Thermal selection drives biodiversity origination across the Atlantic/Indian Ocean boundary

Peter R. Teske[1,*], Jonathan Sandoval-Castillo[2], Tirupathi Rao Golla[1], Arsalan Emami-Khoyi[1], Mbaye Tine[1], Sophie von der Heyden[3], Luciano B. Beheregaray[2]

[1]Centre for Ecological Genomics and Wildlife Conservation, Department of Zoology, University of Johannesburg, Auckland Park 2006, South Africa
[2]Molecular Ecology Lab, College of Science and Engineering, Flinders University, Adelaide 5001, Australia
[3]Evolutionary Genomics Lab, Department of Botany and Zoology, Stellenbosch University, Matieland 7602, South Africa
*Corresponding author

Abstract

Intraspecific genetic structure in widely distributed marine species often mirrors the boundaries between temperature-defined bioregions. This suggests that the same thermal gradients that maintain distinct species assemblages also drive the evolution of new biodiversity. Ecological speciation scenarios are often invoked to explain such patterns, but the fact that adaptation is usually only identified when phylogenetic splits are already evident makes it impossible to rule out the alternative scenario of allopatric speciation with subsequent adaptation. We integrated large-scale genomic and environmental datasets along one of the world's best defined marine thermal gradients (the South African coastline) to test the hypothesis that incipient speciation in the sea is due to divergence linked to the thermal environment. We identified temperature-associated gene regions in a coastal fish species that is spatially homogeneous throughout several temperature-defined biogeographical regions on the basis of selectively neutral markers. Based on these gene regions, the species is divided into geographically distinct regional populations. Importantly, the ranges of these populations are delimited by the same ecological boundaries that define distinct infraspecific genetic lineages in co-distributed marine the species, and biogeographical disjunctions in species assemblages. Our results indicate that ecologically-mediated selection represents an early stage of marine speciation in coastal regions that lack physical dispersal barriers.

Keywords: adaptation, dispersal barrier, ecological speciation, seascape genomics, ddRADseq, ecological genomics, incipient speciation, marine biogeography, selection

Introduction

Molecular phylogenies of marine species present along continuous coastlines have revealed that spatial disjunctions between distinct evolutionary lineages are often associated with the boundaries between different marine biogeographic regions [1,2], but such genetic patterns tend to be present in only a fraction of species [1,3–6] (Fig. 1). This discrepancy is often attributed to life history: actively dispersing species, and those with extended planktonic dispersal phases, cross the boundaries between bioregions more frequently than species with short propagule duration, making them less likely to diverge in spatial isolation [5,7]. However, support for this paradigm is not consistent, as numerous studies from North America [6,8], South Africa [1,9] and Australasia [4,10] have failed to identify a clear link between genetic structure and dispersal potential.

An alternative explanation for this paradox is offered by ecological divergence that preceded the allopatric distribution patterns evident on the basis of selectively neutral genetic markers. This is primarily supported by 'phylogenetic shifts approaches', in which phylogenetic splits coincide with ecological divergence [11]. The evidence for ecological speciation is particularly strong when phylogenetic splits are not associated with physical dispersal barriers that can completely isolate sister lineages [11], when contact zones are located in regions where environmental conditions are intermediate [1,12] and when each lineage displays reduced fitness in the habitat of its sister lineage [13–15]. Phylogenetic splits that are shared by multiple species across the same boundary may differ considerably in age [1,16], and by extension, this supports the hypothesis that species in which phylogenetic divergence is not yet evident have undergone ecological differentiation very recently.

The phylogenetic evidence for ecological drivers of speciation is nonetheless circumstantial, because divergence events mostly occurred during the Pleistocene or earlier [16–18], and it is difficult to extrapolate from contemporary conditions when species' historical distribution patterns are unknown and past oceanographic conditions not well understood. Because of such uncertainties, it is controversial to ascertain whether adaptation to divergent environments that reduced levels of gene flow because of the maladaptation of migrants was the primary driver of divergence, or whether it occurred after a phylogenetic split that may very well have evolved during an extended period of physical isolation.

More compelling evidence for ecological speciation in the sea would come from scenarios in which there is support for genetic differentiation that coincides with biogeography, but in which phylogenetic divergence indicative of speciation has not yet occurred [15]. The fact that phylogeographic breaks tend to be present in only a fraction of the species whose ranges span the boundaries between ecologically distinct marine regions [1,3,4] suggests that the condition of recent divergence may be met by those species that display no genetic divergence on the basis of the selectively neutral datasets typically employed in phylogeographic studies [19,20].

The South African coastline is characterised by ecologically distinct marine bioregions (Fig. 1) that are arranged along a thermal gradient [1]. This provides a unique opportunity for studying the importance of incipient environmentally-driven parapatric speciation in the sea, as biogeography (and, by extension, ecological speciation) is believed to be primarily a function of species' thermal tolerance ranges [21–23]. Numerous species complexes exist along this coastline that comprise cryptic species whose ranges are limited by the boundaries between bioregions [1], and which exhibit distinct temperature preferences [13,14]. This suggests that thermal adaptation contributes towards limiting gene flow between biogeographic regions by reducing migrant fitness and by subjecting migrants to competitive exclusion [1]. However, in some species, a single evolutionary lineage is found across multiple bioregions (Fig. 1). The latter are suitable candidates for determining whether diversifying selection driven by the environment, and corresponding reductions in gene flow, may have preceded phylogenetic splits.

We tested this hypothesis by generating genome-wide data from one of these phylogenetically homogeneous species [24], the Knysna sandgoby, *Psammogobius knysnaensis* (Fig. 1). We expect population divergence that mirrors coastal biogeography to be evident based only on temperature-associated genes. This would present compelling evidence that in coastal regions that lack physical dispersal barriers, thermal selection plays a defining role in the early stages of parapatric ecological speciation.

