



# Insights into the introduction history and population genetic dynamics of the Argentine black-and-white tegu (*Salvator merianae*) in Florida

Jared P. Wood<sup>1</sup> · Stephanie Dowell Beer<sup>2</sup> · Todd S. Campbell<sup>3</sup> · Robert B. Page<sup>4</sup>

Received: 5 April 2018 / Accepted: 15 September 2018 / Published online: 22 September 2018  
© Springer Nature Switzerland AG 2018

## Abstract

Invasive species are a major threat to global biodiversity. The US state of Florida is especially susceptible to the spread of exotic reptiles due to its subtropical climate, disturbed habitats, and robust pet trade. The Argentine black-and-white tegu (*Salvator merianae*) is a large, omnivorous lizard currently established in two different regions of Southern Florida. These two populations pose potential threats to sensitive ground nesting species such as gopher tortoises, American crocodiles, and migratory birds. At present, the introduction histories of these populations and the degree to which they are connected by gene flow are largely unknown. To address these issues, we genotyped *S. merianae* from Hillsborough and Miami-Dade Counties at ten microsatellite loci to assess intrapopulation genetic diversity, the degree of gene flow between populations, and compare the plausibilities of several potential introduction scenarios. Our results indicate that both populations have low genetic diversity [mean number of effective alleles across loci in both populations = 2.09 and are highly differentiated from each other ( $G_{ST}=0.170$ ;  $G''_{ST}=0.545$ )]. In addition, our results suggest that these populations underwent a bottleneck event prior to their divergence. We discuss what our results suggest about the histories of Florida's invasive tegu populations, as well as how they inform ongoing management strategies.

**Keywords** Florida · Founder effect · Genetic structure · Invasive species · Argentine black-and-white tegu

---

Jared P. Wood and Stephanie Dowell Beer contributed equally.

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s10709-018-0040-0>) contains supplementary material, which is available to authorized users.

---

✉ Jared P. Wood  
j.wood@swau.edu

✉ Robert B. Page  
Robert.Page@tamusa.edu

- <sup>1</sup> Department of Biological Sciences, Southwestern Adventist University, 100 W. Hillcrest, Keene, TX 76059, USA
- <sup>2</sup> US Fish and Wildlife Service, Northeast Fishery Center, 227 Washington Ave., Lamar, PA 16848, USA
- <sup>3</sup> Department of Biology, University of Tampa, 401 W. Kennedy Blvd., Tampa, FL 33606, USA
- <sup>4</sup> Department of Science and Mathematics, Texas A&M University-San Antonio, One University Way, San Antonio, TX 78224, USA

## Introduction

One of the greatest threats to global biodiversity is the spread of invasive species (Clavero and Garcí a-Berthou 2005; Wake and Vredenburg 2008; Wilcove et al. 1998). Invasive species can negatively impact native species either directly through competition, predation, and disease or indirectly through alteration of ecosystem structure and function (Klug et al. 2015; Mooney and Cleland 2001). The spread of invasive species has accelerated over the last few centuries due to increases in international trade and transport (Abdelkrim et al. 2005; Di Castri 1989; Mack et al. 2000), and port-rich coastal regions have frequently served as points of entry. Florida is especially susceptible to the proliferation of invasive reptiles largely due to three factors: (1) a subtropical climate, (2) the presence of altered habitats (ponds, canals, levees) that provide suitable migration corridors for invasive species, and (3) an extensive exotic pet industry that imports and produces potentially invasive organisms (Mazzotti et al. 2015; Smith 2006). Consequently,

in Florida, there are more nonnative lizards than native lizard species (Krysko et al. 2011; Pernas et al. 2012).

One of the nonnative species that is of particular concern is the Argentine black-and-white tegu (*Salvator merianae*; Klug et al. 2015). *Salvator merianae* is a large, fecund lizard (Fitzgerald 1994), reaching lengths up to 145 cm total length (TL) and weighing up to 8 kg, with a broad, omnivorous diet that consists of vegetation, fruit, seeds, snails, arthropods, fish, birds, bird eggs, small mammals, amphibians, reptiles, reptile eggs, and carrion (Duarte Valera and Cabrera 2000; Galetti et al. 2009; Kiefer and Sazima 2002; Mocolli and Yanosky 1994). Due to *S. merianae*'s propensity for depredating nests, this species poses a direct threat to Florida's sensitive, ground-nesting species such as American crocodiles (*Crocodylus acutus*), Eastern indigo snakes (*Drymarchon couperi*), Cape Sable seaside sparrow (*Ammodramus maritimus mirabilis*), and gopher tortoises (*Gopherus polyphemus*; Mazzotti et al. 2015). *Salvator merianae* is native to southeastern Brazil, Uruguay, eastern Paraguay, and northern Argentina (Luxmoore et al. 1988). Within their native range, *S. merianae* occupy open habitats such as forest clearings, secondary forests, and other disturbed areas across a broad range of tropical, subtropical, and temperate climates (Cardozo et al. 2012; Chamut et al. 2012; Embert et al. 2010; Fitzgerald 1994; Winck and Cechin 2008). *Salvator merianae* also exhibits dormancy in response to winter temperatures and periods of drought (Abe 1983). Based on these distributional and ecological characteristics, Lanfrie et al. (2013) suggested that *S. merianae* could spread as far north as the US state of West Virginia.

Preventing the spread of harmful species, such as *S. merianae*, is necessary for effective management planning. However, the control of invasive species is often hindered by a lack of information about the histories and origins of the populations in question and the level of connectivity between groups of individuals (Rollins et al. 2009). It is generally assumed that isolated populations are easier to eradicate than populations that are connected by migration and gene flow because connected populations may require simultaneous eradication to prevent recolonization by migrants from neighboring areas (Abdelkrim et al. 2005; Rollins et al. 2009).

*Salvator merianae* was first observed in Hillsborough County, Florida on the Balm Boyette Nature Preserve (Enge 2007). Purportedly, introduction of *S. merianae* to Hillsborough County occurred via an exotic pet dealer that illegally released specimens imported with broken tails or other defects that diminished their market value (Enge 2007). In addition to the Hillsborough population, there is also a self-perpetuating *S. merianae* population approximately 300 km away in Miami-Dade County near Florida City (Pernas et al. 2012). The source of this population is believed to be individuals that escaped from a nearby property operated by an

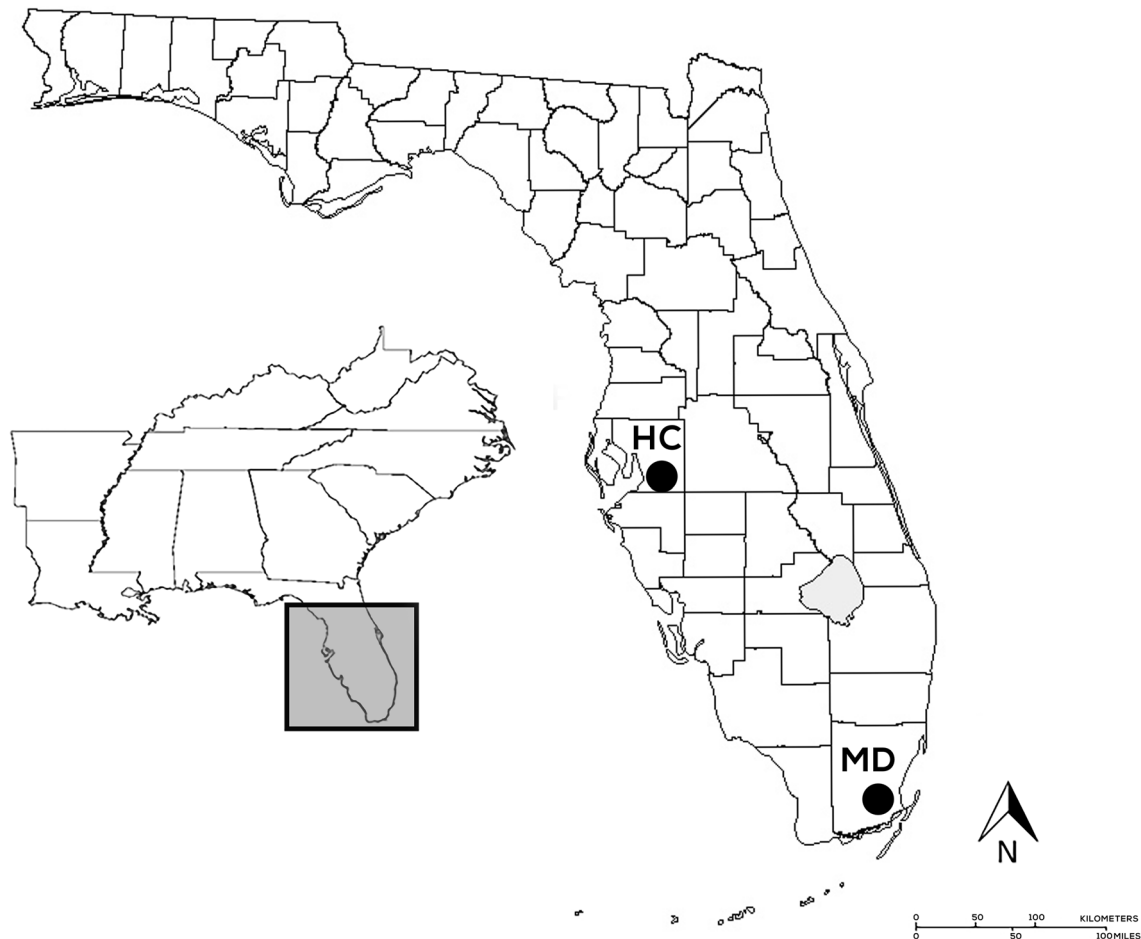
exotic pet importer (Krysko et al. 2011). Both introductions supposedly occurred around 2000, but established breeding populations were not documented in Hillsborough and Miami-Dade Counties until 2006 and 2010, respectively (Enge 2007; Krysko et al. 2011). Importation records from the late 1990s and early 2000s report that live *S. merianae* (formerly known as *Tupinambis merianae*) specimens were being imported from across their native range (<http://trade.cites.org/>). Therefore, the exact native source of the captive populations existing in the United States at that time is unknown. At present, nothing is known about the population genetic dynamics of *S. merianae* populations in Florida. However, surveys of neutral molecular variation could provide insight into the invasion histories of these populations and the extent to which these populations are connected by gene flow. Such information would likely prove useful to wildlife managers and may enable more informed decisions about how to best manage these exotic populations. To this end, we used microsatellite markers to examine genetic diversity within populations, genetic structure between populations, and possible introduction scenarios in Florida's documented *S. merianae* populations.

