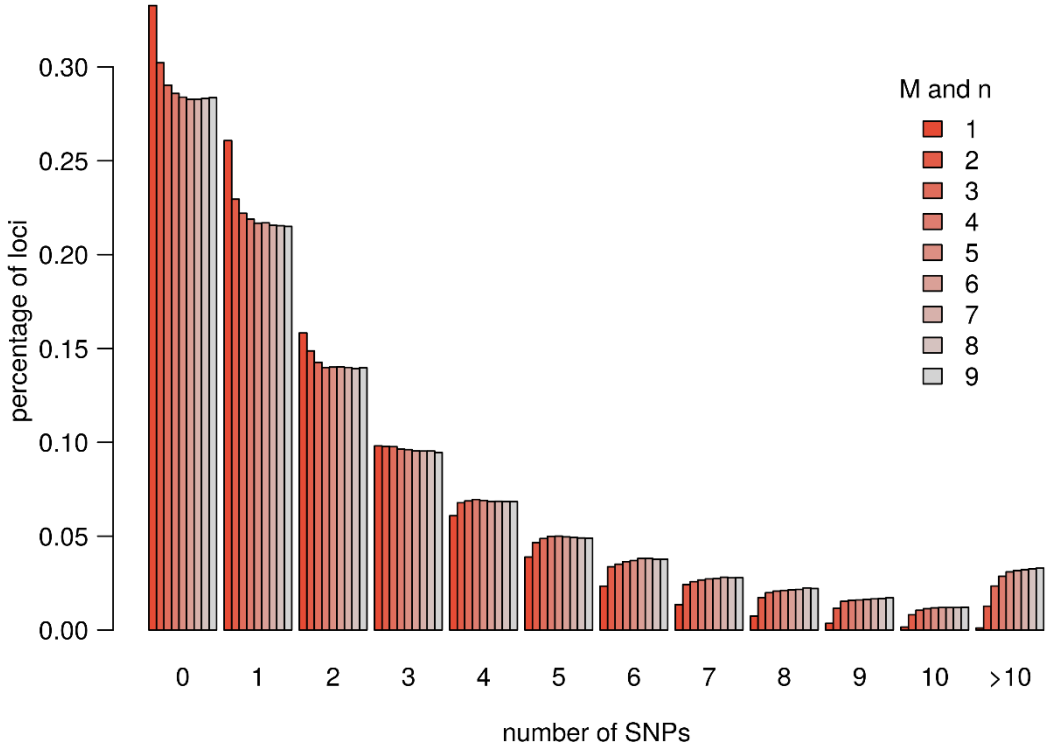



**Fig. S5.1.3.** Number of SNPs per locus shared by 80 % of samples from the *N. rossii* GBS libraries across nine parameters of M and n under constant m = 3 in Stacks v.2.4.

**Supplementary Material S5.2**. Filtering conducted on raw SNP data sets produced from bioinformatics of four GBS libraries of *Notothenia rossii* from the Southern Ocean.

Genotypes of the reference-based and de novo data sets were first pruned using the populations module of Stacks v2.4 (Rochette et al. 2019)., requiring loci to be present in at least 80 % of the individuals of each population, to have a minor allele frequency > 0.05 and heterozygosity < 0.7 (Rochette & Catchen, 2017). Subsequently, genepop files were imported and filtered extensively in R (R Core Team, 2019) using the 'radiator' package v1.1.1 (Gosselin, 2019).

At the start of the radiator filtering pipeline, the data sets contained genotypes from 349individuals at 85,980 (*de novo*) and 73,554 (reference-based) SNPs. First, loci that were not shared across all populations were removed, i.e. 65,269 and 54,850 SNPs. Subsequently, individual genotypes were filtered based on an outlier statistic of missing data and of heterozygosity. These steps removed data from 62 (*de novo*) and 45 (reference) individuals. Then, markers were filtered based on minor allele count (mac), requiring a minimum mac of 10, and on coverage, removing all loci with coverage below 10 or above 100. SNPs were also removed when showing signs of unnormal positioning with the RAD fragment and on short linkage disequilibrium. For the latter, only one SNP per fragment was retained, the one with highest mac. Finally, duplicate genomes were detected and removed and loci significantly ($p < 0.01$) departing from Hardy-Weinberg proportions were removing. After these steps, 277 individuals and 12,400 loci remained in the *de novo* data set and 294 individuals and 4,505 loci in the reference data set. As a last filtering steps, these data were filtered on minor allele frequency (maf, threshold: 0.05), leaving 9,806 and 4,079 SNPs.

