Supplementary materials

Background information on study area and species

South Africa is located at the oceanographic interface between the Atlantic and Indian Oceans, and the number of species recorded particularly along its southern coastline is greater than that of either the south-eastern Atlantic or the south-western Indian Ocean [1]. While this can be partly explained by the fact that the fauna of the two oceans overlaps in South Africa, there is also evidence for large numbers of endemic species [1,2]. Many of these are genetically subdivided and may comprise complexes of evolutionary lineages that represent cryptic or incipient species [3], suggesting that the region's true biodiversity may be much greater than is presently acknowledged. The boundaries between the ranges of many of these lineages coincide with one of the region's four temperature-defined marine provinces, namely the cool-temperate west coast, warm-temperate south coast, subtropical east coast and tropical north-east coast [3], and some species even have ranges that are limited to the transition zones between these provinces [4,5]. However, there are also numerous species in which the range of a particular evolutionary lineage spans multiple marine provinces (Fig. 1). In some species, the amount of gene flow between provinces can be considerable when conditions are favourable, but the migrants are unable to establish themselves in the long term [6,7]. Physiological studies indicate that for some species, sister lineages present in adjacent provinces exhibit distinct temperature preferences [8,9], which suggests that thermal adaptation contributes towards limiting gene flow between biogeographic regions by reducing migrant fitness and, if a conspecific sister lineage is present, by subjecting the migrants to competitive exclusion [3]. The Knysna sandgoby, *Psammogobius knysnaensis* (species 13 in Fig. 1), is one of the most common estuarine fishes in South Africa [10,11]. It has sedentary adults, so connectivity among populations is likely facilitated by its planktonic

larvae, which are transported out of estuaries in large numbers during ebb-tides [12] and use nearshore areas as nurseries [13]. Although the species' planktonic larval duration is not known, it is likely similar to that of other gobies, which is in the order of weeks to months [14,15]. Compared to the other candidate species with similarly wide distributions, the purity of its extracted DNA, which was extracted using the CTAB protocol [16], was much higher. Based on DNA sequence data from the mitochondrial COI gene and the nuclear S72 intron, the species exists as a single lineage whose range spans the west coast, the cool-temperate/warm-temperate transition zone on the south-west coast, the south coast and the warm-temperate/subtropical transition zone on the southern east coast [17]. Although genetic structure was identified between the west coast (site 1 in Fig. 1) and all other regions on the basis of mtDNA, this is not a result of divergent evolution (Fig. 3b), but of lower genetic diversity on the cool-temperate west coast [17].

Supplementary methods

Processing of raw sequences

Raw sequences were processed using the process_radtags.pl script in STACKS 1.19 [18,19]. Briefly, we eliminated reads with ambiguous barcode and/or restriction site, allowing a maximum of two mismatches, and all remaining reads were trimmed to 80 bp by removing the barcode, the restriction site, and the last 8 bp. The dDocent 1.0 software pipeline [20] was then used to remove low quality bases using a quality score limit of 30. Putative RAD reads were identified *de novo* using a minimum depth of 15×, and the selected unique reads where then aligned with a maximum of 8 mismatches to form reference contigs. The reference contigs were then clustered based on an 80% similarity threshold. Single nucleotide polymorphisms (SNPs) were called from the aligned good quality reads of multiple individuals, using a Bayesian-based variant detection approach [21]. After excluding