## Materials and methods

### Field sites, sampling, and tissue collection

*Salvator merianae* specimens were collected from Hillsborough and Miami-Dade Counties, Florida (Fig. 1). In Hillsborough County, *S. merianae* are primarily found in ruderal habitats near Balm Boyette Scrub Preserve located between the cities of Riverview and Lithia. At the time of this study, 38 specimens had been collected from this locale—all of which were used in this study. These samples were collected between 2012 and 2013 by one of us (TSC) from a 43.5 km<sup>2</sup> area centered around approximately 27°47'55"N, 82°11'56"W.

In Miami-Dade County, *S. merianae* are primarily found in the southeastern portion of the County near Florida City (25°23'02"N, 80°30'44"W). Ninety-three tissue samples collected from different specimens captured in Miami-Dade County in 2011 were donated to our study. To keep sample sizes approximately even between locations, we randomly selected 40 of these individuals for inclusion in this study. *Salvator merianae* specimens in Miami-Dade County have been primarily removed from disturbed areas such as ditches, canal levees, and historical wetlands that are comprised of late successional grasslands that are being replaced by shrubs and grasslands (Klug et al. 2015). The *S. merianae* specimens from Miami-Dade County that were used in this study were captured between 2009 and 2011. The Florida Wildlife Commission collected these samples.



**Fig. 1** Map showing the location of the sampling sites (*HC* Hillsborough County; *MD* Miami-Dade County) of *S. merianae* in Southern Florida and the position of Florida within the Southeastern US

The exact ages of the specimens analyzed in this study are unknown. *Salvator merianae* have a potential lifespan of up to 20 years (Brito et al. 2001), with females typically reaching reproductive age around three years of age (Fitzgerald et al. 1993). Therefore, it is possible that our samples represent multiple generations, and could include some of the first introduced individuals.

### DNA isolation and PCR-based genotyping

In 2011, we isolated Genomic DNA from muscle and liver from *S. merianae* specimens collected in Miami-Dade County using the Wizard Genomic DNA Purification Kit (Promega) according to the manufacturer's instructions. In 2012, we used the same procedure and tissues to isolate DNA from *S. merianae* specimens collected in Hillsborough County. These samples were genotyped at 14 microsatellite loci developed using *S. merianae* samples from the Miami-Dade population (Wood et al. 2015). All PCRs had a final volume of 25  $\mu$ l and contained 2  $\mu$ l of template (DNA

concentration between 10 and 100 ng/ $\mu$ l), 1 $\times$  buffer, 1.5 mM MgCl<sub>2</sub>, 0.2 mM of each dNTP, 0.8  $\mu$ M of non-M13(-21)-twinned primer, 0.8  $\mu$ M of 6-FAM labeled M13(-21) primer, 0.2  $\mu$ M of M13(-21)-twinned primer, and 0.625 units of GoTaq polymerase (Promega). Reaction conditions were as follows: 2 min at 94  $^{\circ}$ C followed by 25 cycles of 94  $^{\circ}$ C for 30 s, 30 s at 62  $^{\circ}$ C decreasing by 0.3  $^{\circ}$ C per cycle, and 72  $^{\circ}$ C for 40 s, followed by eight cycles of 94  $^{\circ}$ C for 30 s, 53  $^{\circ}$ C for 30 s, and 72  $^{\circ}$ C for 40 s, followed by a final cleanup step of 30 min at 72  $^{\circ}$ C. Agarose gel electrophoresis (2% gels) was used to confirm successful amplification, and fragment analysis was performed at the Arizona State University DNA Lab using an Applied Biosystems 3730. PEAK SCANNER 1.0 (Applied Biosystems) was used to manually score all loci. In order to identify breaks in the amplicon sizes, allelic bins were determined by graphically examining the rank-ordered fragment size distributions of each locus (Guichoux et al. 2011). Finally, Microsoft EXCEL was used to bin the data from each locus into discrete classes that were defined by each allele's empirically determined size range.

## Summary statistics and quality control

We used MICRO-CHECKER 2.2.3 (Van Oosterhout et al. 2004) to examine each locus for evidence of null alleles, large allele dropout, and scoring errors. GENALEX 6.5 (Peakall and Smouse 2012) was used to calculate summary statistics including number of alleles, effective number of alleles, observed heterozygosity, and expected heterozygosity. Finally, GENEPOP 4.3 (Rousset 2008) was used to test for departures from Hardy–Weinberg proportions, genotypic equilibrium and to calculate the Weir and Cockerham (1984) estimator of  $F_{IS}$ .

## Assessment of population structure

Several approaches were used to determine the degree of genetic differentiation between the *S. merianae* populations in Hillsborough and Miami-Dade Counties. GENALEX 6.5 (Peakall and Smouse 2012) was used to calculate  $G_{ST}$  values based on Nei and Chesser's (1983) unbiased estimators of  $H_S$  (i.e., the Hardy–Weinberg expected heterozygosity averaged across subpopulations) and  $H_T$  (i.e., the Hardy–Weinberg expected heterozygosity in the total population ignoring subdivision). We also used GENALEX to calculate Hedrick's further standardized  $G_{ST}$  ( $G''_{ST}$ ), which is a modified version of Hedrick's  $G'_{ST}$  (a standardized G-statistic that is formulated to equal one when populations have non-overlapping allele sets regardless of the level of allelic richness) that corrects for the tendency of  $G'_{ST}$  to underestimate the true degree of subdivision when only a small number of populations have been sampled (Meirmans and Hedrick 2011). All resampling tests conducted in GENALEX were based on 9999 permutations. We also performed an AMOVA that partitioned genetic variation among populations, among individuals within populations, and within individuals using ARLEQUIN 3.5.1.2 (Excoffier and Lischer 2010).

STRUCTURE 2.3.4 (Pritchard et al. 2000; Falush et al. 2003) was used to estimate the number of genetic clusters ( $K$ ) and to assign individuals to clusters (i.e., clusters). We used the correlated allele frequencies model to allow for the possibility that both populations originated from a common source and allowed for the possibility of admixture. Because we had no a-priori hypotheses about between or within region subdivision, we examined a range of  $K$ -values ( $K = 1$ –6) and inspected the mean  $\ln P(D)$  from replicate runs at each value of  $K$  to assess whether the range of  $K$ -values we considered is likely to contain the optimal value of  $K$  (see results for interpretation). When using STRUCTURE, we conducted 10 replicate runs for each value of  $K$  with a burn-in period of 500,000, followed by 500,000 MCMC steps. STRUCTURE HARVESTER (Earl and vonHoldt 2012) was used to generate mean  $\ln P(D) \pm SD$  and  $\Delta K$  (Evanno et al. 2005) plots, and convergence was

assessed by examining the consistency of parameter estimates and likelihood values across replicate runs. CLUMPP (Jakobsson and Rosenberg 2007) was used to align cluster assignments across replicate runs and STRUCTURE PLOT (Ramasamy et al. 2014) was used to visualize and interpret the CLUMPP output. We also used STRUCTURE 2.3.4, STRUCTURE HARVESTER, CLUMPP, and STRUCTURE PLOT to investigate whether there is evidence of substructure within the Hillsborough and Miami-Dade populations respectively. When performing analyses in STRUCTURE that considered the Hillsborough and Miami-Dade datasets separately, model options, burn-in periods, MCMC parameters, number of replicate runs, and the range of  $K$  values investigated were as described above for our 'global' STRUCTURE analysis.

To further examine genetic differentiation in our sample, we also used Bayesian Analysis of Population Structure (BAPS) v5.3 (Corander et al. 2008) to infer  $K$ , assign individuals to clusters, and perform an admixture analysis. When clustering individuals with BAPS v5.3, the user provides an upper limit on  $K$  and the program identifies the optimal partition of the data assuming a uniform prior distribution in the clustering space between  $K = 1$  and the user-defined upper limit. The BAPS v5.3 manual ([http://www.helsinki.fi/bsg/software/BAPS/macSnow/BAPS5\\_3manual.pdf](http://www.helsinki.fi/bsg/software/BAPS/macSnow/BAPS5_3manual.pdf)) recommends conducting runs based on several user-defined maximum  $K$ -values (including some substantially bigger than the expected optimal value) in order to aid in the detection of outlier genotypes and to reduce the risk of finding only a local mode. In addition, the BAPS v5.3 manual recommends conducting multiple runs for each maximum  $K$ -value to assess whether the stochastic nature of BAPS' optimization algorithm results in different solutions from replicate runs of the same upper limit. In keeping with these recommendations, we ran 20 replicates of maximum  $K = 5$ , maximum  $K = 10$ , and maximum  $K = 15$  (a total of 60 runs). We then used the output from this clustering procedure to perform an admixture analysis in which we discarded clusters with fewer than 5 individuals, used 100 iterations to estimate individual admixture coefficients, used 200 reference individuals from each population, and used 20 iterations to estimate admixture coefficients for reference individuals.

Because introduced populations may not exhibit Hardy–Weinberg or linkage equilibrium, the major assumptions of Bayesian clustering approaches, such as STRUCTURE and BAPS, it is important to examine genetic variation among populations using alternate approaches. Therefore, we performed a principal component analysis (PCA) on the raw genotypic data for the whole data set as well as both respective populations using the *gstudio* package (Dyer 2012) for R 3.1 (R Core Team 2014). Because PCA ordines individuals along axes that summarize variation irrespective of source (i.e., differences between vs. within

populations), we also performed a discriminant analysis of principal components (DAPC), as this approach ordines samples along axes that maximize separation between groups and is therefore better suited than PCA for explicitly examining population structure (Jombart et al. 2010). We used the *adegenet* package (Jombart 2008) for R to infer K via K-means clustering and to conduct a DAPC. When performing K-means clustering, all principal components were retained and the Bayesian information criterion (BIC) was used to determine the most appropriate value of K. DAPC was then performed using the cluster memberships obtained from the K-means algorithm as prior group assignments. Before performing DAPC, the optimal number of principal components to retain in the discriminant analysis was investigated using a cross validation procedure that was scaled to the largest proportion of the sample size that ensured all clusters would be represented in both the training and validation sets (Jombart and Collins 2015).

To assess the possibility that kinship was driving the intraregional structuring observed from the DAPC analysis (see results below), we used the *related* package for R (Pew et al. 2015). In these analyses, resampling without replacement was used to randomly assign individuals from a given geographic region to groups of the same size as the DAPC clusters identified for that region. During each of 1000 rounds of randomization, the mean relatedness of individuals within pseudoclusters was calculated and these values were used to estimate the distribution of mean relatedness within randomly generated groups of the same size as our DAPC clusters. All calculations performed in *related* were based on the Queller and Goodnight (1989) estimator.