Because a bias related to sequencing library was still detectable using principal component analysis in the above data sets, further loci, contributing to this bias, were removed. The final data sets then contain 261 individual genotypes at 7,501 SNPs in the *de novo* data and 272 individual genotypes at 3,503 SNPs in the reference-based data.

**References**

Gosselin T (2019) Radiator: RADseq Data Exploration, Manipulation and Visualization using R. doi: 10.5281/zenodo.1475182. https://thierrygosselin.github.io/radiator/

Rochette, N.C. and Catchen, J.M. (2017) Deriving genotypes from RAD-seq short-read data using Stacks. Nat. Protoc. 12, 2640–2659 https://doi.org/10.1038/nprot.2017.123

Rochette, N.C., Rivera-Colón, A.G. and Catchen, J.M. (2019) STACKS 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. bioRxiv http://dx.doi.org/10.1101/615385

R Core Team (2019) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, v3.5.3. doi: 10.1007/978-3-540-74686-7 http://www.r-project.org/

**Supplementary Material S5.3**. Pairwise genetic differentiation of *Notothenia rossii* in the Southern Ocean using alternative differentiation metrics.

**Table 5.3.1.** Pairwise genetic differentiation of *Notothenia rossii* per sampling locality (see Table 5.1 for codes) based on 7,501 SNP genotypes derived from mapping against a *de novo* assembly. Jost's D (2008) below the diagonal and Hedrick's $G_{ST}$ (2005) above the diagonal, as calculated with R package mmod (Winter 2012).

| | **SSD-06** | **SSK-06** | **SSK-15-16** | **EI-06-07** | **EI-02** | **SO-06** | **SG-02-03** | **SG-05** | **SB-15** | **KI-15** |
|---|---|---|---|---|---|---|---|---|---|---|
| **SSD-06** | | 0.0008 | 0.0015 | 0.0010 | 0.0002 | 0.0030 | 0.0013 | 0.0013 | 0.0015 | 0.0000 |
| **SSK-06** | 0.0002 | | 0.0015 | 0.0017 | 0.0006 | 0.0024 | 0.0015 | 0.0003 | 0.0010 | 0.0013 |
| **SSK-15-16** | 0.0004 | 0.0004 | | 0.0002 | 0.0004 | 0.0013 | 0.0000 | 0.0009 | 0.0012 | 0.0008 |
| **EI-06-07** | 0.0003 | 0.0004 | 0.0001 | | 0.0004 | 0.0023 | 0.0004 | 0.0015 | 0.0010 | 0.0009 |
| **EI-02** | 0.0000 | 0.0001 | 0.0001 | 0.0001 | | 0.0013 | 0.0006 | 0.0006 | 0.0006 | 0.0012 |
| **SO-06** | 0.0008 | 0.0006 | 0.0003 | 0.0006 | 0.0003 | | 0.0016 | 0.0005 | 0.0018 | 0.0014 |
| **SG-02-03** | 0.0003 | 0.0004 | 0.0000 | 0.0001 | 0.0001 | 0.0004 | | 0.0006 | 0.0018 | 0.0014 |
| **SG-05** | 0.0003 | 0.0001 | 0.0002 | 0.0004 | 0.0002 | 0.0001 | 0.0002 | | 0.0005 | 0.0007 |
| **SB-15** | 0.0004 | 0.0003 | 0.0003 | 0.0003 | 0.0001 | 0.0004 | 0.0004 | 0.0001 | | 0.0006 |
| **KI-15** | 0.0000 | 0.0003 | 0.0002 | 0.0002 | 0.0003 | 0.0003 | 0.0004 | 0.0002 | 0.0002 | |

**Table 5.3.2.** Pairwise genetic differentiation of *Notothenia rossii* per sampling locality (see Table 5.1 for codes) based on 3,503 SNP genotypes derived from mapping against the reference genome of *N. coriiceps* (Shin *et al.* 2014). Jost's D (2008) below the diagonal and Hedrick's $G_{ST}$ (2005) above the diagonal, as calculated with R package mmod (Winter 2012).

