Origin and population history of a recent colonizer, the yellow warbler in Galápagos and Cocos Islands

J. A. CHAVES*, P. G. PARKER‡§ & T. B. SMITH*

*Center for Tropical Research, Institute of the Environment, University of California, Los Angeles, CA, USA
‡Department of Ecology & Evolutionary Biology, University of California, Los Angeles, CA, USA
§WildCare Institute, Saint Louis Zoo, St. Louis, MO, USA
*

Center for Tropical Research, Institute of the Environment, University of California, Los Angeles, CA, USA

Introduction

A central goal in evolutionary biology is to understand the origins of biological diversity and the associated factors that promote speciation. The study of island species has provided many important insights, by showing how microevolutionary processes may drive the early stages of diversification (Clegg et al., 2002a,b; Emerson, 2002; Arbogast et al., 2006; Warren et al., 2006; Illera et al., 2007; Ricklefs & Bermingham, 2007; Grant & Grant, 2008; Phillimore et al., 2008; Milá et al., 2010). Some of the most studied insular taxa are those in the Galápagos archipelago. Extensive molecular studies carried out on its endemic birds have provided temporal resolution for divergence times and evolutionary trajectories. For example, the ancestor of Darwin’s finches is estimated to have arrived about two to three million years ago (Grant & Grant, 2008) – Galápagos mockingbirds two millions years ago (Arbogast et al., 2006), whereas the ancestors of Galápagos hawks, magnificent frigatebirds and yellow warblers, are estimated to have colonized the archipelago <300 000 years ago (Bollmer et al., 2006; Browne et al., 2008; Hailey et al., 2010). Yet, the evolutionary outcomes for these groups have been surprisingly variable: at one extreme is the radiation of finches with more than thirteen species and four species of mockingbirds and at the other limited intraspecific variation in morphology among the populations of Galápagos doves (Santiago-Alarcón et al., 2006). This degree of variation in evolutionary outcomes of colonizing taxa is not new to archipelagos. The Hawaiian honeycreepers and thrushes are examples of ancestors arriving to Hawaii at similar
times, yet having undergone completely different evolutionary trajectories, one leading to yet another spectacular radiation of more than 50 species of honeycreepers, whereas thrushes have evolved into a mere four species (Lovette et al., 2002). These patterns raise important questions about island speciation including. How much time after colonization is required for genetic and morphological variation to arise in natural populations? What are the respective roles environments and geographic features play in morphologic diversification? How do different taxonomic groups respond to the same amount of time since isolation and novel environmental conditions? A good starting point to address these questions is to study young island species. Such studies provide insight into microevolutionary processes in diversification, and factors, such as how time since colonization, geographic isolation and environmental conditions might influence genetic and phenotypic divergence (Clegg et al., 2002a,b; Emerson, 2002; Warren et al., 2006; Arbogast et al., 2006; Ricklefs & Bingham, 2007; Illera et al., 2007; Grant & Grant, 2008).

Here, we examine one such recent colonization to the Galápagos, the endemic yellow warbler. The yellow warbler arrived to the Galapagos between 10 000 and 300 000 year (Brown et al., 2008) allowing one to examine whether time since colonization would be sufficient to produce morphological divergence as might be expected based on other young island systems (Clegg et al., 2002a). Yellow warblers are commonly found on almost every island in the archipelago and across steep environmental gradients from dry forest in the lowlands to wet cloud forests in the highlands. Under these circumstances, phenotypic variation might potentially be expected to arise first, by drift alone among isolated islands if gene flow is reduced and second, within islands if selective forces operate, despite gene flow, along an environmental gradient (Endler, 1977; Grant et al., 1985; Smith et al., 1997; Schneider & Moritz, 1999; McCormack & Smith, 2008).

The objectives of this study are threefold: (i) to identify the mainland source populations of aureola yellow warblers and confirm arrival time estimates; (ii) to quantify genetic population differentiation and investigate the probable colonization route that yellow warblers may have used to reach the Galápagos and Cocos Islands; and (iii) to explore the patterns of morphologic variation across islands and habitats to examine evidence for differentiation.

Materials and methods

Geographic sampling, data collection and DNA extraction

Specimens used for genetic analysis originated from two main sources; field trips to the Galápagos and from museum collections such as the Field Museum of Natural History, Chicago (FMC), from which samples were used in the previous yellow warbler phylogeny by Klein & Brown (1994), and Los Angeles County Museum (LACM). All the genetic data were generated at UCLA except for the outgroup Dendroica pensylvanica and three D. petechia aestiva samples for which DNA sequence data were obtained from GenBank (Data S1). Outgroup selection was guided by Klein & Brown (1994). We followed yellow warbler groupings and taxonomy, based on plumage descriptions by Browning (1994) and Olson (1980) and on genetic data produced by Klein & Brown (1994).