### Assessment of gene flow

We assessed the degree of recent gene flow between the Hillsborough and Miami-Dade populations with BAYESASS 1.3 (Wilson and Rannala 2003). This method infers pairwise migration rates during recent generations by utilizing a coalescent-based approach. We performed  $10^8$  iterations, sampling every 2000 iterations, with a burn-in of  $10^7$ . To determine if the runs had reached convergence, we plotted likelihood scores over time and examined the consistency of results across independent runs.

In addition, we used GENECLASS2 (Piry et al. 2004) to perform assignment tests via Paetkau's et al. (1995) frequency-based criterion. We used a default frequency of 0.01 for missing alleles and the Monte-Carlo resampling method described by Paetkau's et al. (2004). Probability computations were based on 10,000 simulated individuals, and the type I error rate was 0.01. GENECLASS2 and Paetkau et al. (1995) frequency-based criterion were also used to test for the presence of first-generation migrants. Since the Hillsborough and Miami-Dade populations represent the only known

*S. merianae* populations in Florida, we used the 'L<sub>home</sub>/L<sub>max</sub>' test statistic because it is most appropriate when all source populations have been sampled (Piry et al. 2004).

### Effective population size and demographic history

To better understand the breeding population size of *S. merianae*, in Florida, we estimated the effective population size ( $N_e$ ) with NeESTIMATOR 2.0 (Do et al. 2014) using the linkage disequilibrium (LD) method, which is based on the frequent occurrence of non-random associations of alleles across independent loci in small populations (Waples and Do 2008). The minimum allele frequency was set to 0.02 to avoid an upward bias caused by rare alleles (Waples and Do 2010) and the 95% confidence interval was calculated using the jackknife method implemented in the program.

We tested for evidence of recent population declines using the program BOTTLENECK 1.2.02 (Piry et al. 1999). This method assesses deviations from expected heterozygosity, indicative of population decline (heterozygote excess) and expansion (heterozygote deficiency). In addition, BOTTLENECK examines the distribution of allele frequencies, which are typically skewed following bottleneck events (Piry et al. 1999). We tested for deviations under the two-phase model (TPM) with 70% stepwise mutation model (SMM) and a variance of 30—the recommended value for microsatellite loci (Di Rienzo et al. 1994). We performed 1,000,000 iterations and tested for significance with the Wilcoxon signed-rank test, recommended for data sets containing fewer than 20 loci (Piry et al. 1999), and a mode-shift test, both of which are implemented in BOTTLENECK. Additionally, we used the program KGTESTS (Bilgin 2007) to test for genetic signatures of population expansion via a within-locus *k* test and an interlocus *g* test. The *k* test is based on the observation that the typical allele distribution at a locus has several modes in a constant-sized population due to a small number of historic splitting events in the genealogy (Reich and Goldstein 1998; Reich et al. 1999). Conversely, an expanding population shows a more peaked allele distribution with a single mode due to many recent splitting events occurring near the time of the expansion (Reich and Goldstein 1998; Reich et al. 1999). Furthermore, expanding populations typically show lower levels of variance in the widths of allele distributions across loci than do constant-sized populations (Reich and Goldstein 1998). Therefore, the *g* test measures the variance in the allele distribution at each locus as well as the variance of these variances across loci to determine if a population shows evidence of expansion (Reich et al. 1999). Finally, we calculated *M*-ratios (Garza and Williamson 2001) in EXCEL using the output from GENALEX. *M*-ratios are defined as the ratio of *k* (total number of alleles) to *r* (overall range in allele size in number of repeat units). These ratios are indicative of recent bottlenecks when they

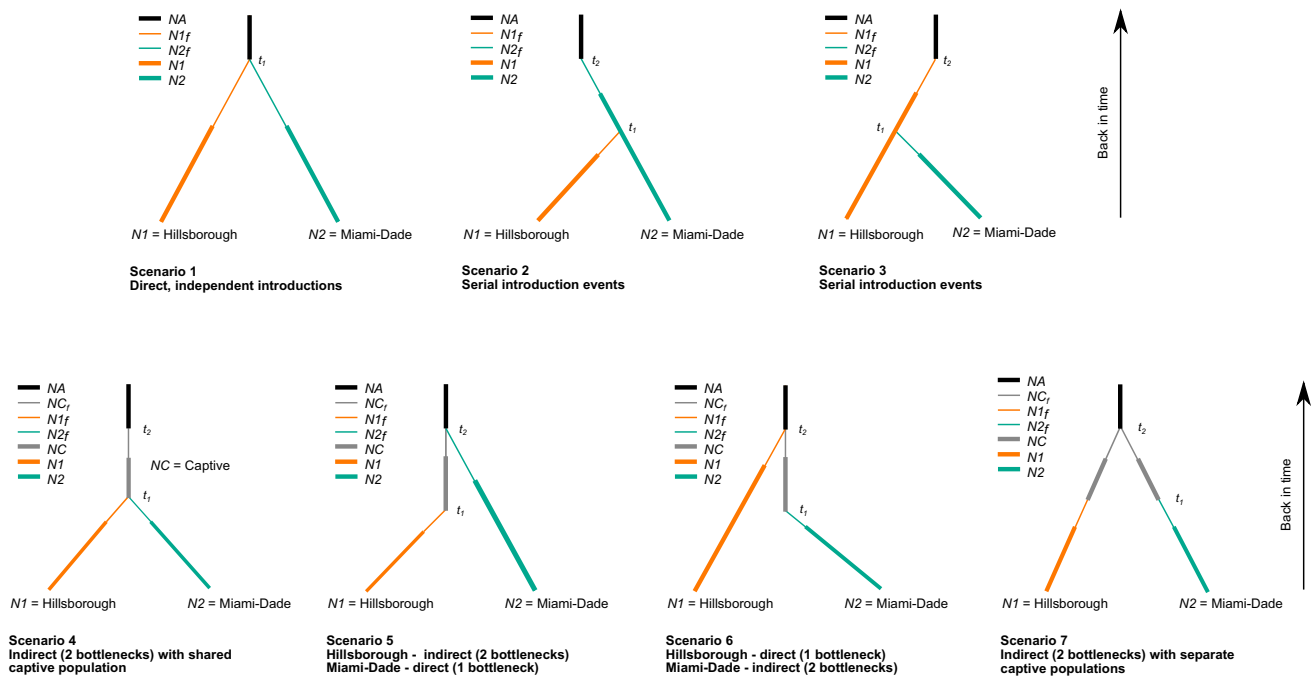
are less than a critical value ( $M_c$ ), determined via simulation under a two-phase mutation model as described by Garza and Williamson (2001). We used the command line executable *Critical\_M* (Garza and Williamson 2001) to obtain  $M_c$  values for the Hillsborough and Miami-Dade populations respectively. These  $M_c$  values were based on the minimum sample size obtained across loci for each respective geographic region and were determined from 10,000 replicate simulations that assumed 90% single-step mutations with a mean of 3.5 repeat units for non-single-step mutations as recommended by Garza and Williamson (2001). In addition, *Critical\_M* requires a user-specified estimate of  $\theta$  in the pre-bottleneck population, where  $\theta = 4N_e\mu$ , and  $\mu$  is the mutation rate. From DIYABC simulations based on our selected scenario (see methods and results), we obtained an  $N_e$  estimate of 9,050 for the ancestral South American population. Thus, assuming mutation rates at the loci we examined are similar to the central tendency for dinucleotide loci ( $5.6 \times 10^{-4}$ ; Bhargava and Fuentes 2010),  $\theta$  in the ancestral South American population is approximately 20.27. As such, we used this value of  $\theta$  when parameterizing *Critical\_M*.

### Tests of alternative introduction scenarios

To infer the introduction history of Florida's *S. merianae* populations, we tested seven competing scenarios with

DIYABC 2.1.0 (Cornuet et al. 2014). Given that this species is prevalent in the pet trade and can be bred in captivity (Hall 1978; Bartlett and Bartlett 1996), we wanted to distinguish between a direct introduction route (defined here as a non-native introduction originating from a wild source population) versus an indirect route (a non-native introduction originating from a captive population). To do this, we included scenarios in which the populations have undergone different numbers of bottleneck events. A direct introduction route is characterized by a single bottleneck event, where the founding population is typically comprised of a few individuals. An indirect route of introduction would encompass two or more bottleneck events, with the first occurring during the formation of the captive population and subsequent bottlenecks associated with the non-native introduction. We acknowledge, however, that multiple successive bottlenecks may be reflective of alternative scenarios, which we discuss below.

The first scenario assumes the introduced *S. merianae* populations were formed from two independent introductions from an ancestral (source) population (direct introduction; Fig. 2). Scenarios 2–3 test hypotheses involving a serial introduction pathway, where the second introduced population originated from the first introduced population, rather than separately from the native range. Scenario 4 describes an indirect introduction where the two



**Fig. 2** Graphical representation of the competing introduction scenarios for *S. merianae* compared with the software DIYABC. In each scenario, thin lines represent bottlenecked populations following introduction events, while thick lines represent the current or post-bottleneck effective population size. NA ancestral (source) effective

population size;  $N_1$  effective population size for the Hillsborough population;  $N_2$  effective population size for the Miami-Dade population;  $N_C$  effective population size for unsampled (captive) population;  $N_f$  the effective number of founding individuals;  $t$  time in generations

*S. merianae* populations originated from the same captive source, i.e. divergence of the populations occurred after the first bottleneck event. Scenarios 5–6 hypothesize that one population originated from a wild source (direct), while the other population originated via a captive source (indirect). Lastly, Scenario 7 describes an indirect introduction where the two *S. merianae* populations originated from different captive populations, i.e. divergence of the populations occurred prior to the first bottleneck event. Because the population structure and gene flow analyses suggested little admixture between the two introduced *S. merianae* populations (see “Results”), we did not test scenarios involving admixture, and instead opted to limit our analysis to a few simple introduction histories that reflect common pathways.

For all analyses, we used broad, uniform prior distributions defined as follows:  $5 < N < 10,000$ ;  $5 < NC < 10,000$ ;  $10 < NA < 50,000$ ;  $1 < N_f < 100$ ;  $1 < db < 50$ ;  $1 < t_1 < t_2 < 100$ ; where ‘*N*’ denotes the current effective population size, ‘*NA*’ denotes the ancestral (wild source) effective population size, ‘*NC*’ denotes the unsampled (captive) effective population size, ‘*N<sub>f</sub>*’ denotes the effective number of founding individuals, ‘*db*’ denotes the bottleneck duration in generations, and ‘*t*’ the time in generations. Priors for the microsatellite mutation model were set to default values, including the Generalized Stepwise Mutation model (Estoup et al. 2002), and a uniform prior distribution for both the mean mutation rate ( $1E-4$  to  $1E-3$ ) and the geometric distribution ( $1E-1$  to  $3E-1$ ). Summary statistics included the mean number of alleles, mean genic diversity, and mean size variance for both the one-sample and two-sample statistics. Additionally, we used the mean Garza-Williamson *M* index (one-sample statistic) as well as pairwise  $F_{ST}$  values and the mean classification index (two-sample statistics). We simulated one million datasets for each scenario, for a total of seven-million, and evaluated the scenario and parameter priors by performing a PCA, as implemented in DIYABC.