| | **SSD-06** | **SSK-06** | **SSK-15-16** | **EI-06-07** | **EI-02** | **SO-06** | **SG-02-03** | **SG-05** | **SB-15** | **KI-15** |
|---|---|---|---|---|---|---|---|---|---|---|
| **SSD-06** | | 0.0020 | 0.0017 | 0.0015 | 0.0006 | 0.0029 | 0.0017 | 0.0012 | 0.0008 | 0.0000 |
| **SSK-06** | 0.0004 | | 0.0036 | 0.0025 | 0.0019 | 0.0048 | 0.0021 | 0.0013 | 0.0010 | 0.0016 |
| **SSK-15-16** | 0.0004 | 0.0008 | | 0.0031 | 0.0016 | 0.0018 | 0.0010 | 0.0015 | 0.0026 | 0.0022 |
| **EI-06-07** | 0.0003 | 0.0006 | 0.0007 | | 0.0012 | 0.0035 | 0.0015 | 0.0014 | 0.0009 | 0.0009 |
| **EI-02** | 0.0001 | 0.0004 | 0.0004 | 0.0003 | | 0.0031 | 0.0006 | 0.0003 | 0.0012 | 0.0019 |
| **SO-06** | 0.0006 | 0.0011 | 0.0004 | 0.0008 | 0.0007 | | 0.0025 | 0.0024 | 0.0033 | 0.0022 |
| **SG-02-03** | 0.0004 | 0.0005 | 0.0002 | 0.0003 | 0.0001 | 0.0005 | | 0.0005 | 0.0012 | 0.0015 |
| **SG-05** | 0.0003 | 0.0003 | 0.0003 | 0.0003 | 0.0001 | 0.0005 | 0.0001 | | 0.0011 | 0.0012 |
| **SB-15** | 0.0002 | 0.0002 | 0.0006 | 0.0002 | 0.0003 | 0.0007 | 0.0003 | 0.0003 | | 0.0009 |
| **KI-15** | 0.0000 | 0.0004 | 0.0005 | 0.0002 | 0.0004 | 0.0005 | 0.0003 | 0.0003 | 0.0002 | |

**Supplementary Material S5.4**. Annotation of loci identified through genome scans for selection in GBS data of *Notothenia rossii*. Candidate outlier SNPs were matched against the NCBI database.

**Table S5.4.1.** BLAST results from 12 candidate SNPs from the de novo data set.

| Nr | Name | Top Hit | Accession Number | Percent Identity | E value |
|---|---|---|---|---|---|
| 1 | >CLocus_130488 | Cottoperca gobio genome assembly, chromosome: 19 | LR131926.1 | 82.353 | 4.72E-23 |
| 2 | >CLocus_138804 | Cottoperca gobio genome assembly, chromosome: 12 | LR131919.1 | 81.818 | 4.73E-04 |
| 3 | >CLocus_221130 | Cottoperca gobio genome assembly, chromosome: 12 | LR597562.1 | 89.474 | 4.42E-17 |
| 4 | >CLocus_223132 | PREDICTED: Notothenia coriiceps ubiquitin specific peptidase 38 (usp38), mRNA | XM_010793758.1 | 100 | 3.88E-24 |
| 5 | >CLocus_237675 | Sparus aurata genome assembly, chromosome: 6 | LR537126.1 | 91.892 | 2.00E-03 |
| 6 | >CLocus_240253 | Gossypioides kirkii chromosome KI_01 | CP032244.1 | 90.909 | 2.40E-01 |
| 7 | >CLocus_251284 | Myripristis murdjan genome assembly, chromosome: 22 | LR597571.1 | 88.889 | 2.29E-33 |
| 8 | >CLocus_263048 | Centromochlus existimatus isolate S1A2_08 ATPase subunit 8 (ATPase 8) gene, complete cds; and ATPase subunit 6 (ATPase 6) gene, partial cds; mitochondrial | JX910183.1 | 88.889 | 8.50E-01 |
| 9 | >CLocus_38304 | Apteryx australis mantelli genome assembly AptMant0, scaffold scaffold1406 | LK066414.1 | 88.095 | 7.00E-02 |
| 10 | >CLocus_60053 | Cottoperca gobio genome assembly, chromosome: 3 | LR131933.1 | 82.353 | 9.74E-13 |
| 11 | >CLocus_6959 | Lateolabrax maculatus linkage group 21 sequence | CP032596.1 | 84.507 | 4.14E-11 |
| 12 | >CLocus_96107 | Cottoperca gobio genome assembly, chromosome: 12 | LR131919.1 | 87.288 | 1.19E-30 |

