PROCEEDINGS B

royalsocietypublishing.org/journal/rspb

Research

**Cite this article:** Clark RD, Aardema ML, Andolfatto P, Barber PH, Hattori A, Hoey JA, Montes Jr HR, Pinsky ML. 2021 Genomic signatures of spatially divergent selection at clownfish range margins. *Proc. R. Soc. B* **288**: 20210407.
https://doi.org/10.1098/rspb.2021.0407

Received: 19 February 2021
Accepted: 11 May 2021

**Subject Category:**
Evolution

**Subject Areas:**
evolution, ecology, genomics

**Keywords:**
spatially divergent selection, range margin, clownfish, local adaptation

**Author for correspondence:**
René D. Clark
e-mail: rene.clark@rutgers.edu

One contribution to a special feature: 'Evolution in Changing Seas'. Guest edited by: Katie E. Lotterhos, Molly Albecker and Geoffrey Trussell.

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.5438761.

THE ROYAL SOCIETY PUBLISHING



# Genomic signatures of spatially divergent selection at clownfish range margins

René D. Clark[1], Matthew L. Aardema[2,3], Peter Andolfatto[4], Paul H. Barber[5], Akihisa Hattori[6], Jennifer A. Hoey[1,7], Humberto R. Montes Jr[8] and Malin L. Pinsky[1]

[1]Department of Ecology, Evolution and Natural Resources, Rutgers University, 14 College Farm Road, New Brunswick, NJ 08901, USA
[2]Department of Biology, Montclair State University, 1 Normal Avenue, Montclair, NJ 07043, USA
[3]Sackler Institute for Comparative Genomics, American Museum of Natural History, 200 Central Park West, New York, NY 10024-5102, USA
[4]Department of Biological Sciences, Columbia University, New York, NY 10026, USA
[5]Department of Ecology and Evolutionary Biology, University of California-Los Angeles, Los Angeles, CA 90095, USA
[6]Faculty of Liberal Arts and Education, Shiga University, 2-5-1 Hiratsu, Otsu, Shiga 520-0862, Japan
[7]Department of Ecology and Evolutionary Biology, University of California-Santa Cruz, 130 McAllister Way, Santa Cruz, CA 95060, USA
[8]Visayas State University, Baybay City, Leyte, Philippines

RDC, 0000-0002-9874-8863; MLA, 0000-0002-7206-6311; PA, 0000-0003-3393-4574; PHB, 0000-0002-1486-8404; AH, 0000-0003-0848-1024; JAH, 0000-0002-9041-915X; MLP, 0000-0002-8523-8952

Understanding how evolutionary forces interact to drive patterns of selection and distribute genetic variation across a species' range is of great interest in ecology and evolution, especially in an era of global change. While theory predicts how and when populations at range margins are likely to undergo local adaptation, empirical evidence testing these models remains sparse. Here, we address this knowledge gap by investigating the relationship between selection, gene flow and genetic drift in the yellowtail clownfish, *Amphiprion clarkii*, from the core to the northern periphery of the species range. Analyses reveal low genetic diversity at the range edge, gene flow from the core to the edge and genomic signatures of local adaptation at 56 single nucleotide polymorphisms in 25 candidate genes, most of which are significantly correlated with minimum annual sea surface temperature. Several of these candidate genes play a role in functions that are upregulated during cold stress, including protein turnover, metabolism and translation. Our results illustrate how spatially divergent selection spanning the range core to the periphery can occur despite the potential for strong genetic drift at the range edge and moderate gene flow from the core populations.

## 1. Introduction

Understanding how environmental heterogeneity drives patterns of selection and partitions adaptive variation into discrete populations is increasingly important in today's changing world. Species with large geographical ranges often span a similarly wide array of environmental conditions, which may result in natural selection favouring distinct sets of alleles across the species range [1]. Such a process is commonly referred to as spatially divergent selection and can shape the evolutionary trajectory of a population by causing allele frequencies at select loci to move away from a global mean and towards local optima [2]. While there exists a large body of work investigating spatially divergent selection, the scale at which such patterns can manifest, and the extent to which populations may become locally adapted, remains an area of intense debate [3,4].

Selection, however, is only one of several evolutionary processes operating in a natural system and rarely acts in isolation. Gene flow and drift also

shape the distribution of genetic variation, and the interplay of these forces can influence population dynamics, patterns of range expansion and evolutionary trajectories [5,6]. Gene flow is commonly thought to decrease the fitness of edge populations through gene swamping, as immigrants from the range core may be suboptimally adapted to edge environments [7,8]. Alternatively, gene flow can increase adaptive potential at the edge by transporting in novel genetic variation and replenishing genetic diversity that is otherwise depleted owing to drift and serial founder events [9,10].

Empirical evidence detailing how migration–selection balance affects adaptation in peripheral populations is rare, especially in marine taxa ([11]; although see [12]). Many marine species are historically thought to have large, well-mixed populations owing to a lack of geographical barriers and high dispersal capabilities [13,14]. Such conditions would tip migration–selection balance away from local adaptation and towards range-wide adaptation to a global trait mean. However, recent studies have found that adaptive and neutral genetic variation can be differentially distributed within marine populations, suggesting that selection may be strong enough to cause differentiation even in the face of heavy gene flow [15,16] and across pronounced environmental gradients [17,18].

Despite such evidence, it is still unclear how selection and gene flow interact to shape adaptation at range peripheries. Populations near the range centre are thought to maintain high enough densities to withstand an influx of maladaptive alleles from elsewhere in the range [7]. Less dense populations at the range edge may not benefit from the same demographic processes, however [6]. Nevertheless, directional selection at the edge may be strong enough to overcome asymmetrical migration rates, especially as peripheral populations tend to inhabit novel environments [6,8]. Thus, our understanding of how evolutionary forces interact to either suppress or replenish genetic diversity at marine range peripheries remains lacking.

At the same time, it is these edge populations that play a critical role in enabling species to adapt to changing environments [19]. Species shift their ranges to track environmental conditions, and it is often individuals at the edge that first colonize novel habitats [20]. The oceans are predicted to warm by 1–3°C over the coming century, presenting serious challenges for marine taxa [21]. Many biological pathways are sensitive to temperature change [22], and water temperature has been shown to influence development, reproduction and survival in many ectotherms [23]. Thus, adaptation to thermal regimes often provides a large fitness advantage [24] and shifting thermal environments may impose strong selective pressures on populations as they respond *in situ*. Such responses may vary across a species range, as adaptive potential is unlikely to be uniform from the core to the periphery. Thus, understanding the roles that evolutionary forces play in partitioning genetic variation among populations is important when trying to predict how species will fare in the coming years.

The yellowtail clownfish, *Amphiprion clarkii*, provides an ideal system for investigating how different evolutionary processes interact to shape genetic diversity and adaptive potential across a species range. *Amphiprion clarkii* occupies one of the broadest latitudinal ranges of all anemonefish species, with populations from the Indo-Pacific tropics to the subtropics (figure 1) [25]. Anemonefish are philopatric,

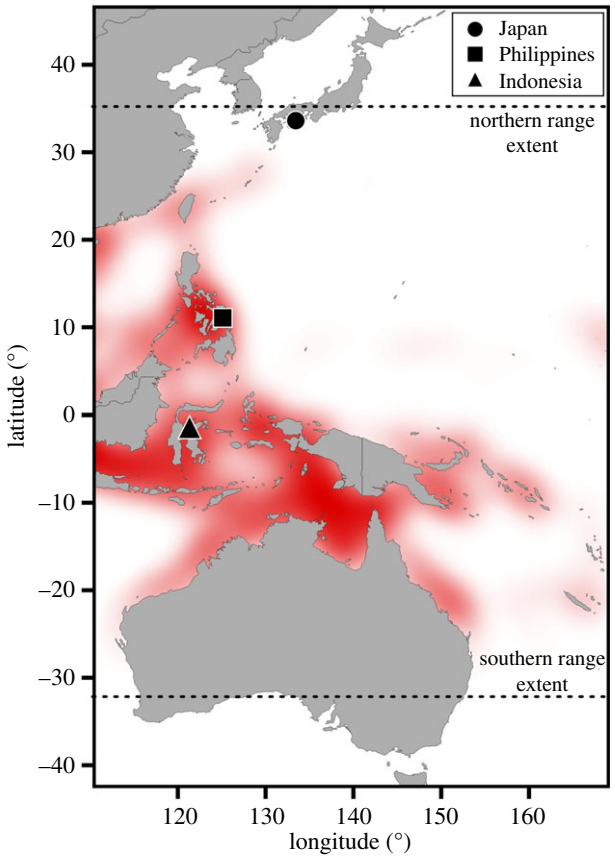


**Figure 1.** Map of the three sampling locations: Japan ($n = 8$), Philippines ($n = 10$) and Indonesia ($n = 7$). Northern and southern range extents are marked with horizontal dashed lines. Red shading indicates the relative probability of occurrence of *A. clarkii* (data from AquaMaps). (Online version in colour.)