We determined the optimal scenario based on posterior probabilities using the logistic regression analysis implemented in DIYABC and the 1% closest simulated data sets. For comparison, we performed a pre-processing step (Linear Discriminant Analysis) on the summary statistics prior to computing the logistic regression. To further evaluate the power of our ABC method in distinguishing among the various competing scenarios, we analyzed 100 simulated pseudo-observed data sets (pods) for each scenario, using parameter values drawn from the same prior distribution as our previous analyses. The relative posterior probabilities of each scenario, estimated for each pod, were then used to calculate the likelihood of excluding the focal scenario when it is actually the true scenario (type I error rate), as well as the likelihood of selecting the focal scenario when it is not the true scenario (type II error rate).

## Results

### Summary statistics and quality control

All 14 of the loci we initially examined were variable when considering both populations. However, upon performing Holm’s (1979) correction for multiple testing via treating the tests associated with each population as a family of tests, we detected significant departures from Hardy–Weinberg proportions for *Teg4*, *Teg14*, and *Teg17* in the Hillsborough population and *Teg4*, *Teg5*, *Teg17*, and *Teg19* in the Miami-Dade population. In addition, MICRO-CHECKER detected evidence of null alleles for *Teg4*, *Teg5*, and *Teg19*, and upon performing Holm’s (1979) correction for multiple testing, there was evidence for genotypic disequilibrium between *Teg14*–*Teg19* in the Miami-Dade population. Due to these quality control issues, we excluded the loci that exhibited significant departures from Hardy–Weinberg equilibrium in both populations (*Teg4* and *Teg17*) and the loci that exhibited evidence of null alleles (*Teg4*, *Teg5* and *Teg19*) from all further analyses. By excluding *Teg19*, we were also able to address the non-independence between *Teg14* and *Teg19* in the Miami-Dade population. Summary statistics for the remaining ten loci that we used for all inferential analyses are given in Table 1.

### Assessment of population structure

Locus-specific  $G_{ST}$  estimates ranged from 0.028 to 0.312 and were statistically significant (maximum  $P = 0.011$ ). Locus-specific estimates  $G''_{ST}$  were also statistically significant (maximum  $P = 0.009$ ), with values ranging from 0.119 to 0.893. The global  $G_{ST}$  estimate that resulted from averaging information across all loci was 0.170 ( $SE = 0.025$ ,  $P = 0.0001$ ). Similarly, the global estimate for  $G''_{ST}$  was 0.545 ( $SE = 0.060$ ,  $P = 0.0001$ ). The AMOVA results computed in ARLEQUIN are also indicative of a high degree of genetic differentiation between the Hillsborough and Miami-Dade populations (Table 2) and suggested moderate heterozygote excess (i.e., produced a negative  $F_{IS}$  estimate).

In the PCA generated from the raw genotypic data, the first two principal components accounted for 35.57% of the overall genetic variation (Fig. 3) and separated the Hillsborough and Miami-Dade populations into distinct clusters, with only two intermediate individuals (individual 24 from Hillsborough and individual 42 from Miami-Dade). One tegu (individual 37) sampled in the Hillsborough population showed a large discrepancy in principal component 2 and did not cluster with the remaining individuals.

**Table 1** Summary statistics and diversity estimates for the ten loci that were used for genotyping

Locus/Pop.	<i>N</i>	<i>k</i>	$H_O$	$H_E$	$F_{IS}$	No. effective alleles	No. private alleles	<i>M</i>
Hillsborough								
<i>Teg1</i>	27	4	0.74	0.66	−0.10	2.96	0	0.80
<i>Teg2</i>	34	3	0.62	0.49	−0.25	1.96	1	0.60
<i>Teg6</i>	33	2	0.55	0.49	−0.10	1.96	1	0.67
<i>Teg7</i>	32	4	0.53	0.61	0.15	2.58	0	0.57
<i>Teg9</i>	30	3	0.63	0.64	0.03	2.78	1	0.50
<i>Teg10</i>	31	3	0.61	0.66	0.09	2.98	0	0.16
<i>Teg12</i>	33	3	0.61	0.63	0.05	2.71	1	0.75
<i>Teg13</i>	35	2	0.63	0.50	−0.24	2.00	1	0.29
<i>Teg14</i>	34	5	0.68	0.59	−0.13*	2.45	2	0.50
<i>Teg20</i>	27	4	0.67	0.63	−0.04	2.71	2	1.00
Pop. Mean	31.60	3.30	0.63	0.59	−0.05	2.51	0.90	0.58
Pop. SEM	0.90	0.30	0.02	0.02	0.04	0.13	0.23	0.08
Miami-Dade								
<i>Teg1</i>	40	4	0.63	0.56	−0.11	2.26	0	0.80
<i>Teg2</i>	40	2	0.03	0.02	N/A	1.03	0	0.40
<i>Teg6</i>	40	2	0.25	0.22	−0.13	1.28	1	1.00
<i>Teg7</i>	40	4	0.33	0.28	−0.14	1.39	0	0.57
<i>Teg9</i>	40	2	0.08	0.12	0.37	1.13	0	1.00
<i>Teg10</i>	39	4	0.59	0.53	−0.11	2.11	1	0.21
<i>Teg12</i>	38	2	0.08	0.08	−0.03	1.08	0	1.00
<i>Teg13</i>	38	2	0.47	0.45	−0.04	1.82	1	0.25
<i>Teg14</i>	40	4	0.68	0.64	−0.04	2.79	1	0.40
<i>Teg20</i>	38	2	0.58	0.43	−0.33	1.76	0	1.00
Pop. Mean	39.30	2.80	0.37	0.33	−0.06	1.67	0.40	0.66
Pop. SEM	0.30	0.33	0.08	0.07	0.06	0.19	0.16	0.11

*N* number of individuals, *k* number of alleles;  $H_O$  observed heterozygosity;  $H_E$  expected heterozygosity;  $F_{IS}$  inbreeding coefficient (Weir and Cockerham 1984), *M* *M*-ratios (Garza and Williamson 2001)

\*Significantly deviated from Hardy–Weinberg equilibrium

**Table 2** AMOVA results

Source of variation	Degrees of freedom	Sum of squares	Variance component	Fixation Index
Among populations	1	67.85	$V_a=0.85$	$F_{ST}=0.32$
Among individuals	76	111.62	$V_b=-0.36$	$F_{IS}=-0.20$
Within individuals	78	171.00	$V_c=2.19$	$F_{IT}=0.18$
Total	155	350.47	2.68	N/A

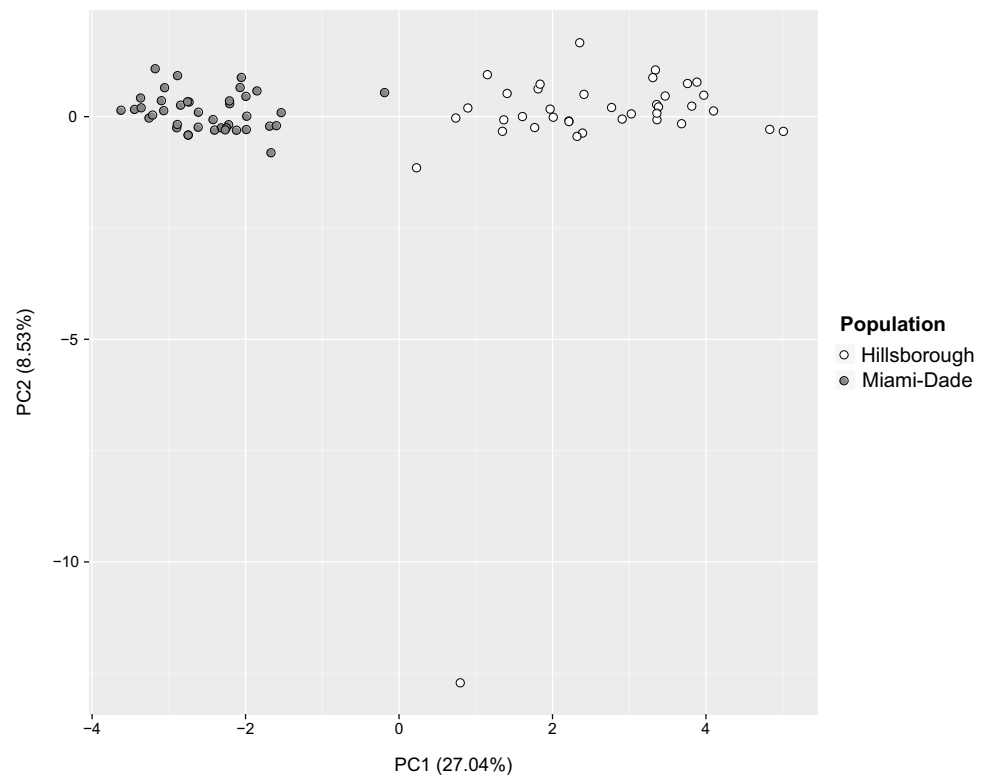
Plotting each population individually, excluding the outliers 37 and 42 to provide better resolution of intrapopulation genetic patterns showed only minor substructuring (Online Resource 1).

Evaluation of K-means clustering solutions for  $K=1-30$  showed that BIC decreased until  $K=5$  at which point BIC essentially leveled off for several values of  $K$  before increasing for  $K \geq 9$  (Online Resource 2). Consequently, we selected  $K=5$  as the optimal value of  $K$  for DAPC. The cross-validation procedure identified retention of eight

principal components in the DAPC as optimal. Therefore, we retained eight principal components in order to avoid overfitting. As can be seen in Fig. 4, the first discriminant function (horizontal plane) effectively distinguishes between the Hillsborough and Miami-Dade regions, whereas the second discriminant function (vertical plane) largely separates clusters from the same geographic region. Examination of the DAPC cluster membership probabilities (Online Resource 3) is consistent with the idea that membership between clusters within geographic regions is much fuzzier



**Fig. 3** Principal component analysis based on raw genotypes of introduced *S. merianae* populations in Florida



than membership between clusters from different geographic regions. However, individual 42 has a probabilistic membership of 0.844 to one of the Miami-Dade clusters, while also having respective probabilistic memberships of 0.092 and 0.061 to two different Hillsborough clusters.

