doi:10.1111/evo.12667

Genomic tests of the species-pump hypothesis: Recent island connectivity cycles drive population divergence but not speciation in Caribbean crickets across the Virgin Islands

Anna Papadopoulou[1,2] and L. Lacey Knowles[1]

[1]*Department of Ecology and Evolutionary Biology, Museum of Zoology, University of Michigan, Ann Arbor, Michigan 48109*

[2]*E-mail: apapadop@umich.edu*

Received November 11, 2014
Accepted April 13, 2015

Harnessing the power of genomic scans, we test the debated "species pump" hypothesis that implicates repeated cycles of island connectivity and isolation as drivers of divergence. This question has gone understudied given the limited resolution of past molecular markers for studying such dynamic phenomena. With an average of 32,000 SNPs from the genome of 136 individuals from 10 populations of a Caribbean flightless ground cricket species (*Amphiacusta sanctaecrucis*) and a complementary set of statistical approaches, we infer a stepping-stone colonization model and high levels of genetic differentiation across the Virgin Islands, which have been periodically interconnected until 8 ka. Estimates of divergence times from models based on the site frequency spectrum coincide with a period of repeated connection and fragmentation of the islands at 75–130 ka. These results are consistent with a role of island connectivity cycles in promoting genomic divergence and indicate that the genetic distinctiveness of island populations has persisted despite subsequent and extended interisland connections identified from bathymetric data. We discuss these findings in the broader context of Caribbean biogeography, and more specifically why high levels of genomic divergence across the Virgin Islands associated with repeated connectivity cycles do not actually translate into species diversification.

KEY WORDS: **Biogeography, Pleistocene Aggregate Island Complex, Rad-seq, sea-level change.**

Sea-level oscillations between Pleistocene glacial maxima and interglacial periods have caused repeated cycles of island connections via land-bridges and subsequent isolation in archipelagos around the globe. Tests of the hypothesized "species pump" action of rising and falling sea levels under the "Pleistocene Aggregate Island Complex" (PAIC) (Heaney 1985; Heaney 1986; Brown and Diesmos 2009) have established mixed support for the model in explaining levels of species richness and endemism (e.g., data on the Philippines: Heaney et al. 2005; Esselstyn and Brown 2009; Siler et al. 2010; Brown et al. 2013). Both the predictions and empirical findings involving tests of the PAIC diversification model have been primarily based upon distributional data coupled with phylogenetic estimates to examine how species diversity is partitioned among different PAICs (for a review see Brown et al. 2013). Moreover, of the handful of studies that have explored whether sea-level oscillations have driven divergence within PAICs at the population level, the results from the genetic analyses are subject to different interpretations (Gorog et al. 2004; Roberts 2006; Esselstyn and Brown 2009; Oaks et al. 2013).

This contrasts with the vast number of biogeographic studies on archipelagos conducted at larger temporal or spatial scales. For example, in the Caribbean, the main focus has been on vicariant diversification and colonization patterns at large geographic scales (Fig. 1A) (Hedges 1996; Glor et al. 2005; Reynolds

EVOLUTION
INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

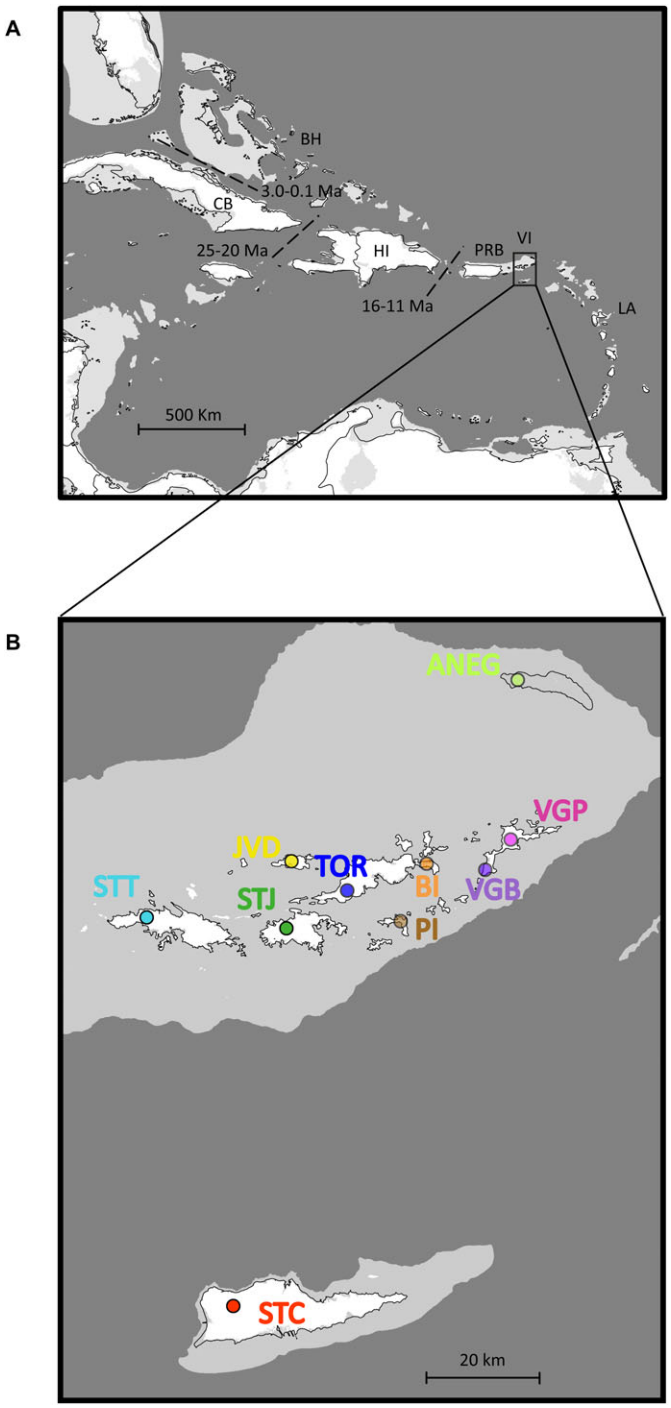

Figure 1. Maps showing (A) the Caribbean with predicted dates of interisland vicariant events (Iturralde-Vinent and MacPhee 1999) (BH, Bahamas; CB, Cuba; HI, Hispaniola; PRB, Puerto Rico Bank; VI, Virgin Islands; LA, Lesser Antilles) and (B) the Virgin Islands with the sampling localities of the ten studied island populations of *A. sanctaecrucis* (STC, St. Croix; STT, St. Thomas; STJ, St. John's; PI, Peter's Island; TOR, Tortola; JVD, Jost Van Dyke; BI, Beef Island; VGB, Virgin Gorda Baths; VGP, Virgin Gorda Peak, and ANEG, Anegada). Current sea-level coastlines are denoted with black lines; topographic and bathymetric information is conveyed by shading. Specifically, elevations 10 m and above current sea levels that would not have been submerged during the last interglacial period are marked in white, whereas lowland areas (elevations <10m and depths <120m below current sea-level) that would have been periodically exposed and submerged during glacial cycles are marked in light gray and deeper areas (depths >120m below current sea-level) are marked in dark gray. All sampled islands are separated by shallow waters (minimum depths separating island pairs range between 19–25 m), apart from St. Croix, which is separated by a very deep sea channel (minimum depth 1900 m). Note that Anegada is light gray because its highest peak is 8.5 m. Topographic and bathymetric data were downloaded as digital elevation models with 1-arc second cell-size (Grothe et al. 2012) from the National Geophysical Data Center. The same population codes are used across all figures and tables.

et al. 2013), with the coalescing-island paradigm largely applied to divergences dating back to the Miocene or the Pliocene (Glor et al. 2004; Thorpe et al. 2010; Sly et al. 2011). Relatively little attention has been paid to patterns of population divergence within PAICs at local geographic scales, despite the existence of such complexes throughout the Bahamas, and Greater and Lesser Antilles (Pregill and Olson 1981; see also Fig. 1A). In fact, tests of the PAIC model, whether based on distributional or genetic data, have been very limited in island systems outside of the Philippine archipelago (e.g., Jordan et al. 2005; Salvi et al. 2014).

The inherent difficulties of studying historical population scenarios that are both recent and dynamic, with repeated cycles of population reconnections, likely contribute to the lack of PAIC studies. For example, the signal of population divergence may be hard to decipher because of patterns arising from incomplete lineage sorting and/or secondary admixture. The limited resolution of traditional genetic markers used in phylogeography (e.g., mtDNA or a few microsatellite loci) may also hamper our ability to distinguish among alternative evolutionary scenarios in such dynamic systems. However, the vast amounts of genome-wide polymorphism data that can now be generated for nonmodel organisms with next generation sequencing (NGS) technologies (Emerson et al. 2010; Peterson et al. 2012) can provide the requisite resolution for addressing fine-scale phylogeographic structure and recent diversification history (Reitzel et al. 2013; Wagner et al. 2013).

Here, we harness the power of NGS to infer the history of divergence of a flightless ground cricket, *Amphiacusta sanctaecrucis*, within a dynamic PAIC using a genome-wide dataset. *Amphiacusta* is a very diverse genus with a predominantly vicariant pattern of diversification across the Caribbean (Oneal et al. 2010). Although there are more than 80 species in the Greater and Lesser Antilles, most of them are single island endemics (Otte and Perez 2009). The species *A. sanctaecrucis* stands out because of its relatively broad distribution across many of the United States and British Virgin Islands (Fig. 1B). These islands, which emerged as part of the Puerto Rican Bank in the late Eocene (Heatwole et al. 1981), have experienced multiple cycles of connections and isolation due to sea level changes up until they became isolated in the Holocene (approx. 7–8 ka; Pregill and Olson 1981), except for St. Croix, which has been isolated since at least the Pliocene (Heatwole et al. 1981). With comprehensive sampling of genomic diversity from populations across the range of *A. sanctaecrucis* we are interested in evaluating different hypotheses derived from the PAIC model that make contrasting predictions about how divergence might proceed. Specifically, the PAIC can be viewed as both a model where divergence would have been promoted, but also one in which diversification may be unlikely. For example, the repeated fragmentation of populations might be expected to promote diversification (Gorog et al. 2004; Esselstyn and Brown 2009). Alternatively, extended gene flow during periods of inter-island connections established by low sea levels (Esselstyn and Brown 2009; Brown et al. 2013), complemented by sporadic oversea dispersal due to short geographic distances, might inhibit diversification across present-day islands. Which process predominates may partly depend on how fast reproductive isolation evolves relative to the tempo of the cycles. Here, we assess the relative roles of island fragmentation and connectivity in driving or inhibiting population divergence of *A. sanctaecrucis* across the Virgin Islands. We compare these results with the respective roles of vicariance and dispersal in the interspecific diversification of the genus *Amphiacusta* across the Caribbean (Oneal et al. 2010), to highlight particular insights the PAIC model provides about the divergence process across local versus larger geographical, as well as temporal, scales.

