Demographic history of a remote ichthyofauna assemblage reveals common processes driving colonization and persistence in endemic coral reef fishes

Erwan Delrieu-Trottin, Nicolas Hubert, Emily C. Giles, Pascaline Chifflet-Belle, Arnaud Suwalski, Valentina Neglia, Cristian Rapu-Edmunds, Stefano Mona, Pablo Saenz-Agudelo

1 Instituto de Ciencias Ambientales y Evolutivas, Universidad Austral de Chile, Valdivia, Chile
2 Institut de Recherche pour le Développement, UMR 5554 (UM-CNRS-IRD-EPHE), ISEM, Montpellier, France
3 Museum für Naturkunde, Leibniz-Institut für Evolutions-und Biodiversitätsforschung an der Humboldt-Universität zu Berlin, Berlin, Germany
4 Institut de Systématique, Evolution, Biodiversité, ISYEB, Muséum national d'Histoire naturelle, CNRS, Sorbonne Université, EPHE, Université des Antilles, Paris, France
5 EPHE, PSL Research University, Paris, France
6 Mike Rapu Diving Center, Caleta Hanga Roa O'tai, Chile

† These authors jointly supervised this work.

Author for correspondence: Erwan Delrieu-Trottin, e-mail: erwan.delrieu.trottin@gmail.com

Keywords: endemism, community assembly, colonization, ddRADs, Easter Island, population genetics
Abstract

Elucidating demographic history during the settlement of ecological communities is crucial to properly inferring the mechanisms that shape species diversity and their persistence through time. We used genomic data to elucidate for the first time the demographic dynamics associated with the settlement by endemic reef fish fauna of one of the most isolated islands of the Pacific Ocean, Rapa Nui (Easter Island). We compared the demographic history of nine endemic species including seven small-range (restricted to Rapa Nui and Motu Motiro Hiva) and two large-range (present in Rapa Nui and other southern subtropical islands of the Pacific) endemic species in order to explore the onset of community settlement and associated demographic history. We found that most Rapa Nui endemic species share a common demographic history, with a demographic expansion initiated during the last interglacial period and related to the last sea level high-stand. The commonality of this pattern suggests that eustatic fluctuations associated with Milankovitch cycles have played a central role in species demographic history and the final stage of contemporary community assembly of Rapa Nui reef fishes. We discuss the potential role of seamounts in the colonization / extinction / recolonization dynamics of populations of the Rapa Nui Archipelago.
1. Introduction

Local and regional processes jointly drive community assembly, i.e. the dynamics of species aggregations in ecological communities and the number of coexisting species at a given place. The importance and nature of local (e.g. environmental filtering, species interactions inside a community) and regional (e.g. dispersal, immigration of new species from outside of a community) processes in the origin and maintenance of biodiversity in communities is one of the oldest question in ecology [1,2]. Phylogenetic analysis of community assembly can give a better understanding of the relative importance of local and regional dynamics in shaping community assembly [3] as they provide a temporal framework [4]. However, the difficulty of reconstructing accurate divergence times between species due to the incompleteness of available phylogenies limit the power of such analyses. In the meantime, community assembly, especially in island systems, is known to be dynamic, with regular colonization / extinction / and recolonization of populations occurring through time [2,5,6], processes which cannot be assessed using species level approaches. Investigating shallower time scale, population genetics can provide insights into both assembly time and species demographic history during community assembly, hence helping understanding the mechanisms by which species persist through time. Thus by comparing the demographic histories of multiple species in a community and estimating the timing of the early stages of community assembly, it is possible to disentangle not only the relative contribution of historical (e.g. tectonic, eustatic changes) and ecological (e.g. competitive exclusion, habitat filtering) dynamics but also the scale of the processes (local vs. regional) in shaping the final stage of contemporary community assembly.

Rapa Nui (Easter Island, Chile, 166 km²) is one of the most remote islands on Earth and hosts an exceptional reef fish community. The mechanisms at the origin of
this diversity and its persistence through time remain enigmatic. Rapa Nui (RN) is the second hotspot of endemism of the Indo-Pacific Ocean, with 21.7% of fishes endemic to these islands [7,8]. This reef fish community is known for being extremely species poor (only 169 fish species; 139 shore fishes) compared to the species-rich islands of the Indo-Australian Archipelago for instance. RN and the islet Motu Motiro Hiva (MMH) (Salas y Gómez, 0.15 km²), located 400 km further east, which shares the same reef fish fauna as RN [9], constitute the only two emerged islands of the Easter Chain. RN and MMH are relatively young; 2.5 and 1.7 My, respectively [10], but they are embedded in a network of numerous seamounts. This seamount chain extends 2,232 km east to the Nazca seamount (23°360 S et 83°300 W) [10,11] and these mounts have emerged to various degrees during periods of low sea level. In this context, an "Ancient Archipelago" hypothesis has been formulated to account for the high level of endemism observed in RN and MMH, young islands, hypothesizing that "most if not all endemic species were acquired from elsewhere in the region" [12]. Seamounts, which were once likely islands, could have provided potentially suitable habitat for at least the past 29 My for endemics of the region to evolve and persist up to present times [12]. However, a recent analysis of the divergence times of endemic species from their closest relatives has shown that small-range endemics are not older that the emergence of RN and MMH; they are thus neoendemics [13], thereby questioning the ancient origin of RN endemics. Yet, the seamounts nearby RN could have played a major role in long-term species persistence through time, especially during the Pleistocene. This era is characterized by notable Milankovitch cycles, historical processes resulting in the alternation of glacial and inter-glacial time that resulted in sea levels 150 m below present level. Such paleo-environmental perturbations might be expected to have left a footprint of bottlenecks and expansions in the demographic history of populations of many marine species [14–17]. These major sea level changes altered shallow water habitats; exposing continental shelves during low sea levels but
also giving access to habitat otherwise too deep to colonize such as relatively shallow
seamounts. Thus, these features of the Pleistocene could have temporally expanded
the range distribution of endemic species and their population sizes[18].

