Universität Bremen

A molecular view on adaptation on local and continental scales in the Sub-Antarctic and Antarctic bivalve *Aequiyoldia*.

Dissertation —— 2021

Mariano Martínez

A molecular view on adaptation on local and continental scales in the Sub-Antarctic and Antarctic bivalve *Aequiyoldia*

Mariano Martínez

Dissertation

in fulfilment of the requirements for the degree of
Doctor of natural sciences (Dr. rer. nat.)

at the Faculty 02 – Biology and Chemistry
of the University of Bremen, Germany

Submitted on August 2nd 2021 and reviewed by
Prof. Dr. Martin Zimmer
Dr. Felix Mark
Colloquium held on 28 September 2021.

Supervised by
Prof. Dr. Doris Abele
Dr. Christoph Held

ALFRED-WEGENER-INSTITUT
HELMHOLTZ-ZENTRUM FÜR POLAR-
UND MEERESFORSCHUNG

Universität Bremen

This cumulative thesis was carried out in the Alfred Wegener Institute between January 2017 and July 2021. It was funded by Agencia Nacional de Investigación e Innovación (ANII, Uruguay), Deutscher Akademischer Austauschdienst (DAAD, Germany), Zempelin-Stiftung - Deutsches Stiftungszentrum and by the Biosciences division of the Alfred Wegener Institute.

Thesis reviewers

First reviewer: Prof. Dr. Martin Zimmer
Mangrove Ecology,
University Bremen, and
Leibniz Centre for Tropical Marine Research
Bremen, Germany

Second reviewer: Dr. Felix Mark
Biosciences – Integrative Ecophysiology
Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research
Bremerhaven, Germany

Examination Commission

Commission chair: Prof. Dr. Christian Wild
Marine Ecology,
University Bremen
Bremen, Germany

Examiner: Prof. Dr. Martin Zimmer
Mangrove Ecology,
University Bremen, and
Leibniz Centre for Tropical Marine Research
Bremen, Germany

Examiner: Dr. Felix Mark
Biosciences – Integrative Ecophysiology
Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research
Bremerhaven, Germany

Examiner: Dr. Charlotte Havermans
Biosciences – Functional Ecology
Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research
Bremerhaven, Germany

Student member: Msc. Onur Karakus
PhD candidate at Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research
and University Bremen
Bremen, Germany

Student member: Luis Kuchenmüller
Master Student (Marine Biology) at University Bremen
Bremen, Germany

Summary

Marine species and populations have three potential responses to climate change: shift their distribution, adapt to the new environmental conditions or go extinct. The persistence of species unable to shift their ranges in response to changing conditions will be determined by their standing phenotypic plasticity or their ability to develop evolutionary adaptive responses. Physiological comparisons of closely related species/populations on latitudinal gradients have proven to be very informative in determining their respective phenotypic plasticity and genetic adaptability. These macro-scale perspectives, however, overlook the role of small-scale environmental variation in the inter-individual physiological and genetic differences. In this thesis, I used the Southern Ocean protobranch bivalve *Aequiyoldia cf. eightsii* (Jay, 1839) from West Antarctic Peninsula (WAP) and southern South America (SSA) as a "model species" to study the genetic and phenotypic traits that support adaptation to current and future environmental change at small (i.e., local or population scale) and large -scale (i.e., continental or species scale).

As recent evidence suggests the possibility of cryptic speciation between *Aequiyoldia* bivalves from WAP and SSA, **Chapter 2** aims at analysing the genetic diversity between and within populations on both sides of the Drake Passage. In this Chapter I report several highly differentiated mitochondrial genomes (h1, h2, h3, h4) within *A. cf. eightsii* coexisting in Antarctic populations but also inside a subset of the individuals sampled (mitochondrial heteroplasmy). The mitochondrial differentiation pattern is mirrored in nuclear Single Nucleotide Polymorphisms (SNPs) only across the Drake Passage, whilst the equally strongly differentiated mitochondrial lineages in the Southern Ocean are part of the same distribution of SNPs. These results suggest that populations on both sides of the Drake are two reproductively isolated species, and refuted the previous suggestions of cryptic speciation in WAP *A. cf. eightsii*. Using SNPs from the entire nuclear and mitochondrial genomes for reference, I demonstrated that mitochondrial heteroplasmy unpredictably misleads classical molecular barcoding procedures using universal *cytochrome c oxidase subunit I* (*COI*) primers, producing wrong taxonomic inferences with high confidence.

The small-scale approach (**Chapter 3**) involved the study of in situ gene expression patterns within an *Aequiyoldia* population in front of a melting glacier in the region of the WAP. This population exhibited strikingly different gene expression pattern under subtly different natural conditions. This pattern was influenced by at least three independent underlying causes: small scale habitat heterogeneity (down to a kilometre scale), and the composition of the mitochondrial and nuclear genomes. Interestingly, the expression of nuclear genes correlated strongly with the mitochondrial genotype, with the highest gene expression differences between homoplasmic and heteroplasmic organisms. This novel mechanism might serve to add another layer of flexibility to respond to the environment and turn out instrumental in the face of the ongoing rapid environmental change in Antarctic fjords.

The large-scale approach (**Chapter 4**) involved an experimental inter-continental comparison of gene expression patterns in response to changes in thermal and oxygen regimes expected under a global warming scenario. In both populations (from WAP and SSA), the experimental temperature implied exposure scenarios simulating a crossing of the Drake Passage, and to an expected near future warming scenario at the WAP. The WAP bivalves showed a moderated physiological response to warming and a remarkable ability to cope with short-term exposure to hypoxia by switching to a metabolic rate depression strategy and activating the alternative oxidation pathway. In SSA, the high prevalence of apoptosis (cell-death)-related differentially expressed genes especially under combined higher temperatures and hypoxia indicated that the SSA *Aequiyoldia* are operating near their physiological limits already. While the effect of temperature *per se* may not represent the single most effective barrier to Antarctic colonization by South American bivalves, the current distribution patterns as well as their resilience to future conditions may be better understood by looking at the synergistic effects of temperature in conjunction with short term exposure to hypoxia.

Overall, this thesis provides a molecular perspective on the adaptive capacity and cross-continental invasibility of two *Aequiyoldia* sibling species inhabiting WAP and SSA under a global change scenario.

Zusammenfassung

Marine Arten und Populationen haben drei mögliche Reaktionen auf den Klimawandel: Sie können ihr Verbreitungsgebiet verlagern, sich an die neuen Umweltbedingungen anpassen oder aussterben. Das Überleben von Arten, die nicht in der Lage sind ihr Verbreitungsgebiet als Reaktion der sich ändernden Bedingungen zu verlagern, wird durch ihre ständige phänotypische Plastizität oder ihre Fähigkeit evolutionär adaptive Reaktionen zu entwickeln, bestimmt. Physiologische Vergleiche von eng verwandten Arten/Populationen auf Breitengradienten haben sich als sehr aufschlussreich erwiesen, um ihre jeweilige phänotypische Plastizität und genetische Anpassungsfähigkeit zu bestimmen. Diese makroskaligen Gegenüberstellungen übersehen jedoch die Rolle der kleinräumigen Umweltvariationen bei den interindividuellen physiologischen und genetischen Unterschieden. In dieser Arbeit habe ich die Protobaummuschel *Aequiyoldia cf. eightsii* (Jay, 1839) von der Westantarktischen Halbinsel (WAP) und dem südlichen Südamerika (SSA) als "Modellart" verwendet, um die genetischen und phänotypischen Merkmale zu untersuchen, die die Anpassung an aktuelle und zukünftige Umweltveränderungen auf kleiner (d.h. lokaler oder Populations-Skala) und großer Skala (d.h. kontinentaler oder Spezies-Skala) unterstützen.

Da neuere Erkenntnisse die Möglichkeit einer kryptischen Artbildung wischen *Aequiyoldia*-Muscheln aus WAP und SSA nahelegen, zielt **Kapitel 2** auf die Analyse der genetischen Diversität zwischen und innerhalb von Populationen auf beiden Seiten der Drake Passage. In diesem Kapitel berichte ich über mehrere hochdifferenzierte mitochondriale Genome (h1, h2, h3, h4) innerhalb der *A. cf. eightsii*, die sowohl in antarktischen Populationen koexistieren, als auch innerhalb einer Teilmenge der beprobten Individuen (mitochondriale Heteroplasmie). Das mitochondriale Differenzierungsmuster spiegelt sich in nuklearen Single Nucleotide Polymorphisms (SNPs) nur auf der anderen Seite der Drake Passage wieder, während die ebenso stark differenzierten mitochondrialen Linien im Südpolarmeer Teil der gleichen SNP-Verteilung sind. Diese Ergebnisse deuten darauf hin, dass es sich bei den Populationen auf beiden Seiten der Drake Passage um zwei reproduktiv isolierte Arten handelt und widerlegen die bisherigen Vermutungen einer kryptischen Artbildung in WAP *A. cf.*

eightsii. Unter Verwendung von SNPs aus dem gesamten Kern- und Mitochondriengenom als Referenz habe ich gezeigt, dass mitochondriale Heteroplasmie klassische molekulare Barcoding-Verfahren mit universellen Cytochrom-c-Oxidase-Untereinheit I (COI)-Primern in unvorhersehbarer Weise in die Irre führt und mit hoher Sicherheit falsche taxonomische Rückschlüsse zulässt.

Der kleinräumige Ansatz (**Kapitel III**) beinhaltete die Untersuchung von in situ Genexpressionsmustern innerhalb einer *Aequiyoldia*-Population vor einem schmelzenden Gletscher in der Region des WAP. Diese Population zeigte ein auffallend unterschiedliches Genexpressionsmuster unter subtil unterschiedlichen natürlichen Bedingungen. Dieses Muster wurde von mindestens drei unabhängigen, zugrundeliegenden Ursachen beeinflusst: kleinräumige Habitat-Heterogenität (bis zu einer Kilometerskala), die Zusammensetzung des mitochondrialen Genoms und die Zusammensetzung des Kern Genoms. Interessanterweise korrelierte die Expression von Kerngenen stark mit dem mitochondrialen Genotyp, wobei die größten Unterschiede in der Genexpression zwischen homoplasmatischen und heteroplasmatischen Organismen auftraten. Dieser neuartige Mechanismus könnte dazu dienen, eine weitere Ebene der Flexibilität zu schaffen, um auf die Umwelt zu reagieren und sich angesichts der anhaltenden schnellen Umweltveränderungen in den antarktischen Fjorden als hilfreich zu erweisen.

Der groß angelegte Ansatz (**Kapitel IV**) beinhaltete einen experimentellen interkontinentalen Vergleich von Genexpressionsmustern als Reaktion auf Veränderungen im thermischen- und Sauerstoffregime, die unter einem globalen Erwärmungsszenario erwartet werden. In beiden Populationen (aus WAP und SSA) implizierte die experimentelle Temperatur Expositionsszenarien, die eine Durchquerung der Drake Passage simulierten, ein erwartetes zukünftiges Erwärmungsszenario am WAP. Die WAP-Muscheln zeigten eine gemäßigte physiologische Reaktion auf die Erwärmung und eine bemerkenswerte Fähigkeit mit einer kurzfristigen Exposition auf Hypoxie zu reagieren indem sie strategisch auf die Senkung der Stoffwechselrate umschalteten und den alternativen Oxidationsweg aktivierten. In SSA deutet die hohe Prävalenz von Apoptose (Zelltod)-bezogenen differentiell exprimierten Genen

besonders unter kombinierten höheren Temperaturen und Hypoxie darauf hin, dass die SSA *Aequiyoldia* bereits nahe ihrer physiologischen Grenzen operieren. Während der Temperatureffekt an sich vielleicht nicht die effektivste Barriere für die Besiedlung der Antarktis durch südamerikanische Muscheln darstellt, können die aktuellen Verbreitungsmuster sowie ihre Widerstandsfähigkeit gegenüber zukünftigen Bedingungen besser verstanden werden, wenn man die synergistischen Effekte der Temperatur in Verbindung mit kurzzeitiger Hypoxie Exposition betrachtet.

Insgesamt zeigt diese Arbeit eine molekulare Perspektive auf die Anpassungsfähigkeit und die Kontinent-übergreifende Invasivität von zwei *Aequiyoldia*-Geschwisterarten, die auf WAP und SSA verbreitet und einem globalen Veränderungszenario ausgesetzt sind.

Table of Contents

Summary i

Zusammenfassung iii

1 General introduction and objectives 1

1.1 From cross-Drake to within-Antarctic fjord spatial scales 5

1.2 Responses of biota to global change: move, adapt or go extinct 7

1.2.1 Move: poleward expansion of lower-latitude species 8

1.2.2 Adaptation to environmental change: ecological niche exploitation 9

1.3 *Aequiyoldia eightsii* as a model species 11

1.3.1 Justification 11

1.3.2 General characteristics 12

1.3.3 Distribution and population genetics 13

1.3.4 Responses to environmental stress 14

1.4 Objectives 15

1.4.1 Specific objectives of chapters 15

2 Publication I 25

3 Publication II 59

4 Publication III 99

5 Synoptic discussion 135

5.1 Gene expression patterns: product of adaptive and neutral evolution 137

5.1.1 The molecular phenotype in phylogenetic inferences 139

5.1.2 Climate adaptation hallmarks of the molecular phenotype 141

5.2 The fate of *A. cf. eightsii* in a warming Southern Ocean 142

5.3 Conclusions and perspectives 148

Declaration on the contribution of the candidate to a multi-author article 159

Versicherung an Eides Statt 160

Acknowledgments 161

CHAPTER 1

1 General introduction and objectives

Understanding the capacity of species to survive and thrive under conditions of environmental change is a fundamental requirement if we want to make predictions concerning ecosystem response to climatic change and human interference. When faced with environmental change, marine species/populations have three potential responses: **(i)** shift their distribution, **(ii)** adapt to the new environmental conditions or **(iii)** become locally or globally extinct. Species may persist by shifting their distribution ranges, either on bathymetric or latitudinal scale. The expansion or displacement of the distribution range of a species or population inevitably implies the invasion of a new habitat, colonized by a native community, which may be affected by the arrival of non-indigenous species (NIS). Poorly dispersing organisms, habitat specialists or those without more suitable regions to migrate may be unable to shift their ranges in response to changing conditions, and instead their persistence will be determined by their standing phenotypic plasticity or their ability to develop adaptive responses. While an adaptive evolutionary response occurs on long timescales and involves a change in allele frequencies (including changes at genotypic level), organisms can develop adaptive responses which only involve changes at the phenotypic level; a phenomenon called phenotypic plasticity (Stearns 2015). Environmental changes can also trigger epigenetic modifications, i.e., heritable changes of gene expression that are not dependent on changes to DNA sequence (or allele frequency). One of the best studied epigenetic mechanisms is DNA methylation, which provides a signal involved in the regulation of gene expression, connected with development, disease, and environmentally induced plasticity (Eirin-Lopez and Putnam 2019).

Thus, the study of species and population genetic diversity and the identification of environmental factors that govern their biogeographic distribution patterns and elicit phenotypic responses (morphological and physiological plasticity) to current and future environmental change, is of fundamental importance for predicting species or population resilience. Such analyses allow us to predict how local communities will respond to scenarios of expected environmental change. This, in turn, matters for assessment of ecosystem health and development of conservation measures, and the management of ecosystem services.

Physiological comparisons of different populations or closely related species on latitudinal gradients exposed to different regimes of variability have proven to be very informative in determining their respective phenotypic plasticity and genetic adaptability (Seebacher and Franklin 2012; Koenigstein et al. 2013). These macro-scale perspectives, however, do not address the importance of within-population variation for selection to act upon, overlooking the role of small-scale environmental variation in the inter-individual physiological and genetic differences. The small-scale environmental heterogeneity may interact with biological sources of variation within populations (e.g., inter-individual genetic diversity) promoting non-uniform responses of organisms, a fact that may be essential for the adaptation of populations to environmental shifts (Tanner and Dowd 2019).

In this thesis, I used the Southern Ocean protobranch bivalve *Aequiyoldia cf. eightsii* (Jay, 1839) from West Antarctic Peninsula (WAP) and Southern South America (SSA) as a “model species” to study the genetic and phenotypic traits that support adaptation to current and future environmental change at small (i.e., local or population scale) and large -scale (i.e., continental or species scale). Although the populations of the WAP and SSA are currently considered the same species, recent evidence suggests the possibility of cryptic speciation between the two groups (González-Wevar et al. 2019), with the Antarctic Circumpolar Current functioning as biogeographic barrier of reproductive/genetic connectivity (Muñoz-Ramírez et al. 2020). The possibility of cryptic speciation biases interpretations and significance of species’ adaptive capacity. In the following I will refer to both *Aequiyoldia* populations as *Aequiyoldia cf. eightsii*, and

based on the results of Chapter 2, they will henceforth be considered as separated species.

The **Chapter 2** of this thesis aims at analyzing the genetic diversity between and within populations on both sides of the Drake to determine the possibility of cryptic speciation between both continents. In addition, this chapter calls into question the reliability of a taxonomic identification system based on a few nuclear or mitochondrial markers such as the widely used *cytochrome c oxidase subunit I* (*COI*). It presents novel results regarding intra-population genetic variation at the WAP, such as the presence of two mitochondrial genomes, a phenomenon observed in some bivalves species (Gusman et al. 2016) but never shown in an Antarctic species (heteroplasmy). The condition of mitochondrial heteroplasmy and its implication for local habitat exploitation and adaptation in a changing world is addressed in the following chapters. The small-scale approach involved the study of gene expression patterns within a population in response to habitat heterogeneity in front of a melting glacier in the region of the WAP (**Chapter 3**) and to intra-population genetic diversity observed in Chapter 2. The large-scale approach involved an experimental inter-continental comparison of gene expression patterns in response to changes in thermal and oxygen regimes expected under a global warming scenario (**Chapter 4**). As the expected temperatures in an upcoming warming scenario in the WAP region overlap with the current temperature range in SSA, this experimental approach focuses on the study of physiological traits that may allow for/restrict distributional range shifts involving cross-continental invasions of organisms from either side of the Drake Passage (Fig. 1.1)

This introductory section (**Chapter 1**) will provide the general theoretical framework for the understanding and linkage of the three chapters of this thesis. It starts with a brief description of the most relevant features at the two different spatial scales (trans-Drake and within-Antarctic fjord – Potter Cove) and the effects of climate change at high latitudes (1.1). Next, it addresses the two alternative responses of populations/species facing environmental changes - to change distribution ranges and to develop novel adaptational capacities (1.2) - with emphasis on the colonization of the Antarctic continent by non-indigenous species from lower latitudes (1.2.1) and on the different

spatial and temporal scales at which adaptive responses can be approached (1.2.2). Last, it presents the study species *Aequiyoldia cf. eightsii* providing information about its biology, distribution, population genetics and responses to thermal stress (1.3). In section 1.4 the specific objectives for each chapter are provided.

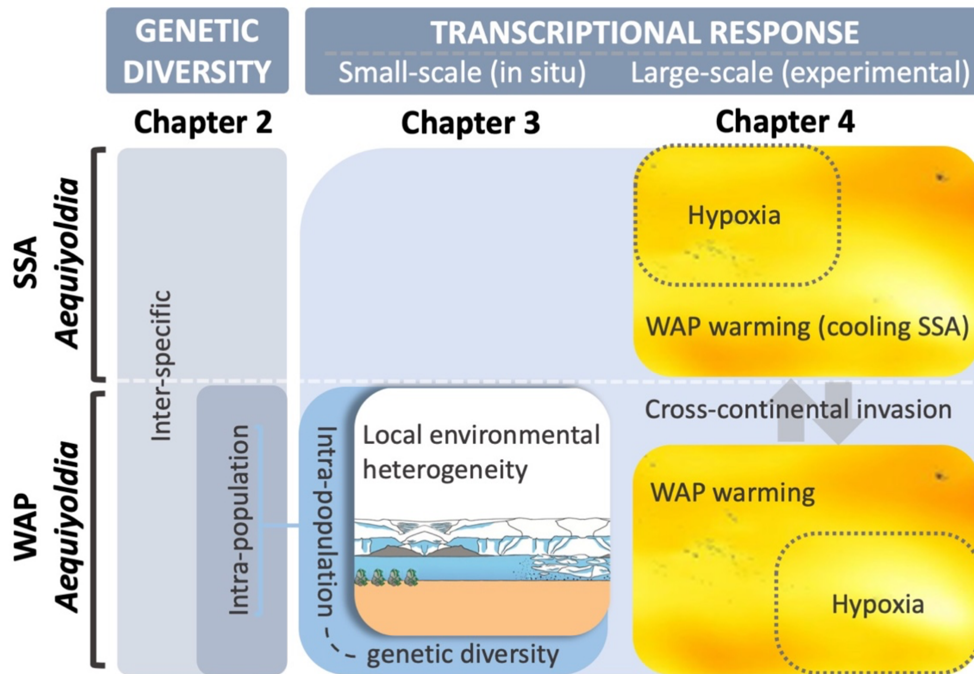


Figure 1.1 Outline of the approaches in Chapters 2, 3 and 4. Chapter 2 is focus on evaluating genetic diversity between and within populations on both sides of the Drake to determine the possibility of cryptic speciation within *Aequiyoldia cf. eightsii*. Through a small-scale approach within a population in front of a melting glacier in the region of the WAP, Chapter 3 analyses in situ transcriptional responses to local environmental heterogeneity and to intra-population genetic diversity. Chapter 4 involves an experimental inter-continental comparison of gene expression patterns in response to two major stressors related to climate change: temperature shifts and hypoxia. In both populations (from WAP and SSA), the experimental temperature conditions simulate a future scenario of warming in WAP (by warming organisms from WAP, and cooling organisms from SSA) to evaluate the effects of warming in WAP *Aequiyoldia* and at the same time the possibility of cross-continental invasions in a global warming scenario.

1.1 From cross-Drake to within-Antarctic fjord spatial scales

Southern South America (SSA, here referring to the Magellan region and Tierra del Fuego) and the West Antarctic Peninsula (WAP) constitute two distinct climatic zones, cold-temperate/sub-polar and polar respectively, with similar coastal systems characterised by islands and peninsulas situated within an extended glacial channel and fjordic system. These two regions were contiguous until severed by south-eastward relative movement of the Antarctic Peninsula away from Tierra del Fuego between 49 to 17 million years ago (Scher and Martin 2006). This continental separation led to the opening of the Drake Passage and the formation of the Antarctic Circumpolar Current (ACC) which today dominates the Southern Ocean circulation, and contributed to cold isolation of the Antarctic (Barker and Burrell 1977) mainly determined by the decrease of atmospheric CO_2 during Cenozoic (DeConto and Pollard 2003). The ACC flows clockwise around Antarctica and is delimited by the Sub-Antarctic Front (SAF) and the Antarctic Polar Front (APF) which have major influence over the distribution of the marine benthic biota acting as a barrier to connectivity between Antarctic and sub-Antarctic species/populations (Barnes et al. 2006; Griffiths and Waller 2016) while at the same time facilitating connectivity in the Subantarctic region (Leese et al. 2010; Moore et al. 2018). In fact, several co-distributed Antarctic and South American invertebrates show a marked molecular divergence, a fact that reveals a separated evolutionary history in isolation for millions of years (Held 2000, 2001; Moon et al. 2017). However, the diversification of several marine groups occurred long after the physical separation of the continental landmasses indicating occasional dispersal across the APF on long timescales (Duenas et al. 2016; González-Wevar et al. 2017). Recently the number of studies reporting the cross invasion of species from both continents, including coastal invertebrates, has increased significantly due to the increment of ship traffic between the two continents (Hughes et al. 2015; Baird et al. 2020) (see section 1.2.1).

High latitudes are particularly sensitive to global warming. Increasing of mean annual temperature, has generated a general recession of most of the Patagonian, Fuegian and WAP glaciers, mostly due to loss of accumulation area, rising temperatures at the glacier snout elevation, and increase of ice calving in lakes or the sea (Rabassa 2009; Cook et al.

2016). Within the glacier-fjord system of Southern Chilean Patagonia, the Magellan Strait separates the American continent from the Tierra del Fuego Island, and connects the Pacific and Atlantic oceans. In this region, freshwater from glacier melting and river discharge mixes with Sub-Antarctic oceanic water, resulting in a significant decrease in salinity and implying a source of stress to stenohaline organisms (Detree et al. 2019). On the other hand, Patagonia's fjords are currently highly threatened by the increment of salmon farming. Together with enhanced organic matter loading from the watershed, primarily related to agriculture and sewage discharge, and the enhanced stratification of water column due to seasonal warming and precipitation, human interference by salmon culture entails an increasing risk of local eutrophication and oxygen deficiency in bottom water layers (Levin et al. 2009; Quinones et al. 2019).

The Antarctic Peninsula exhibits one of the most rapid rates of regional warming anywhere on earth. On King George Island (KGI), the largest of the South Shetland Islands, most of the glaciers that discharge into fjords and the coastal ocean are retreating at an unprecedented speed (Cook et al. 2005; Osmanoğlu et al. 2013; Cook et al. 2016). The accelerated retreat velocity of land and tidewater glaciers and West Antarctic ice sheet result, in turn, in reduced surface salinity an enhanced turbidity in coastal waters (Schloss et al. 2012; Monien et al. 2017). Enhanced pelagic primary production and colonization of macroalgae in newly ice-free areas causes local increase of organic matter deposition and its microbial recycling, reducing oxygen availability in surface sediments and bottom water of coastal shallows (Braeckman et al. 2019). Indeed, in many areas of the WAP, as Potter Cove - a small fjord on KGI - suboxic and even seasonally anoxic zones have been detected in the surface sediment layers (Monien et al. 2014; Wagner et al. 2015). Moreover, human activities in Antarctica including tourism and scientific research have intensified substantially in the last fifty years, with the concomitant impact on the Antarctic ecosystems through the release of wastewater from scientific stations and the use of fossil fuels for energy production and boat traffic. In fact, the presence of petroleum-derived hydrocarbons has been observed in Antarctic coastal sediments of Potter Cove, with highest concentrations recorded near the glacier front (Curtosi et al. 2009).

Thus, SSA and WAP fjords experience comparable effects related to global change and harbour common or closely related species, making this biogeographic and climatic gradient an ideal large-scale model for studying the role of genetic and phenotypic traits of species/populations in adaptation to changing environments. On the other hand, Potter Cove, a small fjord about 4 km long and 2.5 km wide in KGI, has become a showcase area to study the response to climate change of Antarctic coastal biota at local scale, not only for its logistical facilities (Argentinian Carlini Station –Dallmann Ger-Arg Laboratory), but also for its rapid and noticeable environmental change related to the retreat of the Fourcade glacier that surrounds the cove and the availability of decades-old environmental data timeseries (Kim et al. 2018). Its natural characteristics and the differential impact of the melting glacier across the cove (described in detail in Chapter 3) imply a high degree of environmental heterogeneity (Jerosch et al. 2018). This makes Potter Cove ideal to study the interaction between local habitat heterogeneity and intra-population variability of species traits (genetic and phenotypic standing variation) and its potential adaptive implications when populations are confronted with environmental shifts.

1.2 Responses of biota to global change: move, adapt or go extinct

The accelerated rate of climate change is expected to exceed the adaptive potential of species (Bradshaw and Holzapfel 2006), and as such, distribution shifts are considered the most likely response to environmental change (VanDerWal et al. 2012). Predictions based on global temperature suggest that species will shift poleward with warming. For those species unable to migrate, or in the case of polar organisms - without colder places available to migrate - their survival will depend on their adaptive potential to face changes in their current location. Antarctic organisms, however, have been traditionally considered as highly sensitive to environmental changes for inhabiting a continent characterized by great stability over the last 20 million years (Abele 2011).

1.2.1 Move: poleward expansion of lower-latitude species

Global warming has the potential to alter the boundaries of species tolerance ranges and to shift their biogeographic distributions. Utilizing a comprehensive data sets of species' thermal tolerance limits and distributions, Sunday et al. (2012) concluded that marine ectotherms are expected to expand at their poleward range boundaries and contract at their equatorward boundaries with climate warming. Thus, high latitude species/communities are not only threatened by inhabiting the fastest warming zones on earth (Clem et al. 2020), but also by the greater risk of being disrupted by non-indigenous species (NIS) from lower latitudes through competition for resources or trophic interactions (Chan and Briski 2017). Invasive NIS can lead to declines or even extinctions of native species, disrupt ecosystem functions, enhance transmission of viruses and pathogens, and cause substantial damage to natural resources and ecosystem services (Simberloff et al. 2013).

The historic isolation of the Southern Ocean explains the high levels of endemism in this region; up to 70% in some taxa (David and Saucède 2015). This isolation also implies that the Antarctic still has an almost entirely native biota (Convey et al. 2014), which makes this continent extremely sensitive to invasions of NIS. While many areas of the WAP may be vulnerable to invasion by species from comparable cold environments, such as the Arctic or high-altitude areas, the increased temperatures already registered in the WAP region and those predicted for 2100, render it suitable for the colonization of NIS from lower latitudes (Chown et al. 2012). Colonization of NIS can occur by natural processes, such as rafting on kelp or through pelagic life-stages (Fraser et al. 2018) but also via anthropogenic-derived vectors such as the increasing diversity and quantity of anthropogenic substrata to foul, such as plastics (Barnes 2002). Currently, biofouling on ship hulls or larval transport with bilge water are likely to be the most important vectors for transporting species to Antarctica, as vessel traffic between South America and Antarctica has increased significantly in the last decades. McCarthy et al. (2019) estimated that the number of active ships and voyages annually around Antarctic and sub-Antarctic islands has undergone a 5- to 10-fold increase since the 1960s. The most frequently travelled routes in the Southern Ocean connect South American and South

Atlantic ports to the WAP (Bender et al. 2016). Furthermore, the WAP is by far the region with the highest amount of tourism in the Antarctic. For these reasons, but also for the strongest similarities of habitats to SSA, the WAP region has more recordings of marine and terrestrial NIS than anywhere else in continental Antarctica (Hughes et al. 2015; McGeoch et al. 2015). However, regarding marine species, Antarctica has no confirmed established populations of NIS so far, albeit reports of occasional presence of species that were potentially transported by anthropogenic means are increasing, as for example: *Ulva intestinalis* (grass kelp), *Hyas araneus* (great spider crab), *Bugula neritina* (bryozoan), *Ciona intestinalis* (tunicate), and *Mytilus cf. platensis* (bank forming mussels) (McCarthy et al. 2019; Cardenas et al. 2020).

The concern about NIS in Antarctica due to warming and increased human activity has prompted the development of risk analyses to identify most likely invader species and vectors, and to provide stronger prevention and management advice (Hughes et al. 2020). There is a solid body of knowledge in the literature regarding physiological and behavioural adaptations of Antarctic organisms to their environment, as well as comparisons with related species from lower latitudes (see section 1.2.2). However, studies that actually evaluate physiological adaptation capacity/restrictions of potential NIS, especially those of lower latitudes closely related to WAP species in response to near future warming scenarios at the WAP, are still missing in the literature.

1.2.2 Adaptation to environmental change: ecological niche exploitation

When migration is not an option, the persistence of organisms facing environmental changes is determined by their ability to develop adaptive responses. Responses to environmental change range from molecular to community or ecosystem responses which vary with respect to the temporal and spatial scales of change (Convey and Peck 2019). Gene expression and physiological plasticity enable responses over hours to weeks, while alterations of gene frequencies in populations through selection, and behavioral or morphological trait modifications require longer time scales. Over years to centuries, evolutionary genetic responses and speciation are the main drivers

(Somero 2012). The preponderance of one process over the other then depends on the temporal scale but also on the generation time of the species. Species with short generation time respond mainly by genetic modification while species with long generation time (years, decades) depend on the standing genetic variation and phenotypic plasticity to survive long enough for evolutionary processes to take effect (Geerts et al. 2015; Bell 2017). Most of marine Antarctic fish and invertebrates are among the latter group, however, their long-term adaptation to constant cold climate has led to a loss of phenotypic plasticity which limits, to a different extent depending on the species, the possibility for an adaptive response to a thermal stress stimulus (Abele 2011). Best studied cases are the loss of genes centrally important for warm acclimation (including genes encoding for heat shock proteins or the heat shock transcription factor) and the inability to decrease mitochondrial densities and develop efficient antioxidant responses to counteract high levels of reactive oxygen species (ROS) arising from increased metabolic activity under warming conditions (Abele et al. 2001; Morley et al. 2009; Abele 2011). Much of the knowledge in this regard has been generated by comparing populations from the same or closely related species on biogeographical climate gradients (i.e., large spatial scale comparisons, e.g., Antarctic vs. South American populations), scaling-up the observations and inferring adaptive processes occurring on larger temporal scale (evolutionary time scale).

The importance of small-scale processes that may cumulatively operate and interact to determine patterns at biogeographical scales is often ignored. Abiotic factors are spatially and temporally heterogeneous, and as such, organisms do not respond to average changes in the environment but rather to fluctuations in habitat conditions (Helmuth et al. 2010). An increasing number of studies have pointed out that environmental heterogeneity can significantly buffer the sensitivity of organisms to changing oceanic conditions (Boyd et al. 2016; Choi et al. 2019). For example, recent studies highlighted the role of thermal refugia, as microhabitats where microclimatic conditions permit survival of organisms during extreme events (Potter et al. 2013; Dong et al. 2017). On the other hand, the physiological variation among individuals within and between populations is often an aspect not considered when predicting the effects of

temperature (or other variables) on species function and survival. In the intertidal limpet *Cellana toreuma*, the in situ expression of genes related to the heat shock response and energy regulation differ among individuals by more than 1000-fold at the same site and same date (Dong et al. 2014). In another limpet species, the Antarctic *Nacella concinna*, it has been observed that intertidal organisms grow steeper shells and adjust body volume to inner shell volume, in order to prevent desiccation during air exposure (Weihe and Abele 2008), whereas flatter shells can reduce damage of ice scour in the sublittoral environment (Morley et al. 2010). Jimenez et al. (2015) observed in the mussel *Mytilus californianus* a reduction in the magnitude of variation in tissue-level antioxidant capacities by approximately 30% when abolishing micro-scale environmental variation via common garden acclimation, highlighting the role of micro-scale environmental variation in the modulation of patterns of physiological variation among individuals.

Therefore, to accurately predict the ecological impacts of climate change on organisms requires the analysis of large-scale environmental signals broken down to niche-level processes, to elucidate how these larger scale scenarios ultimately translate to physiological responses at population level.

1.3 *Aequiyoldia eightsii* as a model species

1.3.1 Justification

The model species of this thesis, *Aequiyoldia cf. eightsii*, is a bivalve of the subclass Protobranchia (Fig. 1.2) and occurs in high abundance in both biogeographic regions targeted in this study. Given their wide geographical distribution, protobranch bivalves constitute a widely applicable model group to examine genetic traits and plastic responses supporting adaptation to different environmental conditions and regimes of variability, from Antarctic bivalves such as *A. cf. eightsii* to species located in North America and West European Basins such as *Deminucula atacellana* (Allen and Sanders 1996). This group is of particular interest because of its long evolutionary history (Cambrian) and abundant occurrence from shallow waters to the deeper oceans. Protobranch bivalves feed close to the sediment surface and play an important role in

surface sediment carbon turnover (Reed et al. 2013). Despite of this, there are very few studies that delve into the physiological and molecular responses to stress conditions in this group (Wilson and Davis 1984; Abele et al. 2001; Stead and Thompson 2003; Peck et al. 2009; Peck et al. 2010; Clark et al. 2016). Moreover, some particular aspects of protobranch evolutionary biology have only recently been discovered and are currently receiving much greater scientific attention. These include the relatively wide ranging incidence of doubly uniparental inheritance (DUI) in this group (Gusman et al. 2016), i.e., independent heritage of male mitochondrial genomes in gonads of male individuals, which could be extremely important in understanding their remarkable adaptive capacities and explain their widespread occurrence across so many different ecological settings.

1.3.2 General characteristics

A. cf. eightsii is an infaunal nuculanid bivalve dominant in soft-sediment habitats between few meters and *ca.* 800 m depth, but is most commonly found at depths shallower than 100 m. This species has a flexible feeding strategy, being able to switch between deposit and filter feeding strategies and consuming phytoplankton and organic detritus present in the surface layers of sediments. During deposit feeding, the animals carry out vertical feeding migrations in the sediment, involving intensive locomotory activity. Still, *A. cf. eightsii* burrows relatively shallowly compared to other protobranch species (Davenport 1988). In contrast to other protobranch species and many Antarctic invertebrates, the reproductive ecology of *A. cf. eightsii* exhibits continuous oogenesis, with a period of increased reproductive intensity and spawning during the austral winter, and asynchrony between females. This is attributed to an innate, endogenous rhythm related to energetic allocation between somatic and reproductive investment, but also to its flexible feeding strategy (suspension and deposit feeding) in response to seasonal changes in food supply characteristic of the Antarctic marine environment (Lau et al. 2018). Although it is not the case for *A. cf. eightsii,* within the subclass Protobranchia occurrence of doubly uniparental inheritance (DUI) has recently been

suggested (Gusman et al. 2016). DUI of mitochondria is an evolutionarily stable mechanism, alternative to the classical uniparental mitochondrial inheritance, which has so far been confirmed in over 100 bivalve species (Gusman et al. 2016; Ghiselli et al. 2019). In bivalves with DUI, females normally inherit mitochondria from their mothers (F-type) whilst males inherit mitochondria from both parental lines (F-type and M-type); therefore, males may possess two variants of the mitochondrial genome, a condition known as mitochondrial heteroplasmy. In males, the F-type is dominant in somatic tissues whereas the M-type is prominent in the gonads but can sometimes occur in small amounts also in the somatic tissue (Zouros and Rodakis 2019). Sex-associated mitochondrial DNA (mtDNA) can have a high degree of divergence which in some species of bivalves reaches up to 50 % of intraspecific DNA divergence (Doucet-Beaupré et al. 2010). Despite its widespread occurrence in the bivalve class, the adaptive value of this peculiar mechanism is largely unknown (Bettinazzi et al. 2019).

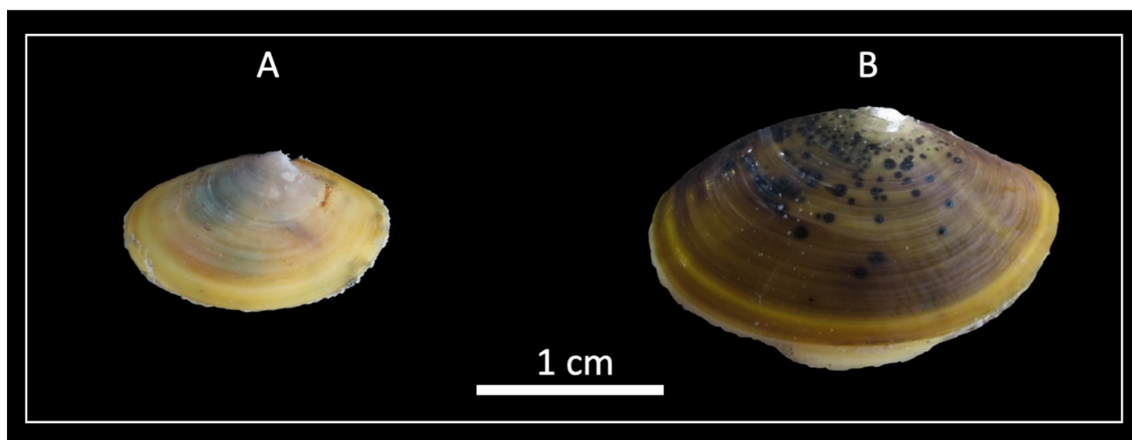


Figure 1.2 *Aequiyoldia cf. eightsii* specimens from southern South America (A) and King George Island – West Antarctic Peninsula (B)

1.3.3 Distribution and population genetics

Aequiyoldia cf. eightsii, previously genus *Yoldia* (Jay, 1839), exhibit a broad distribution across different Southern Ocean provinces such as South America, the Antarctic Peninsula, and the Kerguelen Islands; as based in morphological criteria several

synonymous species have been grouped within *A. cf. eightsii* (originally described from the South Shetland Islands), including *Yoldia kerguelensis* (Thiele, 1931), *Yoldia subaequilateralis* (E.A. Smith, 1875) from the Kerguelen Islands, and *Yoldia woodwardi* (Hanley, 1860) from the Falkland/Malvinas Islands (Linse et al. 2006; Sharma et al. 2013; González-Wevar et al. 2019). Thus, the Antarctic and South American populations were presently formally subsumed under a single species. They are morphologically identical (Fig. 1.2), although Antarctic individuals are larger than those recorded in South America (Chapter 2, Appendix 2.3). This wide distribution range on both sides of the Drake Passage has given rise to several taxonomic analyses (González-Wevar et al. 2012; Sharma et al. 2013). González-Wevar et al. (2019) suggested the presence of several cryptic species comprising two lineages along the Antarctic Peninsula (5.78 % *COI* p-distance), a third lineage in South America (6.5 – 7.5 % *COI* and 1.2 % *ITS* p-distances with respect to the Antarctic), and two additional mitochondrial lineages on the Kerguelen Island and Falkland/Malvinas Islands. Recently, (Muñoz-Ramírez et al. 2020) highlighted the role of the ACC as biogeographic barrier to larval transport between both continents preventing or massively reducing the potential for genetic connectivity between the Antarctic and South American *Aequiyoldia*.

1.3.4 Responses to environmental stress

A. cf. eightsii has been used as a model organism to understand the effects of climate change on Antarctic invertebrates through measurements of metabolism, oxidative stress, gene expression and survival in response to warming. Clark et al. (2016) categorized this species as being comparatively resistant to thermal stress when compared with other Antarctic invertebrates because it displays an extraordinary upper lethal temperature of 25 °C when warmed at a rate of 1 °C h^{-1}. However, Peck et al. (2010) showed that this species did not change its upper lethal temperature after acclimating during 60 days to a temperature 3.5 °C above annual average, which the authors interpreted as poor ability for thermal acclimation in an Antarctic stenotherm. In earlier experimental work in which *A. cf. eightsii* was ramped between -1°C and 5°C

at 1°C increments at 10 hour intervals, Abele et al. (2001) tested the effect of warming on routine and basal metabolism. Especially routine metabolism (RMR) increased moderately between -1 and +2°C and more pronouncedly at temperatures above the 2°C breakpoint temperature, as the animals increased the number of respiratory bouts of vigorous foot movements to dig down into the (unavailable) colder sediment layers. At temperatures between 2° and 5°C, the burrowing bouts became shorter but more frequent, showing exhaustion of energy reserves. Furthermore, the authors showed that the activity of the antioxidant enzyme superoxide dismutase (SOD) was inhibited when animals were maintained for a period of 5 days at a temperature above 2 °C, resulting in a concomitant increased of oxidative damage (membrane lipid peroxidation).

