

## Original Article

TITLE:

Linking micro- and macroevolutionary perspectives to evaluate the role of Quaternary sea-level oscillations in island diversification

Anna Papadopoulou<sup>1,2</sup> and L. Lacey Knowles<sup>2</sup>

<sup>1</sup> *Department of Integrative Ecology, Estación Biológica de Doñana (EBD-CSIC), Avda. Américo Vespucio s/n, 41092 Seville, Spain*

<sup>2</sup> *Department of Ecology and Evolutionary Biology, 1109 Geddes Ave., Museum of Zoology, University of Michigan, Ann Arbor, MI 48109-1079*

*E-mail: apapad08@ucy.ac.cy (A.P.); knowlesl@umich.edu (L.L.K.)*

Corresponding author: Anna Papadopoulou; *apapad08@ucy.ac.cy*

RUNNING TITLE: Micro- to macroevolutionary island dynamics

KEYWORDS:

Puerto Rico Bank; Virgin Islands; ground crickets; island fission-fusion; Rad-seq

WORD COUNT: 7,977 TABLES: 3; FIGURES: 5

DATA ARCHIVAL LOCATION: DOI: DOI:10.5061/DRYAD.CC000

ACKNOWLEDGEMENTS

This work was funded by the Hubbell and Ammerman endowments at the University of Michigan, Museum of Zoology, and by NSF (DEB 1118815 to LLK). This project was made possible by the availability of specimens collected as part of projects sponsored by NSF (including DEB 0715487 to LLK) and in collaboration with Elen Oneal and Dan Otte. AP acknowledges financial support from the Spanish Ministry of Economy and Competitiveness, through the Severo Ochoa Program for Centres of Excellence in R+D+I (SEV-2012-0262). We wish to thank Qixin He, Jen-Pan Huang and

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/evo.13384](https://doi.org/10.1111/evo.13384).

This article is protected by copyright. All rights reserved.

Andrea Thomaz for help with analyses and discussions. We are also grateful to the Associate Editor Matthew Carling, to Rafe Brown and to five anonymous reviewers for their constructive comments that helped us to improve the manuscript. The authors declare no conflicts of interest.

#### ABSTRACT:

With shifts in island area, isolation, and cycles of island fusion-fission, the role of Quaternary sea-level oscillations as drivers of diversification is complex and not well understood. Here we conduct parallel comparisons of population and species divergence between two island areas of equivalent size that have been affected differently by sea-level oscillations, with the aim to understand the micro- and macroevolutionary dynamics associated with sea-level change. Using genome-wide datasets for a clade of seven *Amphiacusta* ground cricket species endemic to the Puerto Rico Bank (PRB), we found consistently deeper interspecific divergences and higher population differentiation across the unfragmented Western PRB, in comparison to the currently fragmented Eastern PRB that has experienced extreme changes in island area and connectivity during the Quaternary. We evaluate alternative hypotheses related to the microevolutionary processes (population splitting, extinction and merging) that regulate the frequency of completed speciation across the PRB. Our results suggest that under certain combinations of archipelago characteristics and taxon traits the repeated changes in island area and connectivity may create an opposite effect to the hypothesized “species pump” action of oscillating sea levels. Our study highlights how a microevolutionary perspective can complement current macroecological work on the Quaternary dynamics of island biodiversity.

#### INTRODUCTION

There is a long history of studying the ecological and evolutionary processes related to the generation and maintenance of biodiversity on oceanic islands (Warren et al. 2015; Patiño et al. 2017). Among a variety of foci on explanatory mechanisms, the potential role of Quaternary sea-level oscillations in shaping the diversity of island biota has been explored and debated in well-studied

archipelagos such as Hawaii or the Philippines (e.g., Jordan et al. 2005; Esselstyn et al. 2009; Siler et al. 2010; Brown et al. 2013). Recent meta-analyses have identified a strong signature of Quaternary shifts in island area (Weigelt et al. 2016) and cycles of island fusion-fission (Rijsdijk et al. 2014) on species diversity patterns of island endemics. In the light of these empirical findings, Quaternary-sensitive macroecological models of island biogeography have been proposed (Fernández-Palacios et al. 2016) that predict how immigration and extinction rates may vary between glacial and interglacial periods due to shifts in island area and isolation. Even though the effects of Quaternary sea-level dynamics on island biodiversity are now undisputable from a macroecological perspective, a thorough understanding of the underlying evolutionary processes is still required.

Several factors make evolutionary generalization about the divergence processes associated with dynamic island histories especially challenging. For example, estimates of the timing of divergence based on a few genetic loci have had limited utility because uncertainty surrounding such time estimates encompasses glacial-interglacial cycles (e.g., Carstens and Knowles 2007), making it unclear what geographic configurations might drive divergence (e.g., past island connections versus isolation). Alternative approaches to testing the hypothesized “species pump” action of rising and falling sea-levels (Heaney 1985; Brown and Diesmos 2009) based on evidence of concordant topologies and divergence times among multiple co-distributed taxa have garnered little empirical support (Esselstyn and Brown 2009; Siler et al. 2010; Oaks et al. 2013; but see Papadopoulou and Knowles 2015b; 2016), while recent refinements of the model (e.g., the “oscillating geography mechanism”; Ali and Aitchison 2014), which predict different evolutionary outcomes depending on dispersal rates and/or speciation duration, remain to be tested. Interestingly, these evolutionary hypotheses have focused primarily on the shifts in island connectivity as potential driver of diversification and have not explicitly considered as an independent variable the simultaneous changes in island area, which have been identified as a key driver from a macroecological perspective (MacArthur and Wilson 1963, 1967; Fernández-Palacios et al. 2016; Weigelt et al. 2016).

Although a current theoretical framework for testing the combined effects of shifting island area and connectivity on the process of divergence is lacking, insights might be gained by studying

simultaneously the microevolutionary processes acting at short evolutionary time-scales and the observed macroevolutionary patterns. In particular, a thorough understanding of the underlying microevolutionary processes is a prerequisite for developing Quaternary-sensitive models of island biogeography, given that processes affecting population persistence and divergence over short evolutionary time-scales (Papadopoulou and Knowles 2015a) can have an important impact on species diversification and diversity patterns (Dynesius and Jansson 2014). Specifically, the effects on species diversity are mediated by the processes that regulate the frequency of completed speciation (Dynesius and Jansson 2014), namely (i) the rate of initiation of within-species lineages (splitting; Fig 1a) and (ii) the degree of persistence of these lineages, with failed persistence occurring either by local extinction or by merging due to gene flow (Fig. 1a). The rates of splitting, extinction and merging are predicted to vary during high-sea level (interglacial) versus low sea-level (glacial) periods. However, the effects of these sea-level changes under each of the two major proposed hypotheses are predicted to differ (Fig. 1b, 1c). Specifically, if shifts in island area are a main driver of island diversity (Fernández-Palacios et al. 2016; see also MacArthur and Wilson 1963; MacArthur and Wilson 1967; Heaney 2000; Losos and Schluter 2000; Kisel and Barraclough 2010; Weigelt et al. 2016), then the rate of lineage splitting is expected to be higher during low sea-level periods (Fig. 1b), whereas if shifts in island connectivity are a main driver of island diversity through a “species pump” mechanism (Heaney 1985; Ali and Aitchison 2014), then the rate of lineage splitting is expected to be higher during high sea-levels (Fig. 1c). Given the contrasting predictions (Fig. 1b, 1c), the overall impact of sea-level oscillations on island biodiversity when both island area and connectivity change through time remains unclear and requires further empirical investigation.

Here we focus on a clade of ground crickets (genus *Amphiacusta*) that has diversified across the Puerto Rico Bank (PRB) in the Caribbean (Oneal et al. 2010). *Amphiacusta* (Gryllidae: Phalangopsinae) is a diverse genus of flightless nocturnal ground crickets restricted to the Caribbean (~80 species across the Greater and Lesser Antilles, with most species being single-island endemics; Desutter-Grandcolas and Otte 1997). Spatial isolation and limited vagility have been shown to play an important role in the diversification of *Amphiacusta* across the Greater Antilles (Oneal et al. 2010).

Although recent evolutionary dynamics may not affect all taxa equally (see literature on the Philippine fauna e.g., Brown and Alcala 1970; Diamond and Gilpin 1983; Brown et al. 2013), they are expected to be particularly important for taxa with more limited dispersal and/or a rapid pace of evolutionary diversification, such as these flightless insects. Moreover, this system provides the opportunity to conduct parallel comparisons of species diversity and genetic divergence to evaluate alternative hypotheses related to the drivers of the diversification process under contrasting island settings. Specifically, we compare population and species divergence patterns in crickets endemic to two regions that span similar geographic areas (Fig. 2a), but contrast in their exposure to sea-level oscillations (Fig. 2b). One region – the Eastern PRB (corresponding to the present-day islands of Vieques, Culebra and most of the US and British Virgin Islands) – has been partially and periodically submerged during high sea-level periods, and therefore it has repeatedly shifted in both area and connectivity (Fig 2b). In contrast, the other region – the Western PRB (corresponding to present-day Puerto Rico) – has remained continuously emerged, with only minor shifts in island area (Fig 2b). Additionally, the two regions of the PRB differ in species richness and endemism (as documented in different groups of organisms, e.g., Heatwole and MacKenzie 1967; Hedges 1999; Scheffrahn et al. 2003), which is also mirrored by diversity patterns in the crickets, with twice the richness on the continuously emerged region of the Western PRB compared to the dynamic Eastern PRB (Table 1).

By analyzing genome-wide SNP data for seven *Amphiacusta* species: (i) we contrast patterns of intraspecific genetic divergence between species confined to each of the two regions, and (ii) to characterize the degree of differentiation within each of the regions in the recent versus the more distant past, given that the spatial distribution of alleles reflects demographic processes since mutation (Slatkin 1985; Barton and Slatkin 1986), we further examine separately the intraspecific genetic structure based on rare versus intermediate polymorphisms. Lastly, (iii) we infer the phylogeny of the entire clade and contrast patterns of interspecific divergence and phylogenetic diversity between the two regions. By comparing the results of the above analyses, we discuss how the island dynamics are likely to influence intra- and interspecific diversification on the Eastern PRB because of the impact of sea-level oscillations. Lastly, we consider diversity and genetic divergence of *Amphiacusta* taxa on

the small adjacent island of Mona, which has not been connected to the Puerto Rico Bank, to evaluate whether contemporary island size, by itself, is sufficient to explain the differences in diversification histories between the Eastern and Western PRB regions. Although there are other possible factors not tested here that might structure divergence processes (see discussion), our study nonetheless highlights how a comparative analysis of population and species divergence can help to understand the evolutionary dynamics associated with the shifts in island area and connectivity due to sea-level oscillations.

## METHODS

### *Sampling and genomic library preparation*

For this study, we selected 121 specimens of *Amphiacusta*, representing seven species sampled from 16 localities across the Puerto Rico Bank (specifically from the islands of Puerto Rico, Vieques, Culebra, St. Thomas, Tortola and Virgin Gorda) and five localities on the adjacent island of Mona (Table 2, Fig. 2), spanning the known distribution of these species. Each of the seven species was represented by at least seven individuals, except for *A. tijicohniae* with three available specimens that are only included in interspecific analyses for reconstructing the evolutionary relationships across the clade. Three of the species (*A. pronauta* and *A. lares* from the Western PRB and *A. sanctaecrucis* from the Eastern PRB; selected based on specimen availability) were sampled more extensively for intraspecific analyses (3–5 populations per species, 3–10 specimens per population; see Table 2 for sample sizes per species and locality). Genomic DNA was extracted from the femur of each individual using the Qiagen DNeasy Blood and Tissue Kit.