Reef fish community assembly in remote islands is dependent on colonization /
extinctions / recolonization processes. The RN reef ichthyofauna is very dynamic;
several species that were reported as being abundant in 1969 in RN waters were then
reported as rare or even absent 16 years later. Along the same line, numerous species
are reported as present in RN waters but are actually vagrants, i.e. species that
colonize RN from time to time without being able to establish a population locally [7].
The endemic RN reef ichthyofauna is composed of two types of endemics, small-range
and large-range endemic species. It is possible to hypothesize that local processes, at
the scale of the RN and MMH community, influence the demographic history of small-
range endemics as these fishes are restricted to these two islands. On the contrary,
large-range endemics are present both in RN and in other southern subtropical islands
of the Pacific. As such, populations of these fishes are embedded in metapopulations
of larger geographic ranges than those of small-range endemics. Thus the population
dynamics of large-range endemics are likely to also be affected by regional processes
such as fragmentation and dispersal outside of the RN and MMH community.

Through genome-wide sequencing (ddRAD), we explore the demographic
history of nine endemic species of RN. The species studied here represent the two
types of endemics (small-range and large-range), seven major reef fish families, and
three different reproductive strategies endemics. We examine whether demographic
histories vary (1) according to the current range distribution of the species; (2) with life
history traits; doing so, we investigate the consequence of historical processes such as
sea level changes or ecological dynamics driven by reproductive strategy on the genetic diversity. If historical processes are more important than ecological processes in the maintenance of endemic species in RN, we expect to find similar population dynamic for the two types of endemic species. In the same way, if local processes are more important than regional processes in the maintenance of endemic species in RN, we expect to find similar assembly time for the two types of endemic species. Implications for the biogeography of the area are discussed.

2. Material and methods

(a) Sampling

A total of 143 reef fishes (nine species) were collected using polespears or an anesthetic (clove oil) in Rapa Nui in October 2016 (Table 1). Species were classified as either small-range or large-range endemics. Small-range endemics only present around RN and MMH and have a maximum range of <500 km in linear distance, see Delrieu-Trottin et al. [19]. Southern subtropical endemics have large-ranges (1,000–8,000 km in linear distance, see Delrieu-Trottin et al. [19], hereafter large-range endemic species) and are distributed from Southern Polynesia to RN (regional endemics in Friedlander et al. [9]). The species analyzed here also possess different reproductive strategies, with five species producing pelagic eggs, three species producing demersal eggs, and one species brooding eggs in their mouths (Table 1).

(b) Library preparation and sequencing

Whole genomic DNA was extracted from fin or gill tissue preserved in 96% ethanol using the GeneJet Genomic DNA purification kit according to the manufacturer’s protocols (Thermo Fisher Scientific). Double-digest restriction-associated DNA (ddRAD) libraries were prepared following Peterson et al.’s protocol [20]. The genomic
libraries obtained were sequenced in 3 lanes of a HiSeq 2500 Illumina sequencer (single end, 125 pb). Illumina reads are available from the Sequence Read Archive (SRA) at NCBI under the Accession nos. XXXXXX–XXXXXXX.

(c) De novo assembly

We used the ‘process_radtags.pl’ pipeline in STACKS version 2.0 [21,22] to demultiplex and quality filter the sequences obtained. In the absence of reference genomes for the species under study, RADSeq loci were assembled de novo using the ‘denovo_map.pl’ pipeline in STACKS. We used the parameter combination recommended by Mastretta-Yanes et al. [23]; this included minimum read depth to create a stack \(m = 3\), number of mismatches allowed between loci within individuals \(M = 3\), number of mismatches allowed between loci within catalogue \(n = 3\) and required a locus to be present in all individuals of each species \(r = 1\). Following de novo mapping, an initial data-filtering step was performed using the population component of STACKS removing all loci with maximum observed heterozygosity higher than 0.8. We first kept all single-nucleotide polymorphisms (SNP) per stack \(i.e.\) locus and did not use any threshold regarding the minor allele frequencies as we obtained a high coverage for each individual. We then removed all loci displaying more than three SNPs to avoid potential paralogs. Summary statistics based on the resulting vcf file (Supplementary material) such as nucleotide diversity \(\pi\) on variable loci, heterozygosity \(H\) and \(F_{is}\), were calculated for every SNP using the populations program in STACKS. The folded Site (Allele) Frequency Spectrum (SFS) was computed in R [24] with the package pegas [25].

