UC Merced
UC Merced Previously Published Works

Title
Genome-Wide Assessment of Diversity and Divergence Among Extant Galapagos Giant Tortoise Species

Permalink
https://escholarship.org/uc/item/4wj350nf

Journal
Journal of Heredity, 109(6)

ISSN
0022-1503

Authors
Miller, Joshua M
Quinzin, Maud C
Edwards, Danielle L
et al.

Publication Date
2018-08-24

DOI
10.1093/jhered/esy031

Peer reviewed

OXFORD

Original Article

Genome-Wide Assessment of Diversity and Divergence Among Extant Galapagos Giant Tortoise Species

**Joshua M. Miller*, Maud C. Quinzin*, Danielle L. Edwards,
Deren A. R. Eaton, Evelyn L. Jensen, Michael A. Russello, James P. Gibbs,
Washington Tapia, Danny Rueda, and Adalgisa Caccone**

From the Department of Ecology and Evolutionary Biology, Yale University, 21 Sachem St. New Haven, CT, 06520 (Miller, Quinzin, Eaton, Caccone); the Life and Environmental Sciences, University of California, Merced, 5200 N Lake Rd, Merced, CA (Edwards); the Department of Ecology, Evolution, and Environmental Biology, Columbia University, New York, NY (Eaton); the Department of Biology, University of British Columbia, Okanagan Campus, Kelowna, BC V1V 1V7, Canada (Jensen, Russello); the College of Environmental Science & Forestry, State University of New York, Syracuse, NY (Gibbs); the Galapagos Conservancy, Fairfax, VA (Tapia); the Galápagos National Park Directorate, Puerto Ayora, Galápagos, Ecuador (Tapia, Rueda).

Address correspondence to Joshua M. Miller and Adalgisa Caccone at the address above, or e-mail: joshua.miller@yale.edu, adalgisa.caccone@yale.edu.

*These authors contributed equally to this work

Abstract

Genome-wide assessments allow for fuller characterization of genetic diversity, finer-scale population delineation, and better detection of demographically significant units to guide conservation compared with those based on "traditional" markers. Galapagos giant tortoises (*Chelonoidis* spp.) have long provided a case study for how evolutionary genetics may be applied to advance species conservation. Ongoing efforts to bolster tortoise populations, which have declined by 90%, have been informed by analyses of mitochondrial DNA sequence and microsatellite genotypic data, but could benefit from genome-wide markers. Taking this next step, we used double-digest restriction-site associated DNA sequencing to collect genotypic data at >26 000 single nucleotide polymorphisms (SNPs) for 117 individuals representing all recognized extant Galapagos giant tortoise species. We then quantified genetic diversity, population structure, and compared results to estimates from mitochondrial DNA and microsatellite loci. Our analyses detected 12 genetic lineages concordant with the 11 named species as well as previously described structure within one species, *C. becki*. Furthermore, the SNPs provided increased resolution, detecting admixture in 4 individuals. SNP-based estimates of diversity and differentiation were significantly correlated with those derived from nuclear microsatellite loci and mitochondrial DNA sequences. The SNP toolkit presented here will serve as a resource for advancing efforts to understand tortoise evolution, species radiations, and aid conservation of the Galapagos tortoise species complex.

Galapagos giant tortoises (*Chelonoidis* spp.) are an icon of the Galapagos Archipelago. As the largest herbivore native to the archipelago, Galapagos giant tortoises are ecosystem engineers and seed dispersers that provide critical ecosystem functions (Gibbs et al. 2010). Phylogenetic and phylogeographic methods have estimated that the ancestor of this radiation is likely a South American species that arrived on the oldest islands about 3–4 million years ago (MYA; Caccone et al. 1999, 2002; Poulakakis et al. 2012; Kehlmaier et al. 2017). From there, tortoises colonized the islands where they now occur starting ~1.75 MYA (Poulakakis et al. 2012). Previous genetic work on this species complex based on both mitochondrial DNA (mtDNA) and microsatellite loci has shown that the current distribution occurred through an interplay of dispersal and vicariance, with little or no natural gene flow after the initial colonization of the different islands (Poulakakis et al. 2012). Thus, distinct genetic lineages have formed through allopatric differentiation, each likely on their own evolutionary trajectory (Caccone et al. 1999, 2002; Ciofi et al. 2002, 2006; Beheregaray et al. 2003a, 2004; Russello et al. 2005, 2007; Edwards et al. 2013; Poulakakis et al. 2015, 2008, 2012; Garrick et al. 2015). Currently there are 15 named species (Turtle Taxonomy Working Group 2017). Most islands host only a single species (Figure 1), except for the 2 largest islands, Santa Cruz and Isabela, where 2 and 5 species occur, respectively.

Since the islands were first discovered in the 1500's, numbers of Galapagos giant tortoises have declined by ~90% (Márquez et al. 2004) due to past exploitation by mariners and early colonists (see Porter 1822; Townsend 1925) and introduction of nonnative species which are predators of tortoise nests and hatchlings as well as competitors for forage (Guézou et al. 2010; Phillips et al. 2012). Population declines have resulted in the extinction of 4 tortoise species, and the placement of 9 on the IUCN Red List of endangered taxa [4 as Vulnerable, 3 as Endangered, and 2 as Critically Endangered (Turtle Taxonomy Working Group 2017)]. One recently

named species (*C. donfaustoi*) has yet to be fully evaluated by the IUCN, although likely is eligible for Critically Endangered status based on its small distribution and low population numbers (Turtle Taxonomy Working Group 2017). In response to these declines, concerted conservation efforts are ongoing throughout the archipelago focused on habitat restoration, pest removal (Campbell et al. 2004; Harper and Carrion 2011), captive breeding, and head-start programs to stimulate population recovery (Fritts et al. 2000; Milinkovitch et al. 2004, 2007; Cayot 2008; Jensen et al. 2015).

Until recently, most estimates of genetic variation and differentiation in this complex have been made using mtDNA or nuclear microsatellite loci. However, with the rapid declines in costs and increases in information generated by high-throughput sequencing methods, it has become feasible to characterize genome-wide diversity in nearly any organism (Davey et al. 2011; Ellegren 2014). Transitioning from mtDNA and microsatellite loci to genomic marker sets has numerous advantages. For example, the increased number of loci spread throughout a genome allows fuller characterization of genetic diversity and detection of fine-scale population differentiation (e.g. Viengkone et al. 2016). In addition, the use of reduced-representation "genotyping-by-sequencing" techniques (e.g. Baird et al. 2008) allow for markers to be simultaneously identified and genotyped in a large set of individuals, thereby reducing the potential of ascertainment bias (Heslot et al. 2013; McTavish and Hillis 2015). These methods are especially useful for detecting differences among lineages resulting from rapid radiations, where there has been relatively little time for genetic differentiation to accumulate (e.g. Wagner et al. 2013; Campagna et al. 2015). The genome-wide assessment of populations can also be a powerful conservation tool, as it provides increased resolution to define management units (Funk et al. 2012), characterize adaptive potential in the face of changing environmental conditions (Eizaguirre and Baltazar-Soares 2014; Harrisson et al. 2014), and identify hybrid individuals that
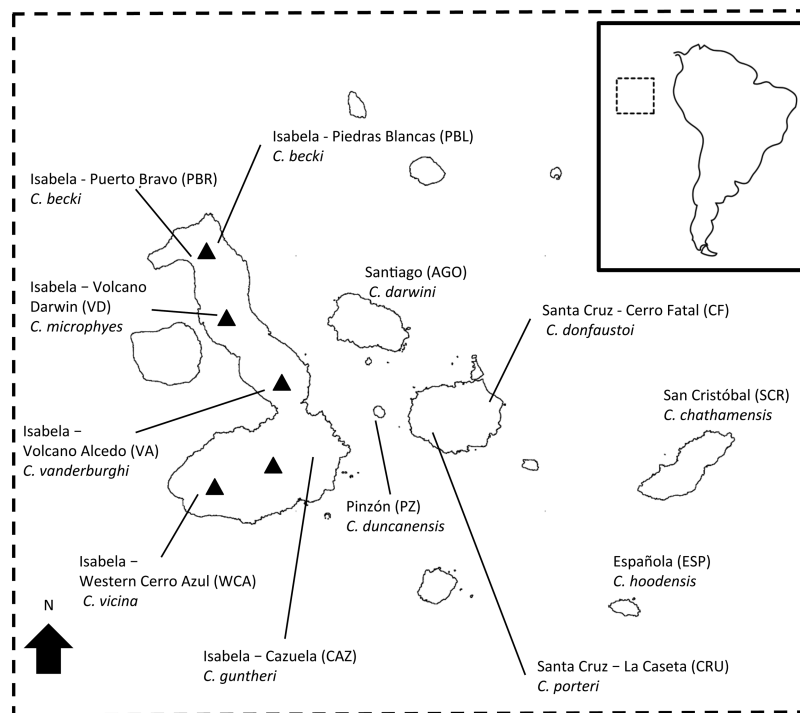

