White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of Leeds.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/194661/

Version: Accepted Version

Article:

Nielsen, ES, Beger, M, Henriques, R et al. (2021) Neither historical climate nor contemporary range fully explain the extant patterns of molecular diversity in marine species. Journal of Biogeography, 48 (10). pp. 2629-2644. ISSN: 0305-0270

https://doi.org/10.1111/jbi.14229

© 2021 John Wiley & Sons Ltd. This is the peer reviewed version of the following article: Nielsen, E. S., Beger, M., Henriques, R., & von der Heyden, S. (2021). Neither historical climate nor contemporary range fully explain the extant patterns of molecular diversity in marine species. Journal of Biogeography, 48, 2629– 2644. , which has been published in final form at https://doi.org/10.1111/jbi.14229. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions. This article may not be enhanced, enriched or otherwise transformed into a derivative work, without express permission from Wiley or by statutory rights under applicable legislation. Copyright notices must not be removed, obscured or modified. The article must be linked to Wiley’s version of record on Wiley Online Library and any embedding, framing or otherwise making available the article or pages thereof by third parties from platforms, services and websites other than Wiley Online Library must be prohibited.

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

Neither historical climate nor contemporary range fully explain extant patterns of molecular diversity in marine species

Running title: Determinants of genetic diversity

Authors: Erica S. Nielsen[1], Maria Beger[2], Romina Henriques[1,3], Sophie von der Heyden[1]

Affiliations:
[1]Evolutionary Genomics Group, Department of Botany and Zoology, University of Stellenbosch, Private Bag X1, Matieland, 7602, South Africa
[2]School of Biology, Faculty of Biological Sciences, University of Leeds, Leeds, LS2 9JT, UK
[3]Technical University of Denmark, National Institute of Aquatic Resources, Section for Marine Living Resources, Velsøvej 39, Silkeborg, 8600, Denmark

Corresponding author: Erica S. Nielsen
Address: Evolutionary Genomics Group, Department of Botany and Zoology, University of Stellenbosch, Private Bag X1, Matieland 7602, South Africa
Email: esnielsen@sun.ac.za
OrcID: 0000-0002-5439-571X

Author: Maria Beger
Address: School of Biology, Faculty of Biological Sciences, University of Leeds, LS2 9JT, UK
Email: m.beger@leeds.ac.uk
OrcID: 0000-0003-1363-3571

Author: Romina Henriques
Address: Technical University of Denmark, National Institute of Aquatic Resources, Section for Marine Living Resources, Velsøvej 39, Silkeborg, 8600, Denmark
Email: romhe@aqua.dtu.dk
OrcID: 0000-0002-6544-5532

Author: Sophie von der Heyden
Address: Evolutionary Genomics Group, Department of Botany and Zoology, University of Stellenbosch, Private Bag X1, Matieland 7602, South Africa
Email: svdh@sun.ac.za
OrcID: 0000-0001-9166-976X

Significance statement: High genetic diversity is important as it can help species adapt to climate change. Thus, we may want to highlight populations with high genetic diversity within a species for conservation. But even more important is identifying the processes which created regions of high diversity, and trying to conserve these processes into the future. Here, we tried to identify the dominant processes potentially shaping genetic diversity within three marine species. We tested two widely acknowledged hypotheses to find that both processes are important in explaining extant genetic diversity patterns, but that neither is likely to be the sole driver of diversity.

Abstract

Aim: Intraspecific diversity is a significant component of adaptive potential, and thus it is important to identify the evolutionary processes which have, and will continue, to shape the molecular diversity of natural populations. This study aims to untangle possible drivers of intraspecific molecular diversity by testing whether patterns of historical climatic stability or contemporary range position correlate with molecular diversity.

Location: South African coastline.

Taxa: The Cape urchin (*Parechinus angulosus*), Common shore crab (*Cyclograpsus punctatus*) and Granular limpet (*Scutellastra granularis*).

Methods: Species distributions were hindcasted to the Last Glacial Maximum to assess the biogeography of the study species. Linear models were built to compare the relationships between historical climatic stability or contemporary distributional ranges with extant genetic and genomic diversity.

Results: We found large differences in the historical ranges among species and time periods. Regions of higher habitat stability corresponded to regions of higher molecular diversity, but historical climatic variability was not a predictor of molecular diversity within linear models. Lower genetic diversity values, and higher genetic differentiation, were detected in edge populations, but this was not consistent across marker type or species

Main conclusions: Both historical and contemporary processes are potentially driving patterns of diversity, but a large portion of the variation in molecular diversity remains unexplained. Our findings suggest that marine species within cool-temperate bioregions in the Southern Hemisphere may have more complex biogeographic and evolutionary histories than terrestrial taxa and/or coastal species within northern, formerly glaciated regions.

Introduction

Characterising the impacts of historical and contemporary processes shaping extant genetic variation is of longstanding interest in evolutionary studies (Avise et al., 1987) and has regained popularity within the context of identifying populations with heightened evolutionary potential (Hoban et al., 2019; Theodoridis et al., 2020). Intraspecific diversity is an important component of resilience to climate change (Phair et al., 2020; Reusch et al., 2005; Wernberg et al., 2018), and as such it is fundamental to characterise the genetic diversity of populations across a species' range (Pauls et al., 2013). Species distributions and genetic diversity are strongly influenced by evolutionary processes (Bowen et al., 2016). Therefore, understanding of how these processes contribute to contemporary patterns may provide valuable insights into the adaptive potential of species and their future persistence in a changing world.

Intraspecific molecular diversity is derived from a complex interplay of spatial and temporal variation, including demographic processes, lineage interactions, local adaptation and genetic drift (Rödin-Mörch et al., 2019). However, studies aiming to describe the biogeographic processes driving genetic diversity have predominantly tested only one of two popular hypotheses (Gougherty et al., 2020; Guo, 2012). The first is the 'abundant-centre', or herein referred as the 'central-margin', hypothesis, which proposes that the core of a species' distribution harbours greater genetic diversity than range edges (Eckert et al., 2008; Lira-Noriega & Manthey, 2014). This hypothesis assumes that range edges experience more variable or potentially unfavourable ecological conditions, as well as more pronounced range shifts over time (Brown, 1984), resulting in lowered genetic diversity (Lawton, 1993). Many empirical studies support this pattern (reviewed in Eckert et al., 2008 and Pironon et al., 2017), but there are some exceptions, especially in the marine environment, possibly caused by high gene flow between core and edge populations or environmental pressures historically selecting on range edges (Maggs et al., 2008). In contrast, several studies provide evidence for an alternative, but not mutually exclusive hypothesis, stating that areas which remained climatically stable and habitable over time should strongly correlate with areas of higher extant genetic diversity (Excoffier et al., 2009; Hewitt, 2000). This hypothesis, herein referred to as 'refugial-persistence', assumes that areas of stability during paleoclimatic oscillations, i.e. glacial refuges, should harbour persistent populations over time (due to their ability to stay within their environmental niches), and lead to higher genetic (and species) diversity, compared to areas that experienced historical range shifts and bottlenecks (Beatty & Provan, 2011; Bennett et al., 1991). Multiple studies have confirmed refugial-persistence within marine species in the Northern Hemisphere, showing higher diversity in southern refugial areas, compared to the northern regions which were formerly glaciated (Assis et al., 2014; Chefaoui et al., 2017; Francisco et al., 2014; Neiva et al., 2014; Waltari & Hickerson, 2013).

Although a multitude of studies have assessed central-margin or refugial-persistence patterns in natural populations, these hypotheses have almost always been investigated separately. Further, many studies reporting central-margin relationships have not accounted for latitudinal gradients, and thus cannot discount the effect of refugial-persistence (Guo, 2012). Although recent work formally testing both hypotheses suggests that contemporary range position may be a better predictor of genetic diversity than historical climatic stability (De Kort et al., 2020; Gougherty et al., 2020), these studies were conducted in terrestrial settings, leaving comprehensive comparisons in marine species unexplored. Different drivers of molecular diversity can be expected in marine than terrestrial species, as global trends of intraspecific diversity differ between the two, with marine taxa showing strong latitudinal trends (Manel et al., 2020) whilst terrestrial taxa do not (De Kort et al., 2020). Many inferences of molecular diversity are also often obtained from a single marker, which has its owns caveats, as markers like mitochondrial DNA (mtDNA) are expected to reflect more historical processes compared to other markers such as genome-wide single nucleotide polymorphisms (SNPs; Wang, 2011). Additionally, processes shaping extant molecular diversity of temperate marine species have predominantly been investigated in the Northern Hemisphere, where glaciation events and variation in ice sheets differ profoundly from Southern Hemisphere coastal systems (Velichko et al., 1997), many of which did not have ice cover at the height of the most recent glaciation. Lastly, drivers of molecular diversity are also usually investigated either visually with habitat suitability maps (Assis et al., 2014; Chefaoui et al., 2017; Neiva et al., 2014) or statistically with linear models (Acevedo-Limón et al., 2020; De Kort et al., 2020; Manel et al., 2020), but rarely with both approaches. The prevailing processes driving intraspecific diversity can be inferred with more certainty when both visual and statistical analyses are combined (Yannic et al., 2014).

South Africa is an ideal region to compare the relative contribution of historical and contemporary processes shaping molecular variation, due to its largely linear coastline exhibiting strong environmental clines over short geographic distances (Branch & Branch, 2018), and high levels of intraspecific genetic variation in the region (Teske et al., 2011; Wright et al., 2015). The coastline is bounded by two contrasting current systems, creating predominantly longitudinal biogeographic gradients (Fig. 1). Thus, in this system, species can broadly shift east or west in response to changes in environmental changes (Branch & Branch, 2018). During the Last Glacial Maximum (LGM) South Africa's coast was not glaciated, but paleoclimatic changes, specifically in sea-level, are thought to have led to shifted species distributions (Toms et al., 2014). Marine phylogeographic studies within South Africa have attributed patterns of genetic variation either to ecological divergence via contemporary selection (Teske et al., 2019; Zardi et al., 2007), or to paleoceanographic changes such as sea-level variability since the LGM (von der Heyden et al., 2011; Phair et al., 2019; Toms et al., 2014), highlighting the need to test the relative contributions of historical and contemporary drivers of genetic variation in the region.

In this study, the processes driving molecular diversity along South Africa's coastal species are tested using a combination of paleoclimatic reconstructions, Species Distribution Models (SDMs), and phylogeographic analyses for three southern African endemic species. Firstly, species distributions are modelled at present day, the Mid-Holocene, and LGM to understand how environmental conditions contribute towards shaping the past and present distributions of southern

African coastal species. Distributions are modelled at the species and lineage level, as recent studies suggest that SDMs based on lineages may lead to better defined distributional patterns when niche divergence is present between populations (Collart et al., 2021). Secondly, this study tests the relationships between historical climatic change, contemporary range position, and molecular diversity for each species, using both visual and quantitative assessments. In summary, the overarching objectives of this study are to: 1) characterise species range shifts since the LGM and the dominant environmental features driving them across three co-distributed southern African species; 2) compare hindcasted species distribution outputs from models run at the species or lineage level; 3) compare extant intraspecific diversity patterns between genetic mtDNA and genomic SNP markers; 4) test the relative importance of historical climatic stability and contemporary range position on intraspecific molecular diversity.

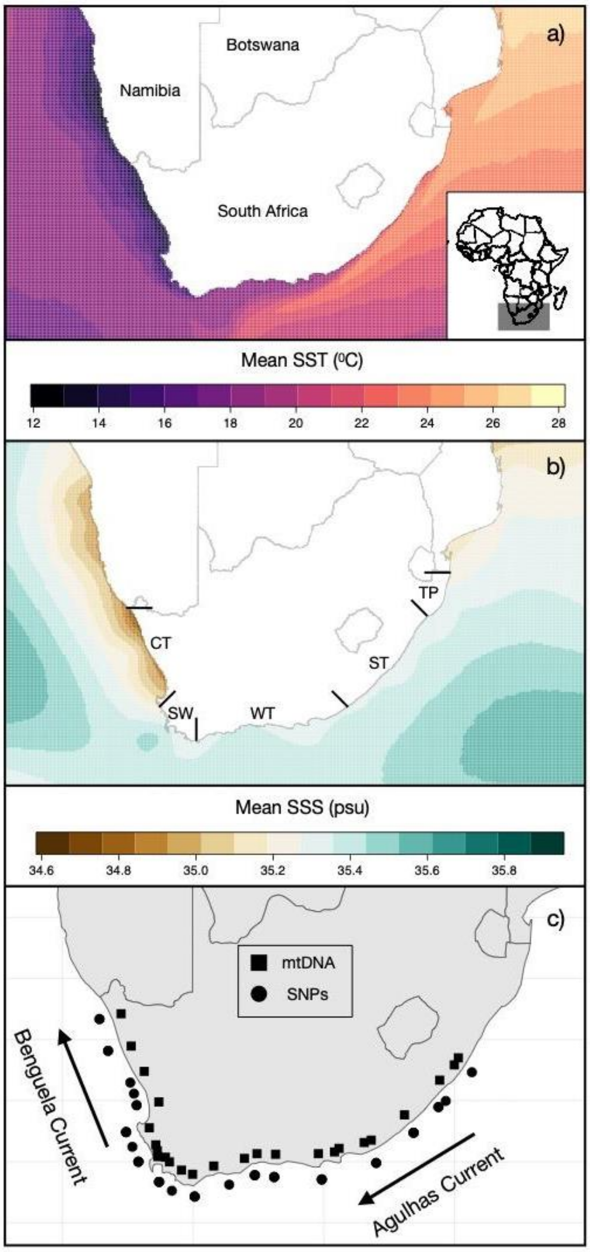

Figure 1- Mean sea surface temperatures (Mean SST; a), mean sea surface salinity (Mean SSS) based on monthly averages from MARSPEC data (Sbrocco, 2014) and the Cool-temperate (CT), Southwest (SW), Warm-temperate (WT), Subtropical (ST) and Tropical (TP) bioregions of South Africa (b), as well as the dominant current systems and the genetic (mtDNA; represented by landward squares) and genomic (SNPs; represented by seaward circles) sample sites (c) are shown. Projection in the maps: WGS84

Materials and Methods

Study species

Study species included the Cape urchin, *Parechinus angulosus*, Granular limpet, *Scutellastra granularis*, and Common shore crab, *Cyclograpsus punctatus*. These are abundant members of South African rocky shore communities, have sympatric distributions, and are found in the low, medium and high intertidal zonations, respectively (Branch & Branch, 2018). The species have differing dispersal abilities that may affect their distributions and genetic diversity. *Scutellastra granularis* and *P. angulosus* are broadcast spawners, while *C. punctatus* is a brooder prior to releasing its larvae. Pelagic larval durations estimates vary from ~ 5-15 days for *S. granularis* and *C. punctatus,* and ~50 days for *P. angulosus* (Wright et al., 2015).