remaining on the same anemone for the duration of their adult lives [25]; as such, populations are only connected by pelagic larval dispersal (PLD). However, with a PLD of approximately two weeks [26], the larval duration of *A. clarkii* is relatively limited, and few larvae regularly disperse farther than 27 km, although rare long-distance migration is possible [27].

Here, we explore the relationship between gene flow, selection and drift across the northern half of the species range of *A. clarkii* to determine if edge populations have responded to spatially divergent selection despite (or perhaps because of) gene flow from the core and the potential for stronger genetic drift at the edge. Specifically, we address three questions: (i) how is genetic diversity, both adaptive and neutral, partitioned across the northern extent of *A. clarkii*'s range; (ii) are there signatures of selection in edge and core populations; and (iii) what are the underlying biological or molecular processes targeted by such selection?

## 2. Material and methods

### (a) Study species and sample collection

We sampled a total of 25 *A. clarkii* individuals (4–11 cm fork length) from three locations. These locations represent the core of the species distribution (near the equator), part-way to the northern edge and near the northern edge of the species range. We collected seven individuals from Sulawesi Tengah, Indonesia (six from 0.652217 °S, 119.739 °E and one from 0.65695 °S,

119.741 °E), 10 from Leyte, Philippines (10.87304 °N, 124.7122 °E) and eight from Shikoku Island, Japan (33.005133 °N, 132.5047 °E) (figure 1). All sampling took place in 2012 (May–August). Upon capture, a sample of heart tissue was taken and immediately preserved in RNAlater. The heart was chosen as it plays an important role in determining thermal sensitivity through oxygen transport and aerobic capacity [28]. As taking heart tissue is lethal, we designed our sampling scheme to capture the largest sample possible while minimizing cost and population impact.

We extracted total RNA using a Qiagen RNAeasy spin column (Qiagen, Hilden, Germany) following the manufacturer's recommendations and made cDNA libraries with an Illumina TruSeq v2 kit (Illumina, San Diego, CA, USA) at half reaction volumes. We assessed concentrations with a Qubit dsDNA HS assay (ThermoFisher Scientific, Waltham, MA, USA), quality with a NanoDrop spectrophotometer (ThermoFisher Scientific), and fragment length with an Agilent 2100 BioAnalyzer and a DNA 1000 kit (Agilent, Santa Clara, CA, USA). We sequenced the libraries on an Illumina HiSeq 2500 (Illumina) with 140 and 187 bp single-end reads at Princeton University's Lewis-Sigler Institute Genomics Core Facility.

## (b) Read mapping and variant calling

We demultiplexed the sequenced reads by Illumina index using a Python script adapted from FASTX Barcode Splitter [29], trimmed to bases with a quality score greater than 20 with TQSfastq.py from the SSAKE assembly pipeline [30], and removed reads less than 30 bases long after trimming. Reads from individual N3 were assembled into a de novo reference transcriptome using Trinity v. 2.2.0 [31] (details in the electronic supplementary material). The final reference transcript contained 103 518 transcripts, varying in length from 201 to 7323 nucleotides. We mapped reads to the reference transcriptome using Stampy v. 1.0.28 [32], filtered for mapping quality greater than 20 using SAMtools v. 1.3.0 [33], marked read duplicates using the MarkDuplicates function in Picard Tools v. 1.119 (https://broadinstitute.github.io/picard/), and realigned indels using IndelRealigner in GATK v. 3.8.1 [34]. Variants were called with HaplotypeCaller in GATK. We removed all variants except biallelic single nucleotide polymorphisms (SNPs) genotyped in at least 24 samples. Additional filtering with VCFtools v. 1.16 [35] removed SNPs with a minor allele count of less than 2. After filtering, we had 4212 SNPs, distributed across 1002 transcripts.

## (c) Genetic diversity

We assessed two measures of genetic diversity and one measure of relatedness. We calculated per-site nucleotide diversity ($\pi$) for each sampling site using VCFtools. The mean inbreeding coefficient ($F_{IS}$) was estimated from observed heterozygosity and expected gene diversity with the *hierfstat* package in R v. 3.4.4 [36,37]. We assessed the mean within-population pairwise relatedness using the *relatedness* R package [38] and the Wang relatedness estimator [39] that has reduced bias with small sample sizes [40]. We calculated 95% confidence intervals (CIs) for all metrics by bootstrapping with replacement across individuals 1000× in R. Tajima's *D* was calculated across each transcript using VCFtools. To include rare variants and avoid potential biases from purifying selection, Tajima's *D* was calculated using only synonymous sites from an SNP dataset unfiltered for minor allele count (1453 SNPs). Synonymous sites were identified using SnpEff [41] (see the annotation section for details).

## (d) Outlier test and environmental association analyses

To identify candidate SNPs under selection, we used an outlier test and two environmental association analyses (EAA). For the outlier analysis, we ran the core model implemented in the program Baypass v. 2.1.1 [42] using default parameters and all 4212 SNPs. This model generates an XtX statistic [43], which is an $F_{ST}$-type measurement that considers population structure. We determined a threshold XtX value for outliers by creating pseudo-observed datasets under a null model and analysed them with the core model [42]. We used the 99% quantile of this empirical XtX distribution under no selection as the selection/neutrality threshold.

For the EAAs, we used two methods. First, we ran the standard covariate model implemented in Baypass which tests for associations between allele frequencies and environmental covariables while accounting for the neutral covariance among localities. We used annual mean sea surface temperature (SST mean), minimum SST (SST min), maximum SST (SST max), latitude and mean sea surface salinity (SSS mean) (electronic supplementary material, table S2) from the MARSPEC database [44]. For every variable, we ran a burn-in of 5000 iterations and then 25 000 Markov chain Monte Carlo (MCMC) steps thinned to every 25. We used the full dataset of 4212 SNPs. Bayes factors (BFs) in deciban (dB) units were used to determine whether an SNP was associated with an environmental variable. As recommended in [42], we considered SNPs with a BF greater than 20 dB to be strongly associated. To assess these associations, we randomly reassigned individuals among locations to create permuted datasets that we then analysed in Baypass (details in the electronic supplementary material).

Our second EAA was a redundancy analysis (RDA). We used the *vegan* v. 2.4.1 R package [45] to perform RDA with the same environmental variables as in Baypass and a centred allele frequency dataset with all 4212 SNPs. We used two methods to identify potential outlier SNPs: (i) those with a *q*-value of greater than 0.1 [46] and (ii) those with scores ± 3 s.d. from the mean axis score for each of the first two constrained axes that also had a $p \leq 0.0001$ when regressed against an environmental variable [47]. Both methods identified similar sets of outlier SNPs. However, the second method was more stringent, so we used only those SNPs for downstream analyses. The mean outlier allele frequencies were calculated for each sampling site, polarized so that the Japanese allele frequency was highest. Because one pair of individuals appeared highly related, we also conducted all outlier analyses without one of the highly related individuals (N4).

