Title: Seascape genomics as a new tool to empower coral reef conservation strategies: an example on north-western Pacific *Acropora digitifera*.

Running Title: Seascape genomics of north-western Pacific *Acropora digitifera*

List of Authors: Oliver Selmoni[1], Estelle Rochat[1], Gael Lecellier[2,3], Veronique Berteaux-Lecellier[2], Stéphane Joost[1].

Institutional Affiliations:

[1] Laboratory of Geographic Information Systems (LASIG), Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland

[2] UMR250/9220 ENTROPIE IRD-CNRS-UR, Labex CORAIL, Promenade Roger-Laroque, Noumea cedex, New Caledonia

[3] UVSQ, Université de Paris-Saclay, 45 avenue des Etats-Unis, Versailles Cedex, France

Contact Information:

Oliver Selmoni: oliver.selmoni@epfl.ch , +41216935787

Abstract

Coral reefs are suffering a major decline due to the environmental constraints imposed by climate change. Over the last 20 years, three major coral bleaching events occurred in concomitance of anomalous heat waves, provoking a severe loss of coral world-wide. Recent works, however, reported of reefs at which recurrent bleaching events resulted in increased resistance to thermal stress, suggesting that adaptation might have occurred. The conservation strategies for preserving the reefs, as they are conceived now, cannot cope with global climatic shifts. In this regard, researchers advocated the set-up of a preservation framework to reinforce coral adaptive potential. The main obstacle to this implementation is that studies on coral adaption are usually hard to generalize at the scale of a reef system. In this work, we combined a seascape genomics and connectivity analysis to characterize the adaptive potential of a flagship coral species of the Ryukyu Archipelago (Japan). By associating genotype frequencies with descriptors of historical environmental conditions, we discovered six genomic regions hosting polymorphisms that might promote resistance against thermal stress. Remarkably, annotations of genes in these regions are consistent with molecular roles in heat responses. Furthermore, we confronted genetic distances between reefs with their spatial separation according to sea currents to predict connectivity patterns across the region. The cross-talk between the results of these analyses portrayed the adaptive potential of this population: we were able to identify reefs carrying potential adaptive traits and to understand how they disperse to neighbouring reefs. This information was summarized by objective, quantifiable and mappable indexes covering the whole region that can be extremely useful for prioritization in conservation plans. This framework is transferable to any coral species and on any reef system and therefore represents a valuable tool for empowering preservation efforts dedicated to the protection of coral reef in warming oceans.

Keywords (6-10)

Coral reef – Local adaptation – climate change – seascape genomics – coral bleaching – conservation genomics – Acropora digitifera – Ryukyu Archipelago

Introduction

Coral reefs are suffering a severe decline due to the effects of climate change (Hughes et al., 2017). Loss of reef is already showing catastrophic consequences for marine wildlife depending on these structures (Pratchett, Thompson, Hoey, Cowman, & Wilson, 2018) and disastrous aftermaths are expected for human economies as well (Moberg & Folke, 1999). One of the major threat to the persistence of this ecosystem is coral bleaching (Bellwood, Hughes, Folke, & Nyström, 2004): a physiological response induced by environmental stress that provokes hard-skeleton coral, the cornerstone of the reef, to separate from the symbiotic microbial algae living within their cells (Mydlarz, McGinty, & Harvell, 2010). These algae, called Symbiodiniacea (LaJeunesse et al., 2018), are crucial for the host sustainment and therefore a persistent bleached state leads coral to death (Mydlarz et al., 2010). Three mass bleaching events occurred over the last 20 years in concomitance with anomalously high sea water temperatures, reason why thermal stress is now considered the main driver of this phenomenon (Hughes et al., 2017). These events struck world-wide and lead to local losses of coral cover up to 50% (Hughes et al., 2018, 2017). Besides, coral bleaching was found to be associated with other stressors linked to human activity, such as ocean acidification, water eutrophication, sedimentation and overfishing (Anthony, Kline, Diaz-Pulido, Dove, & Hoegh-Guldberg, 2008; Ateweberhan et al., 2013; Maina, Venus, McClanahan, & Ateweberhan, 2008).

Conservation efforts facing the plague of coral bleaching aim to restore reefs that underwent severe losses and to limit the impact of future damages (Baums, 2008; Bellwood et al., 2004; Young, Schopmeyer, & Lirman, 2012). For these purposes, two tools are nowadays usually used: the establishment of marine protected areas (MPAs) and the set-up of coral nurseries (Baums, 2008; Bellwood et al., 2004; Young et al., 2012). MPAs are designed zones in which human access and activities are severely restricted in order to alleviate the effects of local

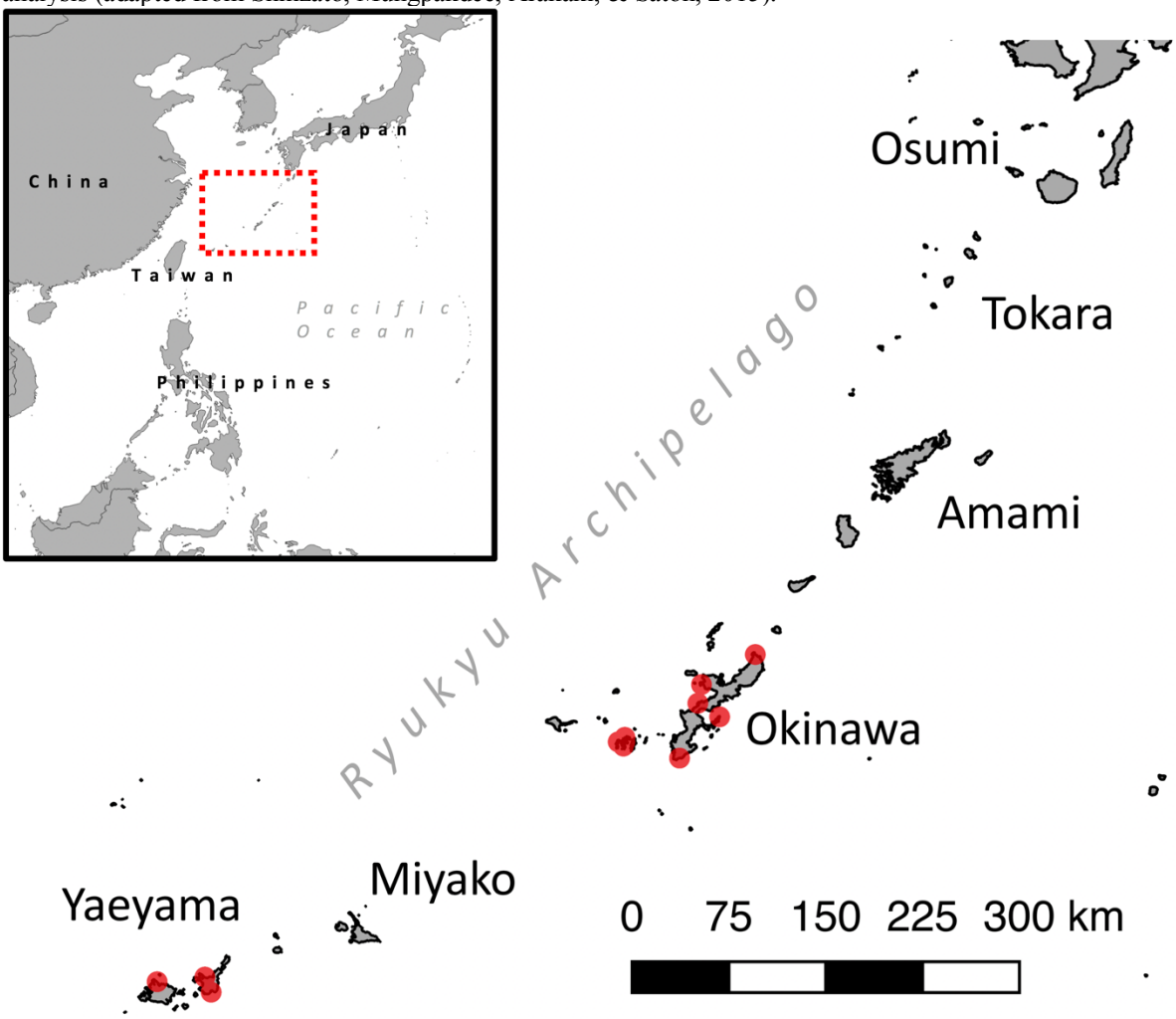
anthropogenic stressors (Lester et al., 2009). Coral nurseries are underwater gardens of transplanted colonies that can then be employed to restore damaged reefs (Baums, 2008; Young et al., 2012). In both cases, researchers advocated the use of design systems accounting for demographic connectivity so that the position of a conservation measure can promote resistance and resilience even on neighbouring sites (Baums, 2008; Krueck et al., 2017; Lukoschek, Riginos, & van Oppen, 2016; Palumbi, 2003; Shanks, Grantham, & Carr, 2003). Even though beneficial effects of this kind of conservation policies have been observed worldwide (Cinner et al., 2016; Rodgers et al., 2017; Selig & Bruno, 2010), these solutions do not confer resistance against the temperature oscillations associated with the last mass bleaching events (Baums, 2008; Hughes et al., 2017). Coral reefs that experienced past thermal stress were found to be more resistant to subsequent heat waves (Hughes et al., 2019; Krueger et al., 2017; Penin, Vidal-Dupiol, & Adjeroud, 2013; Thompson & van Woesik, 2009) but to date this information is neglected in conservation actions (Baums, 2008; Maina, McClanahan, Venus, Ateweberhan, & Madin, 2011). There is an urgent need to understand whether these observations are due to evolutionary processes and, if so, to figure out how the underlying adaptive potential could be included in predictions of climate change responses and in conservation plans (Baums, 2008; Logan, Dunne, Eakin, & Donner, 2014; Maina et al., 2011; van Oppen, Oliver, Putnam, & Gates, 2015).