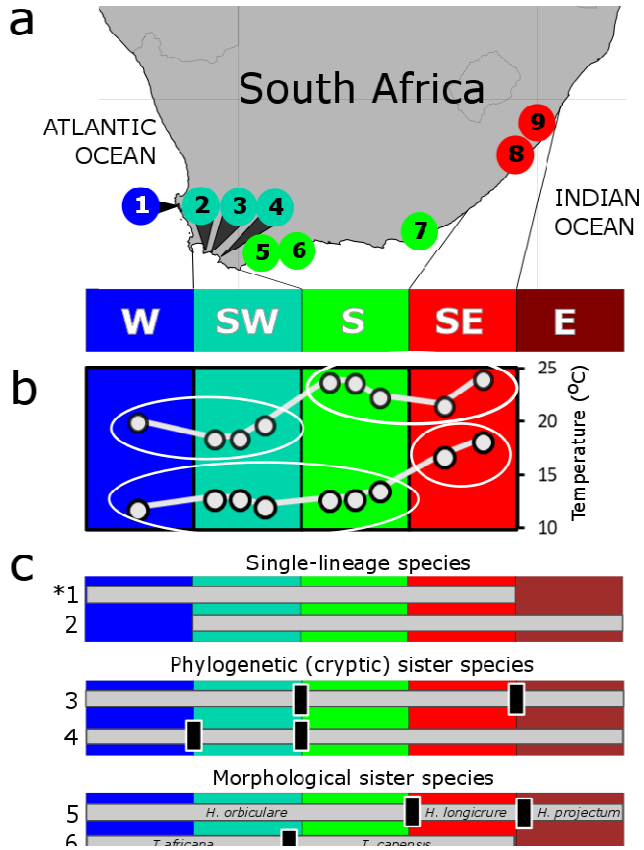

Figure 1. Sampling sites, marine bioregions and examples of genetic breaks in South African coastal animals; **a** A map indicating the location of sampling sites within southern Africa's temperature-defined marine bioregions; **b** maximum and minimum sea surface temperatures at each sampling site; these temperatures divide bioregions into two groups each (indicated by ellipses), and by themselves only partially explain the region's biogeography; **c** examples of distribution ranges (grey horizontal bars) and location of genetic breaks (black vertical bars) in coastal South African animals, arranged hierarchically. The top panel depicts species that occur as a single phylogenetic lineage in multiple bioregions: 1. *Psammogobius knysnaensis* (the study species, marked with an asterisk) [24] and 2. *Scutellastra longicosta* [16]. The middle panel depicts species with phylogenetically distinct sister lineages that are not distinguishable morphologically (cryptic species): 3. *Callichirus kraussi* [25] and 4. *Palaemon peringueyi* [26]. The bottom panel depicts morphologically distinguishable sister species: 5. *Hymenosoma* spp. [27] and 6. *Tricolia* spp. [28]. Abbreviations: W, cool-temperate west coast; SW, transition zone on the south-west coast; S, warm-temperate south coast; SE, transition zone on the south-east coast; E, subtropical east coast.

Methods

(a) Sampling procedure

Tissue samples from a total of 312 individuals of *Psammogobius knysnaensis* were collected from the mouth areas of nine estuaries throughout the species' range (Table 1) using a pushnet. Upon capture, a fin clip was obtained from one of the pectoral fins using sterilised fingernail scissors, and immediately preserved in 100% ethanol. The fish were subsequently released. Samples in the West Coast National Park (Langebaan Lagoon) were collected under research permit no. CRC/2015/033--2015/V1.

(b) Generation and processing of genomic data

Genomic DNA was extracted using the CTAB protocol [29], and double digest restriction site–associated DNA (ddRAD) libraries were constructed for a subset of 129 individuals and 12 replicates with particularly high quality DNA, following the protocol described in [30] and modified as described in Sandoval-Castillo et al. [31]. Libraries were pooled in groups of 48 or 93 samples per lane and sequenced on an Illumina HiSeq 2000 (100 bp paired-end reads) platform at the McGill University and Genome Québec Innovation Centre. Raw sequences were processed as described in the Supporting Information.

(c) Identification of loci under thermal selection and neutral loci

We assessed the contribution of coastal sea surface temperature (SST) to the overall pattern of genetic differentiation using the R package gINLAnd [32]. This software uses a spatial generalized linear mixed model to quantify the correlation between genotypes and environmental variables, while controlling for the effects of spatial population structure and population history. Briefly, gINLAnd estimates the covariance associated with the spatial distribution of the samples and a locus-specific effect of each environment variable; it then estimates the likelihood of two competing models: a model with the environmental effect and a reduced model without the environmental effect. Finally, gINLAnd assesses the strength of genetic dependence on the environmental variable by computing a Bayes factor between the two models. To avoid false positives, we used a conservative approach in which only those loci which showed a log Bayes factor (BF) ≥10 were considered to be under selection (a log BF>4.6 is considered decisive [33]). A plot depicting loci under selection is shown in Fig. S1. We calculated a multidimensional scaling projection of the coastal distance between sampling sites using the R package MASS 7.3 [34], which is more meaningful than using the original geographic coordinates because this would have required connecting sites via terrestrial habitat. As the application of satellite-based

SST data is often problematic when studying coastal biogeography, because it includes data from offshore regions [35], we used southern African temperature data based on *in situ* measurements, as described in the Supporting Information.

To compare the temperature-associated loci with a data set comprising only selectively neutral loci, the following approach was used. Thermal selection may be only one of a number of drivers of selection, so we used BayeScan v. 2.1 [36] to identify markers under selection on the basis of outlier scans rather than temperature data. This method was used because it has a low error rate compared to other tests for the detection of outlier loci [37]. Default settings were applied with prior odds set to 10, but a very high false discovery rate of 20% was applied to create a neutral data set with a low probability of containing any remaining loci under selection. We then excluded 304 outlier loci, together with 27 additional loci identified by gINLAnd that were not found by BayeScan, to create a data set of 8201 selectively neutral loci.

(d) Functional annotation

To identify the possible functions of genes under thermal selection, we blasted the flanking sequences of temperature-associated loci against the NCBI non-redundant nucleotide database. The resulting reads were then annotated against the UniProtKB/Swiss-Prot database [38]. We then performed a gene ontology term analysis in topGO 2.24.0 [39]. Genes whose function indicates an influence of thermal selection were identified by searching the relevant literature.

(e) Population genetic structure

Genetic structure was investigated separately for loci under selection and neutral loci. We employed both clustering and phylogenetic approaches. Discriminant Analysis of Principal Components (DAPC) was performed with the R package ADEGENET v. 2.1.0 [40]. DAPC defines a model with synthetic variables in which the genetic variation is maximized between clusters of individuals (*K*), and minimized within clusters. We used *k*-means clustering and the Bayesian Information Criterion (BIC) to identify the best-supported number of clusters. Patterns of genetic structure were also explored using *fastStructure* 1.0 [41], which uses variational Bayesian inference under a model assuming Hardy-Weinberg equilibrium and linkage equilibrium. We used a simple prior and set all other parameters to the default value, expect for the convergence criterion, which was lowered to 10^{-8}. The programme was run for each value of K = 1-9 independently, and each value was cross-validated 1000 times. The python script *chooseK* was used to identify an optimal range of *K* values, and the resulting barplots were visualised with the R package distruct2.2 [42].