(d) Demographic analyses

Variation in the effective population size \(\Ne\) through time was investigated using the composite likelihood approach implemented in the software stairwayplot [26]. The
stairwayplot is a non-parametric model where $N_e$ is free to vary at each coalescent interval. The composite likelihood is evaluated as the difference between the observed SFS and its expectation under a specific demographic history. To confirm the results obtained by the stairwayplot, we ran an additional approximate Bayesian computation algorithm based on coalescent simulations following Maisano Delser et al [27]. Briefly, we performed 1,000,000 coalescent simulations of a demographic model with three instantaneous changes of $N_e$. The model is therefore defined by seven parameters: four values of $N_e$ and three instantaneous time changes (hereafter, $T$). We set the same uniform distribution for the four $N_e$ values and incremental uniform distribution for the three $T$ parameters, similarly to Maisano Delser et al [27]. Coalescent simulations and SFS computation were performed with fastsimcoal [28]. We used the SFS and the mean pairwise differences (computed with a custom R script) as summary statistics. We retained the best 5,000 simulations to perform a local linear regression [29] and reconstructed the abc skyline at user specified time points starting from the posterior distribution as in Maisano Delser et al [27]. We used a mutation rate ($\mu$) of $1.0 \times 10^{-8}$/site/generation following [30–34] and generation times found in the literature (Table 1) or inferred using maximum standard length.

As data did not meet assumptions of normality, we computed the Wilcoxon-Mann-Whitney and Kruskal-Wallis non parametric tests to examine differences in the Time to the Most Recent Common Ancestor (TMRCA) considering four factors: range size, family, generation time, and reproductive strategies. All statistical analyses were performed in R, using the package vegan [35] and ggplot2 was used for graphical representations [36].

3. Results
Raw sequence filtering, assembly, and SNP calling. A total of 31,543,967 reads of 121 bp each were obtained for the 143 individual samples from the nine species endemic to RN. An average of 3,504,885 reads were found per species (min: 2,983,062, max: 4,430,571). The different filtering steps resulted in the building of an average of 41,470 loci (min: 25,331; max: 63,590) per species out of which an average of 23,693 (min: 19,246; max: 28,855) were variable (Table 2). The variable loci harbored an average of 37,107 SNPs (min: 28,239; max: 46,718). The depth of coverage per SNP ranged from 31x to 50x (43x averaged across species).

Genetic diversity statistics. Different levels of genetic diversity were found for the nine species (Table 2). *Cantherhines rapanui* displayed the highest values of heterozygosity (max: 0.262) and nucleotide diversity (32.73 x 10^-4). *Chrysiptera rapanui* showed the lowest values of heterozygosity (0.136) while *Ostorhinchus chalcius* displayed the lowest values of nucleotide diversity (0.063).

Community assembly time. The mean TMRCA retrieved with the stairway plots for the nine species was very recent 253,871 YBP (+/- 119,482 YBP) and ranged from 105,232 YBP (*Ostorhinchus chalcius*) to 398,058 YBP (*Sargocentron wilhelmi*) (Figure 1). The TMRCAs retrieved did not differ significantly among types of endemics (W = 5, p = 0.6667), families (H6 = 7.4667, p = 0.2798), or reproductive strategy (H2 = 2.56, p = 0.278). The only significant factor was generation time (H2 = 6.1111, p = 0.0471).

Changes in historical effective population size. We recovered very similar demographic patterns with the two methods (stairway plots and the abc skylines; Figure 1, S1). Patterns of population expansions were recovered for seven out of nine endemic species studied, constant population size through time was found for one species (*Sargocentron wilhelmi*) and a decrease in population size was recovered for one.
species (*Cantherhines rapanui*). Expansion times retrieved were highly similar between
the two methods for four out of the seven species, while abc skylines estimated older
expansion times than those inferred with the stairway plots for the three other cases. In
the same way, the bottleneck found for *Cantherhines rapanui* was dated slightly older
with the abc skyline than with the stairway plot analysis.