1.4 Objectives

The main objective of this thesis is to understand the genetic and phenotypic traits that support adaptation to current and future environmental change at small (i.e., local or population scale) and large-scale (i.e., continental or species scale) analyzing patterns of gene expression in the Southern Ocean protobranch bivalve *Aequiyoldia cf. eightsii* from WAP and SSA.

1.4.1 Specific objectives of chapters

<u>Chapter 2</u>

The unresolved taxonomic status of *Aequiyoldia cf. eightsii* and the new findings of recent studies for WAP populations based on a single mitochondrial marker motivated me:

(i) to re-evaluate patterns of genetic divergence between *A. cf. eightsii* populations from SSA (Magellan region) and WAP (Potter Cove - King George Island and Hangar Cove - Adelaide Island)

(ii) to compare the results of a classic barcoding approach based on few selected mitochondrial and nuclear gene fragments with a complete set of Single Nucleotide Polymorphisms (SNPs) of the nuclear and mitochondrial transcriptomes. Particular attention was paid to the question whether and how mitochondrial heteroplasmy may affect population genetic and phylogenetic inference.

Chapter 3

With the "small-scale approach" in Potter Cove (WAP) I aimed to analyse in situ gene expression patterns in one local population of *A. cf. eightsii* in front of the melting Fourcade glacier to evaluate:

(i) the influence of the small-scale habitat heterogeneity on the patterns of gene transcription.

(ii) the influence of the genotype background, mitochondrial and nuclear separately, on the patterns of gene transcription.

Chapter 4

Using the large-scale approach, I aimed to experimentally analyse gene expression patterns in populations of *A. cf. eightsii* on both sides of the Drake passage (SSA and WAP) in response to changes in thermal and oxygen regimes expected under a global warming scenario to specifically evaluate:

(i) physiological responses of both population/species to the major stressors temperature and hypoxia

(ii) physiological traits that may allow for/restrict a cross continental invasion of SSA *Aequiyoldia* in a warming scenario on the WAP.

References

Abele D. 2011. Temperature adaptation in changing climate: marine fish and invertebrates. Temperature Adaptation in a Changing Climate: Nature at Risk: CAB International.67-79.

Abele D, Tesch C, Wencke P, Pörtner HO. 2001. How does oxidative stress relate to thermal tolerance in the Antarctic bivalve *Yoldia eightsi*? Antarct Sci. 13(2):111-118.

Allen J, Sanders H. 1996. The zoogeography, diversity and origin of the deep-sea protobranch bivalves of the Atlantic: the epilogue. Prog Oceanogr. 38(2):95-153.

Baird HP, Moon KL, Janion-Scheepers C, Chown SL. 2020. Springtail phylogeography highlights biosecurity risks of repeated invasions and intraregional transfers among remote islands. Evol. Appl. 13(5):960-973.

Barker P, Burrell J. 1977. The opening of Drake passage. Mar Geol. 25(1-3):15-34.

Barnes DK. 2002. Invasions by marine life on plastic debris. Nature. 416(6883):808-809.

Barnes DK, Hodgson D, Convey P, Allen C, Clarke A. 2006. Incursion and excursion of Antarctic biota: past, present and future. Global Ecol Biogeogr. 15(2):121-142.

Bell G. 2017. Evolutionary rescue. Annu. Rev. Ecol. Evol. Syst. 48:605-627.

Bender NA, Crosbie K, Lynch HJ. 2016. Patterns of tourism in the Antarctic Peninsula region: a 20-year analysis. Antarct Sci. 28(3):194-203.

Bettinazzi S, Rodriguez E, Milani L, Blier PU, Breton S. 2019. Metabolic remodelling associated with mtDNA: insights into the adaptive value of doubly uniparental inheritance of mitochondria. Proc Biol Sci. 286(1896):20182708.

Boyd PW, Cornwall CE, Davison A, Doney SC, Fourquez M, Hurd CL, Lima ID, McMinn A. 2016. Biological responses to environmental heterogeneity under future ocean conditions. Glob Chang Biol. 22(8):2633-2650.

Bradshaw WE, Holzapfel CM. 2006. Evolutionary response to rapid climate change. Science. 312(5779):1477-1478.

Braeckman U, Pasotti F, Vazquez S, Zacher K, Hoffmann R, Elvert M, Marchant H, Buckner C, Quartino ML, Mac Cormack W et al. 2019. Degradation of macroalgal detritus in shallow coastal Antarctic sediments. Limnol Oceanogr. 64(4):1423-1441.

Cardenas L, Leclerc JC, Bruning P, Garrido I, Detree C, Figueroa A, Astorga M, Navarro JM, Johnson LE, Carlton JT et al. 2020. First mussel settlement observed in Antarctica reveals the potential for future invasions. Sci Rep. 10(1):5552.

Chan FT, Briski E. 2017. An overview of recent research in marine biological invasions. Mar Biol. 164(6):121.

Choi F, Gouhier T, Lima F, Rilov G, Seabra R, Helmuth B. 2019. Mapping physiology: biophysical mechanisms define scales of climate change impacts. Conserv Physiol. 7(1):coz028.

Chown SL, Huiskes AH, Gremmen NJ, Lee JE, Terauds A, Crosbie K, Frenot Y, Hughes KA, Imura S, Kiefer K. 2012. Continent-wide risk assessment for the establishment of nonindigenous species in Antarctica. Proc. Natl. Acad. Sci. 109(13):4938-4943.

Clark MS, Sommer U, Sihra JK, Thorne MA, Morley SA, King M, Viant MR, Peck LS. 2016. Biodiversity in marine invertebrate responses to acute warming revealed by a comparative multi-omics approach. Glob Chang Biol. 23(1)318-330

Clem KR, Fogt RL, Turner J, Lintner BR, Marshall GJ, Miller JR, Renwick JA. 2020. Record warming at the South Pole during the past three decades. Nat. Clim. Change. 10(8):762-770.

Convey P, Chown SL, Clarke A, Barnes DK, Bokhorst S, Cummings V, Ducklow HW, Frati F, Green TA, Gordon S. 2014. The spatial structure of Antarctic biodiversity. Ecol Monogr. 84(2):203-244.

Convey P, Peck LS. 2019. Antarctic environmental change and biological responses. Sci. Adv. 5(11):eaaz0888.

Cook A, Fox A, Vaughan D, Ferrigno J. 2005. Retreating glacier fronts on the Antarctic Peninsula over the past half-century. Science. 308(5721):541-544.

Cook AJ, Holland P, Meredith M, Murray T, Luckman A, Vaughan DG. 2016. Ocean forcing of glacier retreat in the western Antarctic Peninsula. Science. 353(6296):283-286.

Curtosi A, Pelletier E, Vodopivez CL, Mac Cormack WP. 2009. Distribution of PAHs in the water column, sediments and biota of Potter Cove, South Shetland Islands, Antarctica. Antarct Sci. 21(4):329-339.

Davenport J. 1988. The feeding mechanism of *Yoldia* (═ *Aequiyoldia*) *eightsi* (Courthouy). Proc. Royal Soc. B. 232(1269):431-442.

David B, Saucède T. 2015. Biodiversity of the Southern Ocean. Elsevier.

DeConto RM, Pollard D. 2003. Rapid Cenozoic glaciation of Antarctica induced by declining atmospheric CO_2. Nature. 421(6920):245-249.

Detree C, Ortiz A, Navarro JM. 2019. Combined effects of warming and freshening on the physiological energetics of the edible whelk *Trophon geversianus*. Mar Environ Res.153:104840.

Dong Yw, Han Gd, Huang Xw. 2014. Stress modulation of cellular metabolic sensors: interaction of stress from temperature and rainfall on the intertidal limpet *Cellana toreuma*. Mol Ecol. 23(18):4541-4554.

Dong YW, Li XX, Choi FMP, Williams GA, Somero GN, Helmuth B. 2017. Untangling the roles of microclimate, behaviour and physiological polymorphism in governing vulnerability of intertidal snails to heat stress. Proc Biol Sci. 284(1854).

Doucet-Beaupré H, Breton S, Chapman EG, Blier PU, Bogan AE, Stewart DT, Hoeh WR. 2010. Mitochondrial phylogenomics of the Bivalvia (Mollusca): searching for the origin and mitogenomic correlates of doubly uniparental inheritance of mtDNA. BMC Evol Biol. 10(1):50.

Duenas LF, Tracey DM, Crawford AJ, Wilke T, Alderslade P, Sanchez JA. 2016. The Antarctic Circumpolar Current as a diversification trigger for deep-sea octocorals. BMC Evol Biol. 16:2.

Eirin-Lopez JM, Putnam HM. 2019. Marine environmental epigenetics. Annu. Rev. Mar. Sci. 11:335-368.

Fraser CI, Morrison AK, Hogg AM, Macaya EC, van Sebille E, Ryan PG, Padovan A, Jack C, Valdivia N, Waters JM. 2018. Antarctica's ecological isolation will be broken by storm-driven dispersal and warming. Nat. Clim. Change. 8(8):704-708.

Geerts A, Vanoverbeke J, Vanschoenwinkel B, Van Doorslaer W, Feuchtmayr H, Atkinson D, Moss B, Davidson TA, Sayer CD, De Meester L. 2015. Rapid evolution of thermal tolerance in the water flea *Daphnia.* Nat. Clim. Change. 5(7):665-668.

Ghiselli F, Maurizii MG, Reunov A, Arino-Bassols H, Cifaldi C, Pecci A, Alexandrova Y, Bettini S, Passamonti M, Franceschini V et al. 2019. Natural Heteroplasmy and Mitochondrial Inheritance in Bivalve Molluscs. Integr Comp Biol. 59(4):1016-1032.

González-Wevar C, Díaz A, Gerard K, Cañete JI, Poulin E. 2012. Divergence time estimations and contrasting patterns of genetic diversity between Antarctic and southern South America benthic invertebrates. Rev Chil Hist Nat. 85(4):445-456.

González-Wevar C, Gérard K, Rosenfeld S, Saucède T, Naretto J, Díaz A, Morley SA, Brickle P, Poulin E. 2019. Cryptic speciation in Southern Ocean *Aequiyoldia eightsii* (Jay, 1839): Mio-Pliocene trans-Drake Passage separation and diversification. Prog Oceanogr. 174:44-54.

González-Wevar CA, Hüne M, Segovia NI, Nakano T, Spencer HG, Chown SL, Saucède T, Johnstone G, Mansilla A, Poulin E. 2017. Following the Antarctic Circumpolar Current: patterns and processes in the biogeography of the limpet *Nacella* (Mollusca: Patellogastropoda) across the Southern Ocean. J Biogeogr. 44(4):861-874.

Griffiths HJ, Waller CL. 2016. The first comprehensive description of the biodiversity and biogeography of Antarctic and Sub-Antarctic intertidal communities. J Biogeogr. 43(6):1143-1155.

Gusman A, Lecomte S, Stewart DT, Passamonti M, Breton S. 2016. Pursuing the quest for better understanding the taxonomic distribution of the system of doubly uniparental inheritance of mtDNA. PeerJ. 4:e2760.

Held C. 2000. Phylogeny and biogeography of serolid isopods (Crustacea, Isopoda, Serolidae) and the use of ribosomal expansion segments in molecular systematics. Mol Phylogen Evol. 15(2):165-178.

Held C. 2001. No evidence for slow-down of molecular substitution rates at subzero temperatures in Antarctic serolid isopods (Crustacea, Isopoda, Serolidae). Polar Biol. 24(7):497-501.

Helmuth B, Broitman BR, Yamane L, Gilman SE, Mach K, Mislan KA, Denny MW. 2010. Organismal climatology: analyzing environmental variability at scales relevant to physiological stress. J Exp Biol. 213(6):995-1003.

Hughes KA, Pertierra LR, Molina-Montenegro MA, Convey P. 2015. Biological invasions in terrestrial Antarctica: what is the current status and can we respond? Biodivers Conserv. 24(5):1031-1055.

Hughes KA, Pescott OL, Peyton J, Adriaens T, Cottier-Cook EJ, Key G, Rabitsch W, Tricarico E, Barnes DK, Baxter N. 2020. Invasive non-native species likely to threaten biodiversity and ecosystems in the Antarctic Peninsula region. Global Change Biol. 26(4):2702-2716.

Jerosch K, Pehlke H, Monien P, Scharf F, Weber L, Kuhn G, Braun MH, Abele D. 2018. Benthic meltwater fjord habitats formed by rapid glacier recession on King George Island, Antarctica. Philos Trans A Math Phys Eng Sci. 376(2122).

Jimenez AG, Jayawardene S, Alves S, Dallmer J, Dowd WW. 2015. Micro-scale environmental variation amplifies physiological variation among individual mussels. Proc Biol Sci. 282(1820):20152273.

Kim H, Ducklow HW, Abele D, Ruiz Barlett EM, Buma AG, Meredith MP, Rozema PD, Schofield OM, Venables HJ, Schloss IR. 2018. Inter-decadal variability of phytoplankton biomass along the coastal West Antarctic Peninsula. Philos. Trans. R. Soc. A. 376(2122):20170174.

Koenigstein S, Pohlmann K, Held C, Abele D. 2013. Ecological comparison of cellular stress responses among populations - normalizing RT-qPCR values to investigate differential environmental adaptations. BMC Ecol. 13:21

Lau SCY, Grange LJ, Peck LS, Reed AJ. 2018. The reproductive ecology of the Antarctic bivalve *Aequiyoldia eightsii* (Protobranchia: Sareptidae) follows neither Antarctic nor taxonomic patterns. Polar Biol. 41(9):1693-1706.

Leese F, Agrawal S, Held C. 2010. Long-distance island hopping without dispersal stages: transportation across major zoogeographic barriers in a Southern Ocean isopod. Naturwissenschaften. 97(6):583-594.

Levin L, Ekau W, Gooday A, Jorissen F, Middelburg J, Naqvi S, Neira C, Rabalais N, Zhang J. 2009. Effects of natural and human-induced hypoxia on coastal benthos. Biogeosciences, 6(10):2063–2098

Linse K, Griffiths HJ, Barnes DK, Clarke A. 2006. Biodiversity and biogeography of Antarctic and sub-Antarctic mollusca. Deep Sea Research Part II. 53(8-10):985-1008.

McCarthy AH, Peck LS, Hughes KA, Aldridge DC. 2019. Antarctica: The final frontier for marine biological invasions. Glob Chang Biol. 25(7):2221-2241.

McGeoch MA, Shaw JD, Terauds A, Lee JE, Chown SL. 2015. Monitoring biological invasion across the broader Antarctic: a baseline and indicator framework. Global Environ Change. 32:108-125.

Monien D, Monien P, Brünjes R, Widmer T, Kappenberg A, Busso AAS, Schnetger B, Brumsack H-J. 2017. Meltwater as a source of potentially bioavailable iron to Antarctica waters. Antarct Sci. 29(3):277-291.

Monien P, Lettmann KA, Monien D, Asendorf S, Wölfl A-C, Lim CH, Thal J, Schnetger B, Brumsack H-J. 2014. Redox conditions and trace metal cycling in coastal sediments from the maritime Antarctic. Geochim Cosmochim Acta. 141:26-44.

Moon KL, Chown SL, Fraser CI. 2017. Reconsidering connectivity in the sub-Antarctic. Biol Rev Camb Philos Soc. 92(4):2164-2181.

Moore JM, Carvajal JI, Rouse GW, Wilson NG. 2018. The Antarctic Circumpolar Current isolates and connects: Structured circumpolarity in the sea star Glabraster antarctica. Ecol. Evol. 8(21):10621-10633.

Morley S, Clark M, Peck L. 2010. Depth gradients in shell morphology correlate with thermal limits for activity and ice disturbance in Antarctic limpets. J Exp Mar Biol Ecol. 390(1):1-5.

Morley SA, Lurman GJ, Skepper JN, Pörtner H-O, Peck LS. 2009. Thermal plasticity of mitochondria: a latitudinal comparison between Southern Ocean molluscs. Comp. Biochem. Physiol. Part A Mol. Integr. 152(3):423-430.

Muñoz-Ramírez CP, Barnes DK, Cárdenas L, Meredith MP, Morley SA, Roman-Gonzalez A, Sands CJ, Scourse J, Brante A. 2020. Gene flow in the Antarctic bivalve *Aequiyoldia eightsii* (Jay, 1839) suggests a role for the Antarctic Peninsula Coastal Current in larval dispersal. R. Soc. Open Sci. 7(9):200603.

Osmanoğlu B, Braun M, Hock R, Navarro F. 2013. Surface velocity and ice discharge of the ice cap on King George Island, Antarctica. Annals of Glaciology. 54(63):111-119.

Peck L, Morley S, Clark M. 2010. Poor acclimation capacities in Antarctic marine ectotherms. Mar Biol. 157(9):2051-2059.

Peck LS, Clark MS, Morley SA, Massey A, Rossetti H. 2009. Animal temperature limits and ecological relevance: effects of size, activity and rates of change. Funct Ecol. 23(2):248-256.

Potter KA, Arthur Woods H, Pincebourde S. 2013. Microclimatic challenges in global change biology. Global Change Biol. 19(10):2932-2939.

Quinones RA, Fuentes M, Montes RM, Soto D, León-Muñoz J. 2019. Environmental issues in Chilean salmon farming: a review. Rev. Aquac. 11(2):375-402.

Rabassa J. 2009. Impact of Global Climate Change on Glaciers and Permafrost of South America, with Emphasis on Patagonia, Tierra del Fuego, and the Antarctic Peninsula. Natural Hazards and Human-Exacerbated Disasters in Latin America. 13:415-438.

Reed AJ, Morris JP, Linse K, Thatje S. 2013. Plasticity in shell morphology and growth among deep-sea protobranch bivalves of the genus *Yoldiella* (*Yoldiidae*) from contrasting Southern Ocean regions. Deep Sea Res. Part I Oceanogr. Res. Pap. 81:14-24.

Scher HD, Martin EE. 2006. Timing and climatic consequences of the opening of Drake Passage. Science. 312(5772):428-430.

Schloss IR, Abele D, Moreau S, Demers S, Bers AV, González O, Ferreyra GA. 2012. Response of phytoplankton dynamics to 19-year (1991–2009) climate trends in Potter Cove (Antarctica). J Mar Syst. 92(1):53-66.

Seebacher F, Franklin CE. 2012. Determining environmental causes of biological effects: the need for a mechanistic physiological dimension in conservation biology. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 367(1596):1607-1614.

Sharma PP, Zardus JD, Boyle EE, Gonzalez VL, Jennings RM, McIntyre E, Wheeler WC, Etter RJ, Giribet G. 2013. Into the deep: a phylogenetic approach to the bivalve subclass Protobranchia. Mol Phylogenet Evol. 69(1):188-204.

Simberloff D, Martin J-L, Genovesi P, Maris V, Wardle DA, Aronson J, Courchamp F, Galil B, García-Berthou E, Pascal M. 2013. Impacts of biological invasions: what's what and the way forward. Trends Ecol Evol. 28(1):58-66.

Somero GN. 2012. The physiology of global change: linking patterns to mechanisms. Annu. Rev. Mar. Sci. 4:39-61.

Stead RA, Thompson RJ. 2003. Physiological energetics of the protobranch bivalve *Yoldia hyperborea* in a cold ocean environment. Polar Biol. 26(2):71-78.

Stearns SC. 2015. The concept of phenotypic plasticity and the evolution of phenotypic plasticity in life history traits. in: Love A. (eds) Conceptual change in biology. boston studies in the philosophy and history of science. 307:131-146.

Sunday JM, Bates AE, Dulvy NK. 2012. Thermal tolerance and the global redistribution of animals. Nat. Clim. Change. 2(9):686-690.

Tanner RL, Dowd WW. 2019. Inter-individual physiological variation in responses to environmental variation and environmental change: Integrating across traits and time. Comp. Biochem. Physiol. Part A Mol. Integr. 238:110577.

VanDerWal J, Murphy HT, Kutt AS, Perkins GC, Bateman BL, Perry JJ, Reside AE. 2012. Focus on poleward shifts in species' distribution underestimates the fingerprint of climate change. Nat. Clim. Change. 3(3):239-243.

Wagner M, Hendy IL, McKay JL, Pedersen TF. 2015. Redox chemistry of West Antarctic Peninsula margin surface sediments. Chem. Geol. 417:102-114.

Weihe E, Abele D. 2008. Differences in the physiological response of inter-and subtidal Antarctic limpets *Nacella concinna* to aerial exposure. Aquat. Biol. 4(2):155-166.

Wilson JG, Davis JP. 1984. The effect of environmental variables on the oxygen consumption of the protobranch bivalve *Nucula turgida* (Leckenby and Marshall). J Molluscan Stud. 50(2):73-77.

Zouros E, Rodakis GC. 2019. Doubly Uniparental Inheritance of mtDNA: An Unappreciated Defiance of a General Rule. Adv Anat Embryol Cell Biol. 231:25-49.

CHAPTER 2

2 Publication I

Mariano Martínez, Lars Harms, Doris Abele, Christoph Held

Mitochondrial heteroplasmy as a systematic bias in molecular species delimitation and barcoding

Manuscript status: Accepted with minor revisions in *Molecular Ecology Resources*

Mitochondrial heteroplasmy as a systematic bias in molecular species delimitation and barcoding

Mariano Martínez[1], Lars Harms[2, 3], Doris Abele[1], Christoph Held*[1]

1) Functional Ecology, Helmholtz Centre for Polar and Marine Research, Alfred Wegener Institute, Am Handelshafen 12, 27570 Bremerhaven, Germany

2) Helmholtz Centre for Polar and Marine Research, Alfred Wegener Institute, Am Handelshafen 12, 27570 Bremerhaven, Germany

3) Helmholtz Institute for Functional Marine Biodiversity at the University of Oldenburg (HIFMB), Ammerländer Herrstrasse 231, 26129 Oldenburg, Germany

*Corresponding author: Christoph Held. Address: Am Handelshafen 12, 27570 Bremerhaven; Phone: +49(471)4831-2005; E-mail: Christoph.Held@awi.de

ABSTRACT

Molecular species delimitation is commonly performed by extrapolating differentiation patterns from small regions to the entire genome. Here we compare different data sources (mitochondrial *COI* sequences and nuclear and mitochondrial SNPs) for the marine bivalve species complex *Aequiyoldia cf. eightsii* in southern South America (SSA) and West Antarctic Peninsula (WAP). Whilst all data suggest that populations on either side of the Drake Passage belong to different species, the picture is less clear within WAP populations which harbor three distinct mitochondrial lineages (p-dist ≈ 6 %) that coexist in populations as well as in a subset of individuals. Haplotype-specific amplification and mitochondrial SNPs data demonstrate that two haplotypes occur in heteroplasmy in some individuals (h1h3 and h2h3). Standard barcoding procedures using universal primers lead to amplification bias favouring either haplotype unpredictably and thus overestimate the species richness with high confidence. In nuclear SNPs, no differentiation akin to the trans-Drake comparison can be observed inside the Southern Ocean, suggesting that 1) the Antarctic populations represent a single species and 2) their mitochondrial haplotypes h1, h2, h3 evolved during periods of temporary allopatry and survived as non-recombining mitochondrial haplotypes whereas recombination eroded similar differentiation patterns in the nuclear genome after secondary contact. Mitochondrial heteroplasmy in combination with amplification bias are possibly more common problems in molecular barcoding but remain undetected because evidence for it is eliminated during quality checking sequencing results. As a countermeasure, an active search in electropherograms for secondary peaks and reads with alternative bases in Illumina data as well as haplotype-specific primers for amplification are recommended.

Key words: Mitochondrial Heteroplasmy, Amplification bias, Mitochondrial DNA, DNA Barcoding, *Aequiyoldia cf. eightsii*.

1.INTRODUCTION

The accurate discrimination and identification of species is pivotal to the understanding of biodiversity patterns and the response of communities to environmental change. Since DNA barcoding was established (Hebert et al. 2003), the sequencing of one or several standard DNA regions has generated almost blind confidence when delimiting species, or when establishing genetic distances between populations in countless studies of phylogeny and phylogeography (DeSalle and Goldstein 2019). A fragment of the mitochondrial gene *cytochrome c oxidase subunit I* (*COI*) is one of the "gold standards" in molecular barcoding in animals and has been applied in a wide variety of taxa (Folmer et al. 1994; Steinke et al. 2016). The utility of mitochondrial DNA (mtDNA) for the purpose of species delimitation is generally attributed to its high mutation rate, maternal inheritance and absence of recombination (Rollins et al. 2016).

Despite the widespread application of this convenient approach to taxonomic identification (as of January 2020 > 8 million barcodes in BOLD database ; Ratnasingham and Hebert (2007)) and an invaluable progress in our understanding of global biodiversity patterns, the simplicity that prompted its success is also the one that limits it. Assuming a strictly uniparental inheritance of mtDNA through the maternal germline can overlook phenomena such as the occurrence of pseudogenes or mitochondrial heteroplasmy paving the way for incorrect phylogenetic and taxonomic inferences. The risk of inadvertently using nuclear mitochondrial pseudogenes (Numts) in barcoding analyses is considered to some degree in the barcoding literature (Kim et al. 2013; Francoso et al. 2019). However, the existence of different mitochondrial genomes in a single organism (heteroplasmy) has scarcely been suggested as a source of bias (Berthier et al. 2011; Mastrantonio et al. 2019). To the best of our knowledge, there are no concrete examples in the literature in which this peculiar feature of the mitochondrial genome has been identified as a source of misinterpretation of patterns of genetic divergence. The reason for this is many fold, but the exclusion of sequencing results characterized by the superposition of several templates in heteroplasmic organisms may be a contributing factor.

Heteroplasmy has been found in a wide range of taxa from mollusks and arthropods (Doublet et al. 2008; Degletagne et al. 2016; Mastrantonio et al. 2019) to vertebrates including fish, birds and mammals (Bentzen et al. 1988; Kvist et al. 2003; Vollmer et al. 2011). The best studied heteroplasmic system in metazoans is the Doubly Uniparental Inheritance (DUI) of mitochondria, an evolutionarily stable mechanism distributed across over 100 bivalve species (Gusman et al. 2016; Ghiselli et al. 2019). In bivalves with DUI, females normally inherit mitochondria from their mothers (F-type) whilst males inherit mitochondria from both their mother and father (F-type and M-type). In males, the F-type is dominant in somatic tissues whereas the M-type is prominent in the gonads but can occur in small amounts in the somatic tissue (Zouros and Rodakis 2019), too. Sex-associated mtDNAs can be extremely divergent with up to 50 % intraspecific DNA divergence (Doucet-Beaupré et al. 2010), well above the divergence level typically observed among bivalve species (Terranova et al. 2007).

DUI is widespread within bivalve subclasses including Protobranchia (Boyle and Etter 2012), the most basal extant bivalves with a poorly resolved phylogeny (Sharma et al. 2013). Protobranchs are of particular interest because of their high abundance in the deep oceans and important role in surface sediment carbon turnover (Reed et al. 2014). The high latitude marine protobranch *Aequiyoldia cf. eightsii* (previously genus *Yoldia*) is widely distributed in the Southern Ocean inhabiting soft-substratum ecosystems of South America, the Antarctic Peninsula, and several Subantarctic islands. This wide distribution range on both sides of the Drake Passage has given rise to several revisions of its taxonomic status (González-Wevar et al. 2012; Sharma et al. 2013). Recently, González-Wevar et al. (2019) suggested the presence of several cryptic species comprising two lineages along the WAP (5.78 % *COI* p-distance), a third linage in SSA (6.5 – 7.5 % *COI* and 1.2 % *ITS* p-distances) with respect to the Antarctic and two additional mitochondrial lineages in Kerguelen Island and Falkland/Malvinas Islands. However, the scarcity of available nuclear information so far hampers a formal taxonomic delimitation, hence *A. cf. eightsii* is being still considered a single species.

The advances in speed and accuracy in tandem with reduced costs of next-generation sequencing (NGS) technologies have made it possible to generate population-scale

genomic surveys. Transcriptome sequencing (RNA-seq) is particularly useful for studying non-model organisms lacking reference genomes (Low et al. 2019). Recently, the number of studies using single nucleotide polymorphisms (SNPs) from RNA-seq data has increased significantly thus establishing transcriptomics as an alternative data source for population genomic studies (De Wit et al. 2015; Yan et al. 2017; Garcia et al. 2019). The key analytical advantage of gaining population genomic insight from expressed sequences is the reduction of complexity by limiting the sequence data to a smaller fraction devoid of repetitive and heterochromatic sequences (Rogier et al. 2018).

Based on the hypothesis of heteroplasmy being a potential confounding factor in population genetics and phylogenetic analyses, this study evaluates patterns of genetic divergence between SSA and WAP *A. cf. eightsii* populations. Our analysis compares the results of a classic barcoding approach based on selected mitochondrial and nuclear gene fragments with SNPs derived from a deep sequencing of expressed genes with particular attention on how heteroplasmy, if any, may affect the population genetic and phylogenetic inference.

2. MATERIALS AND METHODS

2.1 Animal collection

Aequiyoldia cf. eightsii were collected in southern South America (SSA, Magellan region) and West Antarctic Peninsula (WAP) in the austral summer season of October 2017 until February 2018. Bivalves from the Magellan region (n= 100) were collected by SCUBA divers in the shallow subtidal near Punta Arenas (PA; Chile, 53°37'52"S; 70°56'54"W) on a single day in October 2017. Antarctic bivalves (n= 104) were collected between January and February 2018 from three sites in Potter Cove (PC), an 8 km^2 glacial fjord on King George Island - South Shetlands (62°14'11"S; 58°40'14"W - 62°13'32"S; 58°38'31"W - 62°13'35"S; 58°40'58"W), in shallow waters between 6 and 20 meters using a Van Veen grab. Maximal distance between the three collection sites in PC was 1 km, the sites were sampled within approximately one week. An additional subset of samples (n= 34) collected in Hangar Cove in February 2007 (HC; near Rothera station, Adelaide Island)

and Potter Cove in January 2016 and conserved at -20 °C were provided by colleagues from the British Antarctic Survey and the University of Cordoba (Argentina), respectively. These samples were exclusively used for the mitochondrial and nuclear loci analyses. In total, we obtained 238 individual organisms.

During each sampling event, bivalves were transported to the local research facility in insulated containers with seawater and sediment from the sampling site and dissected under a stereomicroscope, after recording individual weight, shell length and width. For each individual, three tissue types were conserved: Foot and mantle tissue samples were conserved separately in RNAlater (SIGMA) and stored at -80 °C, and the rest of the soft body was stored in Ethanol 95 % at 4 °C until further analysis. In total, 168 bivalves were included in the analysis. All of them were utilized for the mitochondrial and nuclear markers analyses, and a subset of 70 individuals were analysed for SNPs on RNA-seq data.

2.2 Mitochondrial and nuclear sequence analysis

The molecular markers included one mitochondrial gene (*cytochrome c oxidase subunit I*, *COI*) amplified using universal primers as well as haplotype-specific primers (see point 2.5 below), one nuclear protein-encoding gene (*histone H3*), and one nuclear ribosomal gene (*18S ribosomal RNA*, *18S*). Total genomic DNA was extracted from muscle tissue with QIAamp DNA Mini Kit (QIAGEN) following the manufacturer's protocol (DNA purification from tissues). *COI* amplification and sequencing amounted to 155 individuals (*n*PA= 50, *n*PC= 82, *n*HC= 24), *histone H3* and *18S* fragments were amplified and sequenced for 72 animals (*n*PA= 24, *n*PC= 24, *n*HC= 24). DNA purity and concentration were assessed using a Nanodrop Spectrophotometer® ND-1000 (NanoDrop Technologies, USA). Purified genomic DNA was diluted to 10 ng/µl and used as template for Polymerase Chain Reaction (PCR) amplification. All PCRs were carried out in 25 µl reaction volumes comprising 5.0 µl 5X Colorless GoTaq® Flexi Buffer (Promega Corp. USA), 2 µl $MgCl_2$ (25 mM), 2.5 µl dNTP mix (2 mM), 0.25 µl Betain (5M), 0.125 of each forward and reverse primer (100 µM), 0.15 µl GoTaq® G2 Flexi DNA Polymerase (5 U/µl,

PROMEGA Corp. USA) and 3 µl DNA extract (10 ng/µl). A fragment of 640 bp corresponding to *COI* was amplified using universal primers (Folmer et al. 1994) through a PCR program which consisted of 2 min at 95 °C followed by 35 cycles at 95 °C for 20 s, 46 °C for 20 s, 72 °C for 40 s, and a final extension of 8 min at 72 °C. For the amplification of *histone H3* and *18S* fragments (257 and 752 bp respectively), new primers were designed based on sequences of species from the same family published in Sharma et al. (2013) (*histone H3*: forward primer: 5'-GAA AAT CTA CCG GTG GCA AG-3', reverse primer: 5'-GTG TCC TCG AAC AAA CCA AC-3'; *18S*: forward primer: 5'-AAG TAC AGA CTC TCA GTA CGG-3', reverse primer: 5'-GAA GGC CAA CAA AAT AGA ACC-3') and the PCR program used for both fragments consisted in the following steps: 2 min at 95 °C, 35 cycles at 95 °C for 20 s, 56 °C for 20 s, 72 °C for 40 s with a final step of 8 min at 72 °C. After the amplification, a first quality assessment of the PCR was done by running the product on a 2 % agarose gel. All PCR products were sequenced for both strands by Eurofins Genomics GmbH (Ebersberg, Germany) using the same primers that were utilized during amplification. *COI*, *histone H3* and *18S* sequences were aligned using the ClustalW algorithm of CodonCode Aligner program (version 5.1.5; CodonCode Corporation, Dedham, MA), and all chromatograms were inspected visually for sequencing mistakes. For *COI*, a haplotype network was constructed using the Neighbour-Joining algorithm and implementing the Hasegawa-Kishino-Yano (HKY) model of substitution in Geneious using default parameters (version 8.1.9, Biomatters Ltd.) and Haplotype Viewer (Center of Integrative Bioinformatics Vienna, http://www.cibiv.at).

2.3 RNA extraction and cDNA library preparation

A subset of 70 individual samples out of the total included in mitochondrial and nuclear loci analysis were analysed for SNPs in RNA-seq data. A total of 70 transcriptomic libraries (one library per individual) were created in this study for animals from Punta Arenas and Potter Cove (*n*PA= 29, *n*PC= 41)

Samples of mantle tissue (5 – 30 mg) were homogenized in Trizol reagent (SIGMA) using a Precellys homogenizer (Precellys24, Bertin Technologies, France). Total RNA was isolated from each sample using the Direct-zol^{TM} RNA MiniPrep Kit (ZYMO Research Corp., USA) according to the manufacturer's instructions. The concentration and quality of the RNA was determined using a Nanodrop Spectrophotometer® ND-1000 (NanoDrop Technologies, USA) and LabChip® GX Touch (PerkinElmer, USA). Libraries were prepared using the Illumina TruSeq® Stranded mRNA Sample Preparation Kit starting from 1 µg of total RNA. The Poly-A containing mRNA molecules were purified using poly-T oligo attached magnetic beads. Subsequently, the mRNA was fragmented using divalent cations under elevated temperature and copied into first strand cDNA using reverse transcriptase and random primers. This was followed by second strand cDNA synthesis using DNA Polymerase I, RNase H and dUTP instead of dTTP to achieve strand specificity and to remove the RNA template. Following adapter ligation, products were enriched by PCR and purified to create the cDNA library. Libraries were validated and quantified using a LabChip® GX Touch (PerkinElmer, USA). All the samples were pooled and cleaned using magnetic beads (AMPure XP, Beckmann Coulter) to remove the remaining primer content. Final cDNA concentration was measure in the LabChip® GX Touch (PerkinElmer, USA). The pool of samples was sequenced on an Illumina NextSeq 500 sequencer using the NextSeq High Output Kit v2 (150 cycles) with a paired-end protocol.

2.4 De novo assembly and SNP analysis of RNA-seq data

Raw reads were quality controlled by FastQC v. 0.11.7 (Babraham Institute, Cambridge, UK). Adapter sequences were removed with bbduk.sh, from the BBtools suite, version 36.38 (Bushnell 2014) with following parameters: ktrim= r, k= 23, mink= 11, hdist= 1, tpe, tbo. Remaining sequences were searched for rRNA sequences by SortMeRNA version 2.1 (Kopylova et al. 2012) which were removed before further processing. To filter the sequences for the common Illumina spikein PhiX, bbduk.sh was used with a kmer size of 31 and a hdist of 1. A final quality trimming was performed with bbduk.sh

using Q10 as minimum quality and 36 bases as the minimum length. All obtained sequences (70 libraries) were normalized using bbnorm.sh (Bushnell 2014) with an average depth of 100x and a minimum depth of 5x before they were *de novo* assembled using the Trinity genome-independent transcriptome assembler version 2.8.4 (Grabherr et al. 2011) with a minimum transcript length of 300 bases and the option for strand specificity (--SS_lib_type RF). To remove duplicate sequences from the assembly, dedube.sh sh (Bushnell 2014) was used with the following parameters: minidenty = 98, arc = t, am = t, ac = t. The read representation and strand specificity of the assembly was assessed using the software Bowtie2 v.2.3.4.1 and the completeness evaluated using the package BUSCO v 3.0.0 (Benchmarking Universal Single-Copy Orthologs) and the orthologs of the public database ‘’eukaryota_odb9’’.

For the SNPs analysis, quality filtered paired-end reads were aligned to the de novo transcriptome using bowtie2 v2.3.4.1 (Langmead and Salzberg 2012). Alignments in SAM format (Sequence Alignment Map) were compressed and indexed with SAMtools v1.8 (Li et al. 2009). Genotype likelihoods were computed using mpileup from SAMtools and variant calling was performed with the BCFtools. In a first filtering, we excluded all variants with a Phred quality score below 30. Since our purpose was to analyze SNPs, we selected only SNPs from the variant calling which were present in at least one individual. A primary dataset of 1,308,131 SNPs in 70 individuals was filtered using VCFtools v0.1.16 (Danecek et al. 2011) and applying an iterative filtering strategy between loci and individuals with a progressive increase of cut-off values (O'Leary et al. 2018). A first filtering for loci quality was performed keeping variants successfully genotyped in 50 % of individuals (max-missing 0.5) and with a minor allele count (MAC) of 3. Subsequently, individuals with more than 37 % of missing data were excluded from the analysis. A second filtering for loci quality implied a max-missing of 0.95, a Minor Allele Frequency (MAF) of 0.05 and minimum number of reads (minDP) of 10. Additionally, variants with more than one allele were discarded (max-alleles 2). Nuclear and mitochondrial SNPs were separated through blasting the sequences containing SNPs against a mitochondrial database obtained from the UniProt Swiss-Prot database. A total of 113,515 SNPs in a final number of 54 individuals were retained. Of these, 4,714 were identified as

mitochondrial SNPs while nuclear SNPs comprised all SNPs in sequences with no hits in the mitochondrial database (108,801). Mitochondrial SNPs included in the analysis (392) were those contained in a single long transcript (*ca* 18 kbp), representing most of the mitochondrial genome (annotated using MITOS following Bernt et. al (2013), Appendix 2.1). To assess genetic differentiation between SSA and WAP animals, but also among animals of the different mitochondrial genotypes (mitotypes, see section 2.5), Principal Component Analyses (PCA), Analysis of Molecular Variance (AMOVA), and pairwise G_{ST}' estimations were carried out separately for nuclear and mitochondrial SNPs with all the individuals that remained after the filtering. G_{ST}' is a corrected version of the coefficient of genetic differentiation G_{ST}, which is a quotient of heterozygosity estimates obtained from a subpopulation and the whole population. A corrected value $G_{ST}' = G_{ST}/ G_{STmax}$ is computed because under common circumstances G_{ST} will never reach the theoretical maximum of 1 (Verity and Nichols 2014). Additional PCAs excluding the South American samples were performed to enhance differences between Antarctic mitotypes resulting from the analysis in our study. AMOVA was conducted considering three strata, location (SSA and WAP), mitotype and individuals, and significance was tested by randomly permutating (n= 1000) the sample matrices as described in Excoffier et al. (1992). All the analyses were performed in R v3.6 (RCore 2016) using the packages vcfR (Knaus and Grünwald 2017), adegenet (Jombart and Ahmed 2011), poppr v2.8.3 (Kamvar et al. 2014), ade4 (Dray and Dufour 2007) and pegas (Paradis 2010).

2.5 Coexistence of mitochondrial variants in a single individual

Observations of raw SNPs data (reads aligned to *de novo* transcriptome, see section 2.4) suggested the coexistence of more than one mitochondrial haplotype (of those previously defined based on *COI* data; see section 3.1) in a subset of individuals from Potter Cove (WAP). These specimens carry the haplotype h1 or h2 (never both) together with haplotype h3. To verify this, haplotype-specific primers were designed to selectively amplify the existing mitochondrial variants in individual samples and avoid artefacts caused by competitive differences of either mitochondrial variant in a

competitive PCR using a single universal primer pair for both amplicons. Primer design was performed on a long transcript (*ca.* 18 kbp) of the *de novo* transcriptome representing most of the mitochondrial genome, positioning them in regions that were variable among haplotypes (3 bp of difference). Thus, two forward primers were designed to amplify haplotype h3 (5´-AAT GTT AAT TTG TTC CAT GAG G-3´; P_b) and haplotypes h1, h2 and h4 (5´-AAT GTT AAT TTG TTC TAT GGC G-3´; P_a), and used in combination with one common reverse primer (5’-AGA AAA TAC AGC CCC CAT TC-3´) designed for both amplifications. The target amplicon (714 bp) intentionally overlapped (by *ca.* 300 bp) with the *COI* sequence amplified with universal primers in order to track the haplotype identification based on *COI*. A subset of samples of the four haplotypes from every location (*n*h1= 31, *n*h2= 27, *n*h3= 5, *n*h4= 16) were amplified with both primer combinations. PCR conditions for both primer combinations were the same as for *COI* with universal primers, with the exception of the annealing temperature (56 °C). The size of PCR products was checked on a 2 % agarose gel, and a representative subset of samples of the mitochondrial haplotypes successfully amplified for one or two variants (n= 32) were sequenced on both strands by Eurofins Genomics GmbH (Ebersberg, Germany). Finally, samples were aligned using the ClustalW algorithm of CodonCode Aligner program (version 5.1.5; CodonCode Corporation, Dedham, MA). This analysis allowed us to group the animals carrying a single or two combined haplotypes in different genetic clusters (mitotypes) which was further considered for SNPs analysis.

