Report

Current Biology

Empty Niches after Extinctions Increase Population Sizes of Modern Corals

Highlights

- Environmental changes causing extinctions also extirpate variation in extant species
- Modern corals expanded as they filled empty niches left by extinct niche competitors
- Corals can rebound from reductions in population size under suitable conditions
- Conservation studies profit by evaluating variation under proper demographic models

Authors

Carlos Prada, Bishoy Hanna, Ann F. Budd, ..., Hiroaki Kitano, Michael DeGiorgio, Mónica Medina

Correspondence

pradac@si.edu (C.P.), mxd60@psu.edu (M.D.), mum55@psu.edu (M.M.)

In Brief

Prada et al. show how populations of corals can rebound from small population sizes under suitable conditions. Their study suggests coral populations harbor rich genetic variation to adapt to climate change. Coral conservation studies would benefit by evaluating genetic variation under appropriate demographic models.

Prada et al., 2016, Current Biology *26*, 1–5
December 5, 2016
http://dx.doi.org/10.1016/j.cub.2016.09.039

CellPress

Please cite this article in press as: Prada et al., Empty Niches after Extinctions Increase Population Sizes of Modern Corals, Current Biology (2016), http://dx.doi.org/10.1016/j.cub.2016.09.039

Current Biology
Report

Empty Niches after Extinctions Increase Population Sizes of Modern Corals

Carlos Prada,[1,12,14,*] Bishoy Hanna,[1] Ann F. Budd,[2] Cheryl M. Woodley,[3] Jeremy Schmutz,[4] Jane Grimwood,[4] Roberto Iglesias-Prieto,[1,5] John M. Pandolfi,[6,7] Don Levitan,[8] Kenneth G. Johnson,[9] Nancy Knowlton,[10] Hiroaki Kitano,[11] Michael DeGiorgio,[1,13,*] and Mónica Medina[1,10,12,13,*]

[1]Department of Biology, The Pennsylvania State University, 208 Mueller Lab, State College, PA 16802, USA
[2]Department of Earth and Environmental Sciences, University of Iowa, 115 Trowbridge Hall, Iowa City, IA 52242, USA
[3]CCEHBR, Hollings Marine Laboratory, NCCOS, National Ocean Service, US National Oceanic and Atmospheric Administration, 331 Fort Johnson Road, Charleston, SC 29412, USA
[4]HudsonAlpha Institute of Biotechnology, 601 Genome Way Northwest, Huntsville, AL 35806, USA
[5]Instituto de Ciencias del Mar y Limnología, Universidad Nacional Autónoma de México, Prol. Av. Niños Héroes, Puerto Morelos C.P. 77580, Q. Roo, Cancún, Mexico
[6]Australian Research Council Centre of Excellence for Coral Reef Studies, The University of Queensland, Brisbane, 4072, Queensland, Australia
[7]School of Biological Sciences, The University of Queensland, Brisbane, 4072, Queensland, Australia
[8]Department of Biological Science, Florida State University, Tallahassee, FL 32306, USA
[9]Department of Earth Sciences, Natural History Museum, Cromwell Road, London SW7 5BD, UK
[10]Department of Invertebrate Zoology, National Museum of Natural History, Smithsonian Institution, 10th and Constitution Avenue, NW Washington, DC 20560-0163, USA
[11]The Systems Biology Institute, Falcon Building 5F, Shirokanedai, Minato, Tokyo 108-0071, Japan
[12]Smithsonian Tropical Research Institute, Smithsonian Institution, 9100 Panama City PL, Washington, DC 20521, USA
[13]Co-senior author
[14]Lead Contact
*Correspondence: pradac@si.edu (C.P.), mxd60@psu.edu (M.D.), mum55@psu.edu (M.M.)
http://dx.doi.org/10.1016/j.cub.2016.09.039

SUMMARY

Large environmental fluctuations often cause mass extinctions, extirpating species and transforming communities [1, 2]. While the effects on community structure are evident in the fossil record, demographic consequences for populations of individual species are harder to evaluate because fossils reveal relative, but not absolute, abundances. However, genomic analyses of living species that have survived a mass extinction event offer the potential for understanding the demographic effects of such environmental fluctuations on extant species. Here, we show how environmental variation since the Pliocene has shaped demographic changes in extant corals of the genus *Orbicella*, major extant reef builders in the Caribbean that today are endangered. We use genomic approaches to estimate previously unknown current and past population sizes over the last 3 million years. Populations of all three *Orbicella* declined around 2–1 million years ago, coincident with the extinction of at least 50% of Caribbean coral species. The estimated changes in population size are consistent across the three species despite their ecological differences. Subsequently, two shallow-water specialists expanded their population sizes at least 2-fold, over a time that overlaps with the disappearance of their sister competitor species *O. nancyi* (the organ-pipe *Orbicella*). Our study suggests that populations of *Orbicella* species are capable of rebounding from reductions in population size under suitable conditions and that the effective population size of modern corals provides rich standing genetic variation for corals to adapt to climate change. For conservation genetics, our study suggests the need to evaluate genetic variation under appropriate demographic models.

RESULTS AND DISCUSSION

Mass extinctions transform biodiversity by altering species composition across habitats [1, 2]. The extinction of species frees up niches that survivors can colonize, which often leads to diversification [3]. While changes in species composition during and after extinctions are well documented, we know little about how populations of species that survive mass extinctions respond to such environmental changes. During extinctions, surviving populations may lose genetic diversity due to declining population sizes. Afterward, populations may become more abundant and increase their phenotypic and genetic diversity as they expand to fill empty niches. However, it is difficult to infer such demographic changes in surviving species deep in the past (>5 million years ago [mya]), when most mass extinctions occurred, because the genealogical density of coalescent events decreases with time, reducing the power to detect changes in population sizes [4]. One solution is to study organisms with a rich fossil record, such as mollusks and corals, and study extinctions that occurred relatively recently (<3 mya).

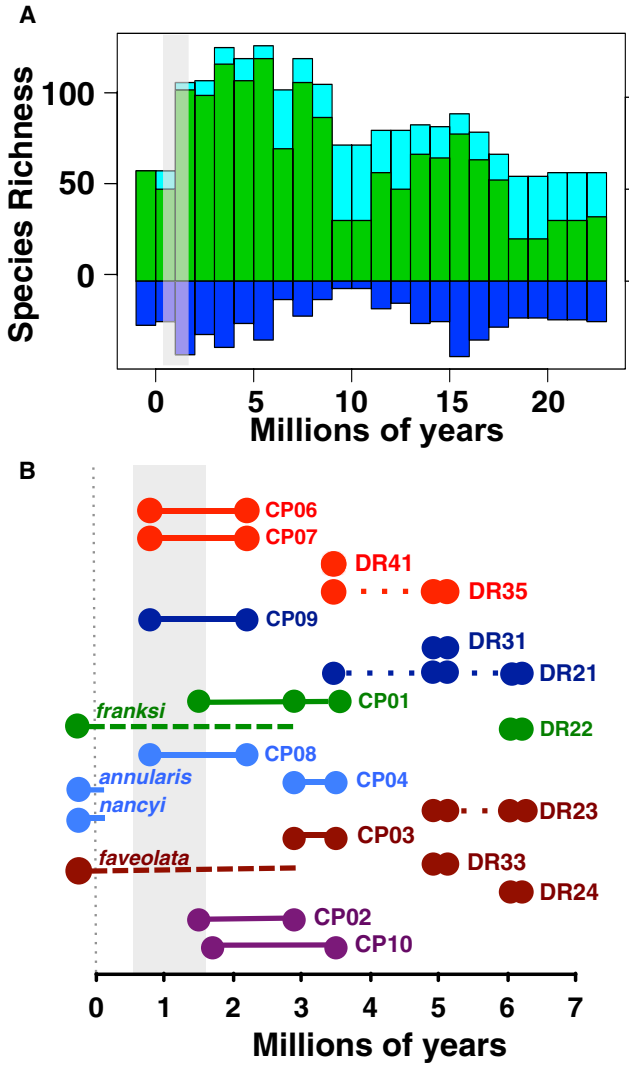

Figure 1. Species Richness through Time in the Caribbean

(A) Species turnover in Cenozoic Caribbean reef corals based on analyses of the Caribbean Cenozoic Coral Occurrence database [8]. Observed (green) and standardized (light blue) species richness and sampling intensity (dark blue) for 1-million-year intervals from early Miocene to present. Species presence for each time interval was weighted against the range of ages estimated from the locality in which a species occurred (standardization).

(B) Occurrences of all recorded *Orbicella* species between Late Miocene and Early Pleistocene time, together with estimated ranges for the three modern species and the extinct Late Pleistocene *O. nancyi*. Codes (DR, Dominican Republic; CP, Costa Rica-Panama) indicate unnamed species described by Budd and Klaus [6, 9] and the extinct *O. nancyi* [5]. Different colors correspond to morphologically defined clades; large dots indicate oldest and youngest age dates estimated for each occurrence. Solid lines indicate age ranges for occurrences, and dotted lines connect occurrences for each species. Estimated ranges of the three modern species and *O. nancyi* are inferred based on phylogenetic analyses [5, 6]. Occurrence data are provided in Table S1. See also Table S2.