Phylogenetic analyses were performed in BEAST v. 2.4.7 [43]. A maximum clade credibility (MCC) tree was reconstructed using a discrete phylogeographic analysis [44]. In this case, the data set comprised individual alleles of each individual that were reconstructed in PHASE v. 2.1.1 [45] using default settings. When more than one pair of haplotypes was possible for an individual, the one with the highest probability was used. In addition to reconstructing a phylogenetic tree, this method can infer the most likely bioregion in which each ancestral node in the MCC tree was present. One hundred million generations were specified, and trees saved every 100 000 generations, and the first 20% of trees were discarded as burn-in. Model and prior settings followed those recommended in the tutorial available at http://hpc.ilri.cgiar.org. For comparison, a corresponding MCC tree was created for previously published mtDNA COI data [24] using the same settings.

Results

The ddRADseq [30] procedure was used to generate a genome-wide dataset of single nucleotide polymorphisms (SNPs) from individuals collected at nine sites that are located within four temperature-defined marine bioregions (Fig. 1). A total of 405,648,596 raw reads were generated on two Illumina lanes. After demultiplexing and quality filtering, an average of 1,560,510 reads were obtained per individual, totalling 224,713,440 reads. The filtered catalogue resulted in 8,532 ddRADseq loci containing 15,633 SNPs. A final dataset was obtained by extracting only the SNPs with the best quality score from each polymorphic ddRADseq locus to remove SNPs that are likely in linkage disequilibrium. After removing individuals with more than 20% missing data, the final data set comprised 109 individuals genotyped for 8,532 SNPs. We then used a spatially explicit generalized linear mixed model to test for direct associations between SNP allele frequencies and temperature-related variables, while controlling for the effects of spatial structure and shared population history, using the program gINLAnd [32]. Unlike F_{ST}-based outlier scans, which identify loci on the basis of population information [36], the identification of loci in in genotype-environment association methods such as gINLAnd is thus not influenced by any regional population structure [46]. We explored various combinations of maximum or minimum temperature as the environmental variable with covariance factors that included geographic distance, biogeographic boundaries and a combination of the two. A simplistic resistance matrix approach [47] was used, where geographic distance and biogeographic boundaries between pairs of sites were given a resistance value (one unit of resistance per km between sites and ten units of resistance if sites were located in different marine bioregions). We then calculated multidimensional scaling projections based on the resistance pairwise matrices, and these were used as covariance factor to control for spatial structure. We further explored the effect of using only SNP data from the coolest and the warmest marine bioregions, using geographic distance as the controlling factor (see Table S1 for detail on number of SNPs identified). Most subsequent analyses were performed with the data set recovered using minimum temperature (with geographic distance as the covariance factor), which is thought to represent a key selective agent of adaptive divergence in the study region because it can limit species' distributions [48,49].

SNPs from ddRADseq originate from all genomic regions, and some may be located on protein-coding genes that are strongly affected by temperature. While such associations may not necessarily imply a causal relationship, identifying their function may contribute towards an improved understanding of possible drivers of genetic divergence between temperature-defined bioregions. Although no fully annotated transcriptome for the family Gobiidae is presently available, nine of the loci (identified using either maximum or minimum temperature, with geographic distance as the covariance factor) could be annotated as genes involved in mitigating thermal stress (Table S2). Three of these (14-3-3 gene, tyrosine protein kinase and tubulin beta chain) are of particularly interest because they were involved in heat stress responses in a species of goby [50], or cold stress adaptation/acclimation in other teleosts [51,52]. In all but two cases, loci that were identified using minimum temperature data were also identified using maximum temperatures (Table S2). This suggests that even though most experimental studies investigated responses to heat stress, genetically fixed differences of these genes between temperature-defined marine bioregions may reflect general adaptations to different thermal environments, and thus play a role in determining thermal tolerance ranges.

To ascertain that the study species does not yet exhibit genetic divergence based on putatively neutral data, which could indicate that geographical isolation preceded thermal adaptation, we created a reduced dataset comprising selectively neutral data. In addition to excluding loci identified as being under thermal selection, we excluded outlier loci from genome scans [36] to better estimate demographic parameters and population differentiation [31].

Two complementary methods of investigating a link between population structure and biogeography were performed on both the temperature-associated loci and the neutral loci, Discriminant Analysis of Principal Components (DAPC) [53] and fastStructure [41]. Statistical support was high for 3-4 clusters (K) when analysing the temperature-associated loci, while a single cluster (K = 1) was found for the neutral loci with both methods (Fig. S2, Fig. S3). Even though temperature alone only accounts for two marine bioregions (Fig. 1), and despite the fact that geographic distances and/or the boundaries between bioregions were controlled for when identifying temperature-associated loci, an affiliation of genetic clusters with up to four bioregions was found when using the temperature-associated loci. The distinctness of subtropical (SE) individuals from the temperate (W, SW and S) sites was evident in all analyses, and most analyses using minimum temperature as the environmental variable also identified the W coast as a distinct cluster. There was even evidence for distinct SW and S coast clusters, although these were comparatively poorly differentiated. This result was robust, and also recovered using fastStructure (Fig. 2).

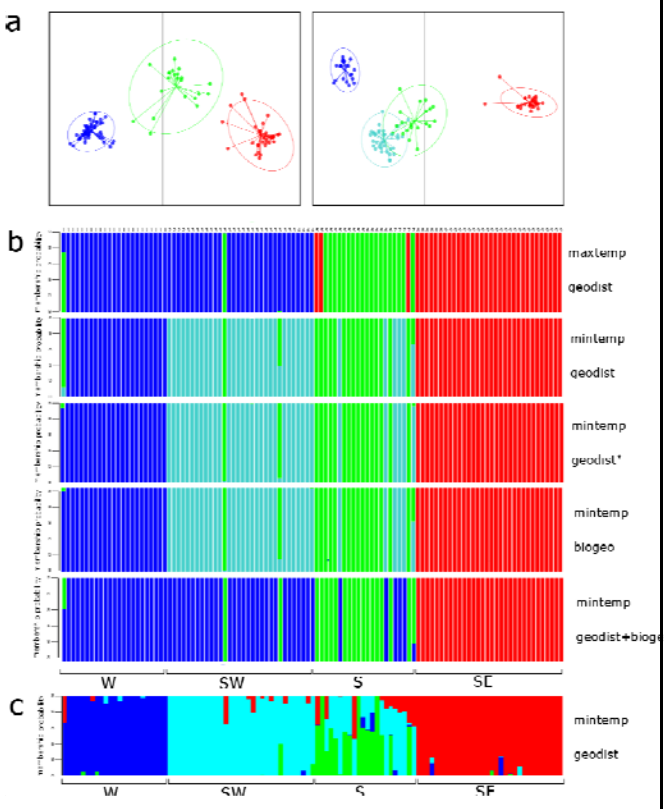

Figure 2. Population genetic structure inferred for temperature-associated loci. **a** DAPC scatterplot with inertia ellipses representing 95% confidence intervals, with colours reflecting the dominant bioregion represented in a particular cluster (left: loci correlated with maximum temperature; right: loci correlated with minimum temperature, in both cases controlling for geographic distance); **b** DAPC compoplots indicating membership probabilities for each individual (vertical bars) within one of four genetic clusters; correlation with temperature and the controlling factor are indicated on the right (maxtemp = maximum temperature, mintemp = minimum temperature, geodist = geographic distance, biogeo = biogeography; *indicates that only sites 1, 8 and 9 were used to find loci correlated with minimum temperature); **c** corresponding concensus fastStructure barplot for four genetic clusters (*K*); for comparison, barplots for *K*=2-5 are shown in Fig. S4. Site numbers and abbreviations correspond to those in Fig. 1 and Table S3.

