Adaptive pathways of coral populations on the Great Barrier Reef

Mikhail V. Matz[1*], Eric A. Treml[2], Galina V. Aglyamova[1], Madeleine J. H. van Oppen[2,3] and Line K. Bay[3]

[1] Department of Integrative Biology, University of Texas at Austin, 205 W 24th St. C0990, Austin, Texas 78712, USA
[2] School of BioSciences, University of Melbourne, Victoria 3010 Australia
[3] Australian Institute of Marine Science, QLD, Australia

* Author for correspondence, matz@utexas.edu

Abstract

Reef-building corals are extremely important for maintenance of marine biodiversity and coastal economy and are currently under severe threat from anthropogenic warming. Warming is predicted to drive preferential survival of warm-adapted genotypes that have migrated to cooler locations and result in an overall decline in genetic diversity due to bleaching-related mortality. To quantify these trends, we analyzed five populations of a common coral *Acropora millepora* along the latitudinal extent of the Great Barrier Reef (GBR). Population genomic analysis revealed that most populations were demographically distinct and that migration was indeed preferential southward, from lower (warmer) to higher (cooler) latitudes. However, no recent increase in southward migration was detectable, and inferred migration rates remained closely correlated with predictions of a biophysical model of larval dispersal based on ocean currents. There was also no evidence of recent declines in genetic diversity: populations of *A. millepora* expanded during the period of rapid sea level changes 100-500 thousand years ago and remained relatively stable since then, with the exception of one population suffering from frequent environmental disturbances. A multi-locus adaptation model indicated that standing genetic variation spread across latitudes should be sufficient to fuel continuous adaptation of *A. millepora* metapopulation to warming over the next 100-200 years. Unexpectedly, we found that naturally low heritability of thermal tolerance in reef-building corals due to contribution from horizontally transmitted algal symbionts would facilitate longer metpopulation persistence. Still, our model also predicted increase in severity of local mortality events induced by thermal anomalies.

Significance statement: Can long-lived organisms such as reef-building corals adapt fast enough to keep up with the historically unprecedented rate of sea surface warming? Here we combine population genomics, biophysical modeling, and evolutionary simulations to argue that populations of a common reef-building coral (*Acropora millepora*) spread across latitudes on the Great Barrier Reef should harbor sufficient genetic variation to fuel efficient adaptation to increasing temperature for another century and perhaps longer. Our study underscores the key role of standing genetic variation in the future persistence of coral reefs and calls for novel reef management strategies to facilitate natural adaptation process.

Hot water coral bleaching, caused by global warming, is devastating coral reefs around the world (*1*) but there is room for hope if corals can adapt to increasing temperatures. Many coral species have wide distributions that span environments that differ dramatically in their thermal regimes, demonstrating that efficient thermal adaptation has occurred in the past (*2*). But can coral adaptation keep up with the unprecedentedly rapid current rate of global warming (*3*)? One way for corals to achieve rapid thermal adaptation is through genetic rescue, involving the spread of existing heat tolerance alleles from low-latitude, warm-adapted populations to higher-latitude, warming regions, via larval migration (*4*, *5*). We have previously demonstrated the presence of genetic variants conferring high thermal tolerance in a low-latitude *A. millepora* population (*4*). It can be expected that global warming will cause preferential survival of warm-adapted poleward migrants because they will be following their thermal optimum, whereas individuals migrating in the opposite direction would find themselves in increasingly mismatched environments (Fig. 1 A, B). Another likely population-level effect of recent declines in coral cover (*6*) is a reduction in overall genetic diversity, potentially limiting both the scope and the rate of adaptation.

Here, we test these predictions in *Acropora millepora*, a common reef-building coral from the most ecologically prominent and diverse coral genus in the Indo-Pacific (staghorn corals, *Acropora*). We have analyzed genome-wide genetic variation using 2bRAD (*7*) in five populations of *A. millepora* along the latitudinal range of the GBR (Fig. 1 A). We genotyped 18-28 individuals per population at >98% accuracy and with a >95% genotyping rate. Analysis of population structure based on ~11,500 biallelic SNPs separated by at least by 2,500 bases agreed with previous microsatellites results (*8*, *9*), and revealed very low levels of genetic divergence, with only the Keppel Islands population being potentially different from the others (Fig. 1 D and Fig. S1). We observed increasing genetic divergence with geographical distance ("isolation by distance", Fig. 1 C) that supports population divergence, however, pairwise F_{ST} were small and did not exceed 0.014 even between the southernmost and northernmost populations (Keppel and Wilkie). To gain a deeper insight into coral demography, we used Diffusion Approximation for Demographic Inference (*dadi,* (*10*)) to more rigorously test for population subdivision and infer pairwise migration rates among populations and population sizes. *dadi* is a coalescent-based method that optimizes parameters of a pre-specified demographic model to maximize the likelihood of generating the observed allele frequency spectrum (for two populations it is essentially a two-dimensional histogram of allele frequencies, Fig. S2). Being a likelihood-based method, *dadi* can be used to compare alternative models using likelihood ratio tests and Akaike Information Criterion (AIC).

We used AIC to confirm that our populations are separate demographic units. For each pair of populations we generated 120 bootstrapped datasets by resampling genomic contigs and performed delta-AIC comparison of two demographic models, a split-with-migration model and a no-split model (Fig. S3 B). The split-with-migration model assumed two populations that have split some time *T* in the past, potentially have different sizes *N1* and *N2*, and exchange migrants at different rates (*m12* and *m21*) depending on direction. The no-split model allowed for ancestral population size to change at time *T* but not for a population split, so the experimental data were

modeled as two random samples from the same population of size N. The majority of bootstrap replicates (88-100%) showed AIC advantage of the split-with-migration model for all but one pair of populations (Sudbury-Magnetic, 39% bootstrap support; Fig. S3). This indicates that the populations are demographically distinct despite very low F_{ST}. This result highlights the power of coalescent analysis relative to classical approaches (such as F_{ST}) that assume genetic equilibrium, i.e., that populations have been stable for thousands of generations.

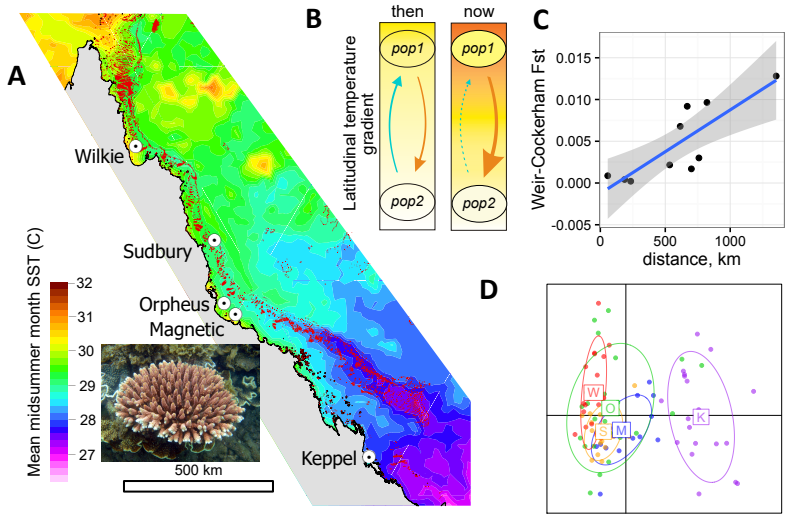

Figure 1. The population setting and background for our study. (A) Locations of sampled populations where mean midsummer month sea surface temperature differed by up to ~3°C. Inset: *Acropora millepora*. (B) Working hypothesis under global warming: Warm-adapted low-latitude genotypes that migrate to higher latitudes would be following their physiological optimum and hence expected to survive better than migrants in the opposite direction. (C) Increase of pairwise F_{ST} with distance, both indicating weak genetic divergence along the GBR, and (D) principal component analysis of genome-wide genetic variation. On panel D, centroid labels are initial letters of population names as in panel A.

We then determined pairwise migration rates from the split-with-migration model and estimated their confidence limits from bootstrap replicates. For all pairwise analyses except Wilkie-Sudbury migration in southward direction exceeded northward migration, and this difference was significant in seven out of nine cases (Fig. 2 A and Fig. S3A). Linear mixed model analysis of direction dependent mean migration rates with a random effect of destination (to account for variation in total migration rate) confirmed the overall significance of this southward trend (P_{MCMC} <1e-4).