3. RESULTS

3.1 Mitochondrial and nuclear sequence analysis

The *18S ribosomal RNA* (*18S*) and *histone H3* loci were more conservative (i.e*.,* less variable) and less informative than *Cytochrome c oxidase subunit I* (*COI*) for the genetic differentiation of South American (PA, Punta Arenas) and Antarctic (PC and HC, Potter Cove and Hangar Cove respectively) populations. There were no differences between *histone H3* sequences, and only one polymorphic site was found in the *18S* fragment, which did not correspond to any geographic pattern in our sampling nor clades

suggested by mitochondrial or nuclear SNPs data (see section 3.3). In most of the cases, Sanger sequencing of *COI* amplified using universal primers yielded clear sequencing results. In some cases (n = 13), however, ambiguous base calls were observed that suggested a co-existence of several *COI* alleles in the template (i.e., heteroplasmy; see 3.2 below).

The results of *COI* sequencing with universal primers show a clear differentiation between Antarctic and South American individuals (h4), but also a division between two major Antarctic clades (h1+h2 vs h3; Fig. 2.1). The overall haplotype network includes 15 distinct haplotypes, four in South America (PA) and nine in Antarctica, of which seven were exclusive to PC or HC, whereas two were present at both Antarctic locations. The most frequent *COI* haplotypes (> 2 individuals) are labelled in Figure 2.1 (h1, h2, h3 and h4) and their genetic distances (p-distances) are shown in Appendix 2.2. The haplotypes h1 and h3 were found in PC and HC, whereas h2 was exclusively observed in PC in our samples. The high p-distances between the South American haplotype h4 and the Antarctic ones (6.6 – 8.6 %), and the distances between the rare h3 and the more frequent Antarctic h1 and h2 haplotypes (5.9 – 6.1 %) were remarkable. These differences represent four amino acid substitutions between the Antarctic haplotypes h1 and h2 and the South American h4, and five between the Antarctic haplotype h3 and h4. No amino acid substitutions were found between Antarctic haplotypes h1 and h2, whereas five substitutions were found between these two haplotypes and h3. No stop codons or indels were found in *COI* sequences of any haplotype/location. The persistence of differences in sympatry, absence of intermediate values and magnitude of differentiation similar to undisputed bivalve species pairs in *COI*, suggest treating the South American (h4), the common (h1 together with h2) and the rare (h3) Antarctic haplotypes as roughly equally differentiated clades, each possibly deserving candidate species status.

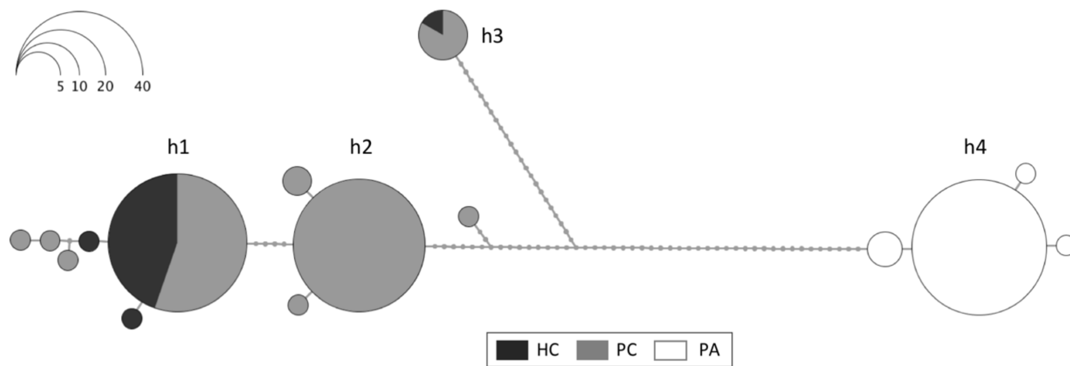


Figure 2.1. Haplotype network based on *COI* sequences of *Aequiyoldia cf.* bivalves (n= 155) amplified with universal primers from Hangar Cove (HC), Potter Cove (PC) and Punta Arenas (PA). Each haplotype is represented by a circle proportional to its frequency (scale top left), and the most frequent haplotypes (> 2 individuals) are label as h1, h2, h3 and h4.

3.2 Coexistence of mitochondrial variants in a single organism

The molecular categorization of individuals and the geographic distribution of *COI* haplotypes changed considerably, however, when haplotype-specific primers were used instead of universal primers. Specimens harbouring two different *COI* alleles were commonly observed (20 out of 63 Antarctic specimens). Amplification and sequencing of the target mitochondrial fragments using haplotype-specific primers designed in this study allowed to identify four *COI* haplotypes (h1, h2, h3, h4). The co-existence of two haplotypes in a single individual was detected in several Antarctic individuals from PC (19 out of 47) and HC (1 out of 16) (h1, h2, h3), but not in South American organisms (0 out of 16) (h4). In all but one individual that had been identified as carrying haplotype "h3-only" with universal primers (see 3.1), the haplotype h1 or h2 were also found. The presence of a second *COI* allele was easily demonstrable when using haplotype-specific primers, but when so-called universal primers were used only a single allele was amplified and sequenced, suggesting that one of the templates would outperform the other in a competitive scenario if a single, universal primer pair was used. A single bivalve from HC, in which variants h1 or h2 could not be amplified, was interpreted as

possessing haplotype h3 only (h3 homozygote). The reverse case, in which the universal primer only detected h1 or h2 but left h3 undetected was also observed. All South American specimens (haplotype h4) and the majority of Antarctic animals analysed qualified as homoplasmic for h1 or h2 with universal primers and also with the newly designed haplotype specific primers (Fig. 2.2). These findings allowed us to group the animals in six different genetic clusters based on their mitochondrial genotypes (mitotypes): animals carrying only one haplotype (h1, h2, h3, h4), from here on are named h1h1, h2h2, h3h3 and h4h4 respectively, and heteroplasmic animals for haplotype h1 or h2 together with the haplotype h3, are from here on called h1h3 and h2h3, respectively. Mitotypes frequencies at each location are shown in Figure 2.2. This grouping based on mitotypes was considered for SNPs analysis.

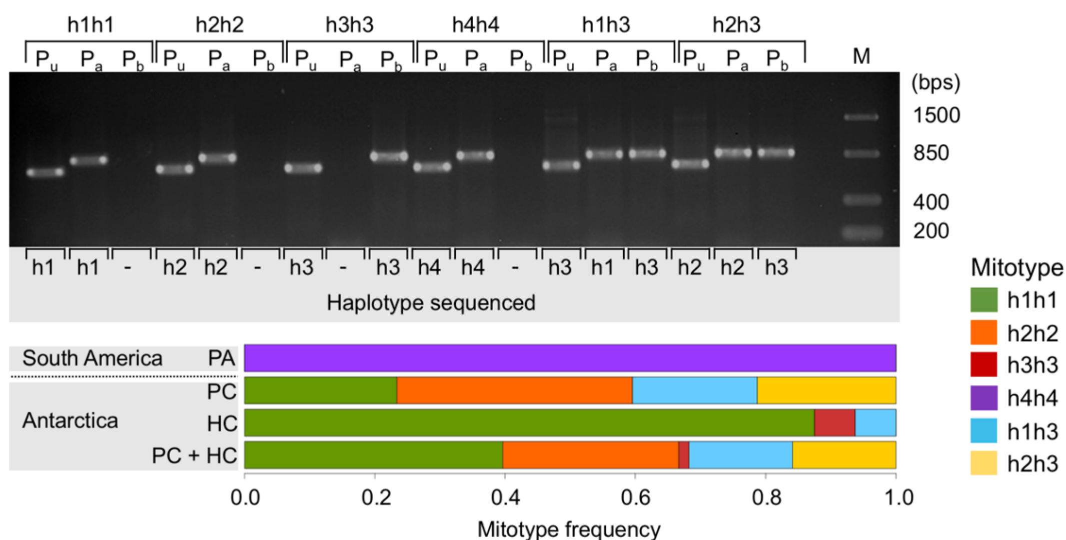


Figure 2.2. Above: Gel electrophoresis (2 % agarose) of PCR amplified products from individual samples of the six mitotypes (h1h1, h2h2, h3h3, h4h4, h1h3, h2h3) using universal primers for *COI* (P_u), and the two haplotype specific primer combinations designed in this study (P_a: h1, h2, h4 specific; P_b: h3 specific). Sanger sequencing results (haplotypes) of each PCR product are shown in the bottom of the gel. Below: Mitotypes frequencies of bivalves from Punta Arenas (PA, n= 16), Potter Cove (PC, n= 47), Hangar Cove (HC, n= 16) and both Antarctic locations together (PC + HC, n= 63).

3.3 De novo assembly and SNPs analysis

The Illumina sequencing resulted in 454.8 million reads (average of reads per library: 6.5 million), of which 388.4 million reads were used for the *de novo* transcriptome assembly, after filtering for quality. A total of 389,929 transcripts were assembled with an average length of 764 bp. The BUSCO assessing of the completeness of the transcriptome showed a very high quality with 100 % (Complete BUSCOs: 98 %, Fragmented BUSCOs: 2 %) of orthologs of the public database ''eukaryota_odb9'' present in the assembly.

The results of the Principal Component Analysis (PCA) for nuclear and mitochondrial SNPs data are shown in Figure 2.3. The graphic representation of the nuclear and mitochondrial PCA includes the assignment of mitotypes obtained by PCR with haplotype-specific primers to each individual plotted. Both the nuclear and mitochondrial SNP composition clearly separate the South American animals (h4h4) from the Antarctic ones along the PC1 which explains 58.6 % and 88.6 % of the variance, respectively. In contrast, nuclear SNPs of Antarctic animals were not congruent with the mitotype grouping, and instead three groups composed of animals of the four Antarctic mitotypes could be discerned. These groups separated by low percentage of variance (PC1: 5.6 %, PC2: 4.8 %) partially matched the sampled sites within Potter Cove indicating close kinship due to proximity. The PCA of the mitochondrial data resulted in two clusters, one containing animals of the mitotype h1h1 and h1h3 and the other animals h2h2 and h2h3 with no signal reflecting the three sampling sites in Potter Cove. Removing South American animals h4h4 from the analysis resulted in a clearer differentiation of Antarctic mitotypes h1h3 and h2h3 vs h1h1 and h2h2 (PC2).

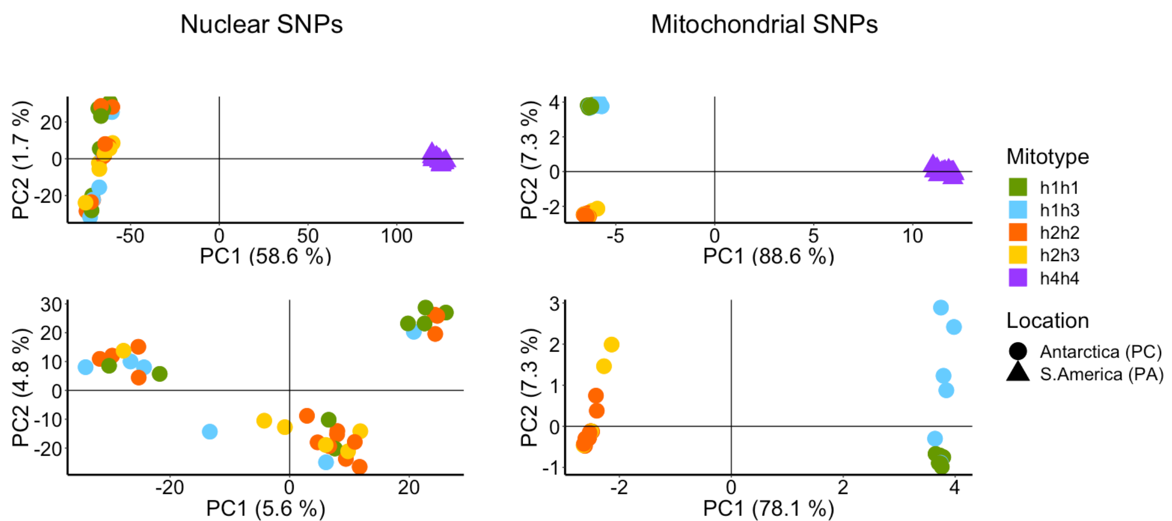


Figure 2.3. Principal component analysis based on 108,801 nuclear and 392 mitochondrial SNPs of Antarctic (circles) and South American (triangles) populations of *Aequiyoldia cf. eightsii.* In panels below, only Antarctic specimens were included in the analysis. Nuclear and mitochondrial SNPs identify the samples from South America as a distinct group. Mitochondrial SNPs identify two groups among Antarctic samples characterized by haplotypes h1 or h2, respectively, either in homoplasmic condition or in heteroplasmic condition together with h3. Nuclear SNPs sort Antarctic samples into three groups that are not congruent with the mitochondrial condition of the samples. Mitochondrial genotypes (Mitotypes) of the organisms are indicated with colours both in PCA based on nuclear and mitochondrial SNPs.

AMOVA results confirmed the clear patterns observed in the PCAs. For mitochondrial SNPs, all three levels of variance partitioning were highly significant ($p < 0.001$), with most of the variation occurring between locations (South America and Antarctica, 89.6 %), and a minor proportion between mitotypes within locations (7.8 %) and among individuals (2.6 %). AMOVA performed on nuclear SNPs resulted in significant differences between locations (67.7 %) and among individuals (31.7 %) but not between mitotypes within locations. This information is summarized in Table 2.1.

Table 2.1. Coefficients of the Analysis of Molecular Variance (AMOVA) of nuclear (Nuc) and mitochondrial (Mit) SNPs considering three strata (location, mitotype and individual)

Source of variation	df		Sum of squares		Percentage of variation		*p* value	
	Nuc.	Mit.	Nuc.	Mit.	Nuc.	Mit.	Nuc.	Mit.
Between location	1	1	4.33	10.33	67.73	89.58	0.001	0.001
Between mitotype, within location	3	3	0.09	0.30	0.57	7.82	0.442	0.001
Within individual	49	49	0.08	0.012	31.70	2.60	0.001	0.001

df: degree of freedom

Both nuclear and mitochondrial pairwise genetic differentiation estimates (G_{ST}') showed a high density of highly differentiated SNPs (G_{ST}'= 1; i.e*.,* present in all animals of one group and in none of the other) between animals from the two continents (Fig. 2.4a, b) supporting AMOVA, PCAs and *COI* results. An important difference between nuclear and mitochondrial G_{ST}' estimations emerged in the form of high density of nuclear SNPs with values close to zero in comparison to the mitochondrial data. Nuclear and mitochondrial pairwise G_{ST}' estimates between Antarctic mitotypes are shown in Figure 2.4 (c, d). Nuclear pairwise G_{ST}' comparisons resulted in similar density distributions with most of the values <0.2 and means ranging from 0.05 to 0.09 indicating low genetic differentiation. Instead, every mitochondrial G_{ST}' estimate comparing animals carrying haplotypes h1 and animals with h2 resulted in a high density of highly differentiated SNPs (G_{ST}' = 1). G_{ST}' estimations comparing h1h1 vs h1h3 and h2h2 vs. h2h3 showed low genetic differentiation (Fig. 2.4d).

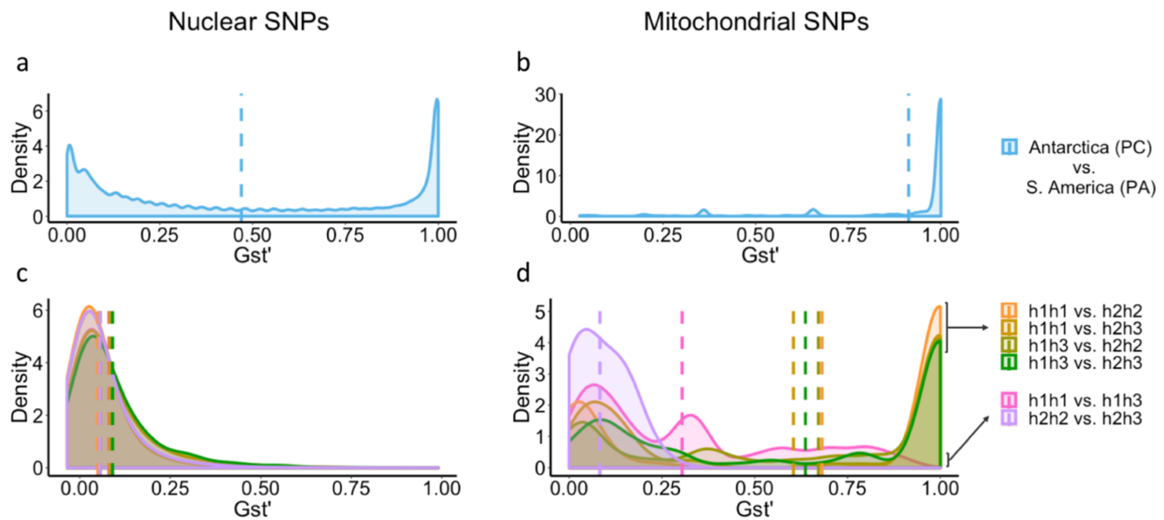


Figure 2.4. Density distributions of G_{ST}' values estimated by pairwise comparison between South American and Antarctic populations of *Aequiyoldia cf. eightsii* based on nuclear and mitochondrial SNPs (a, b), and between Antarctic mitotypes based on nuclear and mitochondrial SNPs (c, d). Dashed lines represent the mean G_{ST}' of the correspondingly coloured distribution. Note that for better visualization, comparisons between continents do not involve comparisons within Antarctic mitotypes, and that comparisons between mitotypes do not include the mitotype from South America (h4h4). A combined graphic representation would illustrate a bimodal distribution for both data sets (mitochondrial and nuclear SNPs); with most of the values close to zero for every comparison within Antarctic mitotypes and a high proportion of values close to one for every comparison between Antarctic and South American groups.

4. DISCUSSION

This study calls into question the reliability of a taxonomic identification system based on a few nuclear or mitochondrial markers such as the widely used *cytochrome c oxidase subunit I* (*COI*). We present a case study in a non-model marine bivalve, in which the standard barcoding approach using so-called universal primers is systematically biased by the occurrence of several distinct mitochondrial genomes in the population, including heteroplasmy, which in the absence of independent data (e.g., nuclear SNPs) would lead to an erroneous interpretation of the number of species and their distribution patterns (Fig. 2.5). Since this systematic bias is not caused by a paucity of data, it is not averted

by increasing the amount of data either but might instead converge, as in our case, on the wrong solution with high confidence. This will be especially common in non-model species for which genome features are mostly unknown, and in which inferences and indiscriminate generalization from model organisms are commonly used procedures (Milani and Ghiselli 2020).

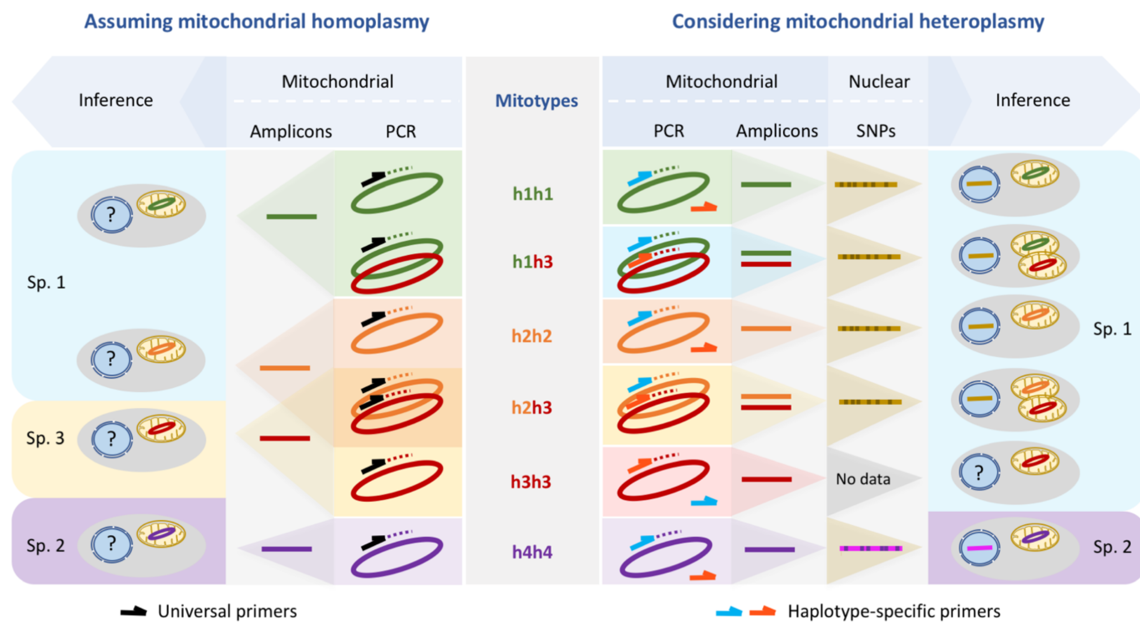


Figure 2.5. The amplification bias of universal primers and a standard barcoding analysis assuming strictly clonal inheritance of mitochondria leads to an overestimate of inferred species numbers (left half) in comparison to haplotype-specific primers and the inclusion of nuclear SNPs data (right half). Note that amplifying h2h3 individuals using universal primers leads to the amplification of either h2 or h3, but never both. Although no nuclear SNP data are available for the single h3h3 individual, we include it in species 1 here on the grounds of h3 mitochondria being associated with nuclear SNP patterns that characterise species 1 (see h1h3 and h2h3).

4.1 Incongruence of methods

4.1.1 Comparison across the Drake Passage

Both mitochondrial and nuclear SNPs analysis showed significant differences between Antarctic and South American bivalves confirming the clear pattern observed in PCAs and in GST' estimations. The bimodal distribution of pairwise differences in both data sources demonstrate that South American and Antarctic populations are genetically much more similar within the limits of either continental waters but highly distinct when compared between them. Viewed in conjunction with marked differences in size (see Appendix 2.3) the most likely conclusion is that populations on either side of the Drake Passage are genetically isolated and represent distinct species. Considering that the *locus typicus* (New South Shetland) for *A. eightsii* (Jay, 1839) is close to Potter Cove in our sampling, it is reasonable to assume that all our Southern Ocean material belongs to *Aequiyoldia eightsii* (Jay, 1839) *sensu strictu* whilst our South American samples (haplotype h4) belong to an as yet undescribed species that has been confused with it. A full taxonomical revision including the formal description of the new species of *Aequiyoldia* is beyond the scope of this paper, however.

4.1.2 Comparison within the Southern Ocean

Whilst the extrapolation from one mitochondrial marker gene to the assumed congruent differentiation of the nuclear genome that the standard barcoding procedure implies seems justified in the comparison across the Drake Passage (Fig. 2.4), the same is not true for emerging patterns of differentiation in the Antarctic.

In the Southern Ocean, the most notable patterns of genetic differentiation in *A. cf. eightsii* are strongly dependent on the source of data used. Although a differentiation pattern of a magnitude similar to the species-level trans-Drake comparison is present also within the Southern Ocean in the mitochondrial *COI* sequence barcoding data (Fig. 2.1; haplotypes [h1+h2] vs h3), the nuclear SNP data show no indication of such differentiation but show a unimodal distribution that is characteristic of a single genetic entity (Fig. 2.4c) for all specimens regardless of their mitochondrial haplotypes (h1, h2,

h3). Furthermore, mitochondrial *COI* barcoding sequences (derived from a 600bp fragment of the mitochondrial *COI* gene) and the mitochondrial SNPs (derived from the 18 kb representing the almost complete mitochondrial genome) suggest a mutually incompatible differentiation pattern. The *COI* barcode sequences suggest a structure [h1+h2] vs h3 whereas the mitochondrial SNPs the most apparent division is orthogonal to the *COI* sequence data: [h1h1+h1h3] vs [h2h2+h2h3], possibly h3h3 by itself; thus, individuals are most strongly characterized by possessing either h1 or h2 with varying admixed proportions of h3 to either of them (Fig. 2.6). This finding is not only different but incongruent and cannot easily be reconciled by assuming a different degree of taxonomic resolution in the various data sources.

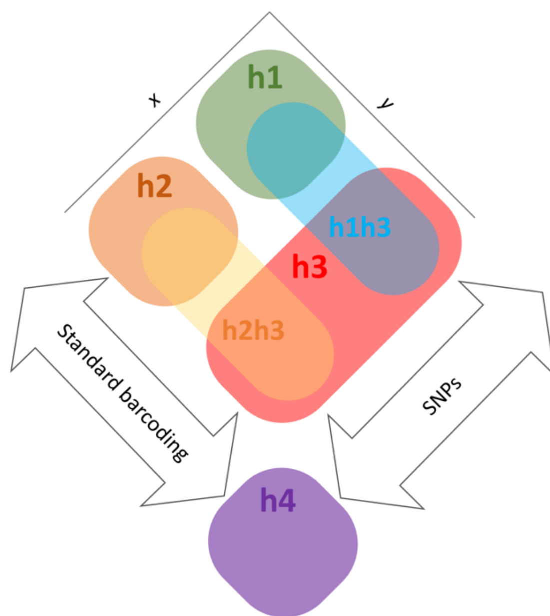


Figure 2.6. All approaches support a clear distinction between South American and Antarctic bivalves (axis *x* and *y*), but within the Southern Ocean mitochondrial *COI* sequence barcoding suggests a structure [h1+h2] vs h3 (axis *y*), whereas SNP data support an alternative structure: [h1h1+h1h3] vs [h2h2+h2h3] (axis *x*).

4.1.3 Taxonomic consequences

The classical molecular barcoding approach based on amplification and sequencing of a fragment of the mitochondrial *COI* gene turned out to be a good first indicator of major genetic diversity in the trans-Drake comparison. Instead, the heteroplasmic mitochondrial condition observed in some Antarctic specimens significantly interferes

with its taxonomic classification when addressed by traditional molecular approaches. Considering the magnitude of the differences found by González-Wevar et al. (2019) based on *COI*, it seems likely that they, too, had identified the division between h3 and [h1+h2] and hence suggested these two lineages as separate species, a conclusion that is not supported by our more complete data set (González-Wevar et al. (2019) molecular data are not publicly available for comparison, however). Our mitochondrial and nuclear SNPs dataset and the coexistence of the haplotype h3 with either the haplotype h1 or h2 in some organisms rule out cryptic speciation between these two groups as an explanation for the observed patterns.

The widespread occurrence of mitochondrial heteroplasmy linked to Doubly Uniparental Inheritance (DUI) in bivalves and the evidence of mitochondrial heteroplasmy in a wide variety of taxa from crustaceans to vertebrates (Doublet et al. 2008; Vollmer et al. 2011) highlight the importance of mitochondrial heteroplasmy as a general, previously underestimated shortcoming of molecular barcoding methodology. There is a consensus around the need of multiple loci (mitochondrial and nuclear) for a correct taxonomic classification, and the use of advanced sequencing techniques is supposed to gain space in molecular systematics to amply covering this need. However, the promise of this new sequencing era has not yet fully materialized, especially not for non-model species, demonstrating a growing taxonomic bias in favour of a small minority of genetic model species (David et al. 2019). DNA barcoding is still the dominant method in non-model species and will probably remain so for a while. It is beyond the purpose of this study to explore how widespread the occurrence of mitochondrial heteroplasmy is across taxa in the animal kingdom but A*. cf. eightsii* serves as a model case that if heteroplasmy occurs it can lead to wrong taxonomic inferences with an inflated sense of confidence.

4.2 True process: high mitochondrial genetic diversity and heteroplasmy

Although our data strongly question the validity of the extrapolation from a fragment of *COI* to the entire (nuclear) genome, the apparent division of *COI* sequences into three

almost equidistant groups [h1+h2], h3 and h4 is likely the result of a separate evolutionary history of the three in isolation of one another. The mitotype h4 is separated by the Drake Passage and the Polar Front from the nearest relatives in the Southern Ocean, a condition that remains so until today. However, divergence time estimations suggest that the haplotype h4 split off around the end of the Miocene *ca.* 8.5 Ma (González-Wevar et al. 2019) thus indicating a dispersal event across the already open Drake Passage between the two continents. The differentiation of the three *Aequiyoldia* clades characterized by major haplotype groups in Antarctica likely happened in the Pliocene, but unlike the South American relatives all three of them occur at least in partial sympatry today. The most likely scenario for the divergent patterns found in mitochondrial and nuclear data today is that the differentiated mitochondrial lineages persisted even after the groups made secondary contact again as consequence of their independent non-recombining, clonal inheritance. The pattern of differentiation in the nuclear genome on the other hand eroded away over time as a result of recombination of the hybridizing lineages, possibly comprising a case of speciation reversal as suggested for other Antarctic invertebrates (Dietz et al. 2015).

The persistence of distinct mitochondrial lineages in the Southern Ocean bivalve *A. cf. eightsii* over millions of years argues in favour of a selective advantage for each of the existing mitochondrial alleles either in space or over time. Several examples of genetic diversification caused by temporary isolation, some of them involving cryptic speciation, were described for other Southern Ocean invertebrates such as crustaceans (Held 2003; Baird et al. 2011), polychaetes (Neal et al. 2014), echinoderms (Hemery et al. 2012), among others.

As of now it remains unclear whether the observed heteroplasmy in *A. cf. eightsii* is the result of doubly uniparental inheritance or whether a different pattern of inheritance is responsible. The bivalves could not be sexed because the gonads could not be separated from all other tissue types in the preserved samples. However, the fact that heteroplasmy was also observed in samples comprising only somatic tissue (mantle), where paternally inherited mitochondria are not normally found in DUI species (but see

Ghiselli et al. (2019)), the possibility of an alternative mechanism of mitochondrial genetic inheritance cannot be ruled out entirely.

4.3 Lessons learned and countermeasures

Population genetics, molecular barcoding and molecular systematics critically depend on extrapolations made from small subsets to larger scales. Our results show that even when the markers are determined with high precision, the confidence in the correctness of the extrapolation from a *COI* fragment to the entire genome as implied in molecular barcoding may be unjustified when mitochondrial heteroplasmy is involved. In our study, an amplification bias with universal primers in the presence of two competing templates available for amplification in a heteroplasmic scenario resulted in the amplification of either of the two existing variants (Fig. 2.5), promoting erroneous inferences with high confidence. While such a volatile amplification bias with unpredictable outcome may seem to require several conditions and hence may be assumed to be rare, it is possibly often overlooked because evidence for it is often eliminated in a quality control step ensuring clean sequences preceding the analysis.

Countermeasures against this include an active search for double peaks in chromatograms in the Sanger sequencing results, the use of haplotype-specific primers whenever there is evidence of sequence competition, and the use of alternative sequencing methods in combination with an active search for segregating sites.

ACKNOWLEDGMENTS

The authors thank Uwe John for advice on experimental design, Andrea Eschbach and Enrique Celemin for their technical support, and Andrea Desiderato for valuable comments in early stages. We also thank to Instituto Antartico Chileno (INACH), Instituto Antartico Argentino (IAA) and Instituto Antartico Uruguayo (IAU) for the logistic support, especially to Marcelo Gonzalez (INACH), Dolores Deregibus (IAA) and Alvaro Soutullo (IAU); and Leyla Cardenas from Centro IDEAL. This work was supported by National

Agency of Investigation and Innovation (ANII, Uruguay) and German Academic Exchange Service (DAAD, Germany) by PhD scholarship to M. Martínez (Grant number POS_EXT_2015_122792).

DATA ACCESSIBILITY

COI amplified with universal and haplotype-specific primers, histone H3 and 18S sequences are available in GenBank (NCBI) under the accession numbers MT176797 - MT176951, MT645995 - MT646032, MT647927 - MT647998 and MT642948 - MT643019, respectively. Raw Illumina reads were deposited in the European nucleotide Archive database (EMBL-EBI) with the accessions ERR4265443 and ERR4276392 – ERR4276460 under the study accession ERP122389. Assembly and SNP data will be available in public repositories by the time of publication.

AUTHOR CONTRIBUTIONS

All the authors were involved in the design of the research. MM and CH performed fieldwork. MM and LH analysed the data and all the authors were involved in data interpretation and writing of the paper.

REFERENCES

Baird HP, Miller KJ, Stark JS. 2011. Evidence of hidden biodiversity, ongoing speciation and diverse patterns of genetic structure in giant Antarctic amphipods. Mol Ecol. 20(16):3439-3454.

Bentzen P, Leggett W, Brown G. 1988. Length and restriction site heteroplasmy in the mitochondrial DNA of american shad (*Alosa sapidissima*). Genetics. 118(3):509-518.

Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsch G, Pütz J, Middendorf M, Stadler PF. 2013. MITOS: improved de novo metazoan mitochondrial genome annotation. Mol Phylogen Evol. 69(2):313-319.

Berthier K, Chapuis M-P, Moosavi SM, Tohidi-Esfahani D, Sword GA. 2011. Nuclear insertions and heteroplasmy of mitochondrial DNA as two sources of intra-individual genomic variation in grasshoppers. Syst Entomol. 36(2):285-299.

Boyle EE, Etter RJ. 2012. Heteroplasmy in a deep-sea protobranch bivalve suggests an ancient origin of doubly uniparental inheritance of mitochondria in Bivalvia. Mar Biol. 160(2):413-422.

Bushnell B. 2014. BBMap. BBMap short read aligner, and other bioinformatic tools.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST. 2011. The variant call format and VCFtools. Bioinformatics. 27(15):2156-2158.

David KT, Wilson AE, Halanych KM. 2019. Sequencing Disparity in the Genomic Era. Mol Biol Evol. 36(8):1624-1627.

De Wit P, Pespeni MH, Palumbi SR. 2015. SNP genotyping and population genomics from expressed sequences - current advances and future possibilities. Mol Ecol. 24(10):2310-2323.

Degletagne C, Abele D, Held C. 2016. A Distinct Mitochondrial Genome with DUI-Like Inheritance in the Ocean Quahog Arctica islandica. Mol Biol Evol. 33(2):375-383.

DeSalle R, Goldstein P. 2019. Review and Interpretation of Trends in DNA Barcoding. Front. Ecol. Evol. 7:302.

Dietz L, Arango CP, Domel JS, Halanych KM, Harder AM, Held C, Mahon AR, Mayer C, Melzer RR, Rouse GW et al. 2015. Regional differentiation and extensive hybridization between mitochondrial clades of the Southern Ocean giant sea spider *Colossendeis megalonyx*. R Soc Open Sci. 2(7):140424.

Doublet V, Souty-Grosset C, Bouchon D, Cordaux R, Marcade I. 2008. A thirty million year-old inherited heteroplasmy. PLoS One. 3(8):e2938.

Doucet-Beaupré H, Breton S, Chapman EG, Blier PU, Bogan AE, Stewart DT, Hoeh WR. 2010. Mitochondrial phylogenomics of the Bivalvia (Mollusca): searching for the origin and mitogenomic correlates of doubly uniparental inheritance of mtDNA. BMC Evol Biol. 10(1):50.

Dray S, Dufour A-B. 2007. The ade4 package: implementing the duality diagram for ecologists. J. Stat. Softw. 22(4):1-20.

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(2):479-491.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol Mar Biol Biotechnol. 3(5):294-299.

Francoso E, Zuntini AR, Ricardo PC, Silva JPN, Brito R, Oldroyd BP, Arias MC. 2019. Conserved numts mask a highly divergent mitochondrial-COI gene in a species complex of Australian stingless bees *Tetragonula* (Hymenoptera: Apidae). Mitochondrial DNA A. 30(7):806-817.

Garcia G, Rios N, Gutierrez V, Serra S, Loureiro M. 2019. Transcriptome-Based SNP Discovery and Validation in the Hybrid Zone of the Neotropical Annual Fish Genus *Austrolebias*. Genes. 10(10).

Ghiselli F, Maurizii MG, Reunov A, Arino-Bassols H, Cifaldi C, Pecci A, Alexandrova Y, Bettini S, Passamonti M, Franceschini V et al. 2019. Natural Heteroplasmy and Mitochondrial Inheritance in Bivalve Molluscs. Integr Comp Biol. 59(4):1016-1032.

González-Wevar C, Díaz A, Gerard K, Cañete JI, Poulin E. 2012. Divergence time estimations and contrasting patterns of genetic diversity between Antarctic and southern South America benthic invertebrates. Rev Chil Hist Nat. 85(4):445-456.

González-Wevar C, Gérard K, Rosenfeld S, Saucède T, Naretto J, Díaz A, Morley SA, Brickle P, Poulin E. 2019. Cryptic speciation in Southern Ocean *Aequiyoldia eightsii* (Jay, 1839): Mio-Pliocene trans-Drake Passage separation and diversification. Prog Oceanogr. 174:44-54.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol. 29(7):644.

Gusman A, Lecomte S, Stewart DT, Passamonti M, Breton S. 2016. Pursuing the quest for better understanding the taxonomic distribution of the system of doubly uniparental inheritance of mtDNA. PeerJ. 4:e2760.

Hebert PD, Cywinska A, Ball SL, deWaard JR. 2003. Biological identifications through DNA barcodes. Proc Biol Sci. 270(1512):313-321.

Held C. 2003. Molecular evidence for cryptic speciation within the widespread Antarctic crustacean *Ceratoserolis trilobitoides* (Crustacea, Isopoda). in: Antarctic biology in a global context. 135-139.

Hemery L, Eléaume M, Roussel V, Améziane N, Gallut C, Steinke D, Cruaud C, Couloux A, Wilson N. 2012. Comprehensive sampling reveals circumpolarity and sympatry in seven mitochondrial lineages of the Southern Ocean crinoid species *Promachocrinus kerguelensis* (Echinodermata). Mol Ecol. 21(10):2502-2518.

Jombart T, Ahmed I. 2011. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics. 27(21):3070-3071.

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.

Kim SJ, Lee KY, Ju SJ. 2013. Nuclear mitochondrial pseudogenes in *Austinograea alayseae* hydrothermal vent crabs (Crustacea: Bythograeidae): effects on DNA barcoding. Mol Ecol Resour. 13(5):781-787.

Knaus BJ, Grünwald NJ. 2017. vcfr: a package to manipulate and visualize variant call format data in R. Mol Ecol Resour. 17(1):44-53.

Kopylova E, Noé L, Touzet H. 2012. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics. 28(24):3211-3217.

Kvist L, Martens J, Nazarenko AA, Orell M. 2003. Paternal leakage of mitochondrial DNA in the great tit (*Parus major*). Mol Biol Evol. 20(2):243-247.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods. 9(4):357.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The sequence alignment/map format and SAMtools. Bioinformatics. 25(16):2078-2079.

Low TY, Mohtar MA, Ang MY, Jamal R. 2019. Connecting Proteomics to Next-Generation Sequencing: Proteogenomics and Its Current Applications in Biology. Proteomics. 19(10):e1800235.

Mastrantonio V, Latrofa MS, Porretta D, Lia RP, Parisi A, Iatta R, Dantas-Torres F, Otranto D, Urbanelli S. 2019. Paternal leakage and mtDNA heteroplasmy in *Rhipicephalus spp*. ticks. Sci Rep. 9:1460.

Milani L, Ghiselli F. 2020. Faraway, so close. The comparative method and the potential of non-model animals in mitochondrial research. Philos Trans R Soc Lond B Biol Sci. 375(1790):20190186.

Neal L, Wiklund H, Muir AI, Linse K, Glover AG. 2014. The identity of juvenile Polynoidae (Annelida) in the Southern Ocean revealed by DNA taxonomy, with notes on the status of *Herdmanella gracilis* Ehlers *sensu* Augener. Mem Mus Vic. 71:203-216.

O'Leary SJ, Puritz JB, Willis SC, Hollenbeck CM, Portnoy DS. 2018. These aren’t the loci you’e looking for: Principles of effective SNP filtering for molecular ecologists. Mol Ecol. 27(16):3193-3206.

Paradis E. 2010. pegas: an R package for population genetics with an integrated–modular approach. Bioinformatics. 26(3):419-420.

Ratnasingham S, Hebert PD. 2007. BOLD: The Barcode of Life Data System (http://www. barcodinglife. org). Mol Ecol Notes. 7(3):355-364.

RCore Team. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Reed AJ, Morris JP, Linse K, Thatje S. 2014. Reproductive morphology of the deep-sea protobranch bivalves *Yoldiella ecaudata*, *Yoldiella sabrina*, and *Yoldiella valettei* (Yoldiidae) from the Southern Ocean. Polar Biol. 37(10):1383-1392.

Rogier O, Chateigner A, Amanzougarene S, Lesage-Descauses M-C, Balzergue S, Brunaud V, Caius J, Soubigou-Taconnat L, Jorge V, Segura V. 2018. Accuracy of RNAseq based SNP discovery and genotyping in *Populus nigra*. BMC Genomics. 19(1):1-12.

Rollins LA, Woolnough AP, Fanson BG, Cummins ML, Crowley TM, Wilton AN, Sinclair R, Butler A, Sherwin WB. 2016. Selection on Mitochondrial Variants Occurs between and within Individuals in an Expanding Invasion. Mol Biol Evol. 33(4):995-1007.

Sharma PP, Zardus JD, Boyle EE, Gonzalez VL, Jennings RM, McIntyre E, Wheeler WC, Etter RJ, Giribet G. 2013. Into the deep: a phylogenetic approach to the bivalve subclass Protobranchia. Mol Phylogenet Evol. 69(1):188-204.

Steinke D, Prosser SW, Hebert PD. 2016. DNA barcoding of marine metazoans. Marine genomics. In: Bourlat S. (eds) Marine Genomics. Methods in Molecular Biology, vol 1452. Humana Press, New York, NY.

Terranova MS, Lo Brutto S, Arculeo M, Mitton JB. 2007. A mitochondrial phylogeography of *Brachidontes variabilis* (Bivalvia: Mytilidae) reveals three cryptic species. J Zool Syst Evol Res. 45(4):289-298.

Verity R, Nichols RA. 2014. What is genetic differentiation, and how should we measure it--GST, D, neither or both? Mol Ecol. 23(17):4216-4225.

Vollmer NL, Viricel A, Wilcox L, Katherine Moore M, Rosel PE. 2011. The occurrence of mtDNA heteroplasmy in multiple cetacean species. Curr Genet. 57(2):115-131.