In more detail, genetic variation from eleven microsatellites was analysed from 149 individuals of the endemic Dendroica petechia aureola species from nine islands in the Galápagos archipelago (Santa Cruz, San Cristobal, Floreana, Isabela, Santiago, Pinta, Genovesa, Fernandina and Pinzon) and 10 individuals from Cocos Island, off the coast of Costa Rica (total n = 159). Sampling spanned geographic populations to encompass genetic variation among but also within the islands across gradients from four islands characterized by marked altitudinal transition zones (Santa Cruz, Isabela, San Cristobal and Santiago).

DNA was obtained from blood samples collected from live birds in the field (UCLA), from previously extracted DNA from field trips by the co-author (P. Parker UMSL) and from toe pads from museum skins (LACM and FMC). Whole genomic DNA was extracted from blood and toe pads using a commercially available kit (Qiagen™, Valencia, CA, USA), following the manufacturer’s protocol.

DNA amplification, sequencing and aligning

A total of 58 samples corresponding to aureola and 59 specimens representing the three groups of D. petechia were used in this study (Data S1). A 330-bp fragment of the mtDNA control region was amplified and sequenced using species-specific primers DPPdL5 and DPPdH4 (Milot et al., 2000). Based on this preliminary analysis, 18 samples were selected and sequenced for two additional mitochondrial genes, ATPase gene (ATPase 6 and ATPase 8: 852 bp treated as one gene) and NADH dehydrogenase subunit 2 (ND2: 1041 bp). Polymerase chain reactions (PCR) were conducted following the previously established protocols for this species (Milot et al., 2000). Sequencing reaction products were resolved on an ABI 3730 automated sequencer. These mitochondrial sequences have been deposited in GenBank (Data S1).

Model selection and phylogenetic reconstruction

Prior to all the phylogenetic reconstruction analyses, the best-fitting models of molecular evolution were determined for each marker individually as well as for concatenated mtDNA with JModeltest v0.1.0. (Posada, 2008), via the Akaike Information Criterion (AIC, Burnham & Anderson, 2002). Phylogenetic reconstruction...
was carried out using maximum parsimony (MP) and maximum likelihood (ML) performed in PAUP* v4.0b10 (Swofford, 2000) and Bayesian (BA) inference in MrBayes v3.1.2 (Ronquist & Huelsenbeck, 2003). MP analyses were performed as heuristic searches with stepwise random addition of taxa with the TBR (tree bisection–tree reconnection) branch-swapping algorithm with all characters equally weighted. The stability of each branch was determined using the nonparametric bootstrap (Felsenstein, 1985), with 1000 replicates and 100 random taxon additions. BA analyses were conducted in MrBayes, with a mixed model with a partition by gene assigning independent model of evolution to each gene with all parameters unlinked between partitions except topology and branch lengths on the mtDNA extended data set. Analysis consisted of two runs of four simultaneous Markov chains each for 3 million generations, sampling a tree every 1000 generations and applying a 25% burn-in after checking for convergence using TRACER v1.4 (Rambaut & Drummond, 2007) and AWTY (Nylander et al., 2008), to confirm that the standard deviation of split frequencies approached zero. The resulting trees were kept to calculate posterior probabilities in a 50% majority-rule consensus tree. As little variation was found within the Galápagos and between the Galápagos and Cocos Islands, the phylogenetic exploration was limited to MP analysis. We also produced haplotype networks for the control region to represent relationships between the haplotypes within aureola and its connections to other yellow warbler populations, using the package pegas (Populations and Evolutionary Genetics Analysis System) (Paradis, 2010) as implemented in R (R Development Core Team 2009). In this case, the haplotype network is constructed between the haplotypes, using a probabilistic approach of the most parsimonious links as given by Templeton et al. (1992). This package was also used to estimate nucleotide diversity (π) of aureola as a whole group for interspecific comparisons, as well as from individual islands with more than one haplotype.

Divergence time estimates

All mtDNA markers were tested independently for clocklike substitution rates as well as for the mtDNA concatenated data set with a likelihood ratio test implemented in PAUP*. A clock-like rate was not rejected for the concatenated data set (P > 0.001); thus, divergence time was estimated under a strict clock phylogenetic framework, using Bayesian Markov Chain Monte Carlo (MCMC) implemented in BEAST v1.4 (Drummond & Rambaut, 2006). To determine the mean substitution rate of the mtDNA markers, we calculated the mean and standard deviation substitution rate of the nine-primaried oscines (Klicka et al., 2000; Ericson & Johansson, 2003; Barker et al., 2004), using the clock data set from the study by Weir & Schluter (2008). Although these estimates are based on cytochrome b sequences, Lovette (2004) found that in Neotropical wood-warblers (Parulidae), cytochrome b sequences evolved at the same rate as other mitochondrial coding regions (ND2, ATPase, COI and COII). This evidence supported the application of the cytochrome b molecular clock in our ND2 mitochondrial data set. A rate of divergence of 1.95% (SD 0.79) was determined corresponding to 0.00975 substitutions/lineage/My for the nine-primaried oscines, supporting the generally accepted molecular rate of 2% corrected sequence divergence/My (Weir & Schluter, 2008).