54 individuals with >20% of missing data, the final data set comprised 109 individuals. Variants
55 were filtered to produce a file containing only the best quality biallelic SNPs per locus, which
56 were present in at least 80% of the individuals sequenced, with a minimum allele frequency
57 equal to or greater than 3% (please see Table S4 for additional detail).
58
59 **Identification of loci under thermal selection and neutral loci**
60 The contribution of coastal sea surface temperature (SST) to the overall pattern of genetic
61 differentiation was assessed using a spatially explicit generalized linear mixed model to test
62 for direct associations between SNP allele frequencies and temperature-related variables,
63 while controlling for the effects of spatial structure and shared population history,
64 implemented in the R package gINLAnd [22]. Unlike F_{ST}-based outlier scans, which identify
65 loci on the basis of population information [23], the identification of loci in in genotype-
66 environment association methods such as gINLAnd is thus not influenced by any regional
67 population structure [24]. Briefly, gINLAnd estimates the covariance associated with the
68 spatial distribution of the samples and a locus-specific effect of each environment variable; it
69 then estimates the likelihood of two competing models: a model with the environmental
70 effect and a reduced model without the environmental effect. Finally, gINLAnd assesses the
71 strength of genetic dependence on the environmental variable by computing a Bayes factor
72 between the two models. To avoid false positives, we used a conservative approach in which
73 only those loci which showed a log Bayes factor (BF) ≥10 were considered to be under
74 selection (a log BF>4.6 is considered decisive [25]). A plot depicting loci under selection is
75 shown in Fig. S1. We calculated a multidimensional scaling projection of the coastal distance
76 between sampling sites using the R package MASS 7.3 [26], which is more meaningful than
77 using the original geographic coordinates because this would have required connecting sites
78 via terrestrial habitat.

To identify 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 [23] 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 [27]. 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.

Temperature data

Temperature data originated from sites that are located in close proximity to the estuaries or lagoons in which samples of *P. knysnaensis* were collected [28] (Table S1). While coastal SST strongly influences thermal conditions within estuaries that are connected to the sea, many southern African estuaries become temporarily disconnected from the sea for prolonged periods of time, and their water temperatures tend to be higher [29]. Moreover, estuaries exhibit much more extensive variation in physico-chemical characteristics compared to subtidal nearshore areas, and their biota are considered to be particularly resilient to cope with this variability [30]. Adult fish in estuaries thus have very wide thermal tolerance ranges [31], and are unlikely to be under strong thermal selection. In contrast, the species' planktonic larvae, which emigrate from estuaries [13] and during their marine dispersal phase will be strongly affected by coastal SST, are likely to be more sensitive to water temperature because fish larvae have limited metabolic scope and thus limited

tolerance to environmental stressors [32]. This suggests that coastal rather than estuarine temperatures are particularly suitable to explore thermal selection in this species.

Environmental variables run with the program gINLAnd were taken from Smit et al.[28] and included minimum and maximum temperatures, as well as the coolest and warmest 5% of temperature measurements (Table S3). We only report results for minimum and maximum temperatures as they showed the trends most clearly (Fig. S4) and differed only slightly from the other temperature readings (Table S1).

Analyses of genetic structure

a) Discriminant Analysis of Principal Components (DAPC)

DAPC analyses were performed with the R package ADEGENET v. 2.1.0 [33]. 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. 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 [34] 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 S3 for detail on number of SNPs identified).

b) *fastStructure* analyses

The program *fastStructure* uses variational Bayesian inference under a model assuming Hardy-Weinberg equilibrium and linkage equilibrium, and was primarily used here to confirm specific results based on DAPC. 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 [35].

c) BEAST analyses

The dataset used for discrete phylogeographic analyses in BEAST comprised data from the individual alleles of each individual, which were reconstructed in PHASE v. 2.1.1 [36] 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.

Supplementary results

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.