Methods

SAMPLED POPULATIONS AND GENOMIC LIBRARY PREPARATION

Genomic data were collected for 158 individuals of *A. sanctaecrucis* sampled from each major, and two of the smaller, United States and British Virgin Islands. For each island population 16–18 individuals were sampled, except for the island Anegada, where only six individuals were available (Fig. 1; Table 1). Additionally, genomic data were collected from 16 individuals of a closely related Puerto Rican taxon (i.e., from three populations of *Amphiacusta. sp.*; Oneal et al. 2010), as an outgroup.

Genomic DNA was extracted from the femur of each individual using the Qiagen DNeasy Blood and Tissue Kit. Two reduced representation libraries were constructed using an amplified restriction fragment approach, following the protocol of Parchman et al. (2012). In summary, DNA was doubly digested with the restriction enzymes EcoR1 and MseI, unique barcodes (10 bp) and illumina adapter sequences were ligated to the digested fragments, and the fragments were amplified by PCR. The individually barcoded products were pooled into two groups of 90 and 84 samples each for gel extraction (selected fragments were between 340 and 460 bp) and the two libraries were sequenced in two lanes of an Illumina *HiSeq2000* at the University of Michigan DNA core facility.

PROCESSING OF ILLUMINA DATA

Raw sequence reads were processed in STACKS v.1.07 (Catchen et al. 2011) using a protocol to maximize the number of loci represented in the final dataset. Given that STACKS does not accept reads of unequal length, raw reads were processed as both (i) full 100 bp reads and (ii) 50 bp reads after trimming 50 bp from the 3' end of the sequences using SEQTK (Heng Li, https://github.com/lh3/seqtk). The number of reads processed with trimmed

Table 1. Total number of sequenced specimens per population (n_T) and number of retained individuals after quality filtering (n_R) for each sampled population, along with geographic information.

Code	Island – Locality	Number of samples (n_T/n_R)	Island size (km^2)	Island's highest peak (m)	Coordinates
ANEG	Anegada	6/6	38	8.5	18°43'38.98"N, 64°23'47.36"W
BI	Beef Island*	18/14	5.2	215	18°26'25.64"N, 64°32'27.75"W
JVD	Jost Van Dyke	16/13	8	321	18°26'46.02"N, 64°44'55.08"W
PI	Peter's Island	17/14	7.2	158	18°21'10.55"N, 64°34'18.12"W
STC	Saint Croix	17/14	214	331	17° 44' 15" N, 64°50' 30"W
STJ	Saint John	17/13	51	389	18° 21' 7.92" N, 64°45' 38"W, 18° 20' 35" N, 64°45'54"W
STT	Saint Thomas	18/14	80.9	474	18° 21' 48" N, 64°58' 27" W, 18° 21' 21" N, 64°58' 27" W
TOR	Tortola	16/16	55.7	523	18° 24' 3" N, 64°39' 40" W
VGB	Virgin Gorda – Baths	17/15	21	417	18° 25' 37.18" N, 64°26' 38" W
VGP	Virgin Gorda – Peak	17/17	As above	As above	18°28'40"N, 64°24'12"W
PR1	Puerto Rico – Cerro las Piñas	7/7	9104	1338	18° 9'7.85"N, 66°5'7.59"W
PR2	Puerto Rico – Patillas	7/4	As above	As above	18° 6'4.70"N, 66°2'27.19"W
PR3	Puerto Rico – Maricao	2/1	As above	As above	18°11'10.33"N, 67°4'31.58"W

*Beef Island is connected to Tortola with a short bridge across a shallow canal today.

ends was larger than the full 100 bp reads because shorter sequences, arising from low quality reads near the 3' end or adapter contamination, are incorporated. In both cases, the reads were demultiplexed and filtered using the *process_radtags.pl* script. Only reads with a Phred score >10 (using a sliding window of 15%), unambiguous barcodes, and individuals with more than 1 million reads were retained. Summaries of the number of reads pre- and postprocessing per individual are given for the 100 bp and 50 bp reads in Fig. S1.

Filtered reads of each individual were assembled de novo into putative loci with the USTACKS program. The "removal algorithm" was used to remove overrepresented 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 multigene families, for example) and the "deleveraging algorithm" was used to resolve overmerged loci (i.e., nonhomologous loci misidentified as a single locus). The minimum stack depth (m) was set to 3 (or to 4 only in the case of three individuals with number of filtered reads > 3.5 million) and the distance allowed between stacks (M) was set to 3 and 2, for the 100 bp and 50 bp datasets, respectively. 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 (ε) set to 0.2. A conservative upper bound was selected for the error parameter, as these models have been developed primarily for higher coverage data. Using a conservative bound was preferred over the unbounded model, which has been shown to underestimate heterozygotes (Catchen et al. 2013). A catalog of consensus homologous loci among individuals was built in CSTACKS, using 106 selected individuals (i.e., the 10 individuals (or six, in the case of the Anegada population) with the highest numbers of reads), with the number of mismatches allowed between individuals (n) set to 2. Each individual was matched against the catalog using SSTACKS, with output files loaded on a MYSQL database and exported using the *export_sql.pl* script and the POPULATIONS program in STACKS. For population genetic analyses, SNP data were 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 2013) and the program PLINK (Purcell et al. 2007). Note that although both datasets were analyzed, we only present results from analyses based on the shorter 50 bp reads given the increased number of assembled loci compared to the 100 bp reads (see Supplementary Text 1) and the results were qualitatively similar between datasets (results not shown).

Only single-SNP loci that were sequenced in at least half of the *A. sanctaecrucis* populations (i.e., five out of the 10 populations) and half of the individuals of each population were used in downstream population genetic analyses. For some analyses (which are identified below) more or less inclusive datasets were also analyzed to evaluate the robustness of analyses to different levels of missing data (see Arnold et al. 2013; Huang and Knowles 2014). Only biallelic SNPs are considered here to comply with the assumptions of the current methods for analyzing SNP data.

Table 2. Results of STRUCTURE analyses on the full dataset (10 populations) and hierarchical analyses of subsets of populations (four hierarchical levels).

Level	Populations	Loci	Inds	Mis (*p*/*in*)	1st K	ΔK	F	2nd K	ΔK
1	All 10 populations	5558	136	0.5/0.5	2	1102.5	0.49/0.49	7	6.1
1.1	STT-STC-STJ-JVD	4428	54	0.5/0.5	2	4.7	0.38/0.36	3	1.4
1.1.1	STC-STJ-JVD	3148	40	0.3/0.5	2	64.49	0.25/0.26	3	0.38
1.1.1.1	STJ-JVD	681	26	0/0.5	2	317.6	0.15/0.37	NA	NA
1.2	PI-TOR-BI-VGB-VGP-ANEG	4724	82	0.5/0.5	2	535.4	0.32/0.41	3	0.3
1.2.1	PI-TOR-BI	3381	44	0.33/0.5	3	12.7	0.31/0.38/0.41	2	3.2
1.2.1.1	TOR-BI	2310	30	0/0.5	2	418.45	0.21/0.34	NA	NA
1.2.2	VGB-VGP-ANEG	1851	38	0/0.7	2	304.2	0.12/0.70	3	134.8
1.2.2.1	VGB-VGP	1221	32	0/0.5	2	25.07	0.78/0.001	NA	NA

The number of loci and individuals included in each dataset and the percentage of missing data allowed in each case are shown (at the level of populations, *p*; and individuals per population, *in*). The two most probable *K*-values are presented, as selected using the Evanno method, with their corresponding Δ*K*-values. Average *F*-values for each of the inferred *K* clusters (shown for the optimal *K*) provide an indication of the amount of genetic drift that has occurred in each lineage since their divergence. See posterior probability plots of assignment of individuals to *K* genetic clusters in Figure 3.

GENETIC DIVERSITY STATISTICS AND POPULATION STRUCTURE

The POPULATIONS program in STACKS and the program GENODIVE (Meirmans and Van Tienderen 2004) were used for calculating population genetic statistics, including nucleotide diversity (π), major allele frequency, observed heterozygosity, and Wright's inbreeding coefficient (F_{IS}) at each locus, and average values across loci. Note that the calculation of the inbreeding coefficient excluded loci fixed within each population. To account for possible biases due to unequal sampling of individuals per population, genetic diversity statistics (i.e., observed and expected heterozygosities, and inbreeding coefficients) were also calculated in GENODIVE from 10 random subsamples of six individuals (i.e., the smallest sample of individuals across populations). Additionally, pairwise F_{ST}-values were estimated among populations and their significance was tested in GENODIVE using 10,000 permutations.

Population genetic structure was assessed using STRUCTURE 2.3.4 (Pritchard et al. 2000). Analyses were conducted initially for all 10 populations, and subsequently for subsets of the data corresponding to the identified genetic clusters in a hierarchical order, given that global analyses in STRUCTURE are not necessarily robust to hierarchical geographic structure (see Ryan et al. 2007; Massatti and Knowles 2014) (see Table 2 for the composition of data subsets). Analyses were performed under the "Admixture model" and the "Correlated allele frequency model" for a series of values of K (1–10 for the 10-population dataset, and 1 to n + 1 for the smaller datasets of n populations). Ten independent runs were performed for each value of K, with 200,000 burnin steps and 1 million MCMC iterations. The optimal K for each dataset was chosen using the delta-K method of Evanno et al. (2005) as implemented in the STRUCTURE HARVESTER program (Earl 2012). The cluster membership coefficients (posterior probabilities of individual assignments to K genetic clusters) were permuted across ten independent runs using CLUMPP (Jakobsson and Rosenberg 2007) and plotted using DISTRUCT (Rosenberg 2004).