Yan J, Song Z, Xu Q, Kang L, Zhu C, Xing S, Liu W, Greimler J, Zust T, Li J et al. 2017. Population transcriptomic characterization of the genetic and expression variation of a candidate progenitor of *Miscanthus* energy crops. Mol Ecol. 26(21):5911-5922.

Zouros E, Rodakis GC. 2019. Doubly Uniparental Inheritance of mtDNA: An Unappreciated Defiance of a General Rule. Adv Anat Embryol Cell Biol. 231:25-49.

[dataset] Martinez, M., Harms, L., Abele, D., & Held, C.; 2020; Metazoan Mitochondrial COX1 / Cytochrome c oxidase Subunit I Aequiyoldia eightsii; GenBank (NCBI); MT176797 - MT176951

[dataset] Martinez, M., Harms, L., Abele, D., & Held, C.; 2020; Metazoan Mitochondrial COX1 / Aequiyoldia eightsii COX1 haplotype-specific primers; GenBank (NCBI); MT645995 - MT646032

[dataset] Martinez, M., Harms, L., Abele, D., & Held, C.; 2020; Eukaryotic Nuclear rRNA/ITS / Aequiyoldia eightsii 18S ribosomal RNA, partial sequence; GenBank (NCBI); MT642948 - MT643019

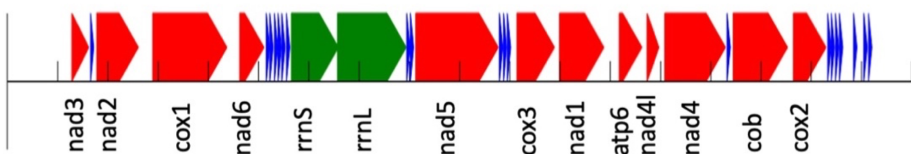
[dataset] Martinez, M., Harms, L., Abele, D., & Held, C.; 2020; Histone H3 Aequiyoldia eightsii, partial sequence; GenBank (NCBI); MT647927 - MT647998

[dataset] Martinez, M., Harms, L., Abele, D., & Held, C.; 2020; Mitochondrial heteroplasmy as a systematic bias in molecular species delimitation and barcoding (ERP122389); European nucleotide Archive database (EMBL-EBI); ERR4265443, ERR4276392 – ERR4276460

APPENDIXES

Appendix 2.1. Annotation of 18 kbp transcript (TRINITY_DN0_c0_g1_i6) using MITOS (Bernt et al. 2013). Above: The table provides the name of the mitochondrial gene annotated, the start and stop position, the direction of the strand and the length of the gene. Below: Mapping of genes on 18kb transcript (red: protein coding gene, green: rRNA gene, blue: tRNA gene)

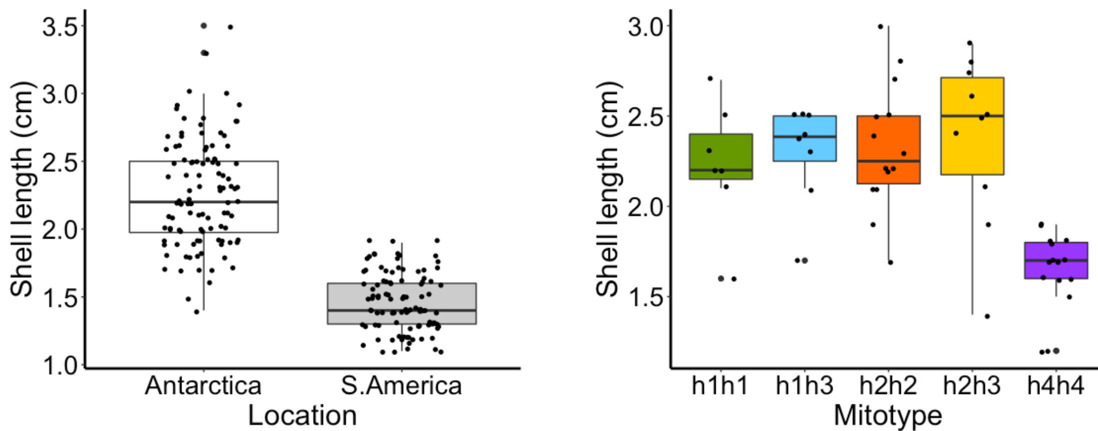
Name	Start	Stop	Strand	Length
nad3	1283	1615	+	333
trnG(gga)	1654	1716	+	63
nad2	1777	2604	+	828
cox1	2890	4368	+	1479
nad6	4625	5107	+	483
trnK(aaa)	5155	5222	+	68
trnL1(cta)	5226	5291	+	66
trnE(gaa)	5317	5385	+	69
trnY(tac)	5390	5454	+	65
trnF(ttc)	5463	5528	+	66
trnH(cac)	5555	5623	+	69
rrnS	5653	6587	+	935
rrnL	6571	7935	+	1365
trnM(atg)	7946	8014	+	69
trnP(cca)	8020	8085	+	66
nad5	8126	9769	+	1644
trnM(atg)	9794	9861	+	68
trnS1(aga)	9877	9937	+	61
trnT(aca)	9951	10019	+	69
cox3	10142	10888	+	747
nad1	10990	11871	+	882
atp6	12176	12625	+	450
nad4l	12731	12964	+	234
nad4	13082	14290	+	1209
trnN(aac)	14322	14389	+	68
cob	14445	15527	+	1083
cox2	15646	16293	+	648
trnQ(caa)	16328	16396	+	69
trnS2(tca)	16401	16463	+	63
trnW(tga)	16482	16550	+	69
trnA(gca)	16562	16629	+	68
trnL2(tta)	16842	16907	+	66
trnR(cga)	17045	17114	+	70
trnV(gta)	17139	17206	+	68



Appendix 2.2. p-distances based on COI between the most frequent haplotypes of *Aequiyoldia cf. eightsii* found in Punta Arenas, Potter Cove and Hangar Cove.

	h1	h2	h3	h4
h1	*			
h2	0.8	*		
h3	6.1	5.9	*	
h4	7.0	6.6	8.6	*

Appendix 2.3 Box plot of Shell length of *Aequiyoldia cf. eightsii* based on locations (Antarctica and South America) (left) and mitotypes (right). Box plot based on location includes all the animals collected and measured, while the boxplot based on mitotypes includes only those animals analyzed with haplotype-specific primers. The thick line indicates the median, the box represents the first and third quartile, and whiskers the 95 % confidence interval



CHAPTER 3

3 Publication II

Mariano Martínez, Lars Harms, Doris Abele, Christoph Held

Mitochondrial heteroplasmy and microhabitat act as main modulators of in situ gene expression in the shallow-water bivalve *Aequiyoldia eightsii* in an Antarctic fjord

Manuscript status: Submitted to *BMC Ecology and Evolution*

Mitochondrial heteroplasmy and microhabitat act as main modulators of in situ gene expression in the shallow-water bivalve *Aequiyoldia eightsii* in an Antarctic fjord

Mariano Martínez*[1], Lars Harms[2, 3], Doris Abele[1], Christoph Held[1]

1) Functional Ecology, Helmholtz Centre for Polar and Marine Research, Alfred Wegener Institute, Am Handelshafen 12, 27570 Bremerhaven, Germany

2) Helmholtz Centre for Polar and Marine Research, Alfred Wegener Institute, Am Handelshafen 12, 27570 Bremerhaven, Germany

3) Helmholtz Institute for Functional Marine Biodiversity at the University of Oldenburg (HIFMB), Ammerländer Herrstrasse 231, 26129 Oldenburg, Germany

*Corresponding author: Mariano Martínez. Address: Am Handelshafen 12, 27570 Bremerhaven; E-mail: mariano.martinez@awi.de

ABSTRACT

Differential regulation of gene expression is a pivotal adaptive mechanism of organisms facing environmental variation, especially when limitations in time and space preclude the evolution of genetically fixed, well-adapted phenotypes. We studied the in situ gene expression of a population of the shallow-water bivalve *Aequiyoldia cf. eightsii* in Potter Cove (King George Island/Isla 25 de Mayo, Antarctica) occupying different habitats in front of a melting glacier on a local scale (1km). We found the expression of nuclear genes was not only determined by 1) variation of the nuclear genome itself (nuclear SNPs) but equally strongly by 2) different environmental conditions characterizing the three locations and 3) the composition of the mitochondrial genotype (mitochondrial SNPs). Mitotype pattern was not reflected in nuclear SNPs, but either of them showed an independent, significant influence on gene expression. Mitochondrial SNPs divided Antarctic animals into two groups, each composed of organisms featuring mitochondrial homoplasmy and heteroplasmy (in total: four mitotypes). We validated our results by contrasting the observed magnitudes of differentially expressed genes (DEG) with magnitudes expected by stochasticity in randomized group comparisons. Habitat comparison revealed further differences in DEG at local scale in Potter Cove suggesting a high adaptive potential to the specific microenvironments. Interestingly, differential expression analysis between mitotypes resulted in a higher number of DEG than the driven by the environment, distinguishing heteroplasmic from homoplasmic organisms. Our results suggest that the mitochondrial genome and heteroplasmy may play a role in the regulation of nuclear gene expression and play a role in adaptation.

Key words: Antarctica, heteroplasmy, phenotypic plasticity, Protobranchia, transcriptomics, Single Nucleotide Polymorphisms.

1. INTRODUCTION

Organisms respond to environmental change via phenotypic plasticity, i.e., the expression of different phenotypes from the same genome, or through genetically-based changes caused by natural selection (adaptive evolution) (Fox et al. 2019). While plasticity manifests itself within the lifetime of a single organism, adaptation occurs on the level of populations and on longer timescales (multiple generations). In the context of rapid environmental change, plasticity has been proposed as a key mechanism for the persistence of organisms in changing environments, acting through 'buying time' for adaptation (Chevin et al. 2010). However, there are not many in situ marine studies looking into how phenotypic plasticity is actually realised, which genetic mechanisms are involved and what the most important determinants are that decide which of the many possible phenotypes actually becomes realized (Donelson et al. 2019); these questions are mostly addressed through experimental approaches (Morel-Journel et al. 2020).

Aequiyoldia eightsii (Jay, 1839) is a common marine bivalve in shallow waters along the West Antarctic Peninsula (WAP), one of the fastest warming regions on Earth (Vaughan et al. 2003; Stammerjohn et al. 2008). Occurring in shallow coastal waters, this species is exposed to rapidly changing conditions driven by fast increase of both annual and seasonal air temperature, upper ocean temperature (Meredith and King 2005; Schloss et al. 2012), accelerated retreat velocity of land and tidewater glaciers, and progressive melting of the West Antarctic ice sheet. Glacial and sea ice melting, in turn, result in reduced and more strongly fluctuating salinity, enhanced turbidity, and the alteration of nutrient and carbon biogeochemistry in coastal areas (Henley et al. 2019). Its frequent occurrence in coastal sedimentary habitats renders this species an excellent model to investigate adaptive responses of Antarctic benthic fauna to the effects of global warming and associated environmental changes. In fact, several studies have evaluated population-specific physiological and behavioral responses to warming and iceberg scouring in this species (Peck and Bullough 1993; Abele et al. 2001; Peck et al. 2004; Clark et al. 2016)

While distribution records of this nominal species extend to both sides of the Drake Passage in southern South America (SSA) and the WAP, recent studies based on mitochondrial and nuclear data suggest cryptic speciation with a pair of two sister species inhabiting South American and Antarctic waters, respectively (González-Wevar et al. 2019, Chapter II) For the sake of simplicity, we will be referring to the Antarctic material as *A. eightsii* on the grounds of this our material being collected near the *locus typicus* of the original species description (South Shetland Islands).

The recent discoveries of the genetic characteristics of this nominal species are not only relevant at inter-continental or population level, but also within the populations down to the level of the individual. In a previous study we reported several highly differentiated mitochondrial genomes coexisting in Antarctic populations but also coexisting in single individuals (heteroplasmy) in a subset of the sampled individuals (Chapter 2). The two major haplotypes in the population (h1 and h2; p-dist. 0.8 %) were mostly observed in homoplasmy (i.e*.,* all mitochondria in one individual being of the same type), but occasionally in combination with a highly divergent haplotype h3 (p-dist. 6 %) in heteroplasmic individuals h1h3 and h2h3. The occurrence of the divergent haplotype in heteroplasmy with either of the two major haplotypes, and the absence of an accompanying pattern at nuclear level rule out the possibility of cryptic speciation as had recently been suggested by González-Wevar et al. (2019). Whether the persistence of distinct mitochondria in the Southern Ocean bivalve *A. eightsii* over millions of years is the result of a selective advantage for each of the existing mitotypes either in space or over time is still an open question. Even more intriguing is the prevalence of heteroplasmic individuals in the population, since this condition is commonly reported as unfavourable and affecting mito-nuclear interactions (Stewart and Chinnery 2015; Ghiselli et al. 2019).

Mitochondrial heteroplasmy has been reported in over 100 bivalve species related to an evolutionarily stable system known as Doubly Uniparental Inheritance (DUI) (Gusman et al. 2016); which implies females inheriting mitochondria from their mothers (F-type) and males from both paternal lines (F-type and M-type). In males, the F-type is mostly present in somatic tissues whereas the M-type is dominant in the gonads but can also

be present in small amounts in the somatic tissue (Zouros and Rodakis 2019). Despite its wide distribution across bivalve species and its long evolutionary history, the functional advantages of this heteroplasmic system, if any, and its evolutionary significance are still unknown (Ghiselli et al. 2019).

Transcriptomic sequencing (RNA-seq) has become useful for identifying biological functions or processes as it uncovers patterns that can only emerge from simultaneously studying a large set of genes. Thus, gene expression patterns can reflect biological states or phenotypes responding to specific environmental conditions, as well as the genotype of individuals, allowing unprecedented mechanistic insight into plasticity and adaptation (King et al. 2018). However, the normally low proportion of annotated genes in non-model species, and the high number of false positives in differential expression analyses (Rajkumar et al. 2015) indicates the need for a careful interpretation of the emerging patterns.

We conducted this study to understand the influence of the genotype background on the patterns of gene transcription in *A. eightsii*, comparing animals across three different and characteristic small-scale habitats in front of a melting glacier in Potter Cove (WAP). The expectation was that the heteroplasmy-homoplasmy condition of the mitochondria affects nuclear gene expression to a yet unknown extent, within different environmental niches. We further assumed that the small-scale habitat variability in Potter Cove also causes adaptive modifications of gene transcription in the bivalves, potentially causing a diversification of genotype-controlled gene expression patterns.

To test our predictions, we explored gene expression patterns and Single Nucleotide Polymorphisms (SNPs) in response to 1) contrasting environmental conditions at local scales (ca. 1 km) considering genetic structuring of the population, as well as in relation to 2) nuclear and mitochondrial genetic diversity. We validated our results by contrasting the observed magnitudes of differentially expressed genes (DEG) with magnitudes expected by randomized group comparisons.

2. MATERIALS AND METHODS

2.1 Study area

Field work was carried out at Potter Cove (PC), a small fjord on the southern coast of King George (South Shetland Islands, Antarctic Peninsula, Fig. 3.1). Potter Cove is about 4 km long and 2.5 km wide opening into the bigger Maxwell Bay. The northern and the eastern coast are surrounded by the Fourcade Glacier, whereas the southern coast has flat gravel and sandy beaches, intersected by meltwater streams that drain the Warzawa Ice Field and its proglacial lakes (Wölfl et al. 2016). During the summer melt season, the inner cove receives freshwater and ice input directly through the fronting line of Fourcade glacier and from two bigger meltwater streams. A major fraction (approximately 50%) of eroded subglacial and terrigenous sediment particles discharging with the meltwater are deposited within the narrow and relatively shallow parts of the inner cove (maximum depth of 50 m) close to the glacier front (Wölfl et al. 2014; Monien et al. 2017). Sedimentation is supported by a prolonged residence time and internal circulation of surface water in the inner cove before the water exits across the internal moraine along the southern coast into Maxwell Bay (Lim et al. 2013).

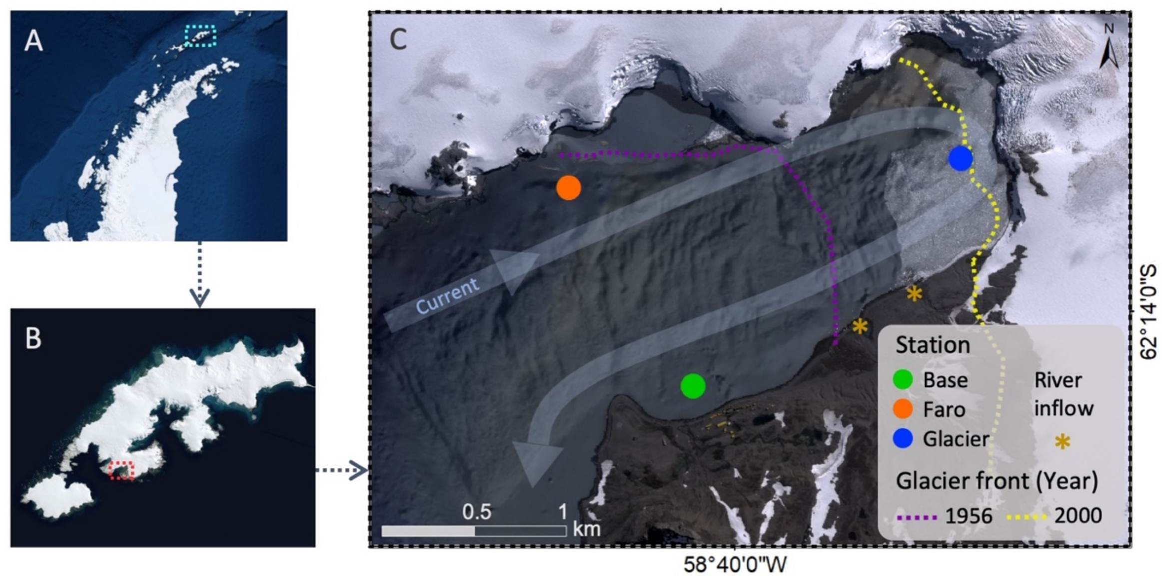


Figure 3.1. Study area: West Antarctic Peninsula (A), King George Island (B) and Potter Cove (C). Sampled stations Base, Glacier and Faro are indicated with colours. Dashed lines indicate the Glacier front position over time (modified from Rückamp et al. (2011)). The curved arrow shows the direction of the water circulation in the Cove, and orange stars indicate meltwater streams inflow in the Cove. Maps obtained from DigitalGlobe (2013)

The three stations sampled in the present study (Base: 62°14'11"S - 58°40'14"W; Glacier: 62°13'32"S - 58°38'31"W; Faro: 62°13'35"S - 58°40'58"W; Fig. 3.1) were selected to compare reference sites of contrasting glacial influence within the inner cove, with confirmed presence of *Aequiyoldia eightsii* in the benthic communities (Pasotti et al. 2015). The three stations also represent distinct benthic sub-habitats defined by Jerosch et al. (2018) based on cluster analysis of 42 environmental variables. The station "Faro" is situated on the northern coastline and characterized by high bed shear stress and wave action of the inflowing water masses (Lim et al. 2013) with large grain size and little fine sediment deposition (Wölfl et al. 2014). In contrast, the station "Glacier" receives fine glacial and meltwater stream deposits and is under a strong influence of glacial calving and ice-scouring (Sahade et al. 1998), both leading to frequent resuspension and re-deposition of the fine sediments. The sediment at station "Base" is also composed of fine-grained, muddy material, but has a much higher content of

organic matter from decaying macroalgal debris. Furthermore "Base" is situated downstream of Carlini station which may also explain part of the organic matter enrichment in the area. "Base" has the longest ice-free history (> 50 years), whereas "Glacier" is the most recently ice-free area, uncovered by glacial ice since the early 2000s.

2.2 Sampling

A. eightsii collection and environmental characterization of stations in Potter Cove were performed in the austral summer season in January and February 2018. Sampling of bivalves and sediment coring, as well as CTD profiling down to the maximum depth of stations were carried out simultaneously at each station. Sampling dates of stations were approximately one week apart.

2.2.1 Environmental characterization

At each station, salinity, temperature and turbidity profiles were performed with a Sea-Bird CTD (SBE19plusV2, Sea-Bird Electronics, Bellevue, WA, USA). Sediment cores were taken by SCUBA divers using cylindrical Plexiglas corer (height: 50 cm, diameter: 8 cm; 3 replicates per station). The open corers were pushed up to 30 cm into the sediment, sealed with a rubber tap and carefully pulled out with the overlying bottom water to preserve the first layer of the sediment undisturbed. Cores were returned to the laboratory in an upright position for further analysis. In the laboratory, the overlying water was removed and the sediment sliced into 1 cm thick slices. The sediment-water slurry of each slice was mechanically homogenized and separated into two subsamples for grain-size analysis and determination of total inorganic carbon (TIC), total organic carbon (TOC), total nitrogen (TN) and total sulphur (TS). All samples were frozen at - 80 °C for 24 hours. Afterwards, the samples were freeze dried for 96 hours for transportation.

2.2.2 Animal collection

A. eightsii were collected in the three stations using a Van Veen grab (n_{Base} = 84, $n_{Glacier}$ = 10, n_{Faro} = 10). Bivalves were transported to the local research facility in insulated containers with water and sediments from the sampling site. At the laboratory, bivalves were dissected on ice under a stereomicroscope after recording individual weight, shell length and width. Bivalves sampled for differential gene expression analysis were dissected within 2 hours after collection whereas animals analysed exclusively for Single Nucleotide Polymorphisms (SNPs) were kept alive in aquaria with water and sediment from the Cove at 1°C and dissected within 20 days after collection. For each individual, foot and mantle tissue samples were conserved separately in RNA later (SIGMA).

2.3 Laboratory and data analysis

2.3.1 Sedimentary parameters

For the grain size determination, freeze-dried samples were dry sieved through a 2 mm strainer and weighed in order to separate gravel from the smaller grain size fractions. Afterwards, sediment aggregation < 2 mm was suspended overnight in demineralised water and wet sieved through a 0.063 mm strainer to separate the sand fraction from the silt and clay fraction. The sand fraction was subsequently dried at 40 °C for one week and weighed to the nearest 1 mg. Silt and clay fractions were placed with water in a 5 L-jar and stood for one week for sedimentation. Subsequently, the overlying water was decanted, and the samples stored at - 80 °C for 24 hours and freeze dried again for 96 hours. Finally, the dry silt-clay-samples were weighed on a fine scale to the nearest 1 mg. With the weights obtained for each grain size fraction, the percentages of gravel and sand, and silt and clay were calculated for each station.

For chemical analyses, samples were pulverized using a sediment mill, and the determination of the total amounts of carbon (TC), nitrogen (TN) and sulphur (TS) was carried out using a gas-phase chromatograph vario EL III (Elementar, Germany). Organic and inorganic carbon (TOC and TIC, respectively) contents were determined by combustion using a Carbon-Sulfur Analyzer CS-2000 (Eltra, Germany). In addition, TOC -

sulphur and nitrogen ratios (C: S, C: N) were calculated and used as proxies of sediment oxygenation conditions (Togunwa and Abdullah 2017) and organic matter origin (marine vs. terrestrial) (Escobar-Briones and García-Villalobos 2009). Based on TIC content, the amount of calcium carbonate ($CaCO_3$) was calculated according to Wölfl et al. (2016).

As *A. eightsii* is normally found within the 2-cm sediment depth (Davenport 1988), results of the sedimentary parameters for each station are presented as the average of the first two layers of sediment (0 -1 cm and 1 -2 cm). For a better visualization of the environmental characteristics of the stations, a Principal Component Analysis Biplot was performed (Fig. 3.2).

2.3.2 RNA extraction, library preparation and de novo assembly

Differential gene expression and Single Nucleotide Polymorphisms (SNPs) on RNA-data analyses were performed using an assembly created from 70 libraries (one library per individual) of *A. eightsii* from South America (Punta Arenas, Chile) and Antarctica (Potter Cove) by Martinez et al (Chapter 2). In the present study, a subset of 31 libraries from Potter Cove (of the total included in the assembly) were included in the transcriptomic analysis (n_{Base} = 21, $n_{Glacier}$ = 5, n_{Faro} = 5), all of them analysed for nuclear and mitochondrial SNPs on RNA-seq data, and a subset of 15 individual samples were analysed for differential gene expression (n_{Base} = 5, $n_{Glacier}$ = 5, n_{Faro} = 5).

RNA was extracted from mantle tissue (5 – 30 mg) by homogenizing the samples in Trizol reagent (SIGMA) using a Precellys homogenizer (Precellys24, Bertin Technologies, France). Total RNA was isolated from each sample using the Direct-zol™ RNA MiniPrep Kit (ZYMO Research Corp., USA) according to the manufacturer's instructions. Libraries were prepared using the Illumina TruSeq® Stranded mRNA Sample Preparation Kit starting from 1 µg of total RNA following the protocol provided by the company. During libraries preparation, each sample was tagged through adapters ligation. Each cDNA library was diluted to 0.8 nM, and subsequently all the libraries were pooled and cleaned using magnetic beads to remove the remaining primer content. Final cDNA concentration was measured in the LabChip® GX Touch (PerkinElmer, USA). The pool of

samples was sequenced on an Illumina NextSeq 500 sequencer using the NextSeq High Output Kit v2 (150 cycles) with a paired-end protocol.

Raw reads were quality controlled by FastQC v. 0.11.7 (Babraham Institute, Cambridge, UK), cleaned and *de novo* assembled using the Trinity genome-independent transcriptome assembler v2.8.4; (Grabherr et al. 2011). For details on *de novo* assembly parameters see Chapter 2.

2.3.3 Single Nucleotide Polymorphisms analysis

The SNPs analysis involved the alignment of quality-filtered paired-end reads to the *de novo* transcriptome using bowtie2 v2.3.4.1 (Langmead and Salzberg 2012). Alignments in SAM format (Sequence Alignment Map) were compressed and indexed with SAMTOOLS v1.8 (Li et al. 2009). Genotype likelihoods were computed using mpileup from SAMTOOLS and variant calling was performed with the BCFTOOLS. In a first filtering step, variants with a Phred quality score below 30 were excluded. A total of 927,305 SNPs in 31 individuals were retained. Subsequently, an iterative filtering strategy between loci and individuals with a progressive increase of cut-off values (O'Leary et al. 2018) was performed using VCFTOOLS v0.1.16 (Danecek et al. 2011). A first filtering for loci quality was performed keeping variants successfully genotyped in 50 % of individuals (max-missing 0.5) and with a minor allele count (MAC) of 3. Afterwards, individuals with more than 27 % of missing data were removed from the analysis. A second filtering for loci quality consisted in a max-missing of 0.95, a Minor Allele Frequency (MAF) of 0.05 and minimum number of reads (minDP) of 10. Additionally, variants with more than two alleles were discarded (max-alleles 2). After filtering, 120,404 in 27 individuals (= libraries) were kept. The retained variants were retrospectively identified as nuclear and mitochondrial SNPs through blasting the sequences containing SNPs against a mitochondrial database obtained from the UniProt Swiss-Prot database. Mitochondrial SNPs included in the analysis were those contained in a single long transcript *ca* 18 kbp, representing most of the mitochondrial genome that had positives hits with the Uniprot mitochondrial database (see Chapter 2, Appendix 2.1); while nuclear SNPs comprised all

SNPs in sequences with no hits in the mitochondrial database. Thus, the analyses were performed with a total of 145 and 116123 mitochondrial and nuclear SNPs, respectively.

To assess genetic differentiation between stations, Analysis of Molecular Variance (AMOVA), Principal Component Analyses (PCA), and pairwise G_{ST}' estimations were carried out separately for nuclear and mitochondrial SNPs with all the individuals that remained after the filtering. AMOVA of mitochondrial and nuclear SNPs was conducted considering two strata, station (Base, Glacier and Faro) and individuals, and significance was tested by randomly permutating (n = 1000) the sample matrices following Excoffier et al. (1992) (Table 3.1). In both PCAs based in mitochondrial and nuclear SNPs, every individual was labeled based on the sampling station and mitochondrial mitotype. Mitotype of individuals were determined in Chapter 2, through the amplification of a 714 bp *Cytochrome c oxidase subunit I* fragment using haplotype-specific primers. Based on that, in Chapter 2 bivalves were grouped in six different genetic clusters: animals carrying only one haplotype (h1, h2, h3, h4), named h1h1, h2h2, h3h3 and h4h4 respectively, and heteroplasmic animals for haplotype h1 or h2 together with the highly divergent haplotype h3 (p-dist. 6 % respect to h1 and h2), called h1h3 and h2h3, respectively. The current study analyzed a subset of animals from Potter Cove only carrying mitotypes h1h1, h1h2, h1h3 and h2h3.

All the analyses were performed in R v3.6 (RCore 2016) using the packages vcfR (Knaus and Grünwald 2017), adegenet (Jombart and Ahmed 2011), poppr v2.8.3 (Kamvar et al. 2014), ade4 (Dray and Dufour 2007) and pegas (Paradis 2010).

2.3.4 Differential gene expression and Gene Ontology analyses

The differential gene expression analysis (DEA) involved the alignment of the short reads of each sample (n = 15) separately against the de novo reference transcriptome using Bowtie2 v 3.3.4.1 (Langmead et al. 2009). The DEA implied pairwise comparisons between clusters of individuals grouped by stations (n_{Base} = 5, $n_{Glacier}$ = 5, n_{Faro} = 5), but also grouped based on their mitotype (n_{h1h1} = 2, n_{h2h2} = 7, n_{h1h3} = 3, n_{h2h3} = 3) and nuclear

genotype (n_{NucA} = 10, n_{NucB} = 5, see section 3.3), using the same set of 15 individual samples; called from here original grouping.

To assess the presence of random or stochastic effects on gene expression, if any, we performed DEAs (n = 1000) using randomized groups of the same 15 individual samples following the same numerical conformation of the original grouping (i.e., stations: n = 5, n = 5, n = 5; mitotype: n = 2, n = 7, n = 3, n = 3; nuclear genotype: n = 10, n = 5). Gamma distribution functions were selected based on Akaike information criterion (AIC) and fitted to each of the three data sets of differential expressed genes (DEG) values generated by randomized grouping, and the quantiles 0.95 ($Q_{0.95}$) and confidence intervals for the $Q_{0.95}$ were calculated following Kulkarni and Powar (2010) and using the R package EnvStats. With this, we tested if the observed values of DEG generated by the original groupings were above the $Q_{0.95}$ of the gamma distribution of DEG values obtained from DEAs (n = 1000) with randomized groups of individuals.

For every DEA, relative abundances of transcripts were estimated by RSEM v1.2.26 (Li and Dewey 2011) and DEG was assessed using a test based on the negative binomial distribution as integrated in the Bioconductor R package DESeq2 (Love et al. 2014), with a standard level of $p \leq 0.001$ and a fold change of at least 2 supporting significance. The tools were executed using the Trinity package v2.8.4. The annotation of DEG was performed using the Trinotate functional annotation suite v3.1.1 (Grabherr et al. 2011) including a homology search against the UniProt Swiss-Prot database and assigning Gene Ontology (GO) terms to annotated transcripts. GO enrichment analyses were carried out using GOseq (Young et al. 2010) by combining the results of DEG between each cluster and its corresponding GO annotation, and using the full list of DEG with GO terms as the background.

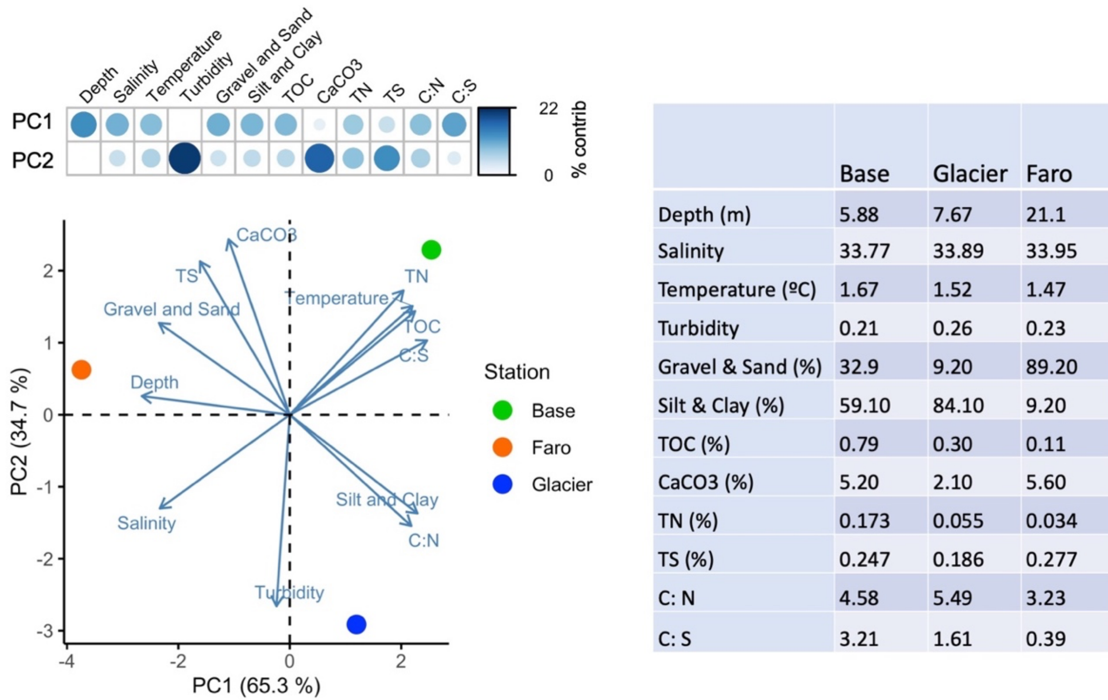
3. RESULTS

3.1. Environmental conditions

The environmental characteristics captured in the water column and sedimentary parameters indicate that the stations Base and Glacier are mostly under the influence

of sediment deposition from subglacial erosion and meltwater streams whereas station Faro is receiving strong currents of inflowing water from the outer Potter Cove and Bransfield Strait. The Principal Component Analysis Biplot (Fig. 3.2) separated the stations Base and Glacier from Faro along the Principal Component axis 1 (PC1), while the axis PC2 mostly discriminated Base and Faro from Glacier. PC1 discriminated stations Base and Glacier from Faro based on higher bottom water temperature, and percentage of total organic carbon (TOC) and nitrogen (TN) in surface sediments at Base, and higher percentage of silt and clay at Glacier; whereas Faro is the deepest station with relatively higher bottom water salinity, and a bottom cover of mainly gravel and sand. Bottom water turbidity, calcium carbonate ($CaCO_3$) and total sulphur (TS) were the main factors represented by PC2 with the highest turbidity at Glacier and higher $CaCO_3$ and TS percentages at Base and Faro compared to Glacier. Bottom water parameters followed the general trend of the water column profiles among stations although with less pronounced differences (Appendix 3.1). The ratio between TOC and TS (C:S) increased from Faro to Glacier and to Base along PC1 with values suggesting, according to Togunwa and Abdullah (2017), near anoxic conditions in Faro and suboxic conditions in Base and Glacier. Instead, the ratio between TOC and TN (C: N) showed slight differences between stations with values within the range characteristic for marine ecosystem components (such as macroalgal debris, phytoplankton deposits and benthic microalgal mats)

Although *Aequiyoldia eightsii* density at each station was not explicitly determined, the higher sampling effort required at Faro for the collection of a lower individual number in comparison to Base and Glacier suggests a very limited habitat availability which was mostly confined to sediment patches within the macroalgal belt. Both mean and maximum shell length of our samples were higher at Base (2.28/3.3 cm) than Glacier (2.18/2.9 cm) and Faro (2.22/2.5 cm).



	Base	Glacier	Faro
Depth (m)	5.88	7.67	21.1
Salinity	33.77	33.89	33.95
Temperature (ºC)	1.67	1.52	1.47
Turbidity	0.21	0.26	0.23
Gravel & Sand (%)	32.9	9.20	89.20
Silt & Clay (%)	59.10	84.10	9.20
TOC (%)	0.79	0.30	0.11
CaCO3 (%)	5.20	2.10	5.60
TN (%)	0.173	0.055	0.034
TS (%)	0.247	0.186	0.277
C: N	4.58	5.49	3.23
C: S	3.21	1.61	0.39

Figure 3.2. Principal Component Analysis Biplot of stations and environmental variables (left), percentage of contribution of each variable to Principal components (top left), and values of each variable for every station (right). Sediment variables are reported as the average of the 0-2 cm layer values of the three Corer replicates and water variables as the deepest point recorded by CTD (Base: 5.88 m, Glacier: 7.67 m, Faro: 21.1 m).

3.2 Genetic structure

Analysis of Molecular Variance (AMOVA) based on mitochondrial SNPs did not show genetic structuring between stations, whereas AMOVA based on nuclear SNPs showed a slight but significant spatial genetic differentiation. In both cases most of the variance (97.84 and 96.53 %, respectively) was among individuals and only 2.16 % and 3.47 % between stations, respectively (Table 3.1).

Table 3.1. Coefficients of the Analysis of Molecular Variance (AMOVA) of mitochondrial (Mit) and nuclear (Nuc) SNPs between stations

Source of variation	df		Sum of squares		Percentage of variation		*p* value	
	Mit.	Nuc.	Mit.	Nuc.	Mit	Nuc.	Mit.	Nuc.
Between stations	2	2	0.41	0.31	2.16	3.47	0.31	0.001
Within individuals	24	24	4.23	2.99	97.84	96.53		

df: degree of freedom

The AMOVA results are in accordance with the Principal Component Analysis of mitochondrial and nuclear SNP data (Fig. 3.3). In Chapter 2, we show that mitochondrial SNPs in *A. eightsii* from Potter Cove fall into two main groups characterized by haplotypes h1 or h2, either in homoplasmic or in heteroplasmic condition with haplotype h3 (Fig. 3.3, left: h1h3, h2h3). This distinction of two major mitotype groups (**h1**h1or **h1**h3) vs. (**h2**h2 or **h2**h3) dominates the separation along PC1 (82.5 %) but no accompanying spatial pattern in Potter Cove is discernible. With the exception of h1h1 and h2h3, which we did not sample at station Faro, all mitotypes were found at every station sampled.

In contrast, the nuclear SNPs divided the animals into groups that are not congruent with the major mitotypes (Fig. 3.3, right). Instead, a weak but significant spatial pattern emerged, discriminating the animals of station Base from a group comprising animals from Faro and Glacier, but the variance explained by each component was low (PC1: 6.7%, PC2: 4.5 %). Pairwise G_{ST}' comparisons (Fig. 3.3, bottom half) also indicated a low level of genetic structuring between stations, with the highest mitochondrial genetic difference between Base and Glacier stations, and rather similar G_{ST}' density distributions at every nuclear pairwise comparison between stations.

Both groupings, the one based on mitochondrial (mitotypes) and the one based on nuclear SNPs (nuclear genotype: NucA: individuals from Base, NucB: individuals from Faro and Glacier) were considered for the gene expression analysis.

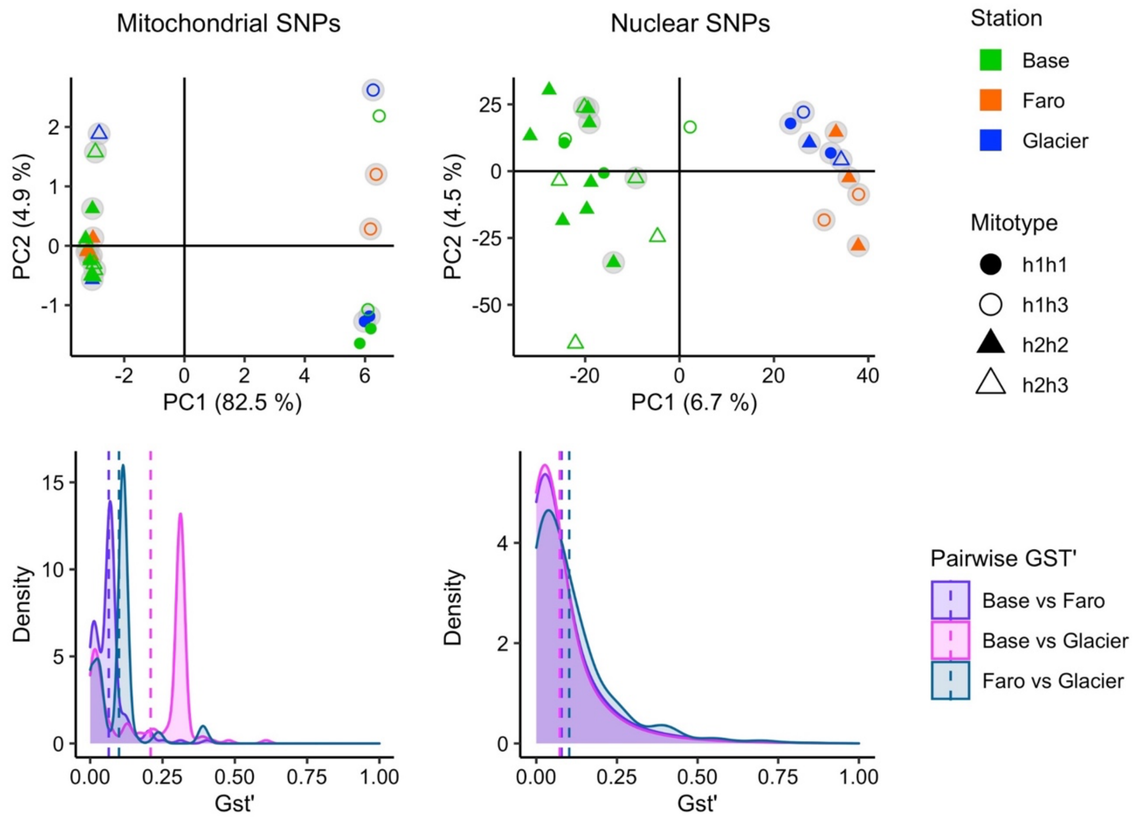


Figure 3.3. Principal component analysis based on mitochondrial and nuclear SNPs of bivalves *Aequiyoldia eightsii* collected in Base, Faro and Glacier stations in Potter Cove (panels above). Sampling stations are indicated with colours, mitotypes with shapes and grey circles highlight individuals that were included in the differential gene expression analysis. Mitochondrial SNPs (left) are dominated by the possession of mitochondrial haplotypes h1 or h2 (PC1) and to a lesser degree also by the amount of admixture of a third mitotype h3 (PC2). Nuclear SNPs (right) do not reflect the mitochondrial genotype of the samples but show a spatial pattern (station Base vs Glacier + Faro) instead. In the panels below, the density distributions (Kernel density estimate) of G_{ST}' values in pairwise comparison of animals from the three stations based on mitochondrial and nuclear SNPs are shown. Dashed lines represent the mean G_{ST}' of the correspondingly coloured distribution.