A GTR+I model was used, and chains for 30 million generations were run under a Yule Process tree prior and sampled every 1000 generations. Good stationarity and high effective sample sizes (ESS > 2000) were observed for all parameters in TRACER v1.4 (Rambaut & Drummond, 2007). A consensus tree with divergence times was obtained from the 30 000 generated trees, after discarding the first 7500 as burn-in.

Compared with cytochrome b, the noncoding mitochondrial control region covers a broader range of substitution rates among the different avian taxa, with estimates of sequences divergence rates ranging from 0.1 to 21% per million years (Ruokonen & Kvist, 2002). As no calibrations for control region have been reported for yellow warblers, a 6% sequence divergence was used, calculated for the closest avian taxa to yellow warblers available, the Old World leaf warblers (Phylloscopus) (Irwin et al., 2001). Given the broad sequence divergence range and the uncertainty in using an external phylogenetic group such as the leaf warblers, this calibration was not included in the BEAST analysis. Instead, GTR+I-corrected average pairwise difference of nucleotide substitutions per site was used between aureola and all other subspecies from mainland and Caribbean sites, to provide an alternative estimate as calculated in ARLEQUIN v3.0 (Excoffier et al., 2005).

Microsatellites and data analysis from Galápagos and Cocos Islands

From a screening of 38 previously published primer sets on birds, six polymorphic microsatellite loci were found in D. petechia aureola: Dpμ01 isolated from a yellow warbler (Dendroica petechia) (Dawson et al., 1997), WpD4, WpD23, WpD30 isolated from a Wilson’s warbler (Wilsonia pusilla) (Clegg et al., 2003), Maμ23 isolated from a brown-headed cowbird (Molothrus ater) (Alderson et al., 1999), and Gf06 isolated from a Medium Ground Finch (Geospiza fortis) (Petren, 1998). To increase the number of loci, we performed shotgun sequencing from one individual yellow warbler on the Roche GS FLX 454 mass sequencer. We analysed the sequences with the software program MSAT COMMANDER (Faircloth, 2008), to identify sequences with tetranucleotide microsatellite
motif repeats. We selected loci with a minimum of four repeats and used the program PRIMERS (Rozen & Skaltsky, 2000), to design primers for amplification. The new primers (YEWA_JC) are presented in Data S2.

We used 11 microsatellites in total and screened 159 individuals for genetic variation from 10 islands (nine islands from Galápagos and Cocos Islands). PCR products were run on an ABI3730 capillary sequencer (Applied Biosystems, Foster City, CA, USA), and alleles were scored using Genemapper software (Applied Biosystems).

For each island population, exact tests were used to examine deviations of each locus from Hardy–Weinberg equilibrium expectations and test for linkage disequilibrium among loci. The method of Evanno et al. (2005) implemented in the online version of STRUCTURE HARVESTER v0.56.3 (http://taylor0.biology.ucla.edu/struct_harvest/) was used to aid in detecting the ‘true K’ values.

Population structure was examined using STRUCTURE version 2.3.1 (Pritchard et al., 2000), a Bayesian clustering program that assigns individuals to clusters (K) using a priori locality assignments for each individuals (10 islands) as implemented in this version, allowing detection of lower levels of divergence, or with less data, than the original STRUCTURE model (Hubisz et al., 2009). An ‘admixture’ prior was assumed allowing mixed ancestry of individuals from K = 1–10 with a burn-in of 50 000 with four runs for each value of K. The method of Evanno et al. (2005) implemented in the online version of STRUCTURE HARVESTER v0.56.3 (http://taylor0.biology.ucla.edu/struct_harvest/) was used to aid in detecting the ‘true K’ by examining ΔK, a measure of the change in likelihood scores between the runs of successive K values.

Relative effective population size (θ) and levels of historical gene flow were estimated between populations (M), using maximum likelihood implemented in MIGRATE 2.4 (Beerli & Felsenstein, 1999). Results from this program are viewed as long-term estimates of gene flow because it assumes mutation-migration-drift equilibrium, constant parameter values and a per-locus mutation rate. The program estimates θ, defined as 4Neμ, where μ denotes mutation rate, and M defined as m/μ, where m denotes migration rate. We designed the runs into separate populations first guided by the STRUCTURE analysis grouping islands into four clusters (Cluster I: Santa Cruz, Pinta, Pinzon, Isabela, Genovesa, Santiago, and Fernandina; Cluster II: San Cristobal; Cluster III: Floreana; Cluster IV: Cocos), as well as each island separately for a total of 10 populations. Runs corresponded to 10 Markov chains of 10 000 steps and three chains of 100 000 steps with and adaptive heating scheme (temperatures 1.0, 1.3, 1.5, 3.0), and were repeated until the confidence intervals for the posterior probabilities of θ and M overlapped.