## (e) Population structure

We analysed population structure with principal component analysis (PCA) and Structure [48]. PCA was performed with all 4212 SNPs using Plink v. 1.9 [49]. In addition, SNPs out of the Hardy–Weinberg proportions (HWP) were identified using VCFtools and PCA was redone without these SNPs. We also performed PCAs with only outlier SNPs and with all outlier SNPs removed. Structure v. 2.3.4 was run assuming admixture and correlated allele frequencies. We ran five replicates of each *K* (number of populations) from 1 to 5 with a burn-in of 100 000 followed by an additional 10 000 MCMC steps. Structure was run on the same four datasets as PCA. The optimal value of *K* was identified using the Evanno method [50]. Results were visualized with Clumpp [51] and the *pophelper* v. 2.3.0 package in R [52]. Finally, pairwise $F_{ST}$ estimates calculated with all 4212 SNPs and 95% CIs were evaluated using the *hierfstat* package in R [36]. We also conducted all analyses without one of the highly related individuals (N4), though we note that removing related individuals can bias inference [53].

## (f) Demographic analyses

We estimated dispersal rates and long-term effective population sizes ($N_e$) with *fastsimcoal2* [54] by fitting a model of population splits and ongoing migration against the multidimensional site frequency spectrum (details in the electronic supplementary

**Table 1.** Mean per-site nucleotide diversity ($\pi$), the inbreeding coefficient ($F_{IS}$), mean within-population pairwise relatedness ($r$), mean Tajima's $D$ (with either all transcripts included or only transcripts with outlier SNPs) and effective population sizes ($N_e$) from *fastsimcoal2* for each sampling location. (The 95% confidence intervals are provided in brackets. s.e. provided after ± for Tajima's $D$.)

| sampling location | $\pi$ ($\times 10^{-4}$) | $F_{IS}$ | $r$ | Tajima's $D$ (all) | Tajima's $D$ (outliers) | $N_e$ |
|---|---|---|---|---|---|---|
| Japan | 8.44 (8.22, 8.67) | −0.189 (−0.201, −0.178) | 0.222 (0.202, 0.243) | −0.28 ± 0.052 | 0.031 ± 0.255 | 2176 (1773, 3117) |
| Philippines | 9.64 (9.45, 9.84) | −0.057 (−0.069, −0.046) | −0.011 (−0.063, 0.04) | −0.29 ± 0.046 | 0.066 ± 0.244 | 2090 (1931, 3527) |
| Indonesia | 9.53 (9.30, 9.73) | −0.101 (−0.113, −0.087) | 0.018 (−0.061, 0.097) | −0.166 ± 0.05 | 0.006 ± 0.244 | 1988 (1552, 3272) |

material). We also ran STAIRWAY PLOT v. 2 [55] to estimate changes in abundance in each location from the folded site frequency spectrum (see the electronic supplementary material), though we acknowledge challenges with inference based on less than $10^6$ SNPs [56].

### (g) Functional and structural annotation

The de novo transcriptome assembly was mapped to *Amphiprion frenatus* with BLASTn searches against *A. frenatus* with an *E*-value cut-off of $10^{-6}$ [57,58]. Structural annotation was performed on these mapped SNPs with SNPEFF [41]. *Amphiprion frenatus* is one of the most closely related and completely annotated transcriptomes available [59]. Gene ontology (GO) terms were annotated to predicted proteins from the best BLASTn match against the SwissProt database.

## 3. Results

### (a) Genetic diversity and relatedness

Per-site nucleotide diversity and the inbreeding coefficient ($F_{IS}$) were both lowest in Japan and highest in the Philippines (table 1). The mean within-site Tajima's $D$ ranged from −0.29 ± 0.046 in the Philippines to −0.166 ± 0.05 in Indonesia and was −0.355 ± 0.041 with all individuals pooled together (table 1; electronic supplementary material, figure, S1). The mean pairwise relatedness ($r$) varied as well, from $r = 0.222$ in Japan to $r = -0.011$ and $r = 0.018$ in the Philippines and Indonesia, respectively (table 1).

### (b) Spatially divergent selection

BAYPASS identified 93 highly diverged SNPs with an XtX value ≥ 6.03, the 99% significance threshold. BAYPASS also revealed 192 SNPs with a strong association with at least one environmental variable (BF > 20 dB). Most SNPs were associated with SST mean. Of those 108 SNPs, 81% were also associated with SST min and latitude. Most of these latter SNPs were more strongly associated with either SST variable than with latitude (electronic supplementary material, figure S2). The empirical cumulative distribution of BFs for each environmental covariate was significantly different from the permuted distribution (Mann–Whitney $U$-test; $p < 0.001$; electronic supplementary material, figure S3). RDA identified 67 SNPs with a significant association with at least one environmental variable ($p \leq 0.0001$) (electronic supplementary material, figure S4). Of these, most were associated with SST min (electronic supplementary material, table S3).

Across all analyses, 56 SNPs had an XtX greater than the 99% threshold and a significant association with at least one environmental variable according to both EAAs (figure 2). Of these, most were associated with SST mean (54 SNPs), SST min (52 SNPs) or latitude (49 SNPs) (electronic supplementary material, table S3). Again, most of these SNPs (80%) were more strongly associated with either temperature variable than with latitude (electronic supplementary material, table S3). The mean polarized outlier allele frequency was 0.847 in Japan, 0.179 in the Philippines and 0.162 in Indonesia (electronic supplementary material, figure S5). When within-site Tajima's $D$ was calculated for only the outlier sequences (transcripts that contained at least one outlier SNP), the mean estimate trended slightly positive (table 1; electronic supplementary material, figure S1). However, the difference between the outlier-only and overall Tajima's $D$ distributions was not significant for any combination of individuals (electronic supplementary material, figure S6). The outlier analyses after removing related individual N4 did not differ substantially from the findings with all individuals included (electronic supplementary material, table S4).

### (c) Population structure

PCA revealed that individuals from Japan clustered more tightly relative to the Philippines or Indonesia (figure 3*a*; electronic supplementary material, figure S7*a*). PC 1 explained 13% of the total variance, while PC 2 explained 10%. However, the PCA with only outlier SNPs revealed that Japanese individuals were much more diverged from an Indonesian and Philippines cluster (figure 3*b*). PC 1 explained 73% of the variance in this latter case, while PC 2 explained 7%. STRUCTURE analyses also revealed population clustering. The Evanno method suggested three clusters ($K = 3$), regardless of whether the full dataset was used, only SNPs in HWP, or only non-outlier SNPs (figure 3*c*; electronic supplementary material, figure S7*c*,*d*). However, with only outlier SNPs, only two clusters were suggested ($K = 2$), one for Japan and one for the Philippines and Indonesia combined (figure 3*d*). Pairwise $F_{ST}$ ranged from 0.0247 to 0.0767 and was highest for the Japan–Indonesia comparison, congruent with patterns of isolation-by-distance (electronic supplementary material, table S5). Population structure analyses without one of the highly related individuals (N4) did not differ substantially (electronic supplementary material, figure S8).