To this end, seascape genomics tools are likely to play an important role. Seascape genomics is the marine counterpart of Landscape genomics, a branch of population genomics that investigates adaptive potential in field-based experiments (S. Joost et al., 2007; Balkenhol et al., 2017; Rellstab, Gugerli, Eckert, Hancock, & Holderegger, 2015). Samples are collected across a landscape and then genotyped using next-generation-sequencing techniques, providing thousands of genetic variants (Rellstab et al., 2015). Simultaneously, the environmental characterization of the study area is carried out mainly based on remote-

98 sensing data describing specific local climatic conditions (Rellstab et al., 2015). Genomics
99 and environmental information are then combined in order to detect genetic polymorphisms
100 associated with peculiar conditions (*i.e.,* potentially adaptive traits; Rellstab et al., 2015).
101 This approach has already been applied to many land species and is increasingly used to
102 analyse marine animals in what is referred to as *seascape genomics* (exhaustively reviewed in
103 Riginos, Crandall, Liggins, Bongaerts, & Treml, 2016). As far as we know, no seascape
104 genomics experiment was applied to reef corals yet. In fact, adaptation of these species has
105 been mostly studied by performing transplantation assays followed by conditioning in
106 aquarium, a time and resource-demanding approach that often focuses on a couple of
107 contrasting reef (Howells, Berkelmans, van Oppen, Willis, & Bay, 2013; Krueger et al.,
108 2017; Palumbi, Barshis, Traylor-Knowles, & Bay, 2014; Sampayo et al., 2016; Ziegler,
109 Seneca, Yum, Palumbi, & Voolstra, 2017). Genotype-environment associations studies have
110 also been conducted on corals but they focused on a limited number of markers (Lundgren,
111 Vera, Peplow, Manel, & van Oppen, 2013) or on a restricted number of locations (Bay &
112 Palumbi, 2014; Thomas, Kennington, Evans, Kendrick, & Stat, 2017). Unlike these works, a
113 seascape genomics approach should cover ecologically meaningful spatial scales, allowing to
114 distinguish pressures of different climatic conditions and to account for confounding effects
115 of demographic processes (Balkenhol et al., 2017). Consequently, the main (adaptive) and
116 collateral (connectivity) information generated by this method provides a valuable
117 contribution for preservation strategies (Baums, 2008; Krueck et al., 2017; Palumbi, 2003).
118 In this work, we applied a seascape genomics framework to detect coral reefs carrying
119 potential adaptive traits to show how conservation policies could implement such findings.
120 Our study focuses on *Acropora digitifera* of the Ryukyu Archipelago in Japan (Fig. 1), an
121 emblematic species of the Indo-Pacific and flagship organism for studies on corals genomics
122 (Shinzato et al., 2011). We first analysed the convergence between genomic and

environmental information i) to detect loci potentially conferring a selective advantage and ii) to develop a model describing connectivity patterns. Next, we took advantage of these findings to indicate which reefs were more likely to be carrying adaptive traits and to evaluate how well such sites are interconnected with the rest of the reef system. Finally, we propose an approach to implement the results obtained in conservation planning. In general, our work provides tools at the interface of conservation genomics and environmental sciences likely to empower preservation strategies for coral reefs.

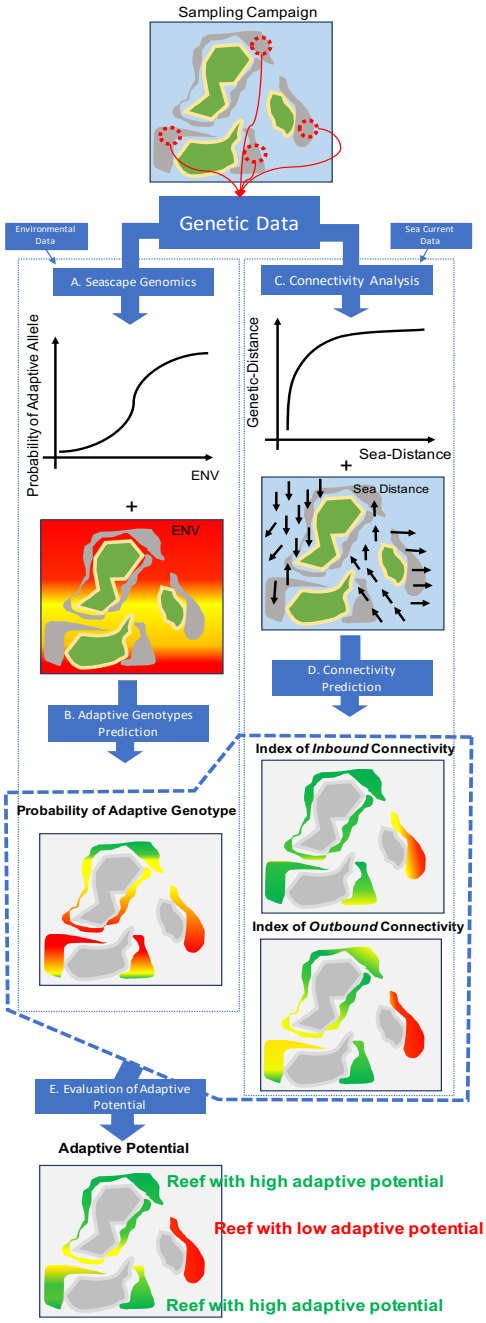
Figure 1. The Study area. The Ryukyu Archipelago extends for more than 1,000 km in the north-western Pacific Ocean. The red circles display the sites where samples were collected for the seascape genomics analysis (adapted from Shinzato, Mungpakdee, Arakaki, & Satoh, 2015).



Materials and methods

The input genomics data used in this paper comes from a published work from Shinzato *et al.* (2015) describing the population structure of *A. digitifera* of the Ryuyku Archipelago. Our framework is structured in two axes of analysis and prediction, one focusing on the presence of adaptive traits (seascape genomics) and the other on connectivity (Fig. 2). The seascape genomics analysis (Fig. 2A) crosses genomics data with environmental information at sampling sites to discover potentially adaptive genotypes. The model describing this relationship is then used to predict, at the scale of the whole study area, the probability of presence of adaptive genotypes (Fig. 2B). In the connectivity study (Fig. 2C), we built a model describing how sea-distances between sites (computed out of remote sensing data) correspond to genetic-distances. This model is then used to predict connectivity at scale of the whole study area (Fig. 2D). Finally, the predictions on where the adaptive traits are more likely to appear and on how the reef system is interconnected allow to assess the adaptive potential across the whole study area (Fig. 2E).

Figure 2. Workflow of the methods. The starting point for the analysis is the generation of genetic data describing the genotypes observed at different sampling locations (in this example, 4 sampling sites). In the seascape genomics analysis (A) this data is crossed with environmental information to discover genotypes whose frequencies associate with specific climatic conditions (*ENV*). Such genotypes are defined a potentially adaptive. The model describing this link is then applied to environmental data at the scale of the whole reef system (B), to predict the probability of presence of the adaptive genotype (green: high probability, red: low probability). The genetic data is also crossed with sea-current information to build a connectivity model (C) describing how sea-distances between sampling sites correspond to genetic-distances. This model is fit with sea-distances between all the reefs of the study area to predict (D) patterns of connectivity from (*outbound*) and to (*inbound*) each reef (green: high connectivity, red: low connectivity). Finally, predictions on the presence of adaptive genotypes and connectivity patterns are combined to assess the adaptive potential across the study area (E): reefs that are connected with sites likely to carry the adaptive traits will be have a high adaptive potential (green), and inversely for those isolated (red).



149

150 *Genomic dataset*

151 The genomic data used come from a publicly available dataset consisting of 155 geo-
152 referenced colonies of *A. digitifera* from 12 sampling locations (13±5 colonies per site) of the
153 Ryukyu Archipelago in Japan (Fig. 1; Bioproject Accession PRJDB4188). These samples
154 were sequenced using a Whole-Genome Sequencing approach in the+ scope of a population
155 genomics study. Details on how samples were collected and processed for genomic analysis
156 can be found in Shinzato et al., (2015).

157 Genomics data were processed using the Genome Analysis Toolkit (GATK; McKenna et al.,
158 2010) framework following the recommended pipeline with modifications necessary to cope
159 with the absence of reliable databases of known variants for this species (Van der Auwera et
160 al., 2013). In short, the *A. digitifera* reference genome (v. 1.1, GenBank accession:
161 GCA_000222465.2; Shinzato et al., 2011) was indexed using bwa (v. 0.7.5a, Li & Durbin,
162 2009), samtools (v. 1.9, Heng Li et al., 2009) and picardtools (v. 1.95,
163 http://broadinstitute.github.io/picard) and raw sequencing reads were aligned using the bwa
164 *mem* algorithm. The resulting alignments were sorted, marked for duplicate reads, modified
165 for read-group headers and indexed using picard-tools. Next, each alignment underwent an
166 independent variant discovery using the GATK HaplotypeCaller tool (using the ERC mode
167 and setting the --minPruning flag to 10) and genotypes were then jointly called by the GATK
168 GenotypeGVCFs tool in random batches of 18 samples to match our computational power
169 (18 CPUs). The variant-calling matrices of the different batches were then joined and filtered
170 in order to keep only bi-allelic Single Nucleotide Polymorphisms (SNPs) using the GATK
171 CombineVariants and SelectVariants tools, respectively. This resulted in a raw genotype
172 matrix counting ~1.2 M of SNPs. Subsequently, we used the GATK VariantAnnotator tool to
173 annotate variants for Quality-by-depth and filtered for this value (<2), read coverage (>5 and