Figure 1. Map of the Galapagos Islands indicating sampling locations used in this genome-wide assessment of diversity and divergence among extant Galapagos giant tortoises. Population abbreviations are in parentheses, the species name is reported below each island and population symbol.

may require special management considerations (Wayne and Shaffer 2016). In addition, genome-wide information can be incorporated into management actions, such as when planning breeding programs or choosing individuals for translocations or reintroductions (Bosse et al. 2015; Whiteley et al. 2015).

Here, we use double-digest restriction-site associated DNA sequencing (ddRAD; Peterson et al. 2012) to identify genome-wide variation using a population dataset including 117 individuals from all extant Galapagos giant tortoise lineages. We then used these >26 000 loci to reconstruct genome-wide patterns and levels of genetic diversity and compare them to those obtained by previous mtDNA and microsatellite loci analyses.

Methods

Sample Collection and Sequencing

In the remainder of the manuscript, we use "species" to refer to named taxa, "population" to refer to genetically distinct, but unnamed groups, and "lineage" when referring to both. Samples used in this study were collected during previous expeditions (Caccone et al. 1999, 2002; Ciofi et al. 2002, 2006; Beheregaray et al. 2003a, 2003b, 2004; Russello et al. 2005, 2007; Poulakakis et al. 2008, 2012; Edwards et al. 2013, 2014; Garrick et al. 2014, 2012). We used approximately 10 samples per lineage (n = 117 individuals in total; Table 1). We grouped the sampling sites Los Pegas and Las Tablas (previously noted as LP and LT) as West Cerro Azul (WCA), given that tortoises at these locations are genetically indistinguishable despite being separated by several kilometers (Edwards et al. 2014). The individuals included in the dataset for this study were chosen from among over 300 samples previously screened at mtDNA (~700 bp of control region) and nuclear microsatellite loci (12 loci) that included all the extant and extinct lineages (Poulakakis et al. 2012; Edwards et al. 2013).

DNA was extracted from blood samples using a DNeasy Blood and Tissue kit (Qiagen) according to the manufacturer's instructions. We then prepared ddRAD libraries following Peterson et al. (2012). For each sample, 500 ng of genomic DNA was digested with the restriction enzymes MluCI and NlaIII (New England BioLabs) and ligated with Illumina-specific adaptors tagged using 18 unique barcodes and 2 index codes. Individual ligation products were quantified using Qubit (Invitrogen) fluorometry and then pooled into 13 libraries that each included between 12 and 18 individuals at equal molar concentrations. Pooled libraries were size-selected ~310 bp (range = 279–341 bp) with a BluePippin (Sage Science). Size-selected libraries were sequenced using 75 bp paired end chemistry on 13 lanes of an Illumina HiSeq 2000 at the Yale Center for Genome Analysis.

SNP Discovery and Screening

We used forward and reverse reads to generate a *de novo* assembly using pyrad v.3.0.3 (Eaton 2014). Reads were de-multiplexed and assigned to individuals based on barcodes, allowing for one mismatch. We replaced base calls of Q < 20 with an ambiguous base (N) and discarded sequences containing more than 4 ambiguities. We used 85% clustering similarity as a threshold to align the reads into loci. Additional filtering parameters were set to allow for a maximum number of SNPs to be called: retaining clusters with a minimum depth of sequence coverage (Mindepth) > 5 and a locus coverage (MinCov) > 10, a maximum number of SNPs per RAD-tag of 15 (maxSNP=15), and a maximum proportion of individuals with

Table 1. Sample information and summary statistics for each lineage used in this study

Species	Island	Population	Abbreviation	N	N_{Loci}	Missing	Ar	H_o	H_e	G_{is}	$G_{is,lo}$	$G_{is,hi}$	π	π Lower 95% CI	π Upper 95% CI
C. darwini	Santiago	Santiago	AGO	10	13,952	6.5	1.461	0.131	0.142	0.136	0.128	0.143	0.1504	0.1504	0.1504
C. guntheri	Isabela	Cazuela	CAZ	10	10,935	8.6	1.375	0.116	0.129	0.162	0.153	0.170	0.1373	0.1373	0.1373
C. donfaustoi	Santa Cruz	Cerro Fatal	CF	7	6,825	7.7	1.233	0.080	0.082	0.112	0.100	0.124	0.0886	0.0886	0.0887
C. porteri	Santa Cruz	La Caseta	CRU	10	10,064	7.5	1.339	0.096	0.111	0.189	0.181	0.198	0.1173	0.1173	0.1173
C. hoodensis	Española	Española	ESP	10	7,560	6.9	1.221	0.076	0.078	0.082	0.071	0.092	0.0820	0.0820	0.0820
C. vicina	Isabela	West Cerro Azul	WCA	10	10,871	7.7	1.369	0.110	0.125	0.179	0.171	0.187	0.1329	0.1329	0.1329
C. becki	Isabela	Piedras Blancas	PBL	10	14,541	8.9	1.507	0.147	0.163	0.162	0.155	0.169	0.1731	0.1730	0.1731
C. becki	Isabela	Puerto Bravo	PBR	10	12,097	8.7	1.409	0.110	0.136	0.249	0.241	0.258	0.1436	0.1436	0.1437
C. duncanensis	Pinzón	Pinzón	PZ	10	8,743	5.9	1.283	0.09	0.097	0.133	0.124	0.143	0.1026	0.1026	0.1026
C. chathamensis	San Cristóbal	San Cristóbal	SCR	10	9,162	6.9	1.315	0.100	0.112	0.166	0.157	0.175	0.1184	0.1184	0.1185
C. vanderburghi	Isabela	Volcano Alcedo	VA	10	11,657	7.3	1.399	0.129	0.137	0.120	0.112	0.127	0.1450	0.1450	0.1450
C. microphyes	Isabela	Volcano Darwin	VD	10	10,723	8.8	1.369	0.112	0.129	0.190	0.181	0.199	0.1364	0.1364	0.1364

Number of individuals (N), polymorphic loci (N_{Loci}), allelic richness (Ar), observed heterozygosity (H_o), expected heterozygosity (H_e), inbreeding coefficient (G_{is}), and nucleotide diversity (π) along with 95% CIs.

shared heterozygote sites of 20% (MaxSH=p.20). This last parameter was used to minimize the effect of paralogous loci being incorrectly grouped together (Eaton 2014).