Congruent with the clustering methods, a maximum-clade credibility tree (Fig. 3a) of temperature-associated loci (minimum temperature with geographic distance as the covariance factor) recovered both the western and south-eastern group as mostly distinct but poorly differentiated clusters nested within a tree whose oldest nodes were inferred to have existed on the south coast. Some branches are nested within clades that mostly have location states from other regions, which may reflect migration between adjacent marine bioregions. For comparison, a maximum-clade credibility tree reconstructed from mtDNA COI sequences [24] shows no clear regional structure (Fig. 3b).

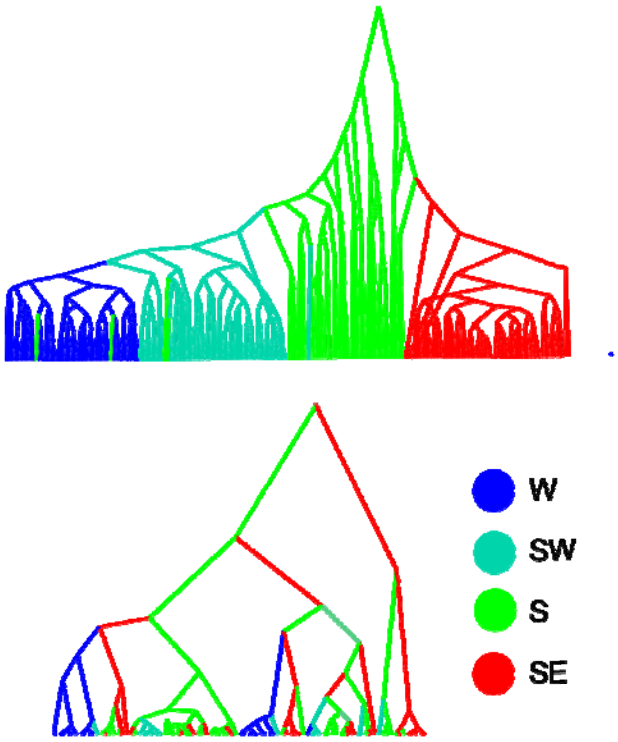

Figure 3. Reconstruction of phylogenetic relationships between individuals of *Psammogobius knysnaensis* from four South African marine bioregions using maximum-clade credibility trees from **a** phased SNP data of temperature-associated loci (correlated with minimum temperature) and **b** mtDNA COI data, with location state reconstructions of ancestral nodes. In **a**, clear regional structure is evident, but there are possible migrants (some branches are nested within clades that mostly have location states from other regions). In contrast, there are no clear regional clades in **b**. Site numbers and abbreviations correspond to those in Fig. 1 and Table 1, and trees are not drawn to scale.

Discussion

Speciation is a continuous process comprising a number of evolutionary stages that range from adaptive differentiation to complete reproductive isolation between populations [54]. Identifying the primary drivers of speciation is challenging because a considerable amount of time has often already passed

by the time incipient speciation becomes evident. This makes it difficult to distinguish ecologically-driven divergence with ongoing gene flow from allopatric divergence and secondary contact [55,56].

Biogeography is often considered to be a function of species' thermal tolerance ranges [21–23]. The fact that South Africa's coastal biogeography is mirrored by intraspecific spatial genetic structure suggests that species present in more than one province should comprise multiple evolutionary lineages that represent cryptic species [1]. The goby *Psammogobius knysnaensis* is one of a number of coastal southern African species that occur in multiple marine bioregions, but which displays no regional divergence on the basis of selectively neutral markers [24] of the type that are primarily used in phylogeographical research [19,20]. Here, we reject the previous finding of genetic homogeneity and show that this species is in fact represented by multiple regional groups delimited by temperature-defined bioregions. The fact that this is only evident for temperature-associated loci, and not for putatively neutral loci, confirms that divergence must have taken place in the absence of an interruption of gene flow due to physical dispersal barriers. Under these conditions, a pattern of isolation-by-adaptation [57] can be expected to eventually evolve, as migrants dispersing into adjacent bioregions will have fewer surviving offspring and reduced survival rates compared to residents. A west-to-east thermal differentiation was evident particularly for the minimum temperatures, where the two easternmost sites had much warmer water than the other sites (Fig. 1). However, there was no indication that marine bioregions could be identified on the basis of high or low temperatures alone, and the identification of loci under thermal selection and subsequent detection of up to four genetic clusters cannot be explained as being an artefact of the temperature variables used in the gINLAnd analyses.

Adaptations to the thermal environment are complex and ubiquitous in nature. Temperature affects many different biological pathways, with strong effects on the integrity of proteins and cellular structures and on the rates of physiological processes, particularly in ectotherms [58,59]. The thermal environment can promote partial reproductive isolation between populations, which might drive them along the speciation continuum [60]. This is particularly true for organisms that (i) have distinct populations with parapatric distributions along the thermal gradient, (ii) do not maintain a stable internal temperature (poikilotherms), and (iii) are found across stable thermal gradients (e.g. aquatic environments), which are regions where exogenous divergent selection is not expected to weaken due to marked temperature fluctuations [60]. Our study system meets all these conditions and represents an example of parapatric ecological divergence with genomic hallmarks of incipient evolutionary divergence driven by the thermal environment.

Unlike previous spatial demographic inferences from coastal southern Africa, which typically reflect the influence of past climatic changes [16,17,61,62], the spatial genetic patterns identified here can be explained by present-day environmental conditions. On the east coast, northward dispersal in the nearshore area is facilitated by wind-driven circulation [63], but this is unlikely to occur beyond site 9 (the northern distribution limit of *P. knysnaensis*) [62,64] because under contemporary conditions, the southward-flowing Agulhas Current flows very close to the coast and causes the parallel southward flow of nearshore circulation [65]. In the western portion of the species' range, gene flow between south and west coast is primarily facilitated by the westward drift of surface water [66]. The limited evidence for gene flow in both cases would be difficult to explain if one exclusively invoked physical isolation, given the high dispersal potential of the species' larvae coupled with the region's strong ocean circulation. It suggests that migrants from a particular bioregion are maladapted to the environmental conditions in adjacent bioregions. For example, the distinctness of the west coast population from those on the south-west and south coast may reflect the influence of cold-water upwelling in the west [67].