### (d) Demographic analyses

Analyses with *fastsimcoal2* revealed moderate Japan–Philippines migration rates of 0.0038 (95% CI: 0.0025, 0.0047) and 0.0056 (95% CI: 0.0038, 0.0072) for Philippines–Indonesia, or the equivalent of approximately 10 individuals per generation. Long-term $N_e$ estimates did not differ substantially

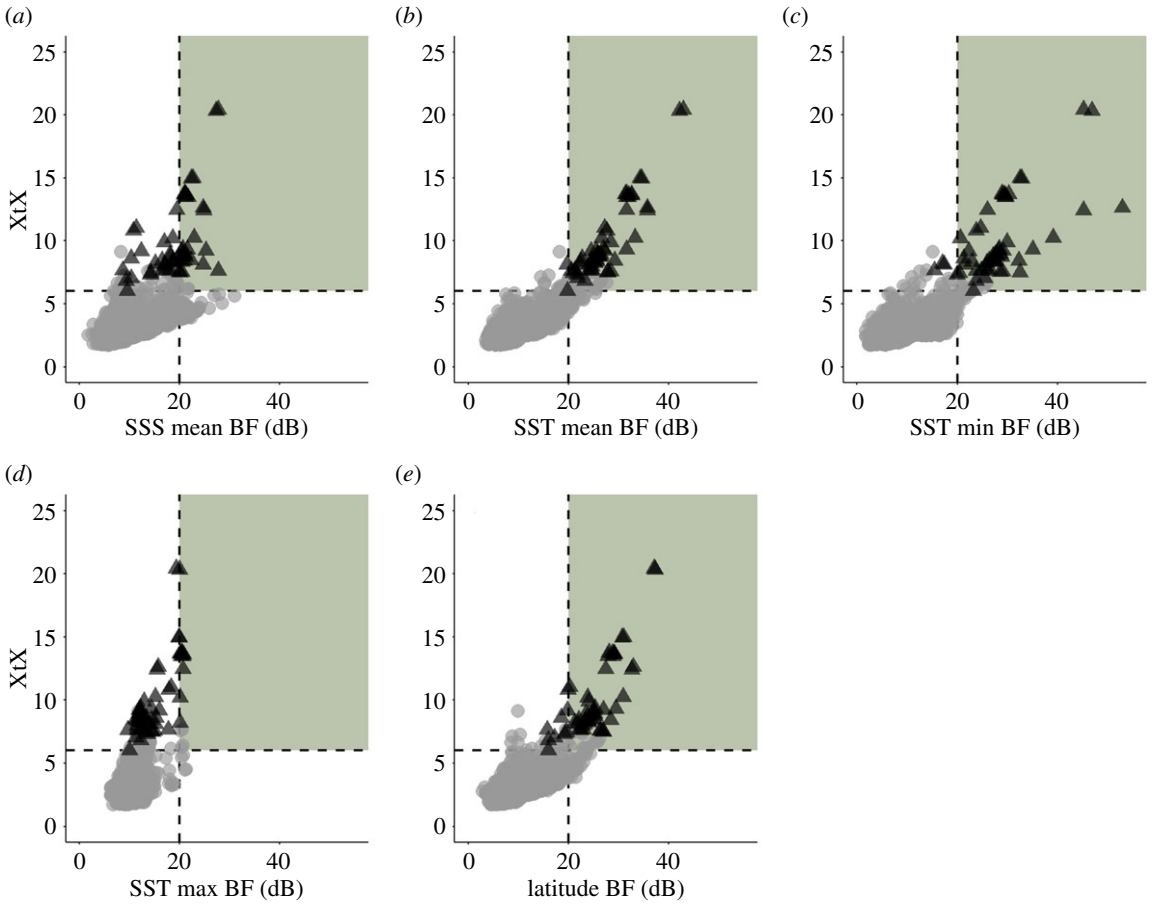


**Figure 2.** Relationships between Bayes factors (BFs) and corresponding XtX values for each SNP–covariate combination (*a*: sea surface salinity (SSS) mean, *b*: sea surface temperature (SST) mean, *c*: SST min, *d*: SST max, *e*: latitude). Lines drawn at BF of 20 deciban units (dB) represent the significance threshold for association with the given covariate. Lines drawn at XtX of 6.03 represent the significance threshold for adaptive divergence among populations. Black triangles represent the 56 candidate SNPs, while grey circles represent the remaining SNPs. (Online version in colour.)

between the three localities (table 1). STAIRWAY PLOT analyses suggested slow declines in the Philippines but did not reveal recent bottlenecks or expansions (electronic supplementary material, figure S9).

### (e) Functional and structural annotation

Of the 4770 SNPs that could be mapped to the *A. frenatus* assembly, most mapped to either coding regions (2418 SNPs) or untranslated regions (UTRs) (1406 SNPs) (electronic supplementary material, table S6). Of the 56 outlier SNPs, 48 could be fully annotated, four could only be annotated functionally and two could only be annotated structurally. These 56 outlier SNPs represented 25 distinct candidate genes. We grouped these candidate genes into general biological categories based on their GO annotations (table 2; electronic supplementary material, table S3). Broadly, most of the candidate genes were involved in protein turnover and translation. The structural annotation suggested that 29 of the SNPs were in coding regions (one nonsense, 11 missense, 17 synonymous) and 19 were in the UTRs (17 in 3′ and two in 5′). One SNP was mapped to the upstream region of a gene and one was mapped to an intergenic region, which may represent either a currently unannotated protein coding gene or RNA gene.

## 4. Discussion

Despite theoretical work predicting how and under what conditions populations at range peripheries may undergo local adaptation [5,8,11], empirical support for these models remains scarce, particularly in the marine realm. Here, we investigated how evolutionary forces interact to shape patterns of adaptive potential and enable local adaptation across the northern half of the range of a common coral reef fish. Comparison of 4212 SNPs from three *A. clarkii* populations revealed clear evidence of spatially divergent selection and reduced levels of genetic variation at the northern periphery of this species' range. In addition, there was substantial population structure despite moderate gene flow between the edge and core populations.

Studies of marine species using putatively neutral genetic markers frequently report genetic structure in the Indo-West Pacific [60–62]. However, this study is unique in that multiple SNPs had strong associations with environmental variables. Moreover, when only examining outlier SNPs, the edge versus core population distinction could explain fully 73% of the genetic variation. These results suggest that selection, in addition to neutral processes, shapes patterns of genetic structure in our study system. Combined, our findings further suggest that neutral and adaptive variation are differently partitioned among the three sampling locations, a pattern that is increasingly reported in marine taxa [16].

### (a) Life on the edge

Climate change is driving range expansions in marine ecosystems [20], renewing interest in the balance between gene flow and adaptation at range margins. Here, population structure