174 <100 within a sample), minor allele frequency (>0.05), major genotype frequency (<0.95)
175 and missing rate of both individuals and SNPs (<0.1) using the GATK VariantFiltrationTool
176 and custom scripts in the R environment (v. 3.5.1, R Core Team, 2016). This pipeline
177 produced the filtered genotype matrix consisting of 136 individuals and 15,516 SNPs.
178
179 *Environmental data*
180 Eleven georeferenced datasets describing atmospheric and seawater conditions were retrieved
181 from publicly available resources (EU Copernicus Marine Service, 2017; NASA, 2016;
182 National Oceanic and Atmospheric Administration, 2017; Tab S1). These remote sensing
183 derived datasets provided monthly or daily variables measured across the whole study area,
184 with a spatial resolution ranging from 25 km to 4 km (Tab S1). We processed these variables
185 in the R environment using the *raster* package (v. 2.8, Hijmans, 2016) to compute monthly
186 averages and standard deviations for the whole studied period. Furthermore, Sea Surface
187 Temperature (SST) measurements were used to compute a Degree Heat Week (DHW)
188 frequency index, representing the % of days during which the bi-weekly accumulated heat
189 stress exceeded the 4 °C (Liu, Strong, & Skirving, 2003; Logan et al., 2014). SST and sea
190 surface salinity records were combined to produce estimates of seawater pH (Covington &
191 Whitfield, 1988), dissolved inorganic carbon (Loukos, Vivier, Murphy, Harrison, & Le
192 Quéré, 2000) and alkalinity (Lee et al., 2006). Bathymetry data (Ryan et al., 2009) were used
193 to retrieve the depth at sampling locations. Finally, population density data (CIESIN
194 Columbia University, 2010) were averaged in a 50 km buffer area to produce a surrogate-
195 variable for anthropogenic pressure (Welle, Small, Doney, & Azevedo, 2017).
196 We used the geographic coordinates associated with each sample to characterize the
197 environmental conditions using the QGIS point sampling tool (v. 2.18.25, QGIS
198 development team, 2009). For predictive step of our study demanding to work at the scale of

199 the whole reef system (Fig. 2C), we retrieved the shapes of the reefs of the region (UNEP-
200 WCMC, WorldFish-Center, WRI, & TNC, 2010) and reported them into a regular grid (cell
201 size of 5x5 km) using QGIS. In this case, we also included reefs from the neighbouring
202 regions (Taiwan and Philippines, Fig. S7) to avoid border-effects in computations.
203 Environmental conditions were assigned to each reef-cell using average function of the QGIS
204 zonal statistics tool.
205
206 *Seascape genomics*
207 We performed the genotype-environment association analysis using the logistic regression
208 method implemented within the SamBada software (v. 0.7; Stucki et al., 2017), customized to
209 fasten multivariate models computation in the Python environment (v. 2.7; Python Software
210 Foundation, 2018) using the *pandas* (v. 0.23.4; McKinney, 2010) and *statsmodels* (v. 0.9;
211 Seabold & Perktold, 2010) libraries. The SamBada approach allows to include proxy
212 variables for genetic structure in the analysis in order to avoid a possible confounding effect
213 (pattern of neutral genetic variation mimicking signals of adaptation to the local
214 environment). Here we performed a discriminant analysis of principal components (DAPC)
215 on the SNPs genotype matrix using the R package *adegenet* (v. 2.1.1; Jombart, 2008). This
216 procedure highlighted a main separation between two groups of samples along the first
217 discriminant function (Fig. S1). The latter was therefore used as co-variable in association
218 models. We then ran the pipeline for the discovery of adaptive signals described hereafter and
219 summarized in a schematic view in Fig. S2. In order to speed up calculations, the 315
220 environmental variables were grouped into 29 clusters of highly correlated (|R|>0.7)
221 descriptors in the R environment (Fig. S2A). For each of these groups, one variable was
222 randomly selected to build logistic models against the three genotypes of each SNP (Fig.
223 S2B). In total, the SamBada instance evaluated 1,349,892 association-models (29

224 environmental groups vs. 3 x 15,516 SNPs; Fig. S2C, S2D) that were subsequently analysed
225 in the R environment. For each association-model related to the same environmental variable,
226 p-values of G-scores (G) and Wald-scores (W) were corrected for multiple testing using the
227 R *q-value* package (v. 2.14, Storey, 2003). Association-models scoring q<0.001 for both
228 statistics were deemed significant (Fig. S2E). If a SNP was found in more than one
229 significant association, only the best model (according to the value of G) was kept. For each
230 significant association retained, we then calculated all the models built with the involved
231 genotype vs. the other environmental variables from the same cluster of correlated descriptors
232 and we looked for the best association model (according to G; Fig. S2F). Finally, we
233 visualized the logistic regression of each significant association-model using the R *popbio*
234 library (v. 2.4.4; Stubben & Milligan, 2007) and manually checked the annotations around
235 significant SNPs on the NCBI genome browser (NCBI, 1988). We set the size of the search
236 window to ± 250 kbs since the gene(s) potentially linked to a mutation can lay up to
237 hundreds of kbs away (Brodie, Azaria, & Ofran, 2016; Visel, Rubin, & Pennacchio, 2009).
238 This window size was chosen because it corresponds approximately to the scaffold N50
239 statistics of the reference genome (*i.e.* half of the genome is contained within scaffolds of this
240 size or longer).
241
242 *Probability of Presence of Adaptive Genotypes*
243 The results of the seascape genomics analysis were then used to predict the probability of
244 presence of adaptive traits at the scale of the whole Ryukyu Archipelago (Fig. 2B). Indeed,
245 for each significant association between a genotype and an environmental variable, the
246 SamBada approach provides parameters of a logistic regression which links the probability of
247 occurrence of the genotype with the value of the environmental variable (Stucki et al., 2017).
248 Those logistic models can therefore be used to estimate the probability of presence of the

genetic variant for any value of the environmental variable (Stéphane Joost, 2006; Rochat, E., Leempoel, K., Vajana, E., Colli, L., Ajmone-Marsan, P., Joost, 2016). We assumed that potentially adaptive traits could reach any reef of the region since several studies reported high gene flow between distant islands of the Archipelago (Nakajima, Nishikawa, Iguchi, & Sakai, 2010; Nishikawa, 2008; Shinzato et al., 2015). Finally, these probabilities where combined in an average probability (the arithmetical mean of the five probabilities) of carrying adaptive traits (P_a) at each reef of the Ryukyu Archipelago.

Sea current data

Daily records of sea surface current were retrieved from publicly available databases (zonal and meridional surface velocities from the *global-reanalysis-phy-001-030* product; EU Copernicus Marine Service, 2017) and used to compute the direction and speed of currents in the R environment using the *raster* library. These day-by-day calculations were then piled-up to retrieve, for each pixel of the study area (~0.083°), the cumulated speed (the conductance) toward each of the eight neighbouring pixels. This information was used to calculate dispersal costs (the inverse of conductance) and was summarized in a transition matrix in the format of the R *gdistance* package (v. 1.2, van Etten, 2018). For the connectivity analysis (Fig. 2C), the transition matrix was used to calculated sea-distances (*i.e.* the least-cost path) between sampling sites of the genotyped colonies. For the connectivity predictions (Fig. 2D), we calculated the sea-distances between all the reefs of the study area, whose coordinates were the centroids of the reef-cells computed as described in the environmental variables section. Importantly, for every couple of reefs (for instance $reef_1$ and $reef_2$) two sea-distances were computed, one per direction (*i.e.* from $reef_1$ to $reef_2$ and from $reef_2$ to $reef_1$).

Connectivity analysis

274 Genomics data was processed by the means of a principal component analysis (PCA; Fig. S3)
275 on the filtered genotype matrix. Prior to PCA, in order to cope with missing genotypes in the
276 SNP matrix, we applied the *random* imputation method of the LEA R package (v.2.4.0;
277 Frichot & François, 2015). The values of the principal components (PCs) were used to
278 calculate the Euclidean distances between each couple of samples that were subsequently
279 averaged by site in order to obtain Genetic Distances (GDs) between reefs. Next, we built a
280 linear model (hereafter referred to as the connectivity model) to estimate GDs out of log-
281 transformed sea-distances and ran a t-test to check for the statistical significance of the
282 relationship. Since we wanted to consider the genetic distances within reef (*i.e.* sea-distance =
283 0) we added +0.005 to all the sea-distances prior to the log-transformation to avoid obtaining
284 non-definite values (*i.e.* the logarithm of zero). As a comparison, we also built a connectivity
285 model using log-transformed aerial-distances (*i.e.* Euclidean distance of coordinates) as
286 independent variable while maintaining GD as response variable. We then confronted models
287 based on sea- and aerial-distance by comparing the coefficients of determination (R2) and the
288 Akaike information criterion (AIC; Bozdogan, 1987).
289
290 *Connectivity Predictions*
291 We then used the connectivity model to predict GDs between any couple of reefs of Ryukyu
292 Archipelago. This was done by translating sea-distances in GDs using the connectivity model
293 trained during the connectivity analysis. Importantly, since there were two sea-distances for
294 every couple of reefs (*i.e.* from $reef_1$ to $reef_2$ and *vice versa*), two GDs were computed as
295 well.
296 Based on these predictions, we were able to calculate, for each reef-cell, two indices (Fig.
297 S4):

298 - *outbound connectivity* (OC; Fig. S4B): the % of cells reachable from the focal cell
299 within a GD-threshold (tGD).
300 - *inbound connectivity* (IC; Fig. S4C): % of cells that can reach the focal cell within a tGD.
301 Since these indices were conceived to display contrasts in connectivity among reefs of the
302 Ryukyu Archipelago, we looked for the tGD value maximizing their variances (Fig. S5). The
303 values corresponding to the 25%-quantile and the 60%-quantile of the GD distribution were
304 those boosting the differences in both connectivity indices. To avoid the artificial bias of
305 border effects, we preferred working at a more local scale and set tGD at the 25%-quantile of
306 the GD distribution (*i.e.* tGD=79.861).
307
308 *Evaluation of the adaptive potential*
309 The adaptive potential was evaluated by combining the predictions on the presence of
310 adaptive traits and on connectivity patterns across the reefs of Ryuyku Archipelago (Fig. 2E).
311 The previously computed average probability of carrying adaptive traits (P_a) is used to
312 identify the putatively adaptive reefs. These are the reefs approaching the maximal values of
313 P_a (P_a>0.6). We then used the GDs calculated during the connectivity analysis to retrieve, for
314 each reef-cell, the GD from the closest putatively adapted reef. This results in an index
315 describing the adaptive potential of every reef of the study area: the distance from the closest
316 adaptive reef.
317
318 *Data Access and Code Availability*
319 All the data and codes used in this article are publicly available on Zenodo
320 (DOI: 10.5281/zenodo.2643454).
321
322 **Results**

323 *Seascape Genomics*

324 We detected 15 significant genotype-environment association-models (A1-A15; q_G and
325 q_W<0.001; Tab.1), referring to genetic markers from 6 distinct scaffolds of the *A. digitifera*
326 reference genome. Among them, one association was related to SST variation in April, one to
327 changes in sea current velocity in May and 13 to DHW frequency (or to a highly correlated
328 variable such as variation in cloud fraction in August).