The rather young coral communities of the Caribbean, where the extinction of habitat specialists allowed previously inferior competitors to spread, diversify, and dominate in habitats where they were formerly excluded [5, 6], provide such an opportunity. We compared richness in the fossil record of these corals with genome-based estimates of population sizes through time from extant species. We show that population sizes decreased as species richness decayed after the onset of the northern hemisphere glaciation. However, populations subsequently expanded as they colonized empty niches left behind by their sister competitors.

Caribbean corals experienced an intense period of speciation between 3.5 and 2.5 mya [7], which was followed by the extinction of >50% of the coral species around 2–1 mya (Figure 1A). While many of these extinctions involved species that were relicts from the Pacific, young Caribbean endemics also disappeared. Among them were lineages in the genus *Orbicella*, which has one of the best-preserved fossil records. Apart from the species that exist today, all species of *Orbicella* that had survived until 2 mya suddenly went extinct over the next million years (Figure 1B). However, we have little knowledge about abundances in *Orbicella* between 1.3 and 0.6 mya [7], even for modern species. Our genomic analyses provide estimates of population sizes of *Orbicella* species through time, covering such gaps in the fossil record.

To investigate demographic variations within *Orbicella* species, we present the first reconstructed genomes of the three modern species (one individual for each species, plus two additional individuals for *O. faveolata* from different locations) and infer effective population size changes through time. We employed both pairwise [10] and conditional sampling models [11] to estimate population size, using genomic data and the density of heterozygous sites across recombining blocks. To generate the diploid genomic sequence, we newly assembled the *O. faveolata* genome and mapped short Illumina reads from the three *Orbicella* species. Mapped reads have a mean depth >70-fold for each diploid genome. Given the uncertainty in allele calling from Illumina reads, we filtered reads by eliminating low-quality nucleotides (<Q20) and suboptimal mapped reads (MQ < 40). To avoid dubious single nucleotide polymorphisms (SNPs) due to co-aligning paralogous copies or to sequencing errors, we constrained SNP calls to areas with at least 20 but fewer than 300 reads and avoided SNPs within 10 bp of indels. To circumvent genotype miscalls due to poor mapping in areas with repetitive elements, we inferred coalescent reconstructions using masked versions of each genome.

All modern *Orbicella* corals increased in effective population size to ~40,000 individuals at ~2 mya, followed by a sharp, 2-fold decline between 1 and 2 mya (Figures 2, S1, and S2). While the pattern of initial increase and sudden decline was consistent across all three *Orbicella* species, populations of *O. faveolata* from Panama and Mexico had more pronounced reductions (~60% decrease) than populations from Florida (<30% decrease). Despite the more recent species-specific demographic trajectories (Figures 2A and 2B), the historical variation in population size was consistent across all three species until ~0.6 mya. Our genomic reconstructions coincide with the extirpation of nearly every other *Orbicella* lineage (Figure 1B).

The population decline at about 1 mya, recovered both with fossil and genetic data, can be understood in the context of changes in species diversity, habitat, and habitat availability during the Pliocene to Pleistocene transition [16]. After the onset of the Northern Hemisphere Glaciation between 2.0 and 1.5 mya, sea-surface temperatures declined in the Caribbean. The

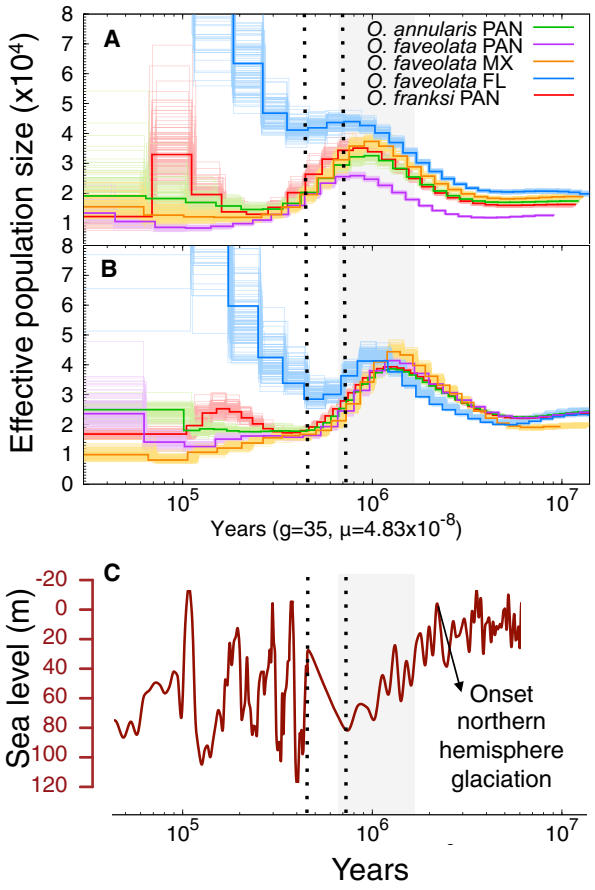

Figure 2. Variation in Effective Population Size of Modern *Orbicella* Species

(A and B) Demographic reconstructions from each sequenced diploid genome as a function of time using *O. faveolata* from Florida (A) and *O. franksi* (B) as reference. Curves are scaled by a generation time of 35 years [12] and per-generation mutation rate of 4.83×10^{-8} [13, 14]. Thin lines indicate bootstrap replicates, and thick lines indicate actual estimates. Variation among bootstrap replicates increases toward the present. The gray bar indicates period of coral extinction, and dotted lines mark the beginning and end of sea level increase after the coral mass extinction.

(C) Global sea level curve relative to the present level (0.1 mya resolution) [15]. Dotted lines indicate onset and end of sea level rise after the coral mass extinction. Genome assembly statistics are shown in Table S1.

Figure S1 shows demographic reconstructions using *O. annularis* as reference. See also reconstructions using diCal v.1.3 in Figure S2.

associated fall in sea level resulted in steeper shelf habitats, thereby reducing mesophotic habitats and shifting reefs to shallower depths [16]. The associated heightened Early Pleistocene extinction resulted in a shift to more modern coral communities comprised of species adapted to the intense climatic changes of the late Pleistocene, communities that have shown remarkable persistence in community structure since then, despite extreme climatic fluctuations (Figure 2C) [17–19].

Following the species decline at about 1 mya, most reconstructed genomes show an increase in population size, which is moderate for *O. annularis* and *O. faveolata* from Mexico and Panama but pronounced (>5-fold) for *O. faveolata* from Florida (Figures 2A and 2B). The population expansion in *O. faveolata* from Florida is associated with an increase in both sea level between 0.75 and 0.45 mya and availability of shallow reef habitat (Figure 2, dotted lines). The broad shallow shelf in Florida provides a more extensive habitat for reef taxa when compared to other Caribbean areas such as Panama or Mexico [20, 21]. Seafloor reconstructions of the upper 60 m from the NOAA ETOPO1 database [22] suggest that the habitat in Florida almost doubled when compared to that available during low sea level [21]. The increase in habitat around Florida also coincides with the time at which hybridization within the complex has been inferred in the fossil record of the nearby Bahamas platform [23].

Unlike the shallow-water specialists, *O. franksi* shows a decline in population size in recent times. Because this could be due to lower mapping efficiency to the *O. faveolata* genome, we tested for any artifacts from mapping to the *O. faveolata* genome by assembling the genome of *O. franksi* and mapping all reads against this new genome. We recovered the same population trajectories with a pronounced increase in population size for both *O. faveolata* and *O. annularis* at ~500 to 50 ka (Figure 2B), but not for *O. franksi*.

The reconstruction of population sizes from the diploid genomes of each species may lack resolution at recent times (<100 ka), as fewer coalescent events accumulate within this timeframe between pairs of genomic copies [10]. To more finely investigate the recent increase in population sizes, we compared inferences from the diploid genomes with those from allele frequencies at microsatellite loci from hundreds of individuals across the Caribbean [24].

Microsatellite variation in *O. annularis* and *O. faveolata* suggest a population expansion of 84- and 14-fold, respectively (Figure 3). In agreement with estimates from diploid genomes, the populations expanded between 100 and 10 ka, after which the population sizes of these species ranged between 10 and 50 thousand individuals. Unlike the two shallow-water specialists, microsatellite variation for *O. franksi* suggests a second population decline. The rather flat likelihood function does not provide strong support for any particular set of population sizes, but even when comparing the extreme values from the broad distribution, modern and ancient population sizes do not overlap (Figure 3), suggesting a population decline rather than an expansion in *O. franksi*.

To quantitatively assess the fit of various evolutionary demographic scenarios, we explicitly tested five models with diverse population size changes (Figure S3). We inferred the derived site frequency spectrum (SFS) for *O. faveolata* by sequencing mRNA from 15 individuals. We used ∂a∂i [25] to estimate parameters of each model from the SFS and Akaike information criterion to perform model selection [26]. To obtain confidence intervals around each parameter, we resampled SNPs from our data and generated 100 different SFSs.

The three-epoch model best fits the *O. faveolata* SFS (Figure 4). It suggests a reduction in population size 2–1 mya, coincident with our earlier reconstructions and with a coral extinction event, followed by a constant population size until ~170 ka (Figure 4), with a more recent (170–27 ka) modest (2-fold) increase in population size. Genomic inferences from the SFS and diploid genomes in the recent past coincide well with estimates based on microsatellite allele frequencies from hundreds of individuals.