RELATIONSHIPS BETWEEN GENETIC DIFFERENTIATION AND GEOGRAPHY

We tested for isolation-by-distance (IBD), that is the correlation between pairwise F_{ST}-values and Euclidean geographic distances among populations with a Mantel test from the R package VEGAN (Oksanen et al. 2013). Significance was assessed with 1 million permutations. Given the widespread concerns about the reliability of Mantel tests (Legendre and Fortin 2010; Guillot and Rousset 2013; but see also Kierepka and Latch 2014) we additionally applied a distance-based redundancy analysis (dbRDA; Legendre and Anderon 1999) using the "*capscale*" and "*anova.cca*" functions in VEGAN, after transforming the geographic Euclidean distance matrix to continuous rectangular vectors via principal coordinates analyses (using the "*pcnm*" function).

An association between genetic differentiation and geography was also assessed considering divergence along both latitudinal and longitudinal axes across populations using Principal Component Analysis (PCA). Specifically, PC1 and PC2 scores were used in a Procrustes transformation approach, which maximizes the similarity between PCA maps of genetic variation and geographic locations of sampled populations (see Wang et al. 2010, 2012). This analysis was performed using the *procrustes* and *protest* functions in VEGAN (Oksanen et al. 2013). Major axes for genome-wide SNP data were visualized using the R package ADEGENET ("*glPCA*" function).

Lastly, we estimated population relationships to evaluate the extent to which the history of population divergence corresponds to the geographic configuration of islands. The program

SVDQUARTETS was used (Chifman and Kubatko 2014). This approach accommodates the differences in the genealogical history of individual loci expected to arise from coalescent variation (in contrast to analyzing a concatenated matrix of SNPs; see supplement for such an analysis with the program RAXML v. 8; Stamatakis 2014). This method, as well as the concatenation, do not account for species divergence with gene flow; however, note this assumption is less likely to be violated in this flightless cricket species (which can be corroborated visually by patterns of genetic variation in the PCAs; see discussion). As the coalescent-based method for analysis of SNP data is robust to missing data (Chifman and Kubatko 2014), a more inclusive dataset was used (i.e., a dataset with all single-SNP biallelic loci sequenced in three or more populations), and included the outgroup from Puerto Rico.

DIVERGENCE TIME ESTIMATION

To provide an approximate time framework for population divergence across the Virgin Island PAIC and between the Virgin Islands and Puerto Rico, we used a composite-likelihood simulation-based approach (Excoffier et al. 2013), implemented in FASTSIMCOAL2 (Excoffier and Foll 2011), which estimates demographic parameters from the site frequency spectrum (SFS). A folded joint SFS (i.e., for the minor allele, in the absence of information for the derived state) was calculated for each population pair based on loci containing one to three SNPs (i.e., monomorphic sites were not considered for parameter inference; see "*removeZeroSFS*" option in FASTSIMCOAL2), representing regions of different mutation rates across the genome; only a single SNP per locus was considered for the calculation to avoid the effects of linkage disequilibrium. To remove all missing data for the calculation of the joint SFS, each population was subsampled using a custom script (available on Dryad) and only loci found in at least 10 individuals per population were retained to minimize errors with allele frequency estimates.

Divergence times were estimated accounting for the possibility of migration. To improve the performance of the models by reducing the number of parameters estimated from the data (Excoffier et al. 2013), one population parameter was calculated directly from the data. Specifically, the effective population size for one population (N_1) was fixed (as explained below), whereas the size of the other population, N_2, the ancestral population size N_A, the divergence time T_{DIV}, and the migration rate m were estimated based on the SFS. The effective population size that was fixed in the model was calculated from nucleotide diversity estimates from loci that contained zero to three SNPs and were present in at least half of the individuals per population, and a mutation rate of 3.5×10^{-9} per site per generation (Keightley et al. 2009). Divergence times were estimated between representative pairs of populations to compare the divergence between Puerto Rico and the Virgin Islands relative to the divergence among islands within the PAIC (see Table 3 for details). Representative populations from the major Virgin Islands and Puerto Rico were chosen based on their relative geographic position, as well as on the percentage of overlap among sequenced loci, in order to maximize the number of available loci without missing data for the calculation of the joint SFS. Note that we had to opt for pairwise population models, rather than a global model based on a history of population divergence estimated phylogenetically because the number of loci without missing data in common to all populations was prohibitively small ($\sim$ 370 loci) for analyses with FASTSIMCOAL2. Likewise, we did not include the population of Anegada as one of the populations explored with the models because of the small number of individuals sampled, which did not allow for the subsampling procedure performed to standardize the number of loci sampled in each population. 40 runs per population pair were conducted and the global maximum likelihood solution is presented. Each run was performed with 100,000–250,000 simulations per likelihood estimation and 10–40 expectation-conditional maximization (ECM) cycles, based on a stopping criterion of a 0.001 relative difference between iterations. Parameter confidence intervals were calculated from 100 parametric bootstrap replicates, by simulating SFS with the same number of SNPs from the maximum composite likelihood estimates and reestimating parameters each time (Excoffier et al. 2013).

Table 3. Results of divergence time estimation with Fastsimcoal2.

Pop pair	Loci	Point estimate	95% CI
PR-STT	1768	711,426 (7.3N)	527,595–687,115
PR-STJ	4330	787,143 (8.1N)	562,878–753,018
STT-TOR	2979	107,328 (0.9N)	92,514–127,626
STJ-VGP	4891	102,336 (0.8N)	90,554–123,944
TOR-VGB	3299	80,914 (0.6N)	71,403–108,317
STJ-STC	3141	18,899 (0.15N)	17,435–24,069

Composite maximum likelihood estimates of divergence times are presented as the number of generations (i.e., number of years ago, with one generation per year) and as a function of the effective population size in parentheses (see Table S4 for estimated population sizes). 95% confidence intervals were calculated from 100 parametric bootstrap replicates. In all cases, apart from the last population pair (St. John—St. Croix) the estimation was performed under an Isolation-with-Migration model. See Table S4 for estimated parameters under alternative models. The number of loci that were used for the calculation of the joint site frequency spectrum for each population pair is indicated. Population codes are the same as in Table 1 and Figure 1.

Results

GENETIC DIVERSITY STATISTICS

Of the approximately 300 million reads generated, after processing of data (see Supplementary Text 1) and applying a

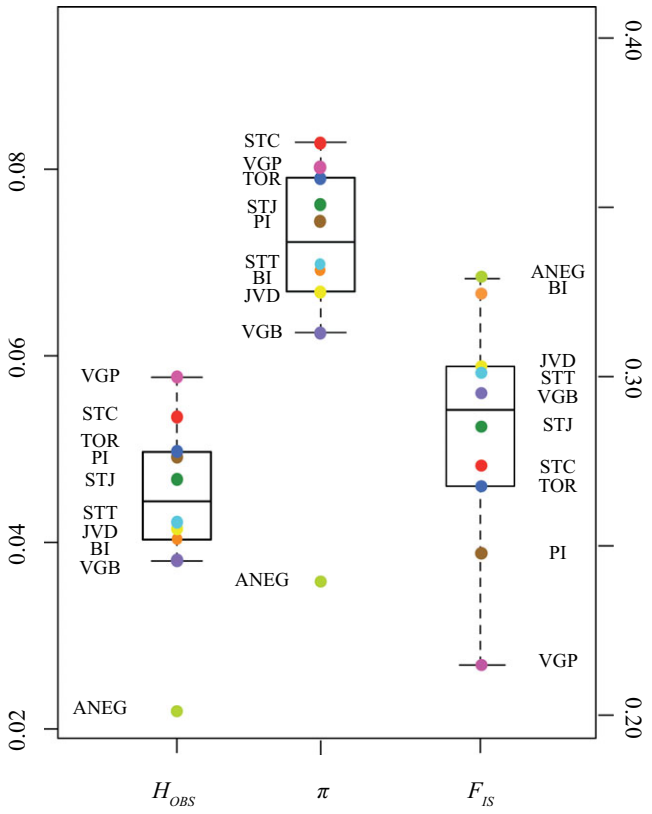

Figure 2. Comparison of genetic diversity statistics across populations, based on 5558 SNPs and averaged across loci. Specifically, observed heterozygosity, H_{obs}, and nucleotide diversity, π, with corresponding values are shown on the *y*-axis on the left, and on the *y*-axis on the right, are corresponding values of inbreeding coefficients, F_{IS}. Each population is color coded, with corresponding abreviated names listed as well (see Fig. 1 for distribution of populations). For each summary statistic, the median, the first, and third quartiles, standard deviation, and range across populations are shown in the box-and-whisker plots.

tolerance threshold of 50% missing data across the ten populations, 5558 loci with a single biallelic SNP were retained for assessing population structure and genetic differentiation. The more inclusive dataset with the higher tolerance of missing data used to estimate population relationships with the coalescent-based program SVDQUARTETS contained 82,443 single-SNP loci, while the dataset used for estimation of divergence times in FASTSIMCOAL2 included 32,096 loci.

The within-population genetic diversity statistics showed that populations did not differ substantially from one another with the exception of the Anegada population (Fig. 2; Table S2); the genetic diversity statistics were also qualitatively similar whether calculated with the POPULATIONS program or GENODIVE. The generally lower nucleotide diversity and heterozygosity observed for the Anegada population (notably 99.15% of the loci are fixed) was not an artifact of the smaller number of individuals sampled from that island; the results were consistent when the genetic diversity statistics were calculated for subsamples of six individuals for all populations (Fig. S3). Island size was not a good predictor of genetic diversity. For example, two of the smallest islands (Jost Van dyke and Peter's Island; Fig. 1) were not particularly genetically depauperate compared to the larger islands.

STRUCTURING OF POPULATION GENETIC VARIATION

The STRUCTURE analyses of the 10-population dataset identified a generally strong correspondence between geography and inferred genetic clusters for each of the different *K*-values ($K = 2$–10). The most probable *K* value based on the Evanno method of ΔK was $K = 2$, while the next probable value ($K = 7$) had a much smaller ΔK (Table 2). Subsequent analyses of separate subgroups of populations identified from the respective genetic clusters to account for the hierarchical structure of populations (i.e., when populations share a nested set of common ancestors; see Ryan et al. 2007; Massatti and Knowles 2014) also identified a strong correspondence between geography and membership to inferred genetic cluster (Fig. 3). The lowest level of the hierarchical analyses identified clusters confined to a single present-day island. High levels of admixture were only apparent between the two populations located on Virgin Gorda (Fig. 3). The *F*-value estimated by STRUCTURE as an indication of the amount of genetic drift that has occurred in each lineage since their divergence (Pritchard et al. 2000; Harter et al. 2004), was particularly elevated for the genetic cluster corresponding to the Anegada population (Table 2).