329 The SNP in the significant association-model A1 is related to SST standard deviation in April
330 and lies within a genomic window (± 250 kb) counting 18 predicted genes, 7 of which are
331 annotated with a known function (Box S1). Association-models A2 to A5 include four SNP
332 lying at two opposite sides of a long scaffold (~1.3 Mb). The SNPs located on one side
333 (involved in association-models A2 and A3) are associated with DHW frequency and are
334 surrounded by 17 predicted genes (of which 12 have a putative encoded protein with known
335 function; Box S2) while the SNPs located on the other side (association-models A4 and A5)
336 are related to Sea current velocity variance in May and are surrounded by twelve putative
337 genes (Box S4). All the remaining association-models (A6 to A15) involved only DHW.
338 Association-model A6 concerns a SNP surrounded by 14 putative genes, but only one is
339 annotated, the related predicted protein being *JNK1/MAPK8-associated membrane protein-*
340 *like* (NCBI accession: LOC107334815; Box S5). In association-model A7, the SNP of
341 interest lies within in a window showing eleven predicted genes (only two with a known
342 match by similarity; Box S6). Association-models A8 to A14 refer to seven SNPs present in
343 the same scaffold (~195 kb long). This region contains five putative genes of which 2
344 annotated by similarity (Box S6). In association-model A15, finally, the SNP associated with
345 DHW frequency is located on a ~75 kb scaffold that contains only one putative gene
346 predicted as *ATP-dependent DNA helicase RecQ-like* (LOC107328703; Box S7).

Table 1. Significant genotype-environment associations. The seascape genomics analysis using the SamBada method detected 15 significant association-models (q_G and q_W<0.001). This table shows, for each association-model, the genomic position (scaffold of the reference genome and relative situation), the genotype involved (0 and 2=homozygous, 1=heterozygous), the concerned environmental variables (in brackets is displayed a variable from the same group of correlated descriptors that resulted in a model with a higher G-score, if any), the q-values associated with the G-score (q_G) and the Wald-score (q_W), the number of predicted genes in a ± 250 kb window around the SNP (with the number of SNPs annotated by similarity with a known gene, in brackets) and the annotations from the reference genome that are in the neighbourhood (± 250 kb) of the SNP and that corroborate an adaptive role against the respective environmental variable.

Association ID	Scaffold ID	Position on Scaffold	GT	Environmental Variable	q_G	q_W	Number of Predicted Genes	Relevant Genome Annotations
A1	NW_015441190.1	582812	1	SST variance in April	1.5E-08	2.6E-04	18 (7)	*Helicase RecQ at 60 kbs*
A2	NW_015441080.1	208387	0	DHW frequency (Cloud Fraction variance in August)	5.8E-08	2.0E-04	17 (12)	*Cytochrome p450 at 150 kbs*
A3	NW_015441080.1	208400	0	DHW frequency	2.3E-09	6.4E-05	17 (12)	*Cytochrome p450 at 150 kbs*
A4	NW_015441080.1	963844	0	DHW frequency (Cloud Fraction variance in August)	4.0E-09	5.8E-04	12 (7)	
A5	NW_015441080.1	963851	0	Sea current velocity variance in May	1.5E-07	9.3E-04	12 (7)	
A6	NW_015441121.1	665651	0	DHW frequency	5.9E-08	1.6E-04	14 (1)	*MAPK8/JNK1 at 200 kbs*
A7	NW_015441282.1	27616	0	DHW frequency	5.5E-06	5.8E-04	7 (2)	*Helicase Q at 250 kbs*
A8	NW_015441785.1	16151	1	DHW frequency	6.1E-06	8.6E-04	5 (2)	*Helicase RecQ at 150kbs*
A9	NW_015441785.1	68415	1	DHW frequency	5.8E-08	1.1E-04	5 (2)	*Helicase RecQ at 90kbs*
A10	NW_015441785.1	68439	1	DHW frequency	3.1E-07	3.0E-04	5 (2)	*Helicase RecQ at 90kbs*
A11	NW_015441785.1	68444	1	DHW frequency	5.2E-06	8.6E-04	5 (2)	*Helicase RecQ at 90kbs*
A12	NW_015441785.1	68515	1	DHW frequency (Cloud Fraction variance in August)	5.9E-08	1.6E-04	5 (2)	*Helicase RecQ at 90kbs*
A13	NW_015441785.1	68521	1	DHW frequency	5.8E-08	2.0E-04	5 (2)	*Helicase RecQ at 90kbs*
A14	NW_015441785.1	68528	1	DHW frequency	6.9E-08	6.7E-04	5 (2)	*Helicase RecQ at 90kbs*
A15	NW_015442197.1	32233	0	DHW frequency	4.4E-06	8.6E-04	1 (1)	*Helicase RecQ at 20kbs*

Probability of Presence of Adaptive Genotypes

The significant association-models of the seascape genomics analysis were then used as starting point for predicting the probability of presence of adaptive genotypes at the scale of the whole Ryukyu Archipelago (Fig. 2B). These genotypes can be theoretically associated to any environmental descriptor but, since most of the putative adaptive genotypes were found to be associated with DHW frequency (Tab. 1), we designated this variable as the environmental pressure of interest (association models A3, A6, A7, A9 and A15, Tab. 1). The average of these probabilities (P_a) ranges from 0.40 to 0.62 (Fig. S7). Potentially adapted reefs (*i.e.* reefs with P_a>0.6) are located in the islands of Okinawa and Miyako, in the remote Japanese islands east of the Ryukyu Archipelago and in the Northern part of Philippines (Fig. S7).

Connectivity modelling

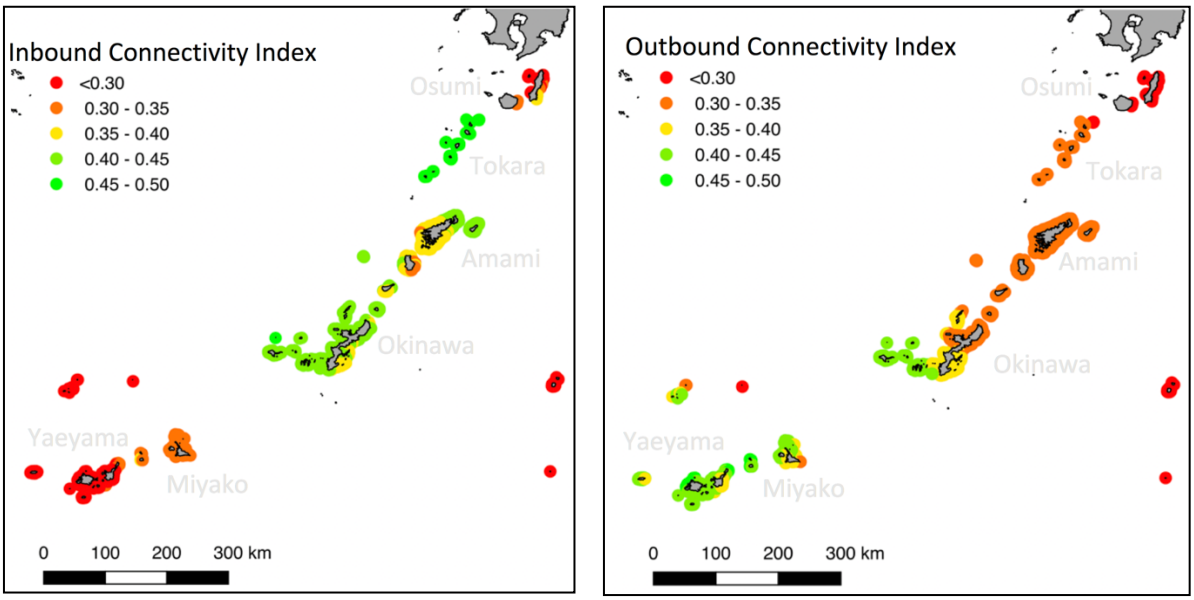
The two dispersal models respectively based on sea- and aerial-distances resulted in a significant association between genetic-distance (GD) and spatial separation ($p<10^{-9}$, Fig. S6). We found that the model based on sea-distances better explained the variation in GD (R2=0.57) than the aerial-distances model (R2=0.45). The AIC is lower for the model based on sea-distances (AIC=315) than for the aerial-distances model (AIC=343). The sea-distances dispersal model was therefore used for the following inbound and outbound connectivity predictions.

The reefs around the islands of Tokara were those with the highest values of Inbound Connectivity index (IC=0.49±0.01; Fig. 3a). The reefs in the central islands of Ryukyu Archipelago (Okinawa and Amami) showed higher values of the IC index (IC=0.41±0.03; Fig. 3a) as compared to those located at the edges of the reef system (Yaeyama in South, Osumi in the North) which showed the lowest values (IC=0.27±0.03). As regards the Outbound Connectivity index, reefs located in the southern part (Yaeyama and Miyako) are those predicted with the highest values (OC=0.42±0.03; Fig. 3b) while those in the northern part (Osumi and Tokara) display the lowest ones (OC=0.20±0.13).