Morphological variation in aureola

Morphological characters were measured from four of the largest islands in the Galápagos archipelago (Isabela, Santa Cruz, Santiago, and San Cristobal), characterized by steep ecological gradients along elevational transects. Continuous transects were all ≤10 km in length covering the habitat range within these islands, from mangrove and sclerophyllous dry forest in the lowlands (sea level to 200 m) to evergreen forests dominated in parts by Scalesia trees in the highlands (200–500 m) (Grant & Grant, 2008). We focused on the two extremes of the gradients where at least 15 individuals from high (H)- and low (L)-elevation zones were targeted. Morphological measurements were taken from a total of 175 individuals caught in the field (males, 101; females, 74). Analyses were conducted only on males because sexual dimorphism is evident in this group, sample sizes were larger, and geographically better distributed (n = 101; Isabela (H:14–L:17), San Cristobal (H:8–L:11), Santa Cruz (H:10–L:13), Santiago (H:15–L:13)). These data were complemented with 10 museum skins from individuals collected on Cocos Island (six LACM, four FMC). All individuals were measured by J.A. Chaves, following the methods described in the study by Chaves & Smith (2011). Morphological data were tested for normality before statistical analyses. Principal components analysis (PCA) on the correlation matrix was used to examine the size and shape variations. To control for the effects of body size on morphological traits, a general linear model (GLM) was used to generate adjusted marginal trait means with a Bonferroni correction to minimize type I errors (Rice, 1989).

Phylogenetic relationships

All methods of tree reconstruction recovered aureola reciprocally monophyletic to two lineages from Central American ‘erithachorides’ group (subspecies erithachorides – Panama– and xanthotera – Costa Rica) (Fig. 1). Analyses also strongly supported the monophyly and sister relationship between ‘aestiva’ from North America and ‘erithachorides’ + ‘petechia’, but the phylogenetic relationship between these two last groups was not supported (Fig. 1). Highly supported monophyly was also evident for some ‘petechia’ lineages (Lesser Antilles: babad, bartholmica, cruciana, gundlachi) and for some ‘erithachorides’
The one major topological discordance between taxonomic and genetic classification is the polyphyly of two ‘erithachorides’ lineages (cienegae, ruficapilla) and ‘petechia’ lineages (melanoptera, aurifrons) between Venezuelan and Lesser Antillean lineages.

**Divergence time estimates in aureola and other lineages**

The yellow warbler aureola lineage and the sister lineages (xanthotera and erithachorides) from mainland Central America were estimated to have diverged around 268 000 years before present (ybp) (height 95% HPD: 88 000–467 000) (Fig. 1). For deeper nodes, the divergence time estimate using BEAST for North American ‘aestiva’ and the rest of the yellow warbler clades were estimated at 1.47 million years ago (Mya) (height 95% HPD: 2 Mya–981 000 ybp) (Data S3). Rough estimates of divergence time from control region between aureola and xanthotera – Costa Rica and erithachorides – Panama were 203 000 and 246 000 ybp, respectively.

**Control region marker variation**

Eight mitochondrial haplotypes were found among 58 samples from Galápagos and Cocos Islands. Figure 2 shows the haplotype network, which describes
relationships among individual haplotypes and locality information. Differentiation between haplotypes was low, with most neighbouring haplotypes differing by a single step. Haplotype A was common to 45 individuals, corresponding to all nine islands including Cocos, except for the four Floreana individuals which were all characterized by one unique haplotype (F). Private haplotypes were found on three islands (Santa Cruz, San Cristobal and Floreana). Nucleotide diversity (\( \pi \)) for aureola based on 58 individuals was 0.0053 (eight haplotypes) and varied from 0.0030 to 0.0045 from four of the islands with more than two haplotypes (Table 1). There was no evidence for haplotype differentiation across the highland and lowland habitat transects within the four major islands.

Table 1 Population genetic estimates for sampled islands based on mtDNA control region.