**Table S5.4.2.** BLAST results from 37 candidate SNPs from the reference data set.

| Nr | Name | Top Hit | Accession Number | Percent Identity | E value |
|---|---|---|---|---|---|
| 1 | >CLocus_207926 [AZAD01004947.1, 742, +] | Chionodraco hamatus Cu/Zn superoxide dismutase (SOD1) mRNA, partial cds | AY736281.1 | 90.244 | 1.38E-18 |
| 2 | >CLocus_417728 [AZAD01011137.1, 181, +] | Gymnodraco acuticeps zona pellucida protein ZPC5 isoform 1 (ZPC5) mRNA, complete cds | KU522427.1 | 89.655 | 1.68E-55 |
| 3 | >CLocus_664045 [AZAD01019142.1, 1235, -] | Cottoperca gobio genome assembly, chromosome: 10 | LR131917.1 | 83.092 | 3.03E-52 |
| 4 | >CLocus_1150506 [AZAD01034986.1, 5571, -] | PREDICTED: Notothenia coriiceps transcription initiation factor IIB-like (LOC104957872), partial mRNA | XM_010785544.1 | 87.903 | 2.49E-34 |
| 5 | >CLocus_1462038 [AZAD01044248.1, 240, +] | Dissostichus mawsoni haplotype 1 AFGP/TLP gene locus, partial sequence | HQ447059.1 | 95.96 | 3.45E-83 |

| | | | | | |
|---|---|---|---|---|---|
| 6 | >CLocus_1491991 [AZAD01045372.1, 5082, +] | PREDICTED: Notothenia coriiceps transcription initiation factor IIB-like (LOC104957872), partial mRNA | XM_010785544.1 | 85.484 | 5.49E-30 |
| 7 | >CLocus_2032174 [AZAD01062243.1, 5021, +] | PREDICTED: Notothenia coriiceps transcription initiation factor IIB-like (LOC104957872), partial mRNA | XM_010785544.1 | 85.484 | 5.49E-30 |
| 8 | >CLocus_2457922 [AZAD01071921.1, 3456, +] | Cottoperca gobio genome assembly, chromosome: 10 | LR131917.1 | 73.973 | 6.69E-10 |
| 9 | >CLocus_51011 [KL662357.1, 5270500, -] | Cottoperca gobio genome assembly, chromosome: 21 | LR131929.1 | 76.136 | 6.26E-23 |
| 10 | >CLocus_139511 [KL662384.1, 326949, -] | Cottoperca gobio genome assembly, chromosome: 10 | LR131917.1 | 74.658 | 1.29E-12 |
| 11 | >CLocus_198071 [KL662552.1, 248239, +] | Cottoperca gobio genome assembly, chromosome: 1 | LR131916.1 | 83.333 | 3.46E-26 |
| 12 | >CLocus_210679 [KL662597.1, 16029, -] | PREDICTED: Notothenia coriiceps uncharacterized LOC104965236 (LOC104965236), mRNA | XM_010794195.1 | 87.5 | 3.95E-19 |
| 13 | >CLocus_210838 [KL662599.1, 5111, +] | Cottoperca gobio genome assembly, chromosome: 14 | LR131921.1 | 83.571 | 2.33E-28 |
| 14 | >CLocus_261220 [KL662789.1, 79206, +] | Dissostichus mawsoni haplotype 2 AFGP/TLP gene locus, partial sequence | HQ447060.1 | 92.405 | 1.38E-56 |
| 15 | >CLocus_292518 [KL662880.1, 318427, +] | Sparus aurata genome assembly, chromosome: 17 | LR537137.1 | 76.829 | 7.63E-22 |
| 16 | >CLocus_313316 [KL662933.1, 117387, -] | Cottoperca gobio genome assembly, chromosome: 9 | LR131939.1 | 76.531 | 6.69E-29 |
| 17 | >CLocus_512277 [KL663578.1, 176442, +] | Cottoperca gobio genome assembly, chromosome: 6 | LR131936.1 | 85.87 | 5.48E-49 |
| 18 | >CLocus_550684 [KL663710.1, 313820, +] | PREDICTED: Notothenia coriiceps symplekin-like (LOC104960710), mRNA | XM_010788863.1 | 82.707 | 8.14E-28 |
| 19 | >CLocus_612378 [KL663896.1, 102980, +] | Cottoperca gobio genome assembly, chromosome: 10 | LR131917.1 | 77.397 | 5.86E-17 |
| 20 | >CLocus_676688 [KL664078.1, 28095, -] | Lateolabrax maculatus chromosome Lm22 | CP027283.1 | 76.301 | 6.69E-29 |
| 21 | >CLocus_984478 [KL665099.1, 640019, +] | Thalassophryne amazonica genome assembly, chromosome: 13 | LR722978.1 | 80 | 8.00E-03 |
| 22 | >CLocus_1083361 [KL665382.1, 29521, +] | Cicer arietinum chromosome Ca2 | CP039332.1 | 85.714 | 3.20E-01 |
| 23 | >CLocus_1090835 [KL665412.1, 447335, +] | Cottoperca gobio genome assembly, chromosome: 1 | LR131916.1 | 84.277 | 9.28E-40 |
| 24 | >CLocus_1092384 [KL665412.1, 1044395, +] | PREDICTED: Aplysia californica calmodulin-like protein 3 (LOC106012422), mRNA | XM_013085361.1 | 82.222 | 4.00E+00 |