We hypothesise that thermal selection, perhaps in combination with factors such as oceanography and primary productivity that covary with temperature to influence local adaptation [60], acts primarily on the sensitive larvae. Under this scenario, ecologically diverging populations are limited in their ability to exchange genes and, as such, reproductive isolation is expected to ensue [60]. There are no known conspicuous phenotypes that differ between the presumably locally-adapted *P. knysnaensis* populations, but this is unsurprising because thermal adaptation often initially creates cryptic changes at the level of cell membranes or thermal stability of enzymes [68]. Studies that combine information from population genomics and controlled laboratory experiments using temperature-defined populations along an evolutionary continuum of speciation are expected to improve the identification of phenotypes enriched for selection signals of thermal adaptation.

Conclusion

Allopatric speciation in the marine environment is often invoked along continuous but ecologically subdivided coastlines, despite evidence that the physical dispersal barriers to whom this is attributed are insufficient to completely isolate regional

populations [69–71]. Our study contributes to the growing evidence that in adjacent, temperature-defined marine provinces, divergence of loci linked to the thermal environment can precede significant spatial divergence of selectively neutral markers [72,73]. This strongly favours a scenario of parapatric ecological divergence over one in which allopatric divergence is followed by thermal adaptation. In the context of larger biogeographical patterns, where range boundaries in the sea often coincide with the boundaries between temperature-defined bioregions [74,75], this evidence suggests that temperature-driven diversifying selection may be an important early-stage factor in the evolution of marine biodiversity.

Authors' contributions. P.R.T. and L.B.B. designed the study, P.R.T., T.R.G. and S.v.d.H. collected the samples, J.S.-C. conducted the experiments, J.S.-C., P.R.T., M.T., A.E.-K. and T.R.G. analysed the data, P.R.T. and L.B.B. wrote the paper, with input from S.v.d.H. and J.S.-C.

Acknowledgements. We are grateful to Robert Schlegel and Albertus Smit for providing the SST data. This study was funded by the PADI Foundation (Grant No. 10981 to PRT), the National Research Foundation (CSUR Grant No. 87702 to PRT), the University of Johannesburg (URC/FRC grant to PRT) and the Australian Research Council (FT130101068 and DP110101275 to L.B.B.). The authors are grateful to the Centre for High Performance Computing (CHPC), particularly Dane Kennedy, for supercomputer resources and bioinformatics support. M.T., T.R.G. and A.E.-K. acknowledge the University of Johannesburg for Global Excellence and Stature (GES) fellowships.

References

1. Teske PR, Von der Heyden S, McQuaid CD, Barker NP. 2011 A review of marine phylogeography in southern Africa. *South Afr. J. Sci.* **107**, 1–11. (doi:10.4102/sajs.v107i5/6.514)
2. Ayre DJ, Minchinton TE, Perrin C. 2009 Does life history predict past and current connectivity for rocky intertidal invertebrates across a marine biogeographic barrier? *Mol. Ecol.* **18**, 1887–1903. (doi:10.1111/j.1365-294X.2009.04127.x)
3. von der Heyden S. 2009 Why do we need to integrate population genetics into South African marine protected area planning? *Afr. J. Mar. Sci.* **31**, 263–269. (doi:10.2989/AJMS.2009.31.2.14.886)
4. Dawson MN. 2012 Parallel phylogeographic structure in ecologically similar sympatric sister taxa. *Mol. Ecol.* **21**, 987–1004. (doi:10.1111/j.1365-294X.2011.05417.x)
5. Pelc RA, Warner RR, Gaines SD. 2009 Geographical patterns of genetic structure in marine species with contrasting life histories. *J. Biogeogr.* **36**, 1881–1890. (doi:10.1111/j.1365-2699.2009.02138.x)
6. Kelly RP, Palumbi SR. 2010 Genetic structure among 50 species of the northeastern pacific rocky intertidal community. *PLoS ONE* **5**, e8594. (doi:10.1371/journal.pone.0008594)
7. Haye PA, Segovia NI, Muñoz-Herrera NC, Gálvez FE, Martínez A, Meynard A, Pardo-Gandarillas MC, Poulin E, Faugeron S. 2014 Phylogeographic structure in benthic marine invertebrates of the southeast pacific coast of Chile with differing dispersal potential. *PLoS ONE* **9**, 1–15. (doi:10.1371/journal.pone.0088613)
8. Weersing K, Toonen RJ. 2009 Population genetics, larval dispersal, and connectivity in marine systems. *Mar. Ecol. Prog. Ser.* **393**, 1–12. (doi:10.3354/meps08287)
9. Wright D, Bishop JM, Matthee CA, 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. *Divers. Distrib.* **21**, 698–710. (doi:10.1111/ddi.12302)
10. Ayre DJ, Minchinton TE, Perrin C. 2009 Does life history predict past and current connectivity for rocky intertidal invertebrates across a marine biogeographic barrier? *Mol. Ecol.* **18**, 1887–1903. (doi:10.1111/j.1365-294X.2009.04127.x)
11. Nosil P. 2012 *Ecological Speciation*. Oxford University Press.
12. Rocha LA, Robertson DR, Roman J, Bowen BW. 2005 Ecological speciation in tropical reef fishes. *Proc. Biol. Sci.* **272**, 573–9.
13. Teske PR, Papadopoulos I, Newman BK, Dworschak PC, McQuaid CD, Barker NP. 2008 Oceanic dispersal barriers, adaptation and larval retention: an interdisciplinary assessment of potential factors maintaining a phylogeographic break between sister lineages of an African prawn. *BMC Evol. Biol.* **8**, 341. (doi:10.1186/1471-2148-8-341)
14. Papadopoulos I, Teske PR. 2014 Larval development reflects biogeography in two formerly synonymised southern African coastal crabs. *Afr. J. Aquat. Sci.* **39**. (doi:10.2989/16085914.2014.938600)
15. Beheregaray LB, Sunnucks P. 2001 Fine-scale genetic structure, estuarine colonization and incipient speciation in the marine silverside fish Odontesthes argentinensis. *Mol. Ecol.* **10**, 2849–2866. (doi:10.1046/j.1365-294X.2001.t01-1-01406.x)
16. Mmonwa KL, Teske PR, McQuaid CD, Barker NP. 2015 Historical demography of southern African patellid limpets: congruence of population expansions, but not phylogeography. *Afr. J. Mar. Sci.* **37**. (doi:10.2989/1814232X.2015.1009165)
17. Toms JA, Compton JS, Smale M, von der Heyden S. 2014 Variation in palaeo-shorelines explains contemporary population genetic patterns of rocky shore species. *Biol. Lett.* **10**, 20140330-. (doi:10.1098/rsbl.2014.0330)
18. Waltari E, Hickerson MJ. 2013 Late Pleistocene species distribution modelling of North Atlantic intertidal invertebrates. *J. Biogeogr.* **40**, 249–260. (doi:10.1111/j.1365-2699.2012.02782.x)
19. Beheregaray LB. 2008 Twenty years of phylogeography: The state of the field and the challenges for the Southern Hemisphere. *Mol. Ecol.* **17**, 3754–3774. (doi:10.1111/j.1365-294X.2008.03857.x)
20. Garrick RC *et al.* 2015 The evolution of phylogeographic data sets. *Mol. Ecol.* **24**, 1164–1171. (doi:10.1111/mec.13108)
21. Pörtner HO, Peck L, Somero G. 2007 Thermal limits and adaptation in marine Antarctic ectotherms: an integrative view. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **362**, 2233–58. (doi:10.1098/rstb.2006.1947)
22. Belanger CL, Jablonski D, Roy K, Berke SK, Krug AZ, Valentine JW. 2012 Global environmental predictors of benthic marine biogeographic structure. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 14046–14051. (doi:10.1073/pnas.1212381109)
23. Bowen BW, Gaither MR, DiBattista JD, Iacchei M, Andrews KR, Grant WS, Toonen RJ, Briggs JC. 2016 Comparative phylogeography of the ocean planet. *Proc. Natl. Acad. Sci.* **113**, 7962–7969. (doi:10.1073/pnas.1602404113)
24. Drost E, Golla TR, Heyden S von der, Teske PR. 2016 No divergent evolution, despite restricted connectivity, between Atlantic and Indian Ocean goby populations. *Mar. Biodivers.* **46**, 465–471. (doi:10.1007/s12526-015-0389-6)
25. Teske PR, Winker H, McQuaid CD, Barker NP. 2009 A tropical/subtropical biogeographic disjunction in southeastern Africa separates two evolutionarily significant units of an estuarine prawn. *Mar. Biol.* **156**, 1265–1275. (doi:10.1007/s00227-009-1168-3)
26. Teske PR, Froneman PW, Barker NP, McQuaid CD. 2007 Phylogeographic structure of the caridean shrimp Palaemon peringueyi in South Africa: Further evidence for intraspecific genetic units associated with marine biogeographic provinces. *Afr. J. Mar. Sci.* **29**. (doi:10.2989/AJMS.2007.29.2.9.192)
27. Teske PR *et al.* 2009 Tri-locus sequence data reject a 'Gondwanan origin hypothesis' for the African/South Pacific crab genus Hymenosoma. *Mol. Phylogenet. Evol.* **53**, 23–33. (doi:10.1016/j.ympev.2009.05.031)
28. Nangammbi TC, Herbert DG, Teske PR. 2016 Molecular insights into species recognition within southern Africa's endemic Tricolia radiation (Vetigastropoda: Phasianellidae). *J. Molluscan Stud.* **82**, 97–103. (doi:10.1093/mollus/eyv037)