<table>
<thead>
<tr>
<th>Island</th>
<th>n</th>
<th>No. haplotypes</th>
<th>No. private haplotypes</th>
<th>Nucleotide diversity (( \pi ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Santa Cruz</td>
<td>11</td>
<td>4</td>
<td>2</td>
<td>0.004545</td>
</tr>
<tr>
<td>San Cristobal</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>0.004040</td>
</tr>
<tr>
<td>Isabela</td>
<td>10</td>
<td>3</td>
<td>0</td>
<td>0.004040</td>
</tr>
<tr>
<td>Santiago</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>0/NA</td>
</tr>
<tr>
<td>Fernandina</td>
<td>6</td>
<td>2</td>
<td>0</td>
<td>0.003030</td>
</tr>
<tr>
<td>Genovesa</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>0/NA</td>
</tr>
<tr>
<td>Pinzon</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0/NA</td>
</tr>
<tr>
<td>Pinta</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0/NA</td>
</tr>
<tr>
<td>Floreana</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0/NA</td>
</tr>
<tr>
<td>Cocos</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0/NA</td>
</tr>
</tbody>
</table>

Table 2 \( F_{ST} \) values and Nei’s genetic distance between pairs of islands. Nei’s standard genetic distances (\( D_s \)) (Nei, 1972) are shown above the diagonal, and \( F_{ST} \) values are shown below the diagonal. Significant values (\( P < 0.01 \)) are indicated by an asterisk. Inbreeding coefficient \( F_{IS} \) per island are shown in bold.

<table>
<thead>
<tr>
<th></th>
<th>Santa Cruz</th>
<th>Santiago</th>
<th>Isabela</th>
<th>San Cristobal</th>
<th>Pinzon</th>
<th>Pinta</th>
<th>Fernandina</th>
<th>Genovesa</th>
<th>Cocos</th>
<th>Floreana</th>
</tr>
</thead>
<tbody>
<tr>
<td>Santa Cruz</td>
<td>0.0116</td>
<td>0.0283</td>
<td>0.0334</td>
<td>0.0885</td>
<td>0.0478</td>
<td>0.0552</td>
<td>0.0676</td>
<td>0.1484</td>
<td>0.0628</td>
<td></td>
</tr>
<tr>
<td>Santiago</td>
<td>-0.0090</td>
<td>0.0227</td>
<td>0.0255</td>
<td>0.0787</td>
<td>0.0523</td>
<td>0.0688</td>
<td>0.0551</td>
<td>0.1392</td>
<td>0.0502</td>
<td></td>
</tr>
<tr>
<td>Isabela</td>
<td>0.0130*</td>
<td>0.0061</td>
<td>0.0192</td>
<td>0.0801</td>
<td>0.0564</td>
<td>0.0732</td>
<td>0.0473</td>
<td>0.1292</td>
<td>0.0835</td>
<td></td>
</tr>
<tr>
<td>San Cristobal</td>
<td>0.0156*</td>
<td>0.0063</td>
<td>0.0044</td>
<td>0.0999</td>
<td>0.0675</td>
<td>0.0737</td>
<td>0.0444</td>
<td>0.1179</td>
<td>0.0888</td>
<td></td>
</tr>
<tr>
<td>Pinzon</td>
<td>0.0197</td>
<td>0.0078</td>
<td>0.0217</td>
<td>0.0356</td>
<td>0.1240</td>
<td>0.1251</td>
<td>0.0879</td>
<td>0.2582</td>
<td>0.1792</td>
<td></td>
</tr>
<tr>
<td>Pinta</td>
<td>-0.0184</td>
<td>-0.0166</td>
<td>-0.0005</td>
<td>0.0049</td>
<td>0.0072</td>
<td>0.0949</td>
<td>0.1361</td>
<td>0.2608</td>
<td>0.1008</td>
<td></td>
</tr>
<tr>
<td>Fernandina</td>
<td>-0.0062</td>
<td>-0.0036</td>
<td>0.0204</td>
<td>0.0149</td>
<td>0.0231</td>
<td>-0.0213</td>
<td>0.1054</td>
<td>0.2168</td>
<td>0.1605</td>
<td></td>
</tr>
<tr>
<td>Genovesa</td>
<td>0.0076</td>
<td>-0.0086</td>
<td>-0.0067</td>
<td>-0.0163</td>
<td>-0.0024</td>
<td>0.0288</td>
<td>0.0150</td>
<td>0.1317</td>
<td>0.1178</td>
<td></td>
</tr>
<tr>
<td>Cocos</td>
<td>0.1033*</td>
<td>0.0930*</td>
<td>0.0963</td>
<td>0.0827*</td>
<td>0.1618</td>
<td>0.1614</td>
<td>0.1399</td>
<td>0.0696</td>
<td>0.1810</td>
<td></td>
</tr>
<tr>
<td>Floreana</td>
<td>0.0124</td>
<td>-0.0034</td>
<td>0.0400</td>
<td>0.0411</td>
<td>0.0882</td>
<td>0.0083</td>
<td>0.1399</td>
<td>0.0467</td>
<td>0.1231</td>
<td></td>
</tr>
</tbody>
</table>