| | | | | | |
|---|---|---|---|---|---|
| 25 | >CLocus_1179376 [KL665586.1, 242786, -] | Gouania willdenowi genome assembly, chromosome: 8 | LR131991.1 | 75 | 1.21E-06 |
| 26 | >CLocus_1224872 [KL665708.1, 68516, -] | Lateolabrax maculatus linkage group 21 sequence | CP032596.1 | 72.358 | 9.93E-08 |
| 27 | >CLocus_1425216 [KL666295.1, 8396428, +] | Cottoperca gobio genome assembly, chromosome: 4 | LR131934.1 | 89.103 | 5.48E-49 |
| 28 | >CLocus_1534498 [KL666587.1, 97216, +] | Sparus aurata genome assembly, chromosome: 9 | LR537129.1 | 89.899 | 6.67E-67 |
| 29 | >CLocus_1551285 [KL666590.1, 947841, +] | Chionodraco hamatus transposon helitron polyprotein (HeliNoto) gene, complete cds | GU014476.2 | 77.236 | 7.14E-16 |
| 30 | >CLocus_1552830 [KL666590.1, 1381214, -] | PREDICTED: Notothenia coriiceps symplekin-like (LOC104960710), mRNA | XM_010788863.1 | 86.667 | 2.49E-34 |
| 31 | >CLocus_1630310 [KL666849.1, 6404, +] | Cottoperca gobio genome assembly, chromosome: 5 | LR131935.1 | 87.879 | 1.57E-11 |
| 32 | >CLocus_1954638 [KL667808.1, 350977, -] | PREDICTED: Notothenia coriiceps ectonucleotide pyrophosphatase/phosphodiesterase family member 2-like (LOC104962317), mRNA | XM_010790748.1 | 99.167 | 3.69E-51 |
| 33 | >CLocus_2033344 [KL668045.1, 3250, +] | PREDICTED: Notothenia coriiceps uncharacterized LOC104955179 (LOC104955179), ncRNA | XR_799431.1 | 90.625 | 2.84E-27 |
| 34 | >CLocus_2067957 [KL668140.1, 732310, +] | Cottoperca gobio genome assembly, chromosome: 6 | LR131936.1 | 75 | 5.14E-05 |
| 35 | >CLocus_2206060 [KL668296.1, 21625648, +] | PREDICTED: Arabidopsis lyrata subsp. lyrata protein SUPPRESSOR OF NIM1 1 (LOC9320118), mRNA | XM_021032701.1 | 89.474 | 9.30E-02 |
| 36 | >CLocus_2245286 [KL668297.1, 4382515, +] | Dicentrarchus labrax chromosome sequence corresponding to linkage group 1, top part, complete sequence | FQ310506.3 | 78.616 | 1.47E-24 |
| 37 | >CLocus_2284480 [KL668297.1, 14524502, -] | Cottoperca gobio genome assembly, chromosome: 10 | LR131917.1 | 76.712 | 2.49E-15 |

**Supplementary Material S5**. Results of the population genomic analysis of *Notothenia rossii* in the Southern Ocean using the reference genome aligned SNP data set.