Three reduced representation libraries (each of 48 individuals, after pooling the samples with 23 individuals from another project) were constructed using a double digest restriction-fragment-based approach (ddRADseq), following the protocol of Peterson et al. (2012). In summary, DNA was double-digested with the restriction enzymes EcoR1 and MseI, unique barcodes (10bp) and Illumina

adapter sequences were ligated to the digested fragments and the individually barcoded products were size-selected between 350-450bp using a Pippin Prep (Sage Science) machine. After size-selection, the fragments were PCR-amplified using high-fidelity DNA polymerase (iProof, BioRad) with 12 cycles. The libraries were sequenced in three lanes of an Illumina HiSeq 2500 (Rapid Run Mode) at The Centre for Applied Genomics (Hospital for Sick Children, Toronto, Canada) to generate 150 base pair, single-end reads.

#### *Processing of Illumina data*

The raw Illumina reads were demultiplexed and quality-filtered using the *process\_radtags.pl* script from the pipeline STACKS v. 1.32 (Catchen et al. 2011; Catchen et al. 2013). Only reads with a Phred score  $>10$  (using a sliding window of 15%) and unambiguous barcodes were retained. Sequence quality inspection using fastQC (Andrews 2010) and examination of the variation of the number of SNPs across the length of the reads based on preliminary data processing suggested a slight increase of sequence error towards the 3' end of the reads. Based on these inspections, 24bp were trimmed off the 3' end of the reads prior to any further processing, using SEQTK (Heng Li, <https://github.com/lh3/seqtk>). After removing barcodes and the enzyme cutsite a read length of 110bp was retained for downstream analyses. An additional subsequent filtering step was performed, where trimmed reads containing more than two sites with a phred score  $< 20$  were removed.

For analyses at the interspecific level, which included all seven species, the reads were processed using the program PYRAD 3.0.5 (Eaton 2014) that accounts for indel variation and is thus appropriate for interspecific data. The reads of each sample were clustered into putative loci using a 90% similarity threshold, a maximum number of indels per cluster of 3 and a minimum depth for making a statistical base call at each site of 5. The same similarity threshold was used for clustering loci across individuals (following the author's recommendation) with a maximum of 6 indels per locus. Note that the selected values are in line with other recent studies focusing on within-genus divergences (e.g., Huang 2016; Razkin et al. 2016; Haponski et al. 2017), and by applying the same

criteria across all samples, our conclusions are not likely to suffer from systematic biases, as they are primarily based on among-species comparisons rather than on the estimated parameters *per se*. Potential paralogs were identified and discarded; specifically, loci containing one or more heterozygous sites shared across > 8 individuals were removed. All remaining loci found in at least four samples were exported for further processing. The 3' ends of the aligned loci were trimmed off, so that all loci had the same final alignment length (110bp). After exporting the data from PYRAD, the number of SNPs per inferred locus was examined to guard against excessive variation arising from the clustering of non-orthologous loci due to errors of the *de novo* assembly procedure. Specifically, loci were removed from the dataset using a custom script (available on GitHub/KnowlesLab; Huang 2016) if: (i) the estimated theta per locus within each species > 0.02 (threshold based on upper limits of biologically realistic effective population sizes and the spontaneous genomic mutation rate of *Drosophila*, Keightley et al. 2009), or (ii) the pairwise genetic divergence between species > 0.15 (threshold based on COI data from these species available on GenBank, Oneal et al. 2010). A dataset containing only putatively unlinked biallelic SNPs (i.e. one SNP per locus, and specifically, the one closest to the 5' end was chosen in order to reduce the possibility for sequence error) sequenced in four or more populations were exported for further analyses at the interspecific level (specifically for species tree reconstruction which is based on quartets and therefore can utilize all loci found in at least four terminal taxa; see more details in the respective section below), while less inclusive datasets were used when required (see details about downstream analyses below).

For intraspecific analyses, which focused on three of the species that were sampled more extensively at the population level, the Illumina reads of *A. lares*, *A. pronauta* and *A. sanctaecrucis* were processed separately in STACKS v. 1.32 (Catchen et al. 2011; Catchen et al. 2013). The reads for each individual were assembled *de novo* into putative loci using the USTACKS program, with a minimum stack depth (m) of 5 and distance allowed between stacks (M) of 2. SNPs were identified at each locus and genotypes were called using a multinomial-based likelihood model that accounts for sequencing error (Hohenlohe et al. 2010; Catchen et al. 2011; Catchen et al. 2013), with the upper bound of the error rate ( $\epsilon$ ) set to 0.1, as the unbounded model has been shown to underestimate



heterozygotes (Catchen et al. 2013). The ‘removal algorithm’ was used to remove unexpectedly deep stacks (i.e., stacks that exceed the expected number of reads for a single locus given the average depth of coverage, as expected when loci are members of multi-gene families) and the ‘deleveraging algorithm’ was used to resolve over-merged loci (i.e., non-homologous loci misidentified as a single locus). A catalog of consensus putative homologous loci among individuals of each species was built in CSTACKS, with the number of mismatches allowed between individuals ( $n$ ) set to 2, and each individual was matched against the respective catalog using SSTACKS. The selected settings have been widely used for intraspecific datasets of similar read length across a range of organisms (e.g., Lanier et al. 2015; Barker et al. 2017) and although varying these parameters may affect the number of recovered loci due to over-splitting of homologous loci (or vice versa merging of non-homologous loci), it has been shown to have a minimum impact on the resulting  $F_{ST}$  and genetic distances among populations (Harvey et al. 2015), therefore the choice of these parameters should not significantly bias the conclusions of the present study. All loci found in at least two populations and a minimum of 0.25 of the sampled individuals per population were exported using the program POPULATIONS. Subsequently, the exported loci were filtered for excessive number of SNPs that could result from merging non-homologous loci using a custom R script (available on GitHub/KnowlesLab; Thomaz et al. 2017) removing loci if per locus estimates of theta were within the 95% quantile of the estimated theta values. The number of SNPs per sequence position was plotted and inspected by eye, and five positions at the 3’ end that had an increased number of SNPs relative to the other sites (attributed to artifacts of the *de novo* assembly algorithms) were trimmed. A whitelist was created including all loci that passed the filtering step and the program POPULATIONS was run again. A single SNP per locus was exported in Variant Call Format (vcf) and then converted to other required file formats using the SNPRELATE package (Zheng et al. 2012) in R (R Core Team 2016) and the program PLINK (Purcell et al. 2007). Different levels of missing data and of minor allele frequencies were used depending on the requirements of the conducted analyses (see respective sections below).

## *Species tree and interspecific divergence*

The interspecific SNP dataset resulting from PYRAD for all seven species was used to infer the phylogeny of the group and contrast patterns of interspecific divergence and phylogenetic diversity between the Eastern and the Western PRB. To infer evolutionary relationships among species and among populations we performed species tree reconstruction using the SVDQUARTETS method (Chifman and Kubatko 2014), as implemented in PAUP\*4.0a146. This approach accounts for the differences in the genealogical history of individual loci expected to arise under a multispecies coalescent model and has been developed specifically for SNP data (see supplement for an analysis of the concatenated dataset using the program RAXML v. 8, Stamatakis 2014). Specifically, SVDQUARTETS was used to estimate the best-supported topology for each possible quartet of taxa in the dataset based on the observed site pattern distribution. These quartets were then assembled using the QFM quartet amalgamation (Reaz et al. 2014) to infer a species/population tree. Clade support was assessed with 100,000 non-parametric bootstrap replicates. As SVDQUARTETS is robust to missing data (Chifman and Kubatko 2014), the most inclusive dataset was used for this analysis (i.e., one biallelic SNP per each locus that was sequenced in at least four populations and was therefore informative for quartet reconstruction). For these analyses we pooled individuals of the neighboring populations on Vieques (VIM-VIB), Culebra (CUV-CUF) and Mona (CDG1-JCS-SRS and CDG2-MHD-PCC-SBC3), after confirming lack of interpopulation differentiation among the respective populations on these three islands in preliminary analyses. The trees were rooted on the clade containing *A. tijiconhia*, *A. mona* and *A. sp.*, based on previous studies on the phylogeny of the genus *Amphiacusta* (Oneal et al. 2010).

In order to provide a relative time component to the diversification of the group and estimate phylogenetic diversity within each of the island regions, we additionally conducted species-tree reconstruction using SNAPP v. 1.2.2 (Bryant et al. 2012), as implemented in BEAST2 v. 2.3.0 (Bouckaert et al. 2014). A reduced dataset was used for this analysis, as this program is very computationally intensive and does not allow missing loci among terminal taxa (i.e., a locus must be

sequenced in at least one representative of each population/species). For these analyses, each species was represented by three individuals with the highest number of reads. Additional analyses at the population level (including three individuals per population) were conducted for a reduced taxon set and more inclusive set of loci, after removing the divergent species (*A. tijicohniae*, *A. mona* and *A. sp.*), which shared a much smaller proportion of loci with the rest of the clade (see Results section). SNAPP analyses do not require defining outgroup, as the program samples the root position along with the other nodes of the tree. An independent theta ( $\theta = 4\mu N_e$ ) was estimated for each branch under (i) the default gamma (11.75, 109.73) prior distribution or (ii) a gamma (2,2000) prior distribution, to ensure that the inferred topology was robust to prior specification (Rheindt et al. 2014). The backward and forward mutation rates,  $u$  and  $v$ , were co-estimated, with the total expected number of mutations per unit time constrained to 1 ( $2u*v/(u+v)=1$ ), using initial values based on the stationary frequencies. Default values were used for all other prior and operator settings. Two independent runs of 6 million generations were performed for each dataset, sampling every 1000 generations. Convergence and mixing of the individual runs was assessed by inspection of the trace plots and the effective sample sizes (>200) in the program TRACER v1.6.0 (Rambaut et al. 2014) and the samples from the two runs were combined after excluding 10% of the generations as burnin. Resulting tree files were visualized using the program DENSITREE (Bouckaert 2010). The total phylogenetic branch length (Phylogenetic Diversity, Faith 1992) for each of three regions (Eastern PRB, Western PRB, and Mona) was estimated using the R package *picante* (Kembel et al. 2010). Note that the species *A. viequesense*, which occurs in both the Western and the Eastern PRB, was considered for the estimation of phylogenetic diversity in both areas.

#### *Genetic diversity and population differentiation*

The three intraspecific SNP datasets of *A. lares*, *A. pronauta* and *A. sanctaecrucis* resulting from processing in STACKS, including all loci found in at least half of the sampled populations and half of the sampled individuals per population, were used to estimate genetic diversity and genetic

differentiation among populations across the Western and the Eastern PRB. The POPULATIONS program in STACKS was used to calculate genetic diversity statistics for each population, including nucleotide diversity ( $\pi$ ) and observed heterozygosity at each locus and average values across loci, as well as pairwise  $F_{ST}$  values among populations. A test of Isolation-by-Distance was performed for *A. pronauta* and *A. sanctaerucis* that were sampled from at least 4 localities. Specifically, we tested for a correlation between pairwise  $F_{ST}$ -values and geographic distances, using a Mantel test (Mantel 1967) as implemented in the R package VEGAN (Oksanen et al. 2013) with significance assessed with 1 million permutations. Given the concerns about the reliability of Mantel tests (e.g., Legendre and Fortin 2010; but see also Kierepka and Latch 2014) we additionally applied a distance-based redundancy analysis (dbRDA, Legendre and Anderson 1999) using the ‘*capscale*’ and ‘*anova.cca*’ functions in VEGAN, after transforming the geographic distance matrix to continuous rectangular vectors via principal coordinates analyses (using the ‘*pcnm*’ function). Geographic distances among populations were calculated using the GEOGRAPHIC DISTANCE MATRIX GENERATOR (Ersts 2011). Additionally, a test of Isolation-by-Resistance was conducted, taking into account the higher elevation (Table 1) and topographic complexity in the Western PRB (Barker et al. 2012). Topographic complexity was approximated using the surface ratio index (Jenness 2004) as implemented in DEM Surface Tools (Jenness 2013) based on Digital Elevation Models of Puerto Rico and the Virgin Islands (Taylor et al. 2008; Grothe et al. 2012) provided by the National Geophysical Data Center. The calculated values of the surface ratio index for each cell (at a resolution of 10 arc-second) were subsequently used in CIRCUITSCAPE 4.0 (Shah and McRae 2008; McRae et al. 2013) to rescale geographic distances among populations under a circuit theory approach (McRae and Beier 2007; McRae et al. 2008).