377

Figure 3. Connectivity indices. The two maps show how dispersal to and from specific reefs is expected to vary across the Ryukyu Archipelago. Using the first quartile of the GD distribution as a threshold value (tGD=80.32), inbound connectivity (IC, a) was calculated as the % of cells that can reach the focal cell within tGD. Conversely, outbound connectivity (OC, b) represents the % of cells reachable from the focal cell within tGD.

a) Inbound Connectivity b) Outbound Connectivity



378

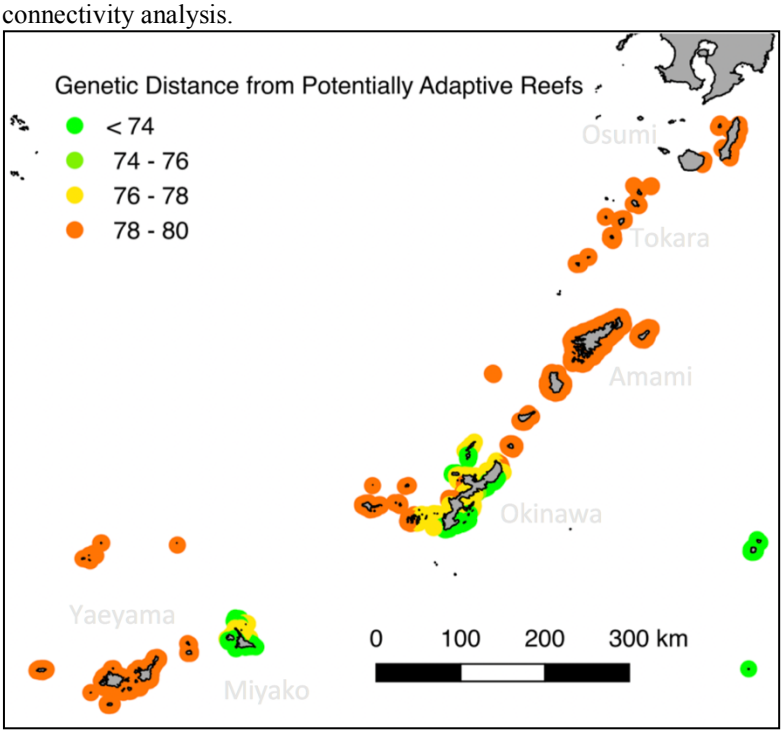
379 *Evaluation of the adaptive potential*

380 Reefs in the northern Islands of the Archipelago (Osumi, Tokara and Amami) and in the
381 southern Islands of Yaeyama were those predicted as the genetically farthest (GD=79.4±0.3
382 and GD=79.0±0.1, respectively) from reefs potentially adapted against thermal stress (P_a>0.6;
383 Fig. 4). Conversely, the southern islands of Miyako and those around Okinawa resulted as the
384 closest ones, despite a considerable variance (GD=73.6±3.1 and GD=76.4±2.8, respectively).

385

386

Figure 4. Distance from Potentially Adaptive Reefs. The map displays the predicted genetic-distance (GD) from the closest reef expected to carry adaptive genotypes against thermal stress. Adaptive sites are designed as those having a cumulative probability of carrying adaptive genotypes (P_a, Fig. S7) higher than 0.6. The GDs are computed out of sea-distances using the dispersal models developed in the connectivity analysis.



387

388 **Discussion**

389 *Adaptation to thermal stress*

390 Thermal stress is expected to be one of the major threat to coral reef survival and the quest
391 for adaptive traits is becoming of paramount importance (Baums, 2008; Logan et al., 2014;
392 Maina et al., 2011). In this work, the seascape genomics analysis of *A. digitifera* of the
393 Ryukyu Archipelago revealed the presence of 6 genomic regions hosting genetic variants that
394 might confer a selective advantage against thermal stress (Tab. 1). None of these SNPs
395 directly lay within a coding sequence of a putative gene, but this is rarely the case for
396 causative-mutations (Brodie et al., 2016). In fact, genetic variants in intergenic regions that
397 play a modulatory action on the expression of neighbouring genes are more frequent and can

398 influence loci at 1-2 Mb of distance (Brodie et al., 2016). The fragmentation of the reference
399 genome forced us to limit our search window to ±250 Kb around each SNP and still we found
400 several annotations corroborating a response to heat stress.
401 The SNP in association-model A6 is related to "JNK1/MAPK8-associated membrane
402 protein", the closest gene with a predicted function (Box S4). Mitogen-activated-protein
403 kinases (MAPKs) are proteins known to be involved in cellular responses to stress in a wide
404 range of taxa (Neupane, Nepal, Benson, Macarthur, & Piya, 2013) and the c-Jun-N-terminal
405 kinase (JNK) was experimentally proved to be activated under thermal stress in coral
406 *Stylopohora pistillata* (Courtial et al., 2017). Of the 26,060 predicted genes in the *A.*
407 *digitifera* genome (Shinzato et al., 2011), 35 are annotated as "mitogen-activated protein
408 kinase" and only one refers to the "JNK1/MAPK8-associated membrane protein". Another
409 annotation supports these findings and concerns gene-models predicted as "ATP-dependent
410 DNA helicase Q" and "RecQ" (A1, A7-15; Box S1, Box S5-7). These are proteins involved
411 in the DNA repairing mechanism to UV-light damage in prokaryotes (Courcelle & Hanawalt,
412 1999) and for which light-stress driven effects were observed in eukaryotic cells as well
413 (Sharma, Doherty, Brosh, & Jr, 2006). Annotations related to these proteins are not
414 uncommon (43 in the reference genome) but it is remarkable to find them around significant
415 SNPs from four distinct genomic regions. A third situation occurs with genes assigned to
416 cytochrome P450 (A2, A3; Box S2), a large family of proteins involved in cellular responses
417 to various types of stresses including heat shocks (Rosic, Pernice, Dunn, Dove, & Hoegh-
418 Guldberg, 2010; Sang, Ma, Qiu, Zhu, & Lei, 2012; Yampolsky et al., 2014) and for which
419 differential patterns of gene expression were observed in *Acropora hyacinthus* under distinct
420 thermal regimes (Palumbi et al., 2014). We detected 4 annotations related to these genes in
421 the proximity of one region with significant SNPs (models A2 and A3) but the relatively high

422 frequency of this gene-family (62 occurrences across the genome) invites to a cautious
423 interpretation.
424 Seascape/Landscape genomics studies are susceptible to high false discovery rates, especially
425 when neutral genetic structure is not accounted for (Rellstab et al., 2015). To take this
426 element into account, we processed data following a conservative pipeline and used models
427 explicitly integrating demographic processes. We also set a restrictive threshold to filter
428 significant association-models (*i.e.* $q<0.001$ in two distinct tests). Most of the significant
429 association-models found were linked to DHW, most probably because this variable
430 represents one of the main constraint to coral survival (Hughes et al., 2017). It is also possible
431 that the initial application of a sampling scheme adapted to seascape genomics (unlike the
432 one used by Shinzato et al. 2015 who did not consider environmental variability) would have
433 increased sensitivity to other types of adaptive signals (Riginos et al., 2016).
434 All the genotypes we found associated to thermal stress also displayed annotations suggesting
435 a functional role in cellular heat responses. Nonetheless, annotations analysis is only a first
436 step in the validation of the association-models detected. Ideally, further assays as reciprocal
437 transplantations (Palumbi et al., 2014), experimental conditioning (Krueger et al., 2017) and
438 molecular analysis (Courtial et al., 2017) could ascertain and help quantifying the link
439 between each genotype and the putative selective advantage it might confer. Nevertheless,
440 the high frequency of SNPs associated to helicase RecQ genes family together with their
441 location on 4 distinct scaffolds is strongly unlikely. This suggests that the maintenance of
442 ability to open the double strand of DNA for transcription, replication and repair is a key
443 point for *A. Digitifera* in Ryukyu archipelago to resist on the pressure applied by DHW.
444
445 *Patterns of connectivity*

446 Understanding connectivity patterns across a reef system is essential in order to plan effective
447 management strategies. This is usually carried out through the quantification of the dispersal
448 trajectories of larvae (Storlazzi, van Ormondt, Chen, & Elias, 2017) or through the study of
449 the population structure using genetic data (*e.g.* Lukoschek et al., 2016). The first approach
450 has the main drawback of being difficult and expensive because it requires an *a priori*
451 knowledge of the dispersal of the studied species (McCook et al., 2009). In population
452 genetics data, on the other hand, sampling often occurs at a few locations and the results are
453 therefore hard to generalize to a whole seascape extent (Beger et al., 2014). The method we
454 applied combines these two approaches. Firstly, genetic data are used to calculate genetic
455 distances between the sampled locations. The latter were then used to train dispersal models,
456 which could in turn predict genetic distances at the scale of the whole reef system.
457 In model training we found that sea-distances clearly represented a better indicator of spatial
458 separation, compared with aerial-distances (Fig. S6). This happens probably because coral
459 dispersal is driven by the water flow (Paris-Limouzy, 2011), which is strongly asymmetrical
460 in this region (north-east oriented) because of the Kuroshio current (Nishikawa, 2008). By
461 fitting this dispersal model with sea-distances among any reef of the region we were able to
462 evaluate the patterns of connectivity. Reefs in the extreme North (Osumi) face both *inbound*
463 and *outbound* isolation (Fig. 3), a situation that might lead to inbreeding depression (Keller &
464 Waller, 2002). Reefs in the extreme south-west of the system (Yaeyama) appear also
465 endangered: despite their proximity to Miyako, the strong north-eastward directionality of
466 sea-currents contrasts the arrival of recruits from the rest of the Archipelago.
467 The dispersal framework we propose here is potentially applicable to any coral species for
468 which geo-referenced genetic data are available and the indices associated can be extremely
469 valuable in management planning (Beger et al., 2014). However, a substantial part of GD