Distributional range shifts since the Last Glacial Maximum

SDMs were used to hindcast species distributions to the LGM, predicting the extent of historical ranges based on statistical relationships between known contemporary species distributions and environmental variation (Guisan & Thuiller, 2005). Models included environmental variables which are thought to influence the genomic variation of the study species (Nielsen et al., 2020b) including mean sea surface temperature (SST) and salinity (SSS), and mean air temperature of the warmest and coolest months (Tmax and Tmin, respectively). Collinearity was assessed from the variance inflation factor (*VIF*). Layers with *VIF* >10 were removed. Environmental variables were obtained from the MARSPEC (Sbrocco, 2014) and WorldClim (Fick & Hijmans, 2017) databases, at a resolution of ~1km. As different General Circulation Models can lead to distinct SDM projections (Vega et al., 2010), models were based on both the Community Climate System Model (CCSM4) and the Model for Interdisciplinary Research on Climate (MIROC5; Assis et al., 2014), for three available time snapshots: present day, Mid-Holocene (MH; 6 thousand years ago - kya) and Last Glacial Maximum (LGM; 21 kya).

SDMs were run independently for each species, using only presence data. Species occurrences were downloaded from the Global Biodiversity Information Facility (GBIF; GBIF, 2020) and Ocean Biogeographic Information System (OBIS; OBIS, 2020) databases, and supplemented/confirmed by fieldwork, including all sample sites. Pseudo-absences were generated with the R package ‘biomod2’ (Thuiller et al., 2009) by randomly generating an equal number of absences to presences per species, replicated over five runs (Bermejo et al., 2018). SDMs were run within *biomod2*, following an ensemble approach that can significantly improve model accuracy over single models (Araújo & New, 2007; Forester et al., 2013). Six standard model types were used within ensemble models; (1) Flexible Discriminant Analysis (FDA), (2) Generalized Additive Model (GAM), (3) Generalized Boosting Model (GBM), (4) Generalized Linear Model (GLM), (5) Multivariate Adaptive Regression Splines (MARS) and (6) Random Forest (RF; Brown & Yoder, 2015; Chefaoui et al., 2017; Engler et al., 2011; Yannic et al., 2014). Default parameters were used for all algorithms, except for GAM models, which were executed with the GAM ‘mgcv’ function, GBM models which were run with 1000 trees, and GLM models which were fitted with a quadratic term.

The relative importance of each environmental predictor variable was assessed by averaging the values from the ‘get_variables_importance’ function of *biomod2* across all models. To assess model performance, a random subset of the data was used for calibration (70%) and validation (30%), with ten evaluation runs and three permutations. Models were scored with an area under the Receiver Operating Characteristic (ROC) curve (AUC; Fielding & Bell, 1997) and True Skill Statistic (TSS; Allouche et al., 2006), with only models having TSS > 0.55 and ROC > 0.8 contributing to the ensemble model. The ensemble models were combined into a weighted mean SDM used to project present day habitat suitability. The ensemble models used both the CCSM4 and MIROC5 models for the Mid-Holocene and Last Glacial Maximum, the outputs of which were merged using the ‘raster’ R package (Hijmans et al., 2015).

Characterising patterns of genetic and genomic diversity

Molecular diversity was derived from both mtDNA and SNP markers. These markers were chosen as slow mutating mtDNA markers are thought to capture relatively historical evolutionary processes compared to markers such as microsatellites, which generally capture more recent events, and SNPs, which have the potential to identify both recent and historical events (Epps & Keyghobadi, 2015). By using mtDNA and SNP data, we are able to assess the influence of marker type on investigations of historical and contemporary drivers of intraspecific diversity. Furthermore, as mtDNA is the dominant marker from which molecular data is available in South Africa (Teske et al., 2011; Tolley et al., 2019), we can identify if these markers are able to capture similar diversity patterns as genomic SNPs, which offer higher statistical resolution but are more costly to produce (Nielsen et al., 2020a).

Diversity was characterised by genetic haplotype diversity (*h*) and genomic expected heterozygosity (H_E), as they represent frequency-weighted variation and are widely used to infer co-ancestry, bottlenecks, and effective population sizes (Nei, 1973, 1987). Genetic data consisted of the mtDNA cytochrome oxidase subunit 1 (CO1) gene region, and was previously generated for 150 and 529 individuals from 11 and 17 sample sites, for *C. punctatus* and *P. angulosus*, respectively (Muller et al., 2012; Wright et al., 2015). Genetic sequences were generated using the universal primers LCO1490 + HCO2198, which resulted in sequences of 790 base pairs (bps) for *P. angulosus* (Muller et al., 2012) and 605 bp for *C. punctatus* (Wright et al., 2015). Mitochondrial genetic data was not included for *S. granularis*, due to the low number (<10) of available individuals sequenced for the majority of the sample sites (Mmonwa et al., 2015). Population-level haplotype diversity was calculated using the program DNAsp v.5 (Librado & Rozas, 2009).

Available genomic data from Nielsen et al. (2020b), was obtained from pooled restriction-site associated DNA sequencing, specifically using the ezRAD approach due to its successful use on non-model marine invertebrates (Toonen et al., 2013). SNP panels were generated in the R program ‘poolfstat’ (Hivert et al., 2018), and after testing multiple parameters, the following were used to call SNPs: minimum coverage = 20, maximum coverage = 400, minimum read count = 4 (Nielsen et al., 2020b). After filtering SNPs for linkage disequilibrium (selecting 1 SNP per 1000 bp), filtering of putatively adaptive alleles, which were those identified by two or more of seven outlier detection methods, was performed (see Nielsen et al., 2020b for details). Final SNP panels consisted of 1175, 804, and 1628 putatively neutral single nucleotide polymorphisms (SNPs) for *C. punctatus*, *P.*

angulosus, and *S. granularis*, respectively. Expected heterozygosity (H_E) was calculated for each site following Fischer et al., (2017; see Eq. S1 in Appendix S1 in Supporting Information).

Influences of past climatic histories on genomic variation at the lineage level

Previous assessments of genetic differentiation in the three study species using mtDNA COI data suggest that each is composed of an east coast and southern/west coast lineage with separate evolutionary histories (Mmonwa et al., 2015; Muller et al., 2012; Wright et al., 2015). Evidence of east and west differentiation was further corroborated by recent genomic work on these species (Nielsen et al., 2020b). SDMs including intraspecific population differentiation, by modelling separate lineages, can potentially portray more accurate habitat suitability outputs (Maia-Carvalho et al., 2018; Zhang et al., 2021). Yet, it has also been argued that the smaller number of observations at the lineage level may lead to issues in model calibration and/or evaluation (Maguire et al., 2018). Recent work suggests that splitting SDMs may only be beneficial when there is evidence of niche divergence (Collart et al., 2021). Here, we tested for niche divergence between the east and west coast populations in each species using an ordination approach (Peñalver-Alcázar et al., 2021; Rodríguez-Rodríguez et al., 2020), conducting redundancy analyses on allele frequencies of the SNP sets stated above, using the same four environmental predictor variables as used in the SDMs. The redundancy analyses were run using the 'vegan' R package (Oksanen et al., 2013).

SDMs were also conducted at the lineage level to compare against those run at the species level. Here the occurrence points were partitioned into east or west coast lineages, and models were run for each lineage independently, following the methodology as above. From the lineage SDMs, the shifts in distributional ranges between the present, Mid-Holocene, and LGM were calculated from the binary species presences outputs, using the 'BIOMOD_RangeSize' function of *biomod2*. As the SDMs at the LGM were mapped onto different map cells, due to lowered sea-level and exposed continental shelf, the habitat suitability values at this time point were interpolated onto map cells of the present and Mid-Holocene models using thine-spline interpolation with the 'fields' R package

(Furrer et al., 2009). The change in available/suitable cells per time point, per lineage, were compared to the H_E values per lineage.

Influence of paleoclimatic changes and current distribution on molecular diversity

To assess whether past climatic changes and/or present day species distributions explain extant intraspecific genetic diversity, linear models were built including three explanatory covariates (Yannic et al., 2014). The first two covariates, climatic variability and sea-level variability, test the hypothesis that past climatic refugia should harbour higher diversity today (i.e. refugial-persistence hypothesis), and the third explanatory variable, distance from the centre of the current distribution, tests whether the distributional centre harbours higher diversity (i.e. central-margin hypothesis). Climatic variability was characterised using paleoclimatic layers projecting mean SST and SSS, precipitation, and air temperature to the LGM, based on the Hadley Centre ocean-atmosphere climate model (HADCM3; Singarayer et al., 2017). These layers have coarse spatial resolutions of >1.25°, and thus were not used for the SDMs. Each variable was projected back 21 kya at 1000-year intervals, and climatic variability was determined by calculating the standard deviation of each variable across the 21 time-steps. Sea-level variability was calculated by measuring the Euclidean distance between the zero to -120m bathymetry lines along the coastline. The distance along the coastline between each sample site and the centre point of the current distribution was calculated per species as the per-site 'marginal distance'.

Linear models were run with the 'lm' R function, comparing eight models including the following predictor variables: 1) Climatic variability, 2) Sea-level variability, 3) Marginal distance, 4) Climatic variability + Sea-level variability, 5) Sea-level variability + Marginal distance, 7) Marginal distance + Climatic variability, 8) Sea-level variability + Climatic variability + Marginal distance. The response variables in the linear models were either genetic or genomic diversity values (h and H_E, respectively). Individual models were ranked using the Akaike Information Criteria corrected for small sample sizes (AICc), specifically ΔAICc and AICc weights (ωi). A test of spatial autocorrelation, using Moran's I of model residuals and pairwise geographic distance, was conducted with the 'testSpatialAutocorrelation' function of the 'DHARMa' R package (Hartig & Hartig, 2017). No significant spatial autocorrelation was detected, meaning that this type of correlation is unlikely to impact model evaluations.

Assessing molecular differentiation across species distributions

The central-margin hypothesis was further tested by comparing pairwise F_{ST} values between central and marginal sites (Kennedy et al., 2020; Wagner et al., 2011). Pairwise F_{ST} values were obtained from the earlier mentioned studies. Weir and Cockerham's pairwise F_{ST} was calculated with the 'poolfstat' R package for the SNP data (see Nielsen et al., 2020b for details), and with Arlequin (Excoffier et al., 2005) for the mtDNA data (see Muller et al., 2012 and Wright et al., 2015 for details). Sample sites were categorized into either the core or the edge of the species' distribution, according to if they were above or below the median range distance from the distributional centre point. Statistical significance of differences in genetic differentiation between core and marginal sites was assessed using permutation tests with 1000 replicates (Karihaloo et al., 2003; Lesbarrères et al., 2006), using the 'coin' R package (Hothorn et al., 2021).

Results

Hindcasting distributions to the Last Glacial Maximum

All environmental variables had *VIFs*<10 at each time-step (Table S1, Appendix S1) and were retained for further analyses. Individual models used in the ensembles had high ROC and TSS values, and thus were included into the ensemble models (Table S2, Appendix S1). Of the four predictor variables, SSS best described the distribution of *C. punctatus* and *P. angulosus*, and minimum air temperature was most important in inferring the distribution of *S. granularis* (Fig. S1, Appendix S1).

The current and hindcasted distributional ranges differed among species and time points (Fig. 2). *Cyclograpsus punctatus* appeared to have experienced a range reduction to the south coast during the Mid-Holocene, which was even more pronounced during the LGM (Fig. 2a-c). The SDMs suggested that, compared to the present day, *P. angulosus* had a reduced range in the Mid-Holocene, with a distributional shift towards the west, and a slightly expanded range during the LGM, with most of the South African west coast and Namibian coastline being habitable (Fig. 2d-f). Compared to *P. angulosus*, *S. granularis* showed a less severe range reduction during the Mid-Holocene, and a slightly larger range expansion during the LGM (Fig. 2g-i).