To visualize the major axes of population genetic variation within *A. lares*, *A. pronauta* and *A. sanctaerucis*, Principal Component Analyses (PCAs) were performed for each of the three species (i.e., individuals were *a priori* assigned to each of the three genetically distinct species, but not to separate populations) using the R package ADEGENET (Jombart 2008; Jombart and Ahmed 2011) (‘*glPCA*’ function). Separate PCAs were run on (i) the full set of loci, (ii) rare variants (Minor Allele

Frequency,  $MAF \leq 0.05$ ), excluding those occurring only in a single individual and (iii) intermediate and common variants ( $MAF > 0.05$ ). Individuals with more than 50% missing data were removed from PCA analyses. We subsequently performed Discriminant Analysis of Principal Components (DAPC, Jombart et al. 2010) to quantify among-cluster differentiation in each of the PCA analyses based on discriminant functions, using the geographic localities as prior groups. The number of retained PC axes for the DAPC analyses was defined based on  $\alpha$ -score optimization. Cluster assignment probabilities, averaged across all individuals of each population, and all populations of each species, were used as an indication of whether the geographic clusters were clearly separated in each analysis. Additionally, mean distances among the centroids of the clusters were calculated for each species and set of loci (with all PCAs plotted on the same axes to allow comparisons).

#### *Intraspecific divergence time estimation*

The three intraspecific SNP datasets of *A. lares*, *A. pronauta* and *A. sanctaecrucis* were additionally used to estimate the timing of population divergence across the Western and the Eastern PRB. We used a composite-likelihood simulation-based approach (Excoffier et al. 2013), implemented in FASTSIMCOAL2 (Excoffier and Foll 2011) to estimate demographic parameters from the site frequency spectrum (SFS) under an Isolation-with-Migration model with three diverging populations. A folded joint SFS (i.e., for the minor allele, in the absence of information for the derived state) was calculated for each species based on a single SNP per locus to avoid the effects of linkage disequilibrium. In order to remove all missing data for the calculation of the joint SFS, each population was subsampled using a custom script (available on GitHub/KnowlesLab) and only loci found in at least five individuals per population (or three in the case of the YUN population) were retained to minimize errors with allele frequency estimates. Divergence times were estimated among three populations of each species (YAU-GUI-LAR; MAR-PAT-YUN; STT-TOR-VGP) rather than including all four populations of *A. sanctaecrucis* or *A. pronauta* in order to maximize the number of loci and make the three datasets comparable; the selection of populations was guided by the population tree analyses conducted in SNAPP, so that the most divergent populations with the highest

number of sampled individuals were included. Since performance of such models is improved by reducing the number of parameters estimated from the data (Excoffier et al. 2013), one effective population size ( $N_1$ ) of each dataset was calculated directly from the data (using nucleotide diversity as estimated by the POPULATIONS program in STACKS and a mutation rate of  $3.5 \times 10^{-9}$  per site per generation, Keightley et al. 2009), whereas the other parameters were allowed to vary. Note that, even though the absolute values of the estimated demographic parameters might not be accurate due to the use of a mutation rate from a distant taxon (which is the only genome-wide mutation rate available for insects), these estimates are certainly valid for relative comparison of demographic parameters among the three focal taxa, as there is no reason to expect differences in genome-wide mutation rates among closely related *Amphiacusta* species. The estimated parameters based on the joint SFS include: the effective population size of two populations ( $N_2$  and  $N_3$ ), the ancestral population size  $N_A$ , the divergence times  $T_1$  and  $T_2$ , and the migration rates among populations,  $m_1$ - $m_3$ . 40 independent runs per species were performed, each run with 100,000 simulations per likelihood estimation and 10–40 Expectation-Conditional Maximization (ECM) cycles, based on a stopping criterion of 0.001 relative difference between iterations. The global maximum likelihood solution across runs is presented. Parameter confidence intervals were calculated from 100 parametric bootstrap replicates, simulating SFS with the same number of SNPs from the maximum composite likelihood estimates and re-estimating parameters each time (Excoffier et al. 2013).

## RESULTS

### *Illumina data processing*

Of the approximately 280 million reads generated across all 121 individuals sequenced for this study, 225 million reads were retained after quality filtering, and the number of reads per individual ranged between 322,629 and 4,718,540 (Fig. S1). All individuals were included in analyses that were less sensitive to missing data (e.g., SVDQUARTETS species tree analysis), while individuals with a high percentage of missing data were removed from PCA, FASTSIMCOAL and SNAPP analyses (see

respective sections, following the criteria detailed in the methods above). After processing the data in PYRAD, the average number of loci per population ranged between 42,492 and 72,579 with an average depth per locus of 14.2–30.3 (Table S1). See individual sections below and Table S2 for the number of loci used for each of the downstream analyses.

### *Species tree and interspecific divergence*

A total of 171,528 putatively unlinked SNPs (i.e., a single SNP from each of 171,528 loci) were used in the SVDQUARTETS species tree analysis, whereas 35,585 SNPs were analyzed in the concatenated RAXML analysis, and 1,155–1,896 putatively unlinked SNPs in the SNAPP species tree analysis, due to different requirements and sensitivity of each method to missing data (see Methods section for details). All analyses produced highly congruent topologies both at the species and population level (Figs. 3a; S2; S3; S4) with high clade support (e.g., all clades supported by bootstrap  $\geq 94$  in SVDQUARTETS analyses and posterior probabilities  $\geq 0.9$  in SNAPP analyses) and each of the seven morphologically recognized species was recovered as monophyletic (Fig. S4). Most sampled populations were also recovered as monophyletic (Fig. S4), with the exception of the neighboring populations on Vieques (VIM-VIB), Culebra (CUV-CUF) and Mona (CDG1-JCS-SRS and CDG2-MHD-PCC-SBC3), which were pooled for subsequent analyses. The selected position of the root for the SVDQUARTETS and RAXML analyses (on the *A. tijiconhia*, *A. mona* and *A. sp.* clade; see Oneal et al. 2010) was also confirmed by the SNAPP analysis (Fig. S3) where there was no *a priori* selection of outgroup *A. sanctaecrucis* (Eastern PRB) and *A. viequesense* (in both Eastern and Western PRB) were recovered as the most recently diverged pair of sister taxa, while within the outgroup clade, the two species from Mona (*A. mona* and *A. sp.*) were not recovered as sister species (Figs. S2, S3).

Although the SNAPP species tree analysis possibly suffered from biases due to its reliance on a small subset of conserved loci (see Huang and Knowles 2016), both SNAPP (Fig. S3), and RAXML (Fig. S4) analyses, based on different subsets of loci and individuals, showed deeper interspecific divergences across the Western PRB and Mona than in the Eastern PRB. Specifically, total phylogenetic branch

length (PD, Faith 1992) was much lower within the Eastern PRB (Tables 1; S3) than in Mona, even though both areas have two *Amphiacusta* species each (*A. viequesense* and *A. sanctaecrucis* on the Eastern PRB and *A. mona* and *A. sp.* on Mona).

#### *Genetic diversity and population differentiation*

Nucleotide diversity ( $\pi$ ) and observed heterozygosity was generally similar across all populations of the three species from the Western and the Eastern PRB (Table S4), with the exception of the St. Thomas, Tortola and Virgin Gorda (Eastern PRB) populations of *A. sanctaecrucis* that had marginally lower observed heterozygosity (Table S4). Population differentiation as measured by  $F_{ST}$  was generally lower across the Eastern PRB (within *A. sanctaecrucis*) than across the Western PRB (within *A. lares* or *A. pronauta*), when controlling for geographic distance separating populations (Fig. 4), or when controlling for resistance due to topographic complexity (Fig. S5), with the exception of a single data point of *A. lares*. Pairwise  $F_{ST}$ -values were highly correlated with geographic distance in *A. sanctaecrucis* from the Eastern PRB (Mantel test  $r=0.92$ ,  $p=0.008$ ; dbRDA  $p=0.025$ ), but not in *A. pronauta* from the Western PRB (Mantel test  $r=0.62$ ,  $p=0.1$ ; dbRDA  $p=0.37$ ). The results were very similar when accounting for topographic complexity (Mantel tests  $r=0.86$ ,  $p=0.008$  and  $r=0.33$ ,  $p=0.3$ ; dbRDA  $p=0.017$  and  $p=0.38$  respectively in the two regions), but the correlations were not significant when controlling for the effect of geographic distance (partial Mantel test  $r=0.03$ ,  $p=0.5$ ), as resistance due to topographic complexity was highly correlated with geographic distance ( $r=0.94$ ,  $p=0.008$  and  $r=0.98$ ,  $p=0.04$  respectively).

PCA analyses of intermediate and common variants showed a strong geographic structuring of genetic variation in all three species, with a clear separation of all populations along the PC1 and PC2 axes (Fig. 5, right column), as confirmed by DAPC analyses (cluster assignment probabilities of 1, Table S5). However, when rare alleles were analyzed, the geographic structuring of genetic variation differed for populations of *A. sanctaecrucis* from the Eastern PRB (Fig. 5c, left panel) compared to *A. pronauta* and *A. lares* population from the Western PRB (Figs. 5a and 5b, left column). Specifically,



all four populations of *A. sanctaecrucis* from the Eastern PRB were highly overlapping (apart from two outlier individuals) based on rare variants (Fig. 5c, left panel) with an average cluster assignment probability of 0.68 and short distances among the centroids of each cluster (Table S5). This pattern did not change radically when two individuals that appeared as outliers, due to high intra-population variance in the St. Thomas and Tortola populations, were removed from the analysis (Fig. S7, although note that the St. Thomas population still showed high intra-population variance and did not strictly overlap with the others). In contrast, the Western PRB populations of *A. lares* clearly showed a higher degree of separation (Fig. 5a; cluster assignment probabilities of 1) and longer distances among clusters (Table S5). In the Western PRB *A. pronauta*, two of the four populations (MAR and PAT) were partially overlapping (Fig. 5b, left panel; cluster assignment probability of 0.85), but mean distance among the centroids of the clusters was even longer than in *A. lares* (Table S5).

#### *Intraspecific divergence times and migration rates*

Estimated intraspecific divergence times in the Eastern PRB species *A. sanctaecrucis* dated back to 83 and 110ka (95% confidence intervals 65–113ka and 106–133ka; see Table S6), coinciding with a period of intense changes in island area and connectivity in the Eastern PRB (Fig. 2, see also Papadopoulou and Knowles 2015a). In the Western PRB species *A. lares* the shallowest intraspecific split (T1) was of a similar age (85ka; 95% CI: 77–97ka), but the deepest split (T2) was more than two-fold older (220ka; 95% CI: 207–309). In the Western PRB species *A. pronauta*, both splits (T1=123ka and T2=156) predated the divergences of *A. sanctaecrucis*, although with partially overlapping confidence intervals in the case of T1 (T1=101–174ka and T2=147–200ka). Estimated migration rates were generally higher (up to ten-fold) among populations of the Eastern PRB species *A. sanctaecrucis* in comparison to the two Western PRB species, *A. lares* and *A. pronauta* (Tables 3, S6).