Pairwise F_{ST}-values were all significantly different from zero based on 10,000 permutations ($P < 0.005$), even after correcting for multiple comparisons using a Bonferroni correction (Table S3). Pairwise F_{ST}-values among populations were not significantly associated with the geographic distances separating populations (Mantel test statistic $r = 0.47$, $P = 0.08$; dbRDA $P = 0.31$), but the association was highly significant if the St. Croix population was excluded (Mantel test $r = 0.82$, $P = 0.0001$; dbRDA $P = 0.001$) (Fig. 4).

The Procrustes-based analysis to quantify the association between genetic variation and geographic positions of populations in multivariate space (Fig. 5) identified a significant similarity score ($t_0 = 0.664$, $P < 0.0001$), which increased to $t_0 = 0.772$ when St. Croix and Anegada populations were excluded (Fig. S4), which are the two populations that deviated the most from the general expectation based on the geographic positioning of the islands (Fig. 5). Specifically, individuals from both St. Croix and Anegada are genetically more similar to the individuals from the other populations than would be expected given the geographic distance separating these two islands from the other islands. Note that the general pattern of association with geography—specifically, the extent and direction of distortions of the individuals in PC space

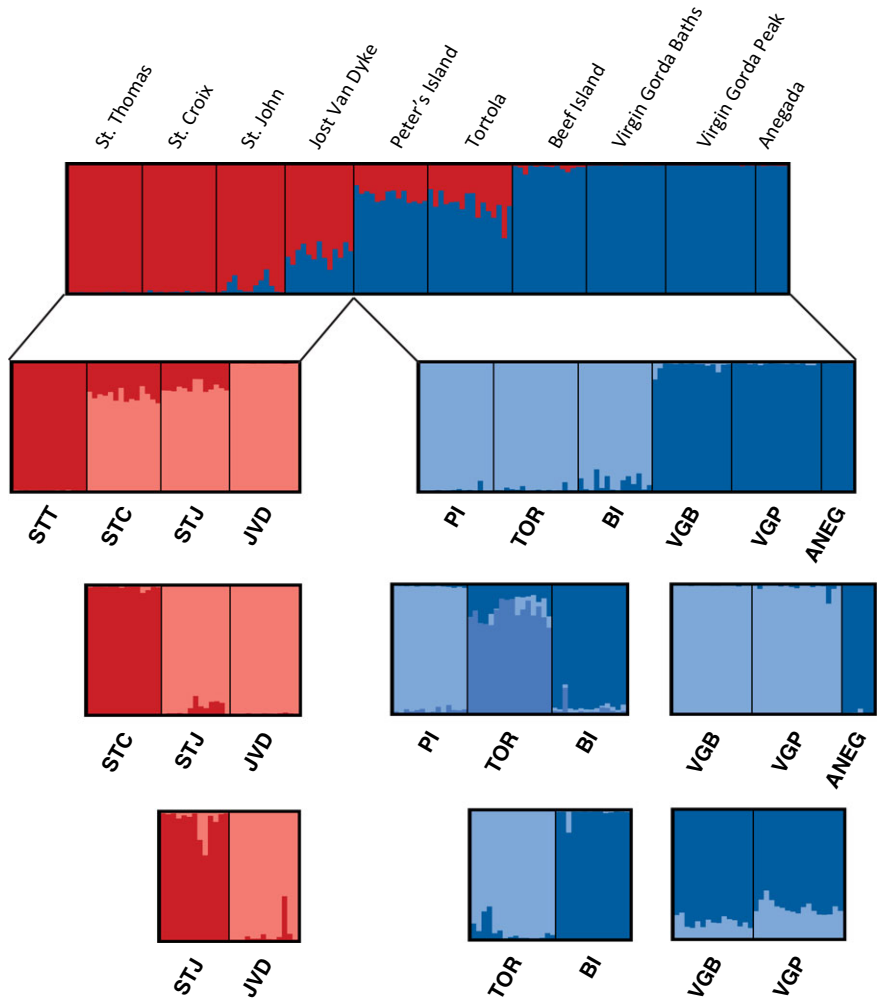

Figure 3. Results of population STRUCTURE analyses with posterior probability plots of individual assignments to the inferred genetic clusters shown for the most probable *K*-value, as well as a series of hierarchical analyses of subsets of populations (defined by the inferred genetic clusters at the previous level). The optimal *K*-value was two for most datasets (see Table 2), apart from the analysis of the Peter's Island, Tortola, and Beef Island populations (PI-TOR-BI) where the optimal $K = 3$. Note that we applied a similar tolerance threshold for missing data to analyses of population subsets, and therefore these respective datasets were generally smaller than the 5558 SNPs identified across all ten populations (see Table 2).

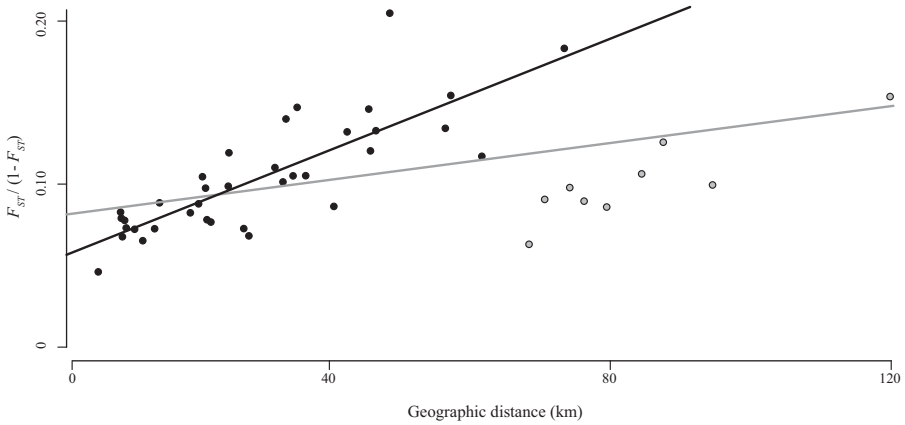

Figure 4. Association between genetic differentiation (measured by pairwise F_{ST}-values) and the Euclidean geographic distance between populations, including all populations (shown in gray; $r = 0.47$, $P = 0.08$), or excluding St. Croix (shown in black; $r = 0.82$, $P = 0.0001$) tested using Mantel tests.

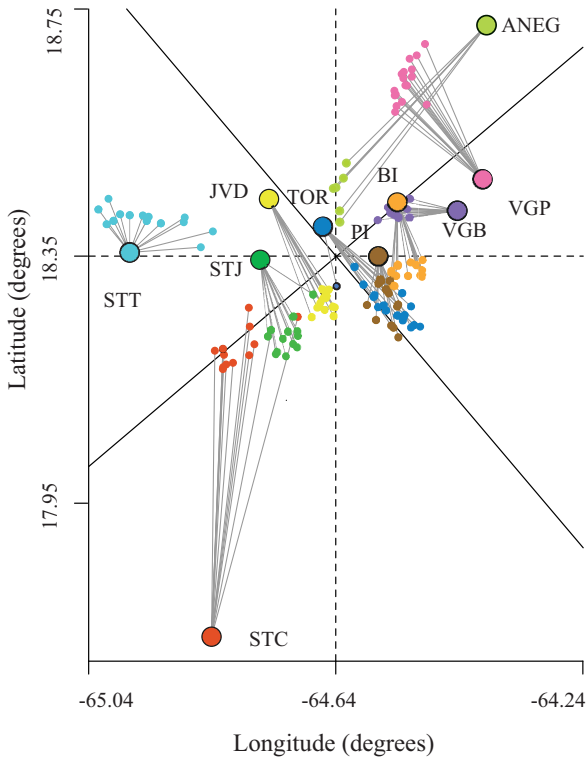

Figure 5. Procrustes-transformed PCA plot of genetic variation with each individual mapped in PC space relative to the geographic location of populations. The length of the line connecting individuals in PC space to their geographic location (shown by the large colored dots; see Fig. 1 for definitions of abbreviated island labels) represent the extent of the deviation from the expected pattern of genetic variation based on geography. Solid axes show the orientation of PC1 and PC2 for the genetic data (explaining 7% and 4% of the genetic variation, respectively) relative to the geographical longitude and latitudinal axes (shown by the dashed lines). The Procrustes similarity score increased from t_0 = 0.664–0.772 when the analysis was repeated without the St. Croix (STC) and Anegada (ANEG) populations (Fig. S4).

based on genetic coordinates—was robust after St. Croix and Anegada were excluded (Fig. S4).

The strong geographic structuring of individuals across the islands was apparent in the estimation of population relationships (Fig. 6, Fig. S5). Most island populations were monophyletic, with the exception of Tortola and St. John's. Specifically, Tortola had one individual that did not cluster with the other individuals from the same island and St. John's was polyphyletic (or paraphyletic if two individuals are excluded; see also Fig. S5). Within the island Virgin Gorda, the two populations were not sorted from each other. The concatenated RAxML tree on the 5558 SNPs dataset estimated the same pattern of relationships among islands (Fig. S6) with the primary difference pertaining to the general distinctiveness of populations sampled within islands. Specifically, with the concatenated dataset Tortola and the Virgin Gorda Peak (VGP) populations are monophyletic, and the Virgin Gorda Baths population (VGB) is paraphyletic in respect to VGP.

DIVERGENCE TIME ESTIMATION

Divergence time estimates among islands of the PAIC largely overlap (Fig. 7), ranging from 71–108 ka (e.g., between Tortola and Virgin Gorda populations; Table 3) to 92–127 ka (e.g., between St. Thomas and Tortola; Table 3), in association with a period of intense sea-level fluctuations causing repeated connection and isolation of the islands (75–115 ka; Fig. 7) following a major sea-level rise of 10 m above current levels during the Eemian interglacial period (115–130 ka). Much older divergence times (i.e., around 700 ka; Table 3) were estimated between the *A. sanctaecrucis* Virgin Island populations and the *Amphiacusta sp*. population from Puerto Rico. Divergence time estimates also confirm the more recent colonization of St. Croix, estimated at 17–24 ka (based on a model of divergence without migration, which is biologically most likely given the geologic history of the island; but note that irrespective of the model used, a recent divergence time is supported; see Table S4). Migration estimates within the Virgin Islands ranged between 0.8 and 4 migrants per generation, which was two to three orders of magnitude higher than between Virgin Islands and Puerto Rico (see Table S4 for estimated parameter values). A divergence without migration model and a model allowing for migration only before the last separation of the islands at 8 ky were also run and produced qualitatively similar estimates for divergence times (Table S4).