29. Doyle J. 1991 CTAB Total DNA Isolation. In *Molecular Techniques in Taxonomy*, pp. 283–293. (doi:10.1007/978-3-642-83962-7_18)
30. Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE. 2012 Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE* **7**. (doi:10.1371/journal.pone.0037135)
31. Sandoval-Castillo J, Robinson NA, Hart AM, Strain LWS, Beheregaray LB. 2018 Seascape genomics reveals adaptive divergence in a connected and commercially important mollusc, the greenlip abalone (Haliotis laevigata), along a longitudinal environmental gradient. *Mol. Ecol.* **27**, 1603–1620. (doi:10.1111/mec.14526)
32. Guillot G, Vitalis R, le Rouzic A, Gautier M. 2014 Detecting correlation between allele frequencies and environmental variables as a signature of selection. A fast computational approach for genome-wide studies. *Spat. Stat.* **8**, 145–155. (doi:10.1016/j.spasta.2013.08.001)
33. Kass RE, Raferty AE. 1995 Bayes factors. *J. Am. Stat. Soc.* **90**, 773–795.
34. Venables WN, Ripley BD. 2002 *Modern applied statistics with S*. 4th editio. New York: Springer. (doi:10.1198/tech.2003.s33)
35. Smit AJ, Roberts M, Anderson RJ, Dufois F, Dudley SFJ, Bornman TG, Olbers J, Bolton JJ. 2013 A coastal seawater temperature dataset for biogeographical studies: large biases between in situ and remotely-sensed data sets around the coast of South Africa. *PLoS ONE* **8**, e81944. (doi:10.1371/journal.pone.0081944)
36. Foll M, Gaggiotti O. 2008 A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics* **180**, 977–993. (doi:10.1534/genetics.108.092221)
37. Narum SR, Hess JE. 2011 Comparison of Fst outlier tests for SNP loci under selection. *Mol. Ecol. Resour.* **11 Suppl 1**, 184–194. (doi:10.1111/j.1755-0998.2011.02987.x)
38. Boutet E, Lieberherr D, Tognolli M, Schneider M, Bairoch A. 2007 UniProtKB/Swiss-Prot. *Methods Mol. Biol. Clifton NJ* **406**, 89–112.
39. Alexa A, Rahnenfuhrer J. 2016 topGO: enrichment analysis for gene ontology. R package version 2.24.0.
40. Jombart T, Ahmed I. 2011 adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics* **27**, 3070–3071. (doi:10.1093/bioinformatics/btr521)
41. Raj A, Stephens M, Pritchard JK. 2014 FastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics* **197**, 573–589. (doi:10.1534/genetics.114.164350)
42. Rosenberg NA. 2004 distruct: a program for the graphical display of population structure. *Mol. Ecol. Notes* **4**, 137–138. (doi:10.1046/j.1471-8286.2003.00566.x)
43. Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D, Suchard MA, Rambaut A, Drummond AJ. 2014 BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLOS Comput. Biol.* **10**, e1003537. (doi:10.1371/journal.pcbi.1003537)
44. Lemey P, Rambaut A, Drummond AJ, Suchard MA. 2009 Bayesian phylogeography finds its roots. *PLoS Comput. Biol.* **5**. (doi:10.1371/journal.pcbi.1000520)
45. Stephens M, Smith NJ, Donnelly P. 2001 A New Statistical Method for Haplotype Reconstruction from Population Data. *Am. J. Hum. Genet.* **68**, 978–989.
46. Forester BR, Lasky JR, Wagner HH, Urban DL. 2018 Comparing methods for detecting multilocus adaptation with multivariate genotype–environment associations. *Mol. Ecol.* **27**, 2215–2233. (doi:10.1111/mec.14584)
47. Knaapen JP, Scheffer M, Harms B. 1992 Estimating habitat isolation in landscape planning. *Landsc. Urban Plan.* **23**, 1–16. (doi:10.1016/0169-2046(92)90060-D)
48. Zardi GI, McQuaid CD, Teske PR, Barker NP. 2007 Unexpected genetic structure of mussel populations in South Africa: indigenous Perna perna and invasive Mytilus galloprovincialis. *Mar. Ecol. Prog. Ser.* **337**, 135–144. (doi:10.3354/meps337135)
49. Maree RC, Whitfield AK, Booth AJ. 2000 Effect of water temperature on the biogeography of South African estuarine fishes associated with the subtropical/warm temperate subtraction zone. *South Afr. J. Sci.* **96**, 184–188.
50. Buckley BA, Gracey AY, Somero GN. 2006 The cellular response to heat stress in the goby Gillichthys mirabilis: a cDNA microarray and protein-level analysis. *J. Exp. Biol.* **209**, 2660–2677. (doi:10.1242/jeb.02292)
51. Detrich HW, Parker SK. 1993 Divergent neural β tubulin from the Antarctic fish Notothenia coriiceps neglecta: potential sequence contributions to cold adaptation of microtubule assembly. *Cytoskeleton* **24**, 156–166. (doi:10.1002/cm.970240303)