4. Discussion

Rapa Nui endemic species share a common history, with population expansions
dominating the demographic histories dated during the last interglacial period for both
small-range and large-range endemics. These expansions indicate that their current
effective population size established in RN during the last glacial maxima (25 000 - 115
000 YBP), a period characterized by climatic cooling and decreased sea levels [37].
These results agree with those found for expansions and bottlenecks of other marine
organisms suggesting that sea level oscillations during the last 800 000 years level
[38,39] has greatly influenced reef fish populations [40–49]. Rapa Nui endemics,
however, differ in their ecology compared to organisms of other reefs in the Pacific
region, as they are far less restricted in terms of depth range. Many of the RN endemic
species have the particularity to be found from shallow waters to the mesophotic zone
(100 - 150 m) in RN water and in seamounts nearby RN (Table 1). Considering the
network of seamounts surrounding RN, low sea levels during glacial maxima would
have produced periods of maximum reef habitat extension (Figure 2). As such,
population expansions during glacial maxima could reflect the colonization /
recolonization history of RN from MMH or other seamounts of the RN archipelago.
While the Ancient Archipelago hypothesis of Newman & Foster [12] has been
invalidated at the species level for this island system [13], this hypothesis could explain
what is observed at the population level for RN endemic species (i.e. a recolonization
from the seamounts rather than from elsewhere in the Pacific). Overall, historical
metapopulation dynamics within the RN Archipelago combined with sea level
fluctuations during glacial maxima could have driven the population dynamics and
hence community assembly of RN endemics.

The common demographic and temporal patterns found for large-range and
small-range endemics provide insights into the mechanisms that shaped the RN
ichthyofauna. Community assembly in remote islands is expected to rely on
colonization / extinction / recolonization processes. The concordant expansions
detected for all species except Cantherines rapanui and Sargocentron wilhelmi suggest
a concomitant establishment of current population sizes during the last glacial cycle.
This pattern does not seem to be influenced by the size of the species range
distribution as both small-range and large-range endemics display similar patterns.
Actually, populations of large-range endemics sampled around RN could potentially
operate at a regional scale, being also present in several other islands of the South
Pacific, but the similar expansion times observed here for both types of endemism
argue in favor of population demography driven by local historical processes.

Local processes rather than regional processes seem to have driven the final stage of
community assembly among endemic species around RN, as indicated by the similar
TMRCAs retrieved here for both types of endemism. This suggests that the
demography of these species has been poorly affected by neighbouring populations.
This is consistent with the isolation of RN, as supported by the lower species richness
observed there compared to in other Pacific islands. The low species diversity in RN
reefs might not only be due to the difficulty for larvae to colonize such remote islands,
but also due to the difficulty to establish a viable population / metapopulation that
functions at such local scales. This hypothesis is reinforced by the fact that the two
large range endemics studied here are among the most abundant species of RN [9],
yet they are rare throughout the rest of their distribution [50]. Assessing the amount of
gene flow between RN populations with those of Pitcairn or Austral Island for these
large range endemics could help to test this hypothesis.

The relatively small range of TMRCAs retrieved here could also reflect the
speciation itself for small-range endemics. This has been hypothesized for French
Polynesian small-range endemics which display similarly population expansions older
than the Last Glacial Maximum [17]. Speciation is not an instantaneous process, and
could takes at least 2 My to complete in various vertebrate clades [51]. Divergence
times, upper bound of the speciation time, are known to be quite recent for several
small-range species of this study (Cantherhines rapanui 0.54 Ma (0.07 - 1.3 95%
HPD), Chrysiptera rapanui 1.39 Ma (0.50 - 2.83 95% HPD), Chromis randalli 2.02 (0.58
- 4.3 95 % HPD), Coris debueni: 2.31 (0.72 - 4.11 95% HPD); see [13]). Only a study in
a genomic framework based on pairs of sister species could provide a better estimate
of the timing of speciation (e.g [33]).

Coalescence time does vary across loci and across sets of individuals. By
screening thousands of loci, the inferences made on the demographic history of the
populations should not be biased by the loci observed. However, uncertainty in both
mutation rates and generation times can potentially bias molecular dating. Mutation
rates for SNP data of fish range from $2.5 \times 10^{-8}$ [52] to $3.5 \times 10^{-9}$ [53] and their selection
for estimating demographic histories is not often justified. As such, we chose the fish
SNP mutation rate most often used in the literature. Finally, estimations of generation
times in the wild for fishes are scarce. Most of the generation times we used (six out of
nine) were selected from the literature from species that are phylogenetically close-
related to the species studied here (same genus). Additionally, we used maximal length
to infer generation times for three species as maximum standard length and generation
time are positively correlated [54].

Conclusion

We elucidate for the first time the community assembly time of the RN endemic
reef fish community using genomic data and demographic inferences. We show that
the RN reef fish endemic community shares a common history, with expansion
occurring during the last glacial cycle for the two different types of endemic species.
Local processes based on the seamounts system around RN should have played a
major role in the foundation and persistence of this endemic reef fish community
through dynamic of colonization / extinction / recolonization throughout the RN
Archipelago.

Ethics. All applicable institutional guidelines for the care and use of animals were
followed. Specimens were collected under permit No. 724, 8 March 2016 obtained from
the Chilean Subsecretary of Fishing. The Universidad Austral de Chile Ethical Care
Committee and Biosecurity Protocol approved our use and handling of animals.

Data accessibility. All Fastq sequence files are available from the GenBank at the
National Center for Biotechnology Information short-read archive database (accession
number: forthcoming).