To reduce the effects of linkage disequilibrium on downstream analyses, we subsampled the loci to have only one (the first) SNP per RAD-tag using VCFtools version v0.1.14 (Danecek et al. 2011). Loci were further screened in PLINK version 1.90b3 (Chang et al. 2015) to have a minimum minor allele frequency (MAF) of 0.05 and be genotyped in at least 80% of samples.

Allelic Diversity and Genetic Differentiation

We characterized genetic diversity summary statistics using GenoDive version 2.0b27 (Meirmans and Van Tienderen 2004) and the package *diveRsity* version 1.9.90 (Keenan et al. 2013) in R version 3.2.2 (R Core Team 2015). Within GenoDive, we measured inbreeding via lineage mean G_{is} values, which are akin to F_{is} values (Weir and Cockerham 1984). Significance of G_{is} values was assessed using 95% confidence intervals calculated through 10 000 bootstrap replicates across loci. To determine pairwise genetic differentiation among the lineages, we calculated F_{st} values with Arlequin version 3.5.2.2 (Excoffier and Lischer 2010). We also used the R package *dartR* version 0.94 (Gruber et al. 2018) to calculate the proportion of fixed differences between pairs of lineages. Finally, we calculated per lineage nucleotide diversity (π) using VCFtools version 0.1.14 (Danecek et al. 2011) and generated confidence intervals (CIs) by bootstrapping per-SNP π estimates 10 000 times in the R package *bootstrap* version 2017.2 (Tibshirani et al. 2017). Note that the estimates of π were based on the polymorphic sites that passed the filtering described above but may be fixed in individual lineages.

Assignment Tests

We performed assignment tests using k-means clustering of individuals in GenoDive. This method implements clustering via an AMOVA framework such that within-group diversity is minimized while among-group diversity is maximized (Meirmans 2012). Between 2 and 20 genetic clusters (K) were examined, leaving all other parameters at default values. The optimal number of clusters was assessed via Bayesian Information Criterion (BIC) such that the optimal K was the one with the lowest BIC score. However, we explored all K within BIC = 2 of the lowest value as BICs within this range suggest that one value is not a substantively better fit to the data than another (Burnham and Anderson 2004).

Genetic Structure

We examined genetic structure among lineages with 3 different methods, 2 of which do not assume a priori grouping of individuals. First, admixture analyses were conducted using the R package *lea* version 1.2.0 (Frichot and François 2015). This package provides least-squares estimates of individual ancestry proportions (Frichot et al. 2014) among genetic clusters without a priori grouping of individuals. We chose this method as it is optimized for large SNP datasets and is robust to departures from traditional population genetic model assumptions, such as Hardy–Weinberg and linkage equilibrium (Frichot et al. 2014), used by programs such as STRUCTURE (Pritchard et al. 2000). We tested K = 1–17 with 20 replicates of each K. Optimal K values were chosen by examining cross-entropy scores for the point at which this value was minimized. Ancestry matrices from *lea* were processed with CLUMPAK (Kopelman et al. 2015) prior to visualization. Secondly, we examined genetic structure by conducting principal component analysis (PCA) with *lea*, which also does not assume a priori grouping of individuals. The percent

variation explained by each principal component was assessed with Tracy–Widom tests (Patterson et al. 2006). Finally, we used discriminant analysis of principal components (DAPC). This method combines discriminant analysis with PCA to maximize differentiation among groups while minimizing within group variation (Jombart et al. 2010), in our case sampling locations. DAPC analyses were undertaken using the R package *adegenet* version 2.0.1 (Jombart 2008; Jombart & Ahmed 2011) and the "optim.a.score" function to determine the number of principal components to retain.

Comparison to Previous Estimates of Genetic Differentiation

Previous studies have estimated differentiation among Galapagos giant tortoises using both mitochondrial DNA and microsatellite markers (Caccone et al. 1999, 2002; Ciofi et al. 2002, 2006; Beheregaray et al. 2003a, 2003b, 2004; Russello et al. 2005, 2007; Poulakakis et al. 2008, 2012; Edwards et al. 2013, 2014; Garrick et al. 2014, 2012). We used a sub-set of the data from Garrick et al. (2015; dryad doi: 10.5061/dryad.7h8q2) to compare the newly derived SNP estimates to previous estimates, which includes the same lineages for which we obtained SNPs. For the microsatellite loci, we calculated diversity statistics using GenoDive and *diveRsity*, and for mtDNA, we calculated nucleotide diversity (π) with DnaSP v5 (Librado and Rozas 2009). In addition, we calculated pairwise F_{st} values among lineages in Arlequin for both mtDNA (~700 bp of the control region; using the distance method of Tamura and Nei (Tamura and Nei 1993)) and microsatellite loci (N = 12). Comparisons among F_{st} values was done using Mantel tests as implemented in the R package *vegan* version 2.4–6 (Oksanen et al. 2018) with significance assessed by 9999 permutations.

Results

After de-multiplexing and filtering for quality and ambiguous barcodes, we retained a total of 3 094 399 092 reads (approximately 15 to 58 million reads per individual). The *de novo* assembly of the data resulted in 48 004 056 RAD-tags (approximately 320 000–465 000 per individual). The mean coverage per locus per individual was 12X (minimum 9; maximum 15). The total number of SNP loci found was 973 321. Following filtering to retain one SNP per RAD-tag (543 761 loci removed), presence in at least 80% of samples (296 675 loci removed; the distribution of missing data per-locus can be seen in Supplementary Figure 1), and minimum MAF of 5% (106 331 loci removed), the final dataset included 26 554 SNPs that were used in all downstream analyses.

The number of polymorphic SNPs within each lineage (Table 1) ranged from 6825 in CF (*C. donfaustoi*) to 14 541 in the PBL population (population of *C. becki*; see Table 1 for lineage abbreviations). Mean observed heterozygosity per lineage ranged from 0.08 in CF to 0.15 in PBL. For all lineages, inbreeding coefficients (G_{is}) ranged from 0.082 for ESP to 0.249 for PBL. As expected when using thousands of markers, G_{is} were significantly >0. The proportion of fixed alleles between lineage pairs ranged from 0% between AGO and the 2 *C. becki* populations to 4.32% between CF and ESP (mean ± SD = 0.93% ± 1.03%; Supplementary Table 1). Per-lineage estimates of π were similar (mean ± SD = 0.1273 ± 0.0265), with the lowest diversity seen in ESP and the highest in the PBL population of *C. becki* (Table 1). However, we note that these estimates of π are based solely on polymorphic sites, so they should be considered as relative values among lineages rather than exact estimates. All pairwise F_{st} estimates were also significantly >0 (Supplementary

Table 2), ranging from 0.108 between AGO (*C. darwini*) and PBR (population of *C. becki*), to 0.660 between CF (*C. donfaustoi*) and ESP (*C. hoodensis*).