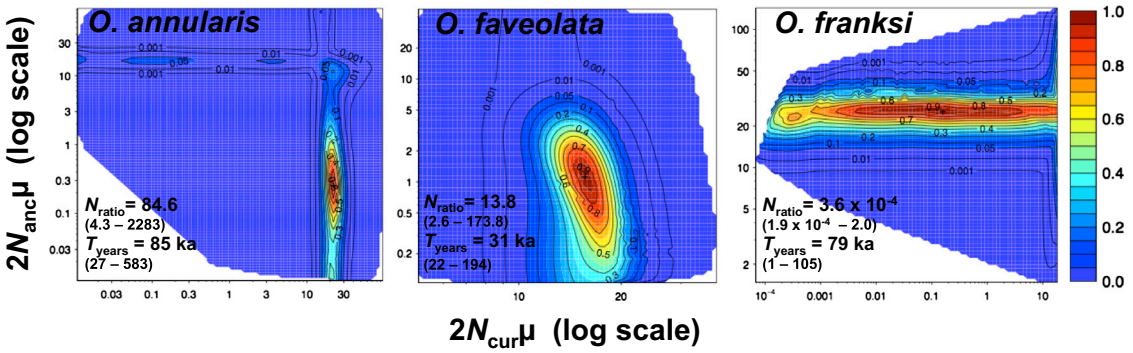

Figure 3. Joint Likelihood Surface for Ancient and Current Effective Population Sizes

The point in the surface with the highest likelihood provides estimates of the ancient (anc) and current (cur) effective population sizes under a model with a single size change. Each point in the heatmap shows the likelihood score for the pair of effective size parameters. Warmer colors indicate higher likelihoods, and colder colors indicate lower likelihoods. White spaces reflect the unsampled parameter space. Effective population sizes are scaled by the mutation rate (μ). Estimates for each species were inferred from microsatellite allele frequencies across hundreds of individuals. Fold change and range of the age of the change are shown. Mean values and 95% confidence intervals are given in Table S2. The maximum likelihood values for the ratio of effective population sizes and the time at which the population changed size are given by $N_{ratio} = N_{cur}/N_{anc}$ and T_{years}, respectively. See full statistics in Table S3.

This recent increase in population size in both *O. faveolata* and *O. annularis*, but not in *O. franksi*, overlaps with the extinction of the organ-pipe coral (*O. nancyi*) [5]. Both *O. annularis* and *O. faveolata* are shallow-water specialists that competed for habitat with the extinct *O. nancyi*. The organ-pipe coral was one of the most common shallow-water corals in the fossil record from ~600 ka to ~100 ka and then disappeared sometime thereafter [5, 6]. The two modern shallow-water specialists, *O. annularis* and *O. faveolata*, have since increased in abundance in shallow water [5]. This observation suggests that the extinction of the faster growing shallow-water *O. nancyi* increased the available habitat for the two modern *Orbicella* shallow-water specialists, which elevated their population sizes as they expanded to fill this vacated niche.

In summary, environmental changes that caused species extinctions not only transformed the ecological structure of coral communities but also changed the demographic trajectories of modern corals. The decrease in shallow habitat availability immediately following the onset of the Northern Hemisphere Glaciation provoked a severe bottleneck across all modern *Orbicella* species. More recently, the habitat space released by the sudden disappearance of the dominant shallow-water specialist organ-pipe (*O. nancyi*) caused a sharp increase in population size of the two modern shallow-water specialists, rebuilding their genetic variation.

The persistence and even expansion of modern corals after Pleistocene species extinctions demonstrates their ability to rebound from drastic environmental fluctuations. The return of modern *Orbicella* species from the devastating event that caused the disappearance of half of the biodiversity on coral reefs suggests that the current worldwide reef deterioration could be reversed if global and local stressors are alleviated. As we continue to study the demographic trajectory of corals under suboptimal conditions, models that incorporate population bottlenecks and expansions, deviating from drift-mutation equilibrium [27], may be better suited for analyzing contemporary genetic variation and the potential of corals to adapt to climate change.

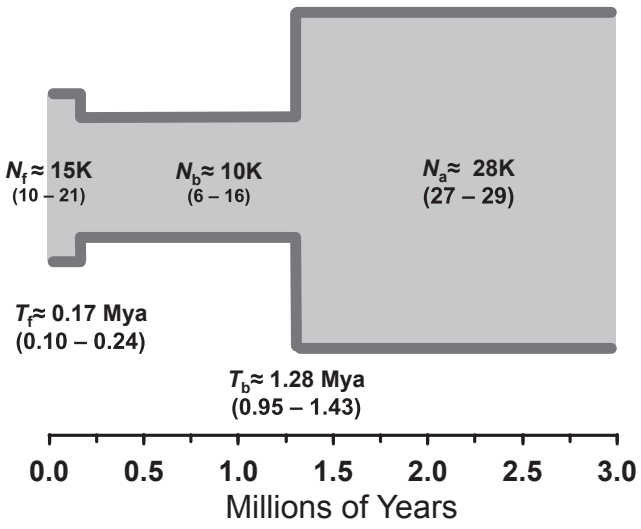

Figure 4. Best-Fit Demographic Model for *O. faveolata* from Puerto Rico and Mexico

Parameters inferred from a derived site frequency spectrum obtained from 75,983 SNPs in 18 individuals. Demographic curves are scaled by a generation time of 35 years [12]. Confidence intervals from 100 non-parametric bootstrap replicates are displayed in parentheses. Likelihood scores and model ranking are shown in Table S4. Best-fitted model is a three-epoch model. See also Figures S3 and S4.

ACCESSION NUMBERS

Illumina reads can be downloaded at www.ncbi.nlm.nih.gov under project number NCBI: PRJNA342412.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, four figures, and four tables and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2016.09.039.

AUTHOR CONTRIBUTIONS

C.P., M.M., and M.D. conceived the study. M.M., D.L., R.I.-P., B.H., N.K., and C.M.W. collected the material. C.P., B.H., M.D., J.S., J.G., and H.K. led the bioinformatic analysis of data. B.H., J.S., and J.G. assembled the genomes. A.F.B., J.M.P., and K.G.J. performed all fossil analysis, and C.P. wrote the paper with input from the other authors. All authors approved the manuscript before submission.

ACKNOWLEDGMENTS

Our project was supported with startup funds by The Department of Biology at The Pennsylvania State University, NSF grants: OCE 1442206 and IOS 0644438, NOAA Coral Reef Conservation Program (CRCP 30022, CDHC 1133), Hudson Alpha, The Smithsonian Tropical Research Institute, and The Cannon Foundation. C.P. has been supported by the Earl S. Tupper Fellowship from the Smithsonian Tropical Research Institute. We thank the governments of Panama, Mexico, and the USA for granting permits for coral collections. We thank Michele Weber, Mary Alice Coffroth, Anastasia Banaszak, and Shinichi Sunagawa for field assistance. We thank Tom Capo and Phil Gillette for keeping coral genome colony stocks at coral hatchery at the University of Miami. The CyberSTAR cluster at Pennsylvania State University provided computing resources. All our procedures are in agreement with the Institutional Animal Care and Use Committee (IACUC) of The Pennsylvania State University.

Received: June 13, 2016
Revised: September 20, 2016
Accepted: September 21, 2016
Published: November 17, 2016