## DISCUSSION

The observed disparity in species and phylogenetic diversity between the two areas of the Puerto Rico Bank that have been differentially affected by the Quaternary sea-level oscillations is coupled by consistent differences in intraspecific genetic differentiation between closely related *Amphiacusta* species that have diversified within each of the two areas. Genomic divergence across the Eastern PRB, which is currently fragmented into separate islands but has experienced intense shifts in area and connectivity during the Quaternary (Fig. 2), appears to be ephemeral over longer evolutionary time (Rosenblum et al. 2012). This contrasts to equivalent comparisons across the Western PRB and the adjacent island of Mona, which have been relatively stable in both area and connectivity throughout the Quaternary. Below we evaluate alternative hypotheses to explain the lower levels of divergence across the Eastern PRB and discuss what insights are provided by this system into the role of sea-level oscillations in shaping island biodiversity. In particular, our study evaluates a specific set of hypotheses about shifting island connectivity and island area, which were selected specifically for the insights they provide within a comparative framework, as opposed to considering a broad array of factors potentially contributing to the “true” history (see Knowles 2009).

### *Linking population divergence and species diversity across the Puerto Rico Bank*

By conducting side-by-side comparisons among population pairs of *A. sanctaecrucis*, *A. lares* and *A. pronauta* across equivalent geographic distances (Fig. 4), we demonstrate that genetic differentiation across the currently fragmented Eastern PRB is lower than within the Western PRB, which has been relatively unaffected by the sea-level changes (i.e., it has not been fragmented during the Quaternary, and has only suffered minor shifts in island area, Fig. 2b). In most cases, population divergences within the Western PRB predate equivalent population splits among separate present-day islands of the Eastern PRB (Table 3, Fig. 3b), with inferred migration rates being generally lower in the Western PRB (Table 3). Shallower intraspecific divergences and higher levels of gene flow across the Eastern PRB, coupled by a similar pattern at the interspecific level (i.e., relatively shallow

divergence between the two *Amphiacusta* species distributed on the Eastern PRB, as opposed to deeper interspecific divergences among the Western PRB and Mona lineages; Figs. S3, S4), point towards a microevolutionary explanation for the low species and phylogenetic diversity (Tables 1, S2) in the Eastern PRB (i.e., reduced rate of population splitting or reduced population persistence through time, assuming that speciation duration should be fairly equivalent among closely related *Amphiacusta* lineages with similar life-history and ecological traits; see Fig. 1a). Given that the entire PRB emerged in the late Eocene (Meyerhoff 1933; Heatwole et al. 1981), and the present-day islands have been continuously emerged since then, there is no *a priori* reason to expect such disparity in the timeframe of diversification and species ages between the Eastern and Western PRB lineages, suggesting that the consistent differences at micro- and macroevolutionary levels (see also Kisel and Barraclough 2010; Rosenblum et al. 2012) could be related to the differential exposure of the two areas to sea-level oscillations.

An alternative factor that could potentially explain differences in population divergence and species diversity between the two areas of the PRB is the higher elevation and subsequently higher topographic complexity of the Western PRB (Table 1; Fig. S6) (see Barker et al. 2012). Island elevation is considered an important driver of population isolation and *in situ* speciation (Irl et al. 2015; Steinbauer et al. 2016), and topographic complexity has been shown to shape genetic structure in other montane Orthoptera species (e.g., Knowles 2000; Noguerales et al. 2016). However, in this system topographic complexity did not appear to provide a better explanation for the observed levels of genetic differentiation across the Western PRB than geographic distance *per se* (i.e., a model of isolation-by-resistance based on topographic complexity did not fit the data better than simply isolation-by-distance; Fig. S5). The lack of significance in these correlations could be affected by the limited population sampling across the Western PRB, and therefore we cannot exclude a potential effect of topography on genetic structure. However, the lack of statistical significance is also consistent with the distribution of these crickets in both montane forest and lower altitudes under current climatic conditions (Figs. 3a; S6) suggesting elevation does not currently appear as a major barrier to gene flow for *Amphiacusta*, although past population connectedness related to shifts in

climatic conditions through time due to glacial cycles may have been mediated by elevation (e.g., Massatti and Knowles 2014; Lanier et al. 2015; Knowles and Massatti 2016). Climatic fluctuations across a topographically complex landscape may repeatedly divide and merge populations (McCormack et al. 2009; Brown et al. 2013; Gillespie and Roderick 2014; Steinbauer et al. 2016) and in that sense elevation might have indeed been an important factor involved in driving population divergence and speciation across the Western PRB (see also Barker et al. 2011; 2015). Moreover, older geological events, such as the mountain uplift in the Western PRB that started approximately 4Ma (Brocard et al. 2016) could have also contributed to the diversification of *Amphiacusta*.

Even if elevation is driving diversification in the Western PRB, it still leaves unexplained the shallow divergences and low species diversity across the repeatedly fragmented Eastern PRB, especially when compared to the small and low-elevation island of Mona (Table S7). Note that the elevation of the Eastern PRB, with several peaks of 400–500m (Table S7), is still sufficient to generate habitat complexity through rain shadow effects, while the highest peak of Mona is only 92m, much lower than any of the major islands of the Eastern PRB. Furthermore, Mona has emerged more recently (Mio-Pliocene, Frank et al. 1998) than the Puerto Rico Bank itself (late Eocene, Heatwole et al. 1981). This comparison suggests that island age or elevation do not explain sufficiently the observed differences in diversity across this system, and enhances the idea that diversity patterns might actually reflect differences in the exposure of each area to the Quaternary sea-level oscillations.

#### *Area vs. connectivity effects of sea-level oscillations*

Given the large area of the Eastern PRB during most of the Quaternary (Fig. 2), as well as the multiple opportunities for population isolation on separate islands during high sea-level periods, there are no obvious reasons to think that the frequency of population splitting (Fig. 1a) has been particularly low across the Eastern PRB. In that sense, the low species diversity and shallow divergences across the Eastern PRB are more likely attributable to processes related to the long-term maintenance of divergence, rather than the drivers of divergence *per se* (Futuyma 1987; Dynesius and

Jansson 2014). Although not a formal test per se of the area hypothesis, we note that there are two divergent *Amphiacusta* lineages that have persisted *in situ* over long evolutionary time on Mona Island (Fig. 3; Figs. S3; S4), which is very similar in size to the majority of the present-day Virgin Islands (Table S7). This circumstantial observation is consistent with the hypothesis that the size of the Eastern PRB islands during high sea-level periods does not appear to be a limiting factor for population persistence.

It is inherently difficult to test the role of extinction in structuring patterns of genetic variation (i.e., to distinguish between incipient divergence frequently wiped out by increased population extinction due to small island size during high sea-level periods, Fig. 1b, versus by increased gene flow due to island merging during low sea-levels, Fig. 1c), and therefore we cannot rule out its potential contribution. However, the current genetic diversity patterns do not show any particular signature of frequent extinction in the Eastern PRB, as both Western and Eastern populations have very similar genetic diversity (Table S3), in contrast to the significantly lower genetic diversity of *A. sanctaecrucis* on the island of Anegada, a flat and low atoll that may have been subject to frequent extinctions (Papadopoulou and Knowles 2015a). On the contrary, there is a clear signature of higher gene flow among the Eastern PRB populations, as suggested by the lower  $F_{ST}$  values and the strong fit to a model of Isolation-by-Distance in *A. sanctaecrucis* (Fig. 4), as well as by the lack of spatial separation of rare genetic variants in this species (Fig. 5c, left panel), which is also reflected in higher estimated migration rates based on the Site Frequency Spectrum (Table 4). Rare genetic variants are expected to be spatially restricted if dispersal and gene flow are limited (Slatkin 1985; Barton and Slatkin 1986), as seen in most Western PRB populations of *A. lares* and *A. pronauta* (Fig. 5a and 5b, left column). Given that the spatial distribution of alleles reflects demographic processes since mutation (Slatkin 1985; Gompert et al. 2014), the lack of spatial separation of the rare (presumably younger) alleles in Eastern PRB *A. sanctaecrucis* (Fig. 5c, left panel) points towards recent gene flow.

The regional differences in recent migration appears rather surprising given the current sea-barriers separating the populations of *A. sanctaecrucis*, in contrast to the Western PRB populations of

*A. lares* and *A. pronauta* that have been continuously connected by land, especially considering that all three species are ecologically similar and presumably have uniform inherent dispersal capabilities (i.e., all are flightless ground crickets with very limited overwater dispersal propensity, pretty uniform morphologically and associated with wet tropical forest, Oneal et al. 2010). Interestingly, this observation also appears compatible with the phylogeographic structure of the frog *Eleutherodactylus antillensis*, which shows shallow divergence and high gene flow across the Eastern PRB islands and the Eastern side of the Western PRB (Barker et al. 2012).

What could explain the higher gene flow across the repeatedly fragmented Eastern PRB in comparison to the continuously emerged and undisrupted Western PRB? One potential hypothesis could relate to the differences in temporal and spatial habitat stability between the two areas. It has been shown both theoretically and empirically that dispersal is favored in spatially and temporally variable habitats (McPeck and Holt 1992; Denno et al. 1996; Dynesius and Jansson 2000; Ribera and Vogler 2000) or under increased rates of patch destruction (Friedenberg 2003), as a mean of tracking patches of favorable habitat and escaping from deteriorating local conditions (for a review see Ronce 2007). In other words, the highly dynamic landscape of the Eastern PRB, subjected to regular local extinctions and therefore variable population densities through space, may select for overall higher dispersal rates through time in comparison to the comparatively more static landscape of the Western PRB, where ecological space is continuously saturated. Even though this scenario is inherently difficult to test and we cannot rule out other alternative hypotheses, it is consistent with the observation that genomic divergence appears ephemeral across the dynamic sea-barriers formed repeatedly along the Eastern PRB.

*Cycles of island fission-fusion: “species pump” or “species vacuum”?*

Incipient genomic divergence of *Amphiacusta* lineages across the dynamic sea-barriers of the Eastern PRB is ephemeral, which contradicts the notion that repeated cycles of island fission-fusion act as “species pump” (Heaney 1985; Brown and Diesmos 2009; Ali and Aitchison 2014). In fact, in

the case of the PRB, it appears as if the island connectivity cycles impede, rather than promote population divergence and speciation. Is this a more general property of island systems exposed to cycles of fission-fusion or is it specific to this system? An important feature of the Eastern PRB is that islands are separated by very shallow waters (19–25 m), hence periods of island connections during the Quaternary have been much longer than periods of island isolation (Papadopoulou and Knowles 2015a). In that sense, conclusions from this system are not directly applicable to other archipelagos where islands have been separated for longer periods during the Quaternary (e.g., the Cycladic plateau in the Eastern Mediterranean; see Papadopoulou and Knowles 2015b), as the periodicity of isolation and connection in relation to the speciation duration of focal organisms may dictate different evolutionary outcomes (see also Gillespie and Roderick 2014; Knowles and Massatti 2016). Even though evolutionary responses to cycles of island fission-fusion might be mediated by such archipelago-specific characteristics, as well as by taxon-specific traits (Papadopoulou and Knowles 2015b; 2016), similar conclusions regarding the role of island connectivity cycles as impediment to speciation, have been recently reached from macroecological analyses of island floras, where past island connectivity appears negatively correlated with species diversity and endemism (Weigelt et al. 2016). Although such correlations do not provide a direct link with the underlying processes, they contribute to a growing body of evidence suggesting that at least for certain organisms and island systems the cycles of island fission-fusion may actually act as a “species vacuum” (as opposed to a “species pump”), which periodically wipes out incipient divergence and hinders *in situ* diversification.