The lowest BIC score for *k*-means clustering of individuals was at *K* = 8 (Supplementary Table 3). However, BIC scores for *K* = 7–12 were all within 2 of one another (Supplementary Table 3), suggesting that any one *K* is not a substantively better fit to the data within this range of values. However, examination of *K* = 12 gave biologically meaningful results, with the majority of samples being assigned to the geographic location in which they were sampled (Supplementary Table 4). The only exception to this pattern was 2 samples from one of the 2 genetically differentiated populations from *C. becki* (PBR), which were assigned to the other population of the same species (PBL).

For the *lea* admixture analyses, cross entropy values were similar for *K* = 7–12 (<0.006 from the lowest value at *K* = 8; Supplementary Table 3). At *K* = 12, however, all sampling locations were differentiated with only 2 individuals sampled in PBR (one of the 2 *C. becki* populations), and one individual each from AGO (*C. darwini*) and CF (*C. donfaustoi*) appearing admixed (Figure 2A). The admixed individuals from PBR were sampled as part of a survey of Wolf Volcano on Isabela in 2008 (Garrick et al. 2012; Edwards et al. 2013) and do not have majority membership to any one cluster, with ~25% of their cluster membership assigned to both AGO and PBL, and the remainder split among other lineages. The admixed sample from AGO shared ancestry with the lineage from VA (*C. vanderburghi*), whereas the one from CF (*C. donfaustoi*) shared ancestry with samples from the other species occurring on the same island Santa Cruz (CRU; *C. porteri*). See Supplementary Material for description of clustering at each *K* from 7 to 12, which showed hierarchical patterns of differentiation among the lineages.

Within the principal component (PC) analysis, PCs 1 and 2 differentiated the majority of lineages explaining 12.4% and 8.5% of the variance in the dataset, respectively (Supplementary Figure 3A). However, whereas species from central and southern Isabela (CAZ, *C. guntheri*; WCA, *C. vicina*; VA, *C. vanderburghi*; and VD, *C. microphyes*) were well separated from the other lineages, they were not clearly distinct from one another on these axes or PCs 3 and 4, when considering all species (Supplementary Figure 3B). Therefore, we repeated the analysis including only samples from southern Isabela species. This analysis clearly differentiated the 4 taxa from one another (Supplementary Figure 4).

Based on the optim.a.score, we retained 9 PCs for the DAPC analyses. Here, the first 2 PCs differentiated most lineages, except for the 5 species from Isabela Island and the species from Santiago Island (*C. darwini*; AGO; Figure 2B). However, across PCs 3 and 4, all species became more clearly distinct from one another, although AGO (*C. darwini*) and PBL (population of *C. becki*) showed overlaps in their grouping (Figure 2B).

Mean observed heterozygosity values per lineage estimated using SNPs were correlated with those from microsatellites (Pearson's $r = 0.70$; $t = 3.09$, df = 10, $P = 0.01$; Supplementary Table 5), but not to nucleotide diversity of the mtDNA control region (Pearson's $r = 0.07$; $t = 0.23$, df = 10, $P = 0.82$; Supplementary Table 5). The Mantel test showed that the pairwise F_{st} values between lineages calculated from microsatellites were strongly correlated with those from SNPs ($r = 0.88$, $P < 0.0001$). However, F_{st} estimates from SNPs were, on average, 72% larger than those from microsatellites, except for between AGO (*C. darwini*) and PBR (population of *C. becki*), where the SNP estimate was 7% smaller (Supplementary Table 2; Figure 3A). Similarly, pairwise F_{st} values calculated from mtDNA control region were all significantly different than zero (Supplementary Table 6), and the magnitudes of the estimates were correlated to those from SNPs (Mantel test $r = 0.60$, $P < 0.0001$; Figure 3B).
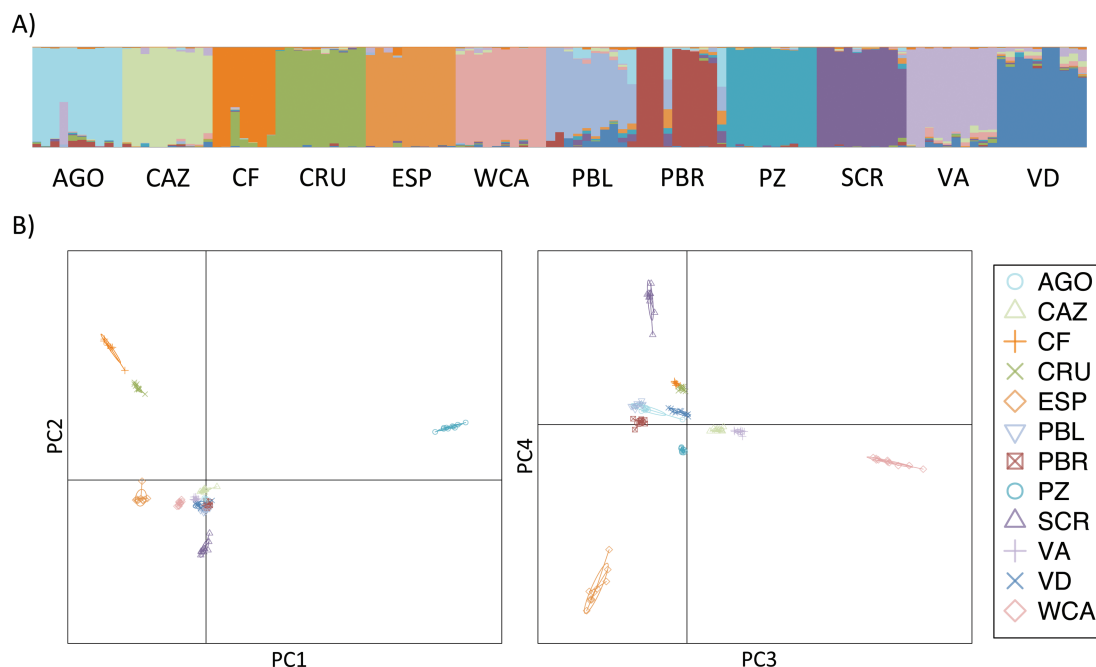

Figure 2. (A) Ancestry bar plot from *lea* (Frichot and François 2015) for *K* = 12 based on 26 554 SNPs for the 117 samples used in this study. The 12 genetic clusters are identified with different colors. Each individual is represented as a vertical bar, with the proportion of colors representing their genetic assignment to the clusters. (B) DAPC plot from *adegenet* for the 117 samples used in this study of diversity and divergence among extant Galapagos giant tortoise lineages using 26 554 SNPs. The left panel shows the results of the first 2 PCs, while the right panel shows the results for PCs 3 and 4. Each individual is represented as a point and populations are surrounded by 95% inertial ellipses. Lineage abbreviations are as in Figure 1 and Table 1.