REFERENCES

1. Jackson, J.B., Jung, P., Coates, A.G., and Collins, L.S. (1993). Diversity and extinction of tropical american mollusks and emergence of the isthmus of panama. Science *260*, 1624–1626.
2. Budd, A.F., Stemann, T.A., and Johnson, K.G. (1994). Stratigraphic distributions of genera and species of Neogene to Recent Caribbean reef corals. J. Paleontology. *68*, 951–977.
3. Meredith, R.W., Janečka, J.E., Gatesy, J., Ryder, O.A., Fisher, C.A., Teeling, E.C., Goodbla, A., Eizirik, E., Simão, T.L., Stadler, T., et al. (2011). Impacts of the Cretaceous terrestrial revolution and KPg extinction on mammal diversification. Science *334*, 521–524.
4. Hudson, R. (1991). Gene genealogies and the coalescent process. In Oxford Surveys of Evolutionary Biology, D. Futuyma, and J. Antonovics, eds. (Oxford University Press), pp. 1–44.
5. Pandolfi, J.M., Lovelock, C.E., and Budd, A.F. (2002). Character release following extinction in a Caribbean reef coral species complex. Evolution *56*, 479–501.
6. Budd, A.F., and Klaus, J.S. (2001). The origin and early evolution of the *Montastraea* "*annularis*" species complex (Anthozoa: Scleractinia). J. Paleontology *75*, 527–545.
7. Budd, A.F., and Johnson, K.G. (1999). Origination preceding extinction during Late Cenozoic turnover of Caribbean reefs. Paleobiology *25*, 188–200.
8. Budd, A.F., Klaus, J.S., and Johnson, K.G. (2011). Cenozoic diversification and extinction patterns in Caribbean reef corals: a review. Paleontological Society Papers *17*, 79–94.
9. Budd, A.F., and Klaus, J.S. (2008). Early evolution of the Montastraea "annularis" species complex (Anthozoa: Scleractinia): evidence from the Mio-Pliocene of the Dominican Republic. In Evolutionary Stasis and Change in the Dominican Republic, R.H. Nehm, and A.F. Budd, eds. (Neogene, Springer), pp. 85–124.
10. Li, H., and Durbin, R. (2011). Inference of human population history from individual whole-genome sequences. Nature *475*, 493–496.
11. Sheehan, S., Harris, K., and Song, Y.S. (2013). Estimating variable effective population sizes from multiple genomes: a sequentially markov conditional sampling distribution approach. Genetics *194*, 647–662.
12. Babcock, R. (1991). Comparative demography of three species of scleractinian corals using age-and size–dependant classifications. Ecol. Monogr. *61*, 225–244.
13. Voolstra, C.R., Sunagawa, S., Matz, M.V., Bayer, T., Aranda, M., Buschiazzo, E., Desalvo, M.K., Lindquist, E., Szmant, A.M., Coffroth, M.A., and Medina, M. (2011). Rapid evolution of coral proteins responsible for interaction with the environment. PLoS ONE *6*, e20392.
14. Prada, C., DeBiasse, M.B., Neigel, J.E., Yednock, B., Stake, J.L., Forsman, Z.H., Baums, I.B., and Hellberg, M.E. (2014). Genetic species delineation among branching Caribbean *Porites* corals. Coral Reefs *33*, 1019–1030.
15. Miller, K.G., Kominz, M.A., Browning, J.V., Wright, J.D., Mountain, G.S., Katz, M.E., Sugarman, P.J., Cramer, B.S., Christie-Blick, N., and Pekar, S.F. (2005). The Phanerozoic record of global sea-level change. Science *310*, 1293–1298.
16. Klaus, J., Lutz, B., McNeill, D., Budd, A., Johnson, K., and Ishman, S. (2011). Rise and fall of Pliocene free-living corals in the Caribbean. Geology *39*, 375–378.
17. Pandolfi, J.M., and Jackson, J.B.C. (2006). Ecological persistence interrupted in Caribbean coral reefs. Ecol. Lett. *9*, 818–826.
18. Renema, W., Pandolfi, J.M., Kiessling, W., Bosellini, F.R., Klaus, J.S., Korpanty, C., Rosen, B.R., Santodomingo, N., Wallace, C.C., Webster, J.M., and Johnson, K.G. (2016). Are coral reefs victims of their own past success? Sci. Adv. *2*, e1500850.
19. Simpson, C., Kiessling, W., Mewis, H., Baron-Szabo, R.C., and Müller, J. (2011). Evolutionary diversification of reef corals: a comparison of the molecular and fossil records. Evolution *65*, 3274–3284.
20. Ludt, W.B., and Rocha, L.A. (2015). Shifting seas: the impacts of Pleistocene sea-level fluctuations on the evolution of tropical marine taxa. J. Biogeogr. *42*, 25–38.
21. Bellwood, D.R., and Wainwright, P.C. (2002). The history and biogeography of fishes on coral reefs. In Coral Reef Fishes, P.F. Sale, ed. (Elsevier Science).
22. Amante, C., and Eakins, B.W. (2009). ETOPO1 1 arc-minute global relief model: procedures, data sources and analysis. NOAA Technical Memorandum NESDros. Inf. Serv. NGDC-24. National Geophysical Data Center, NOAA, doi: http://dx.doi.org/10.7289/V5C8276M.
23. Budd, A.F., and Pandolfi, J.M. (2010). Evolutionary novelty is concentrated at the edge of coral species distributions. Science *328*, 1558–1561.
24. Leblois, R., Pudlo, P., Néron, J., Bertaux, F., Reddy Beeravolu, C., Vitalis, R., and Rousset, F. (2014). Maximum-likelihood inference of population size contractions from microsatellite data. Mol. Biol. Evol. *31*, 2805–2823.
25. Gutenkunst, R.N., Hernandez, R.D., Williamson, S.H., and Bustamante, C.D. (2009). Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. PLoS Genet. *5*, e1000695.
26. Anderson, D.R. (2008). Model Based Inference in the Life Sciences (Springer).
27. Nei, M. (1972). Genetic distance between population. Am. Nat. *106*, 283–292.

Current Biology, Volume *26*

Supplemental Information

Empty Niches after Extinctions Increase Population Sizes of Modern Corals

Carlos Prada, Bishoy Hanna, Ann F. Budd, Cheryl M. Woodley, Jeremy Schmutz, Jane Grimwood, Roberto Iglesias-Prieto, John M. Pandolfi, Don Levitan, Kenneth G. Johnson, Nancy Knowlton, Hiroaki Kitano, Michael DeGiorgio, and Mónica Medina

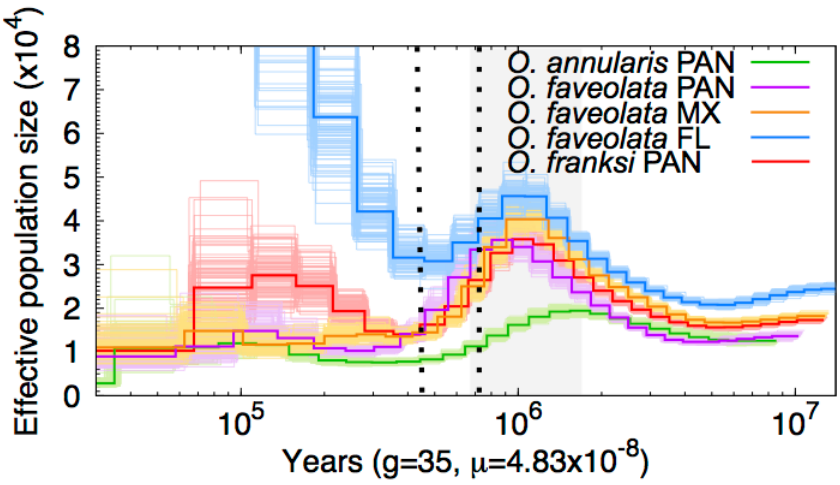

Figure S1. Variation in effective population size of modern *Orbicella* species from each sequenced diploid genome as a function of time using *O. annularis* as reference. Related to Figure 2.
Demographic curves are scaled by a generation time of 35 years [S1] and per generation mutation rate of 4.83×10^{-8} [S2, S3]. Thin lines indicate bootstrap replicates and thick lines indicate actual estimates. Variation among bootstrap replicates increases toward the present. The grey bar indicates period of coral mass extinction. Dotted lines mark onset and end of sea level rise after the coral mass extinction.

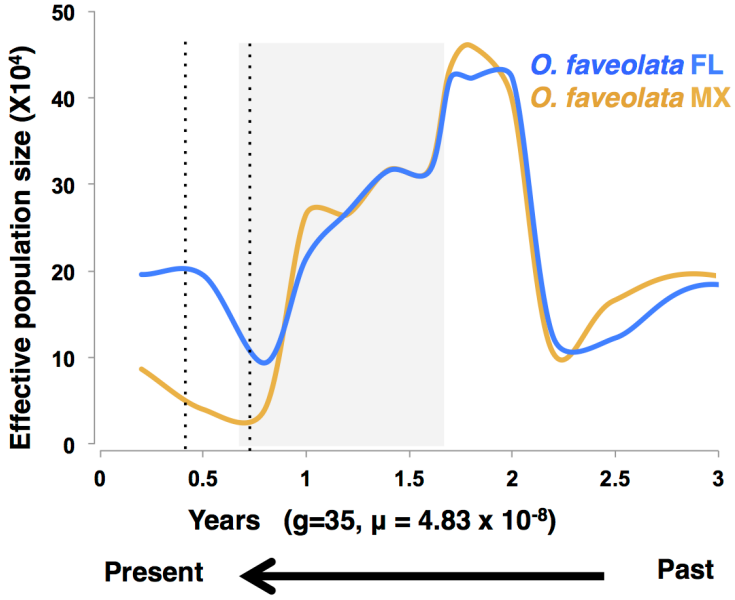

Figure S2. Demographic reconstruction for *O. faveolata* from Mexico using diCal v. 1.3.0 and the *O. faveolata* genome as reference. Related to Figure 2.
Grey bar indicates period of coral mass extinction. Dotted lines indicate onset and end of sea level rise after the coral mass extinction.

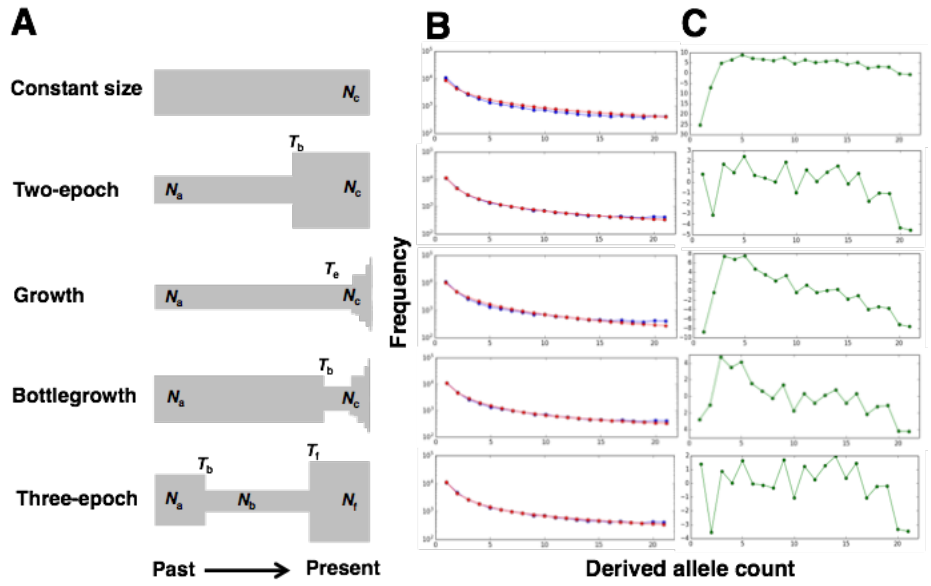

Figure S3. Demographic reconstructions using SNPs derived from RNAseq data. Related to Figure 4. A. Demographic models tested using ∂a∂i applied to the derived site frequency spectrum projected onto 23 alleles with the removal of the last two categories (due to ancestral state misspecification). B. Comparison of the derived site frequency spectrum between the model (red) and the data (blue) and (C) the residuals for each mutational category.

