REPORT

Inter-island local adaptation in the Galápagos Archipelago: genomics of the Galápagos blue-banded goby, *Lythrypnus gilberti*

Giacomo Bernardi¹

Received: 8 March 2021 / Accepted: 27 September 2021 / Published online: 15 October 2021
© The Author(s) 2021

Abstract The Galápagos Archipelago is a place where terrestrial adaptive radiations of finches, mockingbirds, and tortoises have been studied extensively. In contrast, little is known about the potential for marine species to diverge among islands. The overall degradation of coral reefs in the Galápagos makes understanding the mechanisms and factors of speciation, the engine of biodiversity, important, and timely. While speciation in marine archipelagos has been described in the past, such as for cone snails in Cabo Verde Archipelago and limpets in Hawaii, adaptive radiations in the marine environment are still rare and poorly understood. In this study, we focused on the Galápagos blue-banded goby, *Lythrypnus gilberti*, a small endemic fish that is found in shallow subtidal rocky habitats. Using RAD sequencing, we analyzed 19,504 loci that were either neutral, or potentially under directional selection. As expected, considering the small geographic range, population structure based on neutral markers was weak. For loci under directional selection, however, marked differences between islands suggested potential for local adaptation. Our data suggest that for marine species, where dispersal barriers are less apparent, mechanisms of local adaptation may also be at play in the Galápagos Archipelago.

Keywords Galápagos Archipelago · *Lythrypnus gilberti* · Local adaptation · RAD sequencing

Topic Editor Peter Francis Cowman

✉ Giacomo Bernardi
bernardi@ucsc.edu

¹ Department of Ecology and Evolutionary Biology, University of California Santa Cruz, 115 McAllister Way, Santa Cruz, CA 95060, USA

Introduction

The role of isolation has been a central concept to the field of speciation (Coyne and Orr 2004; Price 2007; Rocha and Bowen 2008; Bernardi 2013). In particular, disjunct, fragmented, island-like, and island habitats have been used as models to understand the mechanisms of speciation (e.g., Drovetski et al. 2013; Aguilar et al. 2019; Itescu 2019). In particular, island speciation has received special attention because adaptive radiations, which often accompany speciation events, are best understood in this context (Schluter 2000). Many terrestrial systems are very well-known such as *Drosophila*, spiders, and silverswords, in Hawaii, and tortoises, mockingbirds, and finches in the Galápagos Archipelago (Lack 1947; Kaneshiro 1988; Baldwin and Sanderson 1998; Caccone et al. 1999; Grant 2000; Gillespie 2005; Arbogast et al. 2006; Warren et al. 2015).

In contrast, marine systems, and by extension marine island systems, have traditionally not been seen as conducive to local adaptation and speciation (Pinheiro et al. 2017). While some spectacular examples do exist, such as for Cabo Verde Archipelago *Conus* marine snails (Cunha et al. 2005, 2008; Duda and Rolán 2005; Peters et al. 2016), *Cellana* limpets in Hawaii (Bird et al. 2007, 2011; Bird 2011), Northeastern Pacific *Sebastes* rockfish (Johns and Avise 1998; Alesandrini and Bernardi 1999; Burford and Bernardi 2008), and Antarctic notothenids (Janko et al. 2011), in general, long pelagic larval durations, a common characteristic of marine fishes, have led scientists to assume that island speciation is of limited importance. In fact, even ichthyologist David Starr Jordan suggested to the leader of one of the first scientific expeditions to the Galápagos (1898–1899) not to waste much time on the topic (“The archipelago does not appear to offer a good laboratory for researching the impact of isolation, because

Table 1 Genetic diversity characteristics of populations of *Lythrypnus gilberti* from the Galápagos Archipelago

Sampling site	Number of samples	Number of loci	Variable loci	nA	Ho
San Cristobal (CRI)	10	19,504	6,527	1.47	0.11
Española (ESP)	12	19,504	5,151	1.58	0.11
Floreana (FLO)	19	19,504	5,918	1.66	0.11
Corona del Diablo (CDD)	4	19,504	4,703	1.34	0.11
South Plaza (PLA)	5	19,504	15,368	1.57	0.16
Santa Cruz (ACA)	11	19,504	7,665	1.51	0.11
Los Hermanos (HER)	6	19,504	9,722	1.46	0.13

Columns correspond to (from left to right): Sampling site (see Fig. 1), number of samples, total number of loci, number of usable loci, observed number of alleles (nA), and observed heterozygosity (Ho)

the barren islands do not harbor many isolated fish pools.”, letter to Snodgrass, 1898; In: Larson 2001). Yet, endemic fish species, with narrow geographic range, are interesting models to study the potential for incipient speciation due to their small population sizes, local adaptation, and limited dispersal (Delrieu-Trottin et al. 2014; Tenggardjaja et al. 2016; Crane et al. 2018). In addition, by comparing fish species with different distribution ranges, from very wide-ranging to endemic, evidence of potential inter-island population structure was gathered for endemic Galápagos fishes (Bernardi et al. 2014), leading to the idea that a more powerful approach may reveal patterns relevant to speciation in marine environments.