It is important to note that our pairwise migration rates captured the cumulative effect of genetic exchange between populations, which included direct migration and the spread of alleles via other stepping-stone populations. Such rates do not directly reflect the numbers of larvae exchanged between populations but are very informative in the genetic rescue context. They represent the per-generation rate of replacement of the destination population genotypes by genotypes from the

source population, which is essentially the rate at which genetic rescue could proceed.

To investigate whether the southward migration bias was due to higher survival of southward migrants relative to northward migrants, as predicted under global warming (Fig. 1 B), we developed a biophysical model of coral larval dispersal on the Great Barrier Reef. This model quantified the per-generation migration potential among coral reef habitat patches in the GBR based on ocean currents and parameters of larval biology (*11*, *12*). We found that the genetic and biophysical migration rates were very closely correlated (Mantel test: r = 0.79, P= 0.008, Fig. 2 C). Although the biophysical model explained most of the southward migration bias in the genetic data, the residuals were still in favor of southward migration (Fig. 2 C, inset; P = 0.058).

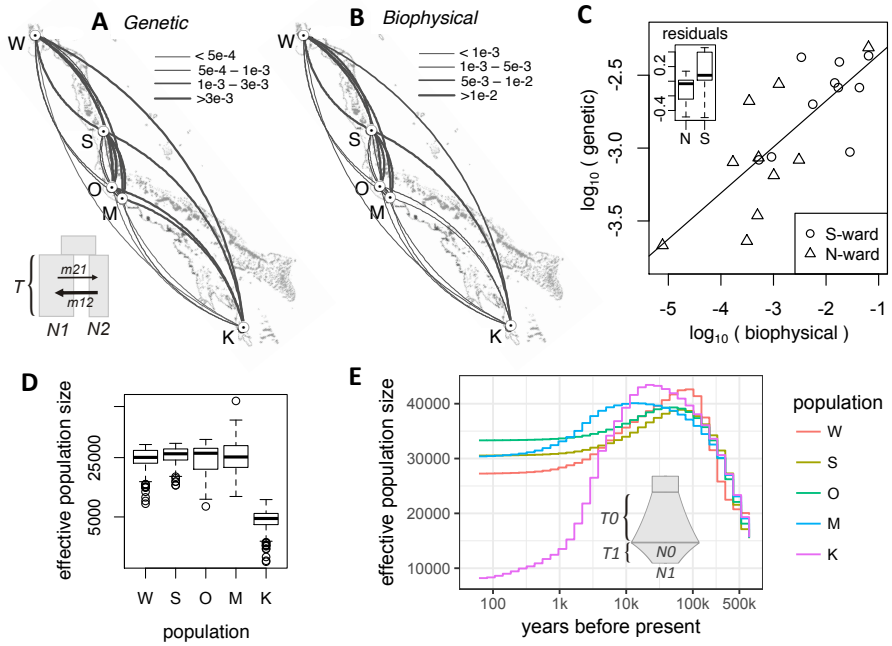

Figure 2. Demography of *A. millepora* populations on the GBR. (A) Arc-plot of migration rates among populations reconstructed from population genetic data. Inset: *dadi* model used: ancestral population splits into two populations of unequal sizes (N1 and N2) some time T in the past, these populations exchange migrants at different rates depending on direction. (B) Migration rates according to the biophysical model. On panels A and B, the arcs should be read clockwise to tell the direction of migration; line thickness is proportional to the migration rate. (C) Correlation between log-transformed biophysical and genetic migration rates (Mantel r = 0.79, P = 0.008). Inset: box-plot of residuals from the linear regression. Southward migration tends to exceed northward migration even after accounting for predictions of the biophysical model (P =0.058), suggesting higher survival of southward migrants. (D) Box plot of effective population sizes inferred by the split-with-migration model (panel A) across all population pairs and bootstrap replicates. (E) Historical changes in effective population sizes inferred using a single-population *dadi* model with two periods of exponential growth (*T0* and *T1*, reaching sizes *N0* and *N1*, inset), averaged across bootstrap replicates.

While this residual excess suggest preferential survival of southward migrants, as predicted by our hypothesis (Fig. 1 B). These genetic predictions represent historical averages since the populations split and did not resolve any potential recent migration changes.

To determine any recent changes in southward migration, we evaluated a similar basic split-with-migration model (Fig. 2A) that allowed for a change in migration over the past 75-100 years. The new model suggested some recent migration changes, but there was no consistent change between northward and southward migration (Fig. S4). Delta-AIC bootstrap analysis favored the new model over the basic one only for two pairs of populations, Wilkie-Orpheus and Wilkie-Magnetic (85 and 60% bootstrap support, respectively). We conclude that with the current data and analysis techniques we cannot yet detect the effect of recent warming on preferential direction of coral migration along the GBR.

The GBR has already warmed by 0.8ºC since the end of last century (*13*) and may have already reduced genetic diversity in *A.millepora* populations. We used *dadi* to infer effective population sizes, which is a measure of genetic diversity and one of the key parameters determining the population's adaptive potential (*14*). The results of the split-with-migration model (Fig. 2 A) were consistent for all population pairs and indicated that Keppel population was about one-fifth the size of others (Fig. 2 D, E). This result was not surprising since the Keppel population frequently suffers high mortality due to environmental disturbances and was therefore is expected to show diminished long-term effective population size (*8*). We also used a single-population *dadi* model that allowed for two consecutive growth/decline periods (Fig. 2 E, inset) to reconstruct effective sizes of individual populations through time (Fig. 2 E and Fig S5). All populations showed evidence of growth prior to the last glaciation, 500-20 thousand years ago (Fig 2 E), which aligned well with the fossil record of rising dominance of *Acropora* corals on Indo-Pacific reefs during this period (*15*). It has been suggested that the fast growth and early sexual maturation of *Acropora* corals gave them an advantage relative to most other reef-building corals during dynamic changes in the reef-forming zone due to the sea level changes accompanying glacial cycles (*15*). Our results suggest that *A. millepora* populations have been in stasis or slow decline since sea level changes abated (Fig. S5), although the inclusion of an additional growth/decline period only improved the model fit significantly for the Keppel population (Fig S6). None of the populations showed evidence of accelerated decline in effective population size over the past few hundred years. Although our samples were collected in the early-mid 2000s, our results are still relevant since they characterize populations only two-three coral generations ago. Disturbances that have affected corals since then would not yet have substantially impacted genetic diversity.

To evaluate whether standing genetic variation contributed by local thermal adaptation could sustain evolution of the *A. millepora* metapopulation in response to warming, we have developed a multi-QTL model of metapopulation adaptation in SLiM (*16*). The model was parameterized with population sizes and migration rates inferred from the genetic analysis (Fig. 2 A, D), and with differences in midsummer monthly mean temperature among populations (Fig. 1 A). The

number of QTLs and their effect sizes, phenotypic plasticity (standard deviation of the Gaussian slope of fitness decline when phenotype mismatches the environment) and heritability (proportion of phenotypic variation attributable to genetics) can all be varied in the model. It can also incorporate climate scenarios with any combination of directional, cyclical and random changes. The model also allows for new mutations but here the new mutation rate was set to zero. This was to assess the contribution of only the standing genetic variation that was introduced into populations at the start of simulation as random QTL effects. The climate scenario started with a pre-adaptation to local thermal conditions for 2,000 generations. Assuming a generation time of of 5 years in *A. millepora* (*17*) this corresponded to the period of stable temperature since the last deglaciation. After pre-adaptation, the temperature was increased at a rate of 0.05^{o}C/generation in all populations, corresponding to the projected 0.1^{o}C warming per decade (*18*). Throughout the simulation temperature was allowed to fluctuate randomly between generations to approximate El Nino Southern Oscillation (ENSO): the temperature deviations were drawn from a normal distribution with a standard deviation of 0.25^{o}C. The size of populations was kept constant throughout the pre-adaptation period and scaled linearly with the populations' relative fitness (mean current fitness divided by the mean fitness at the end of pre-adaptation period) during warming. Migration rates from a population also scaled linearly with the population's fitness. In this way, a population declining in fitness would shrink in size and stop contributing migrants to other populations.

Our model suggested that, with only ten thermal QTLs, under all combinations of heritability and plasticity the pre-adapted metapopulation would be able to persist through the warming for at least 50-100 generations and, in some realistic cases, much longer (Fig. 3 and Figs. S7-S8). Migration in general and southward migration in particular substantially contributed to this persistence (Fig. 3 E, F), underscoring the importance of the spread of warm-adapted genotypes from lower to higher latitudes (*4*).