3.3 Gene expression

3.3.1 Differential gene expression

Differential gene expression analysis (DEA) considering clustering of individuals by station or by genotypes (mitotype and nuclear genotype) resulted in different numbers of differentially expressed genes (DEG), all well above the number of DEG expected by random individual grouping; i.e*.,* the observed numbers of DEG were above the quantile 0.95 ($Q_{0.95}$) of the gamma distribution of DEG values obtained from DEAs (n = 1000) with randomized groups of individuals (Fig. 3.4).

The pairwise comparison between animals grouped by stations resulted in 262 observed DEG, with the $Q_{0.95}$ of the adjusted gamma distribution of 154.5 expected by stochasticity alone (Fig. 3.4 bottom). The comparison between animals grouped by mitotype and nuclear genotype resulted in 806 and 96 DEG, and $Q_{0.95}$ of 524.1 and 17.6, respectively. The DEA based on stations showed Glacier and Faro clustered together with a relatively low number of genes differentially expressed between them (37 DEG), while more than three times as many genes were significantly up- or downregulated between station Base and either of the two others (139 and 121 DEG, respectively). The highest number of DEG in comparisons based on mitotype groups was observed between homoplasmic and heteroplasmic mitochondrial clusters (176 – 600 DEG), whereas the lowest DEG numbers resulted in the pairwise comparison of h1h1 vs h2h2 (18 DEG) and h1h3 vs h2h3 (41 DEG). Thus, clustering based on patterns of gene expression clearly separates bivalves with homoplasmic mitochondrial genome from heteroplasmic individuals.

In summary, this further corroborates our working hypothesis that there is a causal relationship between even small-scale differences in environmental conditions as well as genotypes on the one hand and patterns of gene expression on the other hand. However, these results also show that a significant number of loci made it onto the list of DEG likely as a consequence of random or stochastic effects.

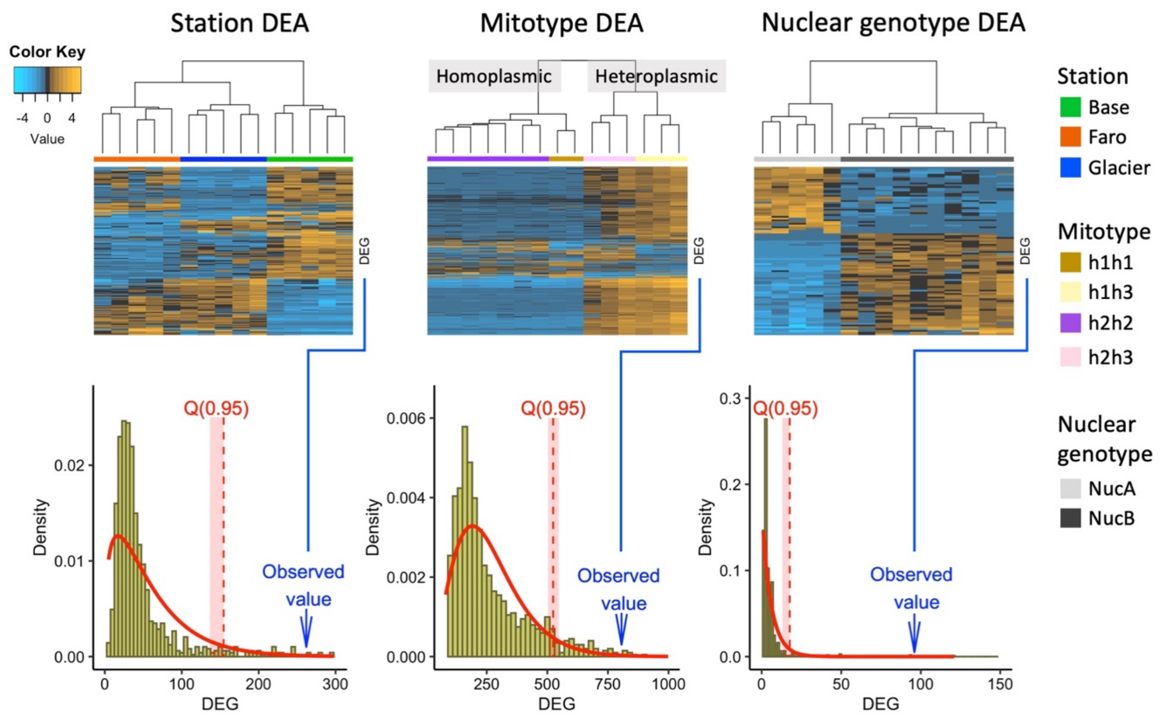


Figure 3.4. Above: heatmaps showing the differentially expressed genes (DEG) in the three differential expression analyses (DEA) considering individuals grouped by station, mitotype and nuclear genotype. For every heatmap, a cluster analysis showing similarities between groups is included on top. DEG are displayed in rows and counts per sample are displayed in columns. Genes included were significant ($p < 0.001$) in at least one pairwise comparison. A colour code indicates up-regulated and down-regulated expression levels (top left). Below: density distributions of DEG values obtained from DEAs (n= 1000) with randomized groups of individuals. The red curve corresponds to the adjusted gamma distribution and the dashed red line and area underlying indicate the Quantile 0.95 ($Q_{0.95}$) and Confidence interval of the $Q_{0.95}$. In every case, the observed DEG value was above the $Q_{0.95}$.

3.3.2 Annotation and GO enrichment analysis

Out of the total number of DEG obtained in the DEA grouped by station, mitotype and nuclear genotype, 15.1 %, 13.4 % and 13.3 % could be functionally annotated against the UniProt Swiss-Prot database. Only in the DEA based on mitotype, the GO enrichment analysis resulted in gene categories enriched over a background gene expression. Grouping samples according to their station or nuclear genotype did not result in

significant enrichment of GO terms. The enriched categories for each mitotype comparison are shown in Figure 3.5 and Appendix 3.2. In agreement with the DEA, the comparisons between individuals characterized by mitotypes h1h1 vs h2h2 and h1h3 vs h2h3 resulted in lower numbers of DEG per enriched GO term category (counts), while every comparison between homoplasmic and heteroplasmic individuals implied a significantly higher number of DEG per enriched category, with the exception of the h2h2 vs h1h3 comparison which did not result in any enriched category. The comparison between the homoplasmic groups (h1h1 vs h2h2) resulted almost exclusively in the enrichment of categories associated with mitochondrial respiration, i.e., energetic functions. However, in comparisons between heteroplasmic groups (h1h3 vs. h2h3) as well as between homoplasmic vs. heteroplasmic individuals, the pattern is less clear. A great diversity of categories involving cytosolic functions were enriched, some of which also involving mitochondrial respiratory functions.

In Appendix 3.3 we show a list of Biological Processes related to each GO term obtained from the DEA based on stations and nuclear genotype which lacks statistical support but summarizes possible differentially activated functions between groups. In these comparisons, genes associated with mitochondrial respiration do not predominate in the same manner as found for the mitotype comparisons. GO terms related to biological processes such us \`lactate oxidation´ and \`response to hypoxia´ were associated to genes differentially expressed in the DEA by stations in the comparison Base vs Faro and Glacier vs. Faro, respectively. These results seem plausible in the context of oxygenation in the upper sediment horizons, the burrowing zone of *A. eightsii,* but the lack of statistical significance precludes a meaningful analysis based on small-scale spatial or nuclear genotype at this point.

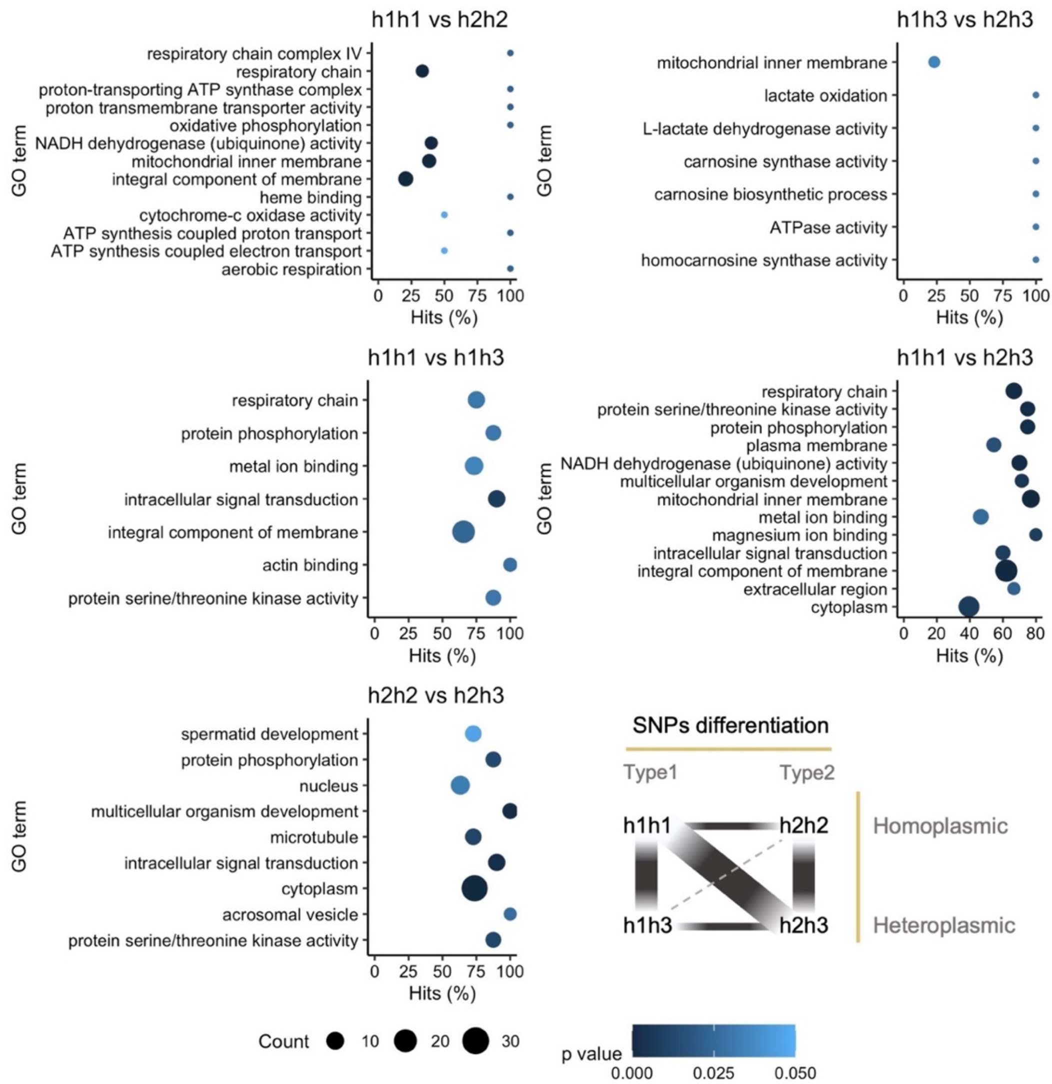


Fig 3.5. Enriched GO term categories between mitotypes. In the *x* axis it is shown the percentage of DEG found in a certain category over the total number of genes in the category (Hits). Size code indicates the number of DEG in the GO term category and a gradient of colour indicates the *p value* of the over represented category. For a better visualization, only the 13 most significant GO terms are shown (higher count and *p value*), as only the comparison h1h1 vs h2h3 involved a higher number (22 GO terms). The diagram at the bottom right summarises GO enrichment results. The thickness of the connectors among mitotypes represents the overall number of genes per enriched category (counts). With the exception of h2h2 vs h1h3 which did not result in any enriched GO terms (dashed line), all comparisons between homoplasmic and heteroplasmic groups resulted in a greater number of genes per category than comparisons between homoplasmic (h1h1 vs h2h2) and heteroplasmic (h1h3 vs h2h3) groups from the two major genetic groups based on SNPs.

4. DISCUSSION

The Antarctic shallow-water marine bivalve *Aequiyoldia eightsii* has a strikingly differentiated gene expression pattern under natural conditions in Potter Cove (King George Island), a pattern that is influenced by at least three independent underlying causes: microclimatic differences down to a kilometre scale (stations), mitotype (especially admixture of a second mitochondrial genotype per individual proved effective) and nuclear genotype. All three determinants of in situ differences in gene expression proved significant if viewed in isolation, even after correcting for spurious correlations due to stochastic effects.

4.1 Gene expression patterns

4.1.1 Microhabitat influence

Physiological responses of organisms are dependent on their perception of the environment. In the attempt to understand organism-environment interactions, we often superimpose our own human perceptions on other species and environment. This ultimately leads to a bias towards analysis on 'human scales' (spatial and temporal) and the overlooking of mechanisms underlying adaptation of species or populations to their closest surrounding environment. Consequently, there is vast empirical evidence showing regional or large scale-environmental heterogeneity promoting plastic responses in widely distributed species (Pöhlmann et al. 2011; Koenigstein et al. 2013; Rivas et al. 2018), but studies showing in situ responses of species to small-scale habitat heterogeneity are very scarce in the literature (Baythavong 2011; Unal et al. 2013).

The environmental characterization of the three sampled stations reflected the small-scale habitat variability previously reported in Potter Cove (Monien et al. 2017; Jerosch et al. 2018), which is shaped by a directional current pattern with characteristic input signatures from different terrigenous sources as well as the dynamic melting of the KGI ice sheet and the recession of Fourcade glacier fronting line (Wölfl et al. 2016). The influence of spatial gradients in Potter Cove on biological communities (composition, functional traits) has been previously addressed (Pasotti et al. 2012; Pasotti et al. 2015);

however, there are no antecedents of studies evaluating in situ gene expression responses for organisms in this area, and in other marine pro-glacial habitats, on such a fine spatial scale. Our spatial differential gene expression analysis revealed that organisms at each station express a distinct and characteristic set of nuclear-encoded genes regardless their genetic background (mitotype and nuclear genotype) enabling plastic responses to specific microenvironmental niches. It is expected that habitats where the spatial scale of environmental variation is less than the dispersal distance of a species favour the expression of adaptive plasticity (Baythavong 2011); this is clearly the case of *A. eightsii* which not only inhabits various microhabitats in Potter Cove but is also widely distributed along the Antarctic continent shelf. The ability of *A. eightsii* to precisely modulate gene transcription according to its intimate physico-chemical environment might confer resilience to the current rapid environmental change in many areas of the Antarctic peninsula.

4.1.2 Genotype influence (nuclear and mitochondrial)

Mitochondrial and nuclear genotypes both have their own influence on gene expression independently from one another. This is remarkable and unexpected as it implies that the composition of the mitochondrial genotype is correlated with the expression of genes that are for the most part located in the nuclear genome, and not merely the result of a co-correlation between nuclear and mitochondrial genotypes. The mechanism for this is unclear for now, but recently Pozzi et al. (2017) have detected the presence of small mitochondrial highly-transcribed RNAs (smithRNAs) in the bivalve *Ruditapes philippinarum*, suggesting RNA interference as a mechanistic explanation for nuclear gene expression regulation by mtDNA.

The peculiar mitochondrial genetic characteristics of *A. eightsii* at both the population and individual level (Chapter 2) raise the question whether the existence of distinct mitochondrial alleles, and their occasional occurrence in heteroplasmy, implies greater functional versatility and spatial adaptive advantages observable at gene transcription level. Even though sample size in our study did not permit nested comparisons (i.e.,

between genotypes/station), which would have allowed to identify potential traits exhibiting different reaction norms between genotypes in response to environment (G X E interactions) (Aubin-Horth and Renn 2009), the DEA by genotypes alone showed gene expression patterns influenced by mitochondrial (and to a lesser extent nuclear) genetic diversification. This transcriptional differentiation based on genotypes was mostly promoted by nuclear differentially expressed genes (DEG), not only in comparisons involving organisms differing in their nuclear genotype but also in comparisons based on mitotypes - from which emerged the most striking results regarding the relation between genetic background and gene expression. Interestingly, whereas SNP data support a structure dominated by the possession of mitochondrial haplotype h1 or h2: [**h1**h1+**h1**h3] vs [**h2**h2+**h2**h3], gene expression patterns were similar among all homoplasmic individuals (with either h1 or h2 mitotype) but were highly distinct from all heteroplasmic individuals (h1h3 and h2h3). The admixture of the mitochondrial h3 haplotype in addition to the major mitotypes h1 and h2, even in low dosage, leads to different gene expression patterns including genes encoded in the nuclear genome (Fig. 3.4).

The genetic and transcriptional diversity of *A. eightsii* might support the success of this species in a wide range of conditions throughout the Antarctic continent. The high diversification of its mitochondrial genome, which probably shaped key aspects of organelle function to meet environmental challenges (via selection), might be compensated by some degree of transcriptional plasticity within the nucleus which adds another layer of flexibility as a means for adaptation.

4.1.2 Randomness and uncontrolled factors quota

Genes are considered as differentially expressed when they simultaneously satisfy a *p value* and a fold change criterion. No matter how small the number of differentially expressed genes resulting from an analysis might be, their statistical significance is normally considered sufficient to support the biological interpretations emerging from them, or at least to refute the ‘no effects’ null hypothesis. We showed in an Antarctic

population of *A. eightsii*, however, that even applying rather strict *p* and fold change criteria and involving comparisons lacking biological/ecological meaning (randomized groups), the probability of a differential gene expression analyses (DEA) yielding null results (i.e., zero DEG) in natural populations is virtually zero.

The reason behind the unexpected numbers of DEG resulting from comparison with randomized groups, which in some cases were even higher than the results obtained with the original grouping (i.e., by station, mitotype and nuclear genotype) has several explanations. Even though DESeq2 is a conservative method with a high number of false negatives and a low number of false positives compared to other methods (Rajkumar et al. 2015), it is reasonable to expect that among ca. 300,000 assembled transcripts, some transcripts appear to be significantly over- or under-expressed merely by chance. The magnitude of this phenomenon likely increases as sample number per group decreases, as it was the case in our study. Additionally, and especially in those cases where the number of DEG obtained from comparisons with randomized groups resulted in numbers similar or higher than those obtained with the original groupings, this phenomenon is likely to be connected to the plethora of non-controlled factors/ conditions (e.g., sex, age, presence of parasites, etc.) triggering uniform gene expression responses across organisms sharing or experiencing them. This issue is applicable to any biological approach, but transcriptomic profiling makes it visible as it provides a broad account of cellular processes responding simultaneously to a wide range of stimuli. The possibility of interference of other factors to the observed responses as well as the presumable stochastic component in the DEAs is a warning sign with respect to analyzing DEG individually, and emphasizes the need for statistical post analysis considering DEG as a whole, such as enrichment analyses. While GO enrichment analysis are based on annotated genes which implies a loss of resolution due to the fact that only between 10 and 20% of DEG are annotated in non-model organisms, this loss of resolution should be random across all genes, i.e., any pattern resulting from it should be robust against the stochastic component of differential gene expression analyses.

4.2 Functional analysis

The enrichment of nuclear GO terms based on differences in mtDNA (mitotypes), highlights the tight coordination between nuclear and mitochondrial genomes. In model organisms, sequence variation within mtDNA affects patterns of gene expression on the same nuclear background across a wide variety of nuclear genes in mice (Dunham-Snary et al. 2018) and key mitochondrial protein-coding genes in *Drosophila* (Camus et al. 2015). In our study, mitochondrial functions were also among the most represented in the enrichment analysis, indicating that intra-population mtDNA variation implies a differential transcriptional regulation of nuclear-encoded mitochondrial complex subunits among mitotypes, in support of adequate mitochondrial functioning.

Remarkably, in most of the cases the comparisons between homoplasmic and heteroplasmic individuals implied the highest number of DEG per enriched category (Fig. 3.5), suggesting that not only the variation in mtDNA *per se* affects the expression of nuclear genes. Instead to an even greater extent, the heteroplasmic condition of the mitochondria seems to alter gene expression. Pozzi et al. (2017) identified distinct smithRNAs associated with each of the F and M mitochondrial types in the DUI species *R. philippinarum*, pointing them out as central in sex determination through the regulation of nuclear gene expression; but also leaving open the possibility of this mechanisms affecting other nuclear genes. Our results lend further support to the latter hypothesis and offer an explanation how the observed correlation of nuclear gene expression patterns and the composition of mitochondrial genomes might be implemented. The same mechanism might explain why comparisons between heteroplasmic and homoplasmic mitotypes in our study resulted in the enrichment of a high diversity of terms involving mitochondrial respiratory and cytosolic functions.

The differential transcriptional regulation based on differences in mtDNA (mitotype) could afford this species a higher adaptive potential to cope with spatial or temporal variability. Marine shallow-water sedimentary environments are in no way homogenous and instead present animals with an enormous variety of redox conditions in space and time (Abele et al. 2017). Cohorts of animals differing in mitochondrial respiratory efficiency might even confer population-wide plasticity of metabolic function that

apparently enables the animals to thrive in different redox environments of the upper sediment horizon. For animals of very reduced horizontal mobility in the adult form as *A. eightsii*, this may be a way to ensure better habitat exploitation and in so doing contribute to the resilience of the fjordic population. The ability of *A. eightsii* to respond to redox conditions was also suggested by the results from DEA based on stations (GO:0019516: lactate oxidation, GO:0001666: response to hypoxia), although as discussed in section *4.1.2*, the interpretations derived from single genes or GO terms, should be taken with caution and ideally corroborated under experimental conditions and GO enrichment analysis.

CONCLUSIONS

The conspicuous genetic diversity of *A. eightsii* in the Southern Ocean and in particular the existence of two mitochondrial haplotypes in heteroplasmy (Chapter 2), in combination with the reported small-scale habitat and genotype-based phenotypic diversification, denote considerable adaptive potential, not normally expected in Antarctic species (Peck et al. 2010). This genomic condition serves to add another layer of flexibility to react to changes in the environment and might turn out instrumental in the face of the ongoing rapid environmental change in Antarctic fjords. Ultimately, it may turn *A. eightsii* into a winner of environmental change (Griffiths et al. 2017), expanding into habitats still dominated by more susceptible filter feeders many of which are on the retreat already (Sahade et al. 2015).

ACKNOWLEDGMENTS

The authors thank Uwe John for advice on experimental design, Andrea Eschbach for technical support, Isabell Viehweger for the analysis of sediment samples and Camila Neder for georeferencing the stations on the map. We also thank to Instituto Antartico Chileno (INACH), Instituto Antartico Argentino (IAA) and Instituto Antartico Uruguayo (IAU) for the logistic support, especially to Marcelo Gonzalez (INACH), Dolores Deregibus (IAA) and Alvaro Soutullo (IAU). This work was supported by National Agency of

Investigation and Innovation (ANII, Uruguay) and German Academic Exchange Service (DAAD, Germany) by PhD scholarship to M. Martínez (Grant number POS_EXT_2015_122792), and by Zempelin-Stiftung - Deutsches Stiftungszentrum (Grant number T0214/33883/2019/sm).

DATA ACCESSIBILITY

Raw Illumina reads were deposited in the European nucleotide Archive database (EMBL-EBI) with the accessions ERR4265443 and ERR4276392 – ERR4276460 under the study accession 'Mitochondrial heteroplasmy as a systematic bias in molecular species delimitation and barcoding' (ERP122389). Assembly and SNP data will be available in public repositories by the time of publication.

AUTHOR CONTRIBUTIONS

All the authors were involved in the design of the research. MM and CH performed fieldwork. MM and LH analysed the data and all the authors were involved in data interpretation and writing of the paper.

REFERENCES

Abele D, Tesch C, Wencke P, Pörtner HO. 2001. How does oxidative stress relate to thermal tolerance in the Antarctic bivalve *Yoldia eightsi*? Antarct Sci. 13(2):111-118.

Abele D, Vazquez S, Buma AG, Hernandez E, Quiroga C, Held C, Frickenhaus S, Harms L, Lopez JL, Helmke E. 2017. Pelagic and benthic communities of the Antarctic ecosystem of Potter Cove: Genomics and ecological implications. Mar genomics. 33:1-11.

Aubin-Horth N, Renn SC. 2009. Genomic reaction norms: using integrative biology to understand molecular mechanisms of phenotypic plasticity. Mol Ecol. 18(18):3763-3780.

Baythavong BS. 2011. Linking the spatial scale of environmental variation and the evolution of phenotypic plasticity: selection favors adaptive plasticity in fine-grained environments. Am Nat. 178(1):75-87.

Camus MF, Wolf JB, Morrow EH, Dowling DK. 2015. Single nucleotides in the mtDNA sequence modify mitochondrial molecular function and are associated with sex-specific effects on fertility and aging. Curr Biol. 25(20):2717-2722.

Chevin L-M, Lande R, Mace GM. 2010. Adaptation, plasticity, and extinction in a changing environment: towards a predictive theory. PLoS Biol. 8(4):e1000357.

Clark MS, Sommer U, Sihra JK, Thorne MA, Morley SA, King M, Viant MR, Peck LS. 2016. Biodiversity in marine invertebrate responses to acute warming revealed by a comparative multi-omics approach. Glob Chang Biol. 23(1)318-330

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST. 2011. The variant call format and VCFtools. Bioinformatics. 27(15):2156-2158.

Davenport J. 1988. The feeding mechanism of *Yoldia* (═ *Aequiyoldia*) *eightsi* (Courthouy). Proc. Royal Soc. B. 232(1269):431-442.

Donelson JM, Sunday JM, Figueira WF, Gaitan-Espitia JD, Hobday AJ, Johnson CR, Leis JM, Ling SD, Marshall D, Pandolfi JM et al. 2019. Understanding interactions between plasticity, adaptation and range shifts in response to marine environmental change. Philos Trans R Soc Lond B Biol Sci. 374(1768):1-14.

Dray S, Dufour A-B. 2007. The ade4 package: implementing the duality diagram for ecologists. J. Stat. Softw. 22(4):1-20.

Dunham-Snary KJ, Sandel MW, Sammy MJ, Westbrook DG, Xiao R, McMonigle RJ, Ratcliffe WF, Penn A, Young ME, Ballinger SW. 2018. Mitochondrial - nuclear genetic interaction modulates whole body metabolism, adiposity and gene expression in vivo. EBioMedicine. 36:316-328.

Escobar-Briones E, García-Villalobos FJ. 2009. Distribution of total organic carbon and total nitrogen in deep-sea sediments from the southwestern Gulf of Mexico. Bol. Soc. Geol. Mex. 61(1):73-86.

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(2):479-491.

Fox RJ, Donelson JM, Schunter C, Ravasi T, Gaitan-Espitia JD. 2019. Beyond buying time: the role of plasticity in phenotypic adaptation to rapid environmental change. Philos Trans R Soc Lond B Biol Sci. 374(1768):20180174.

Ghiselli F, Maurizii MG, Reunov A, Arino-Bassols H, Cifaldi C, Pecci A, Alexandrova Y, Bettini S, Passamonti M, Franceschini V et al. 2019. Natural heteroplasmy and mitochondrial inheritance in bivalve molluscs. Integr Comp Biol. 59(4):1016-1032.

González-Wevar C, Gérard K, Rosenfeld S, Saucède T, Naretto J, Díaz A, Morley SA, Brickle P, Poulin E. 2019. Cryptic speciation in Southern Ocean *Aequiyoldia eightsii* (Jay, 1839): Mio-Pliocene trans-Drake Passage separation and diversification. Prog Oceanogr. 174:44-54.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol. 29(7):644.

Griffiths HJ, Meijers AJ, Bracegirdle TJ. 2017. More losers than winners in a century of future Southern Ocean seafloor warming. Nat. Clim. Change. 7(10):749-754.

Gusman A, Lecomte S, Stewart DT, Passamonti M, Breton S. 2016. Pursuing the quest for better understanding the taxonomic distribution of the system of doubly uniparental inheritance of mtDNA. PeerJ. 4:e2760.

Henley SF, Schofield OM, Hendry KR, Schloss IR, Steinberg DK, Moffat C, Peck LS, Costa DP, Bakker DCE, Hughes C et al. 2019. Variability and change in the west Antarctic Peninsula marine system: Research priorities and opportunities. Prog Oceanogr. 173:208-237.

Jerosch K, Pehlke H, Monien P, Scharf F, Weber L, Kuhn G, Braun MH, Abele D. 2018. Benthic meltwater fjord habitats formed by rapid glacier recession on King George Island, Antarctica. Philos Trans A Math Phys Eng Sci. 376(2122).

Jombart T, Ahmed I. 2011. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics. 27(21):3070-3071.

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.

King NG, McKeown NJ, Smale DA, Moore PJ. 2018. The importance of phenotypic plasticity and local adaptation in driving intraspecific variability in thermal niches of marine macrophytes. Ecography. 41(9):1469-1484.

Knaus BJ, Grünwald NJ. 2017. vcfr: a package to manipulate and visualize variant call format data in R. Mol Ecol Resour. 17(1):44-53.

Koenigstein S, Pohlmann K, Held C, Abele D. 2013. Ecological comparison of cellular stress responses among populations - normalizing RT-qPCR values to investigate differential environmental adaptations. BMC Ecol. 13:21.

Kulkarni H, Powar S. 2010. A new method for interval estimation of the mean of the Gamma distribution. Lifetime Data Anal. 16(3):431-447.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods. 9(4):357.

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10(3):R25.

Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics. 12(1):323.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The sequence alignment/map format and SAMtools. Bioinformatics. 25(16):2078-2079.

Lim CH, Lettmann K, Wolff J-O. 2013. Numerical study on wave dynamics and wave-induced bed erosion characteristics in Potter Cove, Antarctica. Ocean Dynamics. 63(11-12):1151-1174.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15(12):550.

Meredith MP, King JC. 2005. Rapid climate change in the ocean west of the Antarctic Peninsula during the second half of the 20th century. Geophys Res Lett. 32(19).

Monien D, Monien P, Brünjes R, Widmer T, Kappenberg A, Busso AAS, Schnetger B, Brumsack H-J. 2017. Meltwater as a source of potentially bioavailable iron to Antarctica waters. Antarct Sci. 29(3):277-291.

Morel-Journel T, Thuillier V, Pennekamp F, Laurent E, Legrand D, Chaine AS, Schtickzelle N, Marshall D. 2020. A multidimensional approach to the expression of phenotypic plasticity. Funct Ecol. 34(11).2338-2349

O'Leary SJ, Puritz JB, Willis SC, Hollenbeck CM, Portnoy DS. 2018. These aren't the loci you'e looking for: Principles of effective SNP filtering for molecular ecologists. Mol Ecol. 27(16):3193-3206.

Paradis E. 2010. pegas: an R package for population genetics with an integrated–modular approach. Bioinformatics. 26(3):419-420.

Pasotti F, De Troch M, Raes M, Vanreusel A. 2012. Feeding ecology of shallow water meiofauna: insights from a stable isotope tracer experiment in Potter Cove, King George Island, Antarctica. Polar Biol. 35(11):1629-1640.

Pasotti F, Manini E, Giovannelli D, Wölfl A-C, Monien D, Verleyen E, Braeckman U, Abele D, Vanreusel A. 2015. Antarctic shallow water benthos in an area of recent rapid glacier retreat. Mar Ecol. 36(3):716-733.

Peck L, Morley S, Clark M. 2010. Poor acclimation capacities in Antarctic marine ectotherms. Mar Biol. 157(9):2051-2059.

Peck LS, Ansell AD, Webb KE, Hepburn L, Burrows M. 2004. Movements and burrowing activity in the Antarctic bivalve molluscs *Laternula elliptica* and *Yoldia eightsi*. Polar Biol. 27(6):357-367.

Peck LS, Bullough L. 1993. Growth and population structure in the infaunal bivalve *Yoldia eightsi* in relation to iceberg activity at Signy Island, Antarctica. Mar Biol. 117(2):235-241.

Pöhlmann K, Koenigstein S, Alter K, Abele D, Held C. 2011. Heat-shock response and antioxidant defense during air exposure in Patagonian shallow-water limpets from different climatic habitats. Cell Stress and Chaperones. 16(6):621-632.

Pozzi A, Plazzi F, Milani L, Ghiselli F, Passamonti M. 2017. SmithRNAs: Could Mitochondria "Bend" Nuclear Regulation? Mol Biol Evol. 34(8):1960-1973.

Rajkumar AP, Qvist P, Lazarus R, Lescai F, Ju J, Nyegaard M, Mors O, Borglum AD, Li Q, Christensen JH. 2015. Experimental validation of methods for differential gene expression analysis and sample pooling in RNA-seq. BMC Genomics. 16:548.

RCore Team. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rivas MJ, Saura M, Perez-Figueroa A, Panova M, Johansson T, Andre C, Caballero A, Rolan-Alvarez E, Johannesson K, Quesada H. 2018. Population genomics of parallel evolution in gene expression and gene sequence during ecological adaptation. Sci Rep. 8(1):16147.

Rückamp M, Braun M, Suckro S, Blindow N. 2011. Observed glacial changes on the King George Island ice cap, Antarctica, in the last decade. Global Planet Change. 79(1-2):99-109.

Sahade R, Lagger C, Torre L, Momo F, Monien P, Schloss I, Barnes DK, Servetto N, Tarantelli S, Tatián M. 2015. Climate change and glacier retreat drive shifts in an Antarctic benthic ecosystem. Sci. Adv. 1(10):e1500050.

Sahade R, Tatián M, Kowalke J, Kühne S, Esnal G. 1998. Benthic faunal associations on soft substrates at Potter Cove, King George Island, Antarctica. Polar Biol. 19(2):85-91.

Schloss IR, Abele D, Moreau S, Demers S, Bers AV, González O, Ferreyra GA. 2012. Response of phytoplankton dynamics to 19-year (1991–2009) climate trends in Potter Cove (Antarctica). J Mar Syst. 92(1):53-66.

Stammerjohn SE, Martinson DG, Smith RC, Iannuzzi RA. 2008. Sea ice in the western Antarctic Peninsula region: Spatio-temporal variability from ecological and climate change perspectives. Deep Sea Res. Part II Top. Stud. Oceanogr. 55(18-19):2041-2058.

Stewart JB, Chinnery PF. 2015. The dynamics of mitochondrial DNA heteroplasmy: implications for human health and disease. Nat. Rev. Genet. 16(9):530-542.

Togunwa OS, Abdullah WH. 2017. Geochemical characterization of Neogene sediments from onshore West Baram Delta Province, Sarawak: paleoenvironment, source input and thermal maturity. Open Geosci. 9(1):302-313.

Unal E, Bucklin A, Lenz PH, Towle DW. 2013. Gene expression of the marine copepod *Calanus finmarchicus*: Responses to small-scale environmental variation in the Gulf of Maine (NW Atlantic Ocean). J Exp Mar Biol Ecol. 446:76-85.

Vaughan DG, Marshall GJ, Connolley WM, Parkinson C, Mulvaney R, Hodgson DA, King JC, Pudsey CJ, Turner J. 2003. Recent rapid regional climate warming on the Antarctic Peninsula. Clim Change. 60(3):243-274.

Wölfl A-C, Lim CH, Hass HC, Lindhorst S, Tosonotto G, Lettmann KA, Kuhn G, Wolff J-O, Abele D. 2014. Distribution and characteristics of marine habitats in a subpolar bay based on hydroacoustics and bed shear stress estimates—Potter Cove, King George Island, Antarctica. Geo-Mar Lett. 34(5):435-446.

Wölfl A-C, Wittenberg N, Feldens P, Hass HC, Betzler C, Kuhn G. 2016. Submarine landforms related to glacier retreat in a shallow Antarctic fjord. Antarct Sci. 28(6):475-486.

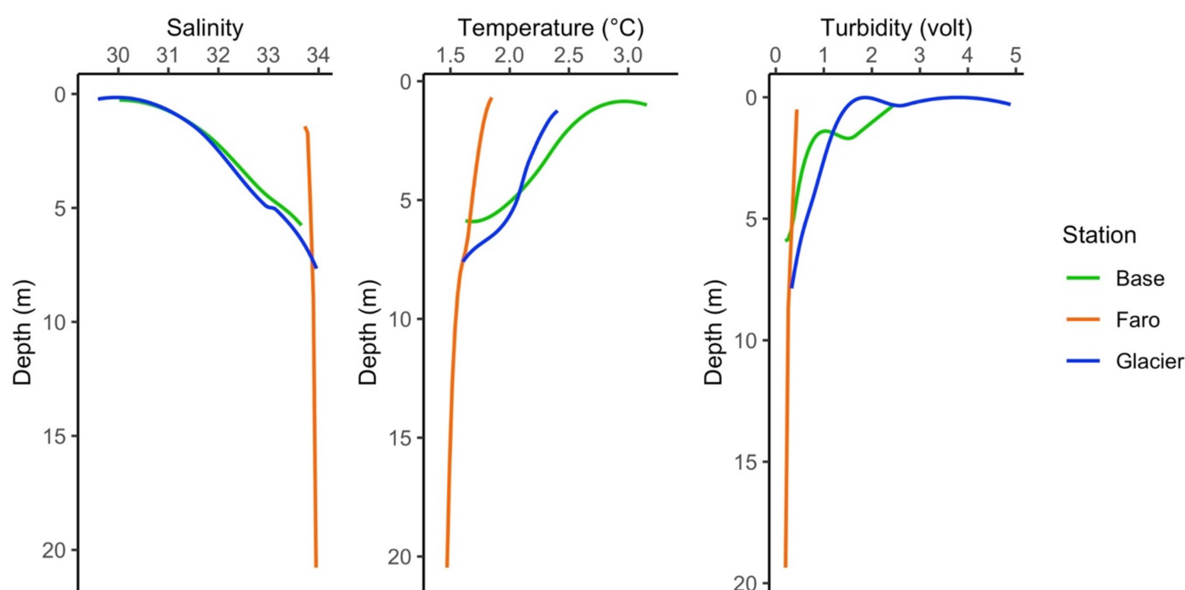
Young MD, Wakefield MJ, Smyth GK, Oshlack A. 2010. Gene ontology analysis for RNA-seq: accounting for selection bias. Genome Biol. 11(2):R14.

Zouros E, Rodakis GC. 2019. Doubly Uniparental Inheritance of mtDNA: An Unappreciated Defiance of a General Rule. Adv Anat Embryol Cell Biol. 231:25-49.

[dataset] Martinez, M., Harms, L., Abele, D., & Held, C.; 2020; Mitochondrial heteroplasmy as a systematic bias in molecular species delimitation and barcoding (ERP122389); European nucleotide Archive database (EMBL-EBI); ERR4265443, ERR4276392 – ERR4276460

APPENDIXES

Appendix 3.1. CTD profiles of salinity, temperature and turbidity at each station in Potter Cove.



Appendix 3.2. GO terms, over-represented *p* value and Counts for every mitotype comparison resulting from the GO enrichment analysis based on mitotypes.

Category	GO term	Over-represented *p* value	Count
	h1h1 vs. h2h2		
GO:0005743	CC mitochondrial inner membrane	5,39E-06	5
GO:0016021	CC integral component of membrane	1,22E-05	6
GO:0008137	MF NADH dehydrogenase (ubiquinone) activity	4,44E-05	4
GO:0070469	CC respiratory chain	1,02E-04	4
GO:0015078	MF proton transmembrane transporter activity	2,33E-02	1
GO:0015986	BP ATP synthesis coupled proton transport	2,33E-02	1
GO:0045263	CC proton-transporting ATP synthase complex	2,33E-02	1
GO:0006119	BP oxidative phosphorylation	2,36E-02	1
GO:0009060	BP aerobic respiration	2,36E-02	1
GO:0020037	MF heme binding	2,36E-02	1
GO:0045277	CC respiratory chain complex IV	2,36E-02	1
GO:0004129	MF cytochrome-c oxidase activity	4,80E-02	1
GO:0042773	BP ATP synthesis coupled electron transport	4,87E-02	1
	h1h3 vs h2h3		

GO:0004459	MF L-lactate dehydrogenase activity	0,03	1
GO:0016887	MF ATPase activity	0,03	1
GO:0019516	BP lactate oxidation	0,03	1
GO:0035499	BP carnosine biosynthetic process	0,03	1
GO:0047730	MF carnosine synthase activity	0,03	1
GO:0102102	MF homocarnosine synthase activity	0,03	1
GO:0005743	CC mitochondrial inner membrane	0,03	3
	h1h1 vs h1h3		
GO:0035556	BP intracellular signal transduction	0,01	9
GO:0016021	CC integral component of membrane	0,02	19
GO:0003779	MF actin binding	0,03	5
GO:0004674	MF protein serine/threonine kinase activity	0,03	7
GO:0006468	BP protein phosphorylation	0,03	7
GO:0070469	CC respiratory chain	0,03	9
GO:0046872	MF metal ion binding	0,03	11
	h1h1 vs h2h3		
GO:0016021	CC integral component of membrane	1,96E-06	18
GO:0005743	CC mitochondrial inner membrane	3,37E-05	10
GO:0070469	CC respiratory chain	1,05E-03	8
GO:0008137	MF NADH dehydrogenase (ubiquinone) activity	1,48E-03	7
GO:0006468	BP protein phosphorylation	1,65E-03	6
GO:0004674	MF protein serine/threonine kinase activity	1,68E-03	6
GO:0007275	BP multicellular organism development	6,35E-03	5
GO:0005737	CC cytoplasm	7,57E-03	15
GO:0000287	MF magnesium ion binding	8,73E-03	4
GO:0035556	BP intracellular signal transduction	8,78E-03	6
GO:0005886	CC plasma membrane	1,57E-02	6
GO:0005576	CC extracellular region	2,25E-02	4
GO:0046872	MF metal ion binding	2,59E-02	7
GO:0001669	CC acrosomal vesicle	3,34E-02	3
GO:0004888	MF transmembrane signaling receptor activity	4,51E-02	2
GO:0051260	BP protein homooligomerization	4,56E-02	2
GO:0005525	MF GTP binding	4,56E-02	4
GO:0007017	BP microtubule-based process	4,56E-02	4
GO:0035036	BP sperm-egg recognition	4,61E-02	2
GO:0060388	CC vitelline envelope	4,61E-02	2
GO:0044877	MF protein-containing complex binding	4,65E-02	2
GO:0042773	BP ATP synthesis coupled electron transport	4,86E-02	2
	h2h2 vs h2h3		
GO:0005737	CC cytoplasm	9,13E-06	28
GO:0007275	BP multicellular organism development	1,38E-03	7

GO:0035556	BP intracellular signal transduction	1,63E-03	9
GO:0005874	CC microtubule	1,11E-02	8
GO:0004674	MF protein serine/threonine kinase activity	1,19E-02	7
GO:0006468	BP protein phosphorylation	1,30E-02	7
GO:0001669	CC acrosomal vesicle	2,63E-02	4
GO:0005634	CC nucleus	3,26E-02	12
GO:0007286	BP spermatid development	4,60E-02	8

Appendix 3.3 GO terms and associated biological processes for single differential expressed genes resulting from differential expression analysis (DEA) by stations and nuclear genotype.