Discussion

High levels of population divergence within the Virgin Island PAIC that predate the last separation of the islands at 7–8 ka, as well as a strong correspondence between genetic differentiation and the geographic configuration of the islands, demonstrate that populations of *A. sanctaecrucis* are not homogenized by gene flow during extended periods of island connections. Instead, the analyses are consistent with the role of PAIC systems in promoting genomic divergence. This local population divergence within the PAIC parallels the inter-island vicariance observed in patterns of diversification across the genus in these Caribbean crickets. Yet, despite the frequency and recency of speciation events within other Caribbean Islands such as Puerto Rico (Oneal et al. 2010), divergence across the PAIC does not translate into species diversification. Unlike the high diversity of endemics found on the larger Caribbean islands, the Virgin Island PAIC is characterized by a single species, *A. sanctaecrucis*. Below we discuss what this pattern implies about the hypothesis of PAICs as species pumps, and speciation in these Caribbean crickets.

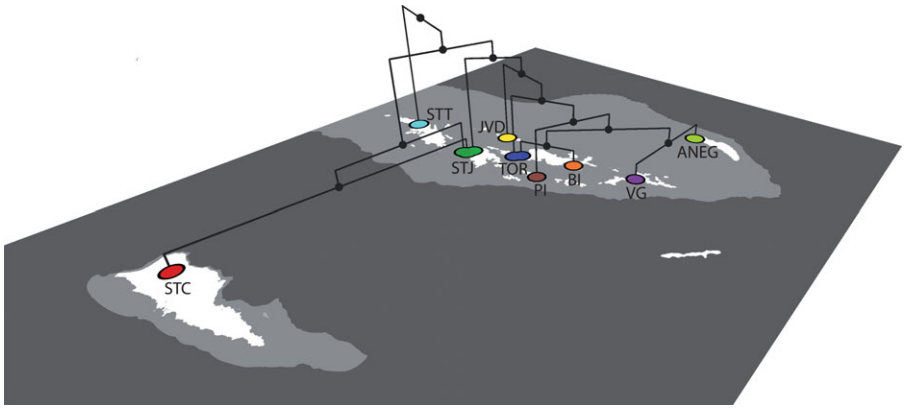

Figure 6. History of population relationships estimated for the 136 *A. sanctaecrucis* individuals based on 82440 SNPs analyzed with SVDQUARTETS (Chifman and Kubatko 2014). The software GENGIS (Parks et al. 2009) was used for plotting the tree against the geographic map of the archipelago. Colored circles represent monophyletic groups of individuals from the same island. A rooted tree with 12 outgroup individuals from Puerto Rico (not shown here; see Fig. S5) supports the relative branching order shown here, with the basal split from Puerto Rico occurring at the node that separates St. Thomas (in light blue) from the rest of the PAIC islands, and suggests a stepping-stone pattern of colonization from the West to the East of the archipelago. Note that the two populations sampled from Virgin Gorda (VGB and VGP) have been merged here into a single group labeled "VG."

DIVERGENCE WITHIN A CARIBBEAN PLEISTOCENE AGGREGATE ISLAND COMPLEX (PAIC)

Despite the short geographic distances separating the individual Virgin Islands that formed a PAIC (Fig. 1), and the recent and extended island connections (Fig. 7), the crickets sampled from the separate islands are genetically distinct (Fig. 3). Moreover, the estimated divergence times within the PAIC (Fig. 7) date back to a period of repeated island connection and isolation (75–115 ka) and the Eemian rise of the sea-level (115–130 ka) (Fig. 7), rather than the most recent separation of the islands (7–8 ka; Pregill and Olson 1981). As such the analyses are generally consistent with the role of PAICs as arenas for divergence. Nevertheless, not all of the genetic differentiation follows directly from the breakup of the PAIC with rising sea levels or fit with a colonization model of isolation-by-distance (IBD).

As noted, patterns of genetic differentiation are not predicted by the geographic position of St. Croix or Anegada in relation to the other populations, and the geographic position of St. Croix does not fit expectations based on IBD (Figs. 4 and 5). Examination of the genomic composition of individuals in multidimensional space fits expectations in relation to the longitudinal position of islands. However, the genetic differentiation among individuals is smaller than expected based on the latitudinal positioning of islands (Fig. 5). This deviation may not be entirely unexpected given that the northernmost island Anegada is a flat and low atoll that lacks the topographic relief of the other islands (Fig. 1B), perhaps making it less likely that its populations would have persisted over time. The significantly lower genetic diversity and increased inbreeding coefficient (Fig. 2; Table S2), and corresponding displacement of the Anegada population in PCA space is consistent with strong population bottlenecks. The island Anegada only recently emerged 119–130 ka (Gore 2013), in contrast to the other islands of the PAIC that have existed throughout the Pleistocene. While a more recent colonization of Anegada may also have contributed to the lower than expected genetic differentiation in the PCA-space from the Procrustes analysis (Fig. 5), unfortunately limited sampling prohibits a definitive test of this possibility (see Methods). Note that the general fit of Anegada with expectations of IBD based on analyses of pairwise population F_{ST} -values are not necessarily inconsistent with the Procrustes analysis because it may reflect a general inflation of F_{ST} due to a high percentage of fixed loci (Table S2).

In contrast to Anegada, the deviation for St. Croix in terms of expected patterns of genetic differentiation based on geography (Figs. 4 and 5) is surprising. With the deep channel separating St. Croix, any incipient differentiation of these crickets should have been protected from the homogenizing effects of recurrent gene flow. This peculiar pattern is not unique to *A. sanctaecrucis*, as recent dispersal between the Virgin Island PAIC and St. Croix has been inferred for other low-vagility organisms too (Hedges and Heinicke 2007; Falk and Perkins 2013). What fosters such dispersal events is not clear. Although hurricanes and storms could contribute to interisland movements, the north-to-south dispersal route involved in the colonization of St. Croix makes dispersal under such mechanisms less probable (Hedges 2006). Moreover, given the level of genetic distinctiveness of individuals from St. Croix and the estimated divergence time at 20 ka (Table 3) it

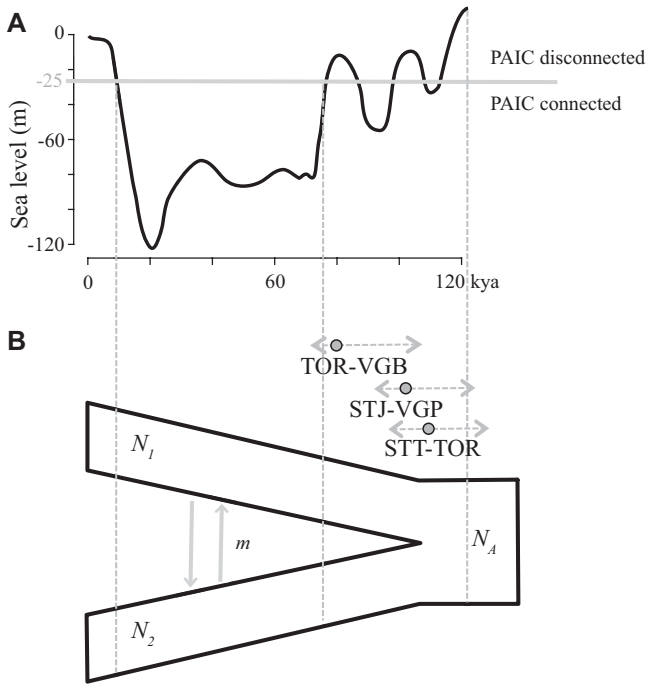

Figure 7. Schematic of (A) sea level change curve in comparison with (B) divergence time estimates within the Virgin Island PAIC under an Isolation-with-Migration model. For the reconstruction of global sea-level change during the last 120 ka (taken from Rijsdijk et al. 2014, based on data from Camoin et al. 2004), the continuous gray line indicates that a 25 m drop in the sea level would connect all the islands within the PAIC to each other. Note that this reconstruction is not specific to the Caribbean, but it serves as an approximation for the timing of island connections, given that tectonic uplift in the region has been minimal during the late Pleistocene and Holocene (Whetten 1966; Hubbard et al. 2008). Point estimates of divergence times under an Isolation-with-Migration model are shown as gray circles with dashed gray arrows marking the 95% confidence intervals.

is unlikely that the pattern of relatedness (Fig. 5) reflects recent human-associated dispersal from St. John, especially given the accuracy of assignment of individuals to their respective populations (Fig. 3).

With respect to the colonization from Puerto Rico to the Virgin Islands and eastwards across the PAIC, it would be reasonable to assume that it occurred through land bridges during periods of low sea-levels, as overwater dispersal from West to East is also unlikely in this region, based on information on hurricane tracks (Hedges 1996). The minimum depth separating any pair of islands (apart from St. Croix) is in fact very similar (Fig. S7), suggesting that the timing of the last interisland connections is not expected to differ greatly across the PAIC, which is confirmed by the largely overlapping estimates of divergence among the three populations pairs (Fig. 7).

The pattern of interisland connections, and not present-day island size per se, appears to best predict genetic patterns. For example, neither Peter's Island nor Jost Van Dyke—the two smallest islands sampled—is genetically less distinct or differentiated from other geographically proximate islands (Fig. 5), and they have comparable genetic diversities relative to the other islands (Fig. 2). It may be that for crickets, the minimal sustainable habitat for population persistence is represented on all islands, irrespective of their size, although we currently lack information to evaluate this hypothesis. A possible alternative may be that total habitat availability during low sea-level periods rather than present-day island size is predictive of genetic diversity patterns (see also Jordan et al. 2005; Salvi et al. 2014). However, the smaller islands fit the isolation-by-distance (Fig. 4) and stepping-stone model of colonization (Fig. 6) and do not deviate in PC space any more than the larger islands of the PAIC from their expected genomic positions based on geography (Fig. 5). This suggests that if past habitat availability contributes to current genetic patterns, island size does not mediate this effect.