As shown in Fig. 5a, the mean Ln P(D) plot as a function of K plateaus around K = 2 before decreasing for values of K > 3, a result indicative of the investigated range of K-values encompassing the optimal value of K. Furthermore, as seen in Fig. 5b,  $\Delta K$  indicates K = 2 as the best partition of the data. STRUCTURE detected evidence of admixture in the same two intermediate individuals (individuals 24 and 42) identified by the PCA and incorrectly assigned individual 42 to the Hillsborough cluster (Fig. 5c). The STRUCTURE analyses that we performed separately on the Hillsborough and Miami-Dade datasets were both indicative of a lack of substructure in each respective geographic region (i.e., K = 1; Online Resources 4–5).

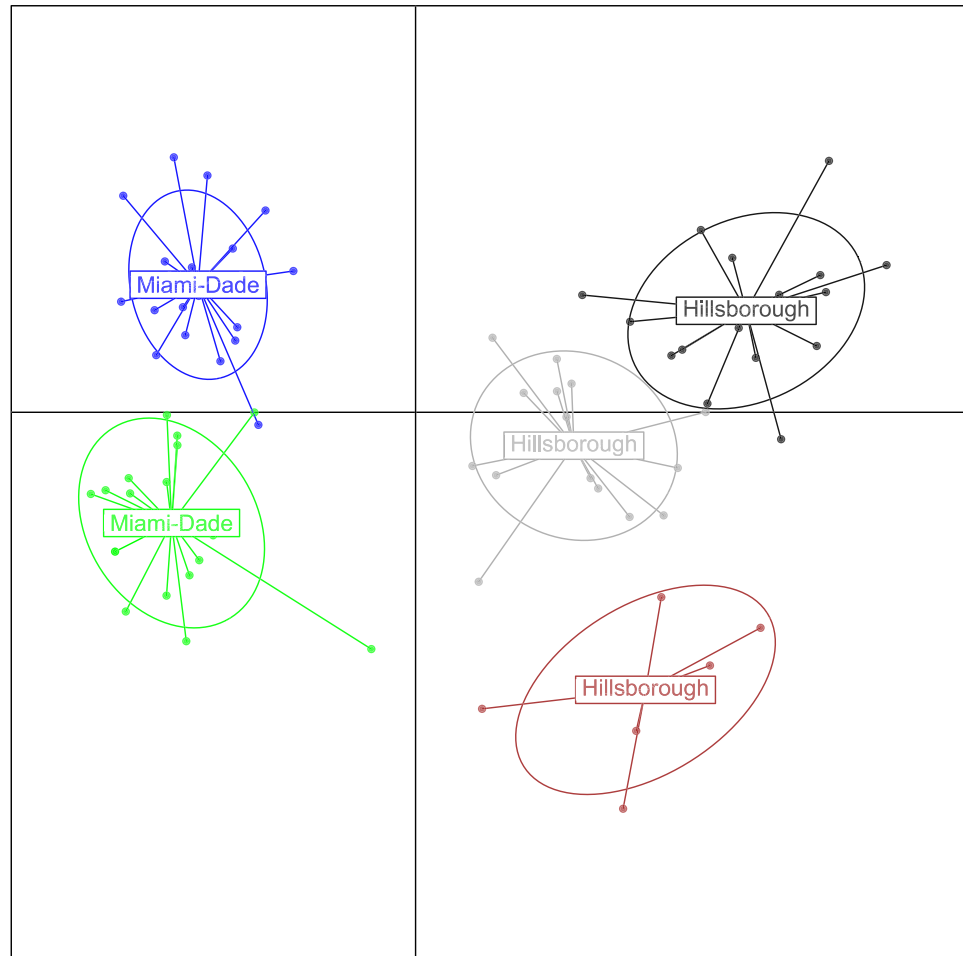
The log(marginal likelihoods) of the 10 best partitions visited by BAPS v5.3 were -1116.5052 (K = 5), -1116.7588 (K = 5), -1116.8369 (K = 5), -1117.1590 (K = 6), -1117.5120 (K = 5), -1117.9614 (K = 6), -1118.2606 (K = 5), -1118.2691 (K = 5), -1118.4195 (K = 5), -1118.4801 (K = 5) and the posterior probabilities for K returned by BAPS v5.3 were 0.0036 (K = 4), 0.7739 (K = 5), 0.1873 (K = 6), and 0.0352 (K = 7). As such, K = 5 was identified as the optimal partition. The five clusters identified by BAPS v5.3 can be summarized as follows: (cluster

1) eight Hillsborough individuals plus individual 42 who was sampled from Miami-Dade, (cluster 2) two Hillsborough individuals, (cluster 3) 27 Hillsborough individuals, (cluster 4) individual 37—the individual previously flagged as an outlier by PCA, and (cluster 5) thirty-nine individuals from Miami-Dade. Kullback–Leibler (KL) divergence values between pairs of clusters generally supported the view that clusters predominantly composed of individuals from the same geographic region are less divergent (mean KL divergence among clusters 1–3 = 0.811) than clusters composed predominantly of individuals from different geographic regions (mean KL divergence between clusters 1–3 and cluster 5 = 2.268). The admixture analysis performed using the output of the individual clustering (i.e., mixture analysis) performed in BAPS v5.3 did not provide evidence of admixture, as all admixture proportions were 1.00.

### Intraregional kinship patterns within DAPC clusters

As can be seen in Table 3, mean relatedness within the Hillsborough DAPC clusters varied between 0.12 and 0.25. For all three of these clusters, mean relatedness was greater than the 0.975 quantile for mean relatedness within randomly generated groups of the same size. Similarly, mean relatedness within the Miami-Dade DAPC clusters ranged from 0.11 to 0.14 and both of these values exceeded the 0.975 quantile for mean relatedness within randomly generated groups of the same size.

**Fig. 4** Discriminant analysis of principal components scatter plot showing ordination of genotypes onto canonical coordinates associated with the first two discriminant functions. The first discriminant function (x-axis) primarily distinguishes between the Miami-Dade and Hillsborough clusters, while the second discriminant function (y-axis) primarily separates clusters within geographic regions



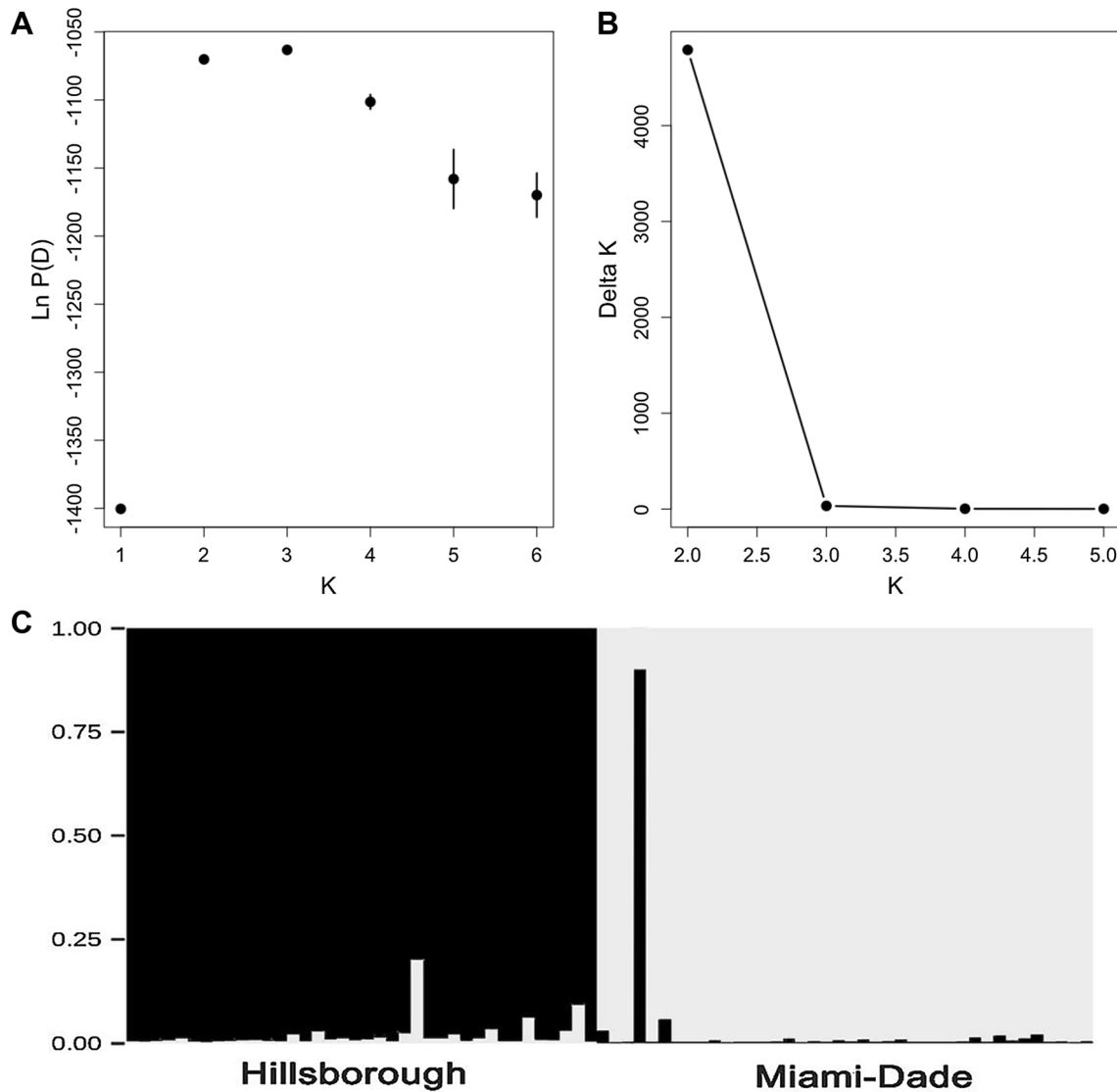
### Assessment of gene flow

The assignment analyses, performed in GENECLASS2, correctly assigned 77 of 78 individuals to the locales from which they were sampled (Fig. 6). The single animal that was incorrectly assigned was individual 42, who was sampled in Miami-Dade County but assigned to the Hillsborough cluster by STRUCTURE. Not surprisingly, GENECLASS2 found evidence that individual 42 was a first-generation migrant from Hillsborough ( $\log(L_{\text{home}}/L_{\text{max}}) = 2.295$ ,  $P = 0.0001$ ). Because the analyses we performed in DIYABC suggested the possibility of an un-sampled, captive population (see below) we repeated the migrant detection analysis in GENECLASS2 using the  $L_{\text{home}}$  likelihood estimation, which produces a more appropriate test statistic when there are populations that have not been sampled (Piry et al. 2004). The results of this analysis suggested that individual 42 from Miami-Dade County [ $-\log(L_{\text{home}}) = 10.721$ ,  $P = 0.0001$ ] and individual 37 from Hillsborough County [ $-\log(L_{\text{home}}) = 11.663$ ,  $P = 0.0001$ ] were both first-generation migrants; individual 37 being the same tegu that did not cluster with any known populations in the PCA.