DEA by stations – biological processes

Base vs. Faro

Term ID	Description
GO:0007155	cell adhesion
GO:0009609	response to symbiotic bacterium
GO:0042391	regulation of membrane potential
GO:0050877	neurological system process
GO:0051260	protein homooligomerization
GO:0070715	sodium-dependent organic cation transport
GO:0006637	acyl-CoA metabolic process
GO:0006006	glucose metabolic process
GO:0070936	protein K48-linked ubiquitination
GO:0007268	chemical synaptic transmission
GO:0009650	UV protection
GO:0035499	carnosine biosynthetic process
GO:0010506	regulation of autophagy
GO:0055114	oxidation-reduction process
GO:0051726	regulation of cell cycle
GO:0043161	proteasome-mediated ubiquitin-dependent protein catabolic process
GO:0019516	lactate oxidation
GO:0042981	regulation of apoptotic process
GO:1990961	drug transmembrane export
GO:0015697	quaternary ammonium group transport
GO:0006544	glycine metabolic process
GO:0071447	cellular response to hydroperoxide
GO:0097340	inhibition of cysteine-type endopeptidase activity
GO:0006814	sodium ion transport
GO:0006468	protein phosphorylation
GO:0046474	glycerophospholipid biosynthetic process

GO:0097435	supramolecular fiber organization
GO:0048251	elastic fiber assembly

Base vs. Glacier

Term ID	Description
GO:0006979	response to oxidative stress
GO:0007155	cell adhesion
GO:0019835	cytolysis
GO:0070177	contractile vacuole discharge
GO:0033298	contractile vacuole organization
GO:0070936	protein K48-linked ubiquitination
GO:0048800	antennal morphogenesis
GO:0042438	melanin biosynthetic process
GO:2001225	regulation of chloride transport
GO:0000281	mitotic cytokinesis
GO:0032849	positive regulation of cellular pH reduction
GO:0006915	apoptotic process
GO:0017148	negative regulation of translation
GO:0000463	maturation of LSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)
GO:0009650	UV protection
GO:0033146	regulation of intracellular estrogen receptor signaling pathway
GO:0031038	myosin II filament organization
GO:0048251	elastic fiber assembly
GO:0031288	sorocarp morphogenesis
GO:0010712	regulation of collagen metabolic process
GO:0006468	protein phosphorylation

Glacier vs. Faro

Term ID	Description
GO:0043627	response to estrogen
GO:0061614	pri-miRNA transcription from RNA polymerase II promoter
GO:0005975	carbohydrate metabolic process
GO:1904184	positive regulation of pyruvate dehydrogenase activity
GO:0010288	response to lead ion
GO:0001666	response to hypoxia
GO:0006412	translation
GO:0042594	response to starvation
GO:0051412	response to corticosterone
GO:0010033	response to organic substance

DEA by nuclear genotype – biological processes

NucA vs NucB

Term ID	Description
GO:0009609	response to symbiotic bacterium
GO:0010506	regulation of autophagy
GO:0030214	hyaluronan catabolic process
GO:0048800	antennal morphogenesis
GO:0070715	sodium-dependent organic cation transport
GO:0000463	maturation of LSU-rRNA from tricistronic rRNA transcript
GO:0070936	protein K48-linked ubiquitination
GO:0071447	cellular response to hydroperoxide
GO:0046653	tetrahydrofolate metabolic process
GO:0097340	inhibition of cysteine-type endopeptidase activity
GO:0055114	oxidation-reduction process
GO:0051726	regulation of cell cycle
GO:0016055	Wnt signaling pathway
GO:1990961	drug transmembrane export
GO:0015697	quaternary ammonium group transport
GO:0060731	positive regulation of intestinal epithelial structure maintenance
GO:0006814	sodium ion transport
GO:0033514	L-lysine catabolic process to acetyl-CoA via L-pipecolate
GO:0009166	nucleotide catabolic process
GO:0006468	protein phosphorylation

CHAPTER 4

4 Publication III

Mariano Martínez, Marcelo González-Aravena, Christoph Held, Doris Abele

A molecular perspective on the invasibility of the Southern Ocean benthos: the impact of hypoxia and temperature shifts on gene expression in South American and Antarctic bivalves

Manuscript status: Submitted to *Global Change Biology*

A molecular perspective on the invasibility of the Southern Ocean benthos: the impact of hypoxia and temperature shifts on gene expression in South American and Antarctic bivalves

Mariano Martínez*[1], Marcelo González-Aravena[2], Christoph Held[1], Doris Abele[1]

1) Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung, Bremerhaven, Germany

2) Departamento Científico, Instituto Antártico Chileno, Puntas Arenas, Chile

*Corresponding author: Mariano Martínez. E-mail: mariano.martinez@awi.de

ABSTRACT

Whenever an organism or species makes a major transition colonizing a new habitat, the associated environmental conditions usually change even more abruptly than ongoing environmental change in situ. Exposing shallow-water marine bivalves (*Aequiyoldia cf. eightsii*) from the Strait of Magellan in southern South America (SSA) and the West Antarctic Peninsula (WAP) to changes in temperature and oxygen availability (isolated and combined), we investigated changes in gene expression patterns in a simulated colonization of the shores of a new continent after crossing of the Drake Passage and in a warming scenario in the WAP. Bivalves from SSA were cooled from 7 °C (in situ) to 4 °C and 2 °C (future warmed WAP conditions), WAP bivalves were warmed from 1.5 °C (in situ) to 4 °C (warmed WAP), gene expression patterns in response to thermal stress by itself and in combination with hypoxia were measured after 10 days. Our results confirm that molecular adaptation (here: gene expression patterns) may play a vital role in local adaptation. In general, hypoxia had a greater effect on the transcriptome than temperature by itself (171 vs 14 and 169 vs 90 differentially expressed genes in SSA and WAP, respectively). The effect was further amplified when hypoxia and temperature acted as combined stressors. The WAP bivalves showed a remarkable ability to cope with short-term exposure to hypoxia by switching to a metabolic rate depression strategy and activating the alternative oxidation pathway, whilst the SSA population had no such response. In SSA, the high prevalence of apoptosis (cell-death)-related differentially expressed genes especially under combined higher temperatures and hypoxia indicated that the SSA *Aequiyoldia* are operating near their physiological limits already. While the effect of temperature *per se* may not represent the single most effective barrier to Antarctic colonization by South American bivalves, the current distribution patterns as well as their resilience to future conditions may be better understood by looking at the synergistic effects of temperature in conjunction with short term exposure to hypoxia.

Key words: warming, hypoxia, non-indigenous species, West Antarctic Peninsula, Alternative oxidase

1. INTRODUCTION

Temperature and oxygen availability are key physiological variables shaping the fitness of marine ectotherms and determine their biogeographical distributions (temperature) and the ecological niche they can locally occupy (oxygen availability) (Poloczanska et al. 2013; Deutsch et al. 2015). Rising global temperatures decrease oxygen solubility in seawater and enhance stratification, reducing downward mixing of oxygen into midwater and deep layers (Shepherd et al. 2017; Li et al. 2020). Increasing oxygen demand and organic matter turnover at higher temperatures intensify spreading of hypoxic or even anoxic zones in coastal sediments and shallow waters, limiting survival of hypoxia-sensitive species while opening new space for hypoxic/anoxic-tolerant species. Hence, global warming has the potential to alter the boundaries of current species distributions (Brown and Thatje 2015).

One approach toward understanding the biological consequences of climate change for marine biodiversity therefore lies in the prediction of future distribution of species by experimentally determining the physiological influence of changes in temperature and oxygen levels, and analysing the molecular mechanisms that alter the responding physiological pathways. The relevance of the experimental results depends strongly on a realistic prediction of future environmental scenarios for current habitats, or the scenario expected in potential new habitats. In this study we used the protobranch bivalve *Aequiyoldia cf. eightsii* from the West Antarctic Peninsula (WAP) and southern South America (SSA) as a model to evaluate combined effects of temperature changes and hypoxia on patterns of gene transcription. Rather than increasing the strength of stressor(s) until a physiological response can be registered (acute stress response), this study employs a different strategy in that we let both the direction and the magnitude of the departure from normal conditions (temperature) be determined by a clearly defined hypothesis: the crossing of an organism across the Drake Passage from SSA to the WAP or vice versa. Contrary to studies that follow an already successful invasive species in its old and new distribution range (Tepolt and Palumbi 2015), this major dispersal event has to our knowledge not happened within recent times in the marine bivalve *A. cf. eightsii* yet, hence we investigate the role (if any) that physiology may have

played in preventing this transition until now. In parallel, we study how a temperature increase (by itself and in combination with hypoxia) leads to stress in Antarctic bivalves, which mimics either a continued warming in situ or a successful transition of the Drake Passage northwards.

Since the early 2000s, the number of Antarctic cruise liners is rapidly increasing, and with the majority of scientific and logistic activities concentrating between SSA and the WAP, the risk of introducing non-indigenous species (NIS) by passive transport onto the shelves of a rapidly warming Antarctic continent is intensifying (Chown et al. 2012; Hughes et al. 2015; McCarthy et al. 2019). In addition, reports of passive introductions of NIS in kelp rafts without direct human assistance have recently increased on the WAP (Fraser et al. 2018; Avila et al. 2020). Thus, polar species/communities are not only threatened by having to cope with the fastest warming zones on earth (Clem et al. 2020) without the possibility of evading to colder higher latitudes, but also by the greater risk of being disrupted by non-indigenous species (NIS) from lower latitudes through competition for resources or trophic interactions (Chan and Briski 2017). Documented invasions of terrestrial and marine organisms at the WAP are increasing, and recently the blue mussel *Mytilus cf. platensis* was added to the list as the first mussel introduced on King-George Island (WAP) (Cardenas et al. 2020), a major turnstile of West Antarctic tourism and logistic shipping.

Aequiyoldia bivalves are a conspicuous genus in the Southern Ocean inhabiting soft-substratum ecosystems of SSA and the WAP, and several Subantarctic islands (González-Wevar et al. 2019). Although the populations of SSA and the WAP are currently formally considered the same species, recent molecular studies based on Single Nucleotide Polymorphisms (SNPs) and mitochondrial and nuclear markers suggest that populations on either side of the Drake should be considered reproductively isolated species (González-Wevar et al. 2019, Chapter 2). For this reason, in the following we will call both species *A. cf. eightsii* but emphasize that Antarctic and South American populations are likely different, but closely related species. Recently, Muñoz-Ramírez et al. (2020) highlighted based on genomic data the role of the Antarctic Circumpolar Current as a biogeographic barrier to larval transport between both continents, preventing or

massively reducing the potential for genetic connectivity between Antarctic and South American *Aequiyoldia* species.

For *A. cf. eightsii* and many Antarctic marine ectotherms the physiological mechanisms that protect cells during acute and chronic thermal stress, oxygen deficiency and ocean acidification (e.g., heat shock response, activities of some antioxidant enzymes, metabolic regulation, etc.) have been studied previously (Abele et al. 2001; Abele and Puntarulo 2004; Weihe et al. 2010; Clark et al. 2013; Johnson and Hofmann 2020). Previous studies of functional responses and gene transcription (Clark et al. 2016) revealed Antarctic *Aequiyoldia* to be relatively tolerant in a temperature ramp experiment compared to six other Antarctic endemic invertebrates from different clades. In an earlier experimental study Abele et al. (2001) showed respiration and locomotory activity of Antarctic *A. cf. eightsii* to increase steadily in response to warming (1°C per 10 h), confirming that *A. cf. eightsii* tolerates temperatures up to 5°C (max. in situ temperature 2°C), with burrowing activity undiminished at these temperatures. If warmed at a slower pace (1°C per 5 days) temperature tolerance was even slightly higher (7°C), although accompanied by loss of antioxidant function and tissue damage in the lipid phase (i.e., membrane damage).

We hypothesized that the predicted warming at the WAP in combination with hypoxia in coastal sediments (Sampaio et al. 2021) represents a physiological stress condition for Antarctic *A. cf. eightsii* and might favour immigration across the Drake Passage from SSA at the northern WAP region. Even though both environments are predicted to warm, waters at the WAP climate are clearly colder relative to SSA, which implies that South American bivalve adults or larvae crossing the Drake would face a rather sudden cooling scenario. The experimental approach on both continents consisted in exposing bivalves to oxygen saturation (21% O_2 saturation, from here: normoxia) and hypoxia (<2% O_2 saturation) under current in situ summer temperature conditions (1.5 °C in WAP and 7°C in SSA). In parallel we tested a near future warming scenario predicted for WAP nearshore bottom water (see Morley et al. 2020) for Antarctic specimens (4°C), as well as a scenario of southward migration of SSA *Aequiyoldia* and their capacity to survive in

shallow water environments of Northern WAP at present (2°C) and in the future (4°C) (Fig. 4.1).

2. METHODS

2.1 Animal collection and maintenance

Aequiyoldia cf. eightsii for the experiments were collected in southern South America (SSA) and West Antarctic Peninsula (WAP) between October 2017 and January 2018. Bivalves from SSA were hand-collected by SCUBA divers in the shallow subtidal (< 3m depth) in the Rinconada Bulnes, Strait of Magellan, Chile (53°37'52"S; 70°56'54"W) on a single day in October 2017. Antarctic bivalves were collected on a single sampling day in January 2018 from one site in Potter Cove, King George Island, South Shetlands (62°14'11"S; 58°40'14"W) in 6 m water depth using a Van Veen grab. At the sampling site, local in situ temperature and salinity were recorded with a Multiparameter HANNA (HI 9828) in SSA and a with a Sea-Bird CTD (SBE19plusV2, Sea-Bird Electronics, Bellevue, WA, USA) in WAP. In addition, sediment cores were taken by SCUBA divers using cylindrical Plexiglas corer (height: 50 cm, diameter: 8 cm; 3 replicates per station) for analyses of total organic carbon (TOC) and total sulphur (TS) in the first 2 cm sediment-layer, and the ratio TOC: TS was calculated as a as proxy of sediment oxygenation conditions (Togunwa and Abdullah 2017). Physico-chemical data at the collection sites are shown in Appendix 4.1.

At each location, bivalves were transported to the local research facility in insulated containers filled with water and sediment from the sampling site. Once in the laboratory, bivalves were immediately sorted from the muddy sediment, checked for shell damage and vitality (foot movements), and transferred to an aquarium supplied with seawater from the collection site and here maintained at in situ temperature. The bottom of the aquarium was covered with a 2 cm thick layer of sediment from the sampling site and allowed to settle for 2 h before adding the animals, and the aquarium was provided with aeration with the help of bubble stones and maintained in the dark during 10 days for acclimation. Incubation water was replaced every 48 h.

2.2 Experimental setup

After the acclimation period, four replicates of five animals each were subjected to each of the four combinations of temperature and oxygen concentration during 18 days in WAP and 21 days in SSA. The animals were randomly placed in 500 ml Kautex jars containing a 1 cm layer of sediment and water from the collection site. Each jar was then assigned to be exposed to treatment under one of the following experimental conditions: in WAP: In situ (1.5 °C) and warming (4 °C) in combination with normoxia ($T_{is}Ox_n$, T_wOx_n) and hypoxia ($T_{is}Ox_{hyp}$, T_wOx_{hyp}); in SSA: In situ (7 °C) and South migration future scenario (4 °C) in combination with normoxia ($T_{is}Ox_n$, T_cOx_n) and hypoxia ($T_{is}Ox_{hyp}$, T_cOx_{hyp}). The water temperature inside the jars was kept constant by submerging the jars in temperature-controlled water baths (Thermo Haake DC10-P21). Hypoxic conditions were created by providing the experimental units with a continuous water flow with 2% O_2 saturation, with flow velocity adjusted to achieve the exchange of the total water volume of the jars (500ml) in the course of one day. The water was supplied from a 10 L tank bubbled with a gas mixture 2% O_2:98% N_2. This design could not be replicated in the normoxic treatment, since a test run with continuous water flow at 21 % O_2 saturation through the hermetically sealed lids showed that animal and sedimentary/microbial respiration decreased oxygen to below 15 % saturation within some hours, with uncontrollable variation between replicates and over time. Therefore, normoxia conditions were ensured by directly bubbling each experimental jar with air and manually replacing the total volume of water in each jar every day (without stirring up the sediment).

In both experiments, one animal per jar (i.e., four replicates per treatment) was collected at day 10 (t_1) for differential gene expression analysis. In the experiment in SSA, an additional step was conducted on day 18, by further decreasing the temperature in the T_cOx_n treatment from 4 °C down to 2 °C ($T_{c+}Ox_n$). On day 21 (t_2), one bivalve per jar was collected from $T_{is}Ox_n$ and $T_{c+}Ox_n$ treatments for differential gene expression analysis (Fig. 4.1). Animals were immediately dissected on ice under a stereomicroscope after collection, and mantle tissue was conserved in RNA later (SIGMA) and stored at -80 °C.

During the total period of both experiments, temperature was recorded continuously in both experimental water baths, and measurements of oxygen were performed daily in each of the four replicates in the four treatments (Fig. 4.1). Oxygen measurements were made approximately one centimetre away from the water-sediment interface, introducing the sensor through a small gate in the lid of the jars.

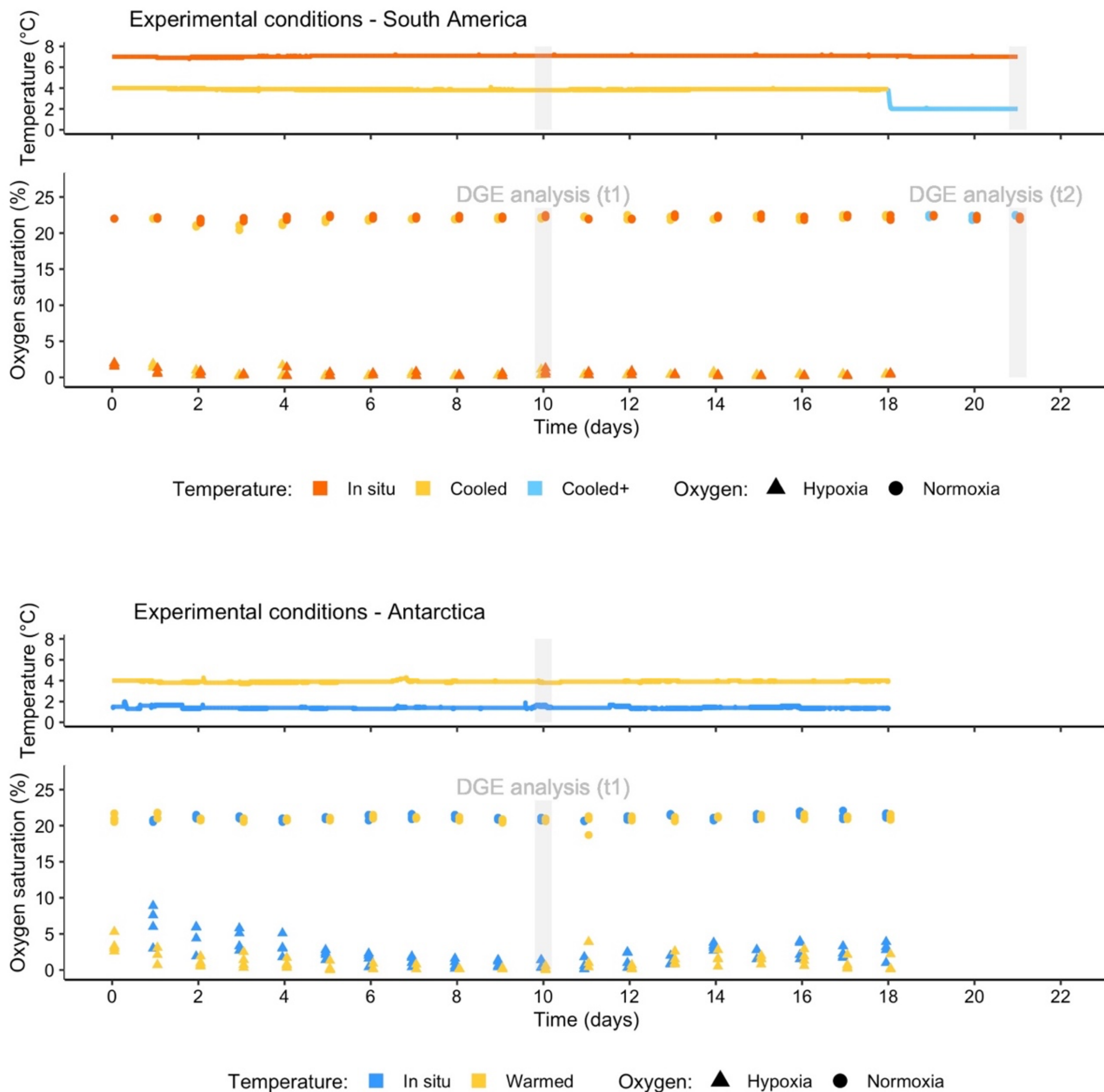


Figure 4.1. Experimental temperature and oxygen recorded during treatments in SSA and WAP: continuous record of temperature (above) and punctual measures of oxygen in each replicate of the four treatments (below). Gray bars indicate the collection day (t_1 and t_2) for differential gene expression (DGE) analysis.

2.3 RNA extraction, library preparation and *de novo* assembly

Differential gene expression analyses were performed using an assembly created from 70 libraries (one library per individual) of *A. cf. eightsii* from SSA and WAP by Martínez et al. (Chapter 2). All the libraries used in the present study (n = 40) for the differential gene expression analysis were used to create the assembly.

Total RNA was isolated from mantle tissue (5 – 30 mg) using the Direct-zol™ RNA MiniPrep kit (ZYMO Research Corp., USA) according to the manufacturer's instructions. A total of 40 individual libraries (one per specimen) were prepared using the Illumina TruSeq® Stranded mRNA Sample Preparation kit starting from 1 µg of total RNA following the protocol provided by the kit. Libraries were tagged through adapter ligation, diluted to 0.8 nM, and subsequently pooled and cleaned using magnetic beads (AMPure XP, Beckmann Coulter) to remove the remaining primer content. Final cDNA concentration was measured utilizing a LabChip® GX Touch (PerkinElmer, USA). The pool of samples was sequenced on an Illumina NextSeq 500 sequencer using the NextSeq High Output Kit v2 (150 cycles) with a paired-end protocol.

Raw reads were quality controlled by FastQC v. 0.11.7 (Babraham Institute, Cambridge, UK), cleaned and *de novo* assembled using the Trinity genome-independent transcriptome assembler v2.8.4; (Grabherr et al. 2011). For details on *de novo* assembly parameters see Chapter 2.

2.4 Differential gene expression and Gene Ontology analyses

The differential gene expression analysis involved the alignment of the short reads of each sample separately against the *de novo* reference transcriptome using Bowtie2 v 3.3.4.1 (Langmead et al. 2009). This involved pairwise differential gene expression analysis among the four treatments of temperature and oxygen at t_1 in the Antarctic (WAP) experiment ($nT_{is}Ox_n$ = 4, nT_wOx_n = 4, $nT_{is}Ox_{hyp}$ = 4, nT_wOx_{hyp} = 4), and separately, among treatments in South America (SSA) at t_1 ($nT_{is}Ox_n$ = 4, nT_cOx_n = 4, $nT_{is}Ox_{hyp}$ = 4,

nT_cOx_{hyp} = 4). Additionally, in South America, a pairwise differential gene expression analysis was carried out between normoxic treatments under in situ (7 °C) and low (2 °C) temperature exposure at t_2 ($nT_{is}Ox_n$ = 4, $nT_{c+}Ox_n$ = 4). In order to contrast the effects of temperature, the comparison between normoxic treatments under in situ (7 °C) and cold exposure scenarios (4 °C) at t_1 ($nT_{is}Ox_n$ = 4, nT_cOx_n = 4) was also included in this analysis.

Relative abundances of transcripts were estimated by RSEM version 1.2.26 (Li and Dewey 2011) and differential expression of genes was assessed using a test based on the negative binomial distribution as integrated in the Bioconductor R package DESeq2 (Love et al. 2014), with a significance threshold set to $p \leq 0.001$ and a fold change of at least 2. Gene expression results are shown graphically in a Principal component Analysis (PCA) performed with the normalized counts per sample of every differentially expressed gene (DEG). The tools were executed using the Trinity package v 2.8.4. The annotation of the DEG was performed using DIAMOND v 0.9.24 (Buchfink et al. 2015) including a homology search against the UniProt Swiss-Prot database and assigning Gene Ontology (GO) terms to annotated transcripts. GO enrichment analyses were carried out using GOseq (Young et al. 2010) by separately considering up- and down-regulated genes of every pairwise differential gene expression analysis and their corresponding GO annotation, and using the full list of DEG with GO terms as the background.

3. RESULTS

3.1 Differential gene expression analysis

In both the WAP and SSA experiments, the effect of oxygen alone (constant in situ temperature, 10 days) resulted in a higher number of differentially expressed genes (DEG) than temperature alone (under normoxia). Moreover, at both SSA and WAP sampling sites, the highest number of DEG clearly resulted from the combination of exposure to higher temperature and hypoxia (T_wOx_{hyp} in WAP and $T_{is}Ox_{hyp}$ in SSA) when compared with the treatment of lower temperature and normoxia (Table 4.1). Although one experiment (WAP) involves warming and the other involves cooling (SSA; both with roughly same magnitude of change), the most notable difference between experiments

from either continent was the influence of the higher experimental temperature, i.e., experimental warming of Antarctic *Aequiyoldia cf. eightsii* caused ca. five times higher number of DEG (90 DEG) than exposure to cooling in South American *A. cf. eightsii* (14 DEG, see Table 4.1). In other words, Antarctic *A. cf. eightsii* respond more vigorously to warming whereas hypoxia affects both populations to approximately the same extent. Even though the stringency of thresholds for the detection of DEG ($p \leq 0.001$) affects the overall number of DEG, the DEG ratios between pairwise comparisons is expected to be less affected, thus making the magnitudes of stressors effect comparable.

Table 4.1. Number of differentially expressed genes (DEG) at each pairwise comparison among the four treatments of temperature and oxygen at time t_1 (10 days) in the experiments performed in West Antarctic Peninsula (WAP) and southern South America (SSA). In bold the highest number of DEG at each experiment, which in both cases resulted from the comparison between higher temperature and hypoxia (T_wOx_{hyp} in WAP and $T_{is}Ox_{hyp}$ in SSA) and lower temperature and normoxia treatments.

	WAP					SSA			
	$T_{is}Ox_n$	$T_{is}Ox_{hyp}$	T_wOx_n	T_wOx_{hyp}		$T_{is}Ox_n$	$T_{is}Ox_{hyp}$	T_cOx_n	T_cOx_{hyp}
$T_{is}Ox_n$		169	90	**298**	$T_{is}Ox_n$		171	14	211
$T_{is}Ox_{hyp}$			93	47	$T_{is}Ox_{hyp}$			**312**	9
T_wOx_n				229	T_cOx_n				202
T_wOx_{hyp}					T_cOx_{hyp}				

The stronger influence of oxygen availability on gene expression patterns becomes obvious in the Principal Component Analysis (PCA) (Fig.4.2A: WAP, 4.3A: SSA). In both PCAs, the Principal Component (PC) 1 mainly separates animals according to oxygenation conditions explaining most of the variance, whereas PC2 distinguishes organisms based on temperature conditions with a lower percentage of explained variance. This pattern is clearer in the PCA of the experiment in SSA, in which the

individual response within each treatment group resulted more homogeneous (Fig. 4.3A).

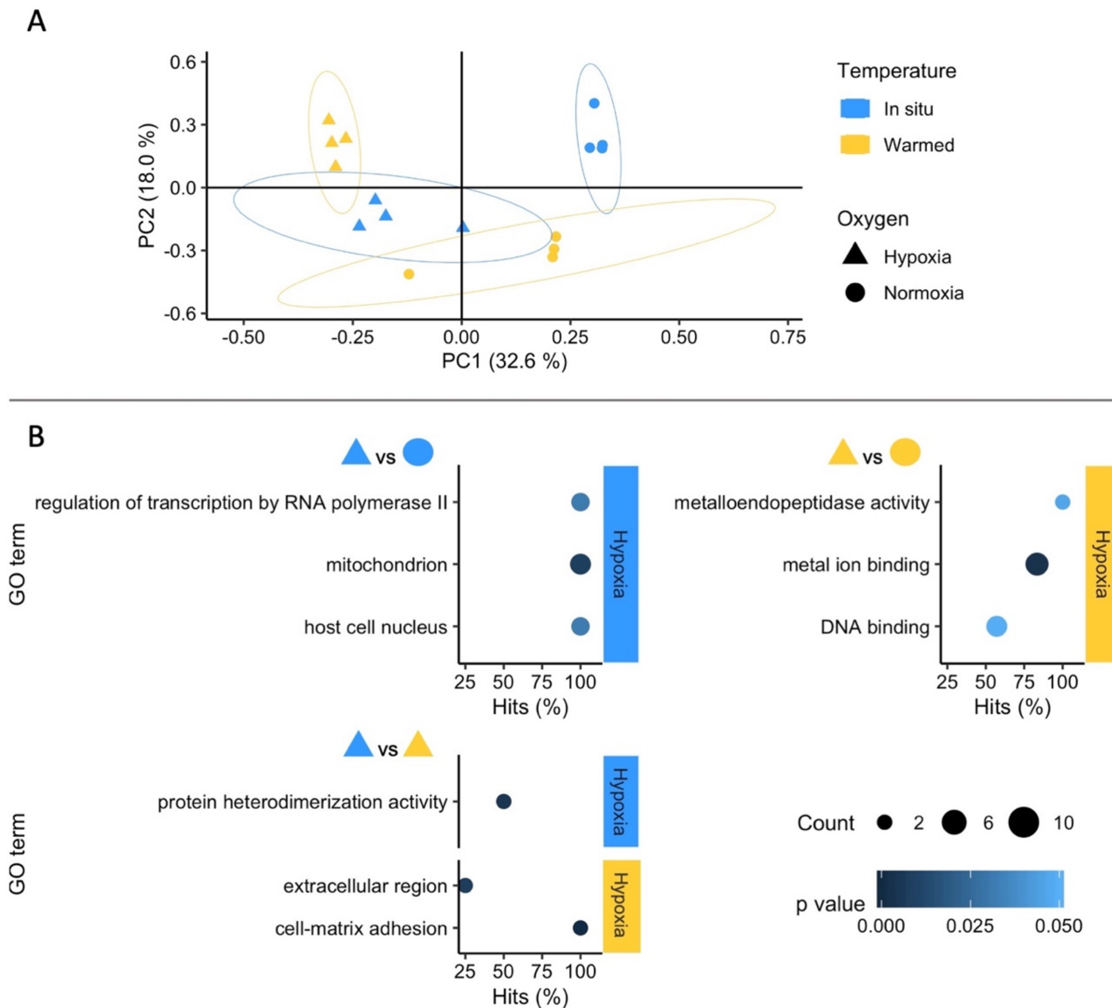


Figure 4.2. A: Principal Component Analysis based on counts of differentially expressed genes (DEG) in the West Antarctic Peninsula experiment (WAP). Colours and shapes identify temperature and oxygen treatments, respectively, and a 95 % confidence ellipse is shown for each of the four treatments. B: GO enrichment analysis of pairwise comparisons. Enriched terms are shown separately for the two compared treatments identifying temperature (colours) and oxygen conditions in the right side of each plot. The *x* axis indicates the percentage of DEG found in a certain category over the total number of genes in this category (Hits). Size code indicates the number of DEG in the GO term category (Count) and a gradient of colour the *p value* of the

over represented category. Only GO terms with more than one DEG in the category are shown in the plots (see full list in Appendix 4.2).

3.2. Gene ontology analysis

Oxygen was also the most influential factor in the enrichment of Gene Ontology (GO) terms on both continents. In SSA bivalves, the combined effects of variation in temperature and oxygen conditions also resulted in highly significant GO terms (Fig. 4.3B). Even though temperature alone also resulted in enriched terms, these were mostly based on single DEG and were hence treated with caution in our main analysis (but see Appendix 4.2). It should be noted that the GO term analysis is based only on annotated genes, i.e., a subset of all DEG, a distinction that is particularly significant here as in the study of all non-model species.

3.2.1 *Functional response to warming and hypoxia of Antarctic* Aequiyoldia

Raising the water temperature to 4° C under normoxic conditions did not elicit a clear-cut response in WAP *A. cf. eigthsii*, because most GO terms were supported by single genes only. Several of these GO terms (3 out of 9, Appendix 4.2) indicate a response of bivalve cellular immune defence.

Furthermore, the functional response to hypoxia in Antarctic *Aequiyoldia* was temperature-dependent. Under in situ temperature the most significant GO term was 'mitochondrion' involving the upregulation of the Alternative Oxidase (AOX), an alternative mitochondrial electron transport system (ETS) component that leads to a metabolic rate reduction, and extracellular metal binding proteins (e.g., Haloacid dehalogenase) (see Appendix 4.3). In individuals exposed to hypoxia at higher temperature the GO term 'metal ion binding' was most significant. Upregulation of (transition) metal binding capacity (in combination with the other enriched GO term "metalloendopeptidase activity") primarily points to activation of metal chaperones (Cu, Zn, Fe) or metal storage/transport proteins such as ferritin or transferrin. This function supports safe storage and translocation of harmful transition metal ions in the cellular

environment. Indeed, one of the up-regulated functions is a transferrin-initiating factor (see Appendix 4.3 which lists DEG supporting the most significant enriched). The GO term “mitochondrion” was not significant at warm temperature hypoxic exposure.

3.2.2 *Functional response to cooling and hypoxia of South American* Aequiyoldia

In SSA, enriched GO terms in response to hypoxia were almost all linked to the regulation of apoptotic process, e.g., programmed cell death, including protein ubiquitination for proteasomal digestion, under in situ and low temperature conditions. The same apoptotic signal was observed when South American *A. cf. eightsii* were exposed to the combination of warmer in situ temperature and hypoxia ($T_{is}Ox_{hyp}$) compared to animals exposed to cooling and normoxia (T_cOx_n). By comparison, the absence of apoptotic GO terms under hypoxic exposure in Antarctic *A. cf. eigthsii* reveals a greater tolerance to oxygen depletion at in situ temperature and moderate warming, presumably supported by mitochondrial uncoupling via alternative oxidase and control of tissue damage through transition metal binding (McDonald and Gospodaryov 2019; Eteshola et al. 2020).

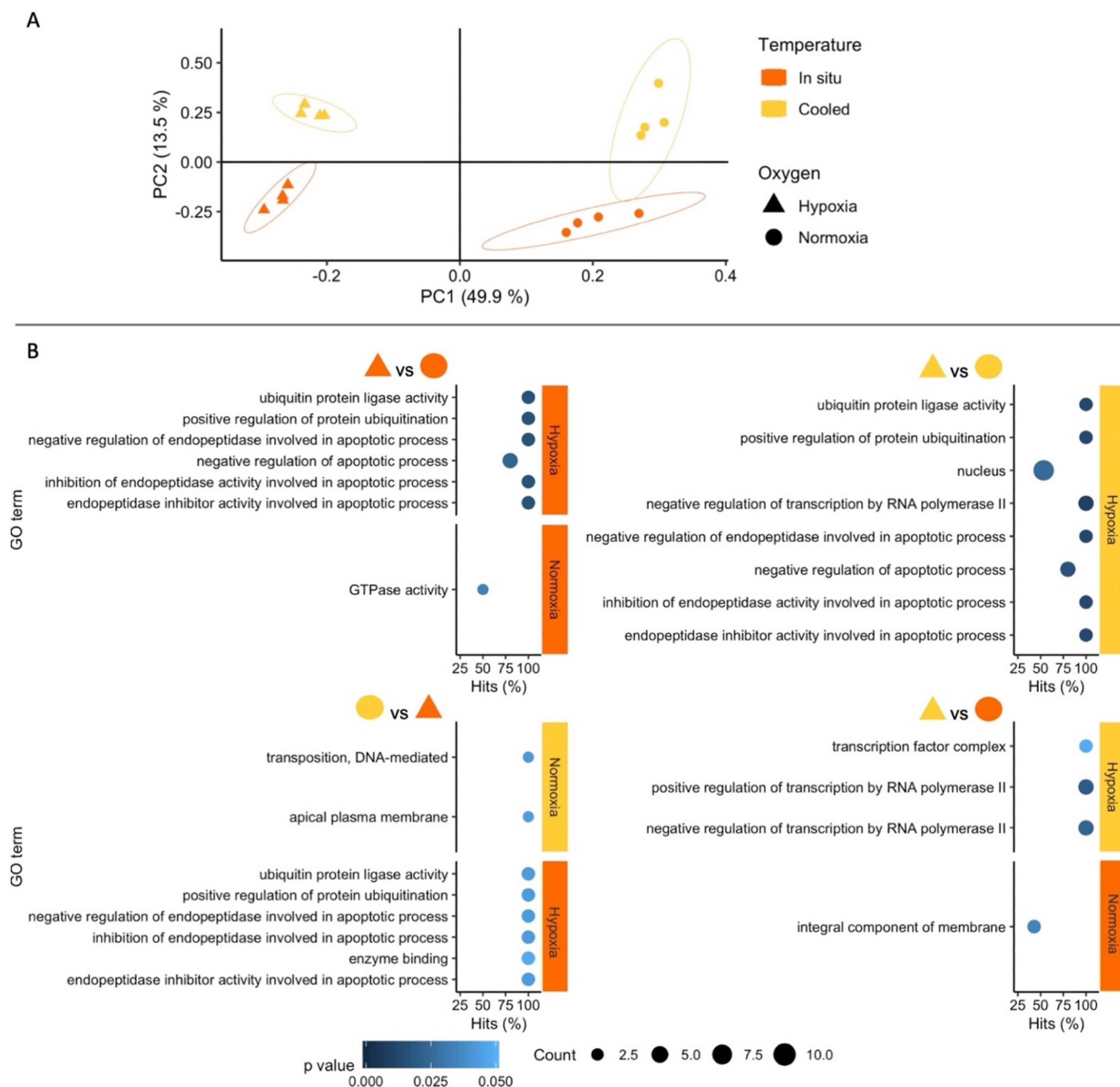


Figure 4.3. A: Principal Component Analysis based on counts of differentially expressed genes (DEG) in the experiment performed in southern South America (SSA). Colours and shapes identify temperature and oxygen treatments, respectively, and a 95 % confidence ellipse is indicated for each of the four treatments. B: GO enrichment analysis of pairwise comparisons. Enriched terms are shown separately for the two treatments compared identifying temperature (colours) and oxygen conditions on the right margin of each plot. The *x* axis indicates the percentage of DEG found in each category over the total number of genes in the category (Hits). Size code indicates the number of DEG in the GO term category (Count) and a gradient of colour the *p value* of the over represented category. Only GO terms with more than one DEG in the category are shown in the plots (see full list in Appendix 4.2).

We also tested the response of South American *A. cf eigthsii* to stepwise cooling to Antarctic shallow water summer temperatures, 2 °C (T_{c+}) without additional hypoxic stress. Results in terms of numbers of DEG between $T_{c+}Ox_n$ (2 °C) and $T_{is}Ox_n$ (7 °C) at t_2 did not differ from those observed between T_cOx_n (4 °C) and $T_{is}Ox_n$ (7 °C) at t_1 (15 and 14 DEG, respectively). This indicates that bivalves from SSA can physiologically tolerate 21 days exposure to current and future Antarctic summer temperatures under normoxic conditions. However, the GO enrichment analysis resulted in the enrichment of rather distinct GO terms between both comparisons and time points. While at t_1 (4 °C vs 7 °C, 10 days) various, not clearly interconnected functions were enriched (see Appendix 4.2), the comparison at t_2 (2 °C vs 7 °C, 21 days) resulted exclusively in enriched GO terms at the in situ temperature. Most of these terms were based on only single DEG and related to RNA binding and translation (Fig. 4.4B). The enrichment of GO terms as RNA binding and translation at the higher in situ temperature suggests a lowered protein synthesis/turnover at the low temperatures awaiting SSA *Aequiyoldia* crossing the Drake.

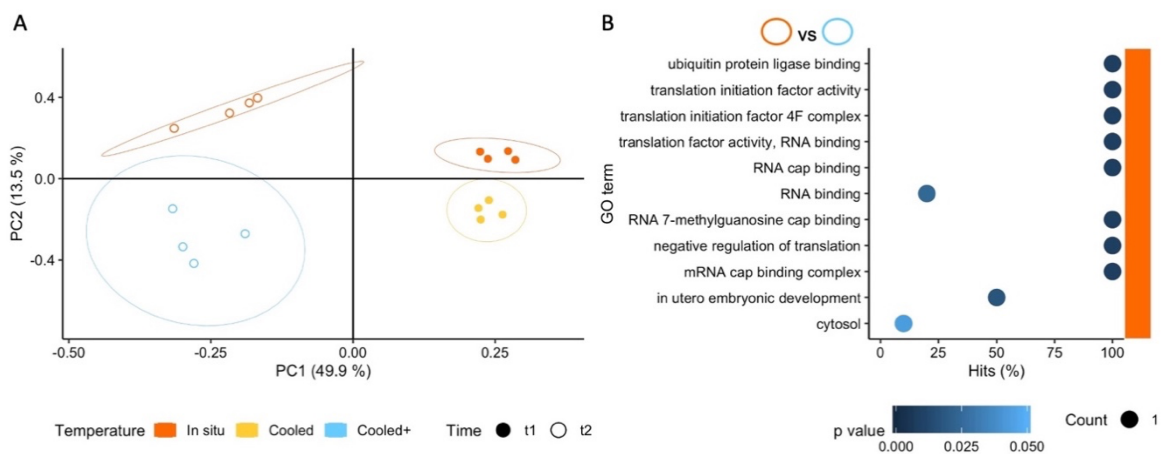


Figure 4.4. A: Principal Component Analysis based on counts of differentially expressed genes (DEG) for all three exposure temperatures applied in southern South America (SSA). Genes included in the analysis were significant ($p < 0.001$) in at least one pairwise comparison. Colours and shapes identify temperature and time of collection, respectively, and a 95 % confidence ellipse is shown for each of the four treatments. B: GO enrichment analysis of pairwise comparisons between $T_{c+}Ox_n$ (2 °C) and $T_{is}Ox_n$ (7 °C) at t_2. The *x* axis indicates the percentage of

DEG found in a certain category over the total number of genes in the category (Hits). Size code indicates the number of DEG in the GO term category (Count) and a gradient of colour indicates the *p value* of the over represented category. Note that only the treatment $T_{is}Ox_{n}$ (7 °C) indicated with colour on the right of the plot resulted in enriched terms.