Predictably, higher phenotypic plasticity promoted population persistence and stability against random thermal anomalies, but we were rather surprised to observe a similar positive effect of lower heritability, set to the values observed in coral quantitative genetics experiments (0.25-0.5, (*17*); Fig 3, Fig. S7). One specific reason why corals are expected to show low heritability of thermal tolerance is that much of natural variation in this trait in corals is due to the type of algal symbionts (*Symbiodinium* spp. (*19*)). Photo-symbionts are not transmitted from parent to offspring in the majority of coral species (*20*), and although host genetics can have some effect on the choice of *Symbiodinium* in the next generation (*21*) environment has a very strong effect on this association (*19*, *22*). Higher persistence under low heritability and high plasticity is most likely explained by the fact that they both allow for higher standing genetic variation to be retained in populations (Fig. S9). During warming, this variation lasts longer as a source of adaptive genetic variants, enabling up to 5°C increase in mean thermal tolerance over 150 generation (Fig. 3 G and Fig. S7). Higher plasticity partially rescued the drop in fitness due to low heritability (Fig. 3 B and D, Fig. S7). Another notable tendency observed with all parameter settings was that during warming the fitness (and hence the size) of adapting populations began to

fluctuate following random thermal anomalies, and the amplitude of these fitness fluctuations increased as the warming progressed even though the amplitude of thermal anomalies did not change (Fig. 3 H). These fluctuations correspond to severe mortality events induced by thermal extremes due to ENSO and affected warm-adapted populations most, which very much resembles the situation currently observed throughout the world (*1*).

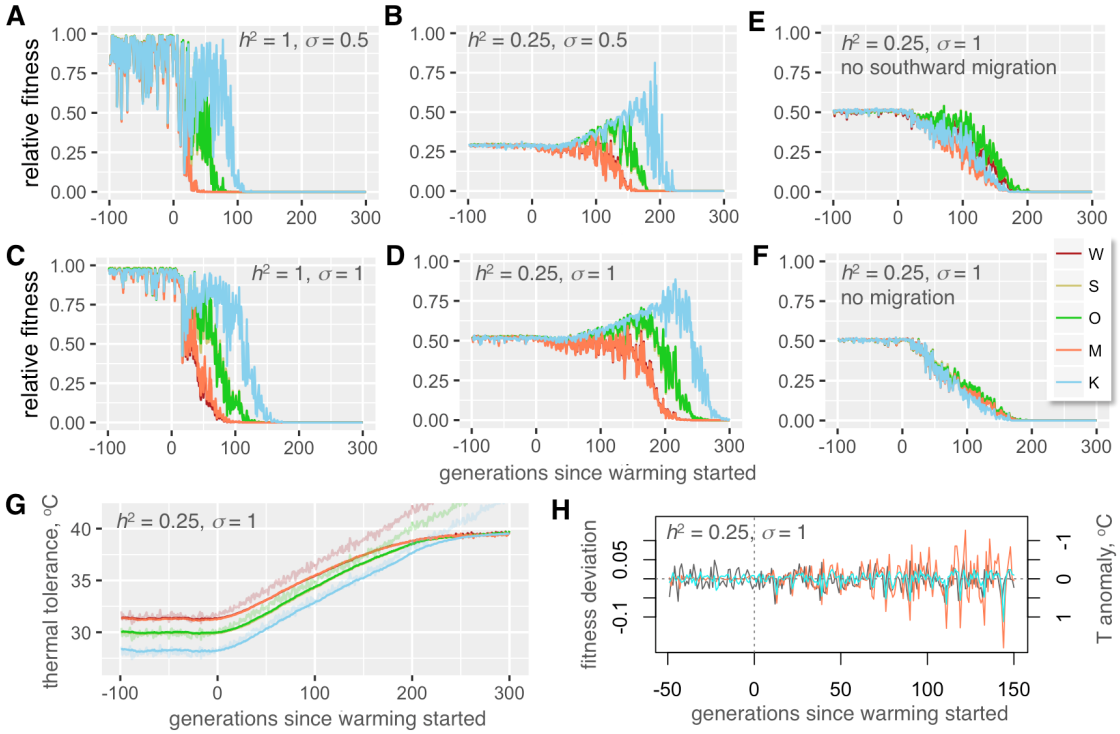

Figure 3. Modeling coral metapopulation persistence under global warming. (A-D) Fitness of modeled populations depending on heritability of thermal tolerance (h^2, proportion of tolerance variation explained by genetics), phenotypic plasticity (σ, standard deviation of the Gaussian slope of fitness decline away from the phenotypic optimum, in degrees C), and presence-absence of migration (E, F). On panels A-F, y-axis is observed fitness relative to maximal fitness at the genetically determined optimum, averaged over all individuals in a population. Warm-adapted populations (W and M) are shown as red-tint traces, populations from mild thermal regime (S and O) are green-tint traces, and the cool-adapted population (K) is the blue trace. Note nearly complete overlap between traces for pairs of populations pre-adapted to the same temperature (W,M and S,O). (G) Thermal tolerances of evolving populations. Thin noisy lines are modeled temperatures at different locations. (H) Modeled random temperature anomalies (grey line) and fluctuations in populations' fitness (the colored lines are residuals from loess regression over fitness traces on panel D; Wilkie: orange line, Keppel: blue line). Note the inverse sign of temperature anomalies: this more clearly shows the correspondence between rise in temperature and drop in fitness in the next generation. As warming progresses, populations (especially originally warm-adapted ones) become increasingly sensitive to random temperature fluctuations.

There are several uncertainties in our model associated with coral biology. Higher number of QTLs and/or their larger effect sizes would promote higher genetic variation and lead to longer

population persistence. To keep the analysis conservative, our model included only ten QTLs, which is likely much fewer that the actual number of thermal QTLs in acroporid corals (*20*). We also kept the distribution of QTL effect sizes narrow: with the current settings and ten QTLs, at the start of simulation only about 2% of corals deviated from the mean thermal tolerance by more than 1.5^{o}C in either direction. Such narrow variation makes adaptation to the thermal gradient of ~3°C along the GBR non-trivial, but still, at present there is no experimental data to evaluate whether even such narrow variation is realistic. Our model was also conservative in using effective population sizes suggested by genetic analysis as census sizes. In highly fecund marine organisms census sizes tend to substantially exceed effective population sizes, sometimes by orders of magnitude (*23*), which would strongly promote higher genetic diversity and population persistence. Moreover, we modeled only our five populations rather than the whole GBR, which would have resulted in much higher standing genetic variation in the metapopulation, promoting longer persistence.

As for phenotypic plasticity, in simulations shown on Fig. 3, $\sigma = 0.5$ and $\sigma = 1$ corresponded to 86% and 40% decline in fitness if the individual's phenotype mismatched the environment by 1°C. The existing data on the issue of coral thermal plasticity are somewhat conflicting. One study shows that acroporid corals can successfully acclimatize to environments differing in maximum temperatures by as much as 2°C (*24*); however, another study found that coral grew 52-80% more slowly when transplanted among locations differing by 1.5ºC average temperature, (*25*). Although it is not possible to directly place these results into our quantitative plasticity framework, the former study supports the higher plasticity setting ($\sigma = 1$) while the latter study supports $\sigma = 0.5$. It must also be noted that both these studies involved *in situ* transplantations and hence the effect of temperature remains confounded with other local fitness-affecting environmental parameters. Also, in adult corals plasticity is likely lower that in larvae and recruits, which are expected to exhibit non-reversible developmental plasticity associated with metamorphosis and establishment within a novel environment (*26*). Future experiments that expose multiple genetically distinct coral individuals to a range of temperatures under controlled laboratory settings are required to rigorously quantify variation in thermal optima and plasticity in natural populations.

In conclusion, we found that genetic diversity and migration patterns of our study species were not yet affected by global warming and were well positioned to facilitate persistence of the GBR metapopulation for a century or more. However, local mortality events induced by thermal anomalies will be increasingly severe, especially among the originally warm-adapted populations. The 10-85% mortality in the Northern GBR as a result of 2016 bleaching event (*27*) could be a particularly sobering recent manifestation of this trend.