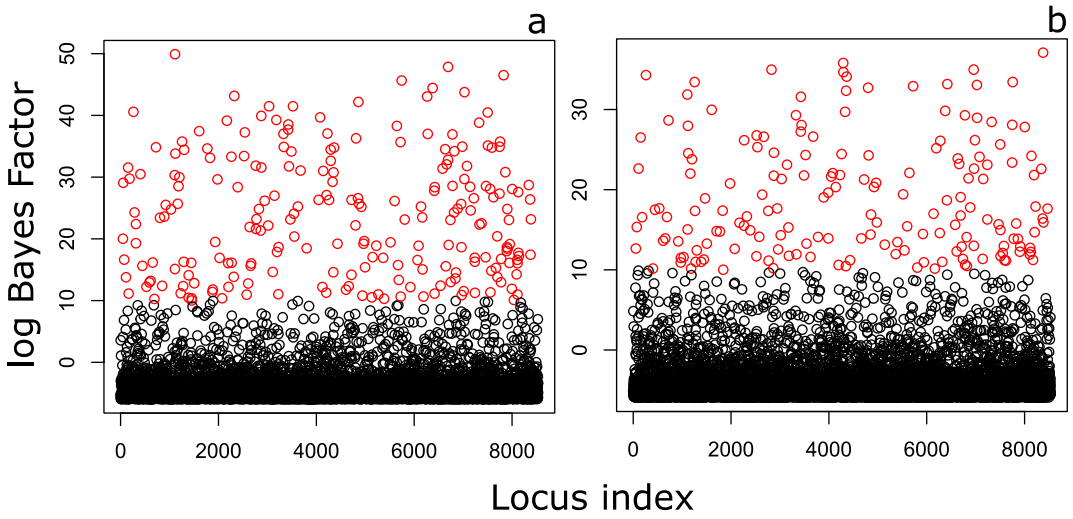

Fig S1. Examples of of loci identified as being linked to genomic regions under thermal selection (red dots) while controlling for geographical distances between sites by gINLAnd analysis. Plots depict results for (**a**) minimum temperatures and (**b**) maximum temperatures. All log Bayes factors (log BF) > 0 indicate potential thermal selection, and the likelihood that these are false positives decreases with increasing log BF. In the loci shown in red (log BF>10), the association between genomic and environmental variables is decisive.