The Galápagos Islands straddle the equator, but due to their isolation and unusual oceanographic conditions, they harbor a somewhat depauperate coral assemblage with only 22 species (Glynn and Wellington 1983). A devastating El Niño event in 1982–1983 resulted in the loss of the majority of corals in the Archipelago, yet they quickly bounced back, giving some hope as to the resilience of Galápagos reefs (Glynn et al. 2009, 2015). Few true coral reefs are present, historically only 17 reefs were fully developed, instead, most corals are found in scattered colonies. Unfortunately, this unique coral reef system is not doing well. Currently, only one reef, Wellington reef, at Darwin Island in the northernmost portion of the Archipelago, is growing (Glynn et al. 2018). Considering the decline of coral reefs in the Galápagos Archipelago, understanding the factors that promote marine speciation is of crucial importance.

In this study, we tackled the issue of inter-island marine speciation by using the power of genomic sequencing on species endemic to the entire Galápagos Archipelago, *Lythrypnus gilberti*, the Galápagos blue-banded goby (Heller and Snodgrass, 1903). The genus *Lythrypnus* comprises 20 species of small subtidal gobies, geographically restricted to the Tropical Eastern Pacific (TEP) and Tropical Western Atlantic (TWA) (Froese and Pauly 2000). Species studied thus far have been shown to have significant gender flexibility and have been proposed to be

bidirectional sex-changers (Muñoz-Arroyo et al. 2019). *Lythrypnus* is a demersal spawner, with males guarding eggs (Patzner et al. 2011). After hatching, goby larvae typically stay in the water column for approximately 20 to 30 d (precise pelagic larval duration is unknown for the genus *Lythrypnus*), and recruit at a total length of approximately 10 mm (Patzner et al. 2011). Recruitment was shown to be limited by shelter availability and the presence of adults and predators in the congeneric *L. dalli* (Behrents 1987). The Galápagos blue-banded goby, *Lythrypnus gilberti*, is endemic to the Galápagos Archipelago, where it is widely distributed on shallow rocky reefs (Allen and Robertson 1994; Grove and Lavenberg 1997). *Lythrypnus gilberti* belongs to a clade that includes the knave goby, *L. cobalus*, a species endemic to Cocos Island, Costa Rica, and the blue-banded goby, *L. dalli*, that ranges from California to Ecuador (Maxfield et al. 2012).

The Galápagos Archipelago, which straddles the equator, is subjected to three main currents. Two currents move westward: the warm Panama current (South Equatorial

Fig. 1 Map of sampling sites of *Lythrypnus gilberti* in the Galápagos Archipelago. Number of samples for each site is described in Table 1. Arrows show the direction of major currents, blue and red arrows showing cold and warm water, respectively

Current) originating in the north and the cold Humboldt Current originating in the south. One additional current moves cold water eastward: the Cromwell Current (Pacific Equatorial Undercurrent) (Fig. 1) (Chavez and Brusca 1991). These currents result in two main factors: first, larvae may be able to ride currents to move among islands, and second, a strong gradient of temperature and salinity is present in the Archipelago, with western islands typically being colder and eastern islands being warmer. This, in turn, results in major oceanographic and nutrient gradients in the region. For marine organisms, this situation translates in significant longitudinal abiotic and biotic environmental gradients.

While this study focuses on a single organism, looking at intraspecific genetic diversity in an endemic species might shed some light as to the putative processes of divergence that lead to speciation. The goal of this study was, therefore, to focus on the endemic Galàpagos blue-banded goby to determine if there were any evidence of inter-island genetic divergence based on both neutral loci and loci potentially under selection, and if present, relate such findings to local adaptation mechanisms.

Materials and methods

Sampling. Samples of the Galàpagos blue-banded goby, *Lythrypnus gilberti*, were collected while scuba diving with hand nets in December 1999 and January 2000. We tried our best to collect as complete a dataset as possible, while respecting the fragility of the environment, thus collecting the minimum possible number of individuals while maintaining proper sampling design. A total of 67 samples were collected along a longitudinal gradient from seven localities in the southern islands of the Galàpagos Archipelago, including four main islands: San Cristobal (CRI), Española (ESP), Floreana (FLO), and Santa Cruz (Academy Bay, ACA), and three minor islands: Corona del Diablo (CDD), South Plaza (PLA), and Los Hermanos (HER) (Fig. 1, Table 1). Whole fish were immediately placed in 95% ethanol and kept at ambient temperature until reaching the laboratory where they were stored at -20°C . While samples were more than 15 years old when this experiment was performed, we were able to obtain a large number of loci ($\sim 20,000$), which is consistent with the results obtained on historical museum specimens (not preserved at -20°C), where 30-year-old samples did not significantly differ from fresh samples (Suchan et al. 2016).