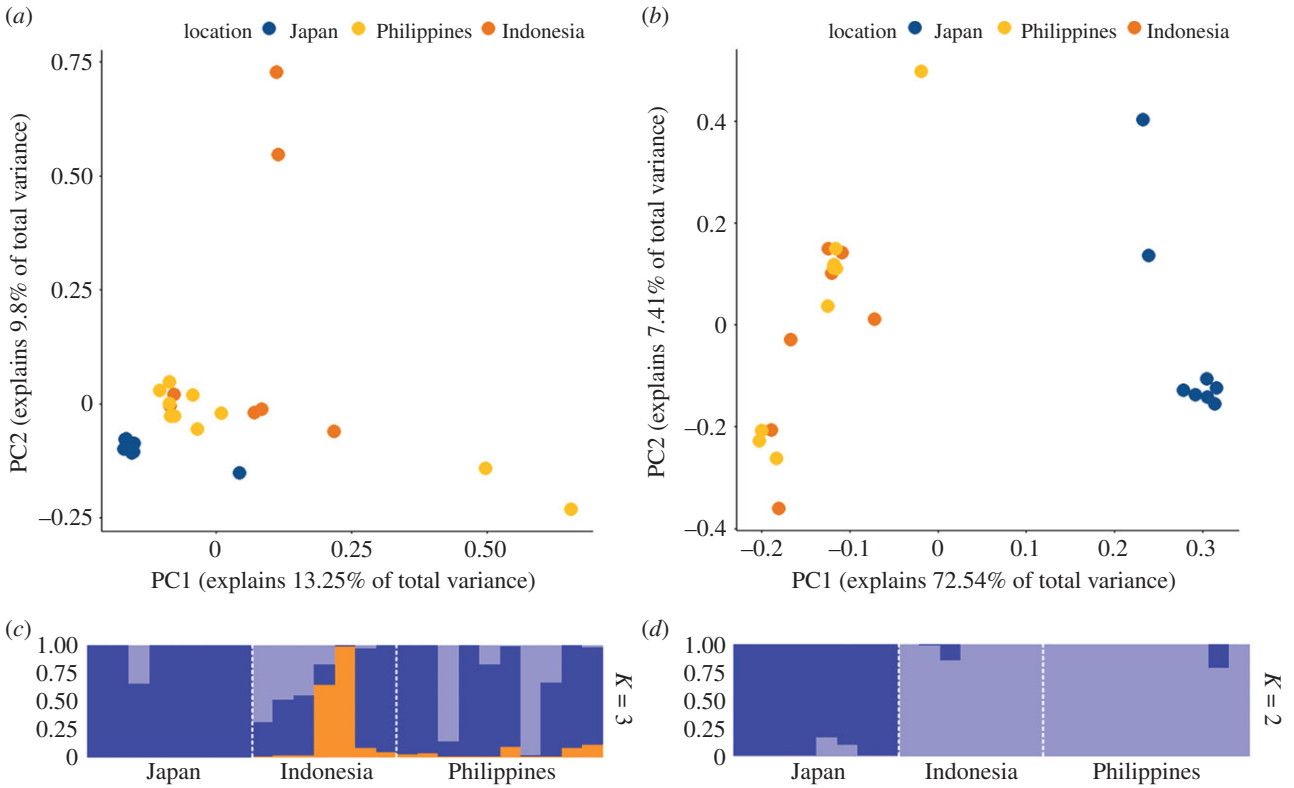


**Figure 3.** (*a*) Results of PCA with all 4212 SNPs. (*b*) Results of PCA with only the 56 outlier SNPs. (*c*) STRUCTURE analysis with all 4212 SNPs ($K = 3$). (*d*) STRUCTURE analysis with only the 56 outlier SNPs ($K = 2$). For both (*c*,*d*), each putative population is represented by a unique colour, and individuals are grouped by sampling sites (divided by the dashed vertical lines). The proportion of assignment is represented by the vertical axis on the left. (Online version in colour.)

analyses provide evidence for at least two genetic clusters, with Japanese individuals at the edge appearing genetically distinct from individuals closer to the core. Clownfish have a relatively short pelagic larval duration (approx. two weeks) [26] and exhibit self-recruitment [63]. These characteristics should limit dispersal, leading to genetic structure as observed in other clownfish species in the region [60]. Given the large physical distance between our sites, a stepping-stone model of gene flow probably explains the observed genetic patterns.

However, a lack of apparent population structure is not unusual in marine populations that span large geographical ranges [14]. Theory also suggests that neutral genetic variation can become homogenized with even minimal gene flow, including rates lower than what we estimated [64]. The Kuroshio Current runs northwards to Japan from the bifurcation of the Northern Equatorial Current off the eastern coast of the Philippines [65] and may provide an avenue for migration, particularly from the core to the edge. Indeed, other damselfish show limited genetic differentiation between Japanese populations and those in the Coral Triangle [66]. However, the strong genetic differentiation seen between *A. clarkii* populations in Japan and the Philippines/Indonesia suggests these distinct genetic lineages may be maintained by local adaptation in addition to gene flow.

Evidence for the role of neutral processes comes from lower genetic diversity at the range edge of *A. clarkii* as well. Nucleotide diversity and $F_{IS}$ were lowest, and relatedness highest, in the peripheral Japanese site. These findings are congruent with the idea that edge populations are subject to higher rates of genetic drift owing to reduced $N_e$ [6,10] and match previous studies that found declining genetic diversity along the Kuroshio Current towards species' northern range margins [67]. However, our analyses also estimated similar long-term $N_e$ in all three sites and no signatures of recent bottlenecks at the range edge. These results suggest the greater levels of diversity observed in the core may be maintained by higher connectivity to other populations in the region, including those we did not sample. Continued gene flow could provide a steady influx of alleles to offset the effects of drift [68]. Populations at the range edge, with fewer connections, may only see reduced benefits of dispersal.

Despite the action of neutral processes, our results also provide strong evidence of spatially divergent selection across the northern half of the range of *A. clarkii*. Fifty-six SNPs were adaptively divergent and significantly associated with environmental variables. Surprisingly, within-population Tajima's *D* for transcripts containing these SNPs was slightly positive, counter to the negative values expected following recent, hard selective sweeps. Positive Tajima's *D* values indicate an over-abundance of intermediate frequency alleles, which are often a signature of balancing selection [69]. While it is unlikely that balancing selection acted separately within each population examined here, balanced polymorphisms may nonetheless be maintained by the push of locally directional selection and the pull of gene flow continuously re-introducing maladaptive alleles and their linked genetic background.

Most of the outlier SNPs were significantly associated with either mean or min SST and latitude. As temperature and latitude are correlated, these findings are not unexpected. However, almost every SNP was more strongly associated with temperature than with latitude, which suggests that temperature, in particular, may be driving the observed differences in allele frequencies. In fact, the pattern of outlier allele frequencies among the three sites closely mirrored

**Table 2.** List of the contigs (from the de novo transcriptome assembly) containing the 56 candidate SNPs and their functional and structural annotations. (Number of candidate SNPs in each contig is listed.)

| contig | no. of SNPs | gene name | GO annotation | SNP effects |
|---|---|---|---|---|
| DN14997_c0_g1_i1 | 3 | MCL1 | apoptotic process | 5′ UTR (1); missense (2) |
| DN2025_c0_g1_i1 | 2 | KRT8 | cell structure | 3′ UTR (1); unannotated (1) |
| DN20310_c1_g1_i1 | 1 | cstb | proteolysis | unannotated (1) |
| DN20701_c0-g1_i1 | 1 | smdt1 | calcium ion transport | synonymous (1) |
| DN22229_c0_g1_i1 | 1 | ATP5H | ATP biosynthetic process | synonymous (1) |
| DN24358_c0_g1_i1 | 1 | Arl6ip1 | protein targeting; cell death | synonymous (1) |
| DN27846_c0_g1_i1 | 3 | KRT8 | cell structure | 3′ UTR (2); unannotated (1) |
| DN28343_c0_g1_i1 | 2 | CRIP1 | signal transduction; protein binding | synonymous (2) |
| DN30912_c0_g1_i1 | 1 | KRT13 | cell structure | unannotated (1) |
| DN33929_c0_g3_i1 | 1 | PSMD12 | proteolysis | 3′ UTR (1) |
| DN34728_c1_g1_i1 | 1 | KRT8 | cell structure | 3′ UTR (1) |
| DN35673_c1_g2_i2 | 2 | rpl9 | translation | missense (1); synonymous (1) |
| DN35709_c0_g2_i1 | 2 | KRT8 | cell structure | missense (1); synonymous (1) |
| DN35710_c0_g1_i1 | 1 | SPCS3 | proteolysis; protein targeting | missense (1) |
| DN36584_c0_g1_i1 | 2 | ANXA5 | calcium ion transport | missense (1); synonymous (1) |
| DN36805_c1_g1_i1 | 3 | PRDX1 | antioxidant activity; response to stress | missense (2); 3′ UTR (1) |
| DN37204_c0_g1_i1 | 3 | Tomm20 | protein targeting | synonymous (1); 3′ UTR (2) |
| DN37469_c0_g3_i1 | 1 | rps3a | translation | synonymous (1) |
| DN37870_c0_g1_i1 | 1 | CHCHD10 | metabolic process; mitochondrial organization | synonymous (1) |
| DN38348_c0_g1_i1 | 1 | CCT5 | protein folding | missense (1) |
| DN38348_c0_g2_i1 | 3 | CCT5 | protein folding | missense (1); synonymous (1); 3′ UTR (1) |
| DN38750_c0_g1_i1 | 1 | GOT2 | metabolic process | synonymous (1) |
| DN39050_c0_g3_i1 | 5 | CCT4 | protein folding | 5′ UTR (1); start lost (1); missense (1); synonymous (2) |
| DN39195_c1_g1_i2 | 1 | YWHAB | signal transduction; protein binding | 3′ UTR (1) |
| DN39216_c0_g1_i1 | 3 | Rab1A | autophagy; protein targeting | 3′ UTR (3) |
| DN39927_c1_g1_i1 | 1 | HMGB2 | transcription; immune response | 3′ UTR (1) |
| DN40382_c0_g2_i1 | 1 | P4hb | isomerase activity; protein folding | 3′ UTR (1) |
| DN40393_c0_g2_i1 | 1 | — | unannotated | upstream (1) |
| DN40479_c0_g2_i1 | 2 | EIF3B | translation | 3′ UTR (2) |
| DN40807_c0_g1_i1 | 3 | — | unannotated | intergenic region (1); unannotated (2) |
| DN4487_c0_g2_i1 | 1 | HMGB2 | transcription; immune response | synonymous (1) |
| DN58176_c0_g1_i1 | 1 | RPL14 | translation | synonymous (1) |

those of average SST temperatures. Similarly, most SNPs associated with the mean SST were also associated with min SST, suggesting that differences in the mean SST are probably reflective of differences in the degree of seasonality (namely, whether a population experiences a winter season). As many biological pathways are temperature-sensitive [22], and winter water temperatures at the Japanese site can regularly drop below 15°C [70], changing seasonality across *A. clarkii*'s range is probably a strong force driving local adaptation.