Authors’ contributions.
E.D.-T. and P.S.-A. conceived the study; E.D.-T., P.S.-A. and S.M. acquired the
funding; E.D.-T., E.C.G., V.N, C.R.E. and P.S.-A., collected the field data; EDT, P.C.-
B. and A.S. produced the data; E.D.-T.; E.D.-T., N.H., S.M, and P.S.-A. analyzed the
data; E.D.-T., N.H., E.C.G., S.M, and P.S.-A wrote the manuscript. All authors gave final approval for publication.

Competing interests. We declare we have no competing interests.

Funding. This work was supported by the FONDECYT Postdoctorado fellowship N°3160692 to E. Delrieu-Trottin and by the ATM grant 2016 from the Museum National d'Histoire Naturelle to Stefano Mona.

Acknowledgements. We thank Rebeca Tepano, Nina, Taveke Olivares Rapu, Liza Garrido Toleado (SERNAPESCA), Ludovic Tuki (Mesa del Mar, Te Mau O Te Vaikava O Rapa Nui), and the people of the Rapa Nui Island for their kind and generous support. We thank Luiz A. Rocha for kindly sharing photographs and information related to maximum depth of several RN endemic species. We are grateful to the Genotoul bioinformatics platform Toulouse Midi-Pyrenees (Bioinfo Genotoul; http://bioinfo.genotoul.fr/) for providing computing resources.

References

1. Macarthur R, Levins R. 1967 The Limiting Similarity, Convergence, and Divergence of Coexisting Species. Am. Nat. (doi:10.1086/282505)

2. MacArthur RH, Wilson EO. 1967 The Theory of Island Biogeography. Princeton university press.

3. Emerson BC, Gillespie RG. 2008 Phylogenetic analysis of community assembly and structure over space and time. Trends Ecol. Evol. 23, 619–630. (doi:10.1016/j.tree.2008.07.005)
4. Graham CH, Fine PVA. 2008 Phylogenetic beta diversity: Linking ecological and evolutionary processes across space in time. Ecol. Lett. 11, 1265–1277. (doi:10.1111/j.1461-0248.2008.01256.x)

5. Whittaker RJ, Triantis KA, Ladle RJ. 2008 A general dynamic theory of oceanic island biogeography. J. Biogeogr. (doi:10.1111/j.1365-2699.2008.01892.x)

6. Ricklefs RE, Bermingham E. 2002 The concept of the taxon cycle in biogeography. Glob. Ecol. Biogeogr. 11, 353–361. (doi:10.1046/j.1466-822x.2002.00300.x)

7. Randall JE, Cea A. 2011 Shore Fishes of Easter Island. University of Hawai'i Press. See http://books.google.com/books?hl=en&lr=&id=iXtCTSRTcb8C&pgis=1.

8. Cea A. 2016 Ika Rapa Nui. Rapa Nui Press.

9. Friedlander AM, Ballesteros E, Beets J, Berkenpas E, Gaymer CF, Gorny M, Sala E. 2013 Effects of isolation and fishing on the marine ecosystems of Easter Island and Salas y Gómez, Chile. Aquat. Conserv. Mar. Freshw. Ecosyst. 23, 515–531. (doi:10.1002/aqc.2333)

10. Clouard V, Bonneville A. 2005 Ages of seamounts, islands, and plateaus on the Pacific plate. Spec. Pap. 388 Plates, plumes Paradig. 388, 71–90. (doi:10.1130/0-8137-2388-4.71)

11. Ray JS, Mahoney JJ, Duncan RA, Ray J, Wessel P, Naar DF. 2012 Chronology and geochemistry of lavas from the Nazca Ridge and Easter Seamount Chain: An ~30 myr hotspot record. J. Petrol. 53, 1417–1448. (doi:10.1093/petrology/egs021)

12. Newman W, Foster B. 1983 The Rapanui faunal district (Easter and Sala y
13. Delrieu-Trottin E, Brosseau-Acquaviva L, Mona S, Neglia V, Giles EC, Rapu-Edmunds C, Saenz-Agudelo P. 2019 Understanding the origin of the most isolated endemic reef fish fauna of the Indo-Pacific: Coral reef fishes of Rapa Nui. *J. Biogeogr.* **46**, 723–733. (doi:10.1111/jbi.13531)

14. Bard E, Hamelin B, Arnold M, Montaggioni L, Cabioch G, Faure G, Rougerie F. 1996 Deglacial sea-level record from Tahiti corals and the timing of global meltwater discharge. *Nature* **382**, 241–244. (doi:10.1038/382241a0)

15. Hewitt GM. 2004 The structure of biodiversity -- insights from molecular phylogeography. *Front. Zool.* **1**, 4. (doi:10.1186/1742-9994-1-4)

16. Woodroffe CD, Brooke BP, Linklater M, Kennedy DM, Jones BG, Buchanan C, Mleczko R, Hua Q, Zhao JX. 2010 Response of coral reefs to climate change: Expansion and demise of the southernmost pacific coral reef. *Geophys. Res. Lett.* **37**, L15602. (doi:10.1029/2010GL044067)

17. Delrieu-Trottin E, Mona S, Maynard J, Neglia V, Veuille M, Planes S. 2017 Population expansions dominate demographic histories of endemic and widespread Pacific reef fishes. *Sci. Rep.* **7**, 40519. (doi:10.1038/srep40519)

18. Pinheiro HT, Bernardi G, Simon T, Joyeux J-C, Macieira RM, Gasparini JL, Rocha C, Rocha LA. 2017 Island biogeography of marine organisms. *Nature* **549**, 82.