4. DISCUSSION

In this work, we analysed the transcriptomic response of two *Aequiyoldia* sibling species from SSA and WAP to exposure scenarios simulating a crossing of the Drake Passage, and to an expected near future warming scenario at the WAP. We showed that the exposure to relatively small temperature changes and hypoxia, both of which are predicted to become more frequent for benthic infauna in high latitudes as SSA and WAP, act as combined stressors (Sampaio et al. 2021) in *Aequiyoldia cf. eightsii* bivalves. Interestingly, the effects of hypoxia in both species surpassed those imposed in an upcoming warming-only scenario forecasted for the WAP and those caused by a thermal transition in a hypothetical cross-continental invasion, with functional responses differing between species. The magnitude of the response was greater under combined hypoxia and elevated temperature, highlighting the threat from the synergistic impact of both parameters in a global warming context. The relatively moderate response to warming as a single stressor of Antarctic *A. cf. eightsii* is noteworthy, as Antarctic invertebrates are usually described as highly sensitive to smallest of thermal increments (Peck et al. 2010). In contrast, cooling down Punta Arenas specimens to summer temperatures already occurring in WAP nearshore environments today (2°C), or are expected in the near future if warming continues (4°C) did not lead to a significant change in the gene expression of South American bivalves. This indicates that water temperature *per se* is unlikely to have formed a physiological barrier to either species introduction or larval migration by South American *A. cf. eightsii* to the WAP.

4.1 Physiological response to thermal shifts and potential cross-continental invasions

In both congeneric species, a moderate shift in temperature by itself (without exposure to hypoxia) did not drastically affect gene expression, and we observed few clearly supported functional adjustments in response to cooling South American or warming Antarctic *A. cf. eightsii*.

The small effect of temperature (cooling) in the South American *A. cf. eightsii* is not surprising. South American *A. cf. eightsii* inhabit both subtidal and intertidal zones in Magellan Strait where annual temperature windows range from 5.9 to 9.9°C in sublittoral and from 1.5 to 18°C in shallow intertidal waters (Cardenas et al. 2020); a range that covers the acclimation temperatures of our cooling experiment and overlaps with current WAP water temperatures. For Magellan bivalves 2° C is, however, a temperature minimum they will experience in the natural environment only rarely and not for prolonged periods of time. The functional analysis indicates lower rates of gene transcription and protein translation, as well as protein degradation (ubiquitination) in bivalves from SSA maintained under cooling conditions (2°C) when compared to the in situ temperature group (7°C). Slow-down of tissue growth and protein turnover are typically observed in marine invertebrates at the lower end of their thermal tolerance spectrum, and support energetic homeostasis during seasonal cold exposure (Fraser and Rogers 2007). Less efficiency of translation and protein synthesis at low temperatures are well-documented phenomena in marine invertebrates and fish (Fraser et al. 2002; Storch et al. 2005) from different climatic zones. Under the current summer temperature conditions at the northern WAP, a successful invasion by South American *A. cf eightsii* may still be thermally constrained, a hindrance that likely becomes less effective as future Antarctic summer temperatures will rise more frequently over 2°C. Interestingly, the functional analysis did not indicate upregulation of anaerobic metabolism or stress signals in cold-exposed *Aequiyoldia* from SSA and thus suggest no acute energetic shortage. As the functional analysis did not reveal the same difference in protein synthesis and turnover capacities upon acclimation to 4°C (comparison T_cOx_n :4 °C vs. $T_{is}Ox_n$:7 °C), successful colonization of Antarctic shallows by South American bivalves appears even more likely once these higher temperatures are reached more frequently

in the WAP.

Moderate warming of the Antarctic *A. cf. eightsii* caused a slightly stronger overall response (90 vs. 14 DEG) which might suggest higher susceptibility, but also adaptive competence toward moderate thermal challenge. Upregulation of immune responsive genes accounts for faster growth of bacteria and viruses under warmer conditions, especially in an experimental set-up. It shows that the bivalves were able to enhance immune competence under mild warming and thus respond to one of the major ecological complications in a climate change scenario (Mackenzie et al. 2014). Notably, both congeneric lineages failed to exhibit other typical responses to thermal stress such as the activation of heat shock proteins (HSP) which is commonly observed in many species, including bivalves, under exposure to a wide range of environmental stressors (Kim et al. 2017; Navarro et al. 2020). The absence of a heat shock response (HSR) in our GO analysis is, however, not completely unexpected. An inducible HSR is often missing (Clark and Peck 2009; Clark et al. 2016), or not very pronounced (Koenigstein et al. 2013) in Antarctic marine invertebrates. It is likely that both species have a constitutive expression of heat shock genes sufficient to withstand the thermal stress to which they were exposed in the experiment. Likewise, we found no indication for an oxidative stress response, e.g., upregulation of antioxidant enzymes or metal chaperones under mild thermal challenge without hypoxic stress. In view of the physiological response to warming, as the decrease in antioxidant response at 5 °C (Abele et al. 2001) and the results of the current transcriptome analysis (relatively high number of DEG involving immune responses), 4°C could be the current threshold beyond which temperature in itself results damaging for Antarctic *A. cf. eightsii*. This clearly puts a limit to potential northward invasions under the present conditions.

4.2 The functional response to hypoxia differs between both species

Exposure to hypoxia (10 days at <5% O_2, Fig. 4.1) had more pronounced effects on *A cf. eightsii* gene expression than moderate warming/cooling alone on both sides of Drake Passage (numbers of DEG). Thus, hypoxia represents a severe challenge, irrespective of

the actual temperature change, cooling or warming in SSA as well as WAP. The dominance of apoptosis-related functions, which occurred only in South American *Aequiyoldia* bivalves, emphasizes a greater susceptibility to oxygen deficiency under the conditions prevailing in SSA. The regulation of apoptosis has been suggested to be an emergency mechanism of tolerance under acute oxygenation stress mainly in organisms adapted to frequent oxygen fluctuations (Falfushynska et al. 2020). Apoptosis mitigates tissue inflammation by reducing prevalence of uncontrolled cell death (necrosis), but this mitigation strategy comes at a high physiological cost and inevitably leads to cell loss and functional impairment; so, the extent of tissue apoptosis correlates negatively with the survival of organisms under hypoxia (Falfushynska et al. 2020).

On the other hand, the response to hypoxia of Antarctic *Aequiyoldia* involved the regulation of genes supporting mitochondrial adjustments including the expression of the Alternative oxidase (AOX) under in situ temperature (Fig. 4.2A and Appendix 4.3). The AOX reduces oxygen to water using electrons from the ubiquinol pool when cytochrome oxidase is inhibited by nitric oxide (Strahl and Abele 2020) or sulfide (Völkel and Grieshaber 1996). This alternative electron pathway forms a shortcut of the "modern" (canonical) mitochondrial ETS and causes uncoupling of complexes III and IV. By doing so, it reduces the efficiency of mitochondrial ATP production and flattens mitochondrial inner membrane potential in a state of transient metabolic suspension. The suggested advantage of this mitigation strategy lies in avoiding an oxidative burst response and the associated excess production of damaging ROS during frequent hypoxia-reoxygenation cycling, typical for marine sedimentary habitats (Abele et al. 2017; McDonald and Gospodaryov 2019). Indeed, genes encoding for AOX are mostly inherited in hypoxia-tolerant marine invertebrates ranging from sponges and cnidarians to ascidians, including marine and freshwater bivalves (McDonald et al. 2009). Inducible expression of AOX has been reported in oysters exposed to hypoxia-reoxygenation (Sussarellu et al. 2012) as well as in hypoxia-exposed freshwater bivalves from Andean lakes (Yusseppone et al. 2018). To the best of our knowledge, this is the first time that AOX upregulation is reported for an Antarctic metazoan. The expression of AOX might explain the ecological success of *A. cf. eightsii* in coastal and fjordic systems such as

Potter Cove, where sedimentary conditions vary at local and temporal scale. Sedimentary redox profiles can vary from a few centimetres of oxic surface layer in glacial vicinity where organic carbon content is low, to sub- and anoxic conditions in areas where debris from macroalgal belts is deposited and microbially recycled. Despite this pronounced and unpredictable temporal and spatial variation in oxygen availability *A. cf. eightsii* is an abundant infaunal component (Monien et al. 2014; Pasotti et al. 2015).

The response to hypoxia of the Antarctic *A. cf. eightsii* under warming condition did not involve this protective mitochondrial adjustment (Fig. 4.2B). Instead, it caused a stress response with enhanced metal binding either to reduce oxidative stress originating from cellular transition metal pools (Eteshola et al. 2020), or to increase the capacity of oxygen binding to intracellular haemoglobin in gills or to hemocyanin in the hemolymphatic fluid of *A. cf. eightsii* (Angelini et al. 1998).

Lower environmental temperatures are known to alleviate the hypoxic stress in marine invertebrates by reducing their metabolic requirements, especially if they can enter a state of suspended animation. Hence, although the TOC:TS ratios indicated similar oxygenation conditions in the sediment at both collection sites (suboxic, Appendix 4.1), at present Antarctic specimens still have better chances of survival. It must also not be forgotten that human influence on the coasts in South America (including small and medium cities, livestock and agriculture) is expected to promote a higher variability in oxygen conditions (Levin et al. 2009), and that pollution can have additional adverse effects that endanger sedimentary fauna.

CONCLUSIONS

Our results suggest that a physiological component that has previously prevented a successful transition of South American bivalves to Antarctica is likely to exist and that sustained warming of the WAP will thus increase the invasibility of the Antarctic benthos by species currently residing in South America. The reverse direction of successful long-distance dispersal (i.e., from the Antarctic to South America), while not completely

impossible, is distinctly less likely as far as physiology as a contributing factor is concerned. By the time the warming on the WAP would have lasted long enough to yield adaptations that would increase the likelihood of survival under temperatures prevailing in South America nowadays, the temperatures there, too, would have increased sufficiently to make establishing a successful foothold of Antarctic bivalves north of the Drake physiologically as difficult as it is already now. In Antarctica, *A. cf. eightsii* appear to have some capacity to withstand exposure to hypoxia as an added stressor, e.g., by switching to the alternative oxidase pathway. The induction of AOX in Antarctic *Aequiyoldia* is remarkable, and the rearrangement of mitochondrial ETS branches may indeed be central for bivalves' hypoxia endurance and more generally for their reported ecological success. The high prevalence of cell death-related functions under hypoxia in South American *A. cf. eightsii* suggests that this species may not be amongst the winners of sustained warming there. While it is possible, even likely, that a temperature rise will alter the composition of faunal communities in SSA, too, our results suggest that the winners of such future change will be recruited from the pool of species already populating the shores of South American now on the grounds of them being already adapted to more variable conditions and having a smaller temperature differential to contend with compared to possible invasive species coming across the Drake from the South.

ACKNOWLEDGMENTS

The authors thank to Lars Harms for the bioinformatic support, Uwe John for advice on experimental design, Andrea Eschbach for technical support and Isabell Viehweger for the analysis of sediment samples. We also thank to Instituto Antartico Chileno (INACH), Instituto Antartico Argentino (IAA) and Instituto Antartico Uruguayo (IAU) for the logistic support. This work was supported by National Agency of Investigation and Innovation (ANII, Uruguay) and German Academic Exchange Service (DAAD, Germany) by PhD scholarship to M. Martínez (Grant number POS_EXT_2015_122792), and by Zempelin-Stiftung - Deutsches Stiftungszentrum (Grant number T0214/33883/2019/sm)

DATA ACCESSIBILITY

Raw Illumina reads were deposited in the European nucleotide Archive database (EMBL-EBI) with the accessions ERR4265443 and ERR4276392 – ERR4276460 under the study accession 'Mitochondrial heteroplasmy as a systematic bias in molecular species delimitation and barcoding' (ERP122389).

REFERENCES

Abele D, Brey T, Philipp E. 2017. Ecophysiology of extant marine Bivalvia. Carter J (editor), In: Treatise Online. Paleontological Institute, University of Kansas, Lindley Hall; p. 1-50.

Abele D, Puntarulo S. 2004. Formation of reactive species and induction of antioxidant defence systems in polar and temperate marine invertebrates and fish. Comp. Biochem. Physiol. Part A Mol. Integr. 138(4):405-415.

Abele D, Tesch C, Wencke P, Pörtner HO. 2001. How does oxidative stress relate to thermal tolerance in the Antarctic bivalve *Yoldia eightsi*? Antarct Sci. 13(2):111-118.

Angelini E, Salvato B, Di Muro P, Beltramini M. 1998. Respiratory pigments of *Yoldia eightsi*, an Antarctic bivalve. Mar Biol. 131(1):15-23.

Avila C, Angulo-Preckler C, Martín-Martín RP, Figuerola B, Griffiths HJ, Waller CL. 2020. Invasive marine species discovered on non–native kelp rafts in the warmest Antarctic island. Sci Rep. 10(1):1-9.

Brown A, Thatje S. 2015. The effects of changing climate on faunal depth distributions determine winners and losers. Glob Chang Biol. 21(1):173-180.

Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. Nat Methods. 12(1):59-60.

Cardenas L, Leclerc JC, Bruning P, Garrido I, Detree C, Figueroa A, Astorga M, Navarro JM, Johnson LE, Carlton JT et al. 2020. First mussel settlement observed in Antarctica reveals the potential for future invasions. Sci Rep. 10(1):5552.

Chan FT, Briski E. 2017. An overview of recent research in marine biological invasions. Mar Biol. 164(6):121.

Chown SL, Huiskes AH, Gremmen NJ, Lee JE, Terauds A, Crosbie K, Frenot Y, Hughes KA, Imura S, Kiefer K. 2012. Continent-wide risk assessment for the establishment of nonindigenous species in Antarctica. Proc. Natl. Acad. Sci. 109(13):4938-4943.

Clark MS, Husmann G, Thorne MA, Burns G, Truebano M, Peck LS, Abele D, Philipp EE. 2013. Hypoxia impacts large adults first: consequences in a warming world. Glob Chang Biol. 19(7):2251-2263.

Clark MS, Peck LS. 2009. HSP70 heat shock proteins and environmental stress in Antarctic marine organisms: A mini-review. Mar Genomics. 2(1):11-18.

Clark MS, Sommer U, Sihra JK, Thorne MA, Morley SA, King M, Viant MR, Peck LS. 2016. Biodiversity in marine invertebrate responses to acute warming revealed by a comparative multi-omics approach. Glob Chang Biol. 23(1)318-330

Clem KR, Fogt RL, Turner J, Lintner BR, Marshall GJ, Miller JR, Renwick JA. 2020. Record warming at the South Pole during the past three decades. Nat. Clim. Change. 10(8):762-770.

Deutsch C, Ferrel A, Seibel B, Pörtner H-O, Huey RB. 2015. Climate change tightens a metabolic constraint on marine habitats. Science. 348(6239):1132-1135.

Eteshola EOU, Haupt DA, Koos SI, Siemer LA, Morris DL. 2020. The role of metal ion binding in the antioxidant mechanisms of reduced and oxidized glutathione in metal-mediated oxidative DNA damage. Metallomics. 12(1):79-91.

Falfushynska H, Piontkivska H, Sokolova IM. 2020. Effects of intermittent hypoxia on cell survival and inflammatory responses in the intertidal marine bivalves *Mytilus edulis* and *Crassostrea gigas.* J Exp Biol. 223(4):jeb217026

Fraser CI, Morrison AK, Hogg AM, Macaya EC, van Sebille E, Ryan PG, Padovan A, Jack C, Valdivia N, Waters JM. 2018. Antarctica's ecological isolation will be broken by storm-driven dispersal and warming. Nat. Clim. Change. 8(8):704-708.

Fraser KP, Clarke A, Peck LS. 2002. Low-temperature protein metabolism: seasonal changes in protein synthesis and RNA dynamics in the Antarctic limpet *Nacella concinna* Strebel 1908. J Exp Biol. 205(19):3077-3086.

Fraser KP, Rogers AD. 2007. Protein metabolism in marine animals: the underlying mechanism of growth. Adv Mar Biol. 52:267-362.

González-Wevar C, Gérard K, Rosenfeld S, Saucède T, Naretto J, Díaz A, Morley SA, Brickle P, Poulin E. 2019. Cryptic speciation in Southern Ocean *Aequiyoldia eightsii* (Jay, 1839): Mio-Pliocene trans-Drake Passage separation and diversification. Prog Oceanogr. 174:44-54.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol. 29(7):644.

Hughes KA, Pertierra LR, Molina-Montenegro MA, Convey P. 2015. Biological invasions in terrestrial Antarctica: what is the current status and can we respond? Biodivers Conserv. 24(5):1031-1055.

Johnson KM, Hofmann GE. 2020. Combined stress of ocean acidification and warming influence survival and drives differential gene expression patterns in the Antarctic pteropod, *Limacina helicina* antarctica. Conserv Physiol. 8(1):coaa013.

Kim B-M, Kim K, Choi I-Y, Rhee J-S. 2017. Transcriptome response of the Pacific oyster, *Crassostrea gigas* susceptible to thermal stress: A comparison with the response of tolerant oyster. Mol. Cell. Toxicol. 13(1):105-113.

Koenigstein S, Pohlmann K, Held C, Abele D. 2013. Ecological comparison of cellular stress responses among populations - normalizing RT-qPCR values to investigate differential environmental adaptations. BMC Ecol. 13:21.

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10(3):R25.

Levin L, Ekau W, Gooday A, Jorissen F, Middelburg J, Naqvi S, Neira C, Rabalais N, Zhang J. 2009. Effects of natural and human-induced hypoxia on coastal benthos. Biogeosciences, 6(10):2063–2098

Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics. 12(1):323.

Li G, Cheng L, Zhu J, Trenberth KE, Mann ME, Abraham JP. 2020. Increasing ocean stratification over the past half-century. Nat. Clim. Change. 10(12):1116-1123.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15(12):550.

Mackenzie CL, Lynch SA, Culloty SC, Malham SK. 2014. Future oceanic warming and acidification alter immune response and disease status in a commercial shellfish species, *Mytilus edulis* L. PLoS One. 9(6):e99712.

McCarthy AH, Peck LS, Hughes KA, Aldridge DC. 2019. Antarctica: The final frontier for marine biological invasions. Glob Chang Biol. 25(7):2221-2241.

McDonald AE, Gospodaryov DV. 2019. Alternative NAD(P)H dehydrogenase and alternative oxidase: Proposed physiological roles in animals. Mitochondrion. 45:7-17.

McDonald AE, Vanlerberghe GC, Staples JF. 2009. Alternative oxidase in animals: unique characteristics and taxonomic distribution. J Exp Biol. 212(16):2627-2634.

Monien P, Lettmann KA, Monien D, Asendorf S, Wölfl A-C, Lim CH, Thal J, Schnetger B, Brumsack H-J. 2014. Redox conditions and trace metal cycling in coastal sediments from the maritime Antarctic. Geochim Cosmochim Acta. 141:26-44.

Muñoz-Ramírez CP, Barnes DK, Cárdenas L, Meredith MP, Morley SA, Roman-Gonzalez A, Sands CJ, Scourse J, Brante A. 2020. Gene flow in the Antarctic bivalve *Aequiyoldia eightsii* (Jay, 1839) suggests a role for the Antarctic Peninsula Coastal Current in larval dispersal. R. Soc. Open Sci. 7(9):200603.

Navarro JM, Detree C, Morley SA, Cardenas L, Ortiz A, Vargas-Chacoff L, Paschke K, Gallardo P, Guillemin ML, Gonzalez-Wevar C. 2020. Evaluating the effects of ocean

warming and freshening on the physiological energetics and transcriptomic response of the Antarctic limpet *Nacella concinna*. Sci Total Environ. 748:142448.

Pasotti F, Manini E, Giovannelli D, Wölfl A-C, Monien D, Verleyen E, Braeckman U, Abele D, Vanreusel A. 2015. Antarctic shallow water benthos in an area of recent rapid glacier retreat. Mar Ecol. 36(3):716-733.

Peck L, Morley S, Clark M. 2010. Poor acclimation capacities in Antarctic marine ectotherms. Mar Biol. 157(9):2051-2059.

Poloczanska ES, Brown CJ, Sydeman WJ, Kiessling W, Schoeman DS, Moore PJ, Brander K, Bruno JF, Buckley LB, Burrows MT. 2013. Global imprint of climate change on marine life. Nat. Clim. Change. 3(10):919-925.

Sampaio E, Santos C, Rosa IC, Ferreira V, Portner HO, Duarte CM, Levin LA, Rosa R. 2021. Impacts of hypoxic events surpass those of future ocean warming and acidification. Nat Ecol Evol. 5:311-321

Shepherd JG, Brewer PG, Oschlies A, Watson AJ. 2017. Ocean ventilation and deoxygenation in a warming world: introduction and overview. Philos. Trans. R. Soc. A. 375:2102

Storch D, Lannig G, Portner HO. 2005. Temperature-dependent protein synthesis capacities in Antarctic and temperate (North Sea) fish (Zoarcidae). J Exp Biol. 208(12):2409-2420.

Strahl J, Abele D. 2020. Nitric oxide mediates metabolic functions in the bivalve *Arctica islandica* under hypoxia. PLoS One. 15(5):e0232360.

Sussarellu R, Fabioux C, Sanchez MC, Le Goïc N, Lambert C, Soudant P, Moraga D. 2012. Molecular and cellular response to short-term oxygen variations in the Pacific oyster *Crassostrea gigas*. J Exp Mar Biol Ecol. 412:87-95.

Tepolt CK, Palumbi SR. 2015. Transcriptome sequencing reveals both neutral and adaptive genome dynamics in a marine invader. Mol Ecol. 24(16):4145-4158.

Togunwa OS, Abdullah WH. 2017. Geochemical characterization of Neogene sediments from onshore West Baram Delta Province, Sarawak: paleoenvironment, source input and thermal maturity. Open Geosci. 9(1):302-313.

Völkel S, Grieshaber MK. 1996. Mitochondrial sulfide oxidation in *Arenicola marina*: evidence for alternative electron pathways. Eur J Biochem. 235(1-2):231-237.

Weihe E, Kriews M, Abele D. 2010. Differences in heavy metal concentrations and in the response of the antioxidant system to hypoxia and air exposure in the Antarctic limpet *Nacella concinna*. Mar Environ Res. 69(3):127-135.

Young MD, Wakefield MJ, Smyth GK, Oshlack A. 2010. Gene ontology analysis for RNA-seq: accounting for selection bias. Genome Biol. 11(2):R14.

Yusseppone MS, Rocchetta I, Sabatini SE, Luquet CM, Rios de Molina MDC, Held C, Abele D. 2018. Inducing the Alternative Oxidase Forms Part of the Molecular Strategy of Anoxic Survival in Freshwater Bivalves. Front Physiol. 9:100.

APPENDIXES

Appendix 4.1. Environmental data registered at the time of animal collection in West Antarctic Peninsula (WAP) and Southern South America (SSA): depth, bottom salinity and temperature, total organic carbon (TOC) and total sulphur (TS) of sediment, and TOC: TS ratio. At both collection sites the ratio TOC: TS indicated suboxic conditions in the sediment (Togunwa and Abdullah 2017).

	Depth (m)	Salinity	Temperature (°C)	TOC (%)	TS (%)	TOC: TS
WAP	5.88	33.80	1.67	0.79	0.25	3.16
SSA	1.50	32.00	7.00	0.95	0.27	3.52

Appendix 4.2. Complete list of GO enriched terms resulting from each pairwise comparisons in a West Antarctic Peninsula (WAP) southern South America (SSA) experiments. The first column indicates the treatment in which the corresponding GO term resulted enriched. The column 'numDEInCat' indicates the number of differentially expressed genes in the category.

WAP experiment

$T_{in}Ox_{hip}$ vs. $T_{in}Ox_{n}$

Enrich. in treatment:	category	numDEInCat	term
$T_{in}Ox_{hip}$	GO:0005739	4	mitochondrion
$T_{in}Ox_{hip}$	GO:0006357	3	regulation of transcription by RNA polymerase II
$T_{in}Ox_{hip}$	GO:0042025	3	host cell nucleus
$T_{in}Ox_{n}$	GO:0005856	1	cytoskeleton

$T_{in}Ox_{hip}$ vs. $T_{w}Ox_{hip}$

Enrich. in treatment:	category	numDEInCat	term
$T_{in}Ox_{hip}$	GO:0046982	2	protein heterodimerization activity
$T_{in}Ox_{hip}$	GO:0003197	1	endocardial cushion development
$T_{in}Ox_{hip}$	GO:0007162	1	negative regulation of cell adhesion
$T_{in}Ox_{hip}$	GO:0007422	1	peripheral nervous system development
$T_{in}Ox_{hip}$	GO:0009968	1	negative regulation of signal transduction
$T_{in}Ox_{hip}$	GO:0014037	1	Schwann cell differentiation
$T_{in}Ox_{hip}$	GO:0014065	1	phosphatidylinositol 3-kinase signaling
$T_{in}Ox_{hip}$	GO:0016328	1	lateral plasma membrane
$T_{in}Ox_{hip}$	GO:0019838	1	growth factor binding

$T_{in}Ox_{hip}$	GO:0021545	1	cranial nerve development
T_wOx_{hip}	GO:0007160	2	cell-matrix adhesion
T_wOx_{hip}	GO:0005112	1	Notch binding
T_wOx_{hip}	GO:0005576	2	extracellular region
T_wOx_{hip}	GO:0003180	1	aortic valve morphogenesis
T_wOx_{hip}	GO:0005581	1	collagen trimer
T_wOx_{hip}	GO:0010629	1	negative regulation of gene expression
T_wOx_{hip}	GO:0010811	1	positive regulation of cell-substrate adhesion
T_wOx_{hip}	GO:0032966	1	negative regulation of collagen biosynthetic process
T_wOx_{hip}	GO:0034668	1	integrin alpha4-beta1 complex
T_wOx_{hip}	GO:0050866	1	negative regulation of cell activation

$T_{in}Ox_n$ vs. T_wOx_n

Enrich. in treatment:	category	numDEInCat	term
T_wOx_n	GO:0002281	1	macrophage activation involved in immune response
T_wOx_n	GO:0003823	1	antigen binding
T_wOx_n	GO:0030277	1	maintenance of gastrointestinal epithelium
T_wOx_n	GO:0030299	1	intestinal cholesterol absorption
T_wOx_n	GO:0030308	1	negative regulation of cell growth
T_wOx_n	GO:0042632	1	cholesterol homeostasis
T_wOx_n	GO:0046790	1	virion binding
T_wOx_n	GO:0099512	1	supramolecular fiber
T_wOx_n	GO:0043231	1	intracellular membrane-bounded organelle

T_wOx_{hip} vs. T_wOx_n

Enrich. in treatment:	category	numDEInCat	term
T_wOx_{hip}	GO:0046872	5	metal ion binding
T_wOx_{hip}	GO:0004222	2	metalloendopeptidase activity
T_wOx_{hip}	GO:0003677	4	DNA binding
T_wOx_n	GO:0000977	1	RNA polymerase II regulatory region DNA binding
T_wOx_n	GO:0002376	1	immune system process
T_wOx_n	GO:0005925	1	focal adhesion
T_wOx_n	GO:0006355	1	regulation of transcription, DNA-templated
T_wOx_n	GO:0007596	1	blood coagulation
T_wOx_n	GO:0051607	1	defense response to virus
T_wOx_n	GO:0060333	1	interferon-gamma-mediated signaling pathway
T_wOx_n	GO:0060337	1	type I interferon signaling pathway

SSA experiment

T_cOx_{hip} vs. T_cOx_n

Enrich. in treatment:	category	numDEInCat	term
T_cOx_{hip}	GO:0000122	4	negative regulation of transcription by RNA polymerase II
T_cOx_{hip}	GO:0031398	3	positive regulation of protein ubiquitination
T_cOx_{hip}	GO:0043027	3	endopeptidase inhibitor activity involved in apoptotic process
T_cOx_{hip}	GO:0043154	3	negative regulation of endopeptidase involved in apoptotic process
T_cOx_{hip}	GO:0061630	3	ubiquitin protein ligase activity
T_cOx_{hip}	GO:1990001	3	inhibition of endopeptidase activity involved in apoptotic process
T_cOx_{hip}	GO:0043066	4	negative regulation of apoptotic process
T_cOx_{hip}	GO:0005634	8	nucleus

$T_{in}Ox_n$ vs. T_cOx_{hip}

Enrich. in treatment:	category	numDEInCat	term
T_cOx_{hip}	GO:0045944	4	positive regulation of transcription by RNA polymerase II
T_cOx_{hip}	GO:0000122	4	negative regulation of transcription by RNA polymerase II
T_cOx_{hip}	GO:0005667	3	transcription factor complex
$T_{in}Ox_n$	GO:0016021	3	integral component of membrane

$T_{in}Ox_h$ vs. T_cOx_n

Enrich. in treatment:	category	numDEInCat	term
T_cOx_n	GO:0006313	2	transposition, DNA-mediated
T_cOx_n	GO:0016324	2	apical plasma membrane
$T_{in}Ox_h$	GO:0031398	3	positive regulation of protein ubiquitination
$T_{in}Ox_h$	GO:0043027	3	endopeptidase inhibitor activity involved in apoptotic process
$T_{in}Ox_h$	GO:0043154	3	negative regulation of endopeptidase involved in apoptotic process
$T_{in}Ox_h$	GO:0061630	3	ubiquitin protein ligase activity
$T_{in}Ox_h$	GO:1990001	3	inhibition of endopeptidase involved in apoptotic process
$T_{in}Ox_h$	GO:0019899	3	enzyme binding

$T_{in}Ox_n$ vs. T_cOx_n

Enrich. in treatment:	category	numDEInCat	term
T_cOx_n	GO:0001674	1	female germ cell nucleus
T_cOx_n	GO:0005534	1	galactose binding
T_cOx_n	GO:0005623	1	cell
T_cOx_n	GO:0033296	1	rhamnose binding
T_cOx_n	GO:0042564	1	NLS-dependent protein nuclear import complex
T_cOx_n	GO:0060473	1	cortical granule
T_cOx_n	GO:1903777	1	melibiose binding

T_cOx_n	GO:0042803	1	protein homodimerization activity
T_cOx_n	GO:0005524	1	ATP binding
$T_{in}Ox_n$	GO:0005506	1	iron ion binding
$T_{in}Ox_n$	GO:0045329	1	carnitine biosynthetic process

$T_{in}Ox_n$ vs. T_cOx_{hip}

Enrich. in treatment:	category	numDEInCat	term
T_cOx_{hip}	GO:0031398	3	positive regulation of protein ubiquitination
T_cOx_{hip}	GO:0043027	3	endopeptidase inhibitor activity involved in apoptotic process
T_cOx_{hip}	GO:0043154	3	negative regulation of endopeptidase involved in apoptotic process
T_cOx_{hip}	GO:0061630	3	ubiquitin protein ligase activity
T_cOx_{hip}	GO:1990001	3	inhibition of endopeptidase activity involved in apoptotic process
T_cOx_{hip}	GO:0043066	4	negative regulation of apoptotic process
$T_{in}Ox_n$	GO:0003924	2	GTPase activity

Appendix 4.3. List of differentially expressed genes (DEG) supporting the most significant enriched GO terms in hypoxia vs. Normoxia treatments at in situ and warming conditions in the experiment performed in West Antarctic Peninsula

$T_{is}Ox_n$ vs. $T_{is}Ox_{hyp}$

GO: Mitochondrion
ETS domain-containing protein Elk-1
Haloacid dehalogenase-like hydrolase domain-containing 5
Alternative oxidase, mitochondrial
AAC-rich mRNA clone AAC4 protein

T_wOx_n vs. T_wOx_{hyp}

GO: metal ion binding
Elastase
Cytochrome b
Transcription initiation factor TFIID subunit 3
Neprilysin-1

GO: metal ion binding
Elastase
Neprilysin-1

CHAPTER 5

5 Synoptic discussion

The main objective of this thesis was to explore genetic and phenotypic traits that support the adaptation of bivalves from both sides of the Drake Passage to environmental changes related to global warming, by evaluating patterns of gene expression/transcription in *Aequiyoldia cf. eightsii*. Transcription is the key transition step between genotype and phenotype, so this approach has the fundamental advantage of looking at both sides, genotype (partially) and 'molecular phenotype' (transcriptome). **Chapter 2** looked exclusively at the genotype (albeit mostly based on coding DNA data) of populations/species on both sides of the Drake passage, to evaluate their genetic diversity within and between sites and to elucidate the possibility of cryptic speciation between continents and within Antarctic populations. One of the major outcomes of Chapter 2 is how the remarkable coexistence of two types of mitochondria can systematically mislead widely used techniques of taxonomic delimitation and identification of species and how these shortcomings can be overcome. Furthermore, this chapter also made a substantial contribution clarifying the patterns of genetic divergence of *Aequiyoldia* in the Southern Ocean. Based on a comprehensive dataset of mitochondrial and nuclear SNPs, it was **(i)** demonstrated that it is very likely that populations on both sides of the Drake are two reproductively isolated species, and **(ii)** refuted the previous suggestions of cryptic speciation in *A. cf. eightsii* in the Antarctic Peninsula, showing instead that the observed diversification pattern denotes high mitochondrial genetic diversity within a single species, with occurrence of mitochondrial heteroplasmy.

These two findings had a big impact in the direction of the following Chapters. The logical line of this thesis continued addressing the following questions in **Chapter 3** in a

small-scale approach based on in situ gene expression patterns: Is the genetic diversity, including the peculiar pattern of mitochondrial inheritance (heteroplasmy), a promoter of intra-population plasticity in Antarctic organisms? In addition, is this population capable of developing plastic responses to environmental heterogeneity shaped by the melting of a glacier in Potter Cove? In answering these questions, I found that gene expression patterns in Potter Cove population were modulated by three independent underlying causes: microclimatic heterogeneity (kilometre scale), mitotype (especially the admixture of a second mitochondrial genotype), and nuclear genotype. Therefore, the results of this chapter show that Antarctic *Aequiyoldia* are not only characterised by a high genetic diversity (mainly mitochondrial as demonstrated in Chapter 2) but also by a transcriptional plasticity shaped by intraspecific genetic variation and by the environment at fine-scale, both features not usually expected in Antarctic organisms. It remains to be explored which further environmental parameters shape local adaptation (differentially expressed genes) in Antarctic *Aequiyoldia*, as extrapolation from controlled laboratory conditions (Chapter 4) to more parameter-rich and complex natural environments is naturally limited.

In **Chapter 4**, I explored whether the genetic differentiation in isolation between *A. cf. eightsii* from both sides of the Drake, which very likely represent two reproductively isolated species as shown in Chapter 2, is accompanied by adaptive differences in the responses to changes in temperature and oxygen regimens; and whether future thermal scenarios could allow cross-continental invasions of *Aequiyoldia* between SSA and WAP. This was addressed by an experimental approach which allowed to establish causal relationships between 'stressors' and transcriptional responses. Among the most relevant, I found WAP bivalves upregulating mitochondrial functions (including the enzyme Alternative oxidase, AOX) under hypoxia, whilst SSA organisms mostly developed apoptotic-related responses to the same lack of oxygen. These physiological responses to hypoxia and those observed in response to temperature shifts (and the combination of both stressors) revealed an ability of the WAP *Aequiyoldia* to cope with mild warming and oxygen deficiency expected in future climate change scenarios, and a greater susceptibility of SSA *Aequiyoldia* which would be operating close to its

physiological limit but would be physiologically capable of a poleward migration to a warmed WAP.

This chapter (**Chapter 5**) presents a synoptic discussion that summarizes the three main chapters, integrating the two main pieces of information gathered in this thesis: **genetic** and **transcriptional** diversity of the Southern Ocean *Aequiyoldia*. In the first section of this synoptic discussion (5.1), I will discuss whether the molecular phenotype (or transcriptional profile) can reinforce the idea of cryptic speciation between *A. cf. eightsii* on both sides of the Drake (5.1.1). For this, I will include results of differential gene expression analysis between SSA and WAP organisms (in situ gene expression) not previously presented along this thesis, further contributing to the discussion of adaptive differences between the two species addressed in Chapter 4 (5.1.2). Second, based on the most relevant genetic and phenotypic findings, and information available in the literature on the species, I will discuss whether *Aequiyoldia* from WAP and SSA will be among the "winners" or "losers" in the expected climate change scenarios (5.2). The last section (5.3) will present the most remarkable conclusions of this thesis and the perspectives or open questions for future research.

5.1 Gene expression patterns: product of adaptive and neutral evolution

Chapter 2 concludes that *Aequiyoldia* populations on both sides of the Drake should be considered separate species. This conclusion is based exclusively on genetic evidence. So far, no morphological characters have been described to differentiate these species, other than their size difference (Chapter 2, Appendix 2.1) - a fact that is commonly observed when comparing populations/sibling species from different climatic regions (Linse et al. 2006; Pöhlmann et al. 2011). Morphological differences may be a response to environmental factors but also, they can be attributed to accumulated genetic differences due to disruption of gene flow among populations during the process of speciation (Prada et al. 2008). The same is true for gene expression: it is shaped by both genetic and environmental components, and therefore it can be considered as a "molecular phenotype" upon which natural selection can act (Pavey et al. 2010).

The molecular phenotype differed between *Aequiyoldia* from both side of the Drake Passage when exposed to changes in regimens of temperature and oxygen (Chapter 4). A direct comparison of gene expression profiles between *Aequiyoldia* from both continents performed between organisms sampled at SSA (Punta arenas, 53°37'52"S; 70°56'54"W) and WAP (Potter Cove, 62°14'11"S; 58°40'14"W), which was not included in Chapters 2-4, is shown in the following. Samples were dissected and preserved within two hours after collection (5 organisms or replicates per site), so that it can be considered as in situ gene expression. The methodology employed at each step from collection to differential gene expression analysis and GO enrichment analysis is exactly the same as the described in Chapter 3 and 4. The analysis resulted in an astonishing number of differentially expressed genes (DEG) between organisms on both sides of the Drake (ca. 7000 DEG, see Fig. 5.1). This number was at least one order of magnitude greater than any of the intra-species/continent comparison in Chapter 3 and 4. The Appendix 5.1 shows the results of the GO enrichment analysis. A total of 149 GO terms where enriched, of which 140 were enriched in SSA and only 9 in WAP. While the WAP-enriched terms consisted of a low number of DEG involving very specific functions, the enriched GO terms in SSA were supported by a large number of DEG and consisted of a wide variety of central cellular functions and organelles. These results will not be discussed extensively and instead I will focus on the evidence that contributes to answering pending questions from the previous chapters.

In view of the results, it is worth asking: could this high difference in gene expression reinforce the idea that populations on both sides of the Drake correspond to separate species? Are among them hallmarks of adaptation to their climatic regions (e.g., thermal regimes) not captured in Chapter 4?

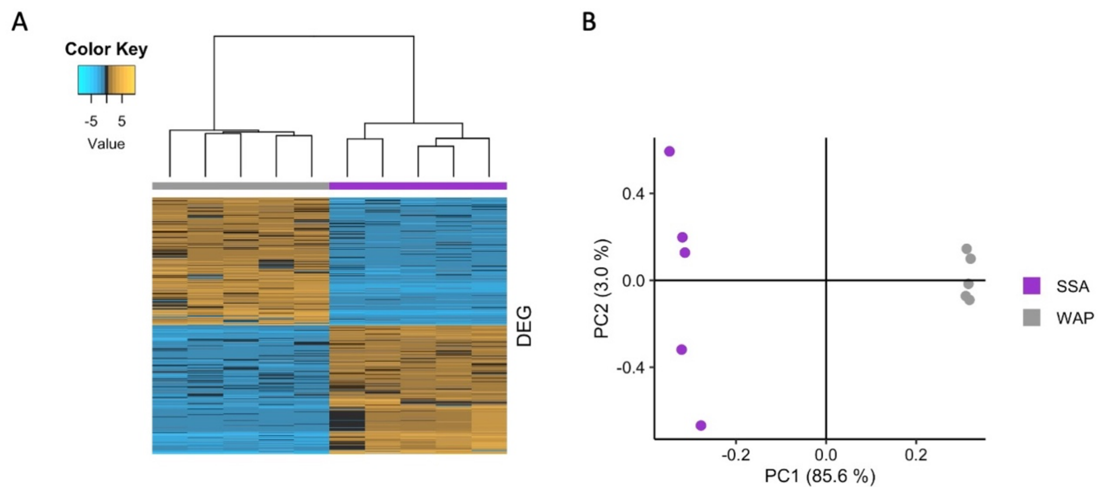


Figure 5.1. A: Heatmap showing differentially expressed genes (DEG) between *Aequiyoldia cf. eightsii* from West Antarctic Peninsula (WAP, grey) and Southern South America (SSA, violet). A cluster analysis showing similarities between samples/replicates is included on top. DEG are displayed in rows and counts per sample are displayed in columns. Genes included were significant ($p < 0.001$) in the pairwise comparison. A colour code indicates up-regulated and down-regulated expression levels (top left). B: Principal Component Analysis based on counts of DEG between *A. cf. eightsii* from WAP and SSA.

5.1.1 The molecular phenotype in phylogenetic inferences

Whether the molecular phenotype correlates phylogenetic patterns, i.e., if there is a correlation between molecular phenotype differentiation and genetic divergence, is still an open question (Khaitovich et al. 2004; Uebbing et al. 2016; Yang et al. 2017). The answer to this question lies in the relative importance between the evolutionary forces that govern the variance of traits between taxa. At the genotype level, most intra and interspecific variations in DNA sequences are neutral or nearly neutral (Kimura 1968; Nei 1987). A smaller proportion of genetic variation involves changes at the phenotypic level with implications for fitness and therefore subjected to selective pressures and adaptive evolution. Therefore, as selection acts at the level of the phenotype while variation is generated at the level of the genotype, the proportion of changes caused by selection can be expected to be largest at the phenotypic level and smallest at the DNA

sequence level (Khaitovich et al. 2004). If a genotypic change has a potential phenotypic effect (with or without implications in fitness), its realization requires gene expression acting as bridge between genotypes and their corresponding organismal phenotypes. In this context, the discussion has focused on whether gene expression level, a molecular phenotype, is more like (organismal) phenotypes or (molecular) genotypes in its evolutionary pattern and mechanism. Specifically, if variations in gene expression levels within and between species result largely from neutral or adaptive evolution (Yang et al. 2017).