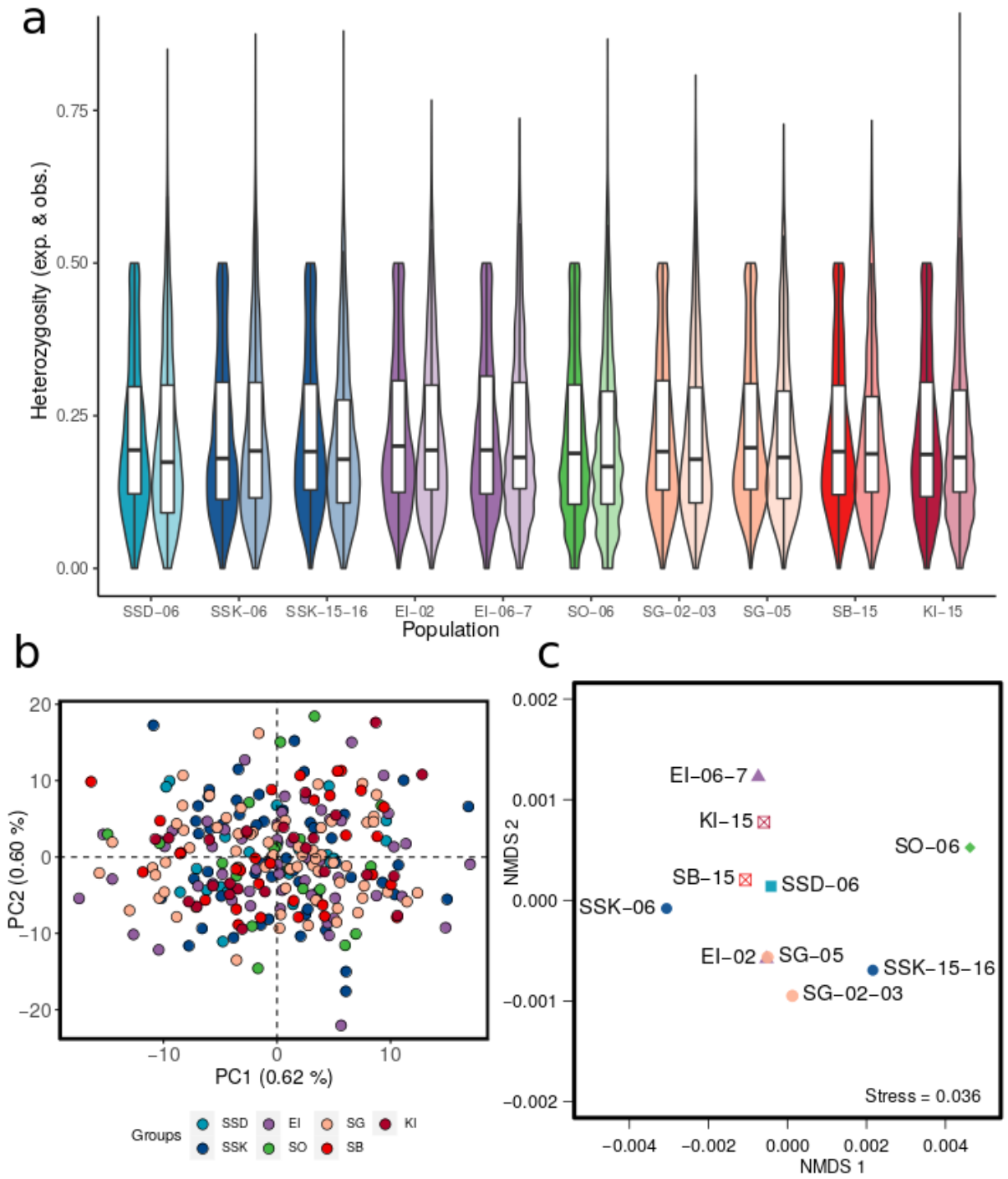



**Fig. S5.5.1.** Genomic diversity of *Notothenia rossii* in the Southern Ocean based on 3,503 SNP loci from reference-based variant calling. Expected (darker shading, left) and observed (light shading, right) heterozygosity is shown as box and violin plots for each genetically screened population (a). Principal component analysis (PCA) reveals little individual-based differentiation (b), while non-metric multidimensional scaling based on $G_{ST}$ distances shows subtle differences between population samples (c). Sample codes as in Table 5.1; samples from different years but same locality are not shown separately on the PCA.