19. Delrieu-Trottin E, Maynard J, Planes S. 2014 Endemic and widespread coral reef fishes have similar mitochondrial genetic diversity. *Proc. R. Soc. B Biol. Sci.* **281**. (doi:10.1098/rspb.2014.1068)

20. Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE. 2012 Double digest
RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One* 7. (doi:10.1371/journal.pone.0037135)

21. Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH. 2011 Stacks: Building and Genotyping Loci De Novo From Short-Read Sequences. *G3 Genes/Genomes/Genetics* 1, 171–182. (doi:10.1534/g3.111.000240)

22. Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA. 2013 Stacks: An analysis tool set for population genomics. *Mol. Ecol.* 22, 3124–3140. (doi:10.1111/mec.12354)

23. Mastretta-Yanes A, Arrigo N, Alvarez N, Jorgensen TH, Piñero D, Emerson BC. 2015 Restriction site-associated DNA sequencing, genotyping error estimation and de novo assembly optimization for population genetic inference. *Mol. Ecol. Resour.* 15, 28–41. (doi:10.1111/1755-0998.12291)

24. R Core Team. 2017 R: A Language and Environment for Statistical Computing. *R Found. Stat. Comput. Vienna, Austria* 0, {ISBN} 3-900051-07-0. (doi:http://www.R-project.org/)

25. Paradis E. 2010 pegas: an {R} package for population genetics with an integrated--modular approach. *Bioinformatics* 26, 419–420.

26. Liu X, Fu YX. 2015 Exploring population size changes using SNP frequency spectra. *Nat. Genet.* 47, 555–559. (doi:10.1038/ng.3254)

27. Maisano Delser P, Corrigan S, Hale M, Li C, Veuille M, Planes S, Naylor G, Mona S. 2016 Population genomics of C. melanopterus using target gene capture data: demographic inferences and conservation perspectives. *Sci. Rep.* 6, 33753. (doi:10.1038/srep33753)
28. Excoffier L, Dupanloup I, Huerta-Sánchez E, Sousa VC, Foll M. 2013 Robust Demographic Inference from Genomic and SNP Data. *PLoS Genet.* 9. (doi:10.1371/journal.pgen.1003905)

29. Beaumont MA, Zhang W, Balding DJ. 2002 Approximate Bayesian computation in population genetics. *Genetics*

30. Jacobs A, Hughes M, Robinson P, Adams C, Elmer K. 2018 The Genetic Architecture Underlying the Evolution of a Rare Piscivorous Life History Form in Brown Trout after Secondary Contact and Strong Introgression. *Genes (Basel).* 9, 280. (doi:10.3390/genes9060280)

31. Le Moan A, Gagnaire PA, Bonhomme F. 2016 Parallel genetic divergence among coastal-marine ecotype pairs of European anchovy explained by differential introgression after secondary contact. *Mol. Ecol.* (doi:10.1111/mec.13627)

32. Souissi A, Bonhomme F, Manchado M, Bahri-Sfar L, Gagnaire PA. 2018 Genomic and geographic footprints of differential introgression between two divergent fish species (Solea spp.). *Heredity (Edinb).* (doi:10.1038/s41437-018-0079-9)

33. Rougeux C, Bernatchez L, Gagnaire PA. 2017 Modeling the multiple facets of speciation-with-gene-flow toward inferring the divergence history of lake whitefish species pairs (Coregonus clupeaformis). *Genome Biol. Evol.* 9, 2057–2074. (doi:10.1093/gbe/evx150)

34. Tine M *et al.* 2014 European sea bass genome and its variation provide insights into adaptation to euryhalinity and speciation. *Nat. Commun.* 5, 5770. (doi:10.1038/ncomms6770)

35. Oksanen J *et al.* 2012 vegan: Community Ecology Package. R package version
2.1-20/r2309. www.R-project.org, 1–255.

36. Wickham H. 2009 ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. See http://ggplot2.org.

37. Miller KG et al. 2005 The Phanerozoic Record of Global Sea-Level Change. Science (80-. ). 310, 1293–1298. (doi:10.1126/science.1116412)

38. Augustin L et al. 2004 Eight glacial cycles from an Antarctic ice core. Nature (doi:10.1038/nature02599)

39. McManus JF. 2004 A great grand-daddy of ice cores. Nature (doi:10.1038/429611a)

40. Robertson DR. 2004 High genetic diversities and complex genetic structure in an Indo-Pacific tropical reef sh (Chlorurus sordidus): evidence of an unstable evolutionary past? Mar. Biol., 757–767. (doi:10.1007/s00227-003-1224-3)

41. Craig MT, Eble JA, Bowen BW, Robertson DR. 2007 High genetic connectivity across the Indian and Pacific Oceans in the reef fish Myripristis berndti (Holocentridae). Mar. Ecol. Prog. Ser. 334, 245–254. (doi:10.3354/meps334245)