Our analysis of recent migration rates in BAYESASS suggested that gene flow between Hillsborough and Miami-Dade is rare. We found that non-migrants accounted for 0.989 (CI 0.963–1.000) and 0.985 (CI 0.960–0.998) of the Hillsborough and Miami-Dade populations, respectively. Migrants from Hillsborough to Miami-Dade comprised 0.015 (0.002–0.040) of the population, while migrants from Miami-Dade to Hillsborough represented 0.011 (CI 0.000–0.037) of the population.

### Effective population size and demographic history

The effective population size for both of the introduced *S. merianae* populations was low, consistent with founder populations. For the Hillsborough population,  $N_e$  was estimated to be 10.8 (95% CI 3.2–34.2), while the Miami-Dade population was estimated to have an “infinite”  $N_e$  (95% CI 22.3– $\infty$ ), which indicates that the population lacks the genetic signature caused by a finite number of parents—a result attributable to sampling error (Do et al. 2014). Decreasing the allele frequency cutoff to 0.01



**Fig. 5** The mean  $\ln P(D) \pm SD$  from replicate STRUCTURE runs at each value of  $K$  investigated ( $n=10$ ) as a function of  $K$  (a). The value of  $\Delta K$  as a function of  $K$  (b). Results of the analysis performed in

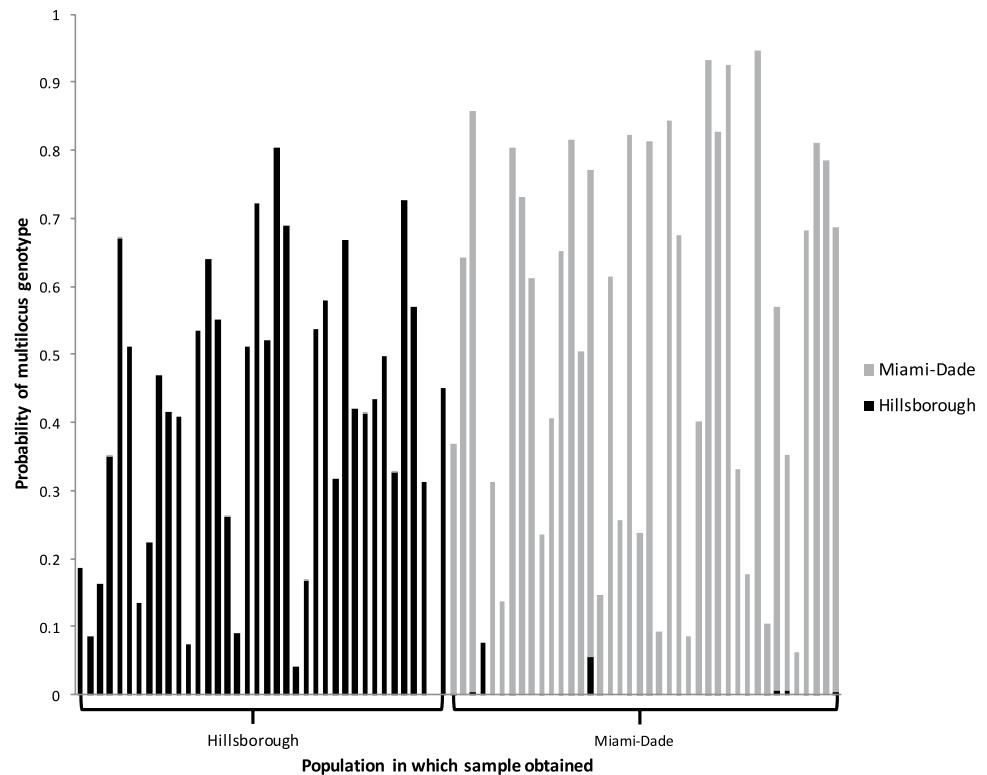
STRUCTURE when  $K=2$ . Bars represent average cluster membership across 10 replicate runs that were aligned using CLUMPP (c)

**Table 3** Results of the intraregional pair-wise relatedness within DAPC clusters analysis

Cluster	No. Individuals	Mean within cluster $r$	Average mean $r$ within random group across 1000 randomizations	0.025 quantile for mean $r$ within random group	0.975 quantile for mean $r$ within random group
Hill 5	6	0.247	-0.024	-0.227	0.140
Hill 2	16	0.138	-0.028	-0.112	0.032
Hill 1	16	0.122	-0.026	-0.114	0.035
MD 3	21	0.105	-0.030	-0.151	0.062
MD 4	19	0.144	-0.036	-0.158	0.066

Relatedness values and resampling analyses were computed separately for the Hillsborough and Miami-Dade populations

**Fig. 6** Stacked bar plots depicting the results of the assignment analysis performed in GENECLASS2. Each individual is represented by a bar that is presented over a label indicating the population in which that individual was sampled. For each individual, GENECLASS2 calculates the probability of that individual's multilocus genotype being derived from Hillsborough (black) and Miami-Dade (light gray). Thus, each bar can consist of as many as two colors, with the height of each color indicating the relative strength of assignment to each of the two populations



yielded  $N_e$  estimates of 6.7 (95% CI 2.7–14.8) and 8.2 (95% CI 2.9–21.2) for Hillsborough and Miami-Dade, respectively.

We found no evidence of population expansion for either population based on the  $k$  test (Hillsborough:  $P=0.93$ ; Miami-Dade:  $P=0.15$ ) and the  $g$  test (Hillsborough:  $g=1.89$ ; Miami-Dade:  $g=2.88$ ). However, the analyses performed in BOTTLENECK suggested that the Hillsborough population has undergone a recent population bottleneck (Wilcoxon test:  $P<0.001$  for heterozygote excess; Mode-shift: shifted), while heterozygote excess was not detected (Wilcoxon test:  $P=0.461$ ; Mode-shift: normal L-shaped distribution) for the Miami-Dade population.

Inspection of repeat size distribution data for each locus in each population revealed that all alleles across all loci and populations were multiples of their respective repeat units and are therefore consistent with stepwise and two-phase mutation models. However, in both populations, *Teg10* had outlier alleles that caused there to be  $>10$  unoccupied repeat positions. Consequently, we calculated mean  $M$ -ratios for both populations that excluded *Teg10* and assessed them for significance against  $M_c$  values obtained from simulations in *Critical\_M* based on 9 loci. The mean  $M$ -ratio, for Hillsborough, excluding *Teg10* ( $M=0.631$ ), was less than the corresponding  $M_c$  of 0.635, suggesting a recent population bottleneck. However, the mean  $M$ -ratio for Miami-Dade, excluding *Teg10* ( $M=0.713$ ), was greater than the corresponding  $M_c$  of 0.665.

## Tests of alternative introduction scenarios

The approximate Bayesian computation found evidence for two successive bottlenecks in both the Hillsborough and Miami-Dade lineages, with divergence of the two populations occurring after the first bottleneck. This demographic pattern is represented in Scenario 4 (Fig. 2), which had the highest posterior probability of 0.6605 (0.6434–0.6775) (Online Resource 6; Online Resource 7). The serial introduction hypotheses, Scenarios 2 and 3, also produced moderate posterior probability values of 0.1744 and 0.1226, respectively. All remaining introduction scenarios showed posterior probability levels below 0.015. Although Scenario 4 was found to have the highest posterior probability, the type I and type II errors were high (Online Resource 6), mainly due to low resolution among Scenario 4 and the serial introduction hypotheses (Scenarios 2 and 3). When Scenarios 2 and 3 were removed from the analysis, the posterior probability for Scenario 4 increased to 0.9370 (0.9298–0.9441), and the type I and type II errors decreased to 0.21 and 0.32, respectively. Further examination of the selected scenario via posterior model checking with all available summary statistics showed that none of the proportions (simulated  $<$  observed) fell outside the 0.05–0.95 range.

## Discussion

### Overview

In this study, we surveyed neutral variation in two invasive Argentine black-and-white tegu populations in Florida—one of which is in Hillsborough County and the other of which is in Miami-Dade County. Our analyses revealed that both of these populations have low genetic richness, and that the Hillsborough population exhibits genetic signatures of having undergone a recent bottleneck. In addition, several analytical approaches suggested that the populations in these respective regions of Florida are well differentiated from one another. Interestingly, use of ABC to infer the most likely introduction pathway suggested that the Hillsborough and Miami-Dade populations underwent a bottleneck event prior to diverging; however, it is unclear whether the first of these bottlenecks happened in a wild, native population or following importation into the United States (see below). In what follows, we elaborate on the implications of our results and make recommendations to wildlife managers who are working to control Florida's self-perpetuating *S. merianae* populations.

### Within population diversity

Heterozygote excess tests and  $M$ -ratios suggested that only the Hillsborough *S. merianae* population showed unequivocal evidence of a recent founder effect. This result is somewhat surprising considering that both populations were likely founded by a small number of individuals within the past 20 years. It is possible that our failure to unambiguously detect a bottleneck in the Miami-Dade population is a Type II statistical error, as bottleneck tests have failed to detect known population collapses, and recent simulation-based power analyses found that heterozygosity excess tests have limited power to detect 10 to 1000-fold population declines (Peery et al. 2012). Unsurprisingly, both *S. merianae* populations had low levels of allelic richness (range of number of alleles per locus: 2–5 in Hillsborough and 2–4 in Miami-Dade) and average numbers of effective alleles across loci that were below three. While definitive conclusions about reductions in genetic diversity would require comparisons to wild populations within the native range, our  $N_e$  estimates were low (point estimates for Hillsborough ranged from 6.7 to 10.8; point estimate for Miami-Dade = 8.2) and are comparable to assessments of  $N_e$  in the Nile monitor (*Varanus niloticus*)—an ecologically similar lizard that is also invasive to Florida (Wood et al. 2016).

### Intraregional genetic structure

Although DAPC and BAPS recovered different patterns of individual assignment, both approaches suggested that there is intraregional genetic structuring. While BAPS indicated that only the Hillsborough population exhibits substructure, the DAPC identified three Hillsborough only clusters and two Miami-Dade only clusters. Thus, contrary to the patterns obtained by STRUCTURE, PCA, BAPS, and GENECLASS2, DAPC correctly assigned all samples to their geographic region of origin. The patterns of assignment recovered by DAPC, and to a lesser extent BAPS, raise the possibility that the population in Hillsborough, and perhaps the population in Miami-Dade as well, have been infused with subsequent introductions since their initial establishment. However, it is also possible that substructuring in Hillsborough and Miami-Dade is fully or partially attributable to kin groups. Our region-specific examinations of relatedness within DAPC clusters, support the idea that these clusters are enriched with individuals who are kin. Indeed, mean relatedness within all five DAPC clusters was statistically greater than mean relatedness within randomly generated groups of the same size and, in most cases, approximated the level of coancestry that would be expected of first cousins.