DNA extraction – Library preparation. DNA was extracted from fin clips using DNeasy Blood & Tissue kits (Qiagen) according to the manufacturer's protocol. We constructed RAD libraries using a variation of the original protocol with the restriction enzyme SbfI (Miller et al.

2007, 2012; Baird et al. 2008; Longo and Bernardi 2015). Initial genomic DNA amount for each individual was 400 ng. Libraries were physically sheared on a Covaris S2 sonicator with an intensity of 5, duty cycle of 10%, cycles/burst of 200 and a cycle time of 30 s. We carried out the final PCR amplification step in 50 μl reaction volumes with 10 amplification cycles. Ampure XP beads (Agencourt) were used for each purification step and size selection. The library was sequenced in a single lane on an Illumina HiSeq 4000 at the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley (USA). Then, we applied Perl scripts to trim the raw reads to 92 base pairs (bp) on the 3' end, quality filtered and demultiplex them according to the 6 bp unique barcodes. Reads with Phred scores of < 33 were discarded. The barcodes and restriction site residues (6 bp) were removed from the 5' end, and this resulted in a final sequence length of 80 bp.

RAD analysis. The software program STACKS version 2.2 (Catchen et al. 2011, 2013) was used to identify orthologous sequences among individuals. In order to optimize the STACKS protocol, we followed published guidelines (Rochette and Catchen 2017). Briefly, we first identified 14 individuals with the highest sequencing coverage and created a catalog of loci based on these samples. (Using the *ustacks* and *cstacks* components of STACKS). In addition, we optimized the values of the parameters *M* and *n*, where the *-M* parameter defines the distance allowed between stacks, whereas the *-n* parameter defines the distance allowed between catalog loci, (*M* and *n* values were kept identical to each other), by varying them from 1 to 9 as recommended. We found that the optimal value for *M* and *n* was 4. We then included the remaining 52 samples using *ustacks*.

Analysis of population structure. Using the population component in STACKS, we further filtered the dataset by retaining loci which aligned in $\geq 80\%$ of individuals (*r* command, *-r* 0.8) in every population. In order to remove paralogs, we used the *minor alleles* function of STACKS following recommendations (*-min_mav* 0.05). We then generated *genepop* files using the *populations* component in STACKS with the *write_single_snp* option (a single SNP is kept for each locus), which were converted afterward using the program PGDSpider V.2.1.0.3 (Lischer and Excoffier 2012) into input files to be used with STRUCTURE.

Population structure was analyzed using STRUCTURE and DAPC approaches. First, structure files from the STACKS populations output were analyzed using a Bayesian approach in STRUCTURE version 2.3.4 (Pritchard et al. 2000). Ten replicate runs were performed for a range of *K* from one to seven, with 10,000 iterations as the burn-in parameter and 100,000 iterations under the admixture model. The highest likelihood for *K* was

estimated according to the Evanno method (Evanno et al. 2005) implemented in Structure Harvester (Earl and VonHoldt 2012). Second, we performed a discriminant analysis of principal components (DAPC) (Jombart et al. 2010; Miller et al. 2020), which combines the benefits of discriminant and principal component analyses, and is particularly useful to study differences between clusters (i.e., sites or populations) as it utilizes a multivariate approach to explore the entire variation in the data while minimizing that within clusters. This analysis was performed using the ADEGENET package in R (Jombart 2008; R CoreTeam 2016) using the genepop file produced by the STACKS populations output used as an input file. The algorithm find.clusters identified the plausible number of clusters by comparing Bayesian information criterion (BIC) values, and the cross-validation tool xvalDapc determined the number of principal components that were retained.

Outlier loci analyses. We examined genetic structure in neutral and outlier loci, separately, to investigate the possibility that different factors have shaped population divergence in unique patterns. Outlier loci were identified using PCADAPT implemented in R (Privé et al. 2020). F_{st} , Φ_{st} values, and their distributions were produced by the STACKS populations output and used to separate loci under directional and stabilizing selection. In order to check for outlier loci found with PCADAPT, we also used the STACKS populations output to determine Φ_{st} (Phi st) values and their distributions to separate loci under directional (Loci with large Φ_{st} , i.e., $\Phi_{st} > 0.1$), and stabilizing selection. Loci identifiers were placed in “whitefiles”, which were used to run population scripts again. Although working with outliers might incorporate a series of shortfalls (Bierne et al. 2011, 2013; Lotterhos and Whitlock 2015), they are commonly used to show evidence of the diverging effects of selection between populations (Gaither et al. 2015; Longo and Bernardi 2015; Bernardi et al. 2016; Stockwell et al. 2016).