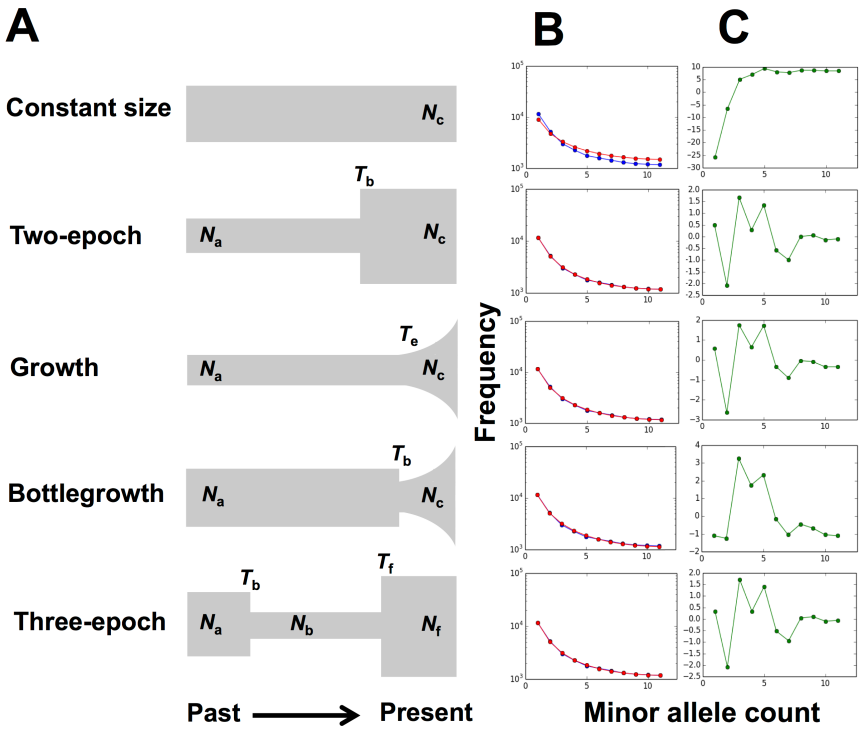

Figure S4. Demographic reconstructions using SNPs derived from RNAseq data. Related to Figure 4. A. Demographic models tested using ∂a∂i applied to the folded site frequency spectrum. B. Comparison of the folded site frequency spectrum between the model (red) and the data (blue) and (C) the residuals for each mutational category.

Table S1. Occurrence Data for Figure 1B [S4, S5]. Occurrence data are dated using the carbonate rocks that contain the fossils. We used the range of dates for each occurrence provided by Don McNeill.

Fossil Species	#colonies	Stratigraphic Unit	Age Date Range (Ma)	Source
CP06	11	p1/2	0.8-2.2	4
CP07	5	p1/2	0.8-2.2	4
CP09	7	p1/2	0.8-2.2	4
DR41	5	level 4	3.48	5
DR35	2	level 3	4.93-5.13	5
DR35	7	level 4	3.48	5
DR31	10	level 3	4.93-5.13	5
DR21	54	level 2	6.06-6.23	5
DR21	37	level 3	4.93-5.13	5
DR21	5	level 4	3.48	5
CP01	31	c1/2	1.5-2.9	4
CP01	2	c3/4	2.9-3.5	4
DR22	14	level 2	6.06-6.23	5
CP08	4	p1/2	0.8-2.2	4
CP04	6	c3/4	2.9-3.5	4
DR23	3	level 2	6.06-6.23	5
DR23	4	level 3	4.93-5.13	5
CP03	9	c3/4	2.9-3.5	4
DR33	11	level 3	4.93-5.13	5
DR24	7	level 2	6.06-6.23	5
CP02	16	c1/2	1.5-2.9	4
CP10	3	p3	1.7-3.5	4

Table S2. Genome assembly statistics at various steps during the assembly with ABySS. Related to reference genome used in Figure 1A.

Assembly	SSPACE Scaffolds Broken at Nx	SSPACE Scaffolds	ABySS-Scaffolds broken at N	ABySS-Scaffolds	SSPACE + gapcloser Broken at N	SSPACE+gap closer
# contigs (≥ 0 bp)	6,180,476	6,076,806	6,183,330	6,111,770	6,110,919	6,076,806
# contigs (≥ 1000 bp)	91,697	35,361	92,808	63,983	64,732	35,367
Total length (≥ 0 bp)	1,038,030,286	1,234,345,766	1,038,108,421	1,044,426,688	1,095,423,029	1,236,531,036
Total length (≥ 1000 bp)	286,630,672	510,028,246	285,566,987	317,139,144	368,810,628	512,216,702
# contigs	155,781	69,703	158,214	101,319	101,180	69,705
Largest contig	56,475	492,874	56,475	241,245	144,459	495,114
Total length	332,586,860	534,300,073	332,525,738	343,784,499	394,676,386	536,485,889
GC (%)	38.90	38.90	38.90	38.90	38.90	38.90
N50	3,356	60,793	3,265	8,134	9,488	61,194
N75	1,532	20,837	1,495	2,714	3,669	20,901
L50	24,500	2,288	25,035	8,866	9,788	2,282
L75	61,633	6,000	63,221	27,886	26,801	5,981
# N's per 100 kbp	24.79	36,757.9	23.97	1,860.92	9.17	26,308.98
# predicted genes (unique)	97,121	64,046	97,760	76,875	86,200	72,623
# predicted genes (≥ 0 bp)	98,803	64,985	99,338	78,061	89,617	75,358
# predicted genes (≥ 300 bp)	49,895	41,875	49,923	44,797	55,054	51,011
# predicted genes (≥ 1500 bp)	12,364	14,330	12,275	14,601	18,943	18,331
# predicted genes (≥ 3000 bp)	4,486	6,930	4,422	6,935	8,869	8,840

Table S3. Summary statistics for demographic reconstructions using allele frequencies at microsatellite loci. Related to Figure 3. Ancestral (θ_{anc}) and current (θ_{cur}) genetic variation estimated as $\theta = 2N\mu$ and $D = T/(2N)$. $N_{ratio} = 2N_{act}\mu/(2N_{anc}\mu) = N_{cur}/N_{anc}$. Population sizes are shown as diploid individuals. Times in years (T_{years}) were estimated using a mutation rate of 5×10^{-4} per locus per generation [S6 - S8], assuming a generation time of 35 years [S1]. Sample sizes are given as the number of individuals in parentheses.

	D	θ_{cur}	θ_{anc}	***N*/2**	T_{years}	N_{ratio}
***O. annularis* (482)**						
Mean estimate	0.20	21.99	0.26	10,995	85,761	84.6
Lower bound of 95% CI	0.08	16.29	0.0105	8,145	27,367	4.376
Upper bound of 95% CI	0.95	30.56	11.26	15,280	583,085	2283
***O. faveolata* (473)**						
Mean estimate	0.10	15.67	1.129	7,835	31653	13.88
Lower bound of 95% low CI	0.10	11.11	0.102	5,555	22,220	2.636
Upper bound of 95% high CI	0.46	21.16	5.299	10,580	194,672	173.8
***O. franksi* (120)**						
Mean estimate	0.19	21.00	29.1	10,500	79,800	0.000362
Lower bound of 95% CI	0.01	13.62	20.65	6,810	1,632	0.000199
Upper bound of 95% CI	0.10	38.71	41.3	19,355	75,949	2

Table S4. Information theoretic statistics for each demographic model run in *∂a∂i* using an unfolded and folded SFS. Related to the best model displayed on Figure 4. *k*, number of parameters; Log(*L*), log likelihood; AIC, Akaike information criterion inferred as (-2×log likelihood)+(2×number of parameters); Δ*i*, difference in AIC score with respect to the best model (three-epoch); model likelihoods, relative likelihood of the model given the data; w_i, model probabilities; evidence ratio, fold difference in model probabilities against the best model. Likelihood scores were inferred after removing the last two categories to avoid SNP ancestral state misspecification.