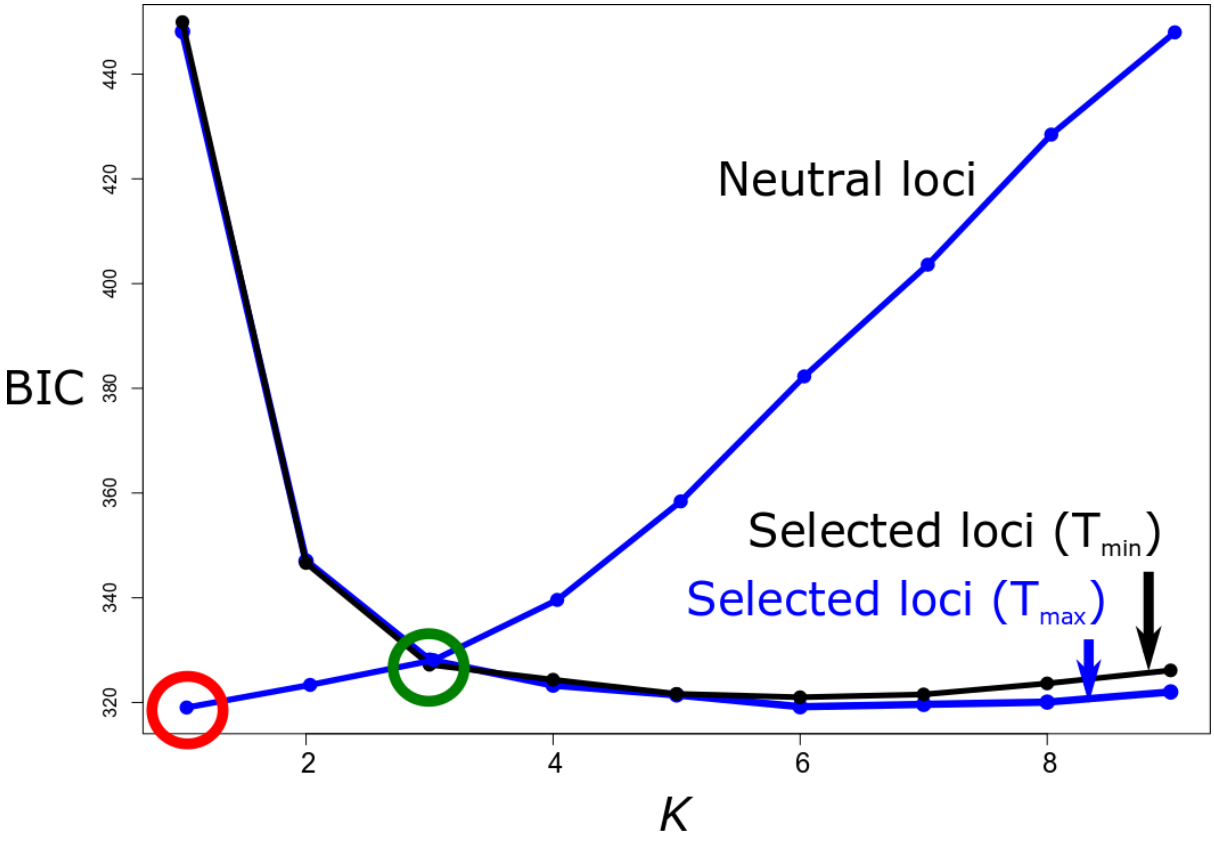

Fig S2. Analyses of candidate loci putatively under thermal selection and selectively neutral loci. The best value of *K* for selectively neutral data is based on the lowest Bayesian Information Criterion (BIC) value (red circle), whereas plots for temperature-associated loci (controlling for geographical distance) for minimum and maximum temperatures (T_{min} and T_{max}, respectively) have a distinct 'elbow' at $K = 3$ beyond which further decrease in BIC is minimal (green circle). Some subsequent analyses were nonetheless conducted with 4 clusters to show differentiation between the pre-defined marine bioregions and to compare the results of DAPC analyses with those of the fastSTRUCTURE analyses.