Discussion

Here, we characterized a novel set of >26000 SNP loci among extant Galapagos giant tortoises using ddRAD sequencing. Analysis of these loci showed that genetic diversity varied among lineages (Table 1), a result consistent with previous mtDNA and microsatellite analyses (Beheregaray et al. 2003b; Poulakakis et al. 2012; Garrick et al. 2015) and not unexpected given the distinct demographic histories of the different lineages. Some lineages have experienced drastic changes in population sizes due to volcanic eruptions and sea level changes, which altered habitats and connectivity among islands (Beheregaray et al. 2003b; Poulakakis et al. 2012), as well as intraspecific competition and niche partitioning (Hunter et al. 2013). Others were dramatically reduced in size in much more recent times due to human exploitation, as well as competition with, and predation by, invasive species (MacFarland et al. 1974; Milinkovitch et al. 2004, 2013; Cayot 2008; Jensen et al. 2015). The newly described species endemic to eastern Santa Cruz Island (*C. donfaustoi*, CF: Table 1) has the lowest number of polymorphic loci and genetic
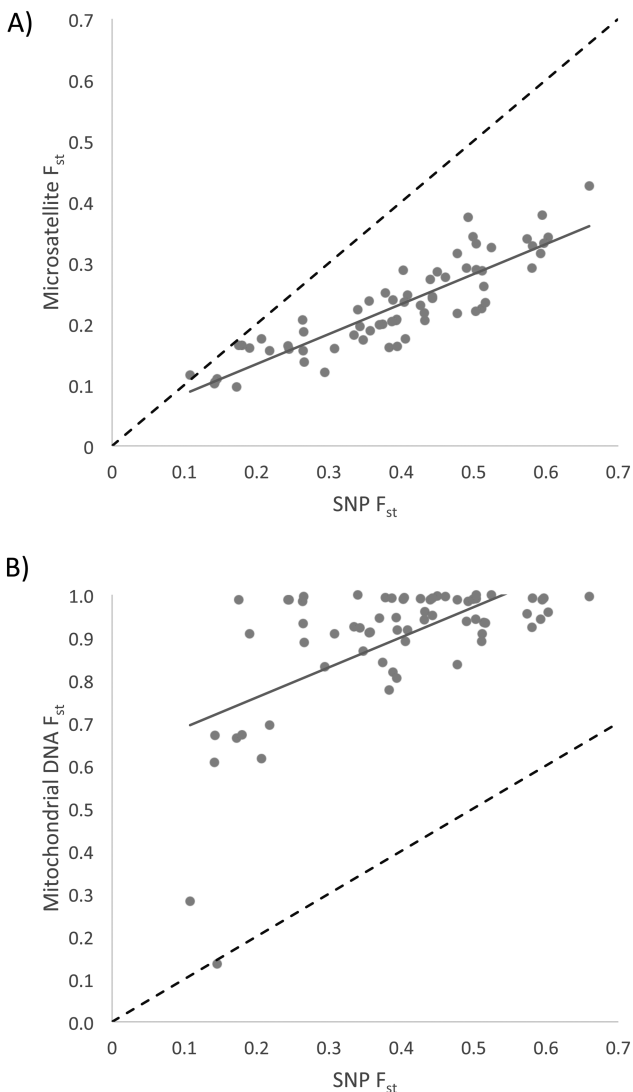
A)

B)

Figure 3. Scatter plot comparison of pairwise F_{st} estimates for 26554 SNPs to 12 microsatellite loci (A) or ~700 bp of the mtDNA control region (B) among extant Galapagos giant tortoise lineages. The solid line represents a linear trend line through the data; the dashed line represents a 1:1 line.

diversity, consistent with genetic estimates from previous work that showed the species diverged only ~0.43 million years ago and persisted as a small, geographically restricted population for most of its history (Poulakakis et al. 2012, 2015; Garrick et al. 2015). The next 2 lineages with the lowest diversity are the species on Espanola (ESP) and Pinzon (PZ) islands (*C. hoodensis* and *C. duncanensis*, respectively), which is again consistent with known population declines that reduced ESP to 15 individuals (MacFarland et al. 1974; Milinkovitch et al. 2004, 2013) and PZ to ~150 individuals (MacFarland et al. 1974; Cayot 2008). In contrast, the 3 lineages with the highest levels of diversity are PBR (population of *C. becki*), AGO (*C. darwini*), and PBL (population of *C. becki*), respectively. These 3 lineages are closely related to one another (Poulakakis et al. 2012), with the PBL and PBR populations founded via independent colonization of Wolf Volcano by individuals from Santiago Island (AGO) approximately 199000 and 53000 years ago, respectively (Garrick et al. 2014). Diversity may have been maintained in these populations due to lower levels of tortoise exploitation and no historical population collapse on Wolf Volcano and Santiago Island.

Our SNP analysis showed positive G_{is} values for all lineages (Table 1). Although this is often taken as an indication of inbreeding, positive G_{is} values can also be generated by processes such as genetic drift or Wahlund effects (Wahlund 1928). In this case, Wahlund effect is an unlikely explanation given the island-based distribution of this complex and the lack of population subdivision within species (Ciofi et al. 2002; Garrick et al. 2015). Rather, population bottlenecks are likely responsible for the positive G_{is} values, given that previous M-ratio tests based on microsatellite loci suggested that demographic contractions have occurred in all lineages of Galapagos giant tortoises (Garrick et al. 2015). These tests, which can detect population reductions within the last 10–50 generations (Peery et al. 2012), date these demographic changes to a time period that coincides with habitat changes and exploitation by mariners and early colonists that started in the late 1700s (Márquez et al. 2004; Garrick et al. 2015).

Across our clustering and assignment analyses, we found consistent evidence for the distinctiveness of 12 genetic lineages of Galapagos giant tortoises. These lineages directly corresponded to 11 recognized species plus the 2 genetically distinct, yet morphologically cryptic, *C. becki* populations (PBL and PBR) on Wolf Volcano (Isabela Island). Within each of these analyses we observed hierarchical clustering (Supplementary Figures 2–4). However, it is important to note that such hierarchical patterns and the order in which different groups separate from each other is largely a function of methods that detect the most prominent structure first, rather than necessarily being reflective of their phylogenetic relationships (Vähä et al. 2007; Janes et al. 2017). Such hierarchical patterns have been also reported for radiations or range expansions, where recovery of all groups in a single analysis is not necessarily expected, especially for groups that have diverged more recently (Præbel et al. 2013; Jensen et al. 2014). In this light, we advocate continued use of probabilistic approaches, such as BIC scores or cross-entropy values, to assess fit of assignments, along with clear presentation of the criterion used (Janes et al. 2017).

The magnitudes of differentiation among lineages were broadly concordant with previous estimates from nuclear microsatellite loci, though SNP-based F_{st} values were higher than the ones based on microsatellite loci (Figure 3A). This pattern is expected when comparing values generated from bi-allelic versus multi-allelic markers (Payseur and Jing 2009; Putman & Carbone 2014). The magnitude of differentiation measured by SNPs was also correlated to that from mtDNA control region sequences (Figure 3B), although estimates from mtDNA were much higher. This pattern is expected

given the reduced effective population size and increased coalescence rate of the mitochondrial genome compared to the nuclear genome. Similar levels of SNP-based genetic differentiation were detected in Gaughran et al. (in press) that conducted genome-wide analyses of 3 of the 11 extant species (PBL, population of *C. becki*; CRU, *C. porteri*; and VA, *C. vandenburghi*) analyzed here, as they had been the sole focus of 2 studies that disputed the occurrence of genomic differentiation among Galapagos giant tortoises based on transcriptomic data collected for a total of 5 individuals (Loire et al. 2013; Loire and Galtier 2017). Interestingly, reanalysis of Loire et al.'s (2013) data by Gaughran et al. (in press) was also able to recover significant population differentiation.