Functional analyses

All outlier loci were compared to GenBank entries with BLAST, where E-values of 0.001 and below were kept and recorded (probability of obtaining the same result by chance < 0.001). When protein coding matching sequences were found, they were classified using KEGG assignments (Ogata et al. 1999; Kanehisa et al. 2008).

Results

Loci and polymorphism statistics

Genomic DNA was sequenced from a total of 67 samples on one Illumina lane producing approximately 100 million reads. Following optimization of the genotyping catalog in STACKS (Rochette and Catchen 2017), the total number of loci was 26,789, with 23,787 loci that passed all filtering criteria and a total of 19,504 loci that were polymorphic (Table 1). While the sample numbers varied between collection localities (populations), they did not significantly correlate with the number of usable loci ($R^2 = 0.124$, $p < 0.05$), average number of alleles (i.e., levels of polymorphism, $R^2 = 0.565$, $p < 0.05$), and observed heterozygosity ($R^2 = 0.253$, $p < 0.05$) (Table 1) for each population, suggesting that performing a comparison between populations is appropriate.

Population structure analysis

We ran our population structure analyses by using neutral and outlier loci separately. Below we first describe the analysis based on neutral loci, with a dataset that comprised the loci remaining after removing all outlier loci. Removing outlier loci was done using the blacklist function of STACKS.

STRUCTURE analysis

Neutral loci

The analyses of population structure were based on all orthologous loci identified by STACKS, after removing all outlier loci identified by PCADAPT, which resulted in a dataset comprising 18,547 neutral loci. An analysis of F_{st} 's showed that the average pairwise F_{st} among populations was very low (average: 0.036, maximum: 0.066), with also low values for Φ_{st} (average: 0.0004, maximum: 0.005). A likelihood assessment of the STRUCTURE results obtained by using ten replicates for each value of K (Evanno method) showed that maximum likelihood was reached when $K = 1$ (panmictic populations). However, considering the potential for an incorrect k assignment (Verity and Nichols 2016), and in order to best illustrate the results, we chose to show results for K values of 2 and 3 (Fig. 2).

Fig. 2 Structure plots with Bayesian assignment of *Lythrypnus gilberti* individuals from the Galápagos Archipelago based on neutral markers (18,547 loci) and loci under directional selection (95 loci). Structure Harvester selected a K of 1 for neutral markers (not shown, as it would simply be a solid color) and K of 2 for outlier loci. Additional values of K (2–3) are presented. Each population is defined by vertical bars, with description of the location at the bottom of the figure. See Fig. 1 for map of the locations, and Table 1 for sample numbers for each location

Outlier loci

Based on PCAdapt, we found that 95 loci were potentially under directional selection (outlier loci). As expected, the average pairwise F_{st} among populations was higher for outlier loci (average: 0.094, maximum: 0.142), and values for Φ_{st} were also higher (average: 0.115, maximum: 0.177). Outlier loci obtained with STACKS ($\Phi_{st} > 0.1$) included all 95 loci identified by PCADAPT. The STRUCTURE analysis based on 95 outlier loci showed that genetic clusters clearly partitioned according to sampling sites (Fig. 2), with a highest likelihood reached with $K = 2$ (both $K = 2$ and $K = 3$ are shown, Fig. 2). One genetic cluster was predominantly found at four locations: San Cristobal, Española, Floreana, and Santa Cruz (Figs. 1, 2), which accounted for an average 93.9% of each individual's genome (orange cluster, Fig. 2, K2 panel). In contrast, this genetic cluster was not very commonly represented in the three remaining localities (Fig. 2). Indeed,

the three minor islands of Corona del Diablo, South Plaza, and Los Hermanos, showed strong divergence. Importantly, the sampling site at Corona del Diablo was less than 5 km from the site in Floreana, yet the genetic makeup of individuals at those two locations was very different.

DAPC analyses

DAPC analyses allow us to study population differentiation at a finer scale and higher resolution than a structure analysis (Jombart et al. 2010). We found that, while STRUCTURE could not identify population structure based on neutral loci, DAPC separated individuals belonging to the South Plaza population, and to a lesser extent the Los Hermanos population (Fig. 3). The same

Fig. 3 DAPC cluster plots of *Lythrypnus gilberti* using putative neutral and outlier loci. Plots were created in R using the adegenet package using six discriminant functions, and retaining ten principal components (selected by the validation tool xvalDapc). Each plot illustrates the first and second axes. Population color codes are provided between the two panels, populations discussed in the text are labeled on the panels

analysis based on outlier loci corroborated those findings (separating South Plaza and Los Hermanos populations from the rest) and, to a lesser extent, Corona de Diablo individuals (Fig. 3).