\( F_{IS} \): 0.100  0.138  0.042  0.129  -0.117  0.311  0.075  -0.036  0.131  -0.132
Microsatellites and gene flow analyses

The genetic characteristics of the eleven microsatellite loci from the ten islands are described in Table 2 and Data S5. Between 2 and 11 alleles per locus were detected among the 159 individuals surveyed, and they do not show significant departures from Hardy–Weinberg equilibrium after Bonferroni corrections. The mean heterozygosity averaged across populations ranged from 0.025 (YEWA_JC20) to 0.897 (WpD30) with an overall mean heterozygosity of 0.453, suggesting high levels of variation at the subspecies level when compared to ‘aestiva’ conspecifics (Gibbs et al., 2000).

Little genetic differentiation among islands is suggested by mostly small and nonsignificant $F_{ST}$ values and Nei’s genetic distance values (Table 2). Significant pairwise comparisons were found between individuals from Isabela and Santa Cruz ($F_{ST} = 0.013$; $P < 0.01$), San Cristobal and Cocos Island ($F_{ST} = 0.082$; $P < 0.01$), San Cristobal and Santa Cruz ($F_{ST} = 0.0156$; $P < 0.01$), Cocos and Santa Cruz ($F_{ST} = 0.1033$; $P < 0.01$), and Cocos and Santiago ($F_{ST} = 0.0930$; $P < 0.01$) (Table 1).

Results from STRUCTURE suggest four distinct groups in *aureola* as the optimal clustering of genetic variation in yellow warblers (Fig. 3). Cluster I corresponded to seven islands within the archipelago (Santa Cruz + Isabela + Santiago + Fernandina + Pinzon + Genovesa + Pinta: Cluster I), where the geographic distance between five of these islands (‘core islands’: Santa Cruz + Isabela + Santiago + Fernandina + Pinzon) is <20 km. The other two islands within the Cluster I (Genovesa and Pinta) are more than 50 km separated from the rest. The other three clusters corresponded to the two southernmost islands sampled within the archipelago San Cristobal and Floreana.
(Cluster II) and Floreana (Cluster III) also found more than 50 km. The last cluster corresponded to Cocos Island (Cluster IV) found more than 800 km north-east from the archipelago (Fig. 3).

Historical effective population sizes \((\theta)\), estimated using MIGRATE, were somewhat homogeneous, but highest in San Cristobal \((\theta = 0.97)\), Santiago \((\theta = 0.813)\) and Floreana \((\theta = 0.811)\), and lowest in Isabela \((\theta = 0.63)\). Historical migration rate \((M)\) showed asymmetric gene flow within the populations grouped into clusters, suggesting a south to north pattern of migration within the Galápagos archipelago (Cluster III to Cluster I, \(M = 1.82\); Cluster II to Cluster I, \(M = 1.82\)) and less a north to south pattern (Cluster I–Clusters II and III, \(M = 0.52\) and 1.42) (Fig. 3 and Table 3). Historical estimates of gene flow between Galápagos and Cocos Islands were the lowest for Cluster III–Cluster IV \((M = 0.012)\), but largest for Cocos to Galápagos Islands (Cluster IV–Cluster II, \(M = 2.2\)). Separate island analysis \((n = 10)\) confirmed the overall south to north pattern of gene flow (Data S6), as well as the high gene flow estimates for Cocos to Galápagos Islands (to Santiago: 1.66 and Pinzon: 1.27) but not so clear into San Cristobal (Cluster II, \(M = 0.9\)) as previously shown. An overall pattern of homogeneous gene flow between the islands at the core of the archipelago (Cluster I) was also observed.

**Morphological variation**

Morphology did not vary significantly among island populations despite the large geographic separation between the Galápagos and Cocos Islands and differences in habitat in which birds were sampled (Fig. 4). The PCA reduced the six morphological measures to four components that explained 67.8% of the total variance in male morphology. The PC1 explained approximately one-third of the variance and was largely a measure of overall body size (PC1, 26.4%) (Data S7). No morphological traits differed among islands in the GLM analysis. Likewise, DFA analyses had low assignment probabilities per island based on PC scores. No differences in male morphology were found between highland and lowland habitats within each island (all traits \(P > 0.05\)) as well as across islands using pooled highland and lowland individuals (all traits, \(P > 0.05\)).

Table 3 Migration estimates \((M)\) across all clusters \((n = 4)\) using MIGRATE. Cluster I, Santa Cruz, Isabela, Fernandina, Pinzon, Santiago; Cluster II, San Cristobal; Cluster III, Floreana; Cluster IV, Cocos.