The correlative nature of our evidence for thermal adaptation cannot rule out alternative selective pressures, however. We did not examine other factors like primary productivity, dissolved oxygen, carbonate chemistry, population densities or resource availability that are also likely to differ between these locations and may drive local adaptation. Similarly, demographic processes like allele surfing during expansion waves and isolation-by-distance may also contribute to the genetic differentiation across *A. clarkii*'s range [71]. However, the outlier detection methods we used are reasonably effective at accounting for these latter sources of evolutionary non-independence [72]; thus, we do not expect drift to be the primary force creating these outlier SNPs.

Further evidence that the allele frequency variation in outlier SNPs is linked to differences in thermal regimes comes from the fact that the genetic pathways represented by these SNPs are similar to those invoked during gene expression changes that accompany acclimation to cold shock. Genes and proteins that are important for temperature acclimation probably also play a large role in temperature

adaptation [73], and differences in gene expression among populations may have a genetic basis in addition to being a plastic response. The candidate genes we identified are largely involved in energy metabolism, protein turnover, cell structure, cell death and oxidative stress response; functional categories that are often upregulated in heart tissue during short-term acclimation to cold stress [74,75]. Cytoskeleton reorganization has been shown to occur during thermal acclimation as well, although this has been at least partially attributed to cold-induced hypertrophy of heart tissue [75]. Interestingly, many candidate SNPs mapped to the 3′ or 5′ UTR (including near CpG islands), regions that help regulate gene expression [76]. Shifting gene expression levels are one of the ways organisms can plastically acclimate to environmental stressors; thus, SNPs in regulatory regions may provide a link between plastic and evolutionary responses involved in thermal adaptation [77]. In addition, most of the candidate SNPs that mapped to coding regions appear to be synonymous substitutions. Recent theory has proposed that silent mutations may not be truly neutral and can undergo weak selection, via codon bias, linkage and translation efficiency [78]. Altered gene expression and enhanced translation accuracy may also be a more feasible route for adaptation, as opposed to modifications of gene structure or the development of novel proteins. Theory suggests that changes in regulatory regions play a major role in adaptation [76,77], and several studies have linked regulatory region mutations with adaptive traits, including temperature response in *Drosophila melanogaster* [79]. While further work is needed to link genetic variation to phenotypes, our results suggest a similar pattern of adaptation within *A. clarkii* populations in response to differences in thermal environments.

## 5. Conclusion

As the oceans continue to change, marine taxa will face substantial shifts in climatological and ecological parameters. The effects of climate change will differ by population depending on their climatic tolerance and local environment [80]. Here, we show how selection, gene flow and drift combined to shape adaptive variation within an edge population, including variation associated with thermal environments. Edge populations are particularly important in predicting species responses, as they are often the first to start shifting as climates change [19,20]. Continued connectivity to warmer core populations may provide an avenue for these typically cooler-climate demes to access novel genetic variation that will enable adaptation to warming temperatures [12]. However, the outcome of adaptation is highly dependent on the extent to which these evolutionary forces interact in a synergistic or antagonistic manner. Species with both the genetic variation to allow local adaptation and the gene flow to transport such variation to novel environments are likely to have particularly strong abilities to adapt to future climatic conditions.

Ethics. This research was done under sampling permit 160/SIP/FRP/SM/IV2012 from the Indonesian Government and the Ministry for Research and Technology (RISTEK-BRIN). The Municipality of Albuera (Leyte) and the Bureau of Fisheries and Aquatic Resources (BFAR) granted prior informed consent and sampling permits in the Philippines.

Data accessibility. Raw reads can be found in the GenBank short reads archive (BioProject PRJNA699463, SRA accession numbers SRR13627993–13628017). All other data and files associated with this project are available at the Dryad Digital Repository: https://doi.org/10.5061/dryad.5x69p8d30 [81].

Authors' contributions. R.D.C.: conceptualization, formal analysis, investigation, methodology, writing—original draft, writing—review and editing; M.L.A.: formal analysis, writing—review and editing; P.A.: resources, validation, writing—review and editing; A.H.: investigation, data collection, data curation; J.A.H.: formal analysis, investigation, writing—review and editing; H.R.M.: data collection, data curation; P.H.B.: data collection, data curation, funding acquisition; M.L.P.: conceptualization, data collection, funding acquisition, investigation, project administration, supervision, validation, writing—review and editing. All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Competing interests. We declare we have no competing interests.

Funding. This work was supported by a David H. Smith Conservation Research Fellowship from the Society for Conservation Biology, a Rutgers EOAS Fellowship, a Rutgers SEBS Excellence Fellowship and National Science Foundation nos OISE-1743711 and OISE-0730256.

Acknowledgements. We thank Dr Irma Arlyza, the Indonesian Government and the Ministry for Research and Technology for sampling permits in Indonesia. We thank Edwyn Alesna, Erlinda de la Victoria, the Municipality of Albuera (Leyte), and BFAR for prior informed consent and sampling permits in the Philippines. We thank Samantha Cheng and Hayley Nuetzel for help sampling in Indonesia; Kazunori Matsumoto for help in Japan; and Gerry Sucano, Beverlito Montalban, Teresita Idara, Rogello Nicanor, Liza Espinosa, Shem San Jose, Noel Alquino, Carlos Balansuna, Froilan Beñas, Danilo Marine, the Municipality of Albuera and Visayas State University for help in the Philippines. We thank Molly Schumer and Julie Peng for help and advice on laboratory work. We also thank Jeffrey Brown for constructing the sampling map.

## References

1. Hare MP, Avise JC. 1996 Molecular genetic analysis of a stepped multilocus cline in the American oyster (*Crassostrea virginica*). *Evolution* **50**, 2305–2315. (doi:10.111/j.1558-5646.1996.tb03618.x)
2. MacColl ADC. 2011 The ecological causes of evolution. *Trends Ecol. Evol.* **26**, 514–552. (doi:10.1016/j.tree.2011.06.009)
3. Kawecki TJ, Ebert D. 2004 Conceptual issues in local adaptation. *Ecol. Lett.* **7**, 1225–1241. (doi:10.111/j.1461-0248.2004.00684.x)
4. Richardson JL, Urban MC, Bolnick DI, Skelly DK. 2014 Microgeographic adaptation and the spatial scale of evolution. *Trends Ecol. Evol.* **29**, 165–176. (doi:10.1016/j.tree.2014.01.002)
5. Bridle JR, Vines TH. 2006 Limits to evolution at range margins: when and why does adaptation fail? *Trends Ecol. Evol.* **22**, 140–147. (doi:10.1016/j.tree.2006.11.002)
6. Kawecki TJ. 2008 Adaptation to marginal habitats. *Annu. Rev. Ecol. Evol. Syst.* **39**, 321–342. (doi:10.1146/annurev.ecolsys.38.091206.095622)
7. Mayr E. 1963 *Animal species and evolution*. Cambridge, MA: Harvard University Press.
8. García-Ramos G, Kirkpatrick M. 1997 Genetic models of adaptation and gene flow in peripheral populations. *Evolution* **51**, 21–28. (doi:10.1111/j.1558-5646.1997.tb02384.x)
9. Kawecki TJ, Holt RD. 2002 Evolutionary consequences of asymmetrical dispersal

rates. *Am. Nat.* **160**, 333–347. (doi:10.1086/341519)