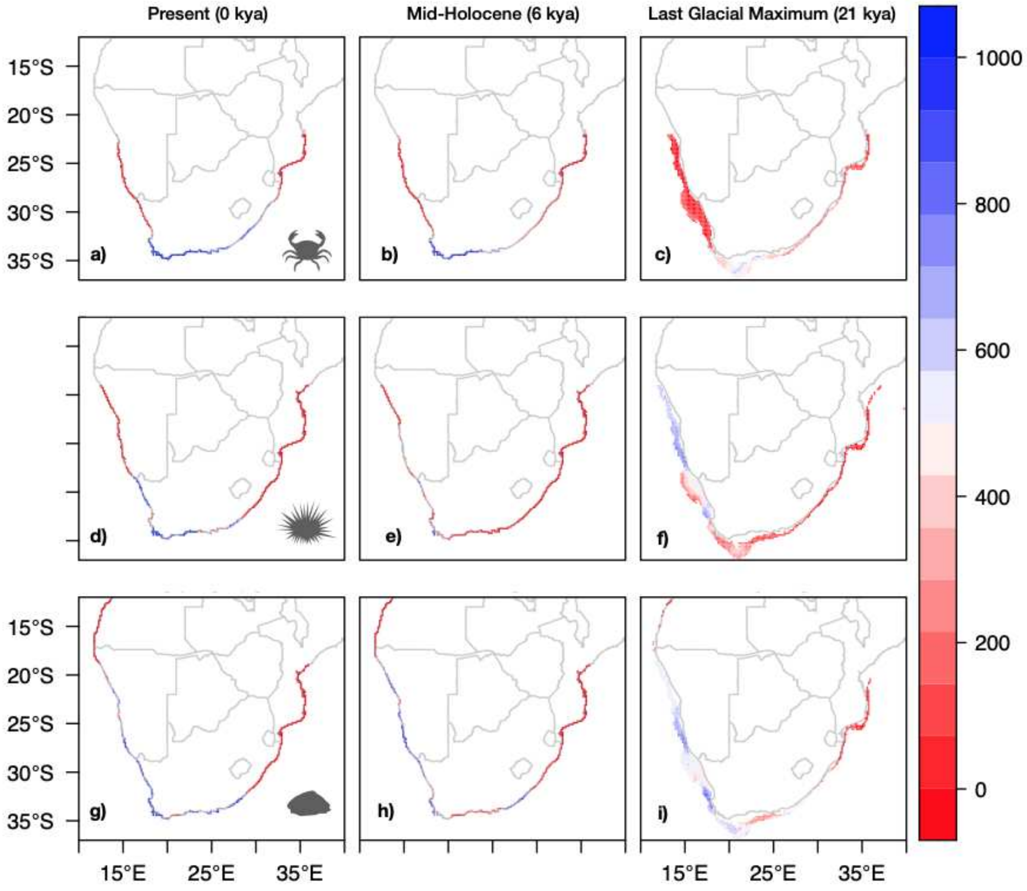

Figure 2 - Species distribution model results projecting distributions within the exposed coastal habitat at the present day (Current; a, d, g), Mid-Holocene (b, e, h), and Last Glacial Maximum (c, f, i), for *C. punctatus* (a-c), *P. angulosus* (d-f), and *S. granularis* (g-i). Habitat suitability ranges from zero to 1000, with areas in blue representing high habitat suitability, and areas in red representing those of low suitability. Projection in the maps: WGS84

Spatial distributions of molecular diversity

Intraspecific spatial patterns of molecular diversity were similar between the mtDNA and SNP datasets (Fig. 3). Broadly, the south coast had highest diversity levels for *C. punctatus*, with the range-edges harbouring lower diversity in both molecular datasets (Fig. 3a; Tables S3-4, Appendix S1). However, both *P. angulosus* and *S. granularis* generally showed higher diversity along the west coast (Fig. 3c,e; Tables S3-4, Appendix S1).

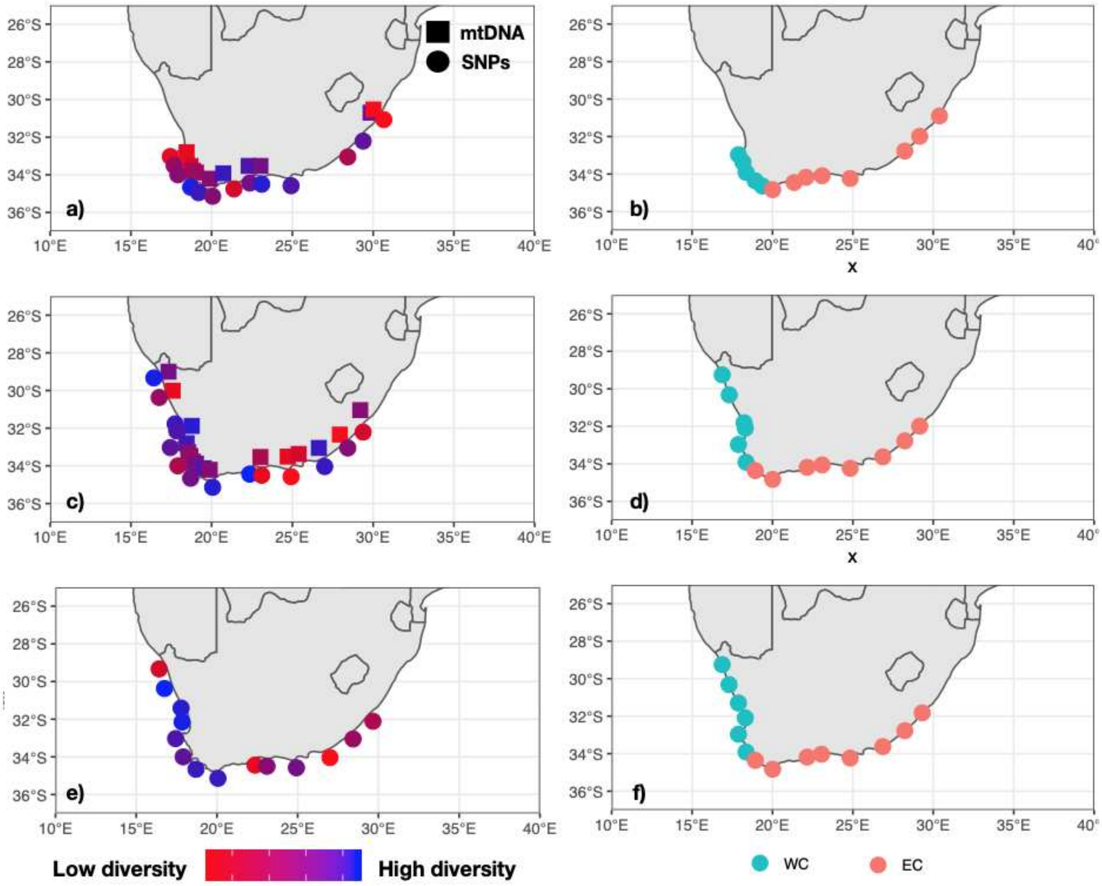

Figure 3- Present day patterns of molecular diversity for *C. punctatus* (a), *P. angulosus* (c), and *S. granularis* (e) with the landward square points representing mtDNA values, and seaward circular points representing SNP values. Sites in red represent lower diversity, and sites in blue represent higher diversity values per species. The genomic sample sites, and whether they fall into the west coast (WC) or east (EC) coast lineages for *C. punctatus* (b), *P. angulosus* (d), and *S. granularis* (f) are also shown. Projection in the maps: WGS84

Comparing region-wide habitat stability and genomic diversity

The redundancy analyses supported niche divergence of eastern and western sites in each species as previously described (Fig. 3b,d,f; Fig. S2, Appendix S1; Mmonwa et al., 2015; Muller et al., 2012; Wright et al., 2015). Broadly, the lineage level SDMs revealed similar hindcasted distributions to the species level SDMs, with *C. punctatus* showing a range reduction through the past, and *P. angulosus* and *S. granularis* expanding up western southern Africa into the LGM (Fig. S3-S5, Appendix S1). The SDMs run per lineage per species showed intraspecific differences in range shifts since the LGM (Fig. 4). Both eastern and western lineages in *C. punctatus* showed range reductions in the past compared to present day, with the eastern population showing larger range losses (Fig. 4a,b). For *P. angulosus* and *S. granularis*, the western lineages displayed larger range sizes during the Mid-Holocene and LGM compared to present day, while the eastern lineages had reduced range sizes (Fig. 4c-f). Eastern and western lineage sites did not significantly differ in H_E for *C. punctatus* or *P. angulosus*, but in *S. granularis*, the western sites had significantly higher H_E values (Fig. 4).

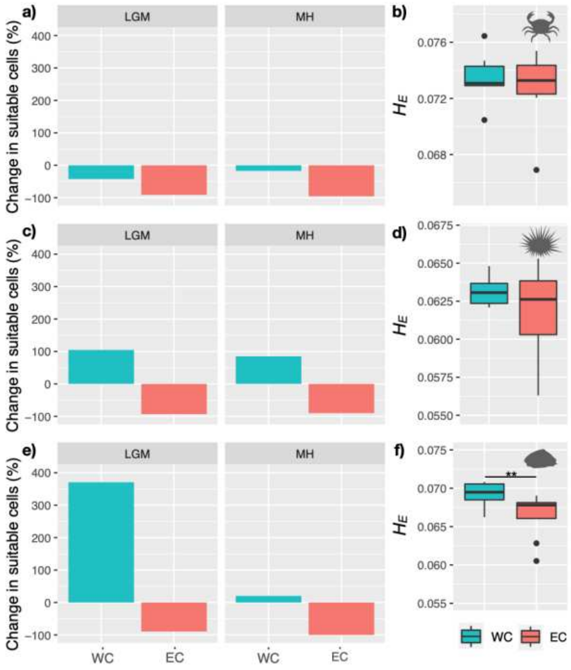

Figure 4- Change in suitable map cells from the lineage SDMs (compared to the present day) is shown for either the west coast (WC) or east coast (EC) sites (see Figure 3 for eastern and western site distinctions), at the Last Glacial Maximum (LGM; 21 kya), Mid-Holocene (MH; 6 kya) for *C. punctatus* (a), *P. angulosus* (c), and *S. granularis* (e). H_E values per lineage are also shown for *C. punctatus* (b), *P. angulosus* (d), and *S. granularis* (f). * = significant at $p < 0.05$, ** = significant at $p < 0.01$.

Relationships between paleoclimatic stability and diversity

Linear models suggested different drivers of molecular variation among and within species across genetic and genomic datasets. The only significant relationship between sea-level variability and molecular diversity was seen in *h* values of *C. punctatus* (P=0.038; R^2=0.37). The relationship between marginal distance and molecular diversity was only significant with *h* for *C. punctatus* (P=0.02; R^2=0.47), and with H_E for *S. granularis* (P=0.04; R^2=0.25).

Marginal distance best explained *C. punctatus* genetic diversity (ωi=0.53), yet the null model ranked highest in explaining genomic diversity in this species (Tables 1 & 2). The null model, followed by climatic variability, was the best ranking model for both genetic and genomic diversity of *P. angulosus* (Tables 1 & 2). Marginal distance best explained genomic diversity of *S. granularis* (ωi=0:43; Tables 1 & 2).

Table 1- Linear models explaining genetic mtDNA haplotype diversity (*h*) for *C. punctatus* and *P. angulosus.* Number of parameters (K), Akaike Information Criteria corrected for small sample sizes (AICc), differences in AICc (ΔAICc), and AICc weights (ωi) are shown per model. Models are listed by rank, with ΔAICc < 2 models highlighted in bold.

Genetic diversity (*h*) model outputs				
Model	K	AICc	ΔAICc	ωi
C. punctatus				
Marginal distance	3	-7.29	0.00	0.56
Sea-level variability	3	-5.48	1.81	0.23
Null	2	-4.04	3.25	0.11
Marginal distance + Sea-level variability	4	-1.76	5.52	0.04
Marginal distance + Climatic variability	4	-1.61	5.67	0.03
Climatic + Sea-level variability	4	0.18	7.47	0.01
Climatic variability	3	0.22	7.50	0.01
Marginal distance + Climatic + Sea-level variability	5	6.82	14.11	0.00
P. angulosus				
Null	2	-45.36	0.00	0.31
Climatic variability	3	-45.34	0.02	0.30
Marginal distance	3	-44.15	1.21	0.17
Sea-level variability	3	-42.38	2.98	0.07
Marginal distance + Climatic variability	4	-42.01	3.35	0.06
Climatic + Sea-level variability	4	-41.85	3.51	0.05
Marginal distance + Sea-level variability	4	-41.01	4.35	0.03
Marginal distance + Climatic + Sea-level variability	5	-37.96	7.40	0.01

Table 2- Linear models explaining genomic SNP expected heterozygosity (H_E) for *C. punctatus*, *P. angulosus*, and *S. granularis*. Number of parameters (K), Akaike Information Criteria corrected for small sample sizes (AICc), differences in AICc (ΔAICc), and AICc weights (ωi) are shown per model. Models are listed by rank, with ΔAICc < 2 models highlighted in bold.

Genomic diversity (H_E) model outputs				
Model	K	AICc	ΔAICc	ωi
C. punctatus				
Null	2	-115.84	0.00	0.36
Marginal distance	3	-115.78	0.05	0.35
Climatic variability	3	-112.62	3.22	0.07
Sea-level variability	3	-112.60	3.23	0.07
Marginal distance + Sea-level variability	4	-112.54	3.30	0.07
Marginal distance + Climatic variability	4	-112.37	3.47	0.06
Marginal distance + Climatic + Sea-level variability	5	-108.98	6.86	0.01
Climatic + Sea-level variability	4	-108.47	7.36	0.01
P. angulosus				
Null	2	-126.17	0.00	0.41
Climatic variability	3	-125.47	0.69	0.29
Sea-level variability	3	-123.27	2.89	0.10
Marginal distance	3	-122.98	3.19	0.08
Climatic + Sea-level variability	4	-122.21	3.96	0.06
Marginal distance + Climatic variability	4	-121.51	4.66	0.04
Marginal distance + Sea-level variability	4	-119.24	6.93	0.01
Marginal distance + Climatic + Sea-level variability	5	-117.77	8.40	0.01
S. granularis				
Marginal distance	3	-121.97	0.00	0.32
Marginal distance + Sea-level variability	4	-121.30	0.67	0.23
Sea-level variability	3	-120.67	1.30	0.17
Null	2	-120.18	1.79	0.13
Marginal distance + Climatic variability	4	-118.88	3.09	0.07
Climatic variability	3	-117.42	4.56	0.03
Marginal distance + Climatic + Sea-level variability	5	-117.37	4.60	0.03
Climatic + Sea-level variability	4	-117.00	4.97	0.03

Testing the central-margin hypothesis with population differentiation

Assessments of molecular differentiation between core and edge populations using permutation tests found both significant and non-significant differences in pairwise F_{ST} values. For *C. punctatus*, mean pairwise F_{ST} values were significantly higher in edge than core sites in the mtDNA data (P=0.003), but not the SNP data. Pairwise F_{ST} values were significantly higher in edge than core sites for *P. angulosus*, but this was again only seen in the mtDNA data (P=6.18e-06). The only SNP dataset that showed significantly higher F_{ST} values in the edge sites was for *S. granularis* (P=4.57e-4).