PAICs AS SPECIES PUMPS OR SIMPLY ISOLATION-BY-DISTANCE?

Our results support the contention that divergence across short geographic scales can occur within PAICs. However, does differentiation necessarily imply a role for the cycles of past island connections and fragmentation among the constituent islands of the PAIC, as opposed to simply isolation-by-distance?

There is a significant isolation-by-distance (IBD) pattern among the islands of the PAIC (i.e., when St. Croix, which is not a member of the PAIC, is excluded; Fig. 4). Moreover, given the east-west configuration of the contemporary islands, the exchange of individuals through past island connections would not necessarily erase a signal of IBD (see also He et al. 2013), thereby making it difficult to rule out a role of isolation-by-distance in structuring genetic variation. Potential comparisons with a similar set of densely sampled populations from a large unfragmented island, such as Puerto Rico, would provide further insights into the role of connectivity cycles in structuring genetic variation of *Amphiacusta*, as opposed to a strictly IBD model, but such data are currently unavailable. Nevertheless, there are particular signals in the existing data suggesting that past island connections, not simply IBD, has contributed to patterns of divergence across the Virgin Islands. Interestingly, some of the most extreme changes in the geographic configuration of islands predicted during periods of low sea level occurred in the easternmost area of the PAIC, which is characterized by a very shallow and broad shelf (Fig. S7). The eastern islands are also the most notable with respect to substantial longitudinal deviations from expected patterns of genetic divergence under an IBD model (Fig. 5), suggesting that exposed island areas may indeed have contributed to spatial

patterns of population differentiation by providing opportunities for colonization or expansion during low sea-level periods, even though a pattern of IBD, which is inherent to the geographic configuration of the Virgin Islands that comprise the PAIC, also structures genetic variation.

Given that levels of differentiation among the constituent islands of the PAIC are greater than those observed between the geologically, and geographically isolated island St. Croix, clearly genetic distinctiveness can be maintained among islands of a PAIC. However, these high levels of genetic distinctiveness do not necessarily confirm a role of PAICs as "species pumps". For speciation to be completed, within-species lineages (i.e., isolated island populations forming separate gene pools) need to persist and differentiate for a sufficiently long period of time to develop reproductive isolation (Allmon 1992; Ricklefs and Bermingham 2007; Dynesius and Jansson 2014). Failure to persist long enough can be caused either by extinction or by merging with other gene pools due to increased gene flow, making genetic differentiation ephemeral (Futuyma 1987; Dynesius and Jansson 2014). Genetic differentiation among populations of the Virgin Island PAIC is not rendered ephemeral in the strictest sense given the observed differentiation has persisted throughout the extended periods of island connections over the last ~100 ka. However, the divergence time estimates are still relatively recent, raising a question about long-term population persistence on these islands. Admittedly, these estimates might be conservative to some extent (e.g., excluding loci with missing data from RAD-seq datasets can favor loci with lower mutation rates; Huang and Knowles 2014) or inaccurate due to the use of the spontaneous mutation rate of *Drosophila* (Keightley et al. 2009). These potential impacts on estimates of the absolute timing of divergence, however, do not compromise the relatively recent divergence times observed within the Virgin Island PAIC (< 1 N) compared to the older divergence between *A. sanctaecrucis* and its closely related species from Puerto Rico (>7 N), which is robust across models (Table S4) and unaffected by mutation rate biases. Could this suggest that more ancient differentiation among the Virgin Island populations has been wiped out due to higher extinction rates during high sea-level periods of interglacial maxima? The small size of the individual Virgin islands in combination with climatic and sea-level changes might drive populations to extinction more often than on larger islands (Ricklefs and Bermingham 2004; Davalos and Russell 2012), such as Puerto Rico. The extent of topographic relief on the large Caribbean islands like Puerto Rico might also have an important effect on levels of local extinction (see also Jordan et al. 2005). In the extreme cases of low and flat islands, such as Anegada, which become completely submerged during interglacial maxima (Fig. 1), island populations would have been subjected to periodic local extinction and recolonization inhibiting diversification (Slatkin 1977; Slatkin 1985; Futuyma 1987). However, both Virgin Gorda and Tortola (Fig. 1) are characterized by relatively high topographic relief (Table 1) that would have afforded refuge to rising sea levels so not all the island populations would have undergone pronounced latitudinal or longitudinal shifts to persist. Yet neither of these islands carries a genetic signature characteristic of local persistence that differs from any of the other islands (Figs. 2, 5).

Apart from extinction rates, an additional factor that may prevent PAICs from acting as "species pumps" is the increased gene flow during periods of interisland connections, which will tend to erase differentiation among island populations (Futuyma 1987; Rhymer and Simberloff 1996). Given the particularly shallow sea-levels separating the Virgin islands (Fig. S7), periods of land-bridge connections via the exposed shelf have been longer than periods of island isolation during the Pleistocene (Fig. 7), and these extended intervals of presumably increased gene flow might be responsible for "reversing" the speciation process (e.g., Seehausen 2006; Taylor et al. 2006). Our results suggest that migration rates during the last ~70 ka of island connections were not high enough to homogenize island populations, but levels of gene flow may have been higher during previous low sea-level periods, as habitat availability on the exposed shelf will depend on climatic conditions which have not been uniform across different glacial cycles.

Additional studies will be needed to generalize the effects of island connectivity cycles on other Virgin Island taxa or other Caribbean PAIC systems, as patterns of gene flow and population persistence might be taxon-specific and will also depend on particular characteristics of each island system. For example, given the short geographic distances and small island sizes, the high levels of population divergence found in these flightless, forest-associated crickets would not be expected for Virgin Island taxa with better over-sea dispersal abilities and habitat generalists with increased propensity to cross the exposed shelf, or taxa with greater area requirements leading to higher extinction-recolonization dynamics, and thus a less static model than the one observed for *Amphiacusta*. However, given where *Amphiacusta* is likely to fall on the spectrum of minimum area requirements for diversification based on comparisons with other taxa (see Discussion below; also Kisel and Barraclough 2010), it is not clear what properties of taxa would be more likely to facilitate the persistence of divergent populations that might eventually lead to speciation under such a dynamic biogeographic setting. This suggests that other intrinsic traits associated with divergence, and specifically, species-specific traits subject to selective divergence, might be essential for maintaining incipient divergences.

Even when focusing on more ancient divergences (e.g. Miocene) the completeness of allopatric speciation after secondary island connections depends on the taxon and island of

interest. While the island fragmentation paradigm has been considered as a main driver of diversification for Cuban anoles (Glor et al. 2004) and the avian fauna of Hispaniola (Sly et al. 2011), the example of the Lesser Antillean anoles shows that even such ancient divergences do not always lead to complete allopatric speciation (Thorpe et al. 2008, 2010). Lineages that have diverged on ancient (6–8 My) precursor islands and have recently come into secondary contact on the island of Martinique are not reproductively isolated, while ecological adaptation is currently playing a more important role in defining levels of genetic differentiation than previous allopatric divergence (Thorpe et al. 2010, 2012). On the other end of the spectrum, rapid bursts of diversification may even exceed the tempo of sea-level oscillations, especially in taxa that are prone to strong ecological selection. For example, in the Bahamas, rapid post-Pleistocene radiations of mosquitofish and pupfish have been driven by ecological adaptation to predator regimes (Langerhans et al. 2007; Heinen et al. 2013) or novel trophic niches (Martin and Feinstein 2014), respectively, since the last inundation of the islands.

SPECIATION IN CARIBBEAN CRICKETS

The relative role of different biogeographic processes in the history of divergence of *A. sanctaecrucis* identified from genomic data (Fig. 6) are similar to those inferred from phylogenetic analyses of species diversity patterns in the genus *Amphiacusta* (Oneal et al. 2010). That is, the results from both population and species-level analyses suggest that these flightless crickets tend to diversify across local geographical scales and do not commonly cross sea-barriers, even though they are apparently capable of occasional long-distance over-water colonization (as in the cases of Jamaica or St. Croix). Despite these strong parallels between the population and species-level analyses that support divergence over localized geographic scales, there is nonetheless a striking difference. *Amphiacusta sanctaecrucis* is the only species distributed across the Virgin Islands. In contrast, multiple endemic taxa that arose in situ (i.e., intraisland speciation) characterize the other islands, many of which are recently diversified species (80–480 ka; Oneal et al. 2010). This difference in species diversity is not simply a function of the size of the area encompassed by the Virgin Island PAIC (~4500 km^2 at low sea level periods), which is more comparable to the size of Puerto Rico (~9100 km^2; see Fig. 1) than to the currently exposed area of the Virgin Islands (~517 km^2). Neither is simply a function of island age, given that the Virgin Islands emerged in the late Eocene as part of the Puerto Rican Bank (Heatwole et al. 1981) and that their surface has been inhabitable for a sufficiently long period for speciation to occur, given the recency of speciation estimated for other *Amphiacusta* taxa (Oneal et al. 2010). Despite the recent and repeated connections between the Virgin Island PAIC and Puerto Rico during low sea-level periods (occurring at a similar tempo as the cycles of connection and isolation among the Virgin Islands; Heatwole and MacKenzie 1967) there are several animal or plant lineages endemic to the Virgin Island PAIC (e.g., Acevedo-Rodríguez, 1996; Henderson and Powell, 1999). However, single-island endemism on individual present-day islands is very low (Heatwole and MacKenzie 1967). Similarly, *A. sanctaecrucis* has persisted as a Virgin Island endemic throughout multiple glacial cycles (~700 ka divergence from the Puerto Rican lineage; Table 3), but regardless of its high levels of genomic divergence and the apparently high intrinsic speciation rate of the genus *Amphiacusta* (Oneal et al., 2010), speciation has not occurred within the Virgin Island PAIC. This contrasting pattern presents somewhat of a paradox about the speciation process. If differentiation apparently occurs at local geographic and short temporal scales (based on phylogenetic analyses and the geographic distribution of other *Amphiacusta* taxa; Oneal et al. 2010), why are individual islands comprising the Virgin Island PAIC not inhabited by a diverse set of closely related endemics?