Functional analyses

When comparing outlier loci with GenBank, we found that out of 95 outlier sequences, one sequence did not have a match, five sequences matched bird or mammal genomic sequences and the remaining 89 matched fish sequences, including 37 (38.9%) protein coding sequences (Supplementary Table 1). Out of the 37 protein coding sequences, 20 were GenBank entries for other goby species (*Periophthalmus magnuspinnatus* and *Boleophthalmus pectinirostris*) (Supplementary Table 1). A KEGG analysis showed that the protein coding sequences partitioned in thirteen different functional groups, however more than half (57%) of the sequences were included in only three functional groups (Genetic information processing – 30%, Environmental information processing – 13%, metabolism – 13%) (Fig. 4).

Discussion

The two main forces of evolution, drift and selection, counter and balance each other. When drift (or gene flow) is stronger than selection (or local adaptation), then speciation is not possible, however, when selection is stronger, then genetic divergence may occur. In that respect, speciation in marine environments has traditionally been seen as very different than its terrestrial counterparts for two main reasons: on one hand, physical barriers are less obvious in the ocean, and on the other hand, most marine organisms have long pelagic larval durations that translate in vast

dispersal capabilities, which in turn result in panmixia (Palumbi 1992; Miglietta et al. 2011). However, recent advances in the fields of recruitment biology and molecular genetics have shown that marine organisms show a large proportion of self-recruitment (despite pelagic larval stages), and strong population structure may be present even in widely distributed species (e.g., Bowen et al. 2013; Pinsky et al. 2017; Salas et al. 2019).

The Galàpagos Archipelago is the cradle of speciation science, where Darwin finches and giant tortoises offer clear examples of terrestrial allopatric speciation (Caccone et al. 1999; Grant 2000). Due to the general idea that marine systems are not prone to the same level of speciation, little work has been done underwater (Bernardi et al. 2014). At first glance, this preconceived idea seems correct. Galàpagos Islands are not very distant from each other, the average distance between adjacent islands is only 29.2 km (range 3.7–46.6 km), except for the two northernmost islands of Darwin and Wolf, that are 140 km from the closest island of the main group, Pinta. It would seem difficult, therefore, for marine organisms not to be able to breach such short distances. However, the Archipelago is also bathed in extreme environments, where currents are strong, and oceanographic differences create fronts and temperature gradients that result in warm, low nutrient, and high salinity southeastern islands and cold, high nutrient, and low salinity northwestern islands. Such differences may create opportunities for marine organisms to locally adapt and differentiate, eventually leading to speciation.

In this study, the use of neutral and loci potentially under directional selection (outlier loci) give us an opportunity to determine if this were the case. We selected a small endemic fish, the blue-banded goby, that does not have vast territories, and is likely to lack large dispersal capabilities, as shown in other endemic species (Tenggardjaja et al. 2018). Data presented here may reflect two

Fig. 4 KEGG assignments of outlier loci that match protein coding genes using the KEGG's orthology and link annotation tool (KOALA). Pie depicts the relative percentage of each category

aspects of the biology of *Lythrypnus gilberti*. Although the species, being endemic, does not disperse beyond the Archipelago, our results show that it does disperse enough within the Archipelago to show a lack of population structure when looking at the STRUCTURE plots based on the neutral loci (and very low overall Φ_{st}). On the other hand, important differences in the local environment are present among islands. Surprisingly, differences in two populations were seen in the DAPC analysis of neutral loci, suggesting that some limitation of gene flow between islands is present. In addition, evidence for local adaptation was found as differences in genetic signatures for outlier loci, both in STRUCTURE and DAPC analyses. This is further evidenced when looking at which genes (when present) are in this group, where several outlier genes correspond to the “Environmental information processing” category (Fig. 4).

Galápagos blue-banded gobies, *Lythrypnus gilberti*, show some level of restricted gene flow, as well as the hallmarks of local adaptation. Abiotic factors in the Galápagos Archipelago are pronounced and very distinct, with major current, temperature and nutrient gradients. Such factors, combined with the use of benthic fish with small adult territories and relatively short pelagic larval durations, together with powerful genomic tools allowed to uncover potential signatures of local adaptation. While the Galápagos Archipelago is best known for terrestrial speciation studies, further studies on marine organisms may provide a possible model to understand the mechanisms of local adaptation in the ocean.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00338-021-02189-5>.