Discussion

Quantifying molecular diversity and identifying the roles of historical and contemporary biogeographic processes shaping intraspecific diversity is of key importance to predict the vulnerability of populations to future environmental change (Gougherty et al., 2020). However, formal tests comparing the effects of past climate and contemporary range position on extant molecular diversity patterns currently remain limited, especially in a marine context. The multi-species, multi-marker analysis used here to distinguish the relative weights of historical versus contemporary biogeographic patterns shaping molecular diversity offered inconsistent results, suggesting that neither refugial-persistence nor central-margin hypotheses may fully explain genetic variation in marine species. While intraspecific molecular diversity showed distinct spatial patterns between species (Fig. 3), areas with high diversity levels within each species roughly correlated with areas identified as past refugia by the SDMs (Fig. 2; Fig. S3-S5, Appendix S1). Moreover, the lineages which demonstrated reduced range losses since the LGM corresponded to regions with higher genomic diversity today, but this was not consistent across species (Fig. 4). The linear models testing refugial-persistence offered little support for past climatic variability driving patterns of molecular diversity (Tables 1 & 2). In contrast, the linear regression models supported the central-margin hypothesis, with significant relationships between marginal distance and molecular diversity, yet this was not consistent across molecular markers or species (Tables 1 & 2). Additional tests for central-margin patterns in genetic differentiation also offered inconsistent results, with no one marker or species having significantly higher genetic differentiation at marginal sites. Together, these factors support the notion that historical and contemporary biogeographic processes cannot be easily untangled within this complex marine environment, and that generating molecular information is invaluable for understanding the evolutionary trajectories of coastal species in the region.

Looking to the past: multi-species range shifts through time

Inferences of habitat stability and paleodistribution modelling of marine species are limited within the Southern Hemisphere (Robinson et al., 2017). During the late Quaternary, southern Africa's coastline experienced unique environmental changes. The Mid-Holocene was characterised by warmer temperatures, a stronger Agulhas Current, with weakened upwelling along the west coast (Cohen & Tyson, 1995). Going back further to the LGM, South Africa experienced cooler SSTs (Stone, 2014), a weakened Agulhas Current (Franzese et al., 2006) and an increase in SSS (Holloway et al., 2016). These climatic changes likely influenced species' distributions in distinct ways. As such, the results suggested that historical habitat suitability differed among species, with the crab *C. punctatus* having the most severe range-shifts between the present and LGM (Fig. 2). The SDMs for *C. punctatus* suggested a range-reduction during this time (Fig. 2), perhaps due to an influx of low salinity waters from the Mozambique Current given the increase in freshwater run-off from the Zambezi River during the Mid-Holocene (Holloway et al., 2016), and to drops in temperature during the LGM (Stone, 2014). Both salinity and temperature have been shown to play an important role in the physiology of adult *C. punctatus* individuals (Winch & Hodgson, 2007), as well as in the larval development of other decapod species (Anger, 2003).

The importance of salinity in predicting the distribution of *C. punctatus*, as well as *P. angulosus*, highlights how marine SDMs based solely on temperature can lead to potentially over- or under-estimating distributions. In contrast, the distribution of *S. granularis* is predominantly influenced by SSTs and minimum air temperature. The broad array of important environmental predictors of species distributions is unexpected, as previous SDMs on coastal species indicate that

SST is the primary driver of marine species distributions (Assis et al., 2014; Chefaoui & Serrão, 2017; Neiva et al., 2014). However, most coastal SDMs fail to include both atmospheric and oceanic predictor variables, likely due to differences in the spatial and temporal characteristics between data types. In addition, variable importance outputs of SDMs has recently been shown to vary by algorithm type (Smith & Santos, 2020), and thus the importance of each predictor variable in the models used here come with some uncertainty.

Similar results were found in the historical distributions of *P. angulosus* and *S. granularis*. Both showed a range reduction and shift toward the west coast of southern Africa in the Mid-Holocene, which is most likely a response to the increase in SST during the last interglacial period (Fig. 2; Crowley & North, 1991). The two species also showed a range expansion along the western coast during the LGM as SSTs decreased (Fig. 2). These findings differ from the majority of other hindcasting SDMs on cool-temperate marine species, which showed range reductions during the LGM (Assis et al., 2014; Neiva et al., 2014; Waltari & Hickerson, 2013). However, these studies were conducted in the Northern Hemisphere, where ice-sheets inhibited northern range edges (Provan, 2013), compared to our study system that lacked permanent ice cover during the last glacial period, and where sea-level change and associated changes in ocean topology and currents are expected to be primary barriers to dispersal (Otto-Bliesner et al., 2006; Toms et al., 2014).

This study also corroborates recent calls for lineage-level SDMs to be used in support of habitat suitability outputs of species-level assessments (Peñalver-Alcázar et al., 2021; Zhang et al., 2021). Here, we did find evidence of niche divergence between lineages (Fig. S2, Appendix S1), but the SDMs run per lineage offered loosely similar suitability patterns as those run on the entire species (Fig. 2; Fig. S3-S5, Appendix S1). However, species-level SDMs over-predicted the range of the east coast lineages, and under-predicted the western lineages for each species. This finding mirrors other studies which suggest that species-level SDMs can both over- and under-predict the ranges of individual lineages, and thus a hierarchical approach should be used to evaluate future or past range shifts (D’Amen et al., 2013; Zhang et al., 2021). Also, it is important to note that the SDMs used here did not account for past changes in oceanographic dynamics such as current and upwelling systems. As changes in ocean currents, such as a reduction in the Agulhas leakage during the LGM, likely influenced population connectivity of other marine species in the region (Silva et al., 2021), these processes could further influence extant levels of genetic diversity.

Comparing historical and contemporary drivers of molecular variation

Characterising the relative importance of historical climatic stability and contemporary range distributions in shaping patterns of molecular diversity is crucial to help predict how intraspecific variation may be altered by future global change. To identify prominent spatio-temporal drivers of diversity in marine species, this study tested whether areas near the centre of distributional ranges or areas of past refugial-persistence correspond to areas of high extant molecular diversity. The output maps from the SDMs and genetic diversity offered support for refugial-persistence, as areas of higher suitability at the MH and LGM corresponded well with areas of contemporary elevated molecular diversity (Fig. 2, Fig. 3; Fig. S3-S5, Appendix S1). The results also showed that lineages of *P. angulosus* and *S. granularis* which experienced higher habitat suitability since the LGM generally had higher molecular diversity at present, although this was only significant in *S. granularis* (Fig. 4). However, the refugial-persistence hypothesis was not strongly supported by the linear models, as climatic and sea-level variability had weak relationships with molecular diversity patterns per species (Tables 1 & 2). The discrepancy between the visual SDM outputs supporting and linear models refuting refugial-persistence in the study species could be driven by fine-scale variation in diversity. For example, while the west coast region generally displayed higher diversity in the urchin and limpet, there are sites within that region which go against this pattern. Thus, sample density within each region may not be sufficient to detect refugial-persistence within the linear models, which is likely why many of the studies testing the relationship between climatic stability and genetic diversity use solely visual outputs (Beatty & Provan, 2011b; Chefaoui et al., 2017; Neiva et al., 2014; Waltari & Hickerson, 2013).

Multiple studies have related hindcasted distributions to genetic patterns, providing evidence for refugial-persistence (Hewitt, 2000; Theodoridis et al., 2020), although contrasting patterns do exist. For example, Bermejo et al., (2018) identified incongruence between LGM suitability and genetic variation in the algae *Cystoseira tamariscifolia*, which they expected was due to range expansions during the Mid-Holocene. Sunday et al. (2012) suggested that compared to terrestrial counterparts, marine species are expected to show more immediate responses to climatic variation. Therefore, the molecular diversity of marine invertebrates with high fecundity and large dispersal ranges may reflect climatic variation from the Mid-Holocene rather than LGM. This is further supported by de De Kort et al. (2020), who found that molluscs, which are often capable of rapid population turnover compared to the other terrestrial taxa, displayed genetic diversity patterns strongly related to Mid-Holocene, rather than LGM, stability. These results, as well as those shown here, provide evidence that multiple historical time points should ideally be included to best estimate the rate which climatic fluctuations imprint on the molecular diversity of marine species.

When testing the central-margin hypothesis, linear models offered inconsistent support across species and marker types (Tables 1 & 2). Patterns of genetic differentiation supported the central-margin hypothesis, but this was only significant in the mtDNA data for *C. punctatus* and *P. angulosus*, and the SNP data for *S. granularis*. Other studies testing central-margin patterns in marine invertebrates also report inconsistencies across either different marker types (Ntuli et al., 2020), or diversity metrics from a single marker (Cahill & Levinton, 2016). This is expected, as different molecular markers and metrics vary in characteristics such as sample size and mutation rate (Morin et al., 2004). However, it is interesting to note that two marker types displayed highly similar spatial arrays of diversity for *C. punctatus* and *P. angulosus*, corroborating other studies which suggest that while different markers may display varying degrees of genetic diversity, the spatial arrangement of diversity is the same (Beatty & Provan, 2011; Kraft et al., 2020).

Previous studies assessing species abundances of southern African sandhoppers (Baldanzi et al., 2013) and North Pacific coastal invertebrates (Sagarin & Gaines, 2002) both offered little support for the central-margin hypothesis, which suggests that the inconsistent findings of genetic central-margin patterns here may be due to lack of a central-margin array in abundance. It is difficult to distinguish if the lack of firm evidence for central-margin patterns here is due to the assumption of a decline in abundances towards range edges not being met (Kennedy et al., 2020), as abundance data is unavailable. However, a recent test of the central-marginal hypothesis in South African mussels (Ntuli et al., 2020) found core-to-edge decreases in abundances and connectivity, yet no clear core-edge patterns in molecular diversity. The authors hypothesized that this is due to central-margin dynamics only imprinting on the genetic composition of species when connectivity levels are below a certain threshold. As the study species here display relatively low levels of genetic differentiation (Nielsen et al., 2020b), it can be assumed that high connectivity between the core and edge sites may lead to genetic diversity patterns that deviate from clear refugial-persistence or central-margin patterns. There was a lack of support for either hypothesis when using linear models, which contrasts with the significant model outputs in recent tests of diversity drivers on a global scale (De Kort et al., 2020; Manel et al., 2020). This could be due to the relatively low number of sample sites and species employed in this study, or due to broad-scale diversity patterns being oversimplified within a fine-scale spatial context. Future studies examining drivers of molecular diversity should ideally include more dense sampling over a broader array of species, to fully test the relative weights of historical and contemporary environmental variation shaping evolutionary patterns at the regional scale.

Furthermore, it is likely that environmental features, such as temperature, are leading drivers of evolutionary patterns compared to dispersal traits, as we found few similarities between species with similar PLDs. For example, the limpet and crab have comparable PLDs of ~10 days, with the urchin having a longer PLD of ~50 days, but the results throughout are most similar between the urchin and the limpet, which share cool-temperate thermal niches, compared to the warm-temperate thermal niche of the crab (Branch & Branch, 2018). We also found that mtDNA, which is expected to capture more historical demographic changes, showed similar diversity patterns to

SNPs, which are expected to generally reflect more recent evolutionary processes (Morin et al., 2004; Wang, 2011). Similar diversity patterns from both markers were consistent for the crab and urchin, even though they differ in their PLDs. This is unexpected, as the time lag (i.e. the time it takes for an environmental occurrence to become detectable in molecular data), is thought to differ between marker types and dispersal levels (Epps & Keyghobadi, 2015). However, as the study species have large effective population sizes, overlapping generation times, and fewer and larger subpopulation structure, all of which increase time lags, these could be driving the similar patterns between marker types (Epps & Keyghobadi, 2015). While outside the scope of this study, future steps to disentangle the roles historical and contemporary processes on molecular diversity can include modelling demographic histories, estimating migration rates, and reconstructing oceanographic currents, all of which can help us understand how the timing of divergence events shapes genetic variation (Dawson 2005; Marko & Hart, 2011).

Concluding remarks

Tests of the central-margin hypothesis on marine species often omit the effects of past climatic change and possible glacial persistence (Liggins et al., 2015). Although our multi-species, multi-marker results offered support for both central-margin and refugial-persistence, we did not find conclusive evidence for either hypothesis. It is likely that both of these hypotheses oversimplify the processes that shape extant patterns of molecular diversity of species (Lal et al., 2017; Maggs et al., 2008). This notion is mirrored by recent global assessments reporting diverse findings in terrestrial systems, suggesting that either contemporary distributional position (De Kort et al., 2020), or historical climatic stability (Theodoridis et al., 2020) is the key driver of molecular diversity. While broad biogeographic theories, such as refugial-persistence and central-margin hypotheses, may significantly align with large-scale global diversity patterns in terrestrial taxa, the results here suggest that they may be unsuitable for marine species, and/or over smaller spatial scales. Overall, the results suggest that marine species within temperate bioregions in the Southern Hemisphere, may have nuanced biogeographic and evolutionary histories. The inconsistent central-margin findings further suggest that patterns of molecular variation in marine invertebrates with high dispersal abilities may be more complex than the ecological gradients they may exist in. Ultimately, further work is needed to assess the processes that have shaped, and will continue to the shape, the molecular diversity of these organisms in this dynamic, and highly vulnerable, marine region.