Such opposing responses to the cycles of island fission-fusion (“species pump” vs. “species vacuum”) might be predicted under different sets of conditions related to the potential for dispersal during low sea levels and for speciation during high sea-level periods (Ali and Aitchison 2014). Dispersal ability is indeed a key trait mediating taxon-specific responses to island connectivity cycles, but other ecological and life-history traits might also be involved (Sukumaran et al. 2016), which are additionally moderated by archipelago-specific characteristics (Papadopoulou and Knowles 2015b). Further empirical studies combining micro- and macroevolutionary perspectives with modeling approaches for generating species-specific predictions under biologically informed hypotheses (see

He et al. 2013; Massatti and Knowles 2016; Papadopoulou and Knowles 2016) may allow to explore further the parameter space and identify specific combinations of archipelago characteristics (e.g., tempo of island connections, island area, topographic relief, bathymetry) and taxon-specific traits (e.g., dispersal ability, habitat/area requirements, speciation duration or traits related to ecological adaptation) where the “species pump” mechanism indeed operates.

### *Conclusions*

Combining a micro- and a macroevolutionary perspective is necessary to gain an understanding on how the Quaternary cycles of island fission-fusion and shifting island area have shaped island biodiversity. By conducting comparisons of population and species divergence between closely related and ecologically similar taxa diversifying *in situ* in two adjacent areas of the Puerto Rico Bank that have been differentially exposed to sea-level oscillations, we show that under certain conditions the repeated changes in island area and connectivity may cause genomic divergence to be ephemeral, thus suggesting an opposing effect to the once hypothesized “species pump” action of oscillating sea levels. Further empirical and simulation studies are needed in order to identify certain archipelago characteristics and taxon-specific traits that may lead to different evolutionary outcomes, as well as to distinguish the presumably contrasting effects of shifting island area *vs.* island connectivity as drivers of diversification. Our approach highlights how consideration of the underlying microevolutionary processes is critical to current efforts for developing Quaternary-sensitive models of island biogeography and more broadly to studies of island diversification.

### LITERATURE CITED

- Ali, J. R. and J. C. Aitchison. 2014. Exploring the combined role of eustasy and oceanic island thermal subsidence in shaping biodiversity on the Galápagos. *J. Biogeogr.* 41:1227-1241.
- Andrews, S. 2010. FastQC A Quality Control tool for High Throughput Sequence Data. Available from: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.



- Barker, B. S., K. Andonian, S. M. Swope, D. G. Luster, and K. M. Dlugosch. 2017. Population genomic analyses reveal a history of range expansion and trait evolution across the native and invaded range of yellow starthistle (*Centaurea solstitialis*). *Mol. Ecol.* 26:1131-1147.
- Barker, B. S., J. A. Rodriguez-Robles, V. S. Aran, A. Montoya, R. B. Waide, and J. A. Cook. 2012. Sea level, topography and island diversity: phylogeography of the Puerto Rican Red-eyed Coqui, *Eleutherodactylus antillensis*. *Mol. Ecol.* 21:6033-6052.
- Barker, B. S., J. A. Rodríguez-Robles, and J. A. Cook. 2015. Climate as a driver of tropical insular diversity: comparative phylogeography of two ecologically distinctive frogs in Puerto Rico. *Ecography* 38:769-781.
- Barker, B. S., R. B. Waide, and J. A. Cook. 2011. Deep intra-island divergence of a montane forest endemic: phylogeography of the Puerto Rican frog *Eleutherodactylus portoricensis* (Anura: Eleutherodactylidae). *J. Biogeogr.* 38:2311-2325.
- Barton, N. H. and M. Slatkin. 1986. A Quasi-equilibrium theory of the distribution of rare alleles in a subdivided population. *Heredity* 56:409-415.
- Bouckaert, R., J. Heled, D. Kühnert, T. Vaughan, C.-H. Wu, D. Xie, M. A. Suchard, A. Rambaut, and A. J. Drummond. 2014. BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Comp. Biol.* 10:e1003537.
- Bouckaert, R. R. 2010. DensiTree: making sense of sets of phylogenetic trees. *Bioinformatics* 26:1372-1373.
- Brocard, G. Y., J. K. Willenbring, T. E. Miller, and F. N. Scatena. 2016. Relict landscape resistance to dissection by upstream migrating knickpoints. *J. Geophys. Res. Earth Surf.* 121:1182-1203.
- Brown, R. and A. Diesmos. 2009. Philippines, biology. Pp. 723-732 in Gillespie R. G. and D. Clague, eds. *Encyclopedia of Islands*. University of California Press, Berkeley.
- Brown, R. M., C. D. Siler, C. H. Oliveros, J. A. Esselstyn, A. C. Diesmos, P. A. Hosner, C. W. Linkem, A. J. Barley, J. R. Oaks, M. B. Sanguila, L. J. Welton, D. C. Blackburn, R. G. Moyle, A. Townsend Peterson, and A. C. Alcala. 2013. Evolutionary Processes of Diversification in a Model Island Archipelago. *Annu. Rev. Ecol. Evol. Syst.* 44:411-435.
- Brown, W. C. and A. C. Alcala. 1970. The zoogeography of the herpetofauna of the Philippine Islands, a fringing archipelago. *Proc. Calif. Acad. Sci.* 38:105-130.
- Bryant, D., R. Bouckaert, J. Felsenstein, N. A. Rosenberg, and A. RoyChoudhury. 2012. Inferring Species Trees Directly from Biallelic Genetic Markers: Bypassing Gene Trees in a Full Coalescent Analysis. *Mol. Biol. Evol.* 29:1917-1932.
- Carstens, B. C. and L. L. Knowles. 2007. Shifting distributions and speciation: species divergence during rapid climate change. *Mol. Ecol.* 16:619-627.
- Catchen, J., P. A. Hohenlohe, S. Bassham, A. Amores, and W. A. Cresko. 2013. Stacks: an analysis tool set for population genomics. *Mol. Ecol.* 22:3124-3140.

- Catchen, J. M., A. Amores, P. Hohenlohe, W. Cresko, and J. H. Postlethwait. 2011. Stacks: building and genotyping loci de novo from short-read sequences. *G3* 1:171-182.
- Chifman, J. and L. Kubatko. 2014. Quartet Inference from SNP Data Under the Coalescent Model. *Bioinformatics* 30:3317-3324.
- Denno, R. F., G. K. Roderick, M. A. Peterson, A. F. Huberty, H. G. Doble, M. D. Eubanks, J. E. Losey, and G. A. Langelotto. 1996. Habitat persistence underlies intraspecific variation in the dispersal strategies of planthoppers. *Ecol. Monogr.* 66:389-408.
- Desutter-Grandcolas, L. and D. Otte. 1997. Revision of the West Indian genus *Amphiacusta* Saussure, 1874, with descriptions of twenty new species (Orthoptera: Grylloidea: Phalangopsidae). *Ann. Soc. Entomol. Fr.* 33:101-128.
- Diamond, J. M. and M. E. Gilpin. 1983. Biogeographic Umbilici and the Origin of the Philippine Avifauna. *Oikos* 41:307-321.
- Dynesius, M. and R. Jansson. 2000. Evolutionary consequences of changes in species' geographical distributions driven by Milankovitch climate oscillations. *Proc. Natl. Acad. Sci. USA* 97:9115-9120.
- Dynesius, M. and R. Jansson. 2014. Persistence of within- species lineages: A neglected control of speciation rates. *Evolution* 68:923-934.
- Eaton, D. A. R. 2014. PyRAD: assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics* 30:1844-1849.
- Ersts, P. J. 2011. Geographic Distance Matrix Generator v. 1.2.3. American Museum of Natural History, Center for Biodiversity and Conservation. Available from: [http://biodiversityinformatics.amnh.org/open\\_source/gdmg](http://biodiversityinformatics.amnh.org/open_source/gdmg).
- Esselstyn, J. A. and R. M. Brown. 2009. The role of repeated sea-level fluctuations in the generation of shrew (Soricidae: Crocidura) diversity in the Philippine Archipelago. *Mol. Phylogenet. Evol.* 53:171-181.
- Esselstyn, J. A., R. M. Timm, and R. M. Brown. 2009. Do geological or climatic processes drive speciation in dynamic archipelagos? The tempo and mode of diversification in Southeast Asian shrews. *Evolution* 63:2595-2610.
- Excoffier, L., I. Dupanloup, E. Huerta-Sánchez, V. C. Sousa, and M. Foll. 2013. Robust demographic inference from genomic and SNP data. *PLoS Genet.* 9:e1003905.
- Excoffier, L. and M. Foll. 2011. Fastsimcoal: a continuous-time coalescent simulator of genomic diversity under arbitrarily complex evolutionary scenarios. *Bioinformatics* 27:1332-1334.
- Faith, D. P. 1992. Conservation evaluation and phylogenetic diversity. *Biol. Conserv.* 61:1-10.
- Fernández-Palacios, J. M., K. F. Rijdsdijk, S. J. Norder, R. Otto, L. de Nascimento, S. Fernández-Lugo, E. Tjørve, and R. J. Whittaker. 2016. Towards a glacial-sensitive model of island biogeography. *Global Ecol. Biogeogr.* 25:817-830.

- Frank, E. F., C. Wicks, J. Mylroie, J. Troester, E. C. Alexander, and J. Carew. 1998. Geology of Isla de Mona, Puerto Rico. *J. Caves Karst Stud.* 60:69-72.
- Friedenberg, N. A. 2003. Experimental evolution of dispersal in spatiotemporally variable microcosms. *Ecol. Lett.* 6:953-959.
- Futuyma, D. J. 1987. On the role of species in anagenesis. *Am. Nat.* 130:465-473.
- Gillespie, R. G. and G. K. Roderick. 2014. Evolution: Geology and climate drive diversification. *Nature* 509:297-298.
- Gompert, Z., L. K. Lucas, C. A. Buerkle, M. L. Forister, J. A. Fordyce, and C. C. Nice. 2014. Admixture and the organization of genetic diversity in a butterfly species complex revealed through common and rare genetic variants. *Mol. Ecol.* 23:4555-4573.
- Grothe, P., L. Taylor, B. Eakins, K. Carignan, R. Caldwell, E. Lim, and D. Friday. 2012. Digital Elevation Models of the US Virgin Islands: Procedures, Data Sources and Analysis, NOAA Technical Memorandum NESDIS NGDC-55. U.S. Dept. of Commerce, Boulder, CO.
- Haponski, A. E., T. Lee, and D. Ó Foighil. 2017. Moorean and Tahitian *Partula* tree snail survival after a mass extinction: New genomic insights using museum specimens. *Mol. Phylogen. Evol.* 106:151-157.
- Harvey, M. G., C. D. Judy, G. F. Seeholzer, J. M. Maley, G. R. Graves, and R. T. Brumfield. 2015. Similarity thresholds used in DNA sequence assembly from short reads can reduce the comparability of population histories across species. *PeerJ* 3:e895.
- He, Q., D. L. Edwards, and L. L. Knowles. 2013. Integrative testing of how environments from the past to the present shape genetic structure across landscapes. *Evolution* 67:3386-3402.
- Heaney, L. 1985. Zoogeographic evidence for middle and late Pleistocene land bridges to the Philippine Islands. *Modern Quaternary Research in Southeast Asia* 9:127-144.
- Heaney, L. R. 2000. Dynamic disequilibrium: a long-term, large-scale perspective on the equilibrium model of island biogeography. *Global Ecol. Biogeogr.* 9:59-74.
- Heatwole, H., R. Levins, and M. Byers. 1981. Biogeography of the Puerto Rican Bank: Geography, geology, vegetation, and general ecology. *Atoll Res. Bull.* 251:1-62.
- Heatwole, H. and F. MacKenzie. 1967. Herpetogeography of Puerto Rico. IV. Paleogeography, Faunal Similarity and Endemism. *Evolution* 21:429-438.
- Hedges, S. B. 1999. Distribution patterns of amphibians in the West Indies. Pp. 211-254 in W. Duellman, ed. *Patterns of distribution of amphibians: a global perspective*. The John Hopkins University Press, Baltimore, MD.
- Hohenlohe, P. A., S. Bassham, P. D. Etter, N. Stiffler, E. A. Johnson, and W. A. Cresko. 2010. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genet.* 6:e1000862.