10. Glémin S, Ronfort J, Bataillon T. 2003 Patterns of inbreeding depression and architecture of the load in subdivided populations. *Genetics* **165**, 2193–2122. (doi:10.1093/genetics/165.4.2193)
11. Eckert CG, Samis KE, Lougheed SC. 2008 Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond. *Mol. Ecol.* **17**, 1170–1188. (doi:/10.1111/j.1365-294X.2007.03659.x)
12. Cure K, Thomas L, Hobbs JA, Fairclough DV, Kennington WJ. 2017 Genomic signatures of local adaptation reveal source-sink dynamics in a high gene flow fish species. *Sci. Rep.* **7**, 8618. (doi:10.1038/s41598-017-09224-y)
13. Waples RS. 1998 Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *Heredity* **89**, 438–450. (doi:10.1093/jhered/89.5.438)
14. Palumbi SR. 1992 Marine speciation on a small planet. *Trends Ecol. Evol.* **7**, 114–118. (doi:10.1016/0169-5347(92)90144-Z)
15. De Wit P, Palumbi SR. 2013 Transcriptome-wide polymorphisms of red abalone (*Haliotis rufescens*) reveal patterns of gene flow and local adaptation. *Mol. Ecol.* **22**, 2884–2897. (doi:10.1111/mec.12081)
16. Hoey JA, Pinsky ML. 2018 Genomic signatures of environmental selection despite near-panmixia in summer flounder. *Evol. Appl.* **11**, 1732–1747. (doi:10.1111/eva.12676)
17. Mach ME, Sbrocco EJ, Hice LA, Duffy TA, Conover DO, Barber PH. 2011 Regional differentiation and post-glacial expansion of the Atlantic silverside, *Menidia*, an annual fish with high dispersal potential. *Mar. Biol.* **158**, 515–530. (doi:10.1007/s00227-010-1577-3)
18. Hamilton AM, Selwyn JD, Hamner RM, Johnson HK, Brown T, Springer SK, Bird CE. 2020 Biogeography of shell morphology in over-exploited shellfish reveals adaptive trade-offs on human-inhabited islands and incipient selectively driven lineage bifurcation. *J. Biogeogr.* **47**, 1494–1509. (doi:10.1111/jbi.13845)
19. Williams JL, Hufbauer RA, Miller TEX. 2019 How evolution modifies the variability of range expansion. *Trends Ecol. Evol.* **34**, 903–913. (doi:10.1016/j.tree.2019.05.012)
20. Pinsky ML, Selden RL, Kitchel ZJ. 2020 Climate-driven shifts in marine species ranges: scaling from organisms to communities. *Annu. Rev. Mar. Sci.* **12**, 153–179. (doi:10.1146/annurev-marine-010419-010916)
21. Bindoff NL *et al*. 2019 Chapter 5: Changing ocean, marine ecosystems, and dependent communities. In *IPCC special report ocean and cryosphere in changing climate* (eds HO Pörtner *et al*.). Cambridge, UK: Cambridge University Press.
22. Somero GN. 2004 Adaptation of enzymes to temperature: searching for basic 'strategies'. *Comp. Biochem. Physiol. B* **139**, 321–333. (doi:10.1016/j.cbpc.2004.05.003)
23. Habary A, Johansen JL, Nay TJ, Steffensen JF, Rummer JL. 2017 Adapt, move or die—how will tropical coral reef fishes cope with ocean warming? *Glob. Change Biol.* **23**, 566–577. (doi:10.1111/gcb.13488)
24. Kavanagh KD, Haugen TO, Gregersen F, Jernvall J, Vøllestad LA. 2010 Contemporary temperature-driven divergence in a Nordic freshwater fish under conditions commonly thought to hinder adaptation. *BMC Evol. Biol.* **10**, 350. (doi:10.1186/1471-2148-10-350)
25. Fautin DG, Allen GR. 1992 *Field guide to anemone fishes and their host sea anemones*. Perth, Australia: Western Australian Museum.
26. Wellington GM, Victor BC. 1989 Planktonic larval duration of one hundred species of Pacific and Atlantic damselfishes (Pomacentridae). *Mar. Biol.* **101**, 557–567. (doi:10.1007/bf00541659)
27. Pinsky ML, Montes HR, Palumbi SR. 2010 Using isolation by distance and effective density to estimate dispersal scales in anemonefish. *Evolution* **64**, 2688–2700. (doi:10.1111/j.1558-5646.2010.01003.x)
28. Wang T, Overgaard J. 2007 The heartbreak of adapting to global warming. *Science* **315**, 49–50. (doi:10.1126/science.1137359)
29. Gordon A. 2011 FASTX-Toolkit. See http://hannonlab.cshl.edu/fastx_toolkit/index.
30. Warren RE, Sutton GG, Jones SJM, Holt RA. 2007 Assembling millions of short DNA sequences using SSAKE. *Bioinformatics* **23**, 500–501. (doi:10.1093/bioinformatics/btl629)
31. Grabherr MG *et al*. 2011 TRINITY: reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nat. Biotechnol.* **29**, 644–652. (doi:10.1038/nbt.1883)
32. Lunter G, Goodson M. 2011 Stampy: a statistical algorithm for sensitive and fast mapping of Illumina sequence reads. *Genome Res.* **21**, 936–939. (doi:10.1101/gr.111120.110)
33. Li H *et al*. 2009 The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079. (doi:10.1093/bioinformatics/btp352)
34. Poplin R *et al.* 2017 Scaling accurate genetic variant discovery to tens of thousands of samples. *bioRxiv*. (doi:10.1101/201178)
35. Danecek P *et al.* 2011 The variant call format and VCFTOOLS. *Bioinformatics* **27**, 2156–2158. (doi:10.1093/bioinformatics/btr330)
36. Goudet J. 2005 HIERFSTAT, a package for R to compute and test hierarchical F-statistics. *Mol. Ecol. Notes* **5**, 184–186. (doi:10.1111/j.1471-8286-2004.00828.x)
37. Core Team R. 2015 *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
38. Pew J, Muir PH, Wang J, Frasier TR. 2014 related: an R package for analysing pairwise relatedness from codominant molecular markers. *Mol. Ecol. Resour.* **15**, 557–561. (doi:10.1111/1755-0998.12323)
39. Wang J. 2002 An estimator for pairwise relatedness using molecular markers. *Genetics* **160**, 1203–1215. (doi:10.1093/genetics/160.3.1203)
40. Wang J. 2017 Estimating pairwise relatedness in a small sample of individuals. *Heredity* **119**, 302–313. (doi:10.1038/hdy.2017.52)
41. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. 2012 A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly* **6**, 80–92. (doi:10.4161/fly.19695)
42. Gautier M. 2015 Genome-wide scan for adaptive divergence and association with population-specific covariates. *Genetics* **201**, 1555–1579. (doi:10.1534/genetics.115.181453)
43. Günther T, Coop G. 2013 Robust identification of local adaptation from allele frequencies. *Genetics* **195**, 205–220. (doi:10.1534/genetics.113.152462)
44. Sbrocco EJ, Barber PH. 2013 MARSPEC: ocean climate layers for marine spatial ecology. *Ecology* **94**, 979. (doi:10.1890/12-1358.1)
45. Okasen J *et al*. 2019 vegan: community ecology package R package version 2.5-5. See https://CRAN.R-project.org/package=vegan.
46. Capblancq T, Luu K, Blum MGB, Bazin E. 2018 Evaluation of redundancy analysis to identify signatures of local adaptation. *Mol. Ecol. Resour.* **18**, 1223–1233. (doi:10.1111/1755-0998.12906)
47. Forester BR, Jones MR, Joost S, Landguth EL, Lasky JR. 2016 Detecting spatial genetic signatures of local adaptation in heterogeneous landscapes. *Mol. Ecol.* **25**, 104–120. (doi:10.1111/mec.13476)
48. Pritchard JK, Stephens M, Donnelly P. 2000 Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959. (doi:10.1093/genetics/155.2.945)
49. Purcell S *et al.* 2007 PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575. (doi:10.1086/519795)
50. Evanno G, Regnaut S, Goudet J. 2005 Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* **17**, 1170–1188. (doi:10.1111/j.1365-294X.2005.02553.x)
51. Jakobsson M, Rosenberg NA. 2007 CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**, 1801–1806. (doi:10.1093/bioinformatics/btm233)
52. Francis RM. 2017 pophelper: an R package and web app to analyse and visualize population structure. *Mol. Ecol. Resour.* **17**, 27–32. (doi:10.1111/1755-0998.12509)
53. Waples RS, Anderson EC. 2017 Purging putative siblings from population genetic data sets: a cautionary view. *Mol. Ecol.* **26**, 1211–1224. (doi:10.1111/mec.14022)
54. Excoffier L, Dunaloup I, Huerta-Sánchez E, Sousa VC, Foll M. 2013 Robust demographic inference from genomic and SNP data. *PLoS Genet.* **9**, e1003905. (doi:10.1371/journal.pgen.1003905)
55. Liu X, Fu Y. 2020 Stairway Plot 2: demographic history inference with folded SNP frequency spectra. *Genome Biol.* **21**, 280. (doi:10.1186/s13059-020-02196-9)