Figure Legends

Figure 1- Mean sea surface temperatures (Mean SST; a), mean sea surface salinity (Mean SSS) based on monthly averages from MARSPEC data (Sbrocco, 2014) and the Cool-temperate (CT), Southwest (SW), Warm-temperate (WT), Subtropical (ST) and Tropical (TP) bioregions of South Africa (b), as well as the dominant current systems and the genetic (mtDNA; represented by landward squares) and genomic (SNPs; represented by seaward circles) sample sites (c) are shown. Projection in the maps: WGS84

Figure 2 - Species distribution model results projecting distributions within the exposed coastal habitat at the present day (Current; a, d, g), Mid-Holocene (b, e, h), and Last Glacial Maximum (c, f, i), for *C. punctatus* (a-c), *P. angulosus* (d-f), and *S. granularis* (g-i). Habitat suitability ranges from zero to 1000, with areas in blue representing high habitat suitability, and areas in red representing those of low suitability. Projection in the maps: WGS84

Figure 3- Present day patterns of molecular diversity for *C. punctatus* (a), *P. angulosus* (c), and *S. granularis* (e) with the landward square points representing mtDNA values, and seaward circular points representing SNP values. Sites in red represent lower diversity, and sites in blue represent higher diversity values per species. The genomic sample sites, and whether they fall into the west coast (WC) or east (EC) coast lineages for *C. punctatus* (b), *P. angulosus* (d), and *S. granularis* (f) are also shown. Projection in the maps: WGS84

Figure 4- Change in suitable map cells from the lineage SDMs (compared to the present day) is shown for either the west coast (WC) or east coast (EC) sites (see Figure 3 for eastern and western site distinctions), at the Last Glacial Maximum (LGM; 21 kya), Mid-Holocene (MH; 6 kya) for *C. punctatus* (a), *P. angulosus* (c), and *S. granularis* (e). H_E values per lineage are also shown for *C. punctatus* (b), *P. angulosus* (d), and *S. granularis* (f). * = significant at $p < 0.05$, ** = significant at $p < 0.01$.

Data availability

Presence data, and species distribution model and linear model R scripts available via Github:

https://github.com/vonderHeydenLab/Nielsen_et_al_2021_JBiogeo

Acknowledgements

Thank you to the two anonymous reviewers for their comments that improved the manuscript. We thank G. Midgley and J. Singarayer for their help in obtaining paleoclimatic reconstructions. We also thank N. Phair for her assistance with the SDMs. We thank G. Branch for his valuable comments on an earlier version of the manuscript. The research was funded by NRF grants 92788 and 105842 to S.V.D.H. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement 837990-DiMaS-H2020-MSCA-IF-2018 to R.H. Permit RES2018/54 from the South African Dept. of Agriculture, Forestry and Fisheries was obtained for all sampling.

References:

Acevedo-Limón, L., Oficialdegui, F. J., Sánchez, M. I., & Clavero, M. (2020). Historical, human, and environmental drivers of genetic diversity in the red swamp crayfish (*Procambarus clarkii*) invading the Iberian Peninsula. *Freshwater Biology*, *65*(8), 1460–1474. https://doi.org/10.1111/fwb.13513

Allouche, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, *43*(6), 1223–1232. https://doi.org/10.1111/j.1365-2664.2006.01214.x

Anger, K. (2003). Salinity as a key parameter in the larval biology of decapod crustaceans. *Invertebrate Reproduction & Development*, *43*(1), 29–45. https://doi.org/10.1080/07924259.2003.9652520

Araújo, M. B., & New, M. (2007). Ensemble forecasting of species distributions. *Trends in Ecology & Evolution*, *22*(1), 42–47. https://doi.org/10.1016/j.tree.2006.09.010

Assis, J., Serrão, E. A., Claro, B., Perrin, C., & Pearson, G. A. (2014). Climate-driven range shifts explain the distribution of extant gene pools and predict future loss of unique lineages in a marine brown alga. *Molecular Ecology*, *23*(11), 2797–2810. https://doi.org/10.1111/mec.12772

Assis, Jorge, Araújo, M. B., & Serrão, E. A. (2018). Projected climate changes threaten ancient refugia of kelp forests in the North Atlantic. *Global Change Biology*, *24*(1), e55–e66. https://doi.org/10.1111/gcb.13818

Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., Reeb, C. A., & Saunders, N. C. (1987). Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, *18*, 489–522. JSTOR.

Baldanzi, S., McQuaid, C. D., Cannicci, S., & Porri, F. (2013). Environmental domains and range-limiting mechanisms: Testing the abundant centre hypothesis using southern African sandhoppers. *PLoS ONE*, *8*(1). https://doi.org/10.1371/journal.pone.0054598

Beatty, G. E., & Provan, J. (2011a). Phylogeographic analysis of North American populations of the parasitic herbaceous plant *Monotropa hypopitys* L. reveals a complex history of range expansion from multiple late glacial refugia. *Journal of Biogeography*, *38*(8), 1585–1599. https://doi.org/10.1111/j.1365-2699.2011.02513.x

Beatty, G. E., & Provan, J. (2011b). Comparative phylogeography of two related plant species with overlapping ranges in Europe, and the potential effects of climate change on their intraspecific genetic diversity. *BMC Evolutionary Biology*, *11*(1), 29. https://doi.org/10.1186/1471-2148-11-29

Bennett, K. D., Tzedakis, P. C., & Willis, K. J. (1991). Quaternary refugia of north European Trees. *Journal of Biogeography*, *18*(1), 103–115. JSTOR. https://doi.org/10.2307/2845248

Bermejo, R., Chefaoui, R. M., Engelen, A. H., Buonomo, R., Neiva, J., Ferreira-Costa, J., Pearson, G. A., Marbà, N., Duarte, C. M., Airoldi, L., Hernández, I., Guiry, M. D., & Serrão, E. A. (2018). Marine forests of the Mediterranean-Atlantic *Cystoseira tamariscifolia* complex show a southern Iberian genetic hotspot and no reproductive isolation in parapatry. *Scientific Reports*, *8*(1), 1–13. https://doi.org/10.1038/s41598-018-28811-1

Bowen, B. W., Gaither, M. R., DiBattista, J. D., Iacchei, M., Andrews, K. R., Grant, W. S., Toonen, R. J., & Briggs, J. C. (2016). Comparative phylogeography of the ocean planet. *Proceedings of the*

National Academy of Sciences, *113*(29), 7962–7969. https://doi.org/10.1073/pnas.1602404113

Branch, G., & Branch, M. (2018). The Living Shores of Southern Africa. *Penguin Random House*, South Africa.

Brown, J. H. (1984). On the relationship between abundance and distribution of species. *The American Naturalist*, *124*(2), 255–279.

Brown, J. L., & Yoder, A. D. (2015). Shifting ranges and conservation challenges for lemurs in the face of climate change. *Ecology and Evolution*, *5*(6), 1131–1142. https://doi.org/10.1002/ece3.1418

Cahill, A. E., & Levinton, J. S. (2016). Genetic differentiation and reduced genetic diversity at the northern range edge of two species with different dispersal modes. *Molecular Ecology*, *25*(2), 515–526. https://doi.org/10.1111/mec.13497

Chefaoui, R. M., Duarte, C. M., & Serrão, E. A. (2017). Palaeoclimatic conditions in the Mediterranean explain genetic diversity of *Posidonia oceanica* seagrass meadows. *Scientific Reports*, *7*(1), 2732. https://doi.org/10.1038/s41598-017-03006-2

Chefaoui, R. M., & Serrão, E. A. (2017). Accounting for uncertainty in predictions of a marine species: Integrating population genetics to verify past distributions. *Ecological Modelling*, *359*, 229–239. https://doi.org/10.1016/j.ecolmodel.2017.06.006

Cohen, A. L., & Tyson, P. D. (1995). Sea-surface temperature fluctuations during the Holocene off the south coast of Africa: Implications for terrestrial climate and rainfall. *The Holocene*, *5*(3), 304–312. https://doi.org/10.1177/095968369500500305

Collart, F., Hedenäs, L., Broennimann, O., Guisan, A., & Vanderpoorten, A. (2021). Intraspecific differentiation: Implications for niche and distribution modelling. *Journal of Biogeography*, *48*(2), 415–426. https://doi.org/10.1111/jbi.14009

Crowley, T. J., & North, G. R. (1991). *Paleoclimatology*. Oxford Monographs on Geology and Geophysics. https://www.osti.gov/biblio/6506929

D'Amen, M., Zimmermann, N. E., & Pearman, P. B. (2013). Conservation of phylogeographic lineages under climate change: African mammals and global warming. *Global Ecology and Biogeography*, *22*(1), 93–104. https://doi.org/10.1111/j.1466-8238.2012.00774.x

Dawson, M. N. (2005). Incipient speciation of *Catostylus mosaicus* (Scyphozoa, Rhizostomeae, Catostylidae), comparative phylogeography and biogeography in south-east Australia. *Journal of Biogeography*, *32*(3), 515-533. https://doi.org/10.1111/j.1365-2699.2004.01193.x

De Kort, H., Baguette, M., Lenoir, J., & Stevens, V. M. (2020). Toward reliable habitat suitability and accessibility models in an era of multiple environmental stressors. *Ecology and Evolution*, *10*(20), 10937-10952. https://doi.org/10.1002/ece3.6753

Eckert, C. G., Samis, K. E., & Lougheed, S. C. (2008). Genetic variation across species' geographical ranges: The central–marginal hypothesis and beyond. *Molecular Ecology*, *17*(5), 1170–1188. https://doi.org/10.1111/j.1365-294X.2007.03659.x

Engler, R., Randin, C. F., Thuiller, W., Dullinger, S., Zimmermann, N. E., Araújo, M. B., Pearman, P. B., Lay, G. L., Piedallu, C., Albert, C. H., Choler, P., Coldea, G., Lamo, X. D., Dirnböck, T., Gégout, J.-C., Gómez-García, D., Grytnes, J.-A., Heegaard, E., Høistad, F., … Guisan, A. (2011). 21st century climate change threatens mountain flora unequally across Europe. *Global Change Biology*, *17*(7), 2330–2341. https://doi.org/10.1111/j.1365-2486.2010.02393.x

Epps, C. W., & Keyghobadi, N. (2015). Landscape genetics in a changing world: Disentangling historical and contemporary influences and inferring change. *Molecular Ecology*, *24*(24), 6021–6040. https://doi.org/10.1111/mec.13454

Excoffier, L., Foll, M., & Petit, R. (2009). Genetic consequences of range expansions. *Annual Review of Ecology and Systematics*, *40*, 481–501. https://doi.org/10.1146/annurev.ecolsys.39.110707.173414

Excoffier, L., Laval, G., & Schneider, S. (2005). Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics*, *1*, 117693430500100. https://doi.org/10.1177/117693430500100003

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*(12), 4302–4315. https://doi.org/10.1002/joc.5086

Fielding, A. H., & Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, *24*(1), 38–49. https://doi.org/10.1017/S0376892997000088

Fischer, M. C., Rellstab, C., Leuzinger, M., Roumet, M., Gugerli, F., Shimizu, K. K., Holderegger, R., & Widmer, A. (2017). Estimating genomic diversity and population differentiation – an empirical comparison of microsatellite and SNP variation in *Arabidopsis halleri*. *BMC Genomics*, *18*(1), 69. https://doi.org/10.1186/s12864-016-3459-7

Forester, B. R., DeChaine, E. G., & Bunn, A. G. (2013). Integrating ensemble species distribution modelling and statistical phylogeography to inform projections of climate change impacts on species distributions. *Diversity and Distributions*, *19*(12), 1480–1495. https://doi.org/10.1111/ddi.12098

Francisco, S. M., Almada, V. C., Faria, C., Velasco, E. M., & Robalo, J. I. (2014). Phylogeographic pattern and glacial refugia of a rocky shore species with limited dispersal capability: The case of Montagu’s blenny (*Coryphoblennius galerita*, Blenniidae). *Marine Biology*, *161*(11), 2509–2520. https://doi.org/10.1007/s00227-014-2523-6

Franzese, A., Hemming, S., Goldstein, S., & Anderson, R. (2006). Reduced Agulhas Leakage during the Last Glacial Maximum inferred from an integrated provenance and flux study. *Earth and Planetary Science Letters*, *250*(1–2), 72–88. https://doi.org/10.1016/j.epsl.2006.07.002

Furrer, R., Nychka, D., Sain, S., & Nychka, M. D. (2009). Package ‘fields’. *R Foundation for Statistical Computing, Vienna, Austria. http://www. idg. pl/mirrors/CRAN/web/packages/fields/fields.*

GBIF. (2020). GBIF.org (20 November 2020) GBIF Occurrence Download. https://doi.org/10.15468/dl.5n4fq3.

Gougherty, A. V., Chhatre, V. E., Keller, S. R., & Fitzpatrick, M. C. (2020). Contemporary range position predicts the range-wide pattern of genetic diversity in balsam poplar (*Populus balsamifera* L.). *Journal of Biogeography*, *47*(6), 1246–1257. https://doi.org/10.1111/jbi.13811

Guisan, A., & Thuiller, W. (2005). Predicting species distribution: Offering more than simple habitat models. *Ecology Letters*, *8*(9), 993–1009. https://doi.org/10.1111/j.1461-0248.2005.00792.x

Guo, Q. (2012). Incorporating latitudinal and central-marginal trends in assessing genetic variation across species ranges. *Molecular Ecology*, *21*(22), 5396–5403. https://doi.org/10.1111/mec.12012

Hartig, F., & Hartig, M. F. (2017). Package 'DHARMa.' *R Package.*

Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, *405*(6789), 907–913. https://doi.org/10.1038/35016000

Heyden, S. von der, Bowie, R. C. K., Prochazka, K., Bloomer, P., Crane, N. L., & Bernardi, G. (2011). Phylogeographic patterns and cryptic speciation across oceanographic barriers in South African intertidal fishes. *Journal of Evolutionary Biology*, *24*(11), 2505–2519. https://doi.org/10.1111/j.1420-9101.2011.02382.x

Hijmans, R. J., van Etten, J., Cheng, J., Mattiuzzi, M., Sumner, M., Greenberg, J. A., Lamigueiro, O. P., Bevan, A., Racine, E. B., Shortbridge, A., & Hijmans, M. R. J. (2015). Package 'raster.' *R Package.*

Hivert, V., Leblois, R., Petit, E. J., Gautier, M., & Vitalis, R. (2018). Measuring genetic differentiation from Pool-seq data. *Genetics*, *210*(1), 315-330. https://doi.org/10.1101/282400

Hoban, S., Dawson, A., Robinson, J. D., Smith, A. B., & Strand, A. E. (2019). Inference of biogeographic history by formally integrating distinct lines of evidence: Genetic, environmental niche and fossil. *Ecography*, *42*(12), 1991–2011. https://doi.org/10.1111/ecog.04327

Holloway, M. D., Sime, L. C., Singarayer, J. S., Tindall, J. C., & Valdes, P. J. (2016). Reconstructing paleosalinity from δ18O: Coupled model simulations of the Last Glacial Maximum, Last Interglacial and Late Holocene. *Quaternary Science Reviews*, *131*, 350–364. https://doi.org/10.1016/j.quascirev.2015.07.007

Hothorn, T., Hornik, K., van de Wiel, M. A., Zeileis, A., & Hothorn, M. T. (2021). Package 'coin'. *Conditional inference procedures in a permutation test framework*.