- Huang, H. and L. L. Knowles. 2016. Unforeseen Consequences of Excluding Missing Data from Next-Generation Sequences: Simulation Study of RAD Sequences. *Syst. Biol.* 65:357-365.
- Huang, J.-P. 2016. Parapatric genetic introgression and phenotypic assimilation: testing conditions for introgression between Hercules beetles (Dynastes, Dynastinae). *Mol. Ecol.* 25:5513-5526.
- Irl, S. D. H., D. E. V. Harter, M. J. Steinbauer, D. Gallego Puyol, J. M. Fernández-Palacios, A. Jentsch, and C. Beierkuhnlein. 2015. Climate vs. topography – spatial patterns of plant species diversity and endemism on a high-elevation island. *J. Ecol.* 103:1621-1633.
- Jenness, J. 2013. DEM Surface Tools. Jenness Enterprises. Available from: [http://www.jennessent.com/arcgis/surface\\_area.htm](http://www.jennessent.com/arcgis/surface_area.htm).
- Jenness, J. S. 2004. Calculating landscape surface area from digital elevation models. *Wildl. Soc. Bull.* 32:829-839.
- Jombart, T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24:1403-1405.
- Jombart, T. and I. Ahmed. 2011. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics* 27:3070-3071.
- Jombart, T., S. Devillard, and F. Balloux. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet.* 11:94.
- Jordan, S., C. Simon, D. Foote, and R. A. Englund. 2005. Phylogeographic patterns of Hawaiian Megalagrion damselflies (Odonata: Coenagrionidae) correlate with Pleistocene island boundaries. *Mol. Ecol.* 14:3457-3470.
- Keightley, P. D., U. Trivedi, M. Thomson, F. Oliver, S. Kumar, and M. L. Blaxter. 2009. Analysis of the genome sequences of three *Drosophila melanogaster* spontaneous mutation accumulation lines. *Genome Res.* 19:1195-1201.
- Kembel, S. W., P. D. Cowan, M. R. Helmus, W. K. Cornwell, H. Morlon, D. D. Ackerly, S. P. Blomberg, and C. O. Webb. 2010. Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* 26:1463-1464.
- Kierepka, E. M. and E. K. Latch. 2014. Performance of partial statistics in individual-based landscape genetics. *Mol. Ecol. Res.* 15:512-525.
- Kisel, Y. and T. G. Barraclough. 2010. Speciation has a spatial scale that depends on levels of gene flow. *Am. Nat.* 175:316-334.
- Knowles, L. L. 2000. Tests of Pleistocene speciation in montane grasshoppers (Genus *Melanoplus*) from the sky islands of Western North America. *Evolution* 54:1337-1348.
- Knowles, L. L. 2009. Statistical Phylogeography. *Ann. Rev. Ecol. Evol. Syst.* 40:593-612.
- Knowles, L. L. and R. Massatti. 2017. Distributional shifts– not geographic isolation – as a probable driver of montane species divergence. *Ecography* (in press).

- Lambeck, K. and J. Chappell. 2001. Sea level change through the last glacial cycle. *Science* 292:679-686.
- Lambeck, K., Y. Yokoyama, and T. Purcell. 2002. Into and out of the Last Glacial Maximum: sea-level change during Oxygen Isotope Stages 3 and 2. *Quat. Sci. Rev.* 21:343-360.
- Lanier, H. C., R. Massatti, Q. X. He, L. E. Olson, and L. L. Knowles. 2015. Colonization from divergent ancestors: glaciation signatures on contemporary patterns of genomic variation in Collared Pikas (*Ochotona collaris*). *Mol. Ecol.* 24:3688-3705.
- Legendre, P. and M. J. Anderson. 1999. Distance-based redundancy analysis: Testing multispecies responses in multifactorial ecological experiments. *Ecol. Monogr.* 69:1-24.
- Legendre, P. and M. J. Fortin. 2010. Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. *Mol. Ecol. Res.* 10:831-844.
- Losos, J. B. and D. Schluter. 2000. Analysis of an evolutionary species-area relationship. *Nature* 408:847-850.
- MacArthur, R. H. and E. O. Wilson. 1963. An Equilibrium Theory of Insular Zoogeography. *Evolution* 17:373-387.
- MacArthur, R. H. and E. O. Wilson. 1967. *The theory of island biogeography*. Princeton University Press.
- Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Res.* 27:209-220.
- Massatti, R. and L. L. Knowles. 2014. Microhabitat differences impact phylogeographic concordance of codistributed species: Genomic evidence in montane sedges (*Carex* L.) from the Rocky Mountains. *Evolution* 68:2833-2846.
- Massatti, R. and L. L. Knowles. 2016. Contrasting support for alternative models of genomic variation based on microhabitat preference: species-specific effects of climate change in alpine sedges. *Mol. Ecol.* 25:3974-3986.
- McCormack, J. E., H. Huang, L. L. Knowles, R. Gillespie, and D. Clague. 2009. Sky islands. Pp. 841-843 in Gillespie R. G. and D. Clague, eds. *Encyclopedia of Islands*. University of California Press, Berkeley.
- McPeck, M. A. and R. D. Holt. 1992. The evolution of dispersal in spatially and temporally varying environments. *Am. Nat.* 140:1010-1027.
- McRae, B., V. Shah, and T. Mohapatra. 2013. *Circuitscape 4 User Guide*. The Nature Conservancy. Available from: <http://www.circuitscape.org>.
- McRae, B. H. and P. Beier. 2007. Circuit theory predicts gene flow in plant and animal populations. *Proc. Natl. Acad. Sci. USA* 104:19885-19890.

- McRae, B. H., B. G. Dickson, T. H. Keitt, and V. B. Shah. 2008. Using circuit theory to model connectivity in ecology, evolution, and conservation. *Ecology* 89:2712-2724.
- Meyerhoff, H. A. 1933. *Geology of Puerto Rico*. Univ. Puerto Rico Monograph Ser. B. 1:1-306.
- Noguerales, V., P. J. Cordero, and J. Ortego. 2016. Hierarchical genetic structure shaped by topography in a narrow-endemic montane grasshopper. *BMC Evol. Biol.* 16:96.
- Oaks, J. R., J. Sukumaran, J. A. Esselstyn, C. W. Linkem, C. D. Siler, M. T. Holder, and R. M. Brown. 2013. Evidence for climate-driven diversification? A caution for interpreting ABC inferences of simultaneous historical events. *Evolution* 67:991-1010.
- Oksanen, J., F. G. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, B. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, and H. Wagner. 2013. *vegan: Community Ecology Package*. R package version 2.0-9.
- Oneal, E., D. Otte, and L. L. Knowles. 2010. Testing for biogeographic mechanisms promoting divergence in Caribbean crickets (genus *Amphiacusta*). *J. Biogeogr.* 37:530-540.
- Papadopoulou, A. and L. L. Knowles. 2015a. Genomic tests of the species-pump hypothesis: Recent island connectivity cycles drive population divergence but not speciation in Caribbean crickets across the Virgin Islands. *Evolution* 69:1501-1517.
- Papadopoulou, A. and L. L. Knowles. 2015b. Species-specific responses to island connectivity cycles: refined models for testing phylogeographic concordance across a Mediterranean Pleistocene Aggregate Island Complex. *Mol. Ecol.* 24:4252-4268.
- Papadopoulou, A. and L. L. Knowles. 2016. Toward a paradigm shift in comparative phylogeography driven by trait-based hypotheses. *Proc. Natl. Acad. Sci. USA* 113:8018-8024.
- Parks, D. H., M. Porter, S. Churcher, S. Wang, C. Blouin, J. Whalley, S. Brooks, and R. G. Beiko. 2009. GenGIS: A geospatial information system for genomic data. *Genome Res.* 19:1896-1904.
- Patiño, J., R. J. Whittaker, P. A. V. Borges, J. M. Fernández-Palacios, C. Ah-Peng, M. B. Araújo, S. P. Ávila, P. Cardoso, J. Cornuault, E. J. de Boer, L. de Nascimento, A. Gil, A. González-Castro, D. S. Gruner, R. Heleno, J. Hortal, J. C. Illera, C. N. Kaiser-Bunbury, T. J. Matthews, A. Papadopoulou, N. Pettorelli, J. P. Price, A. M. C. Santos, M. J. Steinbauer, K. A. Triantis, L. Valente, P. Vargas, P. Weigelt, and B. C. Emerson. 2017. A roadmap for island biology: 50 fundamental questions after 50 years of *The Theory of Island Biogeography*. *J. Biogeogr.* 44:963-983.
- Peterson, B. K., J. N. Weber, E. H. Kay, H. S. Fisher, and H. E. Hoekstra. 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS one* 7:e37135.

- Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. Ferreira, D. Bender, J. Maller, P. Sklar, P. de Bakker, M. Daly, and P. Sham. 2007. PLINK: a toolset for whole-genome association and population-based linkage analysis. *Am. J. Hum. Genet.* 81.
- R Core Team. 2016. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Rambaut, A., M. A. Suchard, D. Xie, and A. J. Drummond. 2014. Tracer version 1.6, Available from: <http://beast.bio.ed.ac.uk/Tracer>.
- Razkin, O., G. Sonet, K. Breugelmanns, M. J. Madeira, B. J. Gómez-Moliner, and T. Backeljau. 2016. Species limits, interspecific hybridization and phylogeny in the cryptic land snail complex *Pyramidula*: The power of RADseq data. *Mol. Phylogen. Evol.* 101:267-278.
- Reaz, R., M. S. Bayzid, and M. S. Rahman. 2014. Accurate Phylogenetic Tree Reconstruction from Quartets: A Heuristic Approach. *PLoS ONE* 9:e104008.
- Rheindt, F. E., M. K. Fujita, P. R. Wilton, and S. V. Edwards. 2014. Introgression and Phenotypic Assimilation in *Zimmerius* Flycatchers (Tyrannidae): Population Genetic and Phylogenetic Inferences from Genome-Wide SNPs. *Syst. Biol.* 63:134-152.
- Ribera, I. and A. P. Vogler. 2000. Habitat type as a determinant of species range sizes: the example of lotic-lentic differences in aquatic Coleoptera. *Biol. J. Linn. Soc.* 71:33-52.
- Rijsdijk, K. F., T. Hengl, S. J. Norder, R. Otto, B. C. Emerson, S. P. Ávila, H. López, E. E. van Loon, E. Tjørve, J. M. Fernández-Palacios, and L. Heaney. 2014. Quantifying surface-area changes of volcanic islands driven by Pleistocene sea-level cycles: biogeographical implications for the Macaronesian archipelagos. *J. Biogeogr.* 41:1242-1254.
- Ronce, O. 2007. How Does It Feel to Be like a Rolling Stone? Ten Questions about Dispersal Evolution. *Annu. Rev. Ecol. Evol. Syst.* 38:231-253.
- Rosenblum, E. B., B. A. J. Sarver, J. W. Brown, S. Des Roches, K. M. Hardwick, T. D. Hether, J. M. Eastman, M. W. Pennell, and L. J. Harmon. 2012. Goldilocks Meets Santa Rosalia: An Ephemeral Speciation Model Explains Patterns of Diversification Across Time Scales. *Evol. Biol.* 39:255-261.
- Scheffrahn, R. H., S. C. Jones, J. Krecek, J. A. Chase, J. R. Mangold, and N.-Y. Su. 2003. Taxonomy, Distribution, and Notes on the Termites (Isoptera: Kalotermitidae, Rhinotermitidae, Termitidae) of Puerto Rico and the U.S. Virgin Islands. *Ann. Entomol. Soc. Am.* 96:181-201.
- Shah, V. and B. McRae. 2008. Circuitscape: a tool for landscape ecology. *Proceedings of the 7th Python in Science Conference* 7: 62-66.
- Siler, C. D., J. R. Oaks, J. A. Esselstyn, A. C. Diesmos, and R. M. Brown. 2010. Phylogeny and biogeography of Philippine bent-toed geckos (Gekkonidae: *Cyrtodactylus*) contradict a prevailing model of Pleistocene diversification. *Mol. Phylogen. Evol.* 55:699-710.