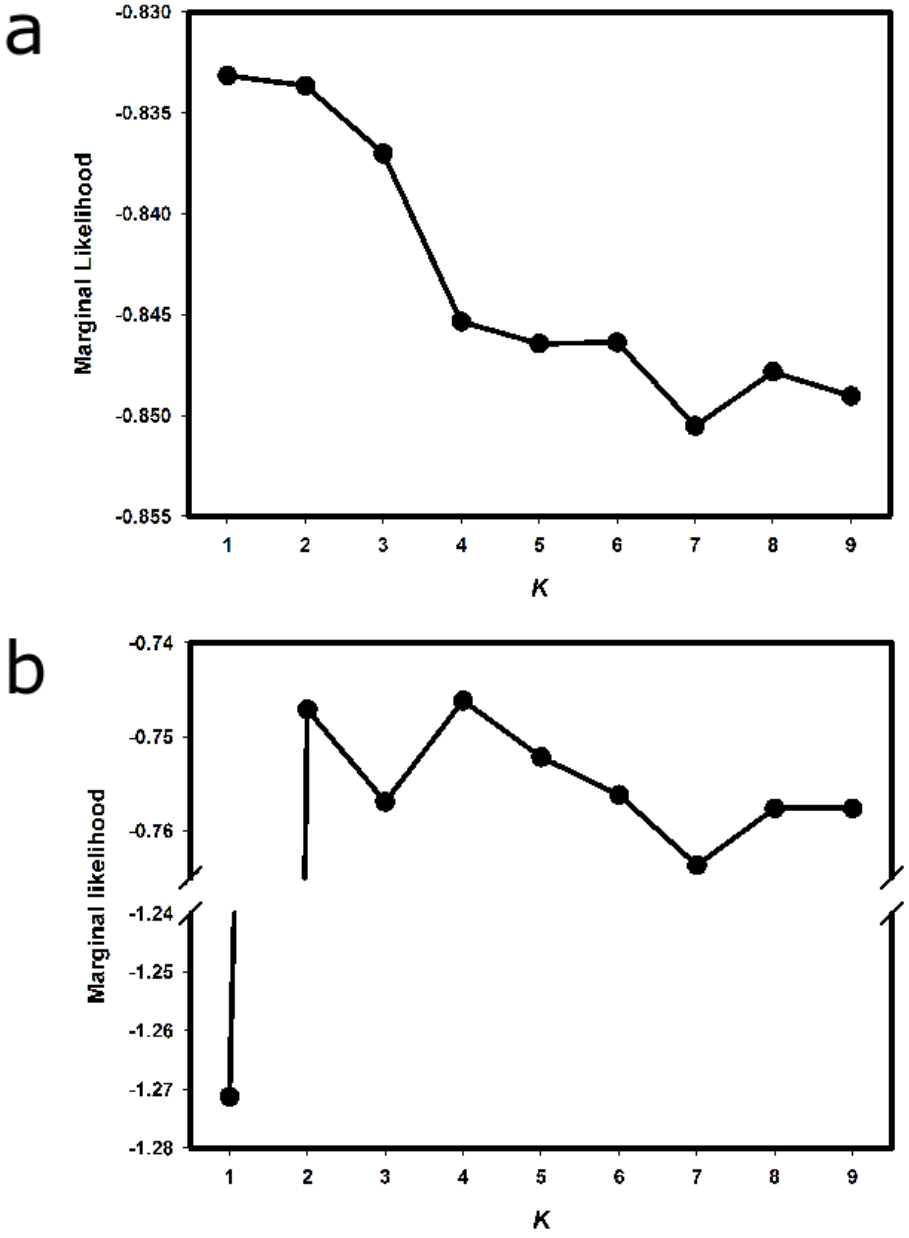

Fig S3. Support for number of clusters in fastStructure. Plots depict marginal likelihoods for number of genetic clusters (*K*) for (**a**) neutral loci, highest likelihood at *K* = 1; (**b**) temperature-associated loci (T_{min}), highest likelihood at *K* = 4.