Karihaloo, J. L., Dwivedi, Y. K., Archak, S., & Gaikwad, A. B. (2003). Analysis of genetic diversity of Indian mango cultivars using RAPD markers. *The Journal of Horticultural Science and Biotechnology*, *78*(3), 285-289. https://doi.org/10.1080/14620316.2003.11511619

Kennedy, J. P., Preziosi, R. F., Rowntree, J. K., & Feller, I. C. (2020). Is the central-marginal hypothesis a general rule? Evidence from three distributions of an expanding mangrove species, *Avicennia germinans* (L.) L. *Molecular Ecology*, *29*(4), 704–719. https://doi.org/10.1111/mec.15365

Kraft, D. W., Conklin, E. E., Barba, E. W., Hutchinson, M., Toonen, R. J., Forsman, Z. H., & Bowen, B. W. (2020). Genomics versus mtDNA for resolving stock structure in the silky shark (*Carcharhinus falciformis*). *PeerJ*, *8*, e10186. https://doi.org/10.7717/peerj.10186

Lal, M. M., Southgate, P. C., Jerry, D. R., Bosserelle, C., & Zenger, K. R. (2017). Swept away: Ocean currents and seascape features influence genetic structure across the 18,000 Km Indo-Pacific distribution of a marine invertebrate, the black-lip pearl oyster *Pinctada margaritifera*. *BMC Genomics*, *18*(1), 66. https://doi.org/10.1186/s12864-016-3410-y

Lawton, J. H. (1993). Range, population abundance and conservation. *Trends in Ecology & Evolution*, *8*(11), 409–413. https://doi.org/10.1016/0169-5347(93)90043-O

Lesbarrères, D., Primmer, C. R., Lodé, T., & Merilä, J. (2006). The effects of 20 years of highway presence on the genetic structure of *Rana dalmatina* populations. *Ecoscience*, *13*(4), 531-538. https://doi.org/10.2980/1195-6860(2006)13[531:TEOYOH]2.0.CO;2

Librado, P., & Rozas, J. (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, *25*(11), 1451–1452. https://doi.org/10.1093/bioinformatics/btp187

Liggins, L., Booth, D. J., Figueira, W. F., Treml, E. A., Tonk, L., Ridgway, T., Harris, D. A., & Riginos, C. (2015). Latitude-wide genetic patterns reveal historical effects and contrasting patterns of turnover and nestedness at the range peripheries of a tropical marine fish. *Ecography*, *38*(12), 1212–1224. https://doi.org/10.1111/ecog.01398

Lira-Noriega, A., & Manthey, J. D. (2014). Relationship of genetic diversity and niche centrality: A survey and analysis. *Evolution*, *68*(4), 1082–1093. https://doi.org/10.1111/evo.12343

Maggs, C. A., Castilho, R., Foltz, D., Henzler, C., Jolly, M. T., Kelly, J., Olsen, J., Perez, K. E., Stam, W., Väinölä, R., Viard, F., & Wares, J. (2008). Evaluating signatures of glacial refugia for north atlantic benthic marine taxa. *Ecology*, *89*(sp11), S108–S122. https://doi.org/10.1890/08-0257.1

Maguire, K. C., Shinneman, D. J., Potter, K. M., & Hipkins, V. D. (2018). Intraspecific niche models for ponderosa ine (*Pinus ponderosa*) suggest potential variability in population-level response to climate change. *Systematic Biology*, *67*(6), 965–978. https://doi.org/10.1093/sysbio/syy017

Maia-Carvalho, B., Vale, C. G., Sequeira, F., Ferrand, N., Martínez-Solano, I., & Gonçalves, H. (2018). The roles of allopatric fragmentation and niche divergence in intraspecific lineage diversification in the common midwife toad (*Alytes obstetricans*). *Journal of Biogeography*, *45*(9), 2146–2158. https://doi.org/10.1111/jbi.13405

Manel, S., Guerin, P.-E., Mouillot, D., Blanchet, S., Velez, L., Albouy, C., & Pellissier, L. (2020). Global determinants of freshwater and marine fish genetic diversity. *Nature Communications*, *11*(1), 1–9. https://doi.org/10.1038/s41467-020-14409-7

Mmonwa, K., Teske, P., McQuaid, C., & Barker, N. (2015). Historical demography of southern African patellid limpets: Congruence of population expansions, but not phylogeography. *African Journal of Marine Science*, *37*(1), 11–20. https://doi.org/10.2989/1814232X.2015.1009165

Marko, P. B., & Hart, M. W. (2011). The complex analytical landscape of gene flow inference. *Trends in Ecology & Evolution*, *26*(9), 448-456. https://doi.org/10.1016/j.tree.2011.05.007

Morin, P. A., Luikart, G., Wayne, R. K., & the SNP workshop group. (2004). SNPs in ecology, evolution and conservation. *Trends in Ecology & Evolution*, *19*(4), 208–216. https://doi.org/10.1016/j.tree.2004.01.009

Muller, C., von der Heyden, S., Bowie, R., & Matthee, C. (2012). Oceanic circulation, local upwelling and palaeoclimatic changes linked to the phylogeography of the Cape sea urchin *Parechinus angulosus*. *Marine Ecology Progress Series*, *468*, 203–215. https://doi.org/10.3354/meps09956

Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences*, *70*, 3321–3323.

Nei, M. (1987). *Molecular Evolutionary Genetics*. Columbia University Press.

Neiva, J., Assis, J., Fernandes, F., Pearson, G. A., & Serrão, E. A. (2014). Species distribution models and mitochondrial DNA phylogeography suggest an extensive biogeographical shift in the high-intertidal seaweed *Pelvetia canaliculata*. *Journal of Biogeography*, *41*(6), 1137–1148. https://doi.org/10.1111/jbi.12278

Nielsen, E. S., Beger, M., Henriques, R., & von der Heyden, S. (2020a). A comparison of genetic and genomic approaches to represent evolutionary potential in conservation planning. *Biological Conservation*, *251*. https://doi.org/10.1016/j.biocon.2020.108770

Nielsen, E. S., Henriques, R., Beger, M., Toonen, R. J., & von der Heyden, S. (2020b). Multi-model seascape genomics identifies distinct environmental drivers of selection among sympatric marine species. *BMC Evolutionary Biology*, *20*(1), 121. https://doi.org/10.1186/s12862-020-01679-4

Ntuli, N. N., Nicastro, K. R., Zardi, G. I., Assis, J., McQuaid, C. D., & Teske, P. R. (2020). Rejection of the genetic implications of the “Abundant Centre Hypothesis” in marine mussels. *Scientific Reports*, *10*(1), 604. https://doi.org/10.1038/s41598-020-57474-0

OBIS. (2020). Ocean Biodiversity Information System. I*ntergovernmental Oceanographic Commission of UNESCO. www.iobis.org.*

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O’hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Wagner, H., & Oksanen, M. J. (2013). Package ‘vegan.’ *Community Ecology Package, Version 2.*, *9*, 1-295.

Otto-Bliesner, B. L., Brady, E. C., Clauzet, G., Tomas, R., Levis, S., & Kothavala, Z. (2006). Last Glacial Maximum and Holocene Climate in CCSM3. *Journal of Climate*, *19*(11), 2526–2544. https://doi.org/10.1175/JCLI3748.1

Pauls, S. U., Nowak, C., Bálint, M., & Pfenninger, M. (2013). The impact of global climate change on genetic diversity within populations and species. *Molecular Ecology*, *22*(4), 925–946. https://doi.org/10.1111/mec.12152

Peñalver-Alcázar, M., Jiménez-Valverde, A., & Aragón, P. (2021). Niche differentiation between deeply divergent phylogenetic lineages of an endemic newt: Implications for Species Distribution Models. *Zoology*, *144*, 125852. https://doi.org/10.1016/j.zool.2020.125852

Phair, N. L., Toonen, R. J., Knapp, I., & Heyden, S. von der. (2019). Shared genomic outliers across two divergent population clusters of a highly threatened seagrass. *PeerJ*, *7*, e6806. https://doi.org/10.7717/peerj.6806

Phair, N. L., Toonen, R. J., Knapp, I. S. S., & von der Heyden, S. (2020). Anthropogenic pressures negatively impact genomic diversity of the vulnerable seagrass *Zostera capensis*. *Journal of Environmental Management*, *255*, 109831. https://doi.org/10.1016/j.jenvman.2019.109831

Pironon, S., Papuga, G., Villellas, J., Angert, A. L., García, M. B., & Thompson, J. D. (2017). Geographic variation in genetic and demographic performance: New insights from an old biogeographical paradigm: The centre-periphery hypothesis. *Biological Reviews*, *92*(4), 1877–1909. https://doi.org/10.1111/brv.12313

Provan, J. (2013). The effects of past, present and future climate change on range-wide genetic diversity in northern North Atlantic marine species. *Frontiers of Biogeography*, *5*(1).

Reusch, T. B. H., Ehlers, A., Hammerli, A., & Worm, B. (2005). Ecosystem recovery after climatic extremes enhanced by genotypic diversity. *Proceedings of the National Academy of Sciences*, *102*(8), 2826–2831. https://doi.org/10.1073/pnas.0500008102

Robinson, N. M., Nelson, W. A., Costello, M. J., Sutherland, J. E., & Lundquist, C. J. (2017). A systematic review of marine-based Species Distribution Models (SDMs) with recommendations for best practice. *Frontiers in Marine Science*, *4*. https://doi.org/10.3389/fmars.2017.00421

Rödin-Mörch, P., Luquet, E., Meyer-Lucht, Y., Richter-Boix, A., Höglund, J., & Laurila, A. (2019). Latitudinal divergence in a widespread amphibian: Contrasting patterns of neutral and adaptive genomic variation. *Molecular Ecology*, *28*(12), 2996–3011. https://doi.org/10.1111/mec.15132

Rodríguez-Rodríguez, E. J., Beltrán, J. F., Tejedo, M., Nicieza, A. G., Llusia, D., Márquez, R., & Aragón, P. (2020). Niche models at inter- and intraspecific levels reveal hierarchical niche differentiation in midwife toads. *Scientific Reports*, *10*(1), 10942. https://doi.org/10.1038/s41598-020-67992-6

Sagarin, R. D., & Gaines, S. D. (2002). Geographical abundance distributions of coastal invertebrates: Using one-dimensional ranges to test biogeographic hypotheses. *Journal of Biogeography*, *29*(8), 985–997. https://doi.org/10.1046/j.1365-2699.2002.00705.x

Sbrocco, E. J. (2014). Paleo-MARSPEC: Gridded ocean climate layers for the mid-Holocene and Last Glacial Maximum. *Ecology*, *95*(6), 1710–1710. https://doi.org/10.1890/14-0443.1

Silva, C. N. S., Young, E. F., Murphy, N. P., Bell, J. J., Green, B. S., Morley, S. A., Duhamel, G., Cockcroft, A. C., & Strugnell, J. M. (2021). Climatic change drives dynamic source–sink relationships in marine species with high dispersal potential. *Ecology and Evolution*, *11*(6), 2535-2550. https://doi.org/10.1002/ece3.7204

Singarayer, J. S., Valdes, P. J., & Roberts, W. H. G. (2017). Ocean dominated expansion and contraction of the late Quaternary tropical rainbelt. *Scientific Reports*, *7*(1), 9382. https://doi.org/10.1038/s41598-017-09816-8

Smith, A. B., & Santos, M. J. (2020). Testing the ability of species distribution models to infer variable importance. *Ecography*, *43*(12), 1801–1813. https://doi.org/10.1111/ecog.05317

Stone, A. E. C. (2014). Last Glacial Maximum conditions in southern Africa: Are we any closer to understanding the climate of this time period? *Progress in Physical Geography: Earth and Environment*, *38*(5), 519–542. https://doi.org/10.1177/0309133314528943

Sunday, J. M., Bates, A. E., & Dulvy, N. K. (2012). Thermal tolerance and the global redistribution of animals. *Nature Climate Change*, *2*(9), 686–690. https://doi.org/10.1038/nclimate1539

Teske, P. R., von der Heyden, S., McQuaid, C. D., & Barker, N. P. (2011). A review of marine phylogeography in southern Africa. *South African Journal of Science*, *107*(5–6), 43–53. https://doi.org/10.4102/sajs.v107i5/6.514

Teske, P. R., Sandoval-Castillo, J., Golla, T. R., Emami-Khoyi, A., Tine, M., Heyden, S. von der, & Beheregaray, L. B. (2019). Thermal selection as a driver of marine ecological speciation. *Proceedings of the Royal Society B*. https://royalsocietypublishing.org/doi/abs/10.1098/rspb.2018.2023

Theodoridis, S., Fordham, D. A., Brown, S. C., Li, S., Rahbek, C., & Nogues-Bravo, D. (2020). Evolutionary history and past climate change shape the distribution of genetic diversity in terrestrial mammals. *Nature Communications*, *11*(1), 2557. https://doi.org/10.1038/s41467-020-16449-5

Thuiller, W., Lafourcade, B., Engler, R., & Araújo, M. B. (2009). BIOMOD – a platform for ensemble forecasting of species distributions. *Ecography*, *32*(3), 369–373. https://doi.org/10.1111/j.1600-0587.2008.05742.x

Tolley, K. A., da Silva, J. M., & Jansen van Vuuren, B. (2019). South African National Biodiversity Assessment 2018 Technical Report Volume 7: Genetic Diversity. *South African National Biodiversity Institute, Pretoria*. http://hdl.handle.net/20.500.12143/6376