470 variation remains unexplained by our models and connectivity models developed at a finer
471 geographical and temporal scales might better describe the situation (Storlazzi et al., 2017).
472
473 *Conservation and Adaptive Potential*
474 Conservation strategies to preserve coral reefs need to account for and even to reinforce the
475 importance of the adaptive potential to resist to thermal stress (Baums, 2008; Logan et al.,
476 2014; Maina et al., 2008; van Oppen et al., 2015). In the perspective of its integration in
477 prioritization programs, adaptive potential needs to be objective and quantifiable (OECD,
478 2017). Previous works suggested to use climatic records (Logan et al., 2014; Maina et al.,
479 2011), but this approach neglects how organisms modulate a response in front of an
480 environmental constraint. The use of genetic data has also been advocated, but findings from
481 population genetics analyses are often hard to be extrapolated at large scales and therefore
482 includes in conservation plans (Beger et al., 2014). Furthermore, the established guidelines
483 (e.g. Beger et al., 2014) discuss how to reinforce conservation planning with genetic data (*i.e.*
484 a few loci), which is less informative for adaptive processes, compared to genomic data. Here
485 we showed how a combination seascape genomics and connectivity analysis makes it
486 possible to assess the adaptive potential at the scale of a reef system.
487 In the specific case of the Ryukyu Archipelago, we found that reef from Miyako and
488 Okinawa islands are those more likely to carry adaptive traits against thermal stress (Fig. 4).
489 Previous works reported indeed that adapted colonies from Okinawa might have resisted
490 during the 1998 and 2001 bleaching events (Van Woesik, Irikawa, & Loya, 2004). In
491 contrast, reefs in the northern part of the Archipelago (Amami, Tokara and Osumi) did not
492 experience the same thermal stress before and therefore are not predicted to show the same
493 adaptive capacities (Fig. 4).

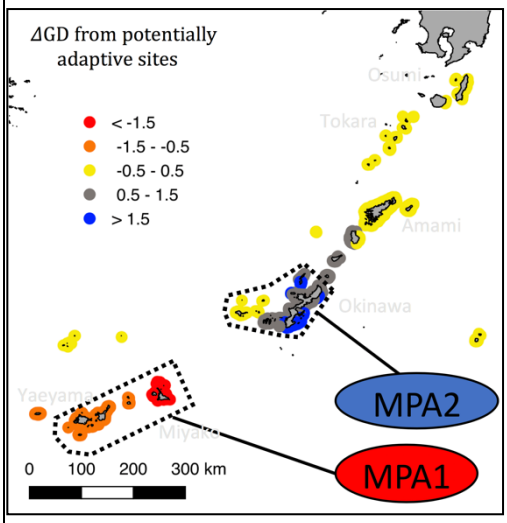
494 The indicators on adaptive potential and connectivity produced in this work provide valuable
495 descriptors of the state of the reef. Furthermore, they can be used for predicting the expected
496 impact of distinct preservation measures and support therefore spatial planning in
497 conservation (Box. 1). It is unknown whether the connectivity patterns and the adaptive
498 signals observed in *A. digitifera* are representative for those of other taxa. In this concern, we
499 advocate that future research should focus on a multi-species implementation of the methods
500 proposed here.
501
502 *Conclusions*
503 This study highlights the value of the seascape genomics approach to support the
504 conservation of corals. We applied this approach to a flagship coral species of Ryukyu
505 Archipelago and identified genetic variants that might underpin adaptation to thermal stress.
506 Coupling this information with a genetic analysis of connectivity enabled the evaluation of
507 the adaptive potential at the scale of the whole study area. The outputs of this analysis are
508 quantitative indices likely to support objective prioritization of reefs in conservation plans.
509 This framework is transferable to any coral species on any seascape and therefore constitutes
510 a useful conservation tool to take into account and evaluate the genomic adaptive potential of
511 coral reefs worldwide.
512

513

Box. 1 Adaptive potential in conservation scenarios.

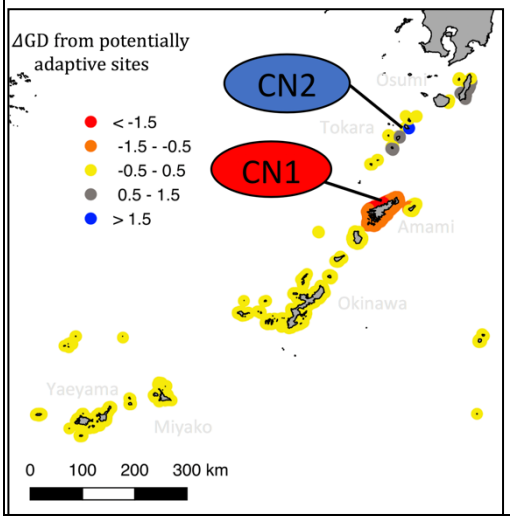
In this box we show how information on adaptive potential can be used to compare different conservation scenarios. We focus on two typical preservation initiatives, marine protected areas (MPAs) and coral nurseries (CNs). For each of them, we compare how the set-up of the measure at two different locations differentially drives the adaptive potential of the reef system.

The starting point for the predictions is the map showing the distribution of the probability of presence of adaptive genotypes (Fig. S7). We modify the values in this map to imitate the effects of preservation strategies. In the case of MPAs, we maintain potentially adaptive reefs (P_a>0.6) only at locations where the protected area is established. The underlying idea is to emulate the way MPAs reduces coral loss due to local stressors (Selig & Bruno, 2010). In the CNs scenarios, potentially adaptive reefs (P_a>0.6) are added where absent. Here we follow the good practices for this kind of measure that suggest to transplant adapted colonies at sites lacking them (Baums, 2008). After modifying the values of P_a, we re-calculate the genetic distance (GD) from adaptive sites (Fig. 4). This is the variable that we use to describe adaptive potential. We then compare differences in this value (ΔGD) between different conservation scenarios to investigate beneficial and negative effects.



Marine Protected Areas

We compare two MPAs established in Miyako-Yaeyama area (MPA1) or in Okinawa (MPA2), both covering the same surface (~15'000 km^2). The overall difference in GD (ΔGD=+0.17) indicates that MPA2 globally reduces GDs from potentially adapted reefs more than MPA1. This is probably due to the fact that MPA2 includes more reefs that are putatively adaptive (Fig. S7) and is better connected to the reef-dense area around Amami. MPA1, on the other hand, is far from this reef-dense area and shows beneficial effects only in Yaeyama.



Coral Nurseries

Two CNs emplacements were compared: in Amami (CN1) or in Tokara (CN2). The overall difference in GD (ΔGD=-0.28) suggests a more beneficial effect on adaptive potential under CN2, rather than CN1. In fact, the existence of a coral nursery in Amami (CN1) could provide adaptive genetic material to a larger number of reefs, compared with a set-up established in Tokara (CN2), because of a higher outbound-connectivity potential (Fig. 3b).

Acknowledgements

The authors received no specific financial support for the research, authorship, and/or publication of this article. The authors have no conflict of interest to declare.

References

Anthony, K. R. N., Kline, D. I., Diaz-Pulido, G., Dove, S., & Hoegh-Guldberg, O. (2008). Ocean acidification causes bleaching and productivity loss in coral reef builders. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(45), 17442–17446. https://doi.org/10.1073/pnas.0804478105

Ateweberhan, M., Feary, D. A., Keshavmurthy, S., Chen, A., Schleyer, M. H., & Sheppard, C. R. C. (2013). Climate change impacts on coral reefs: Synergies with local effects, possibilities for acclimation, and management implications. *Marine Pollution Bulletin*, *74*(2), 526–539. https://doi.org/10.1016/j.marpolbul.2013.06.011

Balkenhol, N., Dudaniec, R. Y., Krutovsky, K. V., Johnson, J. S., Cairns, D. M., Segelbacher, G., … Joost, S. (2017). Landscape Genomics: Understanding Relationships Between Environmental Heterogeneity and Genomic Characteristics of Populations (pp. 1–62). Springer, Cham. https://doi.org/10.1007/13836_2017_2

Baums, I. B. (2008). A restoration genetics guide for coral reef conservation. *Molecular Ecology*, *17*(12), 2796–2811. https://doi.org/10.1111/j.1365-294X.2008.03787.x

Bay, R. A., & Palumbi, S. R. (2014). Multilocus Adaptation Associated with Heat Resistance in Reef-Building Corals. *Current Biology*, *24*(24), 2952–2956.

https://doi.org/10.1016/j.cub.2014.10.044

Beger, M., Selkoe, K. A., Treml, E., Barber, P. H., von der Heyden, S., Crandall, E. D., … Riginos, C. (2014). Evolving coral reef conservation with genetic information. *Bulletin of Marine Science*, *90*(1), 159–185. https://doi.org/10.5343/bms.2012.1106

Bellwood, D. R., Hughes, T. P., Folke, C., & Nyström, M. (2004). Confronting the coral reef crisis. *Nature*, *429*(6994), 827–833. https://doi.org/10.1038/nature02691

Bozdogan, H. (1987). Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions. *Psychometrika*, *52*(3), 345–370. https://doi.org/10.1007/BF02294361

Brodie, A., Azaria, J. R., & Ofran, Y. (2016). How far from the SNP may the causative genes be? *Nucleic Acids Research*, *44*(13), 6046–6054. https://doi.org/10.1093/nar/gkw500

CIESIN Columbia University. (2010). Gridded Population of the World (GPW), v3. Retrieved January 18, 2019, from http://sedac.ciesin.columbia.edu/data/set/gpw-v3-population-count

Cinner, J. E., Huchery, C., MacNeil, M. A., Graham, N. A. J., McClanahan, T. R., Maina, J., … Mouillot, D. (2016). Bright spots among the world's coral reefs. *Nature*, *535*(7612), 416–419. https://doi.org/10.1038/nature18607