Acknowledgements We would like to thank the Galápagos National Park for permission to collect samples under permit Pc-30-99-PNG and the Charles Darwin research station for logistical support. We would also like to thank Nicole Crane, Benjamin Ruttenberg, JB Heiser, Maritza Cardenas, Gabriella Bernardi, and Giorgio Bernardi for help in the field. This research was funded by the University of California Santa Cruz.

Declarations

Conflict of interest The corresponding author states that there is no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted

use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Aguilar C, Miller MJ, Loaiza JR, González R, Krahe R, De León LF (2019) Tempo and mode of allopatric divergence in the weakly electric fish *Sternopygus dardiensis* in the Isthmus of Panama. *Sci Rep* 9:18828
- Alessandrini S, Bernardi G (1999) Ancient species flocks and recent speciation events: What can Rockfish teach us about Cichlids (and vice versa)? *J Mol Evol* 49:814–818
- Allen GR, Robertson DR (1994) Fishes of the tropical eastern Pacific. University of Hawaii Press, Honolulu, Hawaii
- Arbogast BS, Drovetski SV, Curry RL, Boag PT, Grant PR, Grant BR, Anderson DJ, Arbogast BS, Drovetski SV, Curry RL, Boag PT, Grant PR, Grant BR, Anderson DJ (2006) The origin and diversification of Galápagos Mockingbirds. *Evolution* (n Y) 60:370–382
- Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, Johnson EA (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE* 3:e3376
- Baldwin BG, Sanderson MJ (1998) Age and rate of diversification of the Hawaiian silversword alliance (Compositae). *Proc Natl Acad Sci USA* 95:9402–9406
- Behrens KC (1987) The influence of shelter availability on recruitment and early juvenile survivorship of *Lythrypnus dalli* Gilbert (Pisces: Gobiidae). *J Exp Mar Bio Ecol* 107:45–59
- Bernardi G (2013) Speciation in fishes. *Mol Ecol* 22:5487–5502
- Bernardi G, Ramon ML, Alva-Campbell Y, McCosker JE, Bucciarelli G, Garske LE, Victor BC, Crane NL (2014) Darwin's fishes: phylogeography of Galápagos Islands reef fishes. *Bull Mar Sci* 90:533–549
- Bernardi G, Azzurro E, Golani D, Miller MR (2016) Genomic signatures of rapid adaptive evolution in the bluespotted cornetfish, a Mediterranean Lessepsian invader. *Mol Ecol* 25:3384–3396
- Bierne N, Welch J, Loire E, Bonhomme F, David P (2011) The coupling hypothesis: why genome scans may fail to map local adaptation genes. *Mol Ecol* 20:2044–2072
- Bierne N, Roze D, Welch J (2013) Pervasive selection or is it...? why are F_{st} outliers sometimes so frequent...? *Mol Ecol* 22:2061–2064
- Bird CE (2011) Morphological and behavioral evidence for adaptive diversification of sympatric Hawaiian limpets (*Cellana* spp.). *Integr Comp Biol* 51:466–473
- Bird CE, Holland BS, Bowen BW, Toonen RJ (2007) Contrasting phylogeography in three endemic Hawaiian limpets (*Cellana* spp.) with similar life histories. *Mol Ecol* 16:3173–3186
- Bird CE, Holland BS, Bowen BW, Toonen RJ (2011) Diversification of sympatric broadcast-spawning limpets (*Cellana* spp.) within the Hawaiian archipelago. *Mol Ecol* 20:2128–2141
- Bowen BW, Rocha LA, Toonen RJ, Karl SA (2013) The origins of tropical marine biodiversity. *Trends Ecol Evol* 28:359–366
- Burford MO, Bernardi G (2008) Incipient speciation within a subgenus of rockfish (*Sebastes*) provides evidence of recent radiations within an ancient species flock. *Mar Biol* 154:701–717
- Caccone A, Gibbs JP, Ketmaier V, Suatoni E, Powell JR (1999) Origin and evolutionary relationships of giant Galapagos tortoises. *Proc Natl Acad Sci USA* 96:13223–13228

- Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH (2011) Stacks: building and genotyping loci de novo from short-read sequences. *G3* (Bethesda) 1:171–182
- Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool set for population genomics. *Mol Ecol* 22:3124–3140
- Chavez FP, Brusca RC (1991) The Galápagos Islands and their relation to oceanographic processes in the tropical Pacific. In: James M. (eds) *Galápagos Marine Invertebrates*. Plenum Publication Corporation, pp 9–33
- Coyne JA, Orr HA (2004) *Speciation*. Sinauer Associates Inc., Sunderland, MA
- Crane NL, Tariel J, Caselle JE, Friedlander AM, Ross Robertson D, Bernardi G (2018) Clipperton Atoll as a model to study small marine populations: Endemism and the genomic consequences of small population size. *PLoS One* 13
- Cunha RL, Castilho R, Rüber L, Zardoya R (2005) Patterns of cladogenesis in the venomous marine gastropod genus *Conus* from the Cape Verde islands. *Syst Biol* 54:634–650
- Cunha RL, Tenorio MJ, Afonso C, Castilho R, Zardoya R (2008) Replaying the tape: recurring biogeographical patterns in Cape Verde *Conus* after 12 million years. *Mol Ecol* 17:885–901
- 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 London* 281:20141068
- Drovetski SV, Semenov G, Drovetskaya SS, Fadeev IV, Red'kin YA, Voelker G (2013) Geographic mode of speciation in a mountain specialist Avian family endemic to the Palearctic. *Ecol Evol* 3:1518–1528
- Duda TF, Rolán E (2005) Explosive radiation of Cape Verde *Conus*, a marine species flock. *Mol Ecol* 14:267–272
- Earl DA, VonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour* 4:359–361
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14:2611–2620
- Froese R, Pauly D (2000) *FishBase 2000: concepts, design and data sources*. ICLARM, Philippines, Los Baños, Laguna
- Gaither MR, Bernal MA, Coleman RR, Bowen BW, Jones SA, Simison WB, Rocha LA (2015) Genomic signatures of geographic isolation and natural selection in coral reef fishes. *Mol Ecol* 24:1543–1557
- Gillespie RG (2005) Geographical context of speciation in a radiation of Hawaiian *Tetragnatha* spiders (Araneae, Tetragnathidae). *J Arachnol* 33:313–322
- Glynn PW, Wellington GM (1983) *Corals and coral reefs of the Galapagos islands*. University of California Press, Berkeley
- Glynn P, Riegl B, Correa AM, Baums I (2009) Rapid Recovery of a Coral Reef at Darwin Island, Galapagos Island. *Galapagos Res* 66:6–13
- Glynn PW, Riegl B, Purkis S, Kerr JM, Smith TB (2015) Coral reef recovery in the Galápagos Islands: the northernmost islands (Darwin and Wenman). *Coral Reefs* 34:421–436
- Glynn PW, Feingold JS, Baker A, Banks S, Baums IB, Cole J, Colgan MW, Fong P, Glynn PJ, Keith I, Manzello D, Riegl B, Ruttenberg BI, Smith TB, Vera-Zambrano M (2018) State of corals and coral reefs of the Galápagos Islands (Ecuador): Past, present and future. *Mar Pollut Bull* 133:717–733
- Grant PR (2000) *Ecology and Evolution of Darwin's finches*. Princeton University Press, Princeton, New Jersey
- Grove JS, Lavenberg RJ (1997) *The Fishes of the Galápagos Islands*. Stanford University Press
- Heller E, Snodgrass RE (1903) Papers from the Hopkins Stanford Galapagos expedition, 1898–1899. *Proc Wash Acad Sci* 5:189–229
- Itescu Y (2019) Are island-like systems biologically similar to islands? A review of the evidence. *Ecography (cop)* 42:1298–1314
- Janko K, Marshall C, Musilová Z, Van Houdt J, Couloux A, Cruaud C, Lecointre G (2011) Multilocus analyses of an Antarctic fish species flock (Teleostei, Notothenioidei, Trematominae): phylogenetic approach and test of the early-radiation event. *Mol Phylogenet Evol* 60:305–316
- Johns GC, Avise JC (1998) Tests for ancient species flocks based on molecular phylogenetic appraisals of *Sebastes* rockfishes and other marine fishes. *Evolution* (n Y) 52:1135–1146
- Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24:1403–1405
- Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet* 11
- Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T, Yamanishi Y (2008) KEGG for linking genomes to life and the environment. *Nucleic Acids Res* 36:480–484
- Kaneshiro KY (1988) Speciation in the Hawaiian “*Drosophila*”: Sexual selection appears to play an important role. *Bioscience* 38:258–263
- Lack D (1947) *Darwin's finches*. Cambridge University Press, Cambridge
- Larson EJ (2001) *Evolution's Workshop: God and Science on the Galapagos Islands*. Basic Books and London, Penguin, New York
- Lischer HEL, Excoffier L (2012) PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics* 28:298–299
- Longo G, Bernardi G (2015) The evolutionary history of the embiotocid surfperch radiation based on genome-wide RAD sequence data. *Mol Phylogenet Evol* 88:55–63
- Lotterhos KE, Whitlock MC (2015) The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Mol Ecol* 24:1031–1046
- Maxfield JM, Van Tassell JL, St. Mary CM, Joyeux JC, Crow KD, (2012) Extreme gender flexibility: Using a phylogenetic framework to infer the evolution of variation in sex allocation, phylogeography, and speciation in a genus of bidirectional sex changing fishes (*Lythrypnus*, Gobiidae). *Mol Phylogenet Evol* 64:416–427
- Miglietta MP, Faucci A, Santini F (2011) Speciation in the sea: overview of the symposium and discussion of future directions. *Integr Comp Biol* 51:449–455
- Miller MR, Dunham JP, Amores A, Cresko WA, Johnson EA (2007) Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Res* 17:240–248
- Miller MR, Brunelli JP, Pa W, Liu S, Rexroad CE, Palti Y, Doe CQ, Thorgaard GH (2012) A conserved haplotype controls parallel adaptation in geographically distant salmonid populations. *Mol Ecol* 21:237–249
- Miller JM, Cullingham CI, Peery RM (2020) The influence of a priori grouping on inference of genetic clusters: simulation study and literature review of the DAPC method. *Heredity (Edinb)*
- Muñoz-Arroyo S, Rodríguez-Jaramillo C, Balart EF (2019) The goby *Lythrypnus pulchellus* is a bi-directional sex changer. *Environ Biol Fishes* 102:1377–1391
- Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M (1999) KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 27:29–34