56. Lapierre M, Lambert A, Achez G. 2017 Accuracy of demographic inferences from the site frequency spectrum: the case of the Yoruba population. *Genetics* **206**, 439–449. (doi:10.1534/genetics.116.192708)
57. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990 Basic alignment search tool. *J. Mol. Biol.* **215**, 403–410. (doi:10.1016/S0022-2836(05)80360-2)
58. Marcionetti A, Rossier V, Bertrand JAM, Litsios G, Salamin N. 2018 First draft genome of an iconic clownfish species (*Amphiprion frenatus*). *Mol. Ecol. Resour.* **18**, 1092–1101. (doi:10.1111/1755-0998.12772)
59. Litsios G, Salamin N. 2014 Hybridisation and diversification in the adaptive radiation of clownfishes. *BMC Evol. Biol.* **14**, 212. (doi:10.1186/s12862-014-0245-5)
60. Timm J, Kochzius M. 2008 Geological history and oceanography of the Indo-Malay Archipelago shape the genetic population structure in the false clown anemonefish (*Amphiprion ocellaris*). *Mol. Ecol.* **17**, 3999–4014. (doi:10.1111/j.1365-294X.2008.03881.x)
61. DeBoer TS, Naguit MRA, Erdmann MV, Ablan-Lagman MCA, Ambariyanto A, Carpenter KE, Toha AHA, Barber PH. 2014 Concordant phylogenetic patterns inferred from mitochondrial and microsatellite DNA in the giant clam *Tridacna crocea*. *Bull. Mar. Sci.* **90**, 301–329. (doi:10.5343/bms.2013.1002)
62. Jackson AM, Ambariyanto., Erdmann MV, Toha AHA, Stevens LA, Barber PH. 2014 Phylogeography of commercial tuna and mackerel in the Indonesian Archipelago. *Bull. Mar. Sci.* **90**, 471–492. (doi:10.5443/bms.2012.1097)
63. Berumen ML, Almany GR, Planes S, Jones GP, Saenz-Agudelo P, Thorrold SR. 2012 Persistence of self-recruitment and patterns of larval connectivity in a marine protected area network. *Ecol. Evol.* **2**, 444–453. (doi:10.1002/ece3.208)
64. Slatkin M. 1987 Gene flow and the geographic structure of natural populations. *Science* **236**, 787–792. (doi:10.1126/science.3576198)
65. Toole JM, Millard RC, Wang Z, Pu S. 1990 Observations of the Pacific North Equatorial Current bifurcation at the Philippine coast. *J. Phys. Oceanogr.* **20**, 307–318. (doi:10.1175/1520-0485(1990)020<0307:OOTPNE>2.0.CO;2)
66. Liu SYV, Tuanmu M, Rachmawati R, Mahardika GN, Barber PH. 2019 Integrating phylogeographic and ecological niche approaches to delimitating cryptic lineages in the blue-green damselfish (*Chromis viridis*). *PeerJ* **7**, e7384. (doi:10.7717/peerj.7384)
67. Ackiss AS, Bird CE, Akita Y, Santos MD, Tachihara K, Carpenter KE. 2018 Genetic patterns in peripheral marine populations of the fusilier fish *Caesio cunning* within the Kuroshio Current. *Ecol. Evol.* **8**, 11 875–11 886. (doi:10.1002/ece3.4644)
68. Hare MP, Nunney L, Schwartz MK, Ruzzante DE, Burford MO, Waples RS, Ruegg KC, Palstra FP. 2011 Understanding and estimating effective population size for practical application in marine species management. *Conserv. Biol.* **25**, 438–449. (doi:10.1111/j.1523-1739.2010.01637.x)
69. Kreitman M, Di Rienzo A. 2004 Balancing claims for balancing selection. *Trends Genet.* **20**, 300–304. (doi:10.1016/j.tig.2004.05.002)
70. Ochi H. 1985 Temporal patterns of breeding and larval settlement in a temperate population of the tropical anemonefish, *Amphiprion clarkii*. *Jpn J. Ichthyol.* **32**, 248–257. (doi:10.1007/bf02938453)
71. Excoffier L, Ray N. 2008 Surfing during population expansions promotes genetic revolutions and structuration. *Trends Ecol. Evol.* **23**, 347–351. (doi:10.1016/j.tree.2008.04.004)
72. Lotterhos KE, Whitlock MC. 2014 Evaluation of demographic history and neutral parameterization on the performance of $F_{ST}$ outlier tests. *Mol. Ecol.* **23**, 2178–2192. (doi:10.1111/mec.12725)
73. Chen Z, Farrell AP, Matala M, Hoffman N, Narum SR. 2018 Physiology and genomic signatures of evolutionary thermal adaptation in redband trout from extreme climates. *Evol. Appl.* **11**, 1686–1699. (doi:10.1111/eva.12672)
74. Vornanen M, Hassinen M, Koskinen H, Krasnov A. 2005 Steady-state effects of temperature acclimation on the transcriptome of the rainbow trout heart. *Am. J. Physiol. Reg. I* **289**, R1177–R1184. (doi:10.1152/ajpregu.00157.2005)
75. Jayasundara N, Tomanek L, Dowd WW, Somero GN. 2015 Proteomic analysis of cardiac response to thermal acclimation in the eurythermal goby fish *Gillichthys mirabilis*. *J. Exp. Biol.* **218**, 1359–1372. (doi:10.1242/jeb.118760)
76. Wray GA. 2007 The evolutionary significance of *cis*-regulatory mutations. *Nat. Rev. Genet.* **8**, 206–216. (doi:10.1038/nrg2063)
77. Prud'homme B, Gompel N, Carroll SB. 2007 Emerging principles of regulatory evolution. *Proc. Natl Acad. Sci. USA* **104**, 8605–8612. (doi:10.1073/pnas.0700488104)
78. Hershberg R, Perov DA. 2008 Selection on codon bias. *Annu. Rev. Genet.* **42**, 287–299. (doi:10.1146/annurev.genet.42.110807.091442)
79. Kolaczkowski B, Kern AD, Holloway AK, Begun DJ. 2011 Genomic differentiation between temperate and tropical Australian populations of *Drosophila melanogaster*. *Genetics* **187**, 245–260. (doi:10.1534/genetics.110/123059)
80. Miller AD, Coleman MA, Clark J, Cook R, Naga Z, Doblin MA, Hoffmann AA, Sherman CDH, Bellgrove A. 2020 Local thermal adaptation and limited gene flow constrain future climate responses of a marine ecosystem engineer. *Evol. Appl.* **13**, 918–934. (doi:10.1111/eva.12909)
81. Clark RD, Aardema ML, Andolfatto P, Barber PH, Hattori A, Hoey JA, Montes Jr HR, Pinsky ML. 2021 Data from: Genomic signatures of spatially divergent selection at clownfish range margins. Dryad Digital Repository. (https://doi.org/10.5061/dryad.5x69p8d30)