- Slatkin, M. 1985. Gene flow in natural populations. *Annu. Rev. Ecol. Syst.* 16:393-430.
- Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312-1313.
- Steinbauer, M. J., R. Field, J.-A. Grytnes, P. Trigas, C. Ah-Peng, F. Attorre, H. J. B. Birks, P. A. V. Borges, P. Cardoso, C.-H. Chou, M. De Sanctis, M. M. de Sequeira, M. C. Duarte, R. B. Elias, J. M. Fernández-Palacios, R. Gabriel, R. E. Gereau, R. G. Gillespie, J. Greimler, D. E. V. Harter, T.-J. Huang, S. D. H. Irl, D. Jeanmonod, A. Jentsch, A. S. Jump, C. Kueffer, S. Nogué, R. Otto, J. Price, M. M. Romeiras, D. Strasberg, T. Stuessy, J.-C. Svenning, O. R. Vetaas, and C. Beierkuhnlein. 2016. Topography-driven isolation, speciation and a global increase of endemism with elevation. *Global Ecol. Biogeogr.* 25:1097-1107.
- Sukumaran, J., E. P. Economo, and L. L. Knowles. 2016. Machine Learning Biogeographic Processes from Biotic Patterns: A New Trait-Dependent Dispersal and Diversification Model with Model Choice By Simulation-Trained Discriminant Analysis. *Syst. Biol.* 65:525-545.
- Taylor, L. A., B. W. Eakins, K. S. Carignan, R. R. Warnken, T. Sazonova, and Schoolcraft. 2008. Digital Elevation Models of Puerto Rico: Procedures, Data Sources and Analysis, NOAA Technical Memorandum NESDIS NGDC-13, National Geophysical Data Center, Boulder, CO.
- Thomaz, A. T., L. R. Malabarba, and L. L. Knowles. 2017. Genomic signatures of paleodrainages in a freshwater fish along the southeastern coast of Brazil: genetic structure reflects past riverine properties. *Heredity* 119:287-294.
- Warren, B. H., D. Simberloff, R. E. Ricklefs, R. Aguilée, F. L. Condamine, D. Gravel, H. Morlon, N. Mouquet, J. Rosindell, J. Casquet, E. Conti, J. Cornuault, J. M. Fernández-Palacios, T. Hengl, S. J. Norder, K. F. Rijdsdijk, I. Sanmartín, D. Strasberg, K. A. Triantis, L. M. Valente, R. J. Whittaker, R. G. Gillespie, B. C. Emerson, and C. Thébaud. 2015. Islands as model systems in ecology and evolution: prospects fifty years after MacArthur-Wilson. *Ecol. Lett.* 18:200-217.
- Weigelt, P., M. J. Steinbauer, J. S. Cabral, and H. Kreft. 2016. Late Quaternary climate change shapes island biodiversity. *Nature* 532:99-102.
- Zheng, X., D. Levine, J. Shen, S. M. Gogarten, C. Laurie, and B. S. Weir. 2012. A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* 28:3326-3328.

TABLES

Table 1: Island area and elevation of the Western and Eastern Puerto Rico Bank and the adjacent island of Mona and corresponding species diversity of *Amphiacusta* on each of the three island



regions. *Area*, total island area emerged at current sea levels; *Area LGM*, hypothesized area at the Last Glacial Maximum as approximated based on Digital Elevation Models for Puerto Rico and the Virgin Islands (Taylor et al. 2008; Grothe et al. 2012) and considering a 120m drop of sea-level; *Elevation*, highest present-day peak; *Sp*, number of *Amphiacusta* species; *SpE*, number of *Amphiacusta* species endemic to this area; *PD*, phylogenetic diversity (Faith 1992) estimated based on the SNAPP species tree analysis.

Island Region	Area	Area LGM	Elevation	<i>Sp</i> ( <i>SpE</i> )	<i>PD</i>
Western PRB	9,104	~12,600	1,338	4(3)	0.173
Eastern PRB	(~500)	~7,700	523	2(1)	0.002
Mona	55.8	~120	92	2(2)	0.065

Table 2: Geographic information and total number of sequenced individuals per population for each of the seven *Amphiacusta* species.

Species	Island Region	Locality	Code	Coordinates	Ind s
<i>A. lares</i>	West PRB	Guilarte, Puerto Rico	GUI	18°6'34.86"N, 66°44'51.63"W	9
	West PRB	Lares, Puerto Rico	LAR	18°17'12.34"N, 66°51'2.27"W	10
	West PRB	Susua Reserve (Yauco), Puerto Rico	SRY	18°2' 25.20"N, 66°54' 32.67"W	9
<i>A. pronauta</i>	West PRB	Maricao, Puerto Rico	MAR	18°11'10.33"N, 67°4'31.58"W	8
	West PRB	Patillas, Puerto Rico	PAT	18°6'4.70"N, 66°2'27.19"W	7
	West PRB	Cerro de las Pinas, Puerto Rico	PIN	18°9'7.85"N, 66°5'7.59"W	4
	West PRB	El Yunque, Puerto Rico	YUN	18°17'24.65"N, 65°47'53.35"W	3

<i>A. sanctaecrucis</i>	East PRB	Villa Fulladosa, Culebra	CUV	18°18'5.47"N, 65°17'41.15"W	5
	East PRB	Flaminco Beach, Culebra	CUF	18°19'36.66"N, 65°18' 59.17"W	5
	East PRB	St. Thomas	STT	18°21'48.00"N, 64°58'27.00"W	8
	East PRB	Tortola	TOR	18°24'3.00"N, 64°39'40.00"W	10
	East PRB	Virgin Gorda Peak	VGP	18°28'40.00"N, 64°24'12.00"W	10
<i>A. viequesense</i>	East PRB	Monte Pirata, Vieques	VIM	18°5'34.91"N, 65°33'9.73"W	6
	East PRB	Vieques beach, Vieques	VIB	18°5'58.01"N, 65°34'19.08"W	4
	West PRB	San Juan Beach, Puerto Rico	SAJ	18°27' 14.19"N, 65°58' 13.16"W	4
<i>A. tijicohniae</i>	West PRB	Guanica, Puerto Rico	GUA	17°57'11.49"N, 66°53'1.97"W	3
<i>A. mona</i>	Mona	Cueva de Doña Geña	CDG 1	18°5'40.74"N, 67°54'12.84"W	1
	Mona	Jail/Cave at Sardinera Beach Camp	JCS	18°5'5.71"N, 67°56'25.60"W	3
	Mona	Spiny Road at Sardinera Beach Camp	SRS	18°5'25.85"N, 67°56'16.16"W	3
<i>A. sp 2</i>	Mona	Cueva de Doña Geña	CDG 2	18°5'40.74"N, 67°54'12.84"W	2
	Mona	Middle High Desert	MH D	18°3'37.23"N, 67°54'13.89"W	3
	Mona	Pajaros Cave Camp	PCC	18°3'51.27"N, 67°52'4.60"W	2
	Mona	Sardinera Beach Camp	SBC	18°5'19.37"N, 67°56'16.89"W	2

Table 3: Intraspecific divergence times (for two divergence events,  $T1$  &  $T2$ ) and symmetrical migration rates (among three population pairs,  $m1$ - $m3$ ) estimated under an Isolation-with-Migration model of three diverging populations in FASTSIMCOAL2, for each of three *Amphiacusta* species. Composite maximum likelihood estimates of divergence times are presented in number of generations (i.e., number of years ago, considering one generation per year) and as a function of the effective population size in parentheses. 95% confidence intervals on the estimated parameter values are given within brackets. See Table S5 for estimated effective population sizes. The number of loci used for the calculation of the joint site frequency spectrum for each species is indicated.

species	loci	$T1$	$T2$	$m1$	$m2$	$m3$
West PRB						
		85,143 (0.5N)	220,151 (1.3N)	$3 \times 10^{-7}$	$6 \times 10^{-7}$	$3.5 \times 10^{-6}$
<i>A. lares</i>	28,744	[77,412–97,514]	[207,163–309,179]	[ $1.5$ – $3.9 \times 10^{-7}$ ]	[ $5.2$ – $7.8 \times 10^{-7}$ ]	[ $3.1$ – $4.2 \times 10^{-6}$ ]
		123,819 (0.79N)	155,951 (1N)	$6 \times 10^{-7}$	$5 \times 10^{-7}$	$2.2 \times 10^{-6}$
<i>A. pronauta</i>	10,060	[101,235–173,963]	[147,503–200,146]	[ $4.2$ – $8.1 \times 10^{-7}$ ]	[ $3.8$ – $6.2 \times 10^{-7}$ ]	[ $1.9$ – $2.4 \times 10^{-6}$ ]
East PRB						
		83,577 (0.49N)	110,702 (0.65N)	$2.9 \times 10^{-6}$	$1.2 \times 10^{-6}$	$6.1 \times 10^{-6}$
<i>A. sanctaecrucis</i>	22,902	[63,599–113,127]	[106,954–133,409]	[ $2.5$ – $3.4 \times 10^{-6}$ ]	[ $0.9$ – $1.4 \times 10^{-6}$ ]	[ $5.2$ – $6.4 \times 10^{-6}$ ]

## FIGURE LEGENDS

Figure 1: Predictions for the effects of sea-level oscillations on the speciation process from a microevolutionary perspective, and specifically on (a) the three main processes that regulate the frequency of completed speciation (*sensu* (Dynesius and Jansson 2014): the frequency of population splitting (shown in green), the rate of local population extinction (shown in red) and the frequency of population merging (shown in yellow) due to increased gene flow (redrawn from (Dynesius and Jansson 2014). The predictions differ under two proposed hypotheses. If we consider the **shifts in island area** as a main driver of island diversity patterns (MacArthur and Wilson 1963, 1967;

Fernández-Palacios et al. 2016; Weigelt et al. 2016), then (b) the rate of population splitting (in green) is expected to be higher during low sea-level periods (glacial maxima), while local population extinction (in red) and merging due to gene flow (in yellow) is expected to be higher during high sea-levels (present or last interglacial period, LIG). On the contrary, if we consider the **cycles of fusion-fission** as a main driver of island diversity through a “species pump” mechanism (Heaney 1985; Ali and Aitchison 2014), then (c) the rate of population splitting (in green) is expected to be higher during high sea-levels (interglacial periods). Local population extinction (in red) will also be higher during interglacial periods, but merging due to gene flow (in yellow) will be higher during low sea-levels (last glacial maxima, LGM).