- Palumbi SR (1992) Marine speciation on a small planet. *Trends Ecol Evol* 7:114–118
- Patzner RA, Van Tassel MK, Kapoor BG (2011) The biology of Gobies. CRC Press, Taylor & Francis Group, Science Publishers
- Peters H, O’Leary BC, Hawkins JP, Roberts CM (2016) The cone snails of Cape Verde: Marine endemism at a terrestrial scale. *Glob Ecol Conserv* 7:201–213
- 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
- Pinsky ML, Saenz-Agudelo P, Salles OC, Almany GR, Bode M, Berumen ML, Andréfouët S, Thorrold SR, Jones GP, Planes S (2017) Marine dispersal scales are congruent over evolutionary and ecological time. *Curr Biol* 27:149–154
- Price T (2007) Speciation in birds. Roberts & Company Publishers, Greenwood Village, Colorado
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959
- Privé F, Luu K, Vilhjálmsson BJ, Blum MGB, Rosenberg M (2020) Performing highly efficient genome scans for local adaptation with R Package pcadapt Version 4. *Mol Biol Evol* 37:2153–2154
- R CoreTeam R (2016) R: A Language and Environment for Statistical Computing
- Rocha LA, Bowen BW (2008) Speciation in coral-reef fishes. *J Fish Biol* 72:1101–1121
- Rochette NC, Catchen JM (2017) Deriving genotypes from RAD-seq short-read data using Stacks. *Nat Protoc* 12:2640–2659
- Salas EM, Bernardi G, Berumen ML, Gaither MR, Rocha LA (2019) RADseq analyses reveal concordant Indian Ocean biogeographic and phylogeographic boundaries in the reef fish *Dascyllus trimaculatus*. *R Soc Open Sci* 6
- Schluter D (2000) The ecology of adaptive radiation. Oxford University Press, Oxford, UK
- Stockwell BL, Larson WA, Waples RK, Abesamis RA, Seeb LW, Carpenter KE (2016) The application of genomics to inform conservation of a functionally important reef fish (*Scarus niger*) in the Philippines. *Conserv Genet* 17:239–249
- Suchan T, Pitteloud C, Gerasimova NS, Kostikova A, Schmid S, Arrigo N, Pajkovic M, Ronikier M, Alvarez N (2016) Hybridization capture using RAD probes (hyRAD), a new tool for performing genomic analyses on collection specimens. *PLoS ONE* 11:1–22
- Tenggardjaja KA, Bowen BW, Bernardi G (2016) Reef fish dispersal in the Hawaiian Archipelago: comparative phylogeography of three endemic damselfishes. *J Mar Biol* 2016
- Tenggardjaja KA, Bowen BW, Bernardi G (2018) Comparative phylogeography of widespread and endemic damselfishes in the Hawaiian Archipelago. *Mar Biol* 165:1–21
- Verity R, Nichols RA (2016) Estimating the number of subpopulations (K) in structured populations. *Genetics* 203:1827–1835
- Warren BH, Simberloff D, Ricklefs RE, Aguilée R, Condamine FL, Gravel D, Morlon H, Mouquet N, Rosindell J, Casquet J, Conti E, Cornuault J, Fernández-Palacios JM, Hengl T, Norder SJ, Rijdsdijk KF, Sanmartín I, Strasberg D, Triantis KA, Valente LM, Whittaker RJ, Gillespie RG, Emerson BC, Thébaud C (2015) Islands as model systems in ecology and evolution: Prospects fifty years after MacArthur-Wilson. *Ecol Lett* 18:200–217

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.