Toms, J. A., Compton, J. S., Smale, M., & Von Der Heyden, S. (2014). Variation in palaeo-shorelines explains contemporary population genetic patterns of rocky shore species. *Biology Letters*, *10*(6), 20140330. https://doi.org/10.1098/rsbl.2014.0330

Toonen, R. J., Puritz, J. B., Forsman, Z. H., Whitney, J. L., Fernandez-Silva, I., Andrews, K. R., & Bird, C. E. (2013). ezRAD: A simplified method for genomic genotyping in non-model organisms. *PeerJ*, *1*, e203. https://doi.org/10.7717/peerj.203

Vega, R., Fløjgaard, C., Lira-Noriega, A., Nakazawa, Y., Svenning, J.-C., & Searle, J. B. (2010). Northern glacial refugia for the pygmy shrew *Sorex minutus* in Europe revealed by phylogeographic analyses and species distribution modelling. *Ecography*, *33*(2), 260–271. https://doi.org/10.1111/j.1600-0587.2010.06287.x

Velichko, A. A., Kononov, Yu. M., & Faustova, M. A. (1997). The last glaciation of earth: Size and volume of ice-sheets. *Quaternary International*, *41–42*, 43–51. https://doi.org/10.1016/S1040-6182(96)00035-3

Wagner, V., Durka, W., & Hensen, I. (2011). Increased genetic differentiation but no reduced genetic diversity in peripheral vs. Central populations of a steppe grass. *American Journal of Botany*, *98*(7), 1173–1179. https://doi.org/10.3732/ajb.1000385

Waltari, E., & Hickerson, M. J. (2013). Late Pleistocene species distribution modelling of North Atlantic intertidal invertebrates. *Journal of Biogeography*, *40*(2), 249–260. https://doi.org/10.1111/j.1365-2699.2012.02782.x

Wang, I. J. (2011). Choosing appropriate genetic markers and analytical methods for testing landscape genetic hypotheses. *Molecular Ecology*, *20*(12), 2480–2482. https://doi.org/10.1111/j.1365-294X.2011.05123.x

Wernberg, T., Coleman, M. A., Bennett, S., Thomsen, M. S., Tuya, F., & Kelaher, B. P. (2018). Genetic diversity and kelp forest vulnerability to climatic stress. *Scientific Reports*, *8*(1), 1–8. https://doi.org/10.1038/s41598-018-20009-9

Winch, J. J. W. W., & Hodgson, A. N. (2007). The effect of temperature and salinity on oxygen consumption in the brachyuran crab *Cyclograpsus punctatus* (Crustacea: Decapoda: Grapsidae). *African Zoology*, *42*(1), 118–123. https://doi.org/10.1080/15627020.2007.11407384

Wright, D., Bishop, J. M., Matthee, C. A., & von der Heyden, S. (2015). Genetic isolation by distance reveals restricted dispersal across a range of life histories: Implications for biodiversity conservation planning across highly variable marine environments. *Diversity and Distributions*, *21*(6), 698–710. https://doi.org/10.1111/ddi.12302

Yannic, G., Pellissier, L., Ortego, J., Lecomte, N., Couturier, S., Cuyler, C., Dussault, C., Hundertmark, K. J., Irvine, R. J., Jenkins, D. A., Kolpashikov, L., Mager, K., Musiani, M., Parker, K. L., Røed, K. H., Sipko, T., Þórisson, S. G., Weckworth, B. V., Guisan, A., … Côté, S. D. (2014). Genetic diversity in caribou linked to past and future climate change. *Nature Climate Change*, *4*(2), 132–137. https://doi.org/10.1038/nclimate2074

Zardi, G., McQuaid, C., Teske, P., & Barker, N. (2007). Unexpected genetic structure of mussel populations in South Africa: Indigenous *Perna perna* and invasive *Mytilus galloprovincialis. Marine Ecology Progress Series*, *337*, 135–144. https://doi.org/10.3354/meps337135

Zhang, Z., Kass, J. M., Mammola, S., Koizumi, I., Li, X., Tanaka, K., Ikeda, K., Suzuki, T., Yokota, M., & Usio, N. (2021). Lineage-level distribution models lead to more realistic climate change predictions for a threatened crayfish. *Diversity and Distributions*, *27* (4), 684-695. https://doi.org/10.1111/ddi.13225

Biosketches

Erica S. Nielsen is broadly interested in marine biogeography, eco-evolutionary dynamics, and applied conservation genetics/genomics.

Romina Henriques works on comparative population genomics of marine species, being interested in understanding how oceanographic features shape evolutionary trajectories across multiple taxa.

Author contributions: ESN, SVDH, MB, and RH conceived the study. ESN collected and analysed the data, as well as led the writing of the manuscript, to which all authors contributed.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Additional information on the outputs of species distribution models (SDMs) and characteristics of the predictor variables used in the SDMs is given. Molecular diversity values per species per site, redundancy analysis outputs, and the equation used to calculate expected heterozygosity are also shown.

Appendix S1 Additional information on the outputs of species distribution models (SDMs) and characteristics of the predictor variables used in the SDMs is given. Molecular diversity values per species per site, redundancy analysis outputs, and the equation used to calculate expected heterozygosity are also shown.

Equation S1 – Calculation of expected heterozygosity (following Fischer et al. 2017):

$$H_E = (\sum(2\text{x}MAF)\ \text{x}(1-MAF))/nMAF$$

Where MAF is the minor allele frequency per locus, and nMAF is the number of loci per population.

Table S1 – The variance inflation factors (*VIFs*) for mean sea-surface temperature (SST), mean sea-surface salinity (SSS), as well as mean air temperature of warmest and coolest months (Tmax and Tmin*,* respectively), calculated for the present day, Mid-Holocene (MH) and Last Glacial Maximum (LGM). The values for MH and LGM are averaged between the MIROC and CCSM General Circulation Models.

Environmental variable	**Present**	**MH**	**LGM**
SST	7.3	7.6	8.4
SSS	1.1	1.4	1.7
Tmax	4.7	5.3	3.4
Tmin	3.8	4.2	3.8

Table S2 – The average receiver operating characteristic (ROC) and true skill statistic (TSS) values for each individual model included in the ensemble, as well as for the ensemble model per species. Individual models included the following: (1) Flexible Discriminant Analysis (FDA), (2) Generalized Additive Model (GAM), (3) Generalized Boosting Model (GBM), (4) Generalized Linear Model (GLM), (5) Multivariate Adaptive Regression Splines (MARS) and (6) Random Forest (RF).

	C. punctatus		***P. angulosus***		***S. granularis***	
Model	**ROC**	**TSS**	**ROC**	**TSS**	**ROC**	**TSS**
FDA	0.82	0.60	0.83	0.57	0.87	0.63
GLM	0.80	0.59	NA	NA	0.86	0.62
GAM	0.83	0.62	0.84	0.59	0.88	0.66
GBM	0.85	0.65	0.87	0.62	0.88	0.65
RF	0.85	0.64	0.88	0.65	0.90	0.67
MARS	0.83	0.62	0.82	0.57	0.87	0.65
Ensemble	0.92	0.69	0.96	0.75	0.96	0.74

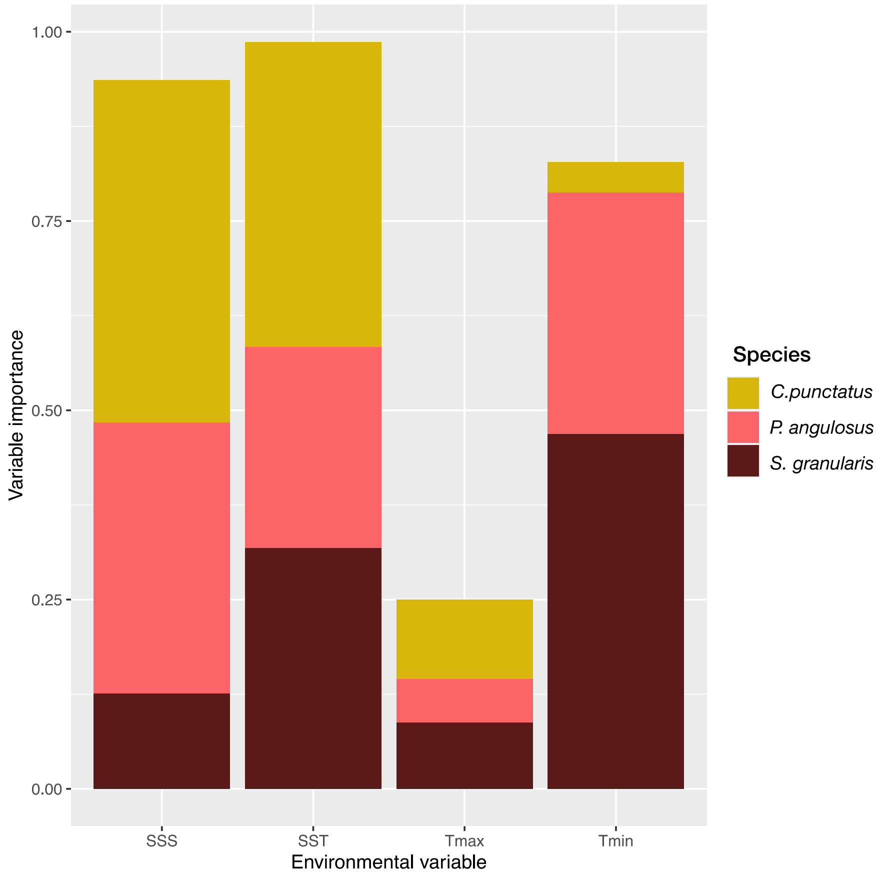

Figure S1 – The relative importance of the four environmental predictor variables (mean sea surface salinity- SSS; mean sea surface temperature - SST; maximum air temperature - Tmax; minimum air temperature- Tmin), averaged across all species distribution models, shown for *C. punctatus, P. angulosus* and *S. granularis*.

Table S3 – The latitude and longitude, and haplotype diversity value (*h*) for each sample site per species.

Site	**Latitude**	**Longitude**	***C. punctatus***	***P. angulosus***
PN	16.86	-29.25	-	0.94
HB	17.26	-30.32	-	0.84
BB	17.88	-31.30	-	-
LB	18.31	-32.09	-	0.98
JB	17.88	-32.97	0.54	0.97
SP	18.38	-33.92	NA	-
MP	18.40	-33.90	-	0.92
KM	18.32	-34.14	-	0.95
WP	18.45	-34.13	0.68	0.90

RO	18.813	-34.30	0.74	0.90
BT	18.93	-34.36	0.79	0.94
GB	19.34	-34.58	-	0.96
CA	20.01	-34.83	0.79	0.92
CI	20.87	-34.44	0.96	-
MB	22.09	-34.21	-	-
HB	22.40	-34.07	0.87	-
SB	25.53	-34.04	-	-
KOS	26.67	-33.70	-	-
KY	23.09	-34.06	0.86	0.89
JFB	24.92	-34.07	-	0.84
PE	25.67	-33.99	-	0.85
PA	26.86	-33.63	-	0.97
HH	28.23	-32.77	-	0.77
PSJ	29.54	-31.65	-	0.92
MU	30.24	-31.06	0.87	-
MG	30.38	-30.90	0.46	-

Table S4 – The latitude and longitude, and expected heterozygosity value (H_E) for each sample site per species.

Site	**Latitude**	**Longitude**	***C. punctatus***	***P. angulosus***	***S. granularis***
PN	16.86	-29.25	-	0.0682	0.0677
HB	17.26	-30.32	-	0.0633	0.0722
BB	17.88	-31.30	-	-	0.0714
DB	18.23	-31.80	-	0.0669	-
LB	18.31	-32.09	-	0.0666	0.0718
JB	17.88	-32.97	0.0719	0.0651	0.0701
YZ	18.16	-33.37	0.0745	-	-
SP	18.38	-33.92	0.0743	0.0649	0.0693

BT	18.93	-34.36	0.0779	0.0651	0.0699
GB	19.34	-34.58	0.0761	-	-
CA	20.01	-34.83	0.0740	0.0651	0.0693
JF	21.34	-34.45	0.0734	-	-
MB	22.09	-34.21	0.0746	0.0667	0.0633
KY	23.09	-34.06	0.0766	0.0600	0.0694
CF	24.83	-34.24	0.0759	0.0560	0.0694
PA	26.86	-33.63	-	0.0653	0.0615
HH	28.23	-32.77	0.0738	0.0633	0.0688
CB	29.15	-31.98	0.0755	0.0623	-
HL	29.31	-31.82	-	-	0.0690
MG	30.38	-30.90	0.0683	-	-

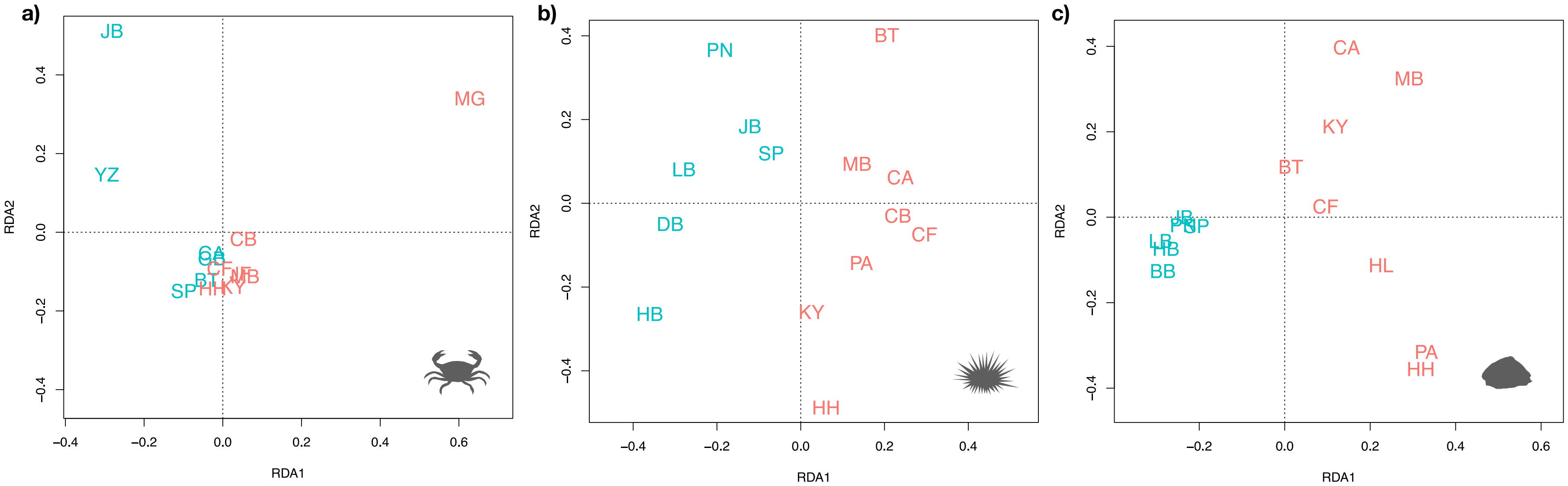

Figure S2 – Redundancy analyses based on the allele frequencies used to generate genomic diversity, using the same four environmental predictor variables used in the SDMs, with western lineage sites shown in blue and eastern sites shown in red for *C. punctatus* (a), *P. angulosus* (b), and *S. granularis* (c).