This apparent paradox might be explained by the overall small size of the individual Virgin Islands that comprise the PAIC. Analyses of species distribution patterns suggest that in situ speciation is rare on small islands (e.g., smaller than 3000 km^2 for Caribbean anoles; Losos and Schluter 2000). However, the minimum area for speciation to occur can vary greatly among taxa with different levels of intraspecific gene flow (0.8–500,000 km^2; Kisel and Barraclough 2010). With a significant within-island genetic differentiation across short geographic distances (e.g., $F_{ST} = 0.15$ between the two Puerto Rican populations PR1 and PR2, which are 7.5 km apart), *Amphiacusta* is expected to require relatively small minimum areas for diversification (Kisel and Barraclough 2010). While the total size of the Virgin Island PAIC is not necessarily below the area limits for speciation for this taxon, the individual islands themselves would appear to fall within the range where in situ diversification is unlikely.

Apart from the effect of geographic distance per se, the presumed relationship between island size and prevalence for speciation is commonly attributed to the greater topographical complexity of larger islands providing more opportunities for allopatric isolation and/or to the greater ecological diversity of larger islands providing more opportunities for niche partitioning and species coexistence (Losos and Ricklefs 2009). Even if the relative contributions of geographical isolation and ecological adaptation in the early stages of population divergence are difficult to disentagle and not mutually exclusive (Wang et al. 2013), several recent studies have emphasized the critical role of ecological divergence in the speciation process (Schluter and Conte 2009; Thorpe et al. 2010), both from a perspective of maintaining incipient divergence through reduction in levels of gene flow due to maladaptation of immigrants and hybrids (Crispo et al. 2006; Lee and Mitchell-Olds 2011; Edelaar et al. 2012), as well as from

a perspective of occupying a finite niche space (Parent and Crespi 2006; Esselstyn et al. 2011; Martin and Feinstein 2014). However, unlike the textbook example of the Caribbean anoles which have diversified into similar sets of habitat specialists ("ecomorphs") on each of the Greater Antilles (Williams 1983; Losos et al. 1998), speciation of *Amphiacusta* crickets on the same islands has notably occurred without consistent ecological divergence, with most taxa occupying wet tropical forest with similar life histories and life styles (i.e., all are flightless ground crickets, in contrast to groups like tree crickets that partition a canopy), and shifts to dry forest are relatively rare (Oneal and Knowles 2013).

Since niche space (or more specifically habitat diversity) does not appear to be the limiting factor in the diversification of the genus *Amphiacusta*, and given the high levels of population differentiation observed in *A. sanctaecrucis* suggesting that the small geographic scale per se or lack of topographic complexity is not the limiting step (i.e., clearly divergence is not inhibited, but rather sea-barriers provide multiple opportunities for allopatric isolation among the Virgin Islands comprising the PAIC), the lack of speciation in association with the PAIC would appear to relate to the ephemerality of genetic divergence. That is, species diversity patterns in the crickets might reflect processes related to the maintenance of divergence (see also Seehausen 2006), not just the drivers of divergence (Futuyma 1987; Dynesius and Jansson 2014). In this sense, the processes associated with the species pump hypotheses might very well be operating; however, the genomic divergence is nonetheless ephemeral due to factors that may not be directly related to the island connectivity cycles per se.

Conclusions

The diversification of *A. sanctaecrucis* in the Virgin Islands highlights the importance of PAIC systems as drivers of divergence at local geographic and short-temporal scale, which has been largely neglected in the Caribbean. However, whether this divergence will lead to speciation under a hypothesized "species pump" depends on factors that determine whether incipient divergence persists, which may be related to taxon-specific traits, geography, and ecological adaptation, as illustrated by comparisons of our findings with other taxa and island systems. We emphasize the role of some island characteristics that might be involved in determining the propensity of a PAIC system to promote diversification under a hypothesis of a species pump, such as island size, overall PAIC size, and topographic relief (including both bathymetry and island elevation). Using this powerful and accessible approach on other taxa and PAIC systems across the Caribbean, will provide a better perspective not only on the factors that contribute to divergence, but also on what determines whether such divergences are likely to be maintained, and therefore contribute to diversification across island systems.

ACKNOWLEDGMENTS

This work was funded by the Hubbell and Ammerman endowments at the University of Michigan, Museum of Zoology, and by NSF (DEB 1118815 to L.L.K.). This project was made possible by the availability of specimens collected as part of projects sponsored by NSF (including DEB 0715487 to L.L.K.) and in collaboration with Elen Oneal and Dan Otte. We are grateful to Raquel Marchan–Rivadeneira for lab work, Rob Massatti and Mark Christie for help with analyses and discussions, as well as to the Associate Editor and two anonymous reviewers for their constructive comments. Special thanks to Qixin He, who wrote the custom script used to subsample each population and calculate the SFS for analyses with fastsimcoal2. The authors declare no conflict of interest.

DATA ARCHIVING

Dryad: http://dx.doi.org/10.5061/dryad.23hs1.

LITERATURE CITED

Acevedo-Rodriguez, P. 1996. Flora of St. John, U.S. Virgin Islands. The New York Botanical Garden Press, Bronx, New York.

Allmon, W. D. 1992. A causal analysis of stages in allopatric speciation. Oxford Surv. Evol. Biol. 8:219–219.

Arnold, B., R. Corbett-Detig, D. Hartl, and K. Bomblies. 2013. RADseq underestimates diversity and introduces genealogical biases due to nonrandom haplotype sampling. Mol. Ecol. 22:3179–3190.

Brown, R., and A. Diesmos. 2009. Philippines, biology. Encyclopedia of Islands. California Univ. Press, Berkeley, pp. 723–732.

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 et al. 2013. Evolutionary processes of diversification in a Model Island Archipelago. Annu. Rev. Ecol. Syst. 44:411–435.

Camoin, G., L. Montaggioni, and C. Braithwaite. 2004. Late glacial to post glacial sea levels in the Western Indian Ocean. Mar. Geol. 206:119–146.

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.

Crispo, E., P. Bentzen, D. N. Reznick, M. T. Kinnison, and A. P. Hendry. 2006. The relative influence of natural selection and geography on gene flow in guppies. Mol. Ecol. 15:49–62.

Davalos, L. M., and A. L. Russell. 2012. Deglaciation explains bat extinction in the Caribbean. Ecol. Evol. 2:3045–3051.

Dynesius, M., and R. Jansson. 2014. Persistence of within-species lineages: a neglected control of speciation rates. Evolution 68:923–934.

Earl, D. A. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genet. Resour. 4:359–361.

Edelaar, P., D. Alonso, S. Lagerveld, J. Senar, and M. Björklund. 2012. Population differentiation and restricted gene flow in Spanish crossbills: not isolation-by-distance but isolation-by-ecology. J. Evol. Biol. 25:417–430.

Emerson, K. J., C. R. Merz, J. M. Catchen, P. A. Hohenlohe, W. A. Cresko, W. E. Bradshaw, and C. M. Holzapfel. 2010. Resolving postglacial phylogeography using high-throughput sequencing. Proc. Nat. Acad. Sci. USA 107:16196–16200.

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., S. P. Maher, and R. M. Brown. 2011. Species interactions during diversification and community assembly in an island radiation of shrews. PLoS One 6:e21885.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 14:2611–2620.

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.

Falk, B. G., and S. L. Perkins. 2013. Host specificity shapes population structure of pinworm parasites in Caribbean reptiles. Mol. Ecol. 22:4576–4590.

Futuyma, D. J. 1987. On the role of species in anagenesis. Am. Nat. 130:465–473.

Glor, R. E., M. E. Gifford, A. Larson, J. B. Losos, L. R. Schettino, A. R. Chamizo Lara, and T. R. Jackman. 2004. Partial island submergence and speciation in an adaptive radiation: a multilocus analysis of the Cuban green anoles. Proc. R. Soc. Lond. B Biol. Sci. 271:2257–2265.

Glor, R. E., J. B. Losos, and A. Larson. 2005. Out of Cuba: overwater dispersal and speciation among lizards in the *Anolis carolinensis* subgroup. Mol. Ecol. 14:2419–2432.

Gore, S. 2013. Anegada: an emergent Pleistocene Reef Island. Pp. 47–60 *in* C. R. C. Sheppard, ed. Coral reefs of the world. Springer, Netherlands.

Gorog, A. J., M. H. Sinaga, and M. D. Engstrom. 2004. Vicariance or dispersal? Historical biogeography of three Sunda shelf murine rodents (*Maxomys surifer, Leopoldamys sabanus and Maxomys whiteheadi*). Biol. J. Linn. Soc. 81:91–109.

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.

Guillot, G., F. Rousset, and L. Harmon. 2013. Dismantling the Mantel tests. Methods Ecol. Evol. 4:336–344.

Harter, A. V., K. A. Gardner, D. Falush, D. L. Lentz, R. A. Bye, and L. H. Rieseberg. 2004. Origin of extant domesticated sunflowers in eastern North America. Nature 430:201–205.

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 Quat. Res. Southeast Asia 9:127–144.

Heaney, L. R. 1986. Biogeography of mammals in SE Asia: estimates of rates of colonization, extinction and speciation. Biol. J. Linn. Soc. 28:127–165.

Heaney, L. R., J. S. Walsh, and A. T. Peterson. 2005. The roles of geological history and colonization abilities in genetic differentiation between mammalian populations in the Philippine archipelago. J. Biogeogr. 32:229–247.

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. 1996. Historical biogeography of West Indian vertebrates. Annu. Rev. Ecol. Syst. 27:163–196.

— S. B. 2006. Paleogeography of the Antilles and origin of West Indian terrestrial vertebrates. Annal. Missouri Bot. Garden 93:231–244.

Hedges, S. B., and M. P. Heinicke. 2007. Molecular phylogeny and biogeography of West Indian frogs of the genus *Leptodactylus* (Anura, Leptodactylidae). Mol. Phylogenet. Evol. 44:308–314.

Heinen, J. L., M. W. Coco, M. S. Marcuard, D. N. White, M. N. Peterson, R. A. Martin, and R. B. Langerhans. 2013. Environmental drivers of demographics, habitat use, and behavior during a post-Pleistocene radiation of Bahamas mosquitofish (*Gambusia hubbsi*). Evol. Ecol. 27:971–991.

Henderson, R. W., and R. Powell. 1999. West Indian herpetoecology. *In* Caribbean amphibians and reptiles. Academic Press, San Diego, California, pp. 223–268.

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. 2014. Unforeseen consequences of excluding missing data from next-generation sequences: simulation study of RAD sequences. Syst. Biol.