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Cluster I</th>
<th>Cluster II</th>
<th>Cluster III</th>
<th>Cluster IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster I</td>
<td>-</td>
<td>1.8</td>
<td>1.82</td>
<td>0.9</td>
</tr>
<tr>
<td>Cluster II</td>
<td>0.52</td>
<td>-</td>
<td>0.68</td>
<td>2.2</td>
</tr>
<tr>
<td>Cluster III</td>
<td>1.42</td>
<td>0.68</td>
<td>-</td>
<td>0.78</td>
</tr>
<tr>
<td>Cluster IV</td>
<td>1.46</td>
<td>1.54</td>
<td>0.012</td>
<td>-</td>
</tr>
</tbody>
</table>

Fig. 4 Yellow warbler morphological traits: (a) across Galápagos and Cocos Islands, and (b) from highland (H) and lowland (L) habitats at four islands in the Galápagos. Means for each trait (mm) are based on size-corrected marginal means after the general linear model (GLM).
Phylogeography of yellow warblers in Galápagos

Discourse

Temporal aspect of genetic and morphological variation

Three major questions in evolutionary biology are: How much time is required for genetic and morphological divergence to arise in natural populations? What is the respective role of ecology and geographic isolation in morphologic diversification? Whether different taxonomic groups respond in the same way to ecological and evolutionary pressures. In the case of yellow warblers, the monophyly of aureola is indicative of a single, relatively recent colonization event from the mainland to these islands, perhaps in the last 268,000–450,000 years. The ‘star-shaped’ haplotype network recovered in the control region and the lack of shared haplotypes between aureola populations and those from the mainland (‘petechia’ and ‘erithachorides’ groups) also suggest a recent colonization followed by haplotype differentiation in situ.

In comparison with other Galápagos avian colonizers of similar estimated arrival time (<300,000 years ago), aureola harbours overall greater genetic diversity (eight haplotypes, $\pi = 0.0053$) than Galápagos hawks (Buteo galapagoensis; five haplotypes, $\pi = 0.0019$) (Bollmer et al., 2006) and magnificent frigatebirds (Fregata magnificens; three haplotypes, $\pi = 0.00012$) (Hailer et al., 2010). A likely explanation is higher effective populations sizes in aureola and the likelihood that larger number of original colonists founded the population, but also that population size may be less affected by past demographic bottlenecks such as El Niño that often decimate seabirds populations (Schreiber & Schreiber, 1984).

Differential selection pressures act upon fitness-related traits along elevation gradients as previously shown in the Galápagos for medium ground-finches Geospiza fortis (Grant et al., 1985) and for other birds (Price, 1991; Soobramoney et al., 2005; McCormack & Smith, 2008; Milá et al., 2010). The analyses performed here found no significant variation in yellow warblers across these gradients. Our small sample size when comparing between elevations within islands may have prevented the detection of effects of such selection. Nevertheless, the analysis of pooled samples from highland and lowland sites across islands was still nonsignificant. Shifts in morphology in insular birds could arise in very short evolutionary timescales (4000 years) since colonization (Clegg et al., 2008). Although we do not provide morphological comparison with mainland counterparts, the overall lack of morphologic variation in yellow warblers in these islands could be the results of many factors such as high levels of gene flow.

The celebrated adaptive radiation in Darwin’s finches and the dramatic interspecific diversity in beak shapes have almost certainly been driven by differential selection resulting from differences in diets (Lack, 1947; Schluter & Grant, 1984). The large differences found among species of Darwin finches evolved over two to three million years, since the ancestor arrived on the archipelago (Grant & Grant, 2008). However, there is also ample evidence of recent and rapid evolutionary change within species of Darwin’s finches. For example, rapid changes in Geospiza fortis and G. scandens on Daphne Major (Grant & Grant, 2002) and G. fortis on Santa Cruz (Hendry et al., 2009), or the very recent shift in bill size of Geospiza fortis on Santa Cruz as a result of anthropogenic changes to the habitats (Hendry et al., 2006). Many of these highly significant changes in bill size in Darwin’s finches have been observed over very short time frames, even on a scale of a few years (Grant & Grant, 2002). Given these marked differences in some other species such as the Galápagos mockingbird, which differ in beak size between islands (Abbott & Abbott, 1978), a phenomenon that Darwin first noted (1836 [1963], Suloway, 1982), and morphological distinctiveness among island populations of Galápagos hawks despite their very recent arrival (Bollmer et al., 2003), why do the yellow warblers show no morphological variation anywhere in the archipelago? It has been argued that the large amount of morphological variation found in Darwin’s finches could be due to their ancestor having intrinsically a higher capacity for morphological change than ancestors of other avian colonizers to the Galápagos (Burns et al., 2002). There is a parallel in the other avian adaptive radiation in the Pacific. This is the case of Hawaiian honeycreepers, in which a single ancestor speciated into at least fifty species. Lovette et al. (2002) found Hawaiian thrushes, whose ancestor arrived on the archipelago around the same time as the ancestor of honeycreepers radiated into only five lineages, which show relatively little variation in beak size. Collectively, this evidence seems to point towards a taxonomic bias in the capacity for divergence in morphology, where lower evolutionary lability in both Hawaiian thrushes and Galápagos yellow warblers could be an alternative explanation.