Courcelle, J., & Hanawalt, P. C. (1999). RecQ and RecJ process blocked replication forks prior to the resumption of replication in UV-irradiated Escherichia coli. *Molecular & General Genetics : MGG*, *262*(3), 543–551. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10589843

Courtial, L., Picco, V., Grover, R., Cormerais, Y., Rottier, C., Labbe, A., … Ferrier-Pagès, C. (2017). The c-Jun N-terminal kinase prevents oxidative stress induced by UV and thermal stresses in corals and human cells. *Scientific Reports*, *7*(1), 45713. https://doi.org/10.1038/srep45713

Covington, A. K., & Whitfield, M. (1988). Recommendations for the determination of pH in sea water and estuarine waters. *Pure and Applied Chemistry*, *60*(6), 865–870. https://doi.org/10.1351/pac198860060865

EU Copernicus Marine Service. (2017). Global Ocean - In-Situ-Near-Real-Time Observations. Retrieved February 2, 2017, from http://marine.copernicus.eu

Frichot, E., & François, O. (2015). LEA : An R package for landscape and ecological association studies. *Methods in Ecology and Evolution*, *6*(8), 925–929. https://doi.org/10.1111/2041-210X.12382

Hijmans, R. J. (2016). raster: Geographic Data Analysis and Modeling. Retrieved from https://cran.r-project.org/package=raster

Howells, E. J., Berkelmans, R., van Oppen, M. J. H., Willis, B. L., & Bay, L. K. (2013). Historical thermal regimes define limits to coral acclimatization. *Ecology*, *94*(5), 1078–1088. https://doi.org/10.1890/12-1257.1

Hughes, T. P., Anderson, K. D., Connolly, S. R., Heron, S. F., Kerry, J. T., Lough, J. M., … Wilson, S. K. (2018). Spatial and temporal patterns of mass bleaching of corals in the Anthropocene. *Science*, *359*(6371), 80–83. https://doi.org/10.1126/science.aan8048

Hughes, T. P., Kerry, J. T., Álvarez-Noriega, M., Álvarez-Romero, J. G., Anderson, K. D., Baird, A. H., … Wilson, S. K. (2017). Global warming and recurrent mass bleaching of corals. *Nature*, *543*(7645), 373–377. https://doi.org/10.1038/nature21707

Hughes, T. P., Kerry, J. T., Connolly, S. R., Baird, A. H., Eakin, C. M., Heron, S. F., … Torda, G. (2019). Ecological memory modifies the cumulative impact of recurrent climate extremes. *Nature Climate Change*, *9*(1), 40–43. https://doi.org/10.1038/s41558-018-0351-2

Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Joost, S. (2006). *The geographical dimension of genetic diversity: a GIScience contribution for the conservation of animal genetic resources*. EPFL. https://doi.org/10.5075/EPFL-THESIS-3454

Joost, S., Bonin, A., Bruford, M. W., Després, L., Conord, C., Erhardt, G., & Taberlet, P. (2007). A spatial analysis method (SAM) to detect candidate loci for selection: Towards a landscape genomics approach to adaptation. *Molecular Ecology*, *16*(18), 3955–3969. https://doi.org/10.1111/j.1365-294X.2007.03442.x

Keller, L., & Waller, D. M. (2002). Inbreeding effects in wild populations. *Trends in Ecology & Evolution*, *17*(5), 230–241. https://doi.org/10.1016/S0169-5347(02)02489-8

Krueck, N. C., Ahmadia, G. N., Green, A., Jones, G. P., Possingham, H. P., Riginos, C., … Mumby, P. J. (2017). Incorporating larval dispersal into MPA design for both conservation and fisheries. *Ecological Applications*, *27*(3), 925–941. https://doi.org/10.1002/eap.1495

Krueger, T., Horwitz, N., Bodin, J., Giovani, M.-E., Escrig, S., Meibom, A., & Fine, M. (2017). Common reef-building coral in the Northern Red Sea resistant to elevated temperature and acidification. *Royal Society Open Science*, *4*(5), 170038. https://doi.org/10.1098/rsos.170038

LaJeunesse, T. C., Parkinson, J. E., Gabrielson, P. W., Jeong, H. J., Reimer, J. D., Voolstra, C. R., & Santos, S. R. (2018). Systematic Revision of Symbiodiniaceae Highlights the Antiquity and Diversity of Coral Endosymbionts. *Current Biology*, *28*(16), 2570–2580.e6. https://doi.org/10.1016/j.cub.2018.07.008

Lee, K., Tong, L. T., Millero, F. J., Sabine, C. L., Dickson, A. G., Goyet, C., … Key, R. M. (2006). Global relationships of total alkalinity with salinity and temperature in surface waters of the world's oceans. *Geophysical Research Letters*, *33*(19), L19605. https://doi.org/10.1029/2006GL027207

Lester, S., Halpern, B., Grorud-Colvert, K., Lubchenco, J., Ruttenberg, B., Gaines, S., … Warner, R. (2009). Biological effects within no-take marine reserves: a global synthesis. *Marine Ecology Progress Series*, *384*, 33–46. https://doi.org/10.3354/meps08029

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … 1000 Genome Project Data Processing Subgroup, 1000 Genome Project Data Processing. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England)*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Liu, G., Strong, A. E., & Skirving, W. (2003). Remote sensing of sea surface temperatures during 2002 Barrier Reef coral bleaching. *Eos, Transactions American Geophysical Union*, *84*(15), 137–141. https://doi.org/10.1029/2003EO150001

Logan, C. A., Dunne, J. P., Eakin, C. M., & Donner, S. D. (2014). Incorporating adaptive responses into future projections of coral bleaching. *Global Change Biology*, *20*(1), 125–139. https://doi.org/10.1111/gcb.12390

Loukos, H., Vivier, F., Murphy, P. P., Harrison, D. E., & Le Quéré, C. (2000). Interannual variability of equatorial Pacific CO_2 fluxes estimated from temperature and salinity data. *Geophysical Research Letters*, *27*(12), 1735–1738. https://doi.org/10.1029/1999GL011013

Lukoschek, V., Riginos, C., & van Oppen, M. J. H. (2016). Congruent patterns of connectivity can inform management for broadcast spawning corals on the Great Barrier Reef. *Molecular Ecology*, *25*(13), 3065–3080. https://doi.org/10.1111/mec.13649

Lundgren, P., Vera, J. C., Peplow, L., Manel, S., & van Oppen, M. J. H. (2013). Genotype - environment correlations in corals from the Great Barrier Reef. *BMC Genetics*, *14*(1), 9.

https://doi.org/10.1186/1471-2156-14-9

Maina, J., McClanahan, T. R., Venus, V., Ateweberhan, M., & Madin, J. (2011). Global gradients of coral exposure to environmental stresses and implications for local management. *PLoS ONE*, *6*(8), e23064. https://doi.org/10.1371/journal.pone.0023064

Maina, J., Venus, V., McClanahan, T. R., & Ateweberhan, M. (2008). Modelling susceptibility of coral reefs to environmental stress using remote sensing data and GIS models. *Ecological Modelling*, *212*(3), 180–199. https://doi.org/10.1016/j.ecolmodel.2007.10.033

McCook, L. J., Almany, G. R., Berumen, M. L., Day, J. C., Green, A. L., Jones, G. P., … Thorrold, S. R. (2009). Management under uncertainty: guide-lines for incorporating connectivity into the protection of coral reefs. *Coral Reefs*, *28*(2), 353–366. https://doi.org/10.1007/s00338-008-0463-7

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., … DePristo, M. A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*(9), 1297–1303. https://doi.org/10.1101/gr.107524.110

McKinney, W. (2010). Data Structures for Statistical Computing in Python. Retrieved from http://conference.scipy.org/proceedings/scipy2010/mckinney.html

Moberg, F., & Folke, C. (1999). Ecological goods and services of coral reef ecosystems. *Ecological Economics*, *29*(2), 215–233. https://doi.org/10.1016/S0921-8009(99)00009-9

Mydlarz, L. D., McGinty, E. S., & Harvell, C. D. (2010). What are the physiological and immunological responses of coral to climate warming and disease? *The Journal of Experimental Biology*, *213*(6), 934–945. https://doi.org/10.1242/jeb.037580

Nakajima, Y., Nishikawa, A., Iguchi, A., & Sakai, K. (2010). Gene Flow and Genetic Diversity of a Broadcast-Spawning Coral in Northern Peripheral Populations. *PLoS*

ONE, *5*(6), e11149. https://doi.org/10.1371/journal.pone.0011149

NASA. (2016). NASA Earth Observations (NEO). Retrieved May 10, 2017, from http://neo.sci.gsfc.nasa.gov/

National Oceanic and Atmospheric Administration. (2017). NOAA DATA CATALOG. Retrieved February 2, 2017, from https://data.noaa.gov/dataset/

NCBI. (1988). National Center for Biotechnology Information. Retrieved January 18, 2019, from https://www.ncbi.nlm.nih.gov/

Neupane, A., Nepal, M. P., Benson, B. V, Macarthur, K. J., & Piya, S. (2013). Evolutionary history of mitogen-activated protein kinase (MAPK) genes in Lotus, Medicago, and Phaseolus. *Plant Signaling & Behavior*, *8*(11), e27189. https://doi.org/10.4161/psb.27189

Nishikawa, A. (2008). Degree and Pattern of Gene Flow in Several Scleractinian Corals in the Ryukyu Archipelago, Southern Japan. *Https://Doi.Org/10.2984/1534-6188(2008)62[413:DAPOGF]2.0.CO;2*, *62*(3), 413–422. https://doi.org/10.2984/1534-6188(2008)62[413:DAPOGF]2.0.CO;2

OECD. (2017). *Marine protected areas : economics, management and effective policy mixes.* Paris: OECD Publishing.