42. Craig MT, Eble JA, Bowen BW. 2010 Origins, ages and population histories: Comparative phylogeography of endemic Hawaiian butterflyfishes (genus Chaetodon). J. Biogeogr. 37, 2125–2136. (doi:10.1111/j.1365-2699.2010.02358.x)

43. Crandall ED, Frey MA, Grosberg RK, Barber PH. 2008 Contrasting demographic history and phylogeographical patterns in two Indo-Pacific gastropods. Mol. Ecol. 17, 611–626. (doi:10.1111/j.1365-294X.2007.03600.x)

44. Fauvelot C, Bernardi G, Planes S. 2003 Reductions in the mitochondrial DNA diversity of coral reef fish provide evidence of population bottlenecks resulting
485 from Holocene sea-level change. *Evolution* **57**, 1571–1583. (doi:10.1111/j.0014-3820.2003.tb00365.x)

486 45. Gaither MR, Toonen RJ, Robertson DR, Planes S, Bowen BW. 2010 Genetic evaluation of marine biogeographical barriers: Perspectives from two widespread Indo-Pacific snappers (Lutjanus kasmira and Lutjanus fulvus). *J. Biogeogr.* **37**, 133–147. (doi:10.1111/j.1365-2699.2009.02188.x)

487 46. Hellberg ME. 2001 Climate-Driven Range Expansion and Morphological Evolution in a Marine Gastropod. *Science* (80-. ). **292**, 1707–1710. (doi:10.1126/science.1060102)

488 47. Horne JB, van Herwerden L, Choat JH, Robertson DR. 2008 High population connectivity across the Indo-Pacific: Congruent lack of phylogeographic structure in three reef fish congeners. *Mol. Phylegenet. Evol.* **49**, 629–638. (doi:10.1016/j.ympev.2008.08.023)

489 48. Klanten OS, Choat JH, Van Herwerden L. 2007 Extreme genetic diversity and temporal rather than spatial partitioning in a widely distributed coral reef fish. *Mar. Biol.* **150**, 659–670. (doi:10.1007/s00227-006-0372-7)

490 49. Winters KL, van Herwerden L, Choat JH, Robertson DR. 2010 Phylogeography of the Indo-Pacific parrotfish Scarus psittacus: Isolation generates distinctive peripheral populations in two oceans. *Mar. Biol.* **157**, 1679–1691. (doi:10.1007/s00227-010-1442-4)

491 50. Randall JE. 2005 *Reef and Shore Fishes of the South Pacific*: New Caledonia to Tahiti and the Pitcairn Islands. University of Hawai’i Press Honolulu.

492 51. Avise JC, Walker D, Johns GC. 1998 Speciation durations and Pleistocene effects on vertebrate phylogeography. *Proc. R. Soc. B Biol. Sci.* **265**, 1707–1712. (doi:10.1098/rspb.1998.0492)
52. Kavembe GD, Kautt AF, Machado-Schiaffino G, Meyer A. 2016 Ecomorphological differentiation in Lake Magadi tilapia, an extremophile cichlid fish living in hot, alkaline and hypersaline lakes in East Africa. *Mol. Ecol.* 25, 1610–1625. (doi:10.1111/mec.13461)

53. Malinsky M, Svardal H, Tyers AM, Miska EA, Genner MJ, Turner GF, Durbin R. 2018 Whole-genome sequences of Malawi cichlids reveal multiple radiations interconnected by gene flow. *Nat. Ecol. Evol.* 2, 1940–1955. (doi:10.1038/s41559-018-0717-x)

54. Froese R, Binohlan C. 2000 Empirical relationships to estimate asymptotic length, length at first maturity and length at maximum yield per recruit in fishes, with a simple method to evaluate length frequency data. *J. Fish Biol.* 56, 758–773. (doi:10.1111/j.1095-8649.2000.tb00870.x)

55. Pante E, Simon-Bouhet B. 2013 marmap: A Package for Importing, Plotting and Analyzing Bathymetric and Topographic Data in R. *PLoS One* 8, 1–4. (doi:10.1371/journal.pone.0073051)

56. Bowen BW, Muss A, Rocha LA, Grant WS. 2006 Shallow mtDNA coalescence in Atlantic pygmy angelfishes (genus Centropyge) indicates a recent invasion from the Indian Ocean. *J. Hered.* 97, 1–12. (doi:10.1093/jhered/esj006)

57. Gajdzik L, Bernardi G, Lepoint G, Frédéric B. 2018 Genetic diversity mirrors trophic ecology in coral reef fish feeding guilds. *Mol. Ecol.* 27, 5004–5018. (doi:10.1111/mec.14936)

58. Easton EE, Sellanes J, Berkenpas E, Gaymer CF, Morales N, Gorny M. 2017 Diversity of deep-sea fishes of the Easter Island Ecoregion. *Deep. Res. Part II* 137, 78–88. (doi:10.1016/j.dsr2.2016.12.006)
**Figure captions**

**Figure 1.** Stairway plots for large-range (red) and small-range endemics (black) representing the variation of effective population size through time and sea level fluctuations for the past 400 000 years using data from [37]. Dark tones indicate glacial era characterized by rises of the sea level while greys tones indicate inter-glacial era with lower sea level.