					Model		**Evidence Ratio**
Model	***k***	**log(*L*)**	**AIC**	Δ*i*,	**Likelihoods**	w_i	**(*w* best model/ w_i)**
unfolded SFS							
three-epoch	4	-110.26	228.53	0.00	1	0.9998	
bottlegrowth	3	-119.68	245.36	16.83	$2.22E^{-04}$	$2.22E^{-04}$	$4.51E^{+03}$
two-epoch	2	-127.83	259.67	31.14	$1.73E^{-07}$	$1.73E^{-07}$	$5.77E^{+06}$
growth	2	-132.28	268.56	40.03	$2.03E^{-09}$	$2.03E^{-09}$	$4.93E^{+08}$

neutral	1	-714.19	1430.39	1201.86	$1.05E^{-261}$	$1.05E^{-261}$	$9.54E^{+260}$
folded SFS							
three-epoch	4	-56.54	121.08	0	1	0.4193	
two-epoch	2	-58.54	121.08	$3.80E^{-05}$	1	$4.19E^{-01}$	1
growth	2	-59.54	123.08	2.00	$3.68E^{-01}$	$1.54E^{-01}$	2.72
neutral	1	-63.69	129.37	8.29	$1.58E^{-02}$	$6.64E^{-03}$	$6.32E^{+01}$
bottlegrowth	3	-64.05	134.11	13.03	$1.48E^{-03}$	$6.21E^{-04}$	$6.75E^{+02}$

Supplemental Experimental Procedures

Fossil data

The species richness of Caribbean reef corals over geologic time was estimated using a compilation of fossil occurrences, the Caribbean Cenozoic Coral Occurrence database, consisting primarily of new collections from outcrop exposures at 811 localities in 15 countries [S9, S10]. These collections comprise a total of >15,000 specimens, which belong to ~230 species. In the analysis, the localities were grouped into faunules, defined as a set of lithologically similar localities from a small geographical area (usually < 1 km) and a restricted stratigraphic interval (usually less than 20 m). Age dates for the faunules were obtained by integrating data using high-resolution chronostratigraphic methods, including nanofossils and planktonic foraminiferal biostratigraphy, paleomagnetics, and strontium isotope analyses (e.g., [S11, S12]), and generally range in accuracy from 2-0.5 million years. To standardize species richness, stratigraphic ranges of species were determined using the oldest and youngest occurrences for each species in the database. These ranges were subdivided into time bins (usually one million years), and both numbers of actual species occurrences (observed) and species ranges (range-through) were counted within each time bin.

Reference genome sequencing and assembly

To generate a reference genome, we extracted high molecular weight genomic DNA from sperm (aposymbiotic) using the QIAGEN Genomic-tip 100/G protocol from an individual of *O. faveolata* sampled from (24.812697° N; Florida 80.66925° W). The individual with the reference genome is available at the RSMAS coral husbandry facility. We generated various libraries with different chemistries, technologies and insert sizes to produce the most comprehensive dataset to assemble the genome. To assemble the genome we used two independent NuGEN standard unamplified libraries sequenced in two HiSeq (2x100) channels, another standard unamplified library sequenced in two channels of a HiSeq 2x100 rapid mode, two Nextera amplified mate-pair libraries of 6-9kb and 2.1 -3.1kb sequenced in two MiSeq runs (one at 2x250 and the other at 2 X 300) and two independent Illumina Tru-Seq PCR-free library sequenced in two MiSeq 2x300 runs.

We initially cleaned all sequences for contaminants and library adapters and trimmed bases at the end of reads when required using Trimmomatic ver. 0.35 [S13], Sickle ver. 1.33 [S14] and FastQC [S15]. We trimmed reads multiple times to assure quality of reads, and retained only proper pairs. We checked read length and quality after clean up with FastQC ver. 0.11.2 [S15]. To evaluate optimal kmer size and increase alignment power during assembly, we error-corrected all reads using RACER [S16]. To generate the best kmer size, we inferred the optimal kmer size before and after error correction using KmerGenie [S17]. The best kmer for the *O. faveolata* genome was 71bp.

Once we had all clean error-corrected sequences and determine best kmer, we tried several assemblers (SOAPdenovo, Hapsambler, SPADES, SGA, Platanus, ALLPATHS, Velvet, IDBA and ABySS).

ABySS Ver. 1.5.0 [S18] produced the best assembly. To generate the assembly in ABySS, we set kmer size to 71 as predicted by KmerGenie, set the bubble size to 10,000 and the popping similarity cutoff to 0.7.

To scaffold the initial contigs generated by ABySS, we used SSPACE 3 [S19] with libraries in increasing size order and adding the information from the long insert mate-pairs. The N50 after scaffolding went from 8,134 to 60,793. We then used Gap-Closer to reduce Ns within scaffolds and produced scaffolds with an N50 of 61,194. Once we generated the best assembly, we ran CEGMA [S20] and recovered 54.4% of the complete Core Eukaryotic Genes (CEG) and 83.4 of the partial CEGs with an approximate genome size of 536 Mb. All assembly statistics as inferred by Quast [S21] are summarized on Table S2.

Additional diploid genome Sequencing

In addition to our assembled referenced genome, we generated three new whole genome sequences—one each for *O. annularis*, *O. franksi*, and *O. faveolata*. We additionally sequenced a genome for *O. faveolata* from Mexico and one for *O. faveolata* from Panama. All samples were collected under CITES permit numbers (SE/A-94-13 and MX-HR-010-MEX). All samples from Panama had been studied for over 20 years by the Levitan and Knowlton labs [S22, S23], and are located at Hospital Point, Bocas Del Toro, Panama (9°19'32.06"N; 82°11'51.69"W). The sample from Mexico was collected around Puerto Morelos (20° 52.49388N; 86° 51.07368W). For these four samples (*O. franksi, O. annularis, O. faveol*ata MX and PAN), we extracted genomic high molecular weight DNA from sperm (aposymbiotic) using QIAGEN Genomic-tip 100/G protocol. PCR-free Illumina libraries were prepared using Illumina Tru-Seq DNA chemistry and protocols. Hudson-Alpha generated and sequenced the gDNA libraries. Each library was sequenced in two independent 150 paired-end lanes of an Illumina HiSeq. Each library produced > 50 Gb of data, and when mapped to the *O. faveolata* reference genome produced a read depth ranging from 74 to 190 reads per site.

As for the reference genome assembly, reads from Illumina sequencing were cleaned from sequencing adapters and primers using Trimmomatic ver. 0.35 [S13], and quality control was performed with FastQC [S15]. For both the *O. annularis* and *O. franksi* samples, we also assembled their genomes following approaches explained above.

Coalescent population size inferences from diploid genomes

To estimate changes in population size through time, we employed coalescent approaches implementing hidden Markov models (HMMs) derived from approximations to the coalescent model with recombination. One approach, PSMC, infers demographic changes by modeling fluctuations in coalescence times in a pair of sequences using an approximation to the ancestral recombination graph [S24]. PSMC infers demographic patterns by estimating the distribution times to the most recent common ancestor across chromosomes. The other approach (diCal v1) uses conditional sampling [S25] and the "trunk genealogy" approximation, and can be used to infer demography with more than two sequences. Further, diCal unlike PSMC, uses a linear approximation to discretize times. However, when making inferences from a pair of sequences both models use heterozygosity across the genome to estimate population size using density of heterozygous sites at various recombining blocks.

To generate the diploid genomic sequence, we used our newly assembled *O. faveolata* genome and mapped short Illumina reads from the three *Orbicella* species using the BWA mem algorithm ver. 0.7.12 [S26]. To avoid SNP errors close to indels, we locally realigned reads using the Genome Analysis Tool Kit (GATK) 3.4 [S27]. Given the uncertainty in allele calling from Illumina reads, we used three approaches to make our reconstructions. We inferred coalescent reconstructions in PSMC from both the whole repeat-masked and unmasked *O. faveolata* genome. We also *de-novo* assembled the *O. franksi* and *O. annularis* genome, and used masked versions as reference to run PSMC.

In all cases, we filtered reads by eliminating low quality nucleotides (< Q20) and reads with low mapping quality (MQ<40), constrained calls to SNPs with at least 20 and fewer than 300 overlapping reads, and eliminated SNPs within 10bp of indels.

After mapping against the *O. faveolata* or *O. franksi* genomes and filtering, we generated a diploid genome sequence using "*vcfutils.pl vcf2fq"* for each of the species to use in PSMC. To translate the values in PSMC to demographic units as time estimates in years and population sizes in numbers of individuals, we used our earlier inferred per year mutation rate of 1.38×10^{-9} for *Porites* and *Acropora* corals [S2, S3]. We also assumed a generation time of 35 years [S1].

We ran PSMC using three different demographic schemes. We initially inferred demography using 28 population size parameters and 64 atomic time intervals using –p '4 + 25*2 + 4 + 6' as suggested by [S24]. Each number separated by a plus, represents a different time interval or a time strata at which PSMC calculates the population size (28 in this case, '1+25+1+1'). The –p '4 + 25*2 + 4 + 6' also contains info

about the number and distribution of atomic values. Here we have 64 (the sum of all numbers) and their distribution. In this case, the first size parameter spans the first four atomic time intervals, the next twenty-five size parameters spans two intervals each, the 27th spans 4 intervals and the last spans six time intervals. To test for deviations assuming different time intervals, we also inferred populations sizes assuming: 1*6+58*1 [S28] and '1*4+1*4+1*3+18*2+1*3+1*4+1*6 [S29]. In all three scenarios, we recover the same evolutionary history for each species. We generated 100 bootstrap replicates by slicing the genome and re-estimating the demography from each chunk.

To generate an alternative approach of estimating population size from diploid genomes, we also applied the method introduced by [S25] as implemented in diCal v1.3. DiCal operates on long genomic segments, so we constrained our demographic inference to scaffolds > 80,000 kb. We followed recommendations in the manual, used default parameters and discretized time into 14 intervals. We discretized parameters across time intervals using "-p "3+2+2+2+2+3"". In this case, our first and last size parameters spans three time intervals and the others parameters overlap only two time intervals.