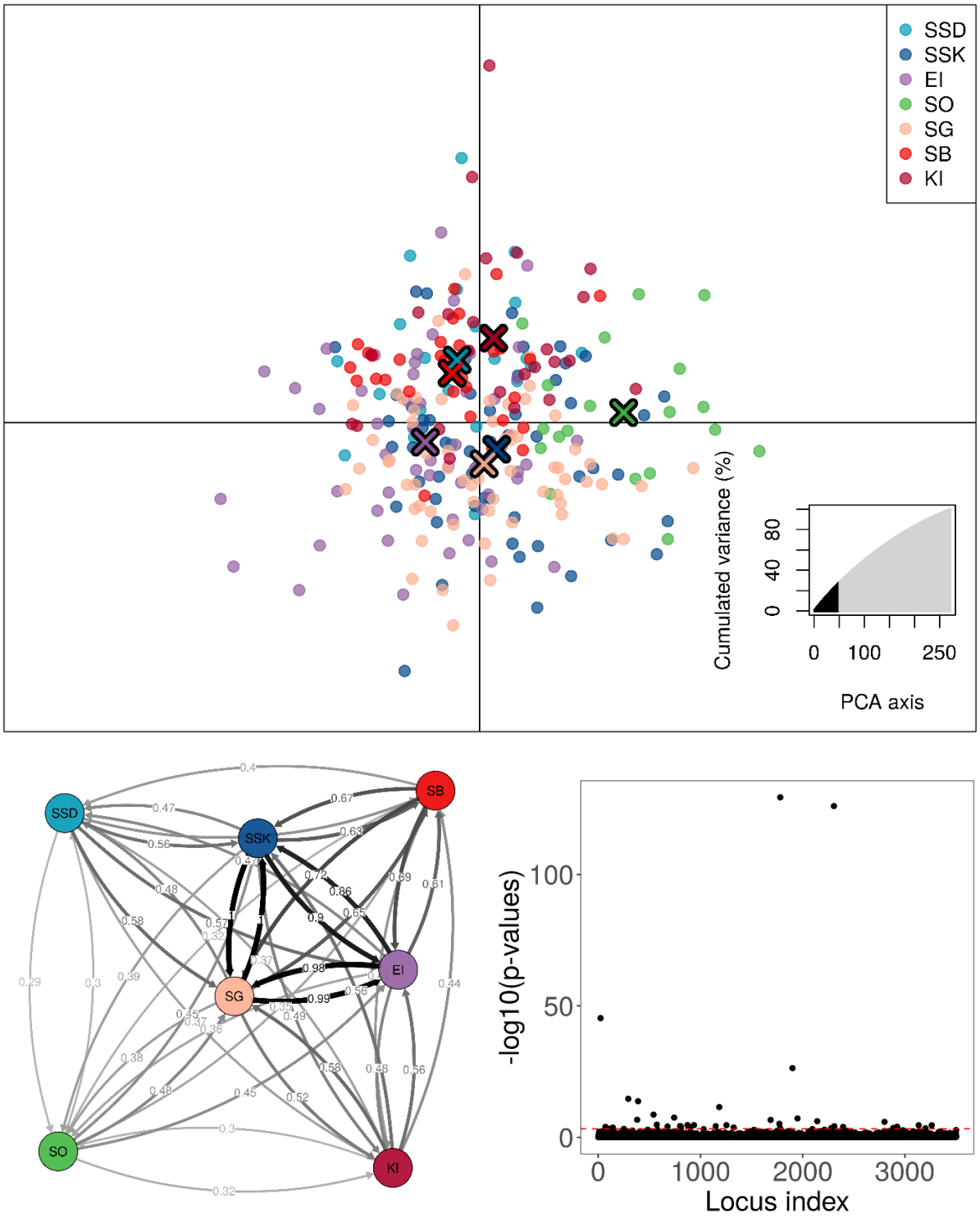



**Fig. S5.5.2.** Genomic differentiation of *Notothenia rossii* in the Southern Ocean based on 3,503 SNP loci from reference-based variant calling. Geographic clustering as attempted through discriminant analysis of principal components is shown along the first two principal components (a). Relative migration as estimated from Nei's $G_{ST}$ reveals overall high and no asymmetric gene flow (b). Genome scans for loci putatively under influence of selection detected 37 outliers at q > 0.05 (c). Sample codes as in Table 5.1; samples from different years but same locality are combined.