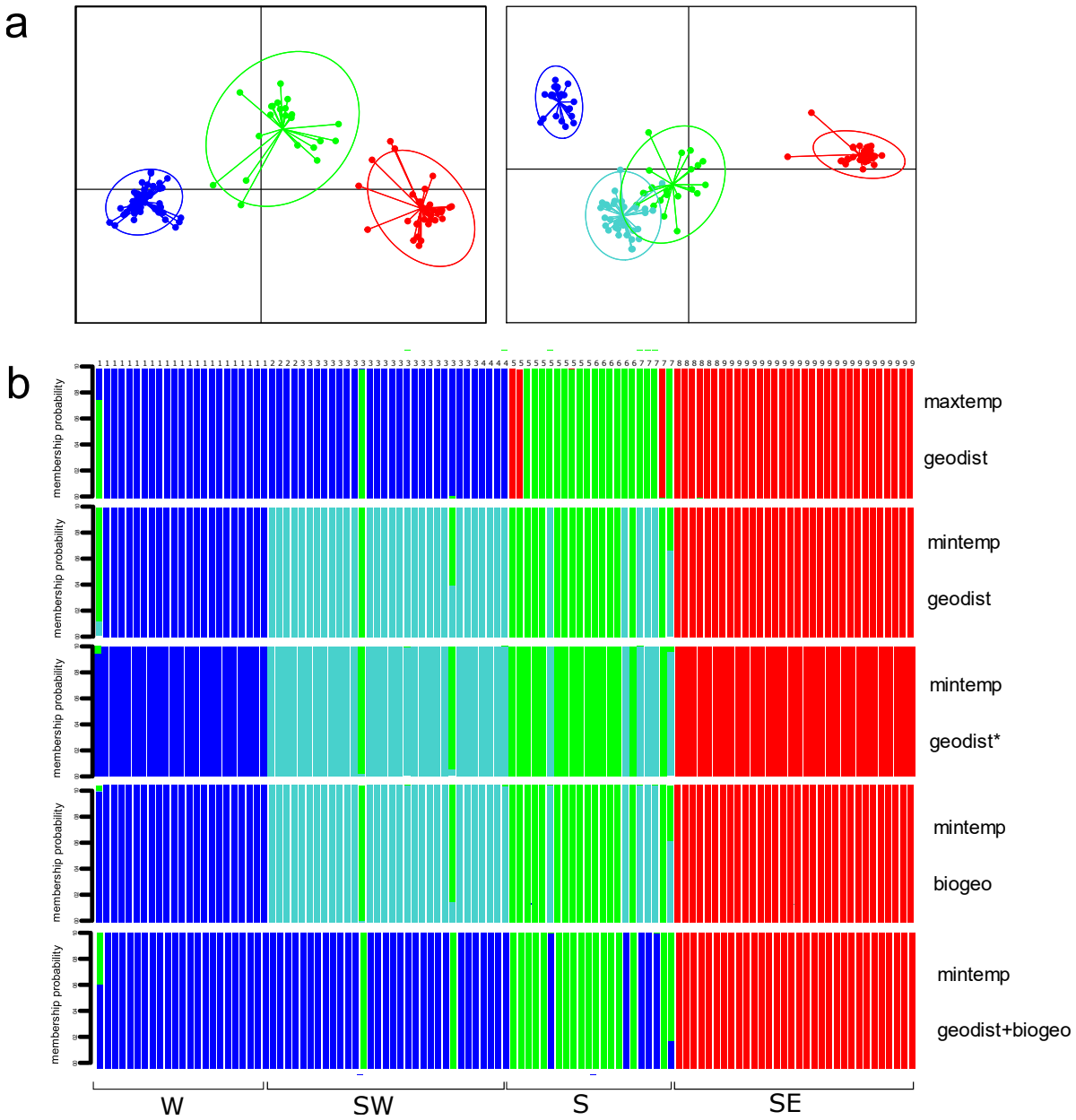

Fig S4. Population genetic structure inferred for temperature-associated loci using DAPC. **a** Scatterplots with inertia ellipses representing 95% confidence intervals, with colours reflecting the dominant bioregion represented in a particular cluster (left: loci correlated with maximum temperature; 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).

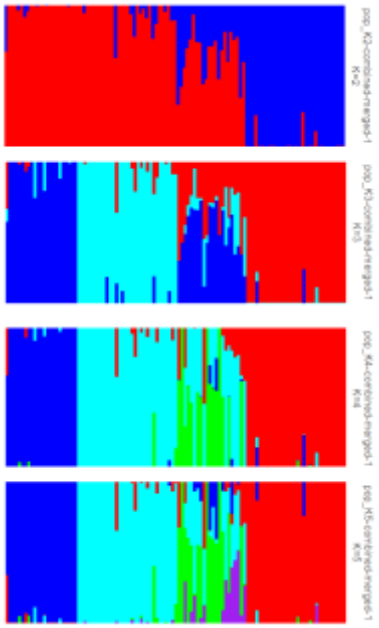

Fig S5. Comparison of concensus fastStructure barplots for *K*=2-5 based on temperature-associated loci (T_{min}).

219

Table S1. Sampling sites, collection dates, number of high quality samples used for the analyses, and temperature data from nearby locations.