More research into phenotypic plasticity and genetic variation in coral thermal tolerance and its genetic architecture (number of QTLs and their effect sizes) is needed to further improve the predictive power of our model. The estimated migration in the order of 10 - 100 migrants per generation could be feasibly facilitated by assisted gene flow efforts (*28*) without risking

disruption of the natural local adaptation patterns (*29*). Corals are declining on reef world-wide and there is an urgent need to develop new solutions to effectively manage the impacts of global processes such as climate change at local management scales. The broad characterization of genetic diversity, local thermal adaptation and migration pathways in multiple reef-building coral species would greatly inform both traditional spatial management and novel assisted gene flow approaches and should therefore be given high priority.

Methods

Genotyping

This study relied predominantly on samples described by van Oppen et al (*9*), with addition of several samples from Orpheus and Keppel islands that were used in the reciprocal transplantation experiment described by Dixon et al (*30*). The samples were genotyped using 2bRAD (*7*) modified for Illumina sequencing platform; the latest laboratory and bioinformatics protocols are available at https://github.com/z0on/2bRAD_GATK. BcgI restriction enzyme was used and the samples retained for this analysis had 2.3-20.2 (median: 7.45) million reads after trimming and quality filtering (no duplicate removal was yet implemented in this 2bRAD version). The reads were mapped to the genome of the outgroup species, *Acropora digitifera* (*31*, *32*), to polarize the allelic states into ancestral (as in *A. digitifera*) and derived, e.g., (*33*, *34*). Genotypes were called using GATK pipeline (*35*).

Preliminary analysis of sample relatedness using vcftools (*36*) revealed that our samples included several clones: four repeats of the same genotype from the Keppel Island (van Oppen et al (*9*) samples K210, K212, K213 and K216), another duplicated genotype from Keppel (samples K211 and K219), and one duplicated genotype from Magnetic Island (samples M16 and M17). All other samples were unrelated. We took advantage of these clonal replicates to extract SNPs that were genotyped with 100% reproducibility across replicates and, in addition, appeared as heterozygotes in at least two replicate pairs (script replicatesMatch.pl with hetPairs=2 option). These 7,904 SNPs were used as "true" SNP dataset to train the error model to recalibrate variant quality scores at the last stage of the GATK pipeline. During recalibration, we used the transition-transversion (Ts/Tv) ratio of 1.438 determined from the "true" SNPs to assess the number of false positives at each filtering threshold (as it is expected that an increase of false positive calls would decrease the Ts/Tv ratio towards unity). We chose the 95% tranche, with novel Ts/Tv = 1.451. After quality filtering that restricted the calls to only bi-allelic polymorphic sites, retained only loci genotyped in 95% or more of all individuals, and removed loci with the fraction of heterozygotes exceeding 0.6 (possible lumped paralogs), we ended up with 25,090 SNPs. In total, 2bRAD tags interrogated 0.18% of the genome. The genotyping accuracy was assessed based on the match between genotyped replicates using script repMatchStats.pl. Overall agreement between replicates was 98.7% or better with the heterozygote discovery rate (fraction of matching heterozygote calls among replicates) exceeding 96%.

Genome-wide genetic divergence

To begin to characterize genome-wide divergence between populations we used pairwise genome-wide Weir and Cockerham's F_{ST} calculated by vcftools (*36*), principal component analysis (PCA) using R package adegenet (*37*), and ADMIXTURE (*38*). For PCA and ADMIXTURE, the data were thinned to keep SNPs separated by 5kb on average and by at least 2.5 kb, choosing SNPs with highest minor allele frequency (script thinner.pl with options 'interval=5000 criterion=maxAF').

Demographic analysis and bootstrapping

Prior to demographic analysis, Bayescan (*39*) was used to identify sites potentially under divergent selection among populations, and 13 such sites with q-value <0.05 were removed. Demographic models were fitted to 120 bootstrapped datasets, which were generated in two stages. First, five alternatively thinned datasets were generated for which SNPs were randomly drawn to be on average 5 kb apart and not closer than 2.5 kb. This time the SNPs were drawn at random to avoid distorting the allele frequency spectrum unlike thinning for PCA and ADMIXTURE where the highest minor allele frequency SNPs were selected. Then, 20 bootstrapped replicates were generated for each thinned dataset by resampling contigs of the reference genome with replacement (script dadiBoot.pl). The fitted model parameters were summarized after excluding bootstrap replicates that fell into the lowest 15% likelihood quantile and the ones where model fitting failed to converge, leading to some parameters being undetermined or at infinity (less than 10% of total number of runs). Delta-AIC values were calculated for each bootstrap replicate that passed these criteria for both compared models, and summarized to obtain bootstrap support value, the percentage of replicates favoring the alternative model. While fitting *dadi* models, the data for each population were projected to sample sizes maximizing the number of segregating sites in the analysis, resulting in 7000-8172 segregating sites per population.

Unit conversion

To convert *dadi*-reported coalescent parameter values (θ, T and M) into time in years (t), effective population sizes in number of individuals (Ne) and migration rates as fraction of new immigrants per generation (m), we estimated the mutation rate (μ) from the time-resolved phylogeny of *Acorpora* genus based on *paxC* intron (*40*), at 4e-9 per base per year. Although *A. millepora* was shown to start reproducing in 3 years (*17*) we assumed the generation time of 5 years reasoning that it would better reflect the attainment of full reproductive potential as the colony grows. Assuming a genome size of 5e+8 bases (*31*) the number of new mutations per genome per generation is 10. Since the fraction 2bRAD-sequenced genome in our experiment was 1.8e-3, the mutation rate per 2bRAD-sequenced genome fraction per generation is $\mu = 0.018$. This value was used to obtain:

- Ancestral effective population size: $Ne = \theta / 2\mu$

- Migration rate: $m = M / 2Ne$
- Time in years: $t = 2TNe \bullet 5$

Biophysical model

A spatially-explicit biophysical modeling framework (*11*, *41*) was used to quantify migration between coral reef habitats of the broader region surrounding the Great Barrier Reef, thereby revealing the location, strength, and structure of a species' potential population connectivity. The model's spatial resolution of ca. 8 km coincides with hydrodynamic data for the broader region (1/12.5 degree; HYCOM+NCODA Reanalysis and Analysis product; hycom.org). Our biophysical dispersal model relies on geographic data describing the seascape environment and biological parameters capturing coral-specific life-histories. Coral reef habitat data are available from the UNEP World Conservation Monitoring Centre (UNEP-WCMC; http://data.unep-wcmc.org/datasets/1) representing a globally-consistent and up-to-date representation of coral reef habitat. To capture specific inter-annual variability, two decades of hydrodynamic data were used from 1992 to 2013 (*42*).

Coral-specific biological parameters for *A. millipora* included relative adult density (dependent on the habitat), reproductive output, larval spawning time and periodicity (e.g., Magnetic Island populations spawn a month earlier than the other GBR sites (*43*)), maximum dispersal duration, pre-competency and competency periods, and larval mortality (*44*, *45*). The spatially explicit dispersal simulations model the dispersal kernel (2-D surface) as a 'cloud' of larvae, allowing it to be concentrated and/or dispersed as defined by the bio-physical parameters. An advection transport algorithm is used for moving larvae within the flow fields (*46*).

Simulations were carried out by releasing a cloud of larvae into the model seascape at all individual coral reef habitat patches and allowing the larvae to be transported downstream by the currents. Ocean current velocities, turbulent diffusion, and larval behavior move the larvae through the seascape at each time-step. Larval competency, behavior, density, and mortality determine when and what proportion of larvae settle in habitat cells at each time step. When larvae encounter habitat, the concentration of larvae settling with the habitat is recorded at that time-step. From the dispersal data, we derived the coral migration matrix representing the proportion of settlers to each destination patch that came from a source patch, which is analogous to the source distribution matrix (*47*) and is equivalent to migration matrices derived from population genetic analysis. It is important to note that migration matrices extracted for the field sites represent the potential migration through all possible stepping-stones.

Metapopulation adaptation model

The model was implemented in SLiM the forward evolutionary simulator, by modifying the provided recipe "Quantitative genetics and phenotypically-based fitness". The model simulates Fisher-Wright populations with discreet generations. At the start of the simulation, populations

were established at specified population sizes and pairwise migration rates (genetic replacement rates), and all QTLs in all individuals were given a mutation with the effect size drawn from a normal distribution with mean zero and specified standard deviation, to create standing genetic variation. The phenotype of each individual was calculated as the sum of QTL effects plus random noise to simulate desired heritability. Then, fitness of each individual was calculated based on the difference between the individual's phenotype (thermal optimum), temperature of the environment, and the setting for phenotypic plasticity, modeled as the standard deviation of the Gaussian slope of fitness decline with increasing distance between phenotype and environment. Then, parents were chosen to produce the next generation according to their fitness; parents for immigrant individuals are chosen from among individuals in the source population. New mutations at QTLs happened at the specified rate when transitioning to the next generation and the effect of a new mutation replaced the previous QTL effect.