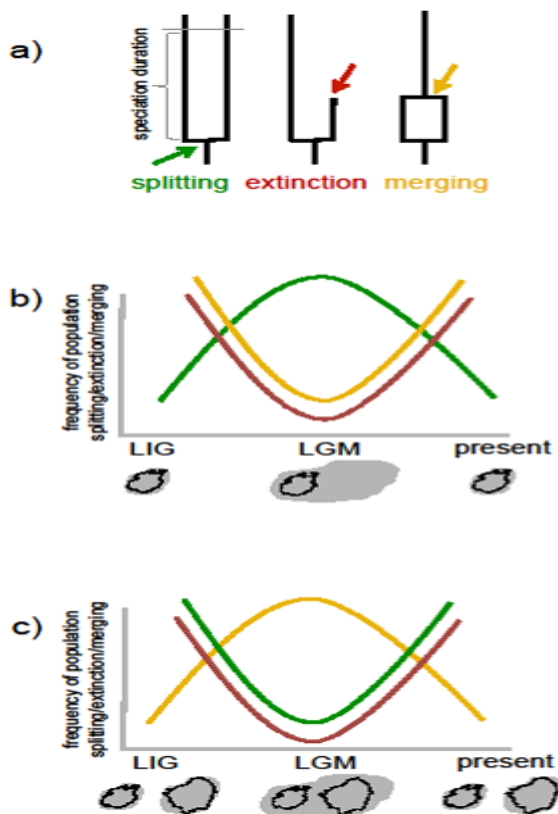


Figure 2: (a) Distribution of seven *Amphiacusta* species and sampling localities across the Puerto Rico Bank (PRB). Dark green areas indicate currently emerged islands, while gray shading represents areas that have been periodically emerged and submerged during the Quaternary (up to 120m depth below current sea-level, with color varying from dark to light grey at 10m intervals). See Table 1 for full locality names. (b) Changes in island area through time since the LGM in the Eastern and Western PRB, as approximated by measuring the area of each region at 5m depth intervals in ArcGIS based on Digital Elevation Models for Puerto Rico and the Virgin Islands (Taylor et al. 2008; Grothe et al. 2012) and curves of temporal variation in sea-level during the last glacial cycle from Lambeck & Chappell (2001) and Lambeck et al. (2002).

Figure 2

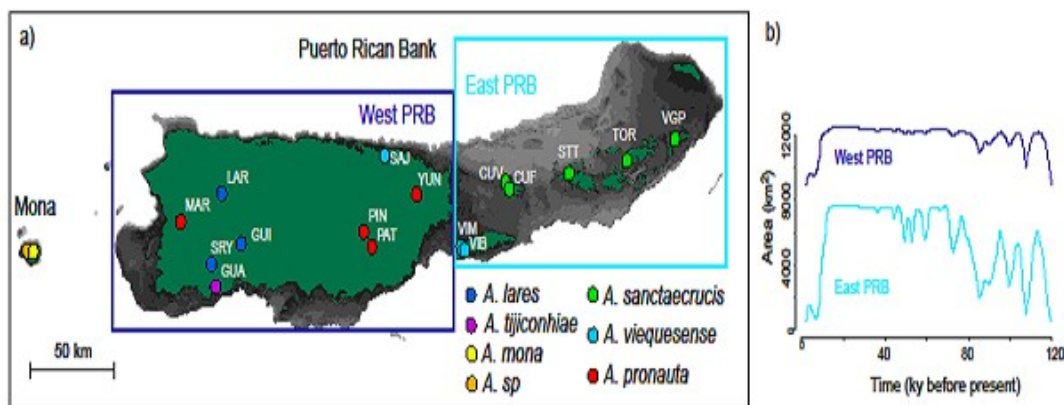


Figure 3: Inter- and intraspecific diversification of *Amphiacusta* species across the Puerto Rico Bank.

(a) Population-level tree estimated for seven *Amphiacusta* species based on 171,528 putatively unlinked SNPs analyzed with SVDQUARTETS. Colored circles represent monophyletic groups of individuals from the same population following the same color code as in Figure 2. (b) Relative intraspecific divergence times for the three *Amphiacusta* species distributed across the Eastern and Western PRB as estimated using FASTSIMCOAL2 under an Isolation-with-Migration model. See Table

4 for demographic parameter values and confidence intervals. The software GENGIS (Parks et al. 2009) was used for plotting the tree against the geographic map of the Puerto Rico Bank.



Figure 3

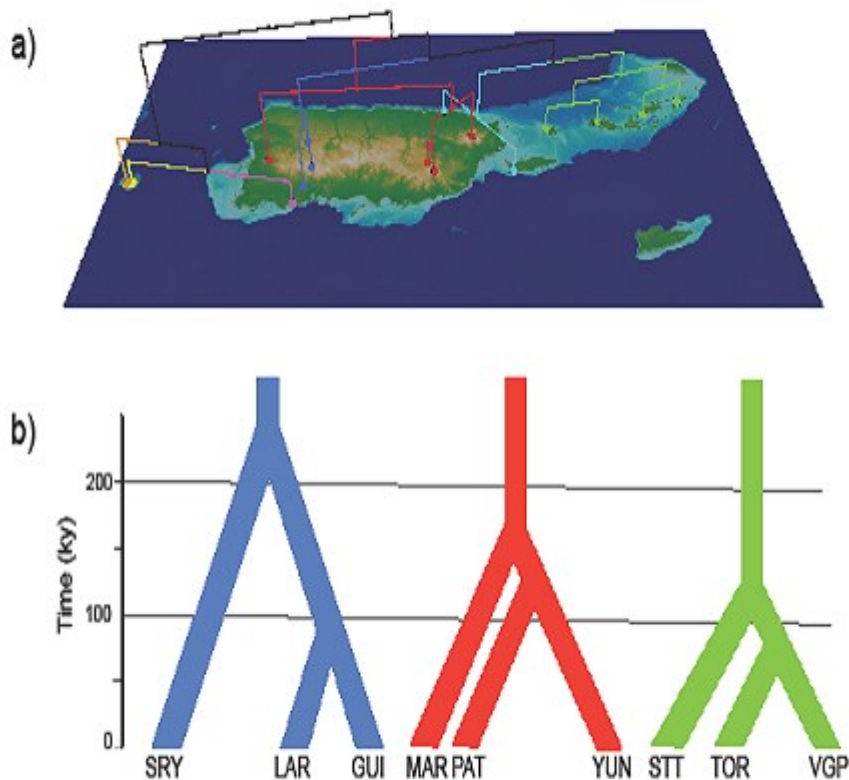


Figure 4: Genetic differentiation (as measured by  $F_{ST}$ ) in three *Amphiacusta* species presented as a function of geographic distance, for the Western PRB (shown in Blue: *A. lares* and Red: *A. pronauta*) and the Eastern PRB (shown in Light Green: *A. sanctaerucis*). There is a significant relationship between  $F_{ST}$  and geographic distance in the Eastern PRB *A. sanctaerucis* (Mantel test  $r=0.92$ ,  $p=0.04$ ) but not in the Western PRB species.

Figure 4

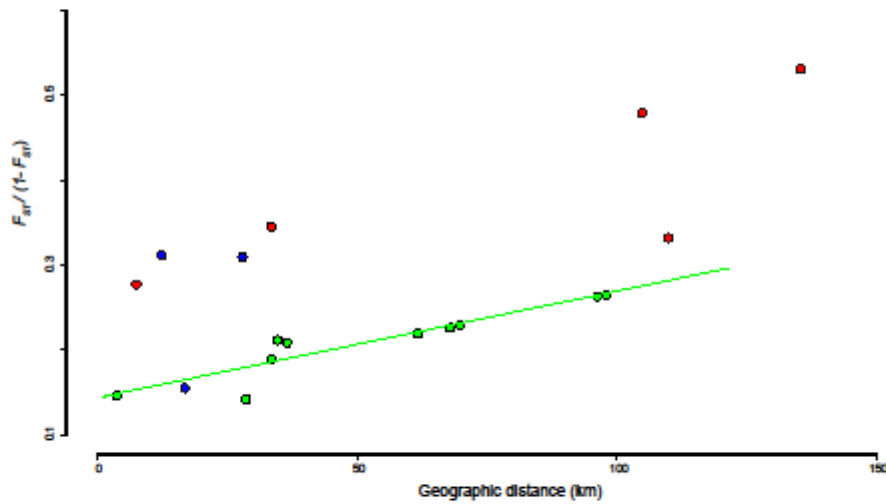
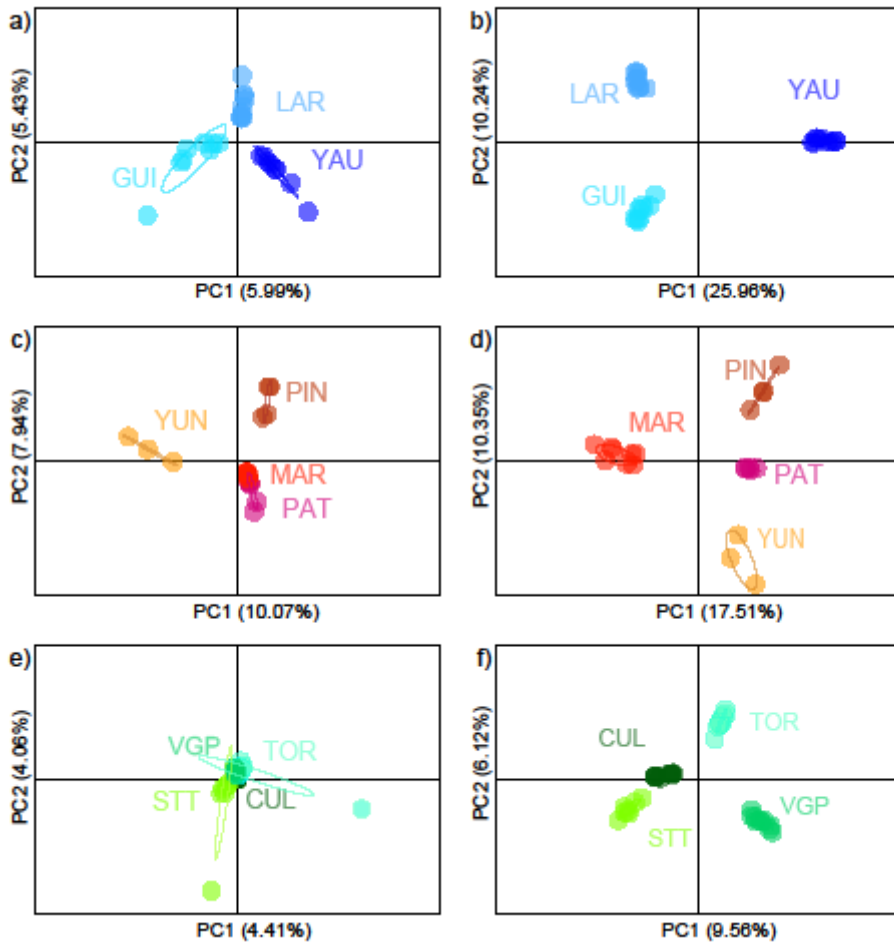


Figure 5: Comparisons of population structure based on the distribution of individuals along principal component 1 (PC1) and PC2 in three *Amphiacusta* species based on rare ( $MAF \leq 0.05$ ) (left column) versus more common polymorphisms ( $MAF > 0.05$ ) (right column) for the Western PRB taxa (a) *A. lares*, (b) *A. pronauta* and the Eastern PRB species (c) *A. sanctaerucis*. The amount of variation explained by PC1 and PC2 in each case is given in parentheses on the corresponding axes and 95% inertia ellipses are drawn for each population. Population codes follow those in Figure 2. Note that the two *A. sanctaerucis* populations from Culebra (CUF and CUV, ~3.5km apart) were pooled into a single population (CUL), as they showed very low genetic differentiation and clustered consistently together in preliminary analyses.



Author 1