52. Castilho PC, Buckley BA, Somero G, Block BA. 2009 Heterologous hybridization to a complementary DNA microarray reveals the effect of thermal acclimation in the endothermic bluefin tuna (Thunnus orientalis). *Mol. Ecol.* **18**, 2092–2102. (doi:10.1111/j.1365-294X.2009.04174.x)
53. Jombart T, Devillard S, Balloux F. 2010 Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet.* **11**, 94–94. (doi:10.1186/1471-2156-11-94)
54. Wu CI. 2001 The genic view of the process of speciation. *J. Evol. Biol.* **14**, 851–865. (doi:10.1046/j.1420-9101.2001.00335.x)
55. Nosil P. 2008 Speciation with gene flow could be common. *Mol. Ecol.* **17**, 2103–2106. (doi:10.1111/j.1365-294X.2008.03715.x)
56. Bowen BW, Rocha LA, Toonen RJ, Karl SA. 2013 The origins of tropical marine biodiversity. *Trends Ecol Evol* **28**, 359–366.
57. Nosil P, Funk DJ, Ortiz-Barrientos D. 2009 Divergent selection and heterogeneous genomic divergence. *Mol. Ecol.* **18**, 375–402. (doi:10.1111/j.1365-294X.2008.03946.x)
58. Kingsolver JG. 2009 The well temperatured biologist. *Am. Nat.* **174**, 755–768. (doi:10.1086/648310)
59. Porcelli D, Butlin RK, Gaston KJ, Joly D, Snook RR. 2015 The environmental genomics of metazoan thermal adaptation. *Heredity* **114**, 502–514. (doi:10.1038/hdy.2014.119)
60. Keller I, Seehausen O. 2012 Thermal adaptation and ecological speciation. *Mol. Ecol.* **21**, 782–799. (doi:10.1111/j.1365-294X.2011.05397.x)
61. Teske PR, Papadopoulos I, McQuaid CD, Newman BK, Barker NP. In press. Climate change, genetics or human choice: Why were the shells of mankind's earliest ornament larger in the Pleistocene than in the Holocene? *PLoS ONE* **7**, e614. (doi:10.1371/journal.pone.0000614)
62. Teske PR, Papadopoulos I, Mmonwa KL, Matumba TG, McQuaid CD, Barker NP, Beheregaray LB. 2011 Climate-driven genetic divergence of limpets with different life histories across a southeast African marine biogeographic disjunction: Different processes, same outcome. *Mol. Ecol.* **20**, 5025–5041. (doi:10.1111/j.1365-294X.2011.05307.x)
63. Assis J, Zupan M, Nicastro KR, Zardi GI, McQuaid CD, Serrão EA. 2015 Oceanographic Conditions Limit the Spread of a Marine Invader along Southern African Shores. *PloS One* **10**, e0128124. (doi:10.1371/journal.pone.0128124)
64. Teske P, Bader S, Rao Golla T. 2015 Passive dispersal against an ocean current. *Mar. Ecol. Prog. Ser.* **539**, 153–163. (doi:10.3354/meps11516)
65. Schumann EH, Cohen AL, Jury MR. 1995 Coastal sea surface temperature variability along the south coast of South Africa and the relationship to regional and global climate. *J. Mar. Res.* **53**, 231–248. (doi:10.1357/0022240953213205)
66. Shannon LV, Chapman P. 1983 Suggested mechanism for the chronic pollution by oil of beaches east of Cape Agulhas, South Africa. *Afr. J. Mar. Sci.* **1**, 231–244.
67. Teske PR, Zardi GI, McQuaid CD, Nicastro KR. 2013 Two sides of the same coin: extinctions and originations across the Atlantic/Indian Ocean boundary as consequences of the same climate oscillation. *Front. Biogeogr.* **5**.
68. Angilletta MJ. 2009 *Thermal adaptation: a theoretical and empirical synthesis*. Oxford: Oxford University Press. (doi:10.1093/acprof:oso/9780198570875.001.1)
69. Ridgway TM, Stewart BA, Branch GM, Hodgson AN. 1998 Morphological and genetic differentiation of Patella granularis (Gastropoda: Patellidae): recognition of two sibling species along the coast of Southern Africa. *J. Zool.* **245**, 317–333.
70. Apte S, Gardner JPA. 2002 Population genetic subdivision in the New Zealand greenshell mussel (Perna canaliculus) inferred from single-strand conformation polymorphism analysis of mitochondrial DNA. *Mol. Ecol.* **11**, 1617–1628. (doi:10.1046/j.1365-294X.2002.01554.x)
71. Ayers KL, Waters JM. 2005 Marine biogeographic disjunction in central New Zealand. *Mar. Biol.* **147**, 1045–1052. (doi:10.1007/s00227-005-1632-7)
72. Gleason LU, Burton RS. 2016 Genomic evidence for ecological divergence against a background of population homogeneity in the marine snail Chlorostoma funebralis. *Mol. Ecol.* **25**, 3557–3573.
73. Chu ND, Kaluziak ST, Trussell GC, Vollmer S V. 2014 Phylogenomic analyses reveal latitudinal population structure and polymorphisms in heat stress genes in the North Atlantic snail Nucella lapillus. *Mol. Ecol.* **23**, 1863–1873. (doi:10.1111/mec.12681)