Palumbi, S. R. (2003). Population genetics, demographic connectivity, and the design of marine reserves. *Ecological Applications*, *13*(1 SUPPL.), 146–158. https://doi.org/10.1890/1051-0761(2003)013[0146:PGDCAT]2.0.CO;2

Palumbi, S. R., Barshis, D. J., Traylor-Knowles, N., & Bay, R. A. (2014). Mechanisms of reef coral resistance to future climate change. *Science*, *344*(6186). Retrieved from http://science.sciencemag.org/content/344/6186/895

Paris-Limouzy, C. B. (2011). Reef Interconnectivity/Larval Dispersal (pp. 881–889). Springer, Dordrecht. https://doi.org/10.1007/978-90-481-2639-2_138

Penin, L., Vidal-Dupiol, J., & Adjeroud, M. (2013). Response of coral assemblages to thermal stress: are bleaching intensity and spatial patterns consistent between events? *Environmental Monitoring and Assessment*, *185*(6), 5031–5042. https://doi.org/10.1007/s10661-012-2923-3

Pratchett, M. S., Thompson, C. A., Hoey, A. S., Cowman, P. F., & Wilson, S. K. (2018). Effects of Coral Bleaching and Coral Loss on the Structure and Function of Reef Fish Assemblages (pp. 265–293). Springer, Cham. https://doi.org/10.1007/978-3-319-75393-5_11

Python Software Foundation. (2018). Python Language Refrence, version 3.5. Retrieved from www.python.org

QGIS development team. (2009). QGIS Geographic Information System. Open Source Geospatial Foundation Project. Retrieved from http://www.qgis.org/

R Core Team. (2016). R: A Language and Environment for Statistical Computing. Retrieved from https://www.r-project.org/

Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., & Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, *24*(17), 4348–4370. https://doi.org/10.1111/mec.13322

Riginos, C., Crandall, E. D., Liggins, L., Bongaerts, P., & Treml, E. A. (2016). Navigating the currents of seascape genomics: how spatial analyses can augment population genomic studies. *Current Zoology*, *62*, doi: 10.1093/cz/zow067. https://doi.org/10.1093/cz/zow067

Rochat, E., Leempoel, K., Vajana, E., Colli, L., Ajmone-Marsan, P., Joost, S. and the N. C. (2016). Map of genotype frequency change in autochtonous Moroccan sheep breeds due to global warming. In *Conference on Conservation genomics, May 3-6, 2016, Vairao, Portugal*.

Rodgers, K. S., Lorance, K., Richards Donà, A., Stender, Y., Lager, C., & Jokiel, P. L. (2017). Effectiveness of coral relocation as a mitigation strategy in Kāne'ohe Bay, Hawai'i. *PeerJ*, *5*, e3346. https://doi.org/10.7717/peerj.3346

Rosic, N. N., Pernice, M., Dunn, S., Dove, S., & Hoegh-Guldberg, O. (2010). Differential regulation by heat stress of novel cytochrome P450 genes from the dinoflagellate symbionts of reef-building corals. *Applied and Environmental Microbiology*, *76*(9), 2823–2829. https://doi.org/10.1128/AEM.02984-09

Ryan, W. B. F., Carbotte, S. M., Coplan, J. O., O'Hara, S., Melkonian, A., Arko, R., … Zemsky, R. (2009). Global Multi-Resolution Topography synthesis. *Geochemistry, Geophysics, Geosystems*, *10*(3), n/a-n/a. https://doi.org/10.1029/2008GC002332

Sampayo, E. M., Franceschinis, L., Roff, G., Ridgway, T., Dove, S., & Hoegh-Guldberg, O. (2016). Coral symbioses under prolonged environmental change: living near tolerance range limits. *Scientific Reports*, *6*(1), 36271. https://doi.org/10.1038/srep36271

Sang, W., Ma, W.-H., Qiu, L., Zhu, Z.-H., & Lei, C.-L. (2012). The involvement of heat shock protein and cytochrome P450 genes in response to UV-A exposure in the beetle Tribolium castaneum. *Journal of Insect Physiology*, *58*(6), 830–836. https://doi.org/10.1016/j.jinsphys.2012.03.007

Seabold, S., & Perktold, J. (2010). Statsmodels: econometric and statistical modeling with Python. In *9th Python in Science Conference* (pp. 57–61). Retrieved from http://statsmodels.sourceforge.net/

Selig, E. R., & Bruno, J. F. (2010). A global analysis of the effectiveness of marine protected areas in preventing coral loss. *PloS One*, *5*(2), e9278. https://doi.org/10.1371/journal.pone.0009278

Shanks, A. L., Grantham, B. A., & Carr, M. H. (2003). Propagule Dispersal Distance and the Size and Spacing of Marine Reserves. *Ecological Applications*, *13*(1).

https://doi.org/https://doi.org/10.1890/1051-0761(2003)013[0159:PDDATS]2.0.CO;2

Sharma, S., Doherty, K. M., Brosh, R. M., & Jr. (2006). Mechanisms of RecQ helicases in pathways of DNA metabolism and maintenance of genomic stability. *The Biochemical Journal*, *398*(3), 319–337. https://doi.org/10.1042/BJ20060450

Shinzato, C., Mungpakdee, S., Arakaki, N., & Satoh, N. (2015). Genome-wide SNP analysis explains coral diversity and recovery in the Ryukyu Archipelago. *Scientific Reports*, *5*(1), 18211. https://doi.org/10.1038/srep18211

Shinzato, C., Shoguchi, E., Kawashima, T., Hamada, M., Hisata, K., Tanaka, M., … Satoh, N. (2011). Using the Acropora digitifera genome to understand coral responses to environmental change. *Nature*, *476*(7360), 320–323. https://doi.org/10.1038/nature10249

Storey, J. D. (2003). The Positive False Discovery Rate: A Bayesian Interpretation and the q-Value. *Annals of Statistics*, *31*(6), 2013–2035.

Storlazzi, C. D., van Ormondt, M., Chen, Y.-L., & Elias, E. P. L. (2017). Modeling Fine-Scale Coral Larval Dispersal and Interisland Connectivity to Help Designate Mutually-Supporting Coral Reef Marine Protected Areas: Insights from Maui Nui, Hawaii. *Frontiers in Marine Science*, *4*, 381. https://doi.org/10.3389/fmars.2017.00381

Stubben, C., & Milligan, B. (2007). Estimating and Analyzing Demographic Models Using the **popbio** Package in *R*. *Journal of Statistical Software*, *22*(11), 1–23. https://doi.org/10.18637/jss.v022.i11

Stucki, S., Orozco-terWengel, P., Bruford, M. W., Colli, L., Masembe, C., Negrini, R., … Consortium, the N. (2017). High performance computation of landscape genomic models integrating local indices of spatial association. *Molecular Ecology Resources*, *17*(5), 1072–1089. https://doi.org/10.1111/1755-0998.12629

Thomas, L., Kennington, W. J., Evans, R. D., Kendrick, G. A., & Stat, M. (2017). Restricted

gene flow and local adaptation highlight the vulnerability of high-latitude reefs to rapid environmental change. *Global Change Biology*, *23*(6), 2197–2205. https://doi.org/10.1111/gcb.13639

Thompson, D. M., & van Woesik, R. (2009). Corals escape bleaching in regions that recently and historically experienced frequent thermal stress. *Proceedings. Biological Sciences / The Royal Society*, *276*(1669), 2893–2901. https://doi.org/10.1098/rspb.2009.0591

UNEP-WCMC, WorldFish-Center, WRI, & TNC. (2010). Global distribution of warm-water coral reefs, compiled from multiple sources including the Millennium Coral Reef Mapping Project. Version 1.3. Retrieved May 9, 2017, from http://data.unep-wcmc.org/datasets/1

Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., … DePristo, M. A. (2013). From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. In *Current Protocols in Bioinformatics* (Vol. 43, p. 11.10.1-11.10.33). Hoboken, NJ, USA: John Wiley & Sons, Inc. https://doi.org/10.1002/0471250953.bi1110s43

van Etten, J. (2018). gdistance: Distances and Routes on Geographical Grids. Retrieved from https://cran.r-project.org/package=gdistance

van Oppen, M. J. H., Oliver, J. K., Putnam, H. M., & Gates, R. D. (2015). Building coral reef resilience through assisted evolution. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(8), 2307–2313. https://doi.org/10.1073/pnas.1422301112

Van Woesik, R., Irikawa, A., & Loya, Y. (2004). Coral Bleaching: Signs of Change in Southern Japan. In *Coral Health and Disease* (pp. 119–141). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-06414-6_5

Visel, A., Rubin, E. M., & Pennacchio, L. A. (2009). Genomic views of distant-acting

enhancers. *Nature*, *461*(7261), 199–205. https://doi.org/10.1038/nature08451

Welle, P. D., Small, M. J., Doney, S. C., & Azevedo, I. L. (2017). Estimating the effect of multiple environmental stressors on coral bleaching and mortality. *PloS One*, *12*(5), e0175018. https://doi.org/10.1371/journal.pone.0175018

Yampolsky, L. Y., Zeng, E., Lopez, J., Williams, P. J., Dick, K. B., Colbourne, J. K., & Pfrender, M. E. (2014). Functional genomics of acclimation and adaptation in response to thermal stress in Daphnia. *BMC Genomics*, *15*(1), 859. https://doi.org/10.1186/1471-2164-15-859

Young, C. N., Schopmeyer, S. A., & Lirman, D. (2012). A review of reef restoration and Coral propagation using the threatened genus Acropora in the Caribbean and western Atlantic. *Bulletin of Marine Science*, *88*(4), 1075–1098. https://doi.org/10.5343/bms.2011.1143

Ziegler, M., Seneca, F. O., Yum, L. K., Palumbi, S. R., & Voolstra, C. R. (2017). Bacterial community dynamics are linked to patterns of coral heat tolerance. *Nature Communications*, *8*, 14213. https://doi.org/10.1038/ncomms14213