Hubbard, D. K., R. B. Burke, I. P. Gill, W. R. Ramirez, and C. Sherman. 2008. Coral-reef geology: Puerto Rico and the US Virgin islands. Pp. 263–302 *in* B.M. Riegl and R.E. Dodge eds. Coral Reefs of the USA. Springer, USA.

Iturralde-Vinent, M., and R. D. MacPhee. 1999. Paleogeography of the Caribbean region: implications for Cenozoic biogeography. Bull. AMNH 238:1–95

Jakobsson, M., and N. A. Rosenberg. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23:1801–1806.

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. Blaxter. 2009. Analysis of the genome sequences of three *Drosophila melanogaster* spontaneous mutation accumulation lines. Genome Res. 19:1195–1201.

Kierepka, E. M., and E. K. Latch. 2014. Performance of partial statistics in individual-based landscape genetics. Mol. Ecol. Res. 15:512–525. doi: 10.1111/1755-0998.12332.

Kisel, Y., and T. G. Barraclough. 2010. Speciation has a spatial scale that depends on levels of gene flow. Am. Nat. 175:316–334.

Langerhans, R. B., M. E. Gifford, and E. O. Joseph. 2007. Ecological speciation in Gambusia fishes. Evolution 61:2056–2074.

Lee, C. R., and T. Mitchell-Olds. 2011. Quantifying effects of environmental and geographical factors on patterns of genetic differentiation. Mol. Ecol. 20:4631–4642.

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., T. R. Jackman, A. Larson, K. D. Queiroz, and L. Rodriguez-Schettino. 1998. Contingency and determinism in replicated adaptive radiations of island lizards. Science 279:2115–2118.

Losos, J. B., and R. E. Ricklefs. 2009. Adaptation and diversification on islands. Nature 457:830–836.

Losos, J. B., and D. Schluter. 2000. Analysis of an evolutionary species-area relationship. Nature 408:847–850.

Martin, C. H., and L. C. Feinstein. 2014. Novel trophic niches drive variable progress towards ecological speciation within an adaptive radiation of pupfishes. Mol. Ecol. 23:1846–1862.

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.

Meirmans, P. G., and P. H. Van Tienderen. 2004. GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. Mol. Ecol. Notes 4:792–794.

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., and L. L. Knowles. 2013. Ecological selection as the cause and sexual differentiation as the consequence of species divergence? Proc. R. Soc. Lond. B Biol. Sci. 280:20122236.

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.

Otte, D., and D. Perez-Gelabert 2009. Crickets of the Caribbean. Publications on orthopteran diversity, The Orthopterists Society, Ann Arbor, MI, 792 p.

Parchman, T. L., Z. Gompert, J. Mudge, F. D. Schilkey, C. W. Benkman, and C. Buerkle. 2012. Genome-wide association genetics of an adaptive trait in lodgepole pine. Mol. Ecol. 21:2991–3005.

Parent, C. E., and B. J. Crespi. 2006. Sequential colonization and diversification of Galapagos endemic land snail genus *Bulimulus* (Gastropoda, Stylommatophora). Evolution 60:2311–2328.

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.

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.

Pregill, G. K., and S. L. Olson. 1981. Zoogeography of West Indian vertebrates in relation to Pleistocene climatic cycles. Annu. Rev. Ecol. Syst.:75–98.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945–959.

Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. Ferreira, D. Bender, J. Maller, P. Sklar, P. de Bakker, M. Daly et al. 2007. PLINK: a toolset for whole-genome association and population-based linkage analysis. Am. J. Hum. Genet. 81:559–575.

Reitzel, A., S. Herrera, M. Layden, M. Martindale, and T. Shank. 2013. Going where traditional markers have not gone before: utility of and promise for RAD sequencing in marine invertebrate phylogeography and population genomics. Mol. Ecol. 22:2953–2970.

Reynolds, R. G., M. L. Niemiller, S. B. Hedges, A. Dornburg, A. R. Puente-Rolon, and L. J. Revell. 2013. Molecular phylogeny and historical biogeography of West Indian boid snakes (*Chilabothrus*). Mol. Phylogenet. Evol. 68:461–470.

Rhymer, J. M., and D. Simberloff. 1996. Extinction by hybridization and introgression. Annu. Rev. Ecol. Syst.27:83–109.

Ricklefs, R. E., and E. Bermingham. 2004. History and the species-area relationship in Lesser Antillean birds. Am. Nat. 163:227–239.

Ricklefs, R. E., and E. Bermingham. 2007. The causes of evolutionary radiations in archipelagoes: passerine birds in the Lesser Antilles. Am. Nat. 169:285–297.

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 et al. 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.

Roberts, T. E. 2006. Multiple levels of allopatric divergence in the endemic Philippine fruit bat *Haplonycteris fischeri* (Pteropodidae). Biol. J. Linn. Soc. 88:329–349.

Rosenberg, N. A. 2004. DISTRUCT: a program for the graphical display of population structure. Mol. Ecol. Notes 4:137–138.

Ryan, P. G., P. Bloomer, C. L. Moloney, T. J. Grant, and W. Delport. 2007. Ecological speciation in South Atlantic island finches. Science 315:1420–1423.

Salvi, D., P. J. Schembri, A. Sciberras, and D. J. Harris. 2014. Evolutionary history of the Maltese wall lizard *Podarcis filfolensis*: insights on the 'Expansion-Contraction' model of Pleistocene biogeography. Mol. Ecol. 23:1167–1187.

Schluter, D., and G. L. Conte. 2009. Genetics and ecological speciation. Proc. Natl. Acad. Sci. USA 106(Suppl 1):9955–9962.

Seehausen, O. 2006. Conservation: losing biodiversity by reverse speciation. Curr. Biol. 16:R334–R337.

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. Phylogenet. Evol. 55:699–710.

Slatkin, M. 1977. Gene flow and genetic drift in a species subject to frequent local extinctions. Theor. Popul. Biol. 12:253–262.

Slatkin, M. 1985. Gene flow in natural populations. Annu. Rev. Ecol. Syst. 16:393–430.

Sly, N. D., A. K. Townsend, C. C. Rimmer, J. M. Townsend, S. C. Latta, and I. J. Lovette. 2011. Ancient islands and modern invasions: disparate phylogeographic histories among Hispaniola's endemic birds. Mol. Ecol. 20:5012–5024.

Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313.

Taylor, E., J. Boughman, M. Groenenboom, M. Sniatynski, D. Schluter, and J. Gow. 2006. Speciation in reverse: morphological and genetic evidence of the collapse of a three-spined stickleback (*Gasterosteus aculeatus*) species pair. Mol. Ecol. 15:343–355.

Team, R. C. 2013. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.

Thorpe, R. S., Y. Surget-Groba, and H. Johansson. 2008. The relative importance of ecology and geographic isolation for speciation in anoles. Philos. Trans. R. Soc. B-Biol. Sci. 363:3071–3081.

Thorpe, R. S., Y. Surget-Groba, and H. Johansson. 2010. Genetic tests for ecological and allopatric speciation in anoles on an island archipelago. PLoS Genet 6:e1000929.

Thorpe, R. S., Y. Surget-Groba, and H. Johansson. 2012. Quantitative traits and mode of speciation in Martinique anoles. Mol. Ecol. 21:5299–5308.

Wagner, C. E., I. Keller, S. Wittwer, O. M. Selz, S. Mwaiko, L. Greuter, A. Sivasundar, and O. Seehausen. 2013. Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and

relationships in the Lake Victoria cichlid adaptive radiation. Mol. Ecol. 22:787–798.

Wang, C., Z. A. Szpiech, J. H. Degnan, M. Jakobsson, T. J. Pemberton, J. A. Hardy, A. B. Singleton, and N. A. Rosenberg. 2010. Comparing spatial maps of human population-genetic variation using Procrustes analysis. Stat. Appl. Genet. Mol. Biol. 9.

Wang, C., S. Zöllner, and N. A. Rosenberg. 2012. A quantitative comparison of the similarity between genes and geography in worldwide human populations. PLoS Genet. 8:e1002886.

Wang, I. J., R. E. Glor, and J. B. Losos. 2013. Quantifying the roles of ecology and geography in spatial genetic divergence. Ecol. Lett. 16:175–182.

Whetten, J. T. 1966. Geology of St. Croix, US Virgin Islands. Geol. Soc. Am. Memoirs 98:177–240.

Williams, E. E. 1983. Ecomorphs, faunas, island size, and diverse end points in island radiations of Anolis. Pp. 326–370 *in* R. B. Huey, E. R. Pianka, and T. W. Schoener, eds. Lizard ecology: studies of a model organism. Harvard Univ. Press, Cambridge, MA.

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.

Associate Editor: Dr. Frank Burbrink
Handling Editor: Dr. Thomas Lenormand

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Table S1. Processing of illumina reads using the full (100 bp) and trimmed (50 bp) reads.
Table S2. Population genetic statistics for ten *A. sanctaecrucis* populations, based on the 5558 loci dataset and calculated only for polymorphic positions.
Table S3. Pairwise *FST*-values as calculated by STACKS (above the diagonal) and GENODIVE (below the diagonal).
Table S4. Results of divergence time estimation with FASTSIMCOAL2 under three alternative models.
Figure S1. Number of reads per individual before and after the quality filtering step, individuals are ordered alphabetically by population code.
Figure S2. Average numbers of (A) polymorphic loci and (B) SNPs per individual, for each population, when the 100 bp and 50 bp reads were analyzed using STACKS.
Figure S3. Comparison of observed (*Hobs*; left box) and expected heterozygosity (*Hexp*; right box) per population averaged across loci, for 10 random samples of six individuals (i.e., the smallest sample of individuals across populations) based on 5558 SNPs (see Fig. 1 for population color codes).
Figure S4. Procrustes-transformed PCA plot of genetic variation for the analysis without the St. Croix and Anegada populations (procrustes similarity score, $t0 = 0.772$).
Figure S5. Coalescent-based tree of 136 *A. sanctaecrucis* and 12 outgroup individuals based on 82,440 SNPs analyzed using SVDQUARTETS.
Figure S6. Maximum likelihood tree based on a concatenated matrix of 5558 SNPs analyzed in RAXML v. 8, using the ASC_GTRGAMMA model, which corrects the likelihood calculations for ascertainment bias when analyzing SNP matrices not containing invariant sites.
Figure S7. Map of the Virgin Island PAIC, with bathymetric data.