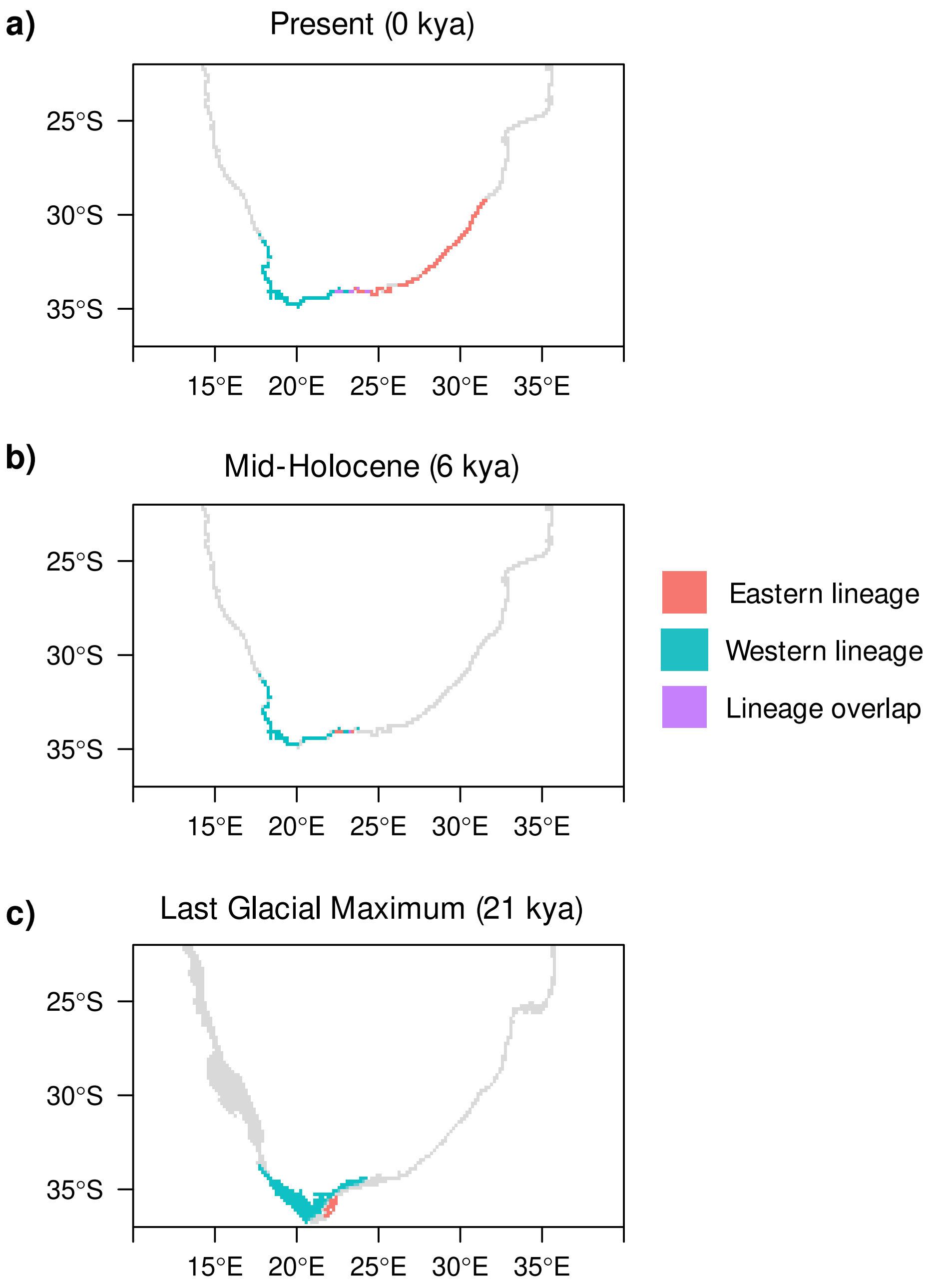

Figure S3 – Binary presence outputs from SDMs run on either western or eastern lineage sites, shown for *C. punctatus*, for the present day (a), Mid-Holocene (b), and Last Glacial Maximum (c), at zero, six, and 21 thousand years ago (kya), respectively.

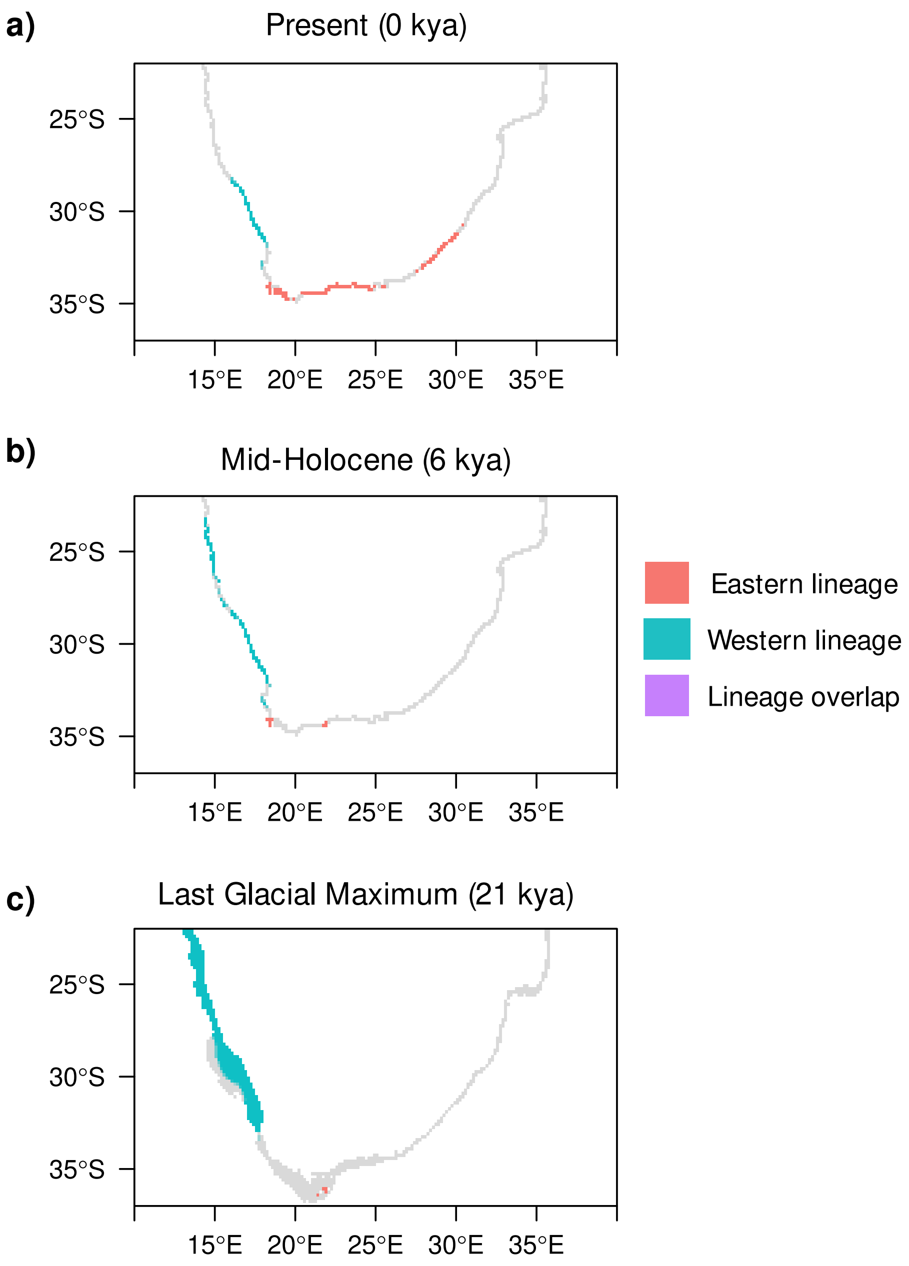

Figure S4 – Binary presence outputs from SDMs run on either western or eastern lineage sites, shown for *P. angulosus*, for the present day (a), Mid-Holocene (b), and Last Glacial Maximum (c), at zero, six, and 21 thousand years ago (kya), respectively.

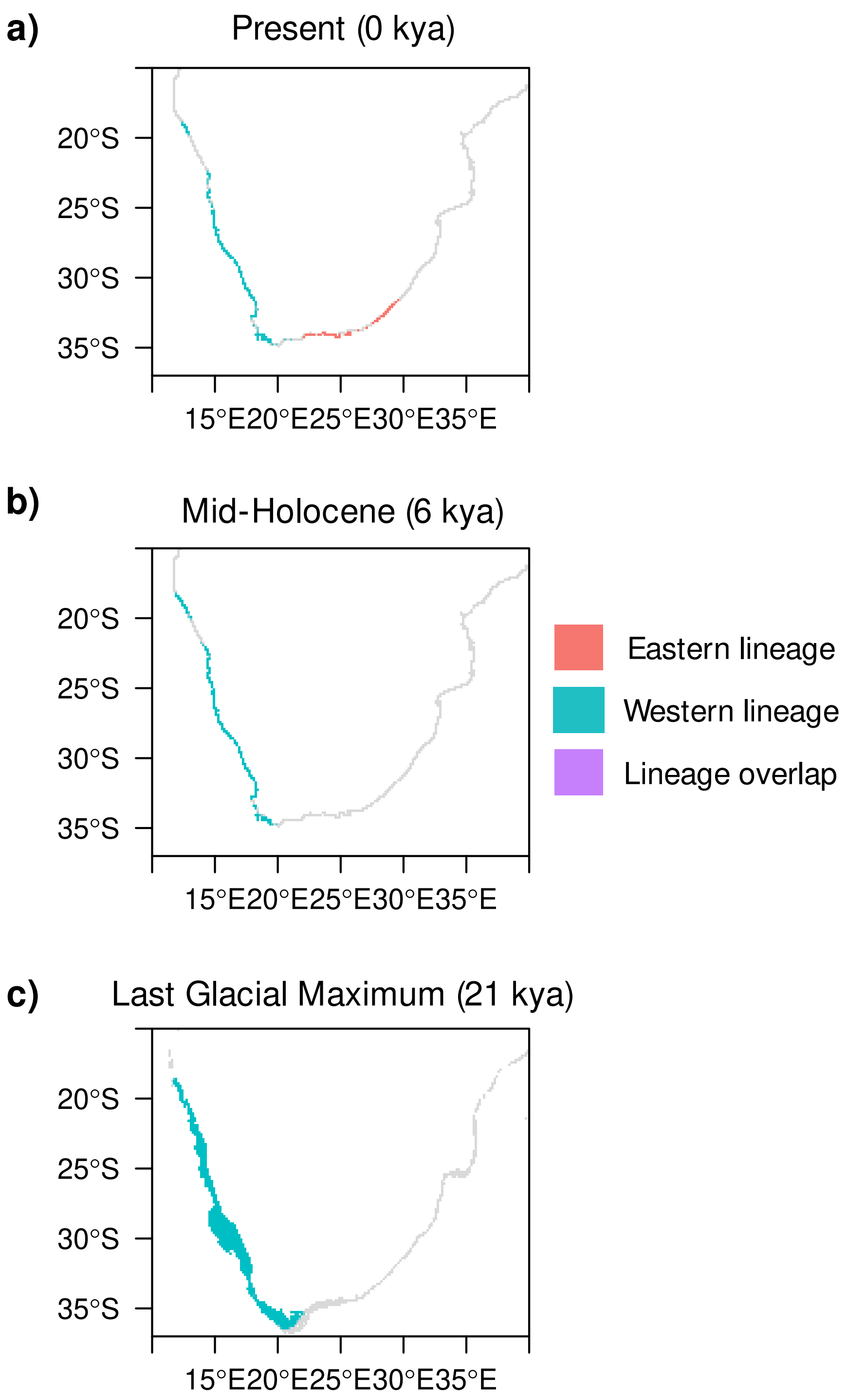

Figure S5 – Binary presence outputs from SDMs run on either western or eastern lineage sites, shown for *S. granularis*, for the present day (a), Mid-Holocene (b), and Last Glacial Maximum (c), at zero, six, and 21 thousand years ago (kya), respectively.

References:

Fischer, M. C., Rellstab, C., Leuzinger, M., Roumet, M., Gugerli, F., Shimizu, K. K., Holderegger, R., & Widmer, A. (2017). Estimating genomic diversity and population differentiation – an empirical comparison of microsatellite and SNP variation in Arabidopsis halleri. *BMC Genomics*, *18*(1), 69. https://doi.org/10.1186/s12864-016-3459-7