The new SNP-based toolkit provides higher resolution to detect individuals with admixed ancestry than previously used genetic markers. Low levels of genetic admixture among species from different islands have been previously reported. Most cases were found among samples from a few geographic locations and attributed to relatively recent human activities, such as translocations (Russello et al. 2007; Poulakakis et al. 2012; Edwards et al. 2013; Garrick et al. 2014; Miller et al. 2017) or changes in land use in the case of allopatric species living on the same island (Russello et al. 2005; Poulakakis et al. 2015). In this study, out of 117 samples, only 4 individuals [PBL (n = 2; population of *C. becki*); AGO (n = 1; *C. darwini*); CF (n = 1; *C. donfaustoi*)] showed consistent signs of admixture (Figure 2A; Supplementary Figure 2). Several processes can generate admixture patterns when using assignment programs such as *lea* or STRUCTURE, including incomplete linage sorting, secondary contact, and isolation by distance (Frantz et al. 2009; Falush et al. 2016). In this analysis we see evidence for the first 2 processes. For the AGO and PBL individuals, the genomic signal of admixture is likely the result of incomplete lineage sorting due to shared ancestry given the recent origin of the PBL population from tortoises from Santiago Island, where AGO currently resides (~200 000 years ago, Poulakakis et al. 2012). For the one admixed individual in the *C. donfaustoi* population from Cerro Fatal (CF), admixture is more likely due to human-mediated land use changes on the island that have facilitated contact between the 2 allopatric species living on the island (*C. porteri* and *C. donfaustoi*; Russello et al. 2005; Poulakakis et al. 2015). Thus, detection of admixture in these 4 individuals speaks to the power of genomic techniques to reveal more detailed ancestry assignments than mtDNA or microsatellites.

Significance and Future Directions

Given that the results presented here largely parallel those found with "traditional genetic markers" one can wonder if the investment to develop genomic resources was necessary (McMahon et al. 2014). Indeed, conservationists often have limited resources and face trade-offs among competing priorities (e.g. genetic/genomic research, habitat conservation, or other on-the-ground management actions) in how to distribute funds. However, we feel that these markers will provide a valuable conservation resource for this iconic reptile group. Tens of thousands of genomic loci offer unprecedented accuracy when inferring patterns and levels of individual and population differentiation, diversity, and genetic admixture that can all be used to inform conservation strategies. In addition, the SNP database we have assembled and present herein will be a resource for future efforts to advance understanding of tortoise evolution and species radiations, including constructing detailed phylogenies (Rubin et al. 2012; Eaton et al. 2017), inferring demographic history (Trucchi et al. 2014), and examining the process of speciation in this rapid radiation (Bultin 2010; Gante et al. 2016).

Future work will include incorporation of population level representation of 2 of the extinct species into the current SNP database. Previous studies have shown that these historical samples are not only key to accurate inferences of the evolutionary history of this radiation (Caccone et al. 1999, 2002; Poulakakis et al. 2012), but are also instrumental in identifying living individuals of high conservation value with mixed ancestry to 2 recently extinct species (Russello et al. 2007, 2010, Poulakakis et al. 2008, 2012; Garrick et al. 2012; Edwards et al. 2013; Miller et al. 2017).

Supplementary Material

Supplementary material can be found at https://academic.oup.com/jhered/.

Data Accessibility

Sequencing reads have been deposited on NCBI Short Read Archive under the bioproject PRJNA407643. SNP genotypes have been deposited on Dryad https://doi.org/10.5061/dryad.dm3hb40. Microsatellite and mtDNA sequences used were previously published and available on Dryad doi: 10.5061/dryad.7h8q2.

Author Contributions

DLE conducted laboratory analyses and prepared the sequencing libraries. JMM, MCQ, DLE and DARE conducted bioinformatic and data analyses. MAR, ELJ, and AC provided technical guidance and conceptual advice for data analyses. JMM drafted the manuscript with assistance from MCQ, DLE, ELJ, MAR, and AC. JMM, MCQ, DLE, DARE, ELJ, MAR, JPG, WT, DR, and AC discussed the results and implications and commented on the manuscript at all stages.

References

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, Johnson EA. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One*. 3:e3376.

Beheregaray LB, Ciofi C, Caccone A, Gibbs JP, Powell JR. 2003a. Genetic divergence, phylogeography and conservation units of giant tortoises from Santa Cruz and Pinzon, Galapagos Islands. *Conservation Genetics*. 4:31–46.

Beheregaray LB, Ciofi C, Geist D, Gibbs JP, Caccone A, Powell JR. 2003b. Genes record a prehistoric volcano eruption in the Galápagos. *Science*. 302:75.

Beheregaray LB, Gibbs JP, Havill N, Fritts TH, Powell JR, Caccone A. 2004. Giant tortoises are not so slow: rapid diversification and biogeographic consensus in the Galapagos. *Proceedings of the National Academy of Sciences of the United States of America*. 101:6514–6519.

Bosse M, Megens HJ, Madsen O, Crooijmans RP, Ryder OA, Austerlitz F, Groenen MA, de Cara MA. 2015. Using genome-wide measures of

coancestry to maintain diversity and fitness in endangered and domestic pig populations. *Genome Res*. 25:970–981.

Bultin RK. 2010. Population genomics and speciation. *Genetica*. 138:409–418.

Burnham KP, Anderson DR. 2004. Multimodel Inference: Understanding AIC and BIC in Model Selection. *Sociological Methods & Research*. 33:261–304.

Caccone A, Gentile G, Gibbs JP, Frirts TH, Snell HL, Betts J, Powell JR. 2002. Phylogeography and history of giant Galápagos tortoises. *Evolution*. 56:2052–2066.

Caccone A, Gibbs JP, Ketmaier V, Suatoni E, Powell JR. 1999. Origin and evolutionary relationships of giant Galapagos tortoises. *Proceedings of the National Academy of Sciences of the United States of America*. 96:13223–13228.

Campagna L, Gronau I, Silveira LF, Siepel A, Lovette IJ. 2015. Distinguishing noise from signal in patterns of genomic divergence in a highly polymorphic avian radiation. *Mol Ecol*. 24:4238–4251.

Campbell K, Donlan CJ, Cruz F, Carrion V. 2004. Eradication of feral goats Capra hircus from Pinta Island, Galápagos, Ecuador. *Oryx*. 38:328–333.

Cayot L. 2008. The restoration of giant tortoise and land iguana populations in Galapagos. *Galapagos Research*. 65:39–43.

Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*. 4:7.

Ciofi C, Milinkovitch MC, Gibbs JP, Caccone A, Powell JR. 2002. Microsatellite analysis of genetic divergence among populations of giant Galápagos tortoises. *Mol Ecol*. 11:2265–2283.

Ciofi C, Wilson GA, Beheregaray LB, Marquez C, Gibbs JP, Tapia W, Snell HL, Caccone A, Powell JR. 2006. Phylogeographic history and gene flow among giant Galápagos tortoises on southern Isabela Island. *Genetics*. 172:1727–1744.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, *et al.*; 1000 Genomes Project Analysis Group. 2011. The variant call format and VCFtools. *Bioinformatics*. 27:2156–2158.

Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML. 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat Rev Genet*. 12:499–510.

Eaton DA. 2014. PyRAD: assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics*. 30:1844–1849.

Eaton DAR, Spriggs EL, Park B, Donoghue MJ. 2017. Misconceptions on Missing Data in RAD-seq Phylogenetics with a Deep-scale Example from Flowering Plants. *Syst Biol*. 66:399–412.

Edwards DL, Benavides E, Garrick RC, Gibbs JP, Russello MA, Dion KB, Hyseni C, Flanagan JP, Tapia W, Caccone A. 2013. The genetic legacy of Lonesome George survives: Giant tortoises with Pinta Island ancestry identified in Galapagos. *Biological Conservation*. 157:225–228.