Our code was designed for general modeling of multilocus adaptation in metapopulations and can be customized easily. It is possible to adjust:

- Number of populations, their sizes, and pairwise migration rates. We modeled our five populations with effective population sizes and pairwise migration rates inferred by *dadi*.
- Number of QTLs and the distribution of their effect sizes. To keep the model conservative, we modeled only ten QTLs with normal distribution of effect sizes with a standard deviation of 0.2^{o}C. With ten QTLs, this setting implied that at the start of simulation only about 2% of corals deviated from mean thermal tolerance by more than 1.5^{o}C in either direction. Since thermal differences between our populations exceeded 3°C, this narrow variation made local adaptation rather non-trivial.
- Dominance of QTLs (set to 0.5 in our simulation).
- Phenotypic plasticity. We modeled three plasticity settings, 0.5, 1 and 2, which corresponded to 86%, 40% and 13% fitness drop when the individual's phenotypic optimum (calculated based on QTLs and heritability setting) mismatched the environment by 1°C.
- Heritability (proportion of phenotypic variation explained by genetics). We examined values 1, 0.5, 0.25 and 1e-5, the latter to confirm that no adaptation or evolution was observed when the trait was not heritable.
- Mutation rate, which was set to zero because we wanted to explore only the role of standing genetic variation.
- Environmental changes, modeled as identical trends across populations with population-specific offsets. The trends can be any combination of linear, cyclical and random components. In the current simulation, environmental trends were offset by +1.6^{o}C in Wilkie and Magnetic populations and by -1.8^{o}C in the Keppel population, to model differences in midsummer monthly mean temperature among populations (Fig. 1). The trends included only random thermal anomalies (drawn from a normal distribution with a standard deviation of 0.25^{o}C, to approximate ENSO events) for the first 2000 generations, after which a linear increase at 0.05^{o}C per generation was added to simulate warming.

To better model population dynamics, we implemented linear scaling of the population size and immigration rates with the population's mean fitness post pre-adaptation period (the initial 2000 generations with no linear change in environment). In this way, a population declining in fitness shrinks in size and stops contributing migrants to other populations.

All combinations of parameter settings were run ten times to ensure consistency. We found that with population sizes in thousands, such as in our case, the results were very consistent among independent runs. We therefore did not aggregate results over many replicated runs but show one randomly chosen run for each tested parameter combination.

Acknowledgements

We wish to thank Ryan Gutenkunst and Benjamin Haller for their continuous support of *dadi* and SLiM users, respectively. The bioinformatics analysis was accomplished using computational resources of the Texas Advanced Computer Center. This study has been supported by NSF (DEB-1054766) grant to M.V.M, ARC (LP120200245) and University of Melbourne ECR grants to E. A.T., a Coral Reef Alliance grant ("Coral Adaptation Challenge") to E.A.T and M.V.M, Queensland Government funding to L.K.B and AIMS funding to L.K.B. and M.J.V.O

Data and code availability

The finalized genotyping dataset in VCF format, detailed bioinformatic walkthrough, accessory formatting and plotting scripts, *dadi* scripts and the SLiM model code are available from https://github.com/z0on/Adaptive-pathways-of-coral-populations-on-the-Great-Barrier-Reef. Raw sequencing data has been deposited to National Center for Biotechnology Information's Short Reads Archive (accession number pending).

References

1. A. C. Baker, P. W. Glynn, B. Riegl, Climate change and coral reef bleaching: An ecological assessment of long-term impacts, recovery trends and future outlook. *Estuar. Coast. Shelf Sci.* **80**, 435–471 (2008).
2. T. P. Hughes *et al.*, Climate change, human impacts, and the resilience of coral reefs. *Science.* **301**, 929–933 (2003).
3. J. M. Pandolfi, S. R. Connolly, D. J. Marshall, A. L. Cohen, Projecting coral reef futures under global warming and ocean acidification. *Science*. **333**, 418–22 (2011).
4. G. B. B. Dixon *et al.*, Genomic determinants of coral heat tolerance across latitudes. *Science.* **348**, 1460–1462 (2015).
5. P. K. Ingvarsson, Restoration of genetic variation lost – the genetic rescue hypothesis. *Trends Ecol. Evol.* **16**, 62–63 (2001).
6. G. De'ath, K. E. Fabricius, H. Sweatman, M. Puotinen, The 27-year decline of coral cover on the Great Barrier Reef and its causes. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 17995–9 (2012).
7. S. Wang, E. Meyer, J. K. McKay, M. V Matz, 2b-RAD: a simple and flexible method for genome-wide genotyping. *Nat. Methods*. **9**, 808–810 (2012).
8. M. J. H. van Oppen, V. Lukoschek, R. Berkelmans, L. M. Peplow, A. M. Jones, A population genetic assessment of coral recovery on highly disturbed reefs of the Keppel Island archipelago in the southern Great Barrier Reef. *PeerJ*. **3**, e1092 (2015).
9. M. J. H. Van Oppen, L. M. Peplow, S. Kininmonth, R. Berkelmans, Historical and contemporary factors shape the population genetic structure of the broadcast spawning coral, Acropora millepora, on the Great Barrier Reef. *Mol. Ecol.* **20**, 4899–4914 (2011).
10. R. N. Gutenkunst, R. D. Hernandez, S. H. Williamson, C. D. Bustamante, Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genet.* **5**, e1000695 (2009).
11. E. A. Treml, J. Roberts, P. N. Halpin, H. P. Possingham, C. Riginos, The emergent geography of biophysical dispersal barriers across the Indo-West Pacific. *Divers. Distrib.* **21**, 465–476 (2015).
12. E. A. Treml *et al.*, Reproductive Output and Duration of the Pelagic Larval Stage Determine Seascape-Wide Connectivity of Marine Populations. *Integr. Comp. Biol.* **52**, 525–537 (2012).
13. J. Lough, A. Hobday, Observed climate change in Australian marine and freshwater environments. *Mar. Freshw. Res.* **62**, 984–999 (2011).
14. B. Charlesworth, Fundamental concepts in genetics: Effective population size and patterns of molecular evolution and variation. *Nat. Rev. Genet.* **10**, 195–205 (2009).
15. W. Renema *et al.*, Are coral reefs victims of their own past success? *Sci. Adv.* **2**, e1500850 (2016).
16. B. C. Haller, P. W. Messer, SLiM 2: Flexible, interactive forward genetic simulations. *Mol. Biol. Evol.* **34**, 230–240 (2017).
17. M. Vanessa, B. Baria, D. W. Dela Cruz, R. D. Villanueva, J. R. Guest, Spawning of three-year-old Acropora millepora corals reared from larvae in Northern Philippines. *Bull. Mar. Sci.* **88**, 61–62 (2012).
18. IPCC, *Climate Change 2007 - The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the IPCC* (Cambridge University Press, New York, NY, 2007).
19. R. Berkelmans, M. J. H. van Oppen, The role of zooxanthellae in the thermal tolerance of corals: a "nugget of hope" for coral reefs in an era of climate change. *Proc. R. Soc. B-Biological Sci.* **273**, 2305–2312 (2006).
20. A. H. Baird, J. R. Guest, B. L. Willis, Systematic and biogeographical patterns in the reproductive biology of scleractinian corals. *Annu. Rev. Ecol. Evol. Syst.* **40**, 551–571 (2009).
21. K. Quigley, B. Willis, L. Bay, Heritability of the Symbiodinium community in vertically- and horizontally-transmitting broadcast spawning corals. *bioRxiv*. doi: https://doi.org/10.1101/100453 (2017).
22. E. J. Howells *et al.*, Coral thermal tolerance shaped by local adaptation of photosymbionts. *Nat. Clim. Chang.* **2** (2011), pp. 116–120.
23. F. P. Palstra, D. J. Fraser, Effective/census population size ratio estimation: a compendium and