Under a neutral drift model, the variation in a trait is a function of genetic distance and has little ecological implications: traits will be more similar among closely related taxa than among more distantly related taxa, regardless the environmental context. Under natural selection pressure, the variation in a trait affects the fitness of organisms and is a function of the ecological setting: traits are conserved or diverge depending on the specific ecological pressures (Whitehead and Crawford 2006). Previous studies suggest that much of the extensive variation in gene expression among individuals and taxa is simply random neutral divergence (Khaitovich et al. 2004; Yanai et al. 2004; Yang et al. 2017), whereas others emphasize the importance of selection, although still smaller than that expected by neutrality (Rifkin et al. 2003). An example of the latter was found by Whitehead and Crawford (2006) in the fish *Fundulus heteroclitus* in which the variation in expression for 22% of the genes correlates with habitat temperature but much of the variation in gene expression fits a null model of neutral drift. In yeast, Yang et al. (2017) found that transcriptome-based clustering of nine strains approximates the genome sequence-based phylogeny irrespective of their ecological environments. Only 0.5 % of genes exhibited similar expression levels among strains from a common ecological environment, no greater than that among strains with comparable phylogenetic relationships but different environments. The authors suggested that most intra- and interspecific variations in yeast gene expression levels result from the accumulation of random mutations rather than environmental adaptations.

Although environmental and physiological stimuli are clearly responsible for changes in the expression levels of many genes, a big amount of evidence shows that the majority

of changes of gene expression fixed during evolution between and within species are caused mainly by stochastic processes rather than Darwinian selection. Answering the initial question posed in this section, this fact implies that gene expression can be used as a proxy of evolutionary distance; therefore, the high number of DEG found in *Aequiyoldia* between continents (ca. 7000 DEG) compared to that found within Antarctic population (between mitotypes: ca. 600 DEG, Chapter 3) add an extra piece of evidence towards the conclusion that the populations of *Aequiyoldia* on either side of the Drake Passage correspond to distinct species.

5.1.2 Climate adaptation hallmarks of the molecular phenotype

Among the genes whose expression is expected to be strongly determined by environmental and physiological stimuli are the heat shock genes coding for molecular chaperones, which help organisms to cope with internal and external stress sources (Gross 2004). Due to its role mediating thermal tolerance, the expression of heat shock proteins (HSP) is expected to feature prominently in local adaptation to climate (Bedulina et al. 2010; Advani et al. 2016). Their induction and utility in Antarctic marine organisms is of great interest, given climate change predictions (Clark and Peck 2009). In Chapter 4, however, no induction of heat shock response (HSR) was observed under thermal stress and hypoxia, neither in SSA nor WAP bivalves. I speculated that the absence of HSR may be due to the fact that these organisms possess constitutive expression levels that are sufficient to cope with the stress induced in the experiments, which is why they do not appear in the output of differentially expressed genes. The absence of an intercontinental comparison in gene expression profiles also left open the question of whether such constitutive expression (if any) is different between polar and cold-temperate *Aequiyoldia*. This has been observed in other species/populations in which organisms from temperate regions tend to have higher HSR than those from higher latitudes due to its natural exposure to higher temperature fluctuations (Pöhlmann et al. 2011; Koenigstein et al. 2013). The intercontinental comparison presented in this section resulted in the differential expression of transcripts annotated

as heat shock proteins (transcript count higher than zero in both species), and the GO enrichment analysis (Appendix 5.1) resulted in the term 'heat shock protein binding' enriched in SSA organisms. These results reinforce what was suggested in Chapter 4 regarding the constitutive expression of HSP, and is in line with previous studies in which Antarctic organisms develop lower HSR for living in a more thermally stable environment. Moreover, these results are relevant in the context that previous studies had not found the expression of HSP in the Antarctic *Aequiyoldia* even when this species was exposed to temperatures of 25°C (Clark et al. 2016). Lastly, the enrichment of a wide variety of GO terms related to central cellular functions, including responses to stress (e.g., responses to oxidative stress) in South American bivalves compared to Antarctic organisms might indicate higher investments into cellular stress compensation which is a typical consequence of warmer climate and restricts growth in ectotherms (Daufresne et al. 2009; Pöhlmann et al. 2011); this is in concordance with differences in size found between the two species, with Antarctic *Aequiyoldia* being significantly larger than South American *Aequiyoldia*.

5.2 The fate of *A. cf. eightsii* in a warming Southern Ocean

In order to predict potential changes in biodiversity patterns as a consequence of climate change, the scientific community has made a great effort to identify species as either climate change 'winners': species that remain stable or expand during these events, or climate change 'losers': species that decline in abundance and distribution or go extinct (Clucas et al. 2014; Brown and Thatje 2015; Griffiths et al. 2017). Many of these contributions involved projections including hundreds of species based on ranges of distribution and thermal tolerances (Griffiths et al. 2017). However, forecasts of species vulnerability to climate change usually do not consider local adaptations despite their importance for determining the potential of populations to respond to future changes (Razgour et al. 2019). Such model improvements are limited by the paucity of observational and experimental studies of local climatic adaptations (Valladares et al. 2014). Another aspect that hampers the accuracy of predictions is the fact that most of

these approaches do not consider the potential synergetic effects of variables that are expected to co-vary with rising temperatures such as oxygen and pH (but see Deutsch et al. (2015)). In some marine ectotherms, physiological stress resulting from warming is intensified in combination with hypoxia (Pörtner 2005). Unfortunately, the frequency of hypoxic/anoxic events is predicted to rise in the course of climate change (Shepherd et al. 2017). Thus, while incorporating information on local adaptations could decrease predicted range loss of species (i.e., lead to less catastrophic predictions), the inclusion of other variables with potential synergetic effects with temperature could result in more pessimistic predictions for the future of species.

Due to its stenothermality, these approximations are especially negative for Antarctic fauna for which a significant reduction in suitable temperature habitat is predicted within the current century. This results in more losers than winners in future warming scenarios of the Southern Ocean (Griffiths et al. 2017). So, having considered synergetic effects of temperature and oxygen and local adaptation, can the results of this thesis suggest whether the Southern Ocean *Aequiyoldia* will be among the winners or losers? Will both species suffer the same fate? The forecast based on information gathered throughout this thesis is not the same for the two species on both continents: the Antarctic species may come out as an eventual winner and the SSA species compromised by synergistic effects of temperature and hypoxia, but with the possibility of poleward migration.

The arguments for the above conclusion with respect to Antarctic organisms are based on the following findings: **(i)** a high genetic diversity at the mitochondrial level that would confer transcriptional plasticity (Chapter 2 and 3), **(ii)** transcriptional plasticity in response to environmental heterogeneity shaped by glacier melting (Chapter 3), **(iii)** a relatively moderate response to warming as a single stressor and the absence of a critical physiological state in response to elevated levels of hypoxia (as did occur in SSA organisms), and instead the upregulation of the AOX enzyme (Chapter 4).

The first argument is, however, the most speculative. Genetic diversity is usually considered as key to promoting population survival and guaranteeing the adaptive potential of populations in the face of rapidly environmental change (Markert et al.

2010). However, genetic diversity per se without associations between genotypes and potential adaptive-phenotypes cannot be conclusive regarding the fate of species/populations in a future changing scenario. In this sense, models including phylogeographic structure based on neutral markers have resulted in more pessimistic forecasts than traditional (without intra-specific variation). These attempts are limited in scope because neutral markers provide information on the species' evolutionary history and barriers to gene flow but not on the adaptive value of the existing genotypes (Razgour et al. 2019). For model species of plants and *Drosophila*, it has been shown that models which incorporate information on genetic structure and local adaptations obtained from common garden experiments are more optimistic in future range projections (Benito Garzón et al. 2011; Bush et al. 2016).

The transcriptional differences based on mitochondrial genetic diversity (i.e., mitotypes) observed in Chapter 3, might support an adaptive value of such genetic diversity. These transcriptional differences mainly involved mitochondrial respiration functions. In Chapter 3, I hypothesised that intra-population differences in mitochondrial respiratory efficiency might confer population-wide metabolic plasticity that enables the animals to distribute across micro-habitats differing in elemental factors such as sediment composition, food availability and redox environment. However, both the small-scale in situ approach (spatial heterogeneity in Potter cove) and the experimental approach did not assess the existence of differential responses of genotypes to the environment (G x E interactions), so a mechanistic explanation of the persistence of such genetic diversity and its adaptive implications is still required. Therefore, whether this genetic diversity, including heteroplasmy, will confer an advantage to *Aequiyoldia* in future climate change scenarios in the WAP remains speculative.

The observed transcriptional diversity of *A. cf. eightsii* in response to environmental heterogeneity in Potter Cove -mostly associated with glacial melting processes (Chapter 3)- reveals a great capacity for sensing and responding to the environment which might confer an advantage in the face of an accelerated retreat of glaciers along the WAP. Previous studies had already reported phenotypic differences of *A. cf. eightsii* in response to glacier influence in Potter cove (spatial heterogeneity): organisms located

in areas exposed to iceberg impact are significantly smaller than those located in more protected areas (Brown, Fraser, et al. 2004; Pasotti et al. 2015). In addition to iceberg impacts, sedimentation related to glacier melt is another variable with a heterogeneous incidence along glacial coves and a source of stress for many pelagic and benthonic groups. It has been pointed out that *A. cf. eightsii* is well adapted to high sedimentation rate which make this species capable of outcompeting the larger bivalve *Laternula elliptica* and dominate disturbed sites in glacial coves (Siciński et al. 2011). Thus, *Aequiyoldia*'s ability to inhabit a wide diversity of habitats in an Antarctic fjord, from near-glacial environments with high levels of sedimentation to rocky areas with sediment pits and macroalgae accumulation involving low oxygen levels, is probably, in part, a consequence of its high transcriptional plasticity, i.e., due to a coupling between niche features and traits expression. An example of this coupling observed in this thesis is the differential expression of genes associated with biological functions such as 'lactate oxidation' and 'response to hypoxia' when stations with different oxygenation conditions were compared (Base vs Faro and Glacier vs. Faro). However, as discussed in Chapter 3, the interpretations derived from single genes or GO terms, should be taken with caution and ideally be validated by an experimental approach and GO enrichment analysis in the future.

The experimental counterpart of this thesis also yielded results that suggest a good performance of the Antarctic *Aequiyoldia* under the predicted climate change scenarios for the Southern Ocean. As discussed in section 5.1.2, the transcriptional response to warming (moderate-low in terms of DEG number) may have been dampened by a constitutive expression of heat shock proteins that prevented other manifestations of physiological stress. On the other hand, the expression of the enzyme AOX under hypoxia is by far one of the outstanding results of Chapter 4, as it is the first time it has been observed in an Antarctic organism. The alternative transfer of electrons via AOX, which involves a metabolic depression, has been mainly identified in organisms sharing the following characteristics: sessile or slow-moving organisms (Matus-Ortega et al. 2011) and highly tolerant to certain stresses, including hypoxia, hypersalinity, dehydration, extreme heat or cold, among others (McDonald and Gospodaryov 2019).

For instance, many bivalves expressing the AOX, including *Arctica islandica*, *Crassostrea gigas*, and *Mytilus edulis*, exhibit oxyconformity, an ability to rapidly acclimate to changes in partial oxygen pressure (McDonald et al. 2009; Sussarellu et al. 2012; Abele et al. 2017). A reduced aerobic activity of an animal at low environmental pO_2 prevent, or at least delay, early onset of fatal tissue hypoxia and cause a reduction of mitochondrial free radicals formation (Abele et al. 2007). The AOX has been pointed to be responsible for oxyconformity and more generally for the stress endurance of sessile benthic fauna, and therefore as a key trait for their response and adaptation to climate change (McDonald and Gospodaryov 2019). The absence of inducible AOX expression in response to hypoxia under warming experimental conditions in the Antarctic *Aequiyoldia* leaves open, however, whether this enzyme will contribute to protect this species during a sustained warming of the Antarctic continent.

While the assessment of individual responses of organisms to environmental stress is highly informative for predicting the future of species in the face of climate change, global warming is also expected to disrupt biological interactions resulting in unforeseeable additional effects on species' distribution and abundances (Van der Putten et al. 2010). Reproduction and feeding-season timing, predator-prey relationships, assemblage structure, interspecific competition for resources are all likely to change in a warming environment adding additional sources of stress for species. The ability of *Aequiyoldia* to alternate its feeding strategy between deposit and suspension-filter feeder (Davenport 1988) and to reproduce throughout the year (Lau et al. 2018) gives this species greater flexibility to cope with biological changes in its environment (e.g., food availability).

The categorization of SSA *Aequiyoldia* as an eventual loser in a future warming scenario arises from the results in Chapter 4; it was experimentally demonstrated that hypoxic conditions under spring-summer temperatures are very unfavorable for the species. The combination of hypoxia and higher temperatures resulted in the highest number of DEG, with induction of apoptosis indicating that these exposures might already be adverse and threaten long-term survival in a population that is likely operating near their physiological limits. The temperature range to which organisms are exposed determines

to a large extent their metabolic rates (Brown, Gillooly, et al. 2004) which means a strong interdependence between temperature and oxygen demand (Abele 2011). The interdependence in the response to hypoxia and temperature has been widely documented in ectotherms. In fish (e.g., *Paralichthys lethostigma*, *Fundulus heteroclitus*) and other bivalve species (e.g., *Pecten maximus*, *Mytilus spp*) for instance, increments in temperature challenge the tolerance thresholds to hypoxia as a result of an increase in oxygen demand (Del Toro-Silva et al. 2008; McBryan et al. 2013; Artigaud et al. 2014). Predictions of global warming point to the generation of marine hypoxic zones as a major side effect (Altieri and Gedan 2015). Also, the increasing salmon farming in South American fjords entails a growing risk of local eutrophication and oxygen deficiency in bottom water layers (Quinones et al. 2019). These threats underline the importance of our results and emphasize the need of multi-stressor studies in the range of forecasted change for the investigation of the impacts of climate change on biological communities.

The status of the SSA *A. cf. eightsii* as a possible loser in its current distribution range under warming scenarios could change by retracting its range towards the pole. Given the current warming at the Northern WAP (Vaughan et al. 2003; Meredith and King 2005; Morley et al. 2020) the range of annual thermal fluctuation in shallow waters of the Magellan region and King-George (KGI) Island are starting to converge, with maximal summer water temperatures of 2°C at KGI corresponding to the lowest winter temperatures in the intertidal of the Magellan strait. The results of the cooling experiment in the SSA *Aequiyoldia* showed that this temperature convergence could imply the introduction of this species in the WAP, and probably others, facilitated by the increase of ships transit between the two continents (McCarthy et al. 2019). However, the apparent absence of adverse physiological responses to cooling in SSA *A. cf. eightsii* adults that emerges from the low number of DEG and the lack of enriched GO terms is not enough to ensure a successful colonization of the Antarctic continent. Instead, the evaluation of a potential establishment of SSA population in WAP requires further experiments with larval development, settlement and subsequent reproduction, the feasibility of which I did not test in this study. While these early stages of development may be the most sensitive to environmental filters restricting immigration from lower

latitudes into the Southern Ocean on an ecological timescale (Aronson et al. 2007; Kaiser et al. 2013), the weakening of these filters in predicted global climate change scenarios is expected to increase the risk of NIS invasion with negative effects for the unusually high levels of endemicity in the Antarctic fauna today.

5.3 Conclusions and perspectives

This thesis provides a molecular perspective on the adaptive capacity and cross-continental invasibility of two *Aequiyoldia* sibling species inhabiting WAP and SSA under a global change scenario. Furthermore, this thesis had a substantial methodological contribution, warning about the interference of mitochondrial heteroplasmy in the classical molecular procedures for species delimitation and identification (known as molecular barcoding) and about the high incidence of stochastic processes behind gene expression patterns that argue against analyses based on the expression pattern of single genes.

The high genetic diversity (Chapter 2) and the transcriptional responses to local environmental heterogeneity and experimental simulations of warming and hypoxic conditions (Chapter 3 and 4) suggest a high adaptive potential of the Antarctic *A. cf. eightsii* that might position this species among the eventual winners of climate change. On the other hand, the critical physiological state (apoptosis) of SSA *Aequiyoldia* in response to hypoxia, worsened by higher temperatures, suggests that this species would be seriously threatened by the predicted warming in South American high latitudes. In addition to the potential direct effects of environmental change on the physiology of organisms in both continents, this thesis demonstrates that the warming scenario forecasted for the WAP will increase the likelihood of introduction of non-native species from lower latitudes. The findings of Chapter 2 suggesting that the populations of *A. cf. eightsii* from either continent do not belong to the same species emphasize the importance of our results regarding the invasibility of SSA *Aequiyoldia* and highlight the need for accurate taxonomy for the understanding of biodiversity patterns and the response of communities to environmental change as species turnover.

Besides the significant gain in knowledge this thesis generated new questions, which raise the need for future investigations:

(i) *Heteroplasmy and doubly uniparental inheritance (DUI) in Aequiyoldia*.

The heteroplasmy of the mitochondrial genome of *A. cf. eightsii* from WAP was a central finding of this thesis. Because DNA was extracted only from somatic tissue (not including gonads), and organisms were not sexed, it could not be determined whether the mechanism behind heteroplasmy is DUI. It is therefore fundamental to find it out, and to determine, through the tissue-specific analysis of a large number of specimens with known sex, whether this phenomenon is expanded to their South American congeners.

(ii) *Intra-population variation in mitochondrial functionality and its role in adaptation to environmental variability*.

I found a population distributed across an environmentally heterogeneous glacial cove (including local redox conditions of the sediment) differing in its mitochondrial functionality, when comparing organisms based on its mitochondrial genotype/mitotype (Chapter 3). Furthermore, under experimental conditions, organisms responded to hypoxia through mitochondrial adjustments. However, since our experimental design did not factor in the possibility of coexisting mitochondrial lineages, future follow-up studies are needed to evaluate the interaction between mitotypes and the environment (G X E), both in situ and experimentally. Thus, it remains to be elucidated whether genetically based functional plasticity of mitochondria mediates organismal responses to environmental stress. Specifically, whether mitotypes respond differentially to the environment, and more interestingly, whether the heteroplasmic/ homoplasmic condition involves a differential response to environmental stress.

(iii) *The Alternative oxidase as a key trait for the adaptation to climate change*.

The analysis of gene expression has the potential for testing the response of numerous traits or genes at the same time, many of which an investigator would not necessarily think to test in advance, particularly so in a non-model species. However, their occurrence or absence in the output of differentially expressed genes generally opens up new questions. This was the case for the enzyme Alternative oxidase (AOX). Its expression was observed in Antarctic organisms exposed for 10 days to hypoxia and in situ temperature, and not in those exposed to hypoxia and warming. But how does the expression of this enzyme vary over time of exposure? Is it an immediate defense mechanism, i.e., is it expressed on the order of hours or a few days after exposure? Does the increase in temperature effectively limit its expression? What is this temperature threshold? Do SSA organisms possess the gene coding for this enzyme? All these questions are important to elucidate the potential role of this enzyme in mediating climate change tolerance in *Aequiyoldia* and other species.

(iv) *Stochasticity in differential gene expression analysis*.

By performing differential gene expression analyses with randomized groups, I demonstrated that there is an important stochastic component behind gene expression patterns (Chapter 3). It is crucial to determine whether the number of DEG arising from the desired comparison is significantly higher than the expected by stochasticity alone (randomized groups) - as it was done for the three comparisons performed in Chapter 3. However, this is not normally tested, and may lead to overestimating the number of differentially expressed genes due to the occurrence of false positives. It is likely that some of the parameters of the differential gene expression analyses such as the strictness (p-value and fold-change), and the number of replicates of compared groups will affect the magnitude of the stochastic effects. It is expected that the incidence of stochasticity increases

with less strictness and fewer replicates per group. Through the use of data sets available in public servers and an R library that we are developing for the automation of groups randomization it would be of great interest to evaluate which are the optimal parameters in the analysis of gene expression to avoid interpretations based on false positives product of stochastic effects.

References

Abele D. 2011. Temperature adaptation in changing climate: marine fish and invertebrates. Temperature Adaptation in a Chnaging Climate: Nature at Risk: CAB International.67-79.

Abele D, Brey T, Philipp E. 2017. Ecophysiology of extant marine Bivalvia. Treatise Online. Paleontological Institute, University of Kansas, Lindley Hall. 1-50.

Abele D, Philipp E, Gonzalez P, Puntarulo S. 2007. Marine invertebrate mitochondria and oxidative stress. Front. Biosci. 12:933-946.

Advani NK, Kenkel CD, Davies SW, Parmesan C, Singer MC, Matz MV. 2016. Variation in heat shock protein expression at the latitudinal range limits of a widely-distributed species, the Glanville fritillary butterfly (*Melitaea cinxia*). Physiol Entomol. 41(3):241-248.

Altieri AH, Gedan KB. 2015. Climate change and dead zones. Global Change Biol. 21(4):1395-1406.

Aronson RB, Thatje S, Clarke A, Peck LS, Blake DB, Wilga CD, Seibel BA. 2007. Climate change and invasibility of the Antarctic benthos. Annu Rev Ecol Evol Syst. 38:129-154.

Artigaud S, Lacroix C, Pichereau V, Flye-Sainte-Marie J. 2014. Respiratory response to combined heat and hypoxia in the marine bivalves *Pecten maximus* and *Mytilus spp.* Comp Biochem Physiol A Mol Integr Physiol. 175:135-140.

Bedulina DS, Zimmer M, Timofeyev MA. 2010. Sub-littoral and supra-littoral amphipods respond differently to acute thermal stress. Comp Biochem Physiol B Biochem Mol Biol. 155(4):413-418.

Benito Garzón M, Alía R, Robson TM, Zavala MA. 2011. Intra-specific variability and plasticity influence potential tree species distributions under climate change. Global Ecol Biogeogr. 20(5):766-778.

Brown A, Thatje S. 2015. The effects of changing climate on faunal depth distributions determine winners and losers. Glob Chang Biol. 21(1):173-180.

Brown JH, Gillooly JF, Allen AP, Savage VM, West GB. 2004. Toward a metabolic theory of ecology. Ecology. 85(7):1771-1789.

Brown KM, Fraser KP, Barnes DK, Peck LS. 2004. Links between the structure of an Antarctic shallow-water community and ice-scour frequency. Oecologia. 141(1):121-129.

Bush A, Mokany K, Catullo R, Hoffmann A, Kellermann V, Sgro C, McEvey S, Ferrier S. 2016. Incorporating evolutionary adaptation in species distribution modelling reduces projected vulnerability to climate change. Ecol Lett. 19(12):1468-1478.

Clark MS, Peck LS. 2009. HSP70 heat shock proteins and environmental stress in Antarctic marine organisms: A mini-review. Mar Genomics. 2(1):11-18.

Clark MS, Sommer U, Sihra JK, Thorne MA, Morley SA, King M, Viant MR, Peck LS. 2016. Biodiversity in marine invertebrate responses to acute warming revealed by a comparative multi-omics approach. Glob Chang Biol. 23(1)318-330

Clucas GV, Dunn MJ, Dyke G, Emslie SD, Naveen R, Polito MJ, Pybus OG, Rogers AD, Hart T. 2014. A reversal of fortunes: climate change 'winners' and 'losers' in Antarctic Peninsula penguins. Sci Rep. 4:5024.

Daufresne M, Lengfellner K, Sommer U. 2009. Global warming benefits the small in aquatic ecosystems. Proc. Natl. Acad. Sci. 106(31):12788-12793.

Davenport J. 1988. The feeding mechanism of *Yoldia* (═ *Aequiyoldia*) *eightsi* (Courthouy). Proc. Royal Soc. B. 232(1269):431-442.

Del Toro-Silva F, Miller J, Taylor J, Ellis T. 2008. Influence of oxygen and temperature on growth and metabolic performance of *Paralichthys lethostigma* (Pleuronectiformes: Paralichthyidae). J Exp Mar Biol Ecol. 358(2):113-123.

Deutsch C, Ferrel A, Seibel B, Pörtner H-O, Huey RB. 2015. Climate change tightens a metabolic constraint on marine habitats. Science. 348(6239):1132-1135.

Griffiths HJ, Meijers AJ, Bracegirdle TJ. 2017. More losers than winners in a century of future Southern Ocean seafloor warming. Nat. Clim. Change. 7(10):749-754.

Gross M. 2004. Emergency services: a bird's eye perspective on the many different functions of stress proteins. Current protein and peptide science. 5(4):213-223.

Kaiser S, Brandão SN, Brix S, Barnes DK, Bowden DA, Ingels J, Leese F, Schiaparelli S, Arango CP, Badhe R. 2013. Patterns, processes and vulnerability of Southern Ocean benthos: a decadal leap in knowledge and understanding. Mar Biol. 160(9):2295-2317.

Khaitovich P, Weiss G, Lachmann M, Hellmann I, Enard W, Muetzel B, Wirkner U, Ansorge W, Paabo S. 2004. A neutral model of transcriptome evolution. PLoS Biol. 2(5):E132.

Kimura M. 1968. Evolutionary rate at the molecular level. Nature. 217(5129):624-626.

Koenigstein S, Pohlmann K, Held C, Abele D. 2013. Ecological comparison of cellular stress responses among populations - normalizing RT-qPCR values to investigate differential environmental adaptations. BMC Ecol. 13:21.

Lau SCY, Grange LJ, Peck LS, Reed AJ. 2018. The reproductive ecology of the Antarctic bivalve *Aequiyoldia eightsii* (Protobranchia: Sareptidae) follows neither Antarctic nor taxonomic patterns. Polar Biol. 41(9):1693-1706.

Linse K, Barnes DKA, Enderlein P. 2006. Body size and growth of benthic invertebrates along an Antarctic latitudinal gradient. Deep Sea Res. Part II Top. Stud. Oceanogr. 53(8-10):921-931.

Markert JA, Champlin DM, Gutjahr-Gobell R, Grear JS, Kuhn A, McGreevy TJ, Roth A, Bagley MJ, Nacci DE. 2010. Population genetic diversity and fitness in multiple environments. BMC Evol Biol. 10(1):1-13.

Matus-Ortega MG, Salmerón-Santiago KG, Flores-Herrera O, Guerra-Sánchez G, Martínez F, Rendón JL, Pardo JP. 2011. The alternative NADH dehydrogenase is present in mitochondria of some animal taxa. Comp. Biochem. Physiol. Part D Genomics Proteomics. 6(3):256-263.

McBryan TL, Anttila K, Healy TM, Schulte PM. 2013. Responses to temperature and hypoxia as interacting stressors in fish: implications for adaptation to environmental change. Integr Comp Biol. 53(4):648-659.

McCarthy AH, Peck LS, Hughes KA, Aldridge DC. 2019. Antarctica: The final frontier for marine biological invasions. Glob Chang Biol. 25(7):2221-2241.

McDonald AE, Gospodaryov DV. 2019. Alternative NAD(P)H dehydrogenase and alternative oxidase: Proposed physiological roles in animals. Mitochondrion. 45:7-17.

McDonald AE, Vanlerberghe GC, Staples JF. 2009. Alternative oxidase in animals: unique characteristics and taxonomic distribution. J Exp Biol. 212(16):2627-2634.

Meredith MP, King JC. 2005. Rapid climate change in the ocean west of the Antarctic Peninsula during the second half of the 20th century. Geophys Res Lett. 32(19).

Morley SA, Abele D, Barnes DKA, Cárdenas CA, Cotté C, Gutt J, Henley SF, Höfer J, Hughes KA, Martin SM et al. 2020. Global Drivers on Southern Ocean Ecosystems: Changing Physical Environments and Anthropogenic Pressures in an Earth System. Front. Mar. Sci. 7:1097

Nei M. 1987. Molecular Evolutionary Genetics. Columbia University Press.

Pasotti F, Manini E, Giovannelli D, Wölfl A-C, Monien D, Verleyen E, Braeckman U, Abele D, Vanreusel A. 2015. Antarctic shallow water benthos in an area of recent rapid glacier retreat. Mar Ecol. 36(3):716-733.

Pavey SA, Collin H, Nosil P, Rogers SM. 2010. The role of gene expression in ecological speciation. Ann N Y Acad Sci. 1206:110-129.

Pöhlmann K, Koenigstein S, Alter K, Abele D, Held C. 2011. Heat-shock response and antioxidant defense during air exposure in Patagonian shallow-water limpets from different climatic habitats. Cell Stress and Chaperones. 16(6):621.

Pörtner HO. 2005. Synergistic effects of temperature extremes, hypoxia, and increases in CO_2 on marine animals: From Earth history to global change. Journal of Geophysical Res. 110(C9).

Prada C, Schizas NV, Yoshioka PM. 2008. Phenotypic plasticity or speciation? A case from a clonal marine organism. BMC Evol Biol. 8:47.

Quinones RA, Fuentes M, Montes RM, Soto D, León-Muñoz J. 2019. Environmental issues in Chilean salmon farming: a review. Rev. Aquac. 11(2):375-402.

Razgour O, Forester B, Taggart JB, Bekaert M, Juste J, Ibanez C, Puechmaille SJ, Novella-Fernandez R, Alberdi A, Manel S. 2019. Considering adaptive genetic variation in climate change vulnerability assessment reduces species range loss projections. Proc Natl Acad Sci U S A. 116(21):10418-10423.

Rifkin SA, Kim J, White KP. 2003. Evolution of gene expression in the *Drosophila melanogaster* subgroup. Nat Genet. 33(2):138-144.

Shepherd JG, Brewer PG, Oschlies A, Watson AJ. 2017. Ocean ventilation and deoxygenation in a warming world: introduction and overview. Philos. Trans. R. Soc. A. 375:2102

Siciński J, Pabis K, Jażdżewski K, Konopacka A, Błażewicz-Paszkowycz M. 2011. Macrozoobenthos of two Antarctic glacial coves: a comparison with non-disturbed bottom areas. Polar Biol. 35(3):355-367.

Sussarellu R, Fabioux C, Sanchez MC, Le Goïc N, Lambert C, Soudant P, Moraga D. 2012. Molecular and cellular response to short-term oxygen variations in the Pacific oyster *Crassostrea gigas*. J Exp Mar Biol Ecol. 412:87-95.

Uebbing S, Kunstner A, Makinen H, Backstrom N, Bolivar P, Burri R, Dutoit L, Mugal CF, Nater A, Aken B et al. 2016. Divergence in gene expression within and between two closely related flycatcher species. Mol Ecol. 25(9):2015-2028.

Valladares F, Matesanz S, Guilhaumon F, Araújo MB, Balaguer L, Benito-Garzón M, Cornwell W, Gianoli E, van Kleunen M, Naya DE. 2014. The effects of phenotypic plasticity and local adaptation on forecasts of species range shifts under climate change. Ecol Lett. 17(11):1351-1364.

Van der Putten WH, Macel M, Visser ME. 2010. Predicting species distribution and abundance responses to climate change: why it is essential to include biotic interactions across trophic levels. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 365(1549):2025-2034.

Vaughan DG, Marshall GJ, Connolley WM, Parkinson C, Mulvaney R, Hodgson DA, King JC, Pudsey CJ, Turner J. 2003. Recent rapid regional climate warming on the Antarctic Peninsula. Clim Change. 60(3):243-274.

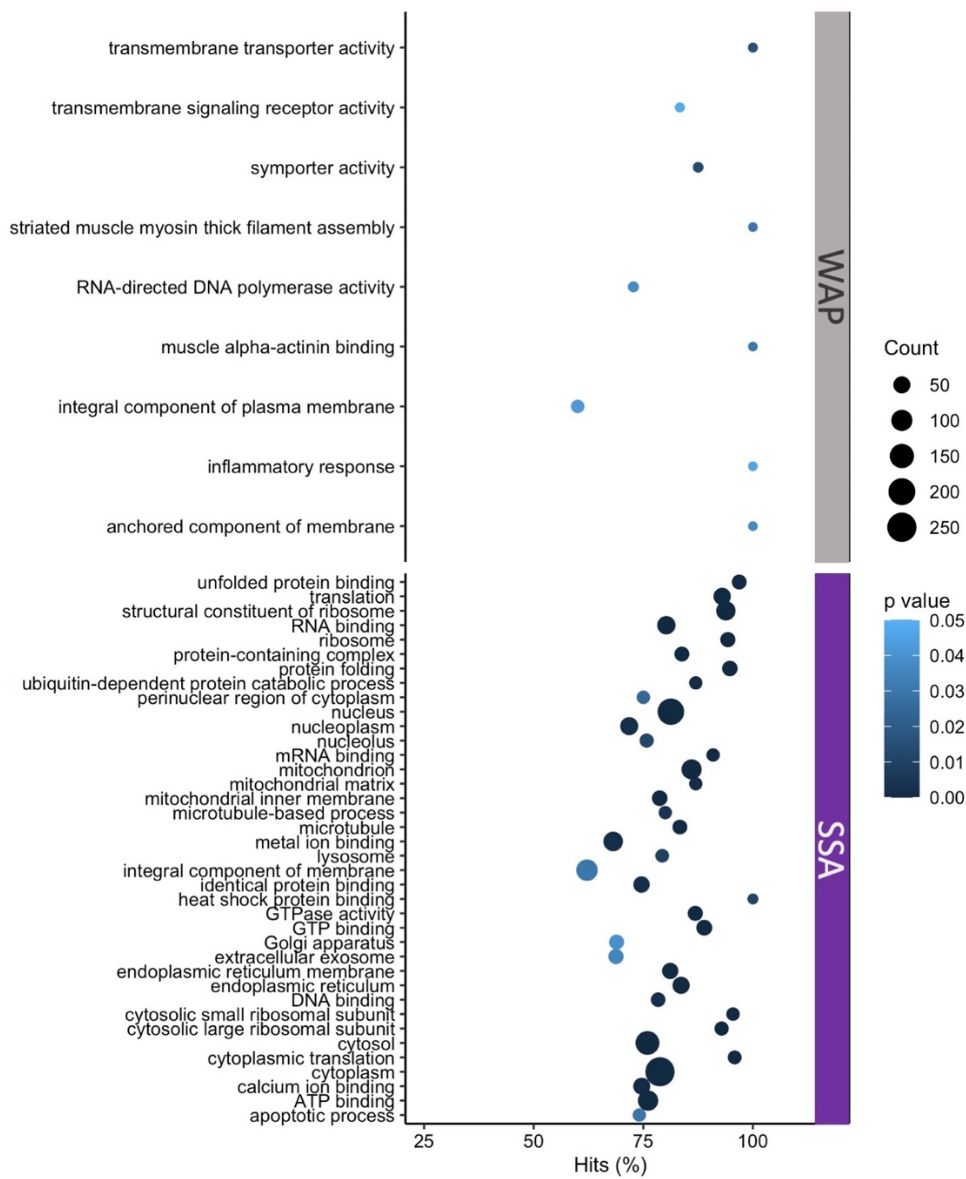
Whitehead A, Crawford DL. 2006. Neutral and adaptive variation in gene expression. Proc Natl Acad Sci. 103(14):5425-5430.

Yanai I, Graur D, Ophir R. 2004. Incongruent expression profiles between human and mouse orthologous genes suggest widespread neutral evolution of transcription control. OMICS: J Integrative Biol. 8(1):15-24.

Yang JR, Maclean CJ, Park C, Zhao H, Zhang J. 2017. Intra and interspecific variations of gene expression levels in Yeast are largely neutral: (Nei Lecture, SMBE 2016, Gold Coast). Mol Biol Evol. 34(9):2125-2139.

APPENDIXES

Appendix 5.1. GO enrichment analysis of pairwise comparisons. Enriched terms are shown separately for West Antarctic peninsula (WAP) and southern South America (SSA). The *x* axis indicates the percentage of DEG found in a certain category over the total number of genes in this category (Hits). Size code indicates the number of DEG in the GO term category (Count) and a gradient of colour the *p value* of the over represented category. Only GO terms with more than 20 differentially expressed genes (DEG) in the category are shown in the plot for SSA group.



Declaration on the contribution of the candidate to a multi-author article/manuscript which is included as a chapter in the submitted doctoral thesis

Chapter: 2

Contribution of the candidate in % of the total work load (up to 100% for each of the following categories):

Experimental concept and design: ca. 80%

Experimental work and/or acquisition of (experimental) data: ca. 90%

Data analysis and interpretation: ca. 80 %

Preparation of Figures and Tables: ca. 100%

Drafting of the manuscript: ca. 75%

Chapter: 3

Contribution of the candidate in % of the total work load (up to 100% for each of the following categories):

Experimental concept and design: ca. 80%

Experimental work and/or acquisition of (experimental) data: ca. 90%

Data analysis and interpretation: ca. 80%

Preparation of Figures and Tables: ca. 100%

Drafting of the manuscript: ca. 90%

Chapter: 4

Contribution of the candidate in % of the total work load (up to 100% for each of the following categories):

Experimental concept and design: ca. 70%

Experimental work and/or acquisition of (experimental) data: ca. 90%

Data analysis and interpretation: ca. 80%

Preparation of Figures and Tables: ca. 100%

Drafting of the manuscript: ca. 75%

Date: July 27th, 2021 Signatures

Versicherung an Eides Statt

Ich, Mariano Martínez

Falkenstrasse 20,

28195 Bremen

Mtr-Nr: 3107414

versichere an Eides Statt durch meine Unterschrift, dass ich die vorstehende Arbeit selbständig und ohne fremde Hilfe angefertigt und alle Stellen, die ich wörtlich dem Sinne nach aus Veröffentlichungen entnommen habe, als solche kenntlich gemacht habe, mich auch keiner anderen als der angegebenen Literatur oder sonstiger Hilfsmittel bedient habe.

Ich versichere an Eides Statt, dass ich die vorgenannten Angaben nach bestem Wissen und Gewissen gemacht habe und dass die Angaben der Wahrheit entsprechen und ich nichts verschwiegen habe.

Die Strafbarkeit einer falschen eidesstattlichen Versicherung ist mir bekannt, namentlich die Strafandrohung gemäß § 156 StGB bis zu drei Jahren Freiheitsstrafe oder Geldstrafe bei vorsätzlicher Begehung der Tat bzw. gemäß § 161 Abs. 1 StGB bis zu einem Jahr Freiheitsstrafe oder Geldstrafe bei fahrlässiger Begehung.

Bremen, 27. Juli 2021 Mariano Martínez

Acknowledgments

First of all, I would like to thank my thesis supervisors **Doris Abele** and **Christoph Held** for their permanent support during the whole period. For every piece of advice, for every interesting discussion (the most memorable one: in that café in Punta Arenas where we adjusted the experimental design), for every message of support during the classic stumbles in a PhD. For being always there, even during very difficult personal situations.

A big thank you to **Andrea Eschbach** for her support (and patience) not only inside the lab, but also outside the lab - with always a solution for all my little daily (mainly bureaucratic) problems. The excellence of her work, and her ability to teach were fundamental to achieving the results I obtained.

I would like to thank **Martin Zimmer** and **Felix Mark** for agreeing reviewing this thesis, and to **Christian Wild** and **Charlotte Havermans** for joing my examining commission. I also thank to **Onur Karakus** and **Luis Kuchenmüller** for agreeing to be a student representation in my commission. I extend my thanks to **Inka Bartsch** who together with Martin Zimmer and Felix Mark was part of my thesis advisory committee.

I thank the funders of this work: **National Agency of Investigation and Innovation** (ANII, Uruguay) and **German Academic Exchange Service** (DAAD, Germany) for PhD scholarship (Grant number POS_EXT_2015_122792), and Zempelin-Stiftung - **Deutsches Stiftungszentrum** (Grant number T0214/33883/2019/sm) and **Biosciences section of AWI** for founding the last period of PhD.

I am grateful for the great support of **Lars Harms** in the bioinformatics analyses, **Stefanie Meyer** in the preparation of the expedition and pilot tests in the laboratory, **Uwe John** in the NGS design, **Nancy Kuehne** in the sample preparation for NGS, **Camila Neder** in the preparation of maps, **Isabell Viehweger** in the analysis of sediment samples and **Enrique Celemin** in the DNA extraction and PCR and for sharing nice moments in the office for half a year.

I thank the whole **Functional Ecology** section, especially **Petra Meyer** for her administrative support, and **Kerstin Beyer** for her help in the lab cutting shells, but mainly both of them for their great sympathy and pleasant daily corridor conversations.

I would also like to thank many people who collaborated and were fundamental to develop a successful expedition for almost 5 months to Punta Arenas and Antarctica: **Dirk Mengedoht** from AWI; **Marcelo González** and **Leyla Cardenas** from **Instituto**

Antártico Chileno; **Ana Laura Machado** and **Alvaro Soutullo** from **Instituto Antártico Uruguayo**; **Dolores Deregibus**, **Walter Nahueltripay**, **Natalia Servetto**, **Divers** and the complete **'Dotación'** of **Carlini** from **Instituto Antártico Argentino**. The expedition was an unforgettable experience, an expedition of life. Thanks to all who shared it with me, especially **Christoph Held** with whom we have unforgettable memories.

I thank my fellows and friends '**gord@s lechones**' with whom I shared countless lunches, coffees, cakes and most importantly - laughs!

I thank from the bottom of my heart all the friends who jumped into this path, at the beginning a hard core of Latinos (**La pandilla**), but then many more, just as important, from far away countries that I have never visited, but that now through them I consider close to me. Gracias amigos, danke, merci, grazie, obrigado, Спасибо, Hvala vam, Teşekkürler, धन्यवाद.

Agradezco a mi familia en Uruguay, especialmente a mis padres y a mi hermano. A 12.000 km de distancia siempre fueron mi respaldo, un lugar seguro. Este trabajo se los dedico a ustedes, viejos. También agradezco a nuestra familia alemana de Bortfeld por hacernos sentir en casa. Por ultimo, a Miri, mi compañera y soporte de todas las horas.

In one of those many meetings with Doris, talking about science and the world, often about Latin America - a passion we share - she gave me a book. It was not a book on physiology. It was a book in Spanish by the Uruguayan author Mario Benedetti.

The book is called *Gracias por el Fuego* (*Thanks for the fire*).

Thank you, Doris, thank you for that fire.

Thank you all. Gracias por el fuego, Danke für das Feuer.