Edwards DL, Garrick RC, Tapia W, Caccone A. 2014. Cryptic structure and niche divergence within threatened Galápagos giant tortoises from southern Isabela Island. *Conservation Genetics*. 15:1357–1369.

Eizaguirre C, Baltazar-Soares M. 2014. Evolutionary conservation—evaluating the adaptive potential of species. *Evolutionary Applications*. 7:963–967.

Ellegren H. 2014. Genome sequencing and population genomics in non-model organisms. *Trends Ecol Evol*. 29:51–63.

Excoffier L, Lischer HE. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour*. 10:564–567.

Falush D, Dorp L van, Lawson D. 2016. A tutorial on how (not) to over-interpret STRUCTURE/ADMIXTURE bar plots. *bioRxiv*. doi:10.1101/066431

Frantz AC, Cellina S, Krier A, Schley L, Burke T. 2009. Using spatial Bayesian methods to determine the genetic structure of a continuously distributed population: Clusters or isolation by distance? *Journal of Applied Ecology*. 46:493–505.

Frichot E, François O. 2015. LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution*. 6:925–929.

Frichot E, Mathieu F, Trouillon T, Bouchard G, François O. 2014. Fast and efficient estimation of individual ancestry coefficients. *Genetics*. 196:973–983.

Fritts TH, Snell HL, Cayot L, MacFarland C, Earsom S, Marquez C, Llerena W, Llerena F. 2000. Progress and priorities in research for the conservation of reptiles. *Bulletin de L'Institut Royal Sciences Naturrelles de Belgique*. 70:39–45.

Funk WC, McKay JK, Hohenlohe PA, Allendorf FW. 2012. Harnessing genomics for delineating conservation units. *Trends Ecol Evol*. 27:489–496.

Gante HF, Matschiner M, Malmstrøm M, Jakobsen KS, Jentoft S, Salzburger W. 2016. Genomics of speciation and introgression in Princess cichlid fishes from Lake Tanganyika. *Mol Ecol*. 25:6143–6161.

Garrick RC, Benavides E, Russello MA, Gibbs JP, Poulakakis N, Dion KB, Hyseni C, Kajdacsi B, Márquez L, Bahan S, *et al.* 2012. Genetic rediscovery of an 'extinct' Galápagos giant tortoise species. *Curr Biol*. 22:R10–R11.

Garrick RC, Benavides E, Russello MA, Hyseni C, Edwards DL, Gibbs JP, Tapia W, Ciofi C, Caccone A. 2014. Lineage fusion in Galápagos giant tortoises. *Mol Ecol*. 23:5276–5290.

Garrick RC, Kajdacsi B, Russello MA, Benavides E, Hyseni C, Gibbs JP, Tapia W, Caccone A. 2015. Naturally rare versus newly rare: demographic inferences on two timescales inform conservation of Galápagos giant tortoises. *Ecol Evol*. 5:676–694.

Gaughran SJ, Quinzin MC, Miller JM, Garrick RC, Edwards DL, Russello MA, Poulakakis N, Ciofi C, Beheregaray LB, Caccone A. In press. Theory, practice, and the conservation of Galápagos giant tortoises in the age of genomics. *Evolutionary Applications*.

Gibbs JP, Sterling EJ, Zabala FJ. 2010. Giant Tortoises as Ecological Engineers: A Long-term Quasi-experiment in the Galápagos Islands. *Biotropica*. 42:208–214.

Gruber B, Unmack PJ, Berry OF, Georges A. 2018. dartr: An r package to facilitate analysis of SNP data generated from reduced representation genome sequencing. *Mol Ecol Resour*. 18:691–699.

Guézou A, Trueman M, Buddenhagen CE, Chamorro S, Guerrero AM, Pozo P, Atkinson R. 2010. An extensive alien plant inventory from the inhabited areas of galapagos. *PLoS One*. 5:e10276.

Harper GA, Carrion V. 2011. Introduced rodents in the Galapagos: colonisation, removal and the future. In: Veitch CR, Clout MN, Towns DR, editors. *Island Invasives: Eradication and Management*. Gland, Switzerland: IUCN. 63–66.

Harrisson KA, Pavlova A, Telonis-Scott M, Sunnucks P. 2014. Using genomics to characterize evolutionary potential for conservation of wild populations. *Evol Appl*. 7:1008–1025.

Heslot N, Rutkoski J, Poland J, Jannink JL, Sorrells ME. 2013. Impact of marker ascertainment bias on genomic selection accuracy and estimates of genetic diversity. *PLoS One*. 8:e74612.

Hunter EA, Gibbs JP, Cayot LJ, Tapia W. 2013. Equivalency of Galápagos giant tortoises used as ecological replacement species to restore ecosystem functions. *Conserv Biol*. 27:701–709.

Janes JK, Miller JM, Dupuis JR, Malenfant RM, Gorrell JC, Cullingham CI, Andrew RL. 2017. The K = 2 conundrum. *Mol Ecol*. 26:3594–3602.

Jensen EL, Govindarajulu P, Russello MA. 2014. When the shoe doesn't fit: applying conservation unit concepts to western painted turtles at their northern periphery. *Conservation Genetics*. 15:261–274.

Jensen EL, Tapia W, Caccone A, Russello MA. 2015. Genetics of a head-start program to guide conservation of an endangered Galápagos tortoise (*Chelonoidis ephippium*). *Conservation Genetics*. 16:823–832.

Jombart T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*. 24:1403–1405.

Jombart T, Ahmed I. 2011. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics*. 27:3070–3071.

Jombart T, Devillard S, Balloux F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet*. 11:94.

Keenan K, McGinnity P, Cross TF, Crozier WW, Prodöhl PA. 2013. diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution*. 4:782–788.

Kehlmaier C, Barlow A, Hastings AK, Vamberger M, Paijmans JLA, Steadman DW, Albury NA, Franz R, Hofreiter M, Fritz U. 2017. Tropical ancient DNA reveals relationships of the extinct Bahamian giant tortoise *Chelonoidis alburyorum*. *Proceedings of the Royal Society B: Biological Sciences*. 284.

Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I. 2015. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Mol Ecol Resour*. 15:1179–1191.

Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*. 25:1451–1452.

Loire E, Chiari Y, Bernard A, Cahais V, Romiguier J, Nabholz B, Lourenço JM, Galtier N. 2013. Population genomics of the endangered giant Galápagos tortoise. *Genome Biol*. 14:R136.

Loire E, Galtier N. 2017. Lacking conservation genomics in the giant Galápagos tortoise. *bioRxiv*.

MacFarland CG, Villa J, Toro B. 1974. The Galápagos giant tortoises (*Geochelone elephantopus*) Part I: Status of the surviving populations. *Biological Conservation*. 6:118–133.

Márquez C, Wiedenfeld D, Snell H, *et al*. 2004. Estado actual de las poblaciones de tortugas terrestres gigantes (Geochelone spp., Chelonia: Testudinae) en las islas Galápagos. *Ecología Aplicada*. 3:98–111.

McMahon BJ, Teeling EC, Höglund J. 2014. How and why should we implement genomics into conservation? *Evol Appl*. 7:999–1007.

McTavish EJ, Hillis DM. 2015. How do SNP ascertainment schemes and population demographics affect inferences about population history? *BMC Genomics*. 16:266.

Meirmans PG. 2012. AMOVA-based clustering of population genetic data. *J Hered*. 103:744–750.