appraisal. *Ecol. Evol.* **2**, 2357–2365 (2012).
24. S. R. Palumbi, D. J. Barshis, N. Traylor-Knowles, R. A. Bay, Mechanisms of reef coral resistance to future climate change. *Science*. **344**, 895–8 (2014).
25. E. J. Howells, R. Berkelmans, M. J. H. van Oppen, B. L. Willis, L. K. Bay, Historical thermal regimes define limits to coral acclimatization. *Ecology*. **94**, 1078–1088 (2013).
26. D. Nettle, M. Bateson, Adaptive developmental plasticity: what is it, how can we recognize it and when can it evolve? *Proc. R. Soc. B Biol. Sci.* **282**, 20151005 (2015).
27. Great Barrier Reef Marine Park Authority, *Interim report on the environmental impacts of the 2016 coral bleaching event*.
28. O. Hoegh-Guldberg *et al.*, Assisted colonization and rapid climate change. *Science.* **321**, 345–346 (2008).
29. S. N. Aitken, M. C. Whitlock, Assisted Gene Flow to Facilitate Local Adaptation to Climate Change. *Annu. Rev. Ecol. Evol. Syst.* **44**, 367–388 (2013).
30. G. B. Dixon, L. K. Bay, M. V. Matz, Bimodal signatures of germline methylation are linked with gene expression plasticity in the coral Acropora millepora. *BMC Genomics*. **15**, 1109 (2014).
31. C. Shinzato *et al.*, Using the Acropora digitifera genome to understand coral responses to environmental change. *Nature*. **476**, 320-U82 (2011).
32. M. J. H. van Oppen, B. J. McDonald, B. Willis, D. J. Miller, The Evolutionary History of the Coral Genus Acropora (Scleractinia, Cnidaria) Based on a Mitochondrial and a Nuclear Marker: Reticulation, Incomplete Lineage Sorting, or Morphological Convergence? *Mol. Biol. Evol.* **18**, 1315–1329 (2001).
33. B. F. Voight, S. Kudaravalli, X. Wen, J. K. Pritchard, A map of recent positive selection in the human genome. *PLoS Biol.* **4**, e72 (2006).
34. I. K. Jordan *et al.*, A universal trend of amino acid gain and loss in protein evolution. *Nature*. **433**, 633–8 (2005).
35. A. Mckenna *et al.*, The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
36. P. Danecek *et al.*, The variant call format and VCFtools. *Bioinformatics*. **27**, 2156–8 (2011).
37. T. Jombart, adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*. **24**, 1403–5 (2008).
38. D. H. Alexander, J. Novembre, K. Lange, Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–64 (2009).
39. T. Günther, G. Coop, Robust identification of local adaptation from allele frequencies. *Genetics*. **195**, 205–20 (2013).
40. Z. T. Richards, D. J. Miller, C. C. Wallace, Molecular phylogenetics of geographically restricted Acropora species: Implications for threatened species conservation. *Mol. Phylogenet. Evol.* **69**, 837–851 (2013).
41. E. A. Treml, P. N. Halpin, D. L. Urban, L. F. Pratson, Modeling population connectivity by ocean currents, a graph-theoretic approach for marine conservation. *Landsc. Ecol.* **23**, 19–36 (2008).
42. E. P. Chassignet *et al.*, The HYCOM (HYbrid Coordinate Ocean Model) data assimilative system. *J. Mar. Syst.* **65**, 60–83 (2007).
43. R. C. Babcock *et al.*, Synchronous Spawnings of 105 Scleractinian Coral Species on the Great-Barrier-Reef. *Mar. Biol.* **90**, 379–394 (1986).
44. S. W. Davies, E. A. Treml, C. D. Kenkel, M. V. Matz, Exploring the role of Micronesian islands in the maintenance of coral genetic diversity in the Pacific Ocean. *Mol. Ecol.* **24**, 70–82 (2015).
45. S. R. Connolly, A. H. Baird, Estimating dispersal potential for marine larvae: dynamic models applied to scleractinian corals. *Ecology*. **91**, 3572–3583 (2010).
46. P. K. Smolarkiewicz, J. Szmelter, An MPDATA-based solver for compressible flows. *Int. J. Numer. Methods Fluids*. **56**, 1529–1534 (2008).
47. R. K. Cowen, G. Gawarkiewicz, J. Pineda, S. Thorrold, F. Werner, Population Connectivity in Marine Systems: An Overview. *Oceanography*. **20**, 14–21 (2007).

Supplemental Figures

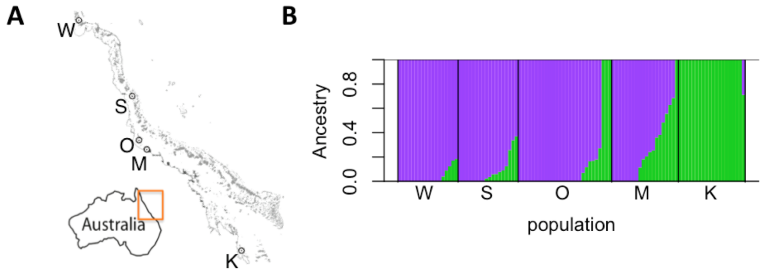

Figure S1. ADMIXTURE analysis of genetic differentiation between populations. (A) Map of sampled locations with one-letter population identifiers. (B) ADMIXTURE plot of ancestry proportions with $K = 2$ (optimal K was 1).

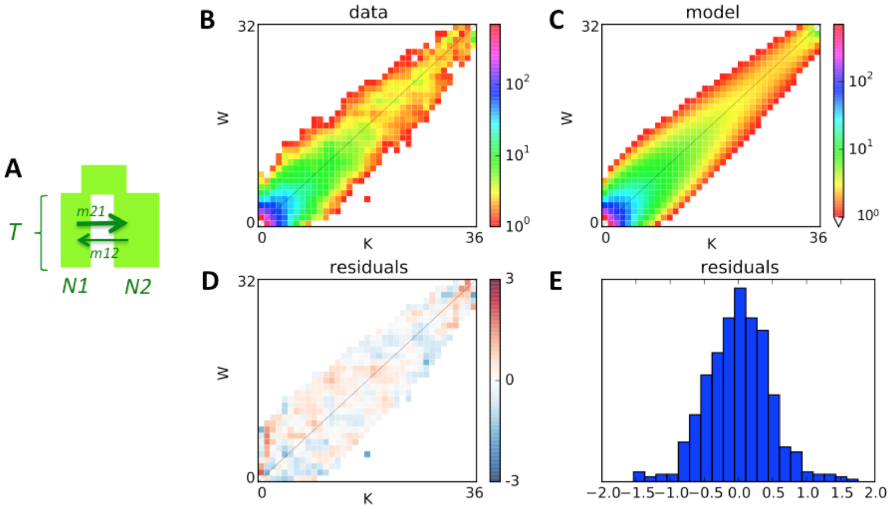

Figure S2. Example of two-population *dadi* model fit. (A) The model: ancestral population splits into two populations of unequal sizes (N1 and N2) some time T in the past, which exchange migrants with different rates depending on direction. (B) Observed allele frequency spectrum comparing Wilkie (W) and Keppel (K) populations. (C) Allele frequency spectrum generated by the fitted model. (D, E) Map and histogram of residuals (absolute scale).