### Assessment of gene flow

We used several independent analyses to examine the degree of genetic structure between the *S. merianae* populations in Hillsborough and Miami-Dade counties. While our results are generally consistent with the view that there is marked genetic differentiation between these two populations, we did detect intermediate and outlier genotypes. In particular, STRUCTURE identified two individuals (tegu 24 collected in Hillsborough County and tegu 42 collected in Miami-Dade County) with substantial admixture proportions, and both of these individuals were intermediate to the Hillsborough and Miami-Dade clusters identified by PCA. Moreover, PCA identified another individual (tegu 37) with a PC2 coordinate that markedly differed from all other individuals in our sample. BAPS also corroborated this result by assigning tegu 37 to a cluster that contained no other individuals. While it is possible, if not likely, that these rare genotypes were sampled by chance and are not indicative of population or anthropogenic phenomena, these results raise the possibility that there has been human-mediated gene flow between the Hillsborough and Miami-Dade populations and that individuals from an unknown population are migrating or being introduced into Florida's wild *S. merianae* populations. Therefore, it may be worthwhile to conduct periodic genetic surveys of these populations to assess whether genetic diversity and admixture increase over time.

## Introduction scenario testing

Our analysis of possible introduction pathways yielded multiple competing scenarios. Although Scenario 4 displayed the highest posterior probability overall, the lack of resolution between this scenario and the serial introduction hypotheses (Scenarios 2 and 3), produced high type I and type II errors. In a serial introduction pathway, one population originally becomes established in the non-native range and subsequently acts as a source for a secondary invasion, forming a new population. Based on our knowledge of the introduced *S. merianae* populations and the high level of genetic differentiation between them, a serial introduction pathway is less probable than the alternative scenario. Because Hillsborough and Miami-Dade are approximately 300 km apart, it is unlikely that individuals could migrate this distance to form a secondary population, with no known intermediate populations. However, as indicated previously, human-mediated transport is a possible explanation.

Excluding the serial introduction hypotheses from the analysis increased the posterior probability and decreased the error rate for Scenario 4. In this scenario, both of the *S. merianae* populations have undergone multiple successive bottleneck events. One possible explanation is that the two introduced *S. merianae* populations originated from the same captive-bred population. The first bottleneck event occurred when the initial captive population was formed, followed by a second bottleneck occurring at the time of introduction. In the United States, *S. merianae* is one of the most commonly bred tegu species (Bartlett and Bartlett 1996). Additionally, the number of reported *S. merianae* imported into the United States is relatively low, compared to other reptiles in the pet trade, with an average of less than 600 live individuals imported annually, primarily from Argentina (<http://trade.cites.org/>). By comparison, Nile monitors (*Varanus niloticus*), another common lizard species in the pet trade, are imported into the United States at rates averaging around 7,000 live individuals per year (<http://trade.cites.org/>). The low rate of tegu imports could be a reflection of the predominance of captive-bred individuals in the pet market, suggesting a higher likelihood that a captive population was the source of the introduction.

Alternatively, the population fluctuations identified in the *S. merianae* lineage could reflect events occurring in the species' native range. For example, recent overexploitation or habitat destruction could have led to a population decline prior to their introduction. Aside from the pet industry, tegus are commonly harvested for their skins, and represent one of the most heavily exploited lizard species in Latin America (Fitzgerald 1994; Chardonnet et al. 2002; Mieres and Fitzgerald 2006). Therefore, another potential explanation for Scenario 4 is that the native source population underwent a population decline, followed by two independent

introductions into Florida. Future studies examining the genetic patterns of South American and captive-bred *S. merianae* could further elucidate the source and introduction history of the Florida populations.

We limited the scope of this analysis to compare common routes of species introduction and did not examine complex patterns involving admixture or additional divergence events. Therefore, the true introduction pathway could be more complicated than those described and would not be captured by our analysis. Additionally, genetic signatures are not always reflective of underlying demographic patterns, and therefore, the results of this study must be viewed cautiously. Positive selection can alter the allele frequencies of neighboring genomic regions, even those thought to be neutral, and can be difficult to distinguish from demographic processes (Maynard-Smith and Haigh 1974; Kaplan et al. 1989; Stephan et al. 1992). Selective sweeps, i.e. strong selection for an advantageous mutation, have been shown to produce an excess of rare alleles compared to equilibrium expectations (Tajima 1989b; Braverman et al. 1995; Fu 1997), which can also result from a recent population bottleneck (Tajima 1989a) as well as from population expansion (Fu and Li 1993). Although selection is often thought to act on specific regions of the genome, and therefore affect few loci in a given analysis, some studies suggest that natural selection can produce genome-wide perturbations (Zayed and Whitfield 2008; Sella et al. 2009; Corbett-Detig et al. 2015), leading to incorrect inferences of demographic patterns (Schridder et al. 2016).

## Conclusion

In this study, we have shown that the *S. merianae* populations present in Hillsborough and Miami-Dade have low standing molecular genetic variation and are well differentiated from one another. However, because of the possibilities of human-mediated movement, the existence of unknown populations, and repeated escapes/releases from captive sources, we propose that interregional coordination is likely to be an important component to the effective management of Florida's wild *S. merianae* populations. According to Florida EddMaps (<http://www.eddmaps.org/florida/distribution/viewmap.cfm?sub=18346>), verified *S. merianae* specimens have already been documented via photograph near the cities of Port Charlotte, Naples, and Port St. Lucie—all of which are over 150 km from the nearest known breeding population. If more self-perpetuating populations become established, natural dispersal between populations will become increasingly likely. As such, it is important for wildlife managers to follow-up on credible sightings and monitor potential migration corridors. Finally, our results are consistent with the idea that both of Florida's

*S. merianae* populations originated from a common captive source. While other scenarios are also compatible with our data (see above), it is important for the Florida Wildlife Commission to continue closely monitoring the exotic pet trade, as it is likely to be the culprit primarily responsible for the introduction and establishment of *S. merianae* in Florida.

**Acknowledgements** JPW and RBP would like to thank their late friend and mentor, David H. Reed, for his advice and support during the earliest phases of this project. We also gratefully acknowledge Jennifer Ketterlin Eckles, Colin Hughes, Tessie Offner, Kenneth L. Krysko, Kevin Enge, Balm Boyette Scrub Preserve, USDA-APHIS, and the South Florida Water Management Group for providing samples, field resources and assistance, lodging, logistics, and site access.

## Compliance with ethical standards

**Conflict of interest** Jared Wood was directly supported by the Columbus Zoo and Aquarium and by the Wallace Endowment to the University of Louisville, Department of Biology. Stephanie Dowell Beer declares that she has no conflict of interest. Todd Campbell received funding for field work from the Charlotte Harbor National Estuary Program, the National Fish and Wildlife Foundation, the Institute for Biological Invasions at the University of Tennessee, and the University of Tampa Dana grant. Robert Page declares that he has no conflict of interest.

## References

- Abdelkrim J, Pascal M, Calmet C, Samadis S (2005) Importance of assessing population genetic structure before eradication of invasive species: examples from insular Norway rat populations. *Conserv Biol* 19(5):1509–1518
- Abe AS (1983) Observation on dormancy in tegu lizards—*Tupinambis teguixin* (Reptilia, Teiidae). *Naturalia* 8:135–139
- Bartlett RD, Bartlett P (1996) Monitors, tegus, and related lizards: everything about selection, care, nutrition, diseases, breeding, and behavior. Barron's Educational Series, Inc, Hauppauge
- Bhargava A, Fuentes FF (2010) Mutational dynamics of microsatellites. *Mol Biotechnol* 44:250–266
- Bilgin R (2007) kgtests: a simple Excel Macro program to detect signatures of population expansion using microsatellites. *Mol Ecol Notes* 7:416–417
- Braverman JM, Hudson RR, Kaplan NL, Langley CH, Stephan W (1995) The hitchhiking effect on the site frequency spectrum of DNA polymorphisms. *Genet* 140:783–796
- Brito SP, Abe AS, Andrade DV (2001) *Tupinambis merianae* (tegu lizard). *Longev Herpetol Rev* 32:260–261
- Cardozo G, Naretto S, Zak M, Chiaraviglio M (2012) The role of landscape in contact zones of sister species of lizards. In: Tiefenbacher J (ed) *Perspectives on nature conservation—patterns, pressures and prospects*. James and Marilyn Lovell Center for Environmental Geography and Hazards Research, Department of Geography Texas State University, San Marcos, pp 161–176
- Chamut S, Jahn GA, Arce OEA, Manes ME (2012) Testosterone and reproductive activity in the male tegu lizard, *Tupinambis merianae*. *Herpetol Conserv Biol* 7(3):299–305
- Chardonnet P, Des Clers B, Fisher J, Gerhold R, Jori F, Lamarque F (2002) The value of wildlife. *Revue Scientifique et Technique de l'Office International des Epizooties* 21:15–51
- Clavero M, Garcí a-Berthou E (2005) Invasive species are a leading cause of animal extinction. Institute of Aquatic Ecology, University of Girona, Girona
- Corander J, Marttinen P, Siren J, Tang J (2008) Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. *BMC Bioinform* 9:539
- Corbett-Detig RB, Hartl DL, Sackton TB (2015) Natural selection constrains neutral diversity across a wide range of species. *PLoS Biol* 13:e1002112
- Cornuet J-M, Pudlo P, Veyssier J, Dehne-Garcia A, Gautier M, Leblois R, Marin J-M, Estoup A (2014) DIYABC v2. 0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinform* 30:1187–1189
- Di Castri F (1989) History of biological invasions with special emphasis on the old world. In: Drake JA (ed) *Biological invasions: a global perspective*. John Wiley, New York
- Di Rienzo A, Peterson AC, Garza JC, Valdes AM, Slatkin M, Freimer MB (1994) Mutational processes of simple-sequence repeat loci in human populations. *Proc Natl Acad Sci USA* 91(8):3166–3170
- Do C, Waples RS, Peel D, Macbeth G, Tillett BJ, Ovenden JR (2014) NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size ( $N_e$ ) from genetic data. *Mol Ecol Resour* 14:209–214
- Duarte Varela CF, Cabrera MR (2000) Testing skeletochronology in black tegu lizards (*Tupinambis merianae*) of known ages. *Herpetol Rev* 31:224–226
- Dyer RJ (2012) The gstudio package. Virginia Commonwealth University, Virginia
- Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour* 4(2):359–361
- Embert D, Fitzgerald L, Waldez F (2010) *Salvator merianae*. IUCN 2013. IUCN red list of threatened species. Version 2013.2
- Enge KM (2007) FWC bioprofile for the Argentine black and white tegu (*Tupinambis merianae*). Florida Fish and Wildlife Conservation Commission Report, Tallahassee, 27 pp
- Estoup A, Jarne P, Cornuet JM (2002) Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Mol Ecol* 11:1591–1604
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14:2611–2620
- Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetic analyses under Linux and Windows. *Mol Ecol Resour* 10:564–567
- Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164:1567–1587
- Fitzgerald LA (1994) The interplay between life history and environmental stochasticity: implication for the management of exploited lizard populations. *Am Zool* 34:371–381
- Fitzgerald LA, Cruz FB, Perotti MG (1993) The reproductive cycle and size at maturity of *Tupinambis rufescens* (Sauria: Teiidae) in the dry Chaco of Argentina. *J Herpetol* 27:70–78
- Fu Y-X (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147:915–925
- Fu Y-X, Li W-H (1993) Statistical tests of neutrality of mutations. *Genetics* 133:693–709
- Galetti M, Bovendorp RS, Fadini RF, Gussoni COA, Rodrigues M, Alvarez AD, Guimaraes PR, Alves K (2009) Hyper-abundant mesopredators and bird extinction in an Atlantic forest island. *Zoologia* 26:288–298