Meirmans PG, Van Tienderen PH. 2004. GENOTYPE and GENODIVE: Two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes*. 4:792–794.

Milinkovitch MC, Kanitz R, Tiedemann R, Tapia W, Llerena F, Caccone A, Gibbs JP, Powell JR. 2013. Recovery of a nearly extinct Galápagos tortoise despite minimal genetic variation. *Evol Appl*. 6:377–383.

Milinkovitch MC, Monteyne D, Gibbs JP, Fritts TH, Tapia W, Snell HL, Tiedemann R, Caccone A, Powell JR. 2004. Genetic analysis of a successful repatriation programme: giant Galapagos tortoises. *Proceedings of the Royal Society B-Biological Sciences*. 271:341–345.

Milinkovitch MC, Monteyne D, Russello M, Gibbs JP, Snell HL, Tapia W, Marquez C, Caccone A, Powell JR. 2007. Giant Galápagos tortoises; molecular genetic analyses identify a trans-island hybrid in a repatriation program of an endangered taxon. *BMC Ecol*. 7:2.

Miller JM, Quinzin MC, Poulakakis N, Gibbs JP, Beheregaray LB, Garrick RC, Russello MA, Ciofi C, Edwards DL, Hunter EA, *et al*. 2017. Identification of genetically important individuals of the rediscovered Floreana Galápagos giant tortoise (*Chelonoidis elephantopus*) provide founders for Species Restoration Program. *Sci Rep*. 7:11471.

Oksanen J, Blanchet GF, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, *et al*. 2018. *vegan: Community Ecology Package*. R package version 2.4–6.

Patterson N, Price AL, Reich D. 2006. Population structure and eigenanalysis. *PLoS Genet*. 2:e190.

Payseur BA, Jing P. 2009. A genomewide comparison of population structure at STRPs and nearby SNPs in humans. *Mol Biol Evol*. 26:1369–1377.

Peery MZ, Kirby R, Reid BN, Stoelting R, Doucet-Bëer E, Robinson S, Vásquez-Carrillo C, Pauli JN, Palsbøll PJ. 2012. Reliability of genetic bottleneck tests for detecting recent population declines. *Mol Ecol*. 21:3403–3418.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE. 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One*. 7:e37135.

Phillips RB, Wiedenfeld DA, Snell HL. 2012. Current status of alien vertebrates in the Galápagos Islands: invasion history, distribution, and potential impacts. *Biological Invasions*. 14:461–480.

Porter D. 1822. *Journal of a cruise made to the Pacific Ocean*. New York: Wiley & Halsted.

Poulakakis N, Edwards DL, Chiari Y, Garrick RC, Russello MA, Benavides E, Watkins-Colwell GJ, Glaberman S, Tapia W, Gibbs JP, *et al*. 2015. Description of a new Galapagos giant tortoise species (Chelonoidis; Testudines: Testudinidae) from Cerro Fatal on Santa Cruz Island. *PLoS One*. 10:e0138779.

Poulakakis N, Glaberman S, Russello M, Beheregaray LB, Ciofi C, Powell JR, Caccone A. 2008. Historical DNA analysis reveals living descendants of an extinct species of Galapagos tortoise. *Proceedings of the National Academy of Sciences of the United States of America*. 105:15464–15469.

Poulakakis N, Russello M, Geist D, Caccone A. 2012. Unravelling the peculiarities of island life: vicariance, dispersal and the diversification of the extinct and extant giant Galápagos tortoises. *Mol Ecol*. 21:160–173.

Præbel K, Knudsen R, Siwertsson A, Karhunen M, Kahilainen KK, Ovaskainen O, Ostbye K, Peruzzi S, Fevolden SE, Amundsen PA. 2013. Ecological speciation in postglacial European whitefish: rapid adaptive radiations into the littoral, pelagic, and profundal lake habitats. *Ecol Evol*. 3:4970–4986.

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

Putman AI, Carbone I. 2014. Challenges in analysis and interpretation of microsatellite data for population genetic studies. *Ecol Evol*. 4:4399–4428.

R Core Team. 2015. *R: A language and environment for statistical computing, reference index version 3.2.2*.

Rubin BE, Ree RH, Moreau CS. 2012. Inferring phylogenies from RAD sequence data. *PLoS One*. 7:e33394.

Russello MA, Beheregaray LB, Gibbs JP, Fritts T, Havill N, Powell JR, Caccone A. 2007. Lonesome George is not alone among Galápagos tortoises. *Curr Biol*. 17:R317–R318.

Russello MA, Glaberman S, Gibbs JP, Marquez C, Powell JR, Caccone A. 2005. A cryptic taxon of Galápagos tortoise in conservation peril. *Biol Lett*. 1:287–290.

Russello MA, Poulakakis N, Gibbs JP, Tapia W, Benavides E, Powell JR, Caccone A. 2010. DNA from the past informs ex situ conservation for the future: an "extinct" species of Galápagos tortoise identified in captivity. *PLoS One*. 5:e8683.

Tamura K, Nei M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol*. 10:512–526.

Tibshirani R, Leisch F, Kostyshak S. 2017. *bootstrap: Functions for the Book "An Introduction to the Bootstrap."*

Townsend CH. 1925. The Galápagos tortoises in relation to the whaling industry: A study of old logbooks. *Zoologica*. 4:55–135.

Trucchi E, Gratton P, Whittington JD, Cristofari R, Le Maho Y, Stenseth NC, Le Bohec C. 2014. King penguin demography since the last glaciation inferred from genome-wide data. *Proceedings of the Royal Society B: Biological Sciences*. 281.

Turtle Taxonomy Working Group. 2017. Turtles of the world: annotated checklist and atlas of taxonomy, synonymy, distribution, and conservation status. In: Rhodin AGJ, Iverson JB, van Dijk PP, Saumure RA, Buhlmann KA, Pritchard PCH, Mittermeier RA, editors. *Conservation biology of freshwater turtles and tortoises: a compilation project of the IUCN/SSC tortoise and freshwater turtle specialist group*. 8th ed. Chelonian Research Monographs 7. pp. 1–292. doi: 10.3854/crm.7.checklist.atlas.v8.2017.

Vähä JP, Erkinaro J, Niemelä E, Primmer CR. 2007. Life-history and habitat features influence the within-river genetic structure of Atlantic salmon. *Mol Ecol*. 16:2638–2654.

Viengkone M, Derocher AE, Richardson ES, Malenfant RM, Miller JM, Obbard ME, Dyck MG, Lunn NJ, Sahanatien V, Davis CS. 2016. Assessing polar bear (*Ursus maritimus*) population structure in the Hudson Bay region using SNPs. *Ecol Evol*. 6:8474–8484.

Wagner CE, Keller I, Wittwer S, Selz OM, Mwaiko S, Greuter L, Sivasundar A, Seehausen O. 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.

Wahlund S. 1928. Zusammensetzung von populationen und korrelationserscheinungen vom standpunkt der vererbungslehre aus betrachtet. *Hereditas*. 11:65–106.

Wayne RK, Shaffer HB. 2016. Hybridization and endangered species protection in the molecular era. *Mol Ecol*. 25:2680–2689.

Weir BS, Cockerham CC. 1984. Estimating f-statistics for the analysis of population structure. *Evolution*. 38:1358–1370.

Whiteley AR, Fitzpatrick SW, Funk WC, Tallmon DA. 2015. Genetic rescue to the rescue. *Trends Ecol Evol*. 30:42–49.