74. Briggs JC. 1974 *Marine zoogeography*. McGraw-Hill.
75. Hutchins LW. 1947 The bases for temperature zonation in geographical distribution. *Ecol. Monogr.* **17**, 325–335.
76. Griffiths CL, Robinson TB, Lange L, Mead A. 2010 Marine biodiversity in South Africa: an evaluation of current states of knowledge. *PLoS ONE* **5**, e12008.
77. Awad AA, Griffiths CL, Turpie JK. 2002 Distribution of South African marine benthic invertebrates applied to the selection of priority conservation areas. *Divers. Distrib.* **8**, 129–145. (doi:10.1046/j.1472-4642.2002.00132.x)
78. Qhaji Y, van Vuuren BJ, Papadopoulos I, McQuaid CD, Teske PR. 2015 A comparison of genetic structure in two low-dispersal crabs from the Wild Coast, South Africa. *Afr. J. Mar. Sci.* **37**. (doi:10.2989/1814232X.2015.1077474)
79. Teske PR, Papadopoulos I, Zardi GI, McQuaid CD, Edkins MT, Griffiths CL, Barker NP. 2007 Implications of life history for genetic structure and migration rates of southern African coastal invertebrates: Planktonic, abbreviated and direct development. *Mar. Biol.* **152**, 697–711. (doi:10.1007/s00227-007-0724-y)
80. Branch GM. 1984 Changes in intertidal and shallow-water communities of the south and west coasts of South Africa during the 1982/1983 temperature anomaly. *South Afr. J. Sci.* **80**, 61–65.
81. Beckley LE. 1986 The ichthyoplankton assemblage of the Algoa Bay nearshore region in relation to coastal zone utilization by juvenile fish. *South Afr. J. Zool.* **21**, 244–252.
82. Whitfield AK. 1989 Ichthyoplankton in a southern african surf zone: Nursery area for the postlarvae of estuarine associated fish species? *Estuar. Coast. Shelf Sci.* **29**, 533–547. (doi:10.1016/0272-7714(89)90009-7)
83. Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH, De Koning D-J. 2011 Stacks: building and genotyping loci de novo from short-read sequences. *Genes|Genomes|Genetics* **1**, 171–182. (doi:10.1534/g3.111.000240)
84. Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA. 2013 Stacks: An analysis tool set for population genomics. *Mol. Ecol.* **22**, 3124–3140. (doi:10.1111/mec.12354)
85. Puritz JB, Hollenbeck CM, Gold JR. 2014 dDocent: a RADseq, variant-calling pipeline designed for population genomics of non-model organisms. *PeerJ* **2**, e431. (doi:10.7717/peerj.431)
86. Garrison E, Marth G. 2012 Haplotype-based variant detection from short-read sequencing. *ArXiv Prepr. ArXiv12073907* , 9. (doi:arXiv:1207.3907 [q-bio.GN])
87. Harrison TD. 2004 Physico-chemical characteristics of South African estuaries in relation to the zoogeography of the region. *Estuar. Coast. Shelf Sci.* **61**, 73–87. (doi:10.1016/j.ecss.2004.04.005)
88. Madeira D, Narciso L, Cabral HN, Vinagre C. 2012 Thermal tolerance and potential impacts of climate change on coastal and estuarine organisms. *J. Sea Res.* **70**, 32–41. (doi:10.1016/j.seares.2012.03.002)
89. Smith EM, Hoi JT, Eissenberg JC, Shoemaker JD, Neckameyer WS, Ilvarsonn AM, Harshman LG, Schlegel VL, Zempleni J. 2007 Feeding Drosophila a biotin-deficient diet for multiple generations increases stress resistance and lifespan and alters gene expression and histone biotinylation patterns 1,. *J. Nutr.* **137**, 2006–2012.
90. Sengupta P, Garrity P. 2013 Sensing temperature. *Curr. Biol. CB* **23**, R304–R307. (doi:10.1016/j.cub.2013.03.009)
91. Park HG, Han SI, Oh SY, Kang HS. 2005 Cellular responses to mild heat stress. *Cell. Mol. Life Sci. CMLS* **62**, 10–23. (doi:10.1007/s00018-004-4208-7)
92. Akashi HD, Díaz AC, Shigenobu S, Makino T, Kawata M. 2016 Differentially expressed genes associated with adaptation to different thermal environments in three sympatric Cuban Anolis lizards. *Mol. Ecol.* **25**, 2273–2285. (doi:10.1111/mec.13625)
93. Zhao H, Yang H, Zhao H, Chen M, Liu S. 2011 Heat stress-mediated gene expression in the body wall of the Japanese sea cucumber Apostichopus japonicus. *Aquat. Biol.* **12**, 23–31. (doi:10.3354/ab00315)
94. Wang Q, Zhao X, Zhang Z, Zhao H, Huang D, Cheng G, Yang Y. 2017 Proteomic analysis of physiological function response to hot summer in liver from lactating dairy cows. *J. Therm. Biol.* **65**, 82–87. (doi:10.1016/j.jtherbio.2017.02.010)
95. Bellantuono AJ, Granados-Cifuentes C, Miller DJ, Hoegh-Guldberg O, Rodriguez-Lanetty M. 2012 Coral thermal tolerance: tuning gene expression to resist thermal stress. *PloS One* **7**, e50685. (doi:10.1371/journal.pone.0050685)
96. Dougherty JJ, Rabideau DA, Iannotti AM, Sullivan WP, Toft DO. 1987 Identification of the 90 kDa substrate of rat liver type II casein kinase with the heat shock protein which binds steroid receptors. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* **927**, 74–80. (doi:10.1016/0167-4889(87)90067-X)
97. Buckley BA, Somero GN. 2009 cDNA microarray analysis reveals the capacity of the cold-adapted Antarctic fish Trematomus bernacchii to alter gene expression in response to heat stress. *Polar Biol.* **32**, 403–415. (doi:10.1007/s00300-008-0533-x)
98. Izadi F, Zarrini HN, Kiani G, Jelodar NB. 2017 Data mining approaches highlighted transcription factors that play role in thermo-priming. *Plant Omics* **10**, 139–145.
99. Chiappori F, Pucciarelli S, Merelli I, Ballarini P, Miceli C, Milanesi L. 2011 Structural thermal adaptation of β-tubulins from the Antarctic psychrophilic protozoan Euplotes focardii. *Proteins Struct. Funct. Bioinforma.* **80**, 1154–1166. (doi:10.1002/prot.24016)
100. Bedulina D, Meyer MF, Gurkov A, Kondratjeva E, Baduev B, Gusdorf R, Timofeyev MA. 2017 Intersexual differences of heat shock response between two amphipods (Eulimnogammarus verrucosus and Eulimnogammarus cyaneus) in Lake Baikal. *PeerJ* **5**, e2864. (doi:10.7717/peerj.2864)