Region	Site no.	Site name	Date	GPS coordinates	N	Temperature location*	min	5%	95%	max
W	1a	Langebaan (Kraalbaai)	Feb 2016	33°08'31.32"S 18°01'26.26"E	12	Langebaan	11.9	12.1	19.6	20.2
W	1b	Langebaan (Shark Bay)	Dec 2016	33°06'42.42"S 18°02'24.47"E	11	Langebaan	11.9	12.1	19.6	20.2
SW	2	Rooiels	Oct 2015	34°17'53.48"S 18°49'11.10"E	4	Betty's Bay	12.9	13.0	18.2	18.6
SW	3	Palmiet	Dec 2016	34°20'36.33"S 18°59'39.57"E	23	Betty's Bay	12.9	13.0	18.2	18.6
SW	4	Klein	Dec 2016	34°25'08.06"S 19°18'23.78"E	4	Hermanus	12.2	13.6	19.1	19.9
S	5	Breede	Dec 2016	34°23'47.19"S 20°50'32.45"E	11	Stilbaai	12.8	13,0	23.5	23.9
S	6	Klein Brak	Oct 2015	34°05'29.82"S 22°08'45.12"E	6	Mossel Bay	12.9	13.7	23.4	23.8
S	7	Swartkops	Oct 2015	33°52'09.06"S 25°37'43.13"E	6	Humewood	13.7	14.4	22.3	22.5
SE	8	Nenga	Oct 2015	31°59'02.81"S 29°09'03.83"E	6	Cwebe	16.9	17.0	21.6	21.7
SE	9a	Bululo	Oct 2015	31°39'02.05"S 29°30'59.14"E	6	Mzamba	18.3	18.6	24.0	24.1
SE	9b	Bululo	May 2016	31°39'02.05"S 29°30'59.14"E	20	Mzamba	18.3	18.6	24.0	24.1

*Smit et al. 2013; min = minimum temperature reading; 5% = coldest 5% of temperature readings; 95% = warmest 5% of temperature readings; max = maximum temperature reading.

220

Table S2. Genes identified among temperature-associated loci whose activity may be directly related to temperature stress. Loci marked with an asterisk were found for T_{min} only, those without asterisk for both T_{min} and T_{max}.

Locus	Uniprot	Gene	Temperature related activity
Contig_69368	A0A1S3GQV0_DIPOR	Histone H3	Biotinylation of Histone H3 is related resistance to heat stress in flies[37]. Related with control of expression of other proteins.
Contig_56520	F1QC71_DANRE	ANO5b	Part of a family of proteins (Anoctamin) that are key mediators of thermal sensing, most do not act as temperature sensors themselves but rather support the behavior in some other capacity, such as by controlling the development or function of the thermo-sensory cell or circuitry[38].
Contig_22921*	A0A1A8EID4_9TELE	SHC4	Moderate heat stress has been shown to induce maximal gene activity to promote tissue repair[39]. The gene was also indicative of adaptation to different thermal environments in *Anolis* lizards[40].
Contig_247440	G3TCM9_LOXAF	MBD2	Regulates the expression of other genes during heat stress in sea cucumbers[41].
Contig_168847	A0A147A7G3_FUNHE	14-3-3 protein gamma	Part of a protein family that has been associated with heat shock responses in gobies[42]. Increased expression in cows exposed to high temperatures, possibly to protect against liver cell apoptosis[43]. Differential expression in corals in response to a thermal challenge[44].

Contig_38433	A0A0F8AG31_LARCR	EH28_02053 (Tyrosine-protein kinase)	Interacts with Hsp90[45] and its activity increases during heat stress in a goby species[42]. Altered gene expression in response to heat stress was also recorded in the Antarctic fish *Trematomus bernacchii*[46].
Contig_93683	A0A1A8MJC6_9TELE	PRSS12	Expressed under heat stress condition in sea cucumbers[41].
Contig_330391	A0A146TQG3_FUNHE	Reverse transcriptase-like protein	Upregulated as early response to heat shock in plants[47].
Contig_351821*	S7PGQ6_MYOBR	Tubulin beta chain	Divergence on tubulin genes has been associated to adaptation to cold temperatures in Antartic fishes[48] and protozoans[49]. It also changes expression levels during thermal acclimation in bluefin tuna[50] and is differentially expressed in response to heat shock in crustaceans[51].

221

222

Table S3. Number of loci identified for gINLAnd analyses.

Temperature	Controlling factor	No. loci
Maximum	Geographic distance	175
Minimum	Geographic distance	226
Minimum	Geographic distance*	196
Minimum	Biogeography	239
Minimum	Geographic distance + Biogeography	169

*Only individuals from the W coast (1) and SE coast (8, 9) were included.