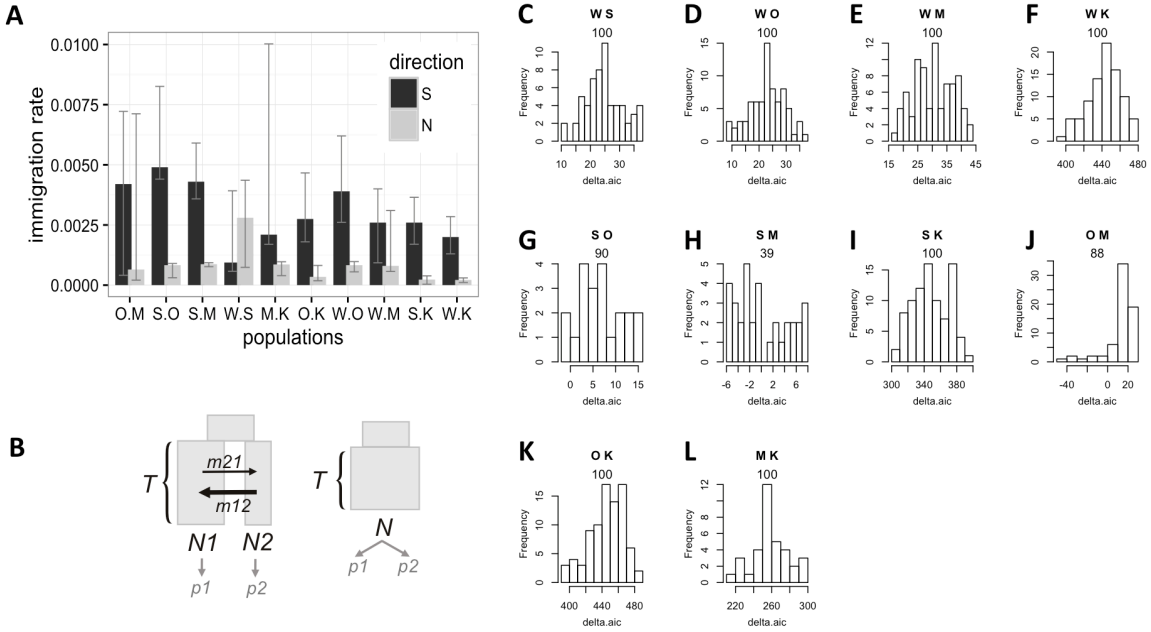

Figure S3. Bootstrap analysis of migration rates and population subdivision using *dadi*. (A) Migration among population pairs, with bootstrap-derived 95% confidence intervals. The pairs are identified on the x-axis and sorted by increasing geographical distance. Black bars – southward migration, grey bars – northward migration. (B) Models being compared: the full model (left) implies populations' split into two different sizes (N1 and N2) at time T in the past, since when they exchanged migrants at unequal rates depending on direction. Reduced model allows for population size change at time T in the past but does not include population split: the two genotyped groups (p1 and p2) are regarded as two samples from the same population. (C-L) Histograms of delta-AIC values for 100 bootstrap replicates (bootstrap was performed over genomic contigs of the draft genome of *A. digitifera*. Positive numbers indicate support for the full model. The letters on top of each panel identify compared populations, the number is the proportion of positive bootstrap replicates (i.e., bootstrap support for the full model). The only comparison that did not receive >50% bootstrap support for population split is between S and M populations (panel H).

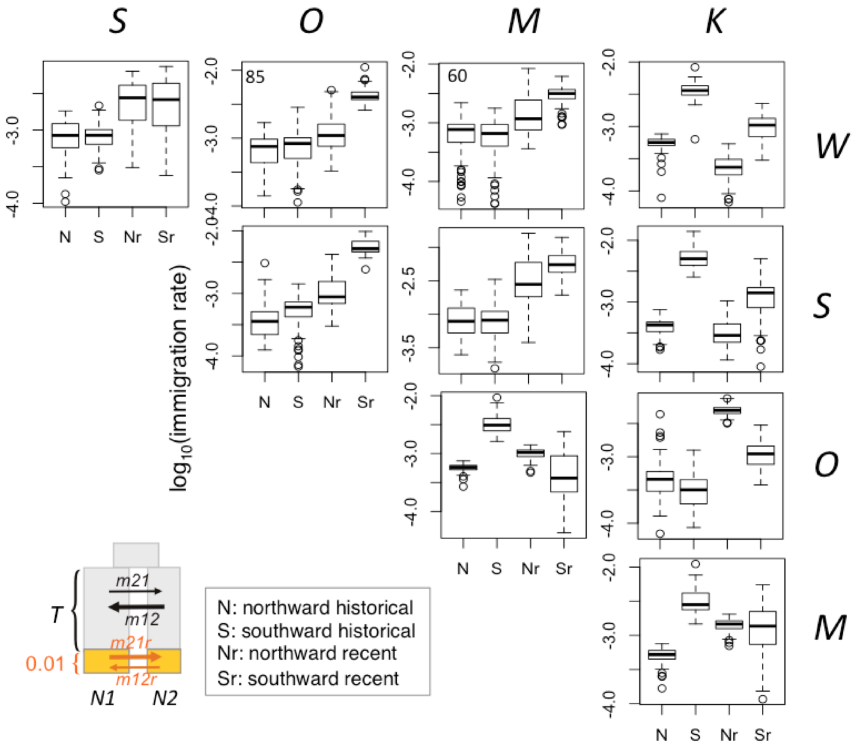

Figure S4. Migration rates inferred by the *dadi* model allowing for the change in migration rates over the last 0.01 T units (15-20 generations or 75-100 years, in our case). Box plots show historical (N, S) and recent (Nr, Sr) migration rates inferred among pairs of population across 100 bootstrap replicates. Numbers in the top left corner of the WO and WM plots are delta-AIC bootstrap support values for the model with the recent change in migration when compared to the split-with-migration model with no recent change (Fig. 1A). All other pairs had less than 50% delta-AIC bootstrap support. There is no consistent recent change in the preferential direction of migration.

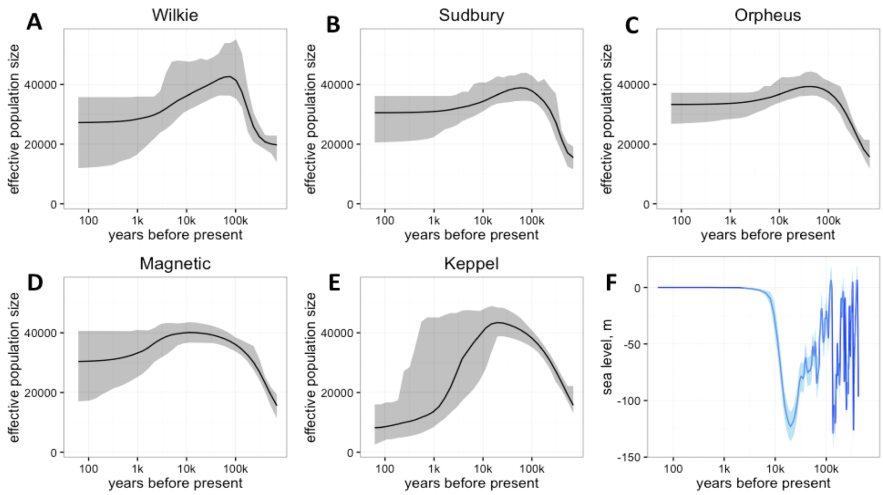

Figure S5. Population history. (A-E) Historical population sizes with bootstrap-derived 95% confidence intervals, according to the two-growth model (Fig. S6 A). (F) Sea level with shaded area corresponding to standard error (*41*).

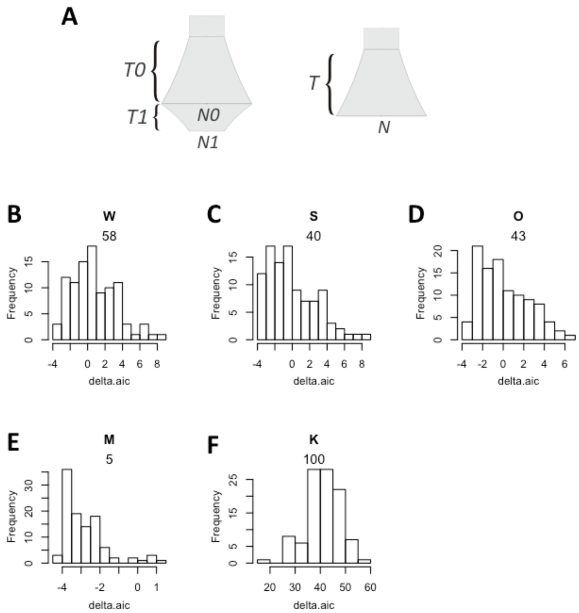

Figure S6. Delta-AIC bootstrap analysis of single-population models. (A) Models compared. The full model (left) includes two exponential growth periods (any of which could be growth or decline), the reduced model (right) has only one growth period. (B-F) Histograms of delta-AIC values for 100 bootstrap replicates. Positive numbers indicate support for the full model. The letter on top of each panel identify the population, the number is the proportion of positive bootstrap replicates (i.e., bootstrap support for the full model). The two-growth model is strongly supported for population K (panel F) and marginally supported for population W (panel B).