Maximum Likelihood reconstructions from microsatellite data

To better understand changes in population size in more recent times (< 0.5 Ma), we employed microsatellite data for each species along with the sampling algorithm developed by [S30], a stepwise-mutation model [S31] and a recent implementation for historical variation in population size coded in Migraine ver. 0.4.1 [S32]. Recent simulations [S32] suggest that the model implemented in Migraine outperforms other methods and that it is particularly robust when changes in population size are within the ranges estimated here for the three *Orbicella* species. Migraine estimates population size from genetic diversity as $\theta = 2N_e\mu$, assuming a model of a single population size change from the past to the present. If a difference in effective size (N_e) is detected, then Migraine estimates the N_e ratio ($\theta_{act}/\theta_{anc}$). A ratio of less than one indicates a population decline, and a ratio greater than one a population expansion.

We ran Migraine on all three species using published datasets. For *O. faveolata*, we employed data from [S33] consisting of 473 individuals genotyped at five microsatellite loci. For *O. annularis,* we utilized the data from [S34] that included 871 individuals screened at six microsatellite loci. For *O. franksi*, we used the data from [S22] composed of 120 individuals sampled at six microsatellite loci.

All runs in Migraine were performed for microsatellites using at least 20,000 runs per point, 3,000 points and at least five iterations. To estimate the time of the population size change, Migraine estimates the parameters T and $D=T/(2N)$ in generations using the formula $T = D \times 2N$ [S32]. To obtain estimates of population sizes and times from Migraine parameters we utilized the microsatellite mutation rate of 5×10^{-4} per generation [S6-S8]. To quantify the variation around our population size and time estimates, we computed the upper and lower bounds of the 95% confidence intervals from Migraine along with our chosen mutation rate.

RNAseq library preparation

To generate RNAseq data, we collected four individuals of *O. faveolata* from Puerto Morelos, Mexico (20° 52.49388N; 86° 51.07368W) under permit MX-HR-010-MEX. We collected samples by scuba diving, and removed samples with a hammer and a chisel. Upon collection, individuals were immediately flash-frozen and kept at -80°C. We extracted total RNA using the RNA Qiagen extraction kit and followed manufacture instructions. We shipped tRNA samples to the Joint Genome Institute (JGI). Samples were then rRNA depleted using Epicentre's Ribo-Zero rRNA removal kit. The JGI generated libraries using Illumina Tru-Seq chemistry and sequenced them in one lane of Illumina HiSeq (100 PE reads). As for the genomes, reads from Illumina sequencing were cleaned from adapters using Trimmomatic ver 0.35 [S13], and quality checked using FastQC ver. 0.11.2 [S15]. After cleaning, our libraries range between 4.6 and 9.6 Gb (19 to 41 million reads) of data.

Along with the in-lab generated RNAseq libraries, we added data from two published studies. The first dataset, is composed of six samples for a study on coral disease (Bioproject number PRJNA236103) [S35], while the other has five individuals and was designed to test for variation in bleaching susceptibility in *O. faveolata* (PRJNA203198) [S36]. In both studies, coral expert Ernesto Weil identified and sampled the individuals in the field. In all cases we pooled reads from the same individual following the SRA designation from the NCBI files.

Site frequency spectrum: SNP calling and demographic model testing

To quantitatively assess the fit of various evolutionary models with diverse degrees of expansion and bottlenecks, we tested them explicitly using the derived site frequency spectrum (SFS) from RNAseq data of 15 individuals for *O. faveolata*.

To estimate the SFS, we initially mapped reads against the reference genome using BWA mem algorithm version 0.7.12 [S26] and then generated sorted bam files using SAMtools 1.1. To avoid biases from misalignments and mapping, we conservatively used The Genome Analysis Tookit (GATK) version 3.3 [S27] and followed their best practices [S37]. The sorted BAM files generated from BWA were used to feed GATK. In GATK, we initially realigned variants around indels and used the UnifiedGenotyper in GATK to generate putative variants. To avoid false positives, we retained SNPs that were present in all samples at > 50 fold coverage and each allele is present in at least 30% of the reads at that site (i.e., a 30 to 70% balance between variants for each SNP). We captured 6,370 'true' SNPs that we then used to train the GATK Variant Quality Score Recalibration (VQSR) algorithm [S37] and test all 97,181 SNPs. Our VQSR filtering step has a Gaussian mixture model with mapping quality (MQ), sample depth (DP) and inbreeding coefficient (or deviations from Hardy-Weinberg equilibrium). We retained all SNPs with a VQSR sensitivity score > 90% with 10X coverage per sample and present in at least 12 samples. To avoid link sites, we only kept sites separated by at least 2000 bp.

To transform the filtered 75,983 SNPs from the VQSR to ∂a∂i format, we used the *vcf2dadi.pl* script (available: https://groups.google.com/forum/#!searchin/dadi-user/vcf2dadi.pl/dadi-user/kvzhF4XSyng/idVM5lLUpt0J). Using the *O. annularis* genome as an outgroup, we then generated the unfolded SFS using ∂a∂i ver. 1.7.0 [S38]. To capture information from most SNPs and compensate for missing data, we projected the SFS to 23 alleles. ∂a∂i's projection uses a hypergeometric distribution to average across all results from sampling all 23 alleles from the total number of genotype calls at each SNP [S38]. To account for variability in the SFS and ∂a∂i's estimation, we used a non-parametric bootstrapping by resampling 100 times the SNP file generated from *vcf2dadi.pl* using dadiBoot.pl (available: https://groups.google.com/forum/#!searchin/dadi-user/vcf2dadi.pl/dadi-user/kvzhF4XSyng/idVM5lLUpt0J).

In ∂a∂i, we explicitly tested five models of demographic evolution: constant size, two-epoch model, growth, bottlegrowth and three-epoch (Figs. S3 and S4). Though for some of the models we depict a population expansion or contraction in Figure. S3A the inference in ∂a∂i does not assume a specific type of size change (expansion or contraction), only that a size change occurred.

For each model, we ran ∂a∂i 100 times for each of the 100 bootstrap replicates and extracted the inferred model parameters with the lowest likelihood score across the 10,000 (100 starting replicates for each of the 100 bootstrap datasets) combinations for each model. We then inferred the Akaike information criterion (AIC) scores, differences between each model against the best model, the relative likelihood of each model given the data, the model probability and the evidence ratio in favor of each model. We used the evidence ratio to rank each model and followed Anderson [S39] for model selection. Once we selected the best-fit model, we generated confidence intervals around each parameter by averaging the second and third largest and second and third smallest values across the 100 bootstrap replicates to obtain the upper and lower bounds, respectively.

We noticed a bias in the SFS due to ancestral allele misspecification in the last two categories (the most ancestral states). To avoid inflated likelihood scores due to ancestral state misspecification, we removed these last two categories and re-estimated the likelihood scores (Table S4). These two last categories are unlikely to have any effect in the demographic models tested. In addition, we estimated likelihood scores with the folded site frequency spectrum with the same 23-allele projection (Table S4; Figure. S4). In both approaches the three-epoch model was the best (Table S4).

Script used to test models of demographic evolution in δaδi

```
"""
Single population demographic models.
"""
import numpy

from dadi import Numerics, PhiManip, Integration
from dadi.Spectrum_mod import Spectrum

def neutral (notused, ns, pts_l):
    """
```

```
    Standard neutral model.

    ns = (n1,)

    n1: Number of samples in resulting Spectrum
    pts_l: Number of grid points to use in integration.
    """
    xx = Numerics.default_grid(pts_l)
    phi = PhiManip.phi_1D(xx)

    fs = Spectrum.from_phi(phi, ns, (xx,))
    return fs

def two_epoch (params, ns, pts_l):
    """
    Instantaneous size change some time ago.

    params = (nu,T)
    ns = (n1,)

    nu: Ratio of contemporary to ancient population size
    T: Time in the past at which size change happened (in units of 2*Na
       generations)
    n1: Number of samples in resulting Spectrum
    pts_l: Number of grid points to use in integration.
    """
    nu,T = params

    xx = Numerics.default_grid(pts_l)
    phi = PhiManip.phi_1D(xx)

    phi = Integration.one_pop(phi, xx, T, nu)

    fs = Spectrum.from_phi(phi, ns, (xx,))
    return fs

def growth (params, ns, pts_l):
    """
    Exponential growth beginning some time ago.

    params = (nu,T)
    ns = (n1,)

    nu: Ratio of contemporary to ancient population size
    T: Time in the past at which growth began (in units of 2*Na
       generations)
    n1: Number of samples in resulting Spectrum
    pts_l: Number of grid points to use in integration.
    """
    nu,T = params

    xx = Numerics.default_grid(pts_l)
    phi = PhiManip.phi_1D(xx)

    nu_func = lambda t: numpy.exp(numpy.log(nu) * t/T)
    phi = Integration.one_pop(phi, xx, T, nu_func)

    fs = Spectrum.from_phi(phi, ns, (xx,))
```

```
    return fs

def bottlegrowth (params, ns, pts_l):
    """
    Instantanous size change followed by exponential growth.

    params = (nuB,nuF,T)
    ns = (n1,)

    nuB: Ratio of population size after instantanous change to ancient
         population size
    nuF: Ratio of contemporary to ancient population size
    T: Time in the past at which instantaneous change happened and growth began
       (in units of 2*Na generations)
    n1: Number of samples in resulting Spectrum
    pts_l: Number of grid points to use in integration.
    """
    nuB,nuF,T = params

    xx = Numerics.default_grid(pts_l)
    phi = PhiManip.phi_1D(xx)

    nu_func = lambda t: nuB*numpy.exp(numpy.log(nuF/nuB) * t/T)
    phi = Integration.one_pop(phi, xx, T, nu_func)

    fs = Spectrum.from_phi(phi, ns, (xx,))
    return fs

def three_epoch (params, ns, pts_l):
    """
    params = (nuB,nuF,TB,TF)
    ns = (n1,)

    nuB: Ratio of bottleneck population size to ancient pop size
    nuF: Ratio of contemporary to ancient pop size
    TB: Length of bottleneck (in units of 2*Na generations)
    TF: Time since bottleneck recovery (in units of 2*Na generations)

    n1: Number of samples in resulting Spectrum
    pts_l: Number of grid points to use in integration.
    """
    nuB,nuF,TB,TF = params

    xx = Numerics.default_grid(pts_l)
    phi = PhiManip.phi_1D(xx)

    phi = Integration.one_pop(phi, xx, TB, nuB)
    phi = Integration.one_pop(phi, xx, TF, nuF)

    fs = Spectrum.from_phi(phi, ns, (xx,))
    return fs
```

Supplemental References

S1. Babcock, R. (1991). Comparative demography of three species of scleractinian corals using age- and size–dependant classifications. Ecol. Monogr. *61*, 225–244.