- Garza JC, Williamson EG (2001) Detection of reduction in population size using data from microsatellite loci. *Mol Ecol* 10:305–318
- Guichoux E, Lagache L, Wagner S, Chaumeil P, Leger P, Lepais O, Lepoittevin C, Malausa T, Revardel E, Salin F, Petit RJ (2011) Current trends in microsatellite genotyping. *Mol Ecol Resour* 11:591–611
- Hall BJ (1978) Notes on the husbandry, behaviour and breeding of captive tegu lizards. *Int Zoo Yearb* 18:91–95. <https://doi.org/10.1111/j.1748-1090.1978.tb00229.x>
- Holm S (1979) A simple sequentially rejective multiple test procedure. *Scand J Stat* 6:65–70
- Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinform* 23:1801–1806
- Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinform* 24(11):1403–1405
- Jombart T, Collins C (2015) A tutorial for discriminant analysis of principal components (DAPC) using adegenet 2.0.0. Imperial College London, MRC Centre for Outbreak Analysis and Modelling, London
- Jombart T, Sebastien D, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet* 11:94
- Kaplan NL, Hudson R, Langley C (1989) The “hitchhiking effect” revisited. *Genetics* 123:887–899
- Kiefer MC, Sazima I (2002) Diet of juvenile tegu lizard *Tupinambis merianae* (Teiidae) in southeastern Brazil. *Amphib-Reptil* 23:105–108
- Klug PE, Reed RN, Mazzotti FJ, McEachern MA, Vinci JJ, Craven KK, Yackel Adams AA (2015) The influence of disturbed habitat on the spatial ecology of Argentine black and white tegu (*Tupinambis merianae*), a recent invader in the Everglades ecosystem (Florida, USA). *Biol Invasion* 17:1785–1797
- Krysko KL, Burgess JP, Rochford MR, Gillette CR, Cueva D, Enge KM, Somma LA, Stabile JL, Smith DC, Wasilewski JA, Kieckhefer GN III, Granatosky MC, Nielson SV (2011) Verified non-indigenous amphibians and reptiles in Florida from 1983 through 2010: outlining the invasion process and identifying invasion pathways and stages. *Zootax* 3028:1–64
- Lanfri S, Di Cola V, Naretto S, Chiaraviglio M, Cardozo G (2013) Understanding the ecological niche to elucidate spatial strategies of the southernmost *Tupinambis* lizards. *Amphib Reptil* 34:551–565
- Luxmoore R, Groombridge B, Broods S (1988) Significant trade in wildlife: a review of selected species in CITES Appendix II. Volume 2: reptiles and invertebrates. IUCN Conservation Monitoring Centre, Cambridge, 306 pp
- Mack RN, Simberloff D, Lonsdale WM, Evans H, Clout M, Bazzaz FA (2000) Biotic invasions: causes, epidemiology, global consequences, and control. *Ecol Appl* 10:689–710
- Maynard-Smith J, Haigh J (1974) The hitch-hiking effect of a favourable gene. *Gene Res* 23:23–35
- Mazzotti FJ, McEachern M, Rochford M, Reed RN, Eckles K, Vinci J, Edwards J, Wasilewski J, J (2015) *Tupinambis merianae* as nest predators of crocodylians and turtles in Florida, USA. *Biol Invas* 17:47–50
- Meirmans P, Hedrick PW (2011) Assessing population structure:  $F_{ST}$  and related measures. *Mol Ecol Resour* 11:5–18
- Mercolli C, Yanosky A (1994) The diet of adult *Tupinambis teguixin* (Sauria: Teiidae) in the eastern chaco of Argentina. *Herpetol J* 4:15–19
- Mieres MM, Fitzgerald LA (2006) Monitoring and managing the harvest of tegu lizards in Paraguay. *J Wildl Manag* 70:1723–1734
- Mooney HA, Cleland EE (2001) The evolutionary impact of invasive species. *PNAS* 98(10):5446–5451
- Nei M, Chesser RK (1983) Estimation of fixation indices and gene diversities. *Ann Hum Genet* 47:253–259
- Paetkau D, Calvert W, Stirling I, Strobeck C (1995) Microsatellite analysis of population structure in Canadian polar bears. *Mol Ecol* 4:347–354
- Paetkau D, Slade R, Burden M, Estoup A (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Mol Ecol* 13:55–65
- Peakall R, Smouse PE (2012) GenA1Ex 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinform* 28:2537–2539
- Peery MZ, Kirby R, Reid BN, Stoelting R, Doucet-Béer E, Robinson S, Vásquez-Carrillo C, Pauli JN, Palsbøll PJ (2012) Reliability of genetic bottleneck tests for detecting recent population declines. *Mol Ecol* 14:3403–3418
- Pernas T, Giardina DJ, McKinley A, Parns A, Mazzotti FJ (2012) First observations of nesting by the Argentine black and white tegu, *Tupinambis merianae*, in South Florida. *Southeast Nat* 11(4):765–770
- Pew J, Muir PH, Wang J, Frasier TR (2015) Related: a R package for analyzing pairwise relatedness from codominant molecular markers. *Mol Ecol Res* 15:557–561
- Piry S, Luikart G, Cornuet J-M (1999) BOTTLENECK: a program for detecting recent effective population size reductions from allele data frequencies. Montpellier, France
- Piry S, Alapetite A, Cornuet JM, Paetkau D, Baudouin L, Estoup A (2004) GENECLASS2: a software for genetic assignment and first-generation migrant detection. *J Hered* 95:536–539
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genet* 155:945–959
- Queller DC, Goodnight KF (1989) Estimating relatedness using molecular markers. *Evol* 43:258–275
- R Core Team (2014) R: a language and environment for statistical computing. R Foundation for Statistical Computing
- Ramasamy RK, Ramasamy S, Bindroo BB, Naik VG (2014) STUCTURE PLOT: a program for drawing elegant STRUCTURE bar plots in user friendly interface. *Springer* 3:431
- Reich DE, Goldstein DB (1998) Genetic evidence for a Paleolithic human population expansion in Africa. *Proc Natl Acad Sci* 95:8119–8123
- Reich DE, Feldman MW, Goldstein DB (1999) Statistical properties of two tests that use multilocus data sets to detect population expansion. *Mol Biol Evol* 16(4):453–466
- Rollins LA, Woolnough AP, Wilton AN, Sinclair RON, Sherwin WB (2009) Invasive species can't cover their tracks: using microsatellites to assist management of starling (*Sturnus vulgaris*) populations in Western Australia. *Mol Ecol* 18:1560–1573
- Rousset F (2008) GENEPOP 007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Mol Ecol Resour* 8:103–106
- Schrider DR, Shanku AG, Kern AD (2016) Effects of linked selective sweeps on demographic inference and model selection. *Genetics* 116:190223
- Sella G, Petrov DA, Przeworski M, Andolfatto P (2009) Pervasive natural selection in the *Drosophila* genome? *PLoS Genet* 5:e1000495
- Smith KG (2006) Patterns of nonindigenous herpetofaunal richness and biotic homogenization among Florida counties. *Biol Conserv* 127:327–335
- Stephan W, Wiehe TH, Lenz MW (1992) The effect of strongly selected substitutions on neutral polymorphism: analytical results based on diffusion theory. *Theor Pop Biol* 41:237–254
- Tajima F (1989a) The effect of change in population size on DNA polymorphism. *Genet* 123:597–601



- Tajima F (1989b) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genet* 123:585–595
- Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol Notes* 4:535–538
- Wake DB, Vredenburg VT (2008) Are we in the midst of the sixth mass extinction? A view from the world of amphibians. *PNAS*. <https://doi.org/10.1073/pnas.0801921105>
- Waples RS, Do C (2008) LDNE: a program for estimating effective population size from data on linkage disequilibrium. *Mol Ecol Resour* 8:753–756
- Waples RS, Do C (2010) Linkage disequilibrium estimates of contemporary  $N_e$  using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evol Appl* 3:244–262
- Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evol* 38:1358–1370
- Wilcove DS, Rothstein D, Dubow J, Phillips A, Losos E (1998) Assessing the relative importance of habitat destruction, alien species, pollution, over-exploitation, and disease. *BioSci* 48:607–616
- Wilson GA, Rannala B (2003) Bayesian inference of recent migration rates using multilocus genotypes. *Genet* 163:1117–1119
- Winck GR, Cechin SZ (2008) Hibernation and emergence pattern of *Tupinambis merianae* (Squamata: Teiidae) in the Taim Ecological Station, southern Brazil. *J Nat Hist* 42:239–247
- Wood JW, Campbell TS, Page RB (2015) Characterization of 14 novel microsatellite loci in the Argentine black and white tegu (*Salvator merianae*) via 454 pyrosequencing. *Amphib Reptil* 36:444–449
- Wood JW, Dowell SA, Campbell TS, Page RB (2016) Insights into the introduction history and population genetic dynamics of the Nile monitor (*Varanus niloticus*) in Florida. *J Hered*. <https://doi.org/10.1093/jhered/esw014>
- Zayed A, Whitfield CW (2008) A genome-wide signature of positive selection in ancient and recent invasive expansions of the honey bee *Apis mellifera*. *Proc Natl Acad Sci USA* 105:3421–3426