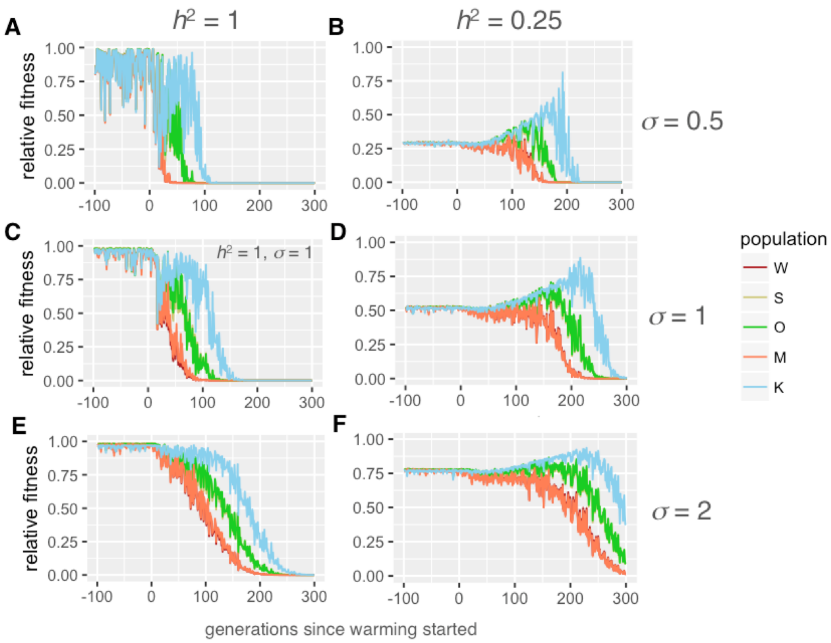

Figure S7. Fitness of modeled populations after pre-adaptation period and under warming, depending on heritability of thermal tolerance (h^2, proportion of phenotypic variation explained by genetics) and phenotypic plasticity (σ, standard deviation of the Gaussian slope of fitness decline away from the phenotypic optimum, in degrees C). X-axis is generations; warming starts at generation 0. Y-axis is fitness relative to maximal fitness at the genetically determined optimum. Warm-adapted populations (W and M) are shown as red-tint traces, populations from mild thermal regime (S and O) are green-tint traces, and the cool-adapted population (K) is the blue trace. Pairs of traces for warm- and mild-adapted populations largely overlap. (A, C, E): h^2=1. (B, D, F): h^2=0.25. (A, B): σ = 0.5. (C, D): σ = 1. (E, F): σ = 2. Higher plasticity facilitates metapopulation persistence during warming and confers stability against random fluctuations. Higher plasticity also partially rescues the drop in fitness achievable under low heritability (compare pre-warming generations, from -100 to 0, on panels B, D and F).

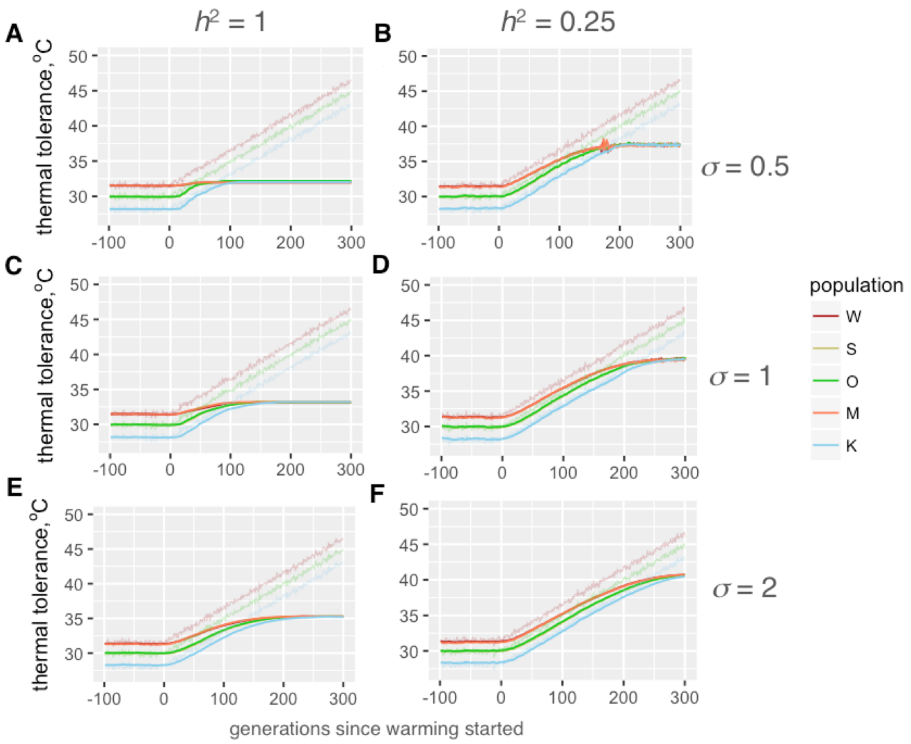

Figure S8. Higher plasticity and lower heritability promote longer and more extensive evolution in response to warming. The graphs show mean thermal tolerance of modeled populations after pre-adaptation period and under warming, depending on heritability of thermal tolerance (h^2, proportion of phenotypic variation explained by genetics) and phenotypic plasticity (σ, standard deviation of the Gaussian slope of fitness decline away from the phenotypic optimum, in degrees C). X-axis is generations; warming starts at generation 0. Y-axis is thermal tolerance (mean phenotype of the population). Warm-adapted populations (W and M) are shown as red-tint traces, populations from mild thermal regime (S and O) are green-tint traces, and the cool-adapted population (K) is the blue trace. Thin noisy lines are modeled temperatures at the corresponding locations. Pairs of traces for warm- and mild-adapted populations largely overlap. (A, C, E): h^2=1. (B, D, F): h^2=0.25. (A, B): σ = 0.5. (C, D): σ = 1. (E, F): σ = 2.

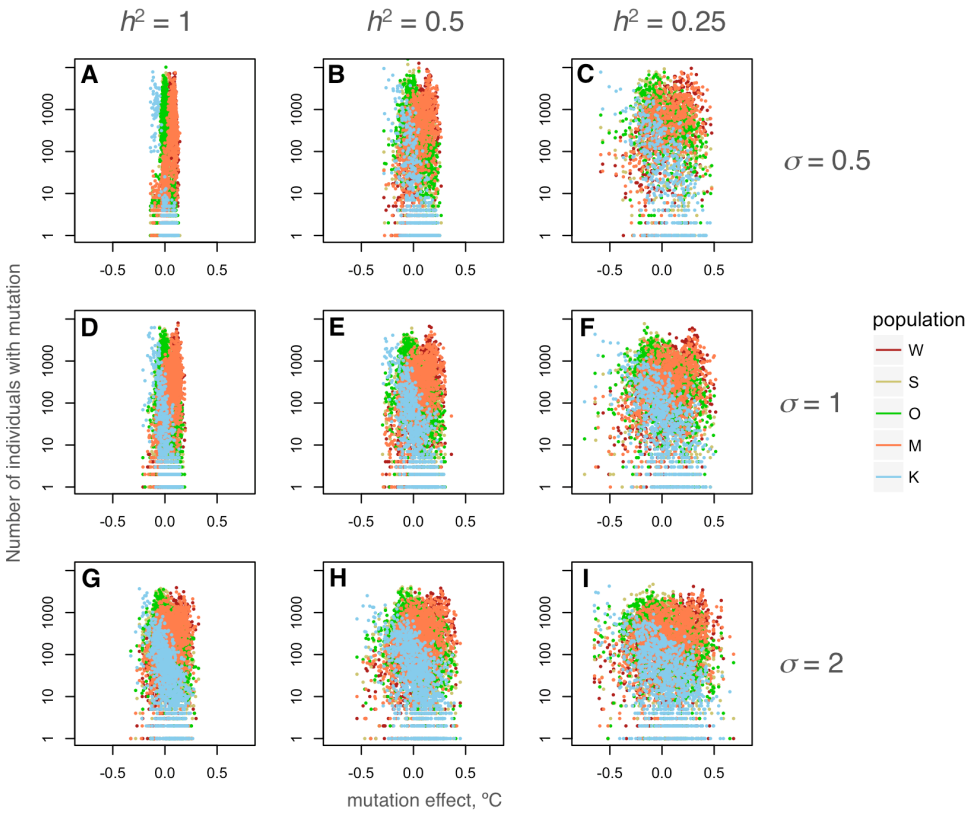

Figure S9. Higher plasticity (σ) and lower heritability (h^2) promote retention of higher genetic variation in thermal tolerance. The scatterplots show the dependence of the number of individuals in a population bearing a mutation at a thermal QTL locus on the mutation's effect size (change in thermal tolerance, in ºC) at the end of the pre-adaptation period (2000 generations with no directional change in temperature). The starting standing genetic variation was the same in all simulations. (A,D,E): h^2=1. (B,E,H): h^2=0.5. (C,F,I): h^2=0.25. (A-C): σ = 0.5. (D-F): σ = 1. (G-I): σ = 2. Populations are colored according to the color scheme used in Figures 3, S7 and S8 (see legend).