S2. Voolstra, C., Sunagawa, S., Matz, M., Bayer, T., Aranda, M., Buschiazzo, E., DeSalvo, M., Lindquist, E., Szmant, A., Coffroth, M., et al. (2011). Rapid evolution of coral proteins responsible for interaction with the environment. PLoS ONE *6*, e20392.

S3. Prada, C., M. B. DeBiasse, J. E. Neigel, B. Yednock, J. L. Stake, Z. H. Forsman, I. B. Baums, and Hellberg, M.E. (2014). Genetic species delineation among branching Caribbean *Porites* corals. Coral Reefs *33*, 1019-1030.

S4. Budd, A.F., and Klaus, J.S. (2001). The origin and early evolution of the *Montastraea* "*annularis*" species complex (Anthozoa: Scleractinia). J. Paleontol. *75*, 527-545.

S5. Budd, A.F., and Klaus, J.S. (2008). Early evolution of the Montastraea "annularis" species complex (Anthozoa: Scleractinia): Evidence from the Mio-Pliocene of the Dominican Republic. In Evolutionary Stasis and Change in the Dominican Republic, R.H. Nehm and A.F. Budd, eds. (New York: Neogene, Springer), pp. 85-124.

S6. Sun, J.X., and A. Helgason, G.M., S.S. Ebenesersdóttir, H. Li, S. Mallick, S. Gnerre, N. Patterson, A. Kong, D. Reich, K. Stefansson (2012). A direct characterization of human mutation based on microsatellites. Nat. Genet. *44*, 1161–1165.

S7. Whittaker, J., Harbord, R., Boxall, N., Mackay, I., Dawson, G., and Sibly, R. (2003). Likelihood-based estimation of microsatellite mutation rates. Genetics *164*, 781-787.

S8. Schlötterer, C., Ritter, R., Harr, B., and Brem, G. (1998). High mutation rate of a long microsatellite allele in *Drosophila melanogaster* provides evidence for allele-specific mutation rates. Mol. Biol. Evol. *15*, 1269–1274.

S9. Johnson, K.G., Budd, A.F., Klaus, J.S., and McNeill, D.F. (2008). The impact of fossils from the northern Dominican Republic on origination estimates for Miocene and Pliocene Caribbean reef corals. In Evolutionary Stasis and Change in the Dominican Republic Neogene, R.H. Nehm and A.F. Budd, eds. (New York: Springer), pp. 253-280.

S10. Budd, A.F., Klaus, J.S., and Johnson, K.G. (2011). Cenozoic diversification and extinction patterns in Caribbean reef corals: A review. Paleontological Society Papers *17*, 79-94.

S11. McNeill, D.F., Coates, A.G., Budd, A.F., and Borne, P.F. (2000). Integrated paleontologic and paleomagnetic stratigraphy of the upper Neogene deposits around Limon, Costa Rica: A coastal emergence record of the Central American Isthmus. Geological Society of America Bulletin *112*, 963-981.

S12. McNeill, D.F., Klaus, J.S., Budd, A.F., Lutz, B., and Ishman, S. (2012). Late Neogene Chronology and Sequence Stratigraphy of Mixed Carbonate-Siliciclastic Deposits of the Cibao Basin, Dominican Republic. Geological Society of America Bulletin *124*, 35-58.

S13. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina Sequence Data. Bioinformatics *30*, 2114-2120.

S14. Joshi, N., and Fass, J.N. (2011). Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33)

S15. Andrews, S., Krueger, F., Seconds-Pichon, A., Biggins, F., and Wingett, S. (2014). FastQC. A quality control tool for high throughput sequence data. Babraham Bioinformatics.

S16. Ilie, L., and Molnar, M. (2013). RACER: rapid and accurate correction of errors in reads. Bioinformatics *29*, 2490-2493.

S17. Chikhi, R., and Medvedev, P. (2013). Informed and Automated k-Mer Size Selection for Genome Assembly. HiTSeq.

S18. Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. *29*, 644-652.

S19. Boetzer, M., C.V. Henkel, H.J. Jansen, D. Butler, and Pirovano, W. (2011). Scaffolding pre-assembled contigs using SSPACE. Bioinformatics *27*, 578–579.

S20. Parra, G., Bradnam, K., and Korf, I. (2007). CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. Bioinformatics *23*, 1061-1067.

S21. Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. Bioinformatics *29*, 1072-1075.

S22. Levitan, D., Fogarty, N., Jara, J., Lotterhos, K., and Knowlton, N. (2011). Genetic, spatial, and temporal components to precise spawning synchrony in reef building corals of the *Montastraea annularis* species complex. Evolution *65*, 1254–1270.

S23. Levitan, D.R. (2004). Density-dependent sexual selection in external fertilizers: variances in male and female fertilization success along the continuum from sperm limitation to sexual conflict in the sea urchin *Strongylocentrotus franciscanus*. Am. Nat. *164*, 298–309.

S24. Li, H., and Durbin, R. (2011). Inference of human population history from individual whole-genome sequences. Natue *475*, 493–496.
S25. Sheehan, S., Harris, K., and Song, Y.S. (2013). Estimating Variable Effective Population Sizes from Multiple Genomes: A Sequentially Markov Conditional Sampling Distribution Approach. Genetics *194*, 647–662.
S26. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler Transform. Bioinformatics *25*, 1754-1760.
S27. McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, et al. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. *20*, 1297–1303.
S28. Freedman, A.H., Gronau, I., Schweizer, R.M., Ortega-Del Vecchyo, D., Han, E., Silva, P.M., Galaverni, M., Fan, Z., Marx, P., Lorente-Galdos, B., et al. (2014). Genome Sequencing Highlights the Dynamic Early History of Dogs. PLoS Genet. *10*, e1004016.
S29. Brandvain, Y., Kenney, A., Flagel, L., Coop, G., and Sweigart, A. (2014). Speciation and Introgression between *Mimulus nasutus* and *Mimulus guttatus*. PLoS Genet *10*, e1004410.
S30. de Iorio, M., and Griffiths, R. (2004). Importance sampling on coalescent histories. Advanced in Applied Probabilities *36*, 417–433.
S31. de Iorio, M., Griffiths, R., Leblois, R., and Rousset, F. (2005). Stepwise mutation likelihood computation by sequential importance sampling in subdivided population models. Theor. Popul. Biol. *68*, 41–53.
S32. Leblois, R., Pudlo, P., Néron, J., Bertaux, F., Beeravolu, C.R., Vitalis, R., and Rousset, F. (2014). Maximum-likelihood inference of population size contractions from microsatellite data. Mol. Biol. Evol. *31*, 2805–2823.
S33. Porto-Hannes, I., Zubillaga, A.L., Shearer, T.L., Bastidas, C., Salazar, C., Coffroth, M.A., and Szmant, A.M. (2014). Population structure of the corals Orbicella faveolata and Acropora palmata in the Mesoamerican Barrier Reef System with comparisons over Caribbean basin-wide spatial scale. Mar. Biol. *162*, 81-98.
S34. Foster, N., Paris, C., Kool, J., Baums, I., Stevens, J., Sanchez, J., Bastidas, C., Agudelo, C., Bush, P., Day, O., et al. (2012). Connectivity of Caribbean coral populations: complementary insights from empirical and modelled gene flow. Mol. Ecol.
S35. Anderson, D.A., Walz, M.E., Weil, E., Tonellato, P., and Smith, M.C. (2015). A transcriptome resource for the coral, Orbicella faveolata (Scleractinia-Meruliniidae) – an emerging model of coral innate immunity. PeerJ PrePrints *3*, e1508.
S36. Pinzón, J., Kamel, B., Burge, C., Harvell, C., Medina, M., Weil, E., and Mydlarz, L. (2015). Whole transcriptome analysis reveals changes in expression of immune related genes during and after bleaching in a reef-building coral. Royal Society Open Science *2*, 140214
S37. DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M., et al. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet *43*, 491-498.
S38. Gutenkunst, R., Hernandez, R., Williams, S., and Bustamante, C. (2009). Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. PLoS Genet. *5*, e1000695.
S39. Anderson, D.R. (2008). Model Based Inference in the Life Sciences, (New York: Springer).