227

Table S4. Summary of loci obtained following raw data processing.

Region	Loci present with less than 20% missing data	Variable loci	Percentage of variable loci	H_O	H_E	Genotyped samples
W	8532	4122	48.3	0.2834	0.2703	23
SW	8532	4206	49.3	0.2849	0.2710	31
S	8532	4288	50.3	0.2592	0.2855	23
SE	8532	4251	49.8	0.2659	0.2628	32

228

229

Supplementary references

1. 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.

2. 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)

3. 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)

4. 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)

5. 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)

6. 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.

7. 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.

8. 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)

9. 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)

10. Strydom NA. 2015 Patterns in larval fish diversity, abundance, and distribution in temperate South African estuaries. *Estuaries Coasts* **38**, 268–284. (doi:10.1007/s12237-014-9801-x)

11. Kruger M, Strydom NA. 2010 Spatial and temporal variability in the larval fish assemblage of a warm temperate South African estuary, with notes on the effects of artificial channelling. *Afr. Zool.* **45**.

12. Beckley LE. 1985 Tidal exchange of ichthyoplankton in the Swartkops estuary mouth, South Africa. *Afr. Zool.* **20**, 15-20–20.

13. 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)

14. Sponaugle S, Cowen RK. 1994 Larval durations and recruitment patterns of two Caribbean gobies (Gobiidae): contrasting early life histories in demersal spawners. *Mar. Biol.* **120**, 133–143. (doi:10.1007/BF00381949)

15. Beldade R, Pedro T, Gonçalves EJ. 2007 Pelagic larval duration of 10 temperate cryptobenthic fishes. *J. Fish Biol.* **71**, 376–382. (doi:10.1111/j.1095-8649.2007.01491.x)

16. Doyle J. 1991 CTAB Total DNA Isolation. In *Molecular Techniques in Taxonomy*, pp. 283–293. (doi:10.1007/978-3-642-83962-7_18)

17. 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)

18. 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)

19. 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)

20. 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)

21. Garrison E, Marth G. 2012 Haplotype-based variant detection from short-read sequencing. *ArXiv Prepr. ArXiv12073907* , 9. (doi:arXiv:1207.3907 [q-bio.GN])

22. 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)

23. 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)

24. 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)

25. Kass RE, Raferty AE. 1995 Bayes factors. *J. Am. Stat. Soc.* **90**, 773–795.

26. Venables WN, Ripley BD. 2002 *Modern applied statistics with S.* 4th editio. New York: Springer. (doi:10.1198/tech.2003.s33)

27. 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)

28. 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)

29. 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)

30. Elliott M, Quintino V. 2007 The estuarine quality paradox, environmental homeostasis and the difficulty of detecting anthropogenic stress in naturally stressed areas. *Mar. Pollut. Bull.* **54**, 640–645. (doi:10.1016/j.marpolbul.2007.02.003)

31. 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)

32. Killen SS, Costa I, Brown JA, Gamperl AK. 2007 Little left in the tank: metabolic scaling in marine teleosts and its implications for aerobic scope. *Proc. R. Soc. B Biol. Sci.* **274**, 431–438. (doi:10.1098/rspb.2006.3741)

33. 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)

34. 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)

35. 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)

36. Stephens M, Smith NJ, Donnelly P. 2001 A New Statistical Method for Haplotype Reconstruction from Population Data. *Am. J. Hum. Genet.* **68**, 978–989.

37. 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.

38. Sengupta P, Garrity P. 2013 Sensing temperature. *Curr. Biol. CB* **23**, R304–R307. (doi:10.1016/j.cub.2013.03.009)

39. 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)

40. 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)

41. 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)

42. 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)

43. 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)

44. 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)

45. 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)

46. 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)

47. 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.

48. 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)

49. 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)

50. 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)

51. 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)