**Figure 2.** Surface of potential habitat for Rapa Nui endemic reef fish species computed using the package marmap [55] in R [24]. Present conditions (red) and past conditions during the glacial maximums (orange).

**Supplementary Material.**

**Figure S1.** Median of effective population sizes (Ne) based on stairway plots and abc skylines analyses.
Table 1 Ecological data on the species of interest of this study. Range size (S: small-range size; L: large-range size), Family, Maximum Depth where the species have been observed (Max. depth) and reference associated (¹: specimens photographed at those depths by Luiz A. Rocha; ²: personal observation from Cristian Rapu-Edmunds freediving); generation time used for this study and references associated, maximum size of the species. Codes for reproductive strategy are: [p] pelagic eggs, [b] demersal eggs and [m] mouth brooding.

| Range Size | Family               | Species                | Max. depth (m) | Generation time (years) | Max. size (TL, cm) |
|------------|----------------------|------------------------|----------------|-------------------------|--------------------|
| S          | Monacanthidae [b]    | *Cantherhines rapanui* | 20 [7]         | 3                       | 20                 |
| S          | Apogonidae [m]       | *Ostorhinchus chalcius*| 25 [7]         | 2                       | 16                 |
| L          | Pomacanthidae [p]    | *Centropyge hotumatu*  | 50 [7]         | 1 [56]                  | 9                  |
| S          | Pomacentridae [b]    | *Chrysiptera rapanui*  | 72¹            | 1 [57]                  | 7.8                |
| S          | Labridae [p]         | *Coris debueni*        | 70¹            | 3                       | 27                 |
| L          | Holocentridae [p]    | *Myripristis tiki*     | 80²            | 3 [41]                  | 26.5               |
| S          | Pomacentridae [b]    | *Chromis randalli*     | 105¹           | 2 [57]                  | 15                 |
| S          | Chaetodontidae [p]   | *Chaetodon litus*      | 105            | 2 [42]                  | 15.5               |
| S          | Holocentridae [p]    | *Sargocentron wilhimi* | 157 [58]      | 3 [41]                  | 19.5               |
Table 2. Summary of genetic data for each species. Sample size and molecular metrics for each species of the study. Observed (Ho) and expected (He) heterozygosity computed on variable loci; $F_{IS}$, inbreeding coefficient. $\pi$ nucleotide diversity across all loci ($\pi$ across all loci = $\pi$ on variable loci x No. SNP / No. of variable RADSeq loci x length of RAD loci). Maximum (in bold) and minimum (italics) values are highlighted for all columns.

| Species                  | No. of ind | No. of Sites TOTAL | No. of RADSeq loci | No. of variable RADSeq loci | No. SNP | $\pi$ across variable sites | Ho   | He   | $F_{IS}$ | $\pi$ across all sites $^1$ (10^-4) |
|--------------------------|-----------|--------------------|--------------------|------------------------------|---------|-----------------------------|------|------|---------|-----------------------------------|
| Cantherhines rapanui    | 15        | 5 373 281          | 44 233             | 20 052                       | 28 239  | 0.271                       | 0.262| 0.260| 0.023   | 32.73                             |
| Centropyge hotumatua    | 15        | 3 563 265          | 29 195             | 21 684                       | 38 816  | 0.163                       | 0.156| 0.158| 0.032   | 25.21                             |
| Chaetodon litus         | 15        | 5 024 341          | 41 439             | 21 905                       | 33 421  | 0.227                       | 0.220| 0.219| 0.022   | 29.86                             |
| Chromis randalli       | 13        | 4 247 760          | 34 893             | 24 673                       | 43 844  | 0.172                       | 0.161| 0.165| 0.043   | 26.34                             |
| Chrysiptera rapanui    | 18        | 3 104 817          | 25 331             | 19 246                       | 35 735  | 0.145                       | 0.137| 0.141| 0.039   | 23.25                             |
| Coris debueni           | 14        | 5 817 876          | 47 949             | 28 594                       | 45 870  | 0.216                       | 0.208| 0.209| 0.029   | 29.94                             |
| Myripristis tiki       | 18        | 5 522 762          | 45 432             | 28 855                       | 46 718  | 0.165                       | 0.159| 0.160| 0.025   | 23.00                             |
| Ostorhinchus chalcius   | 19        | 7 702 975          | 63 590             | 24 019                       | 32 224  | 0.166                       | 0.165| 0.162| 0.008   | 19.19                             |
|                   |   |    |   |   |   |    |    |    |   |    |
|-------------------|---|----|---|---|---|----|----|----|---|----|
| **Sargocentron wilhmi** | 16| 5 009 870 | 41 234 | 24 207 | 37 860 | 0.224 | 0.217 | 0.217 | 0.026 | 30.98 |