Origin and population history of yellow warblers

The mitochondrial data indicate that aureola is most closely related to the Central American group of yellow warblers (‘erithachorides’). This suggests that the source populations were most likely from surrounding continental regions rather than from Caribbean islands (West Indies – ‘petechia’ lineages). This finding is in contrast to the origins proposed for other Galápagos birds which are believed to have phylogeographic affinities with populations in the Caribbean (Darwin’s finches: Burns et al., 2002; and to some extent Galápagos mockingbirds: Arbogast et al., 2006). Olson (1980) reported that the populations from Galápagos and Cocos Islands appear more similar in plumage (chestnut-capped birds) to some of the subspecies in the West Indian ‘petechia’ group than to adjacent ‘erithachorides’ group, which is not supported.
at the DNA level. He also noted that *au reola* could be viewed as the endpoint in a southward increase of yellow on the head and checks from mainland Pacific coast populations and suggested that the plumage similarity to West Indian subspecies may have evolved independently. The preliminary phylogenetic reconstruction we report here supports this latter hypothesis.

Haplotype reconstruction and genetic estimates within the archipelago suggest that haplotype A most likely represents the colonizer haplotype, which subsequently spread rapidly across the islands either from Cocos Island to the archipelago or vice versa despite the more than 800-km distance between them. In contrast, Floreana populations, which are closer geographically to the rest of the islands, not only lack haplotype A but harboured unique private haplotype (F). If we take into consideration genetic diversity estimates (π) where higher diversity represents longer time for differences to accumulate, then Santa Cruz Island should have been the first island to be colonized, followed by other islands in the Galápagos and Cocos Island. This is contrary to the haplotype network reconstruction in which the connection between the Galápagos and mainland populations ('erith-achorides' and 'petechia') is through private haplotype D from San Cristobal Island populations (Fig. 2 and Data S4). This alternative scenario suggests a colonization event to the Galápagos first, with subsequent colonization of Cocos Island. Finally, a third scenario suggests haplotypes A and D arrived to the Galápagos via Cocos Island in a stepping-stone fashion, with the subsequent extinction of (or unsampled) haplotype D in Cocos Island (Fig. 2). Although different from yellow warblers’ point of origin, Darwin’s finches colonization route is consistent with the first scenario with a progression back to the Cocos Island once populations reached the Galápagos archipelago (Petren et al., 1999; Grant & Grant, 2008).

Similar to the results from mtDNA, microsatellite analyses show a moderate but significant amount of differentiation among *au reola* populations, which suggests a genetic substructure between islands. The Bayesian clustering method (structure) grouped Floreana individuals in their own category, largely corroborating the mitochondrial results. Similarly, San Cristobal individuals were clustered in a separate group, but this population exhibited haplotype sharing (type A) and mixed ancestry in mtDNA. Both microsatellite-based FST and FIS pairwise values and migrate estimates of historical gene flow suggest that restricted gene flow occurs, especially among Floreana, San Cristobal and Cocos Islands, the latter also separated in its own genetic cluster. Although high estimates were calculated between Cocos and San Cristobal Island populations, a general trend of unidirectional gene flow within the Galápagos Islands corresponds to a south-east to north-west direction of historical migration. This pattern could be explained by the prevailing south-south-east trade winds in the Galápagos (Power, 1975; Colinvaux, 1984; Geist, 1992), which have been implicated in the pattern of distribution of mockingbirds (*Nesomimus*: Arbogast et al., 2006; but see Zenaida doves: Santiago-Alarcón et al., 2006), giant tortoises (*Geochelone nigra*: Caccone et al., 1999, 2002) and lava lizards (*Microlophus*: Benavides et al., 2009).

The results of this study suggest that yellow warbler populations in the Galápagos and Cocos Islands are at an early stage of diversification after a single colonization event from mainland populations from Central America. Genetic variation was evident across several islands characterized by high genetic diversity compared to other recent avian colonizers. However, the genetic structuring in this group was not paralleled by fitness-related traits when quantified across steep ecological gradients or across islands.

**Acknowledgments**

We thank David Willard and John Bates for allowing the use of Klein and Brown DNA samples stored at the Field Museum of Natural History Chicago, and Kimball Garrett at the Natural History Museum of Los Angeles County. We thank the Galápagos National Park and the Charles Darwin Foundation, for providing support to this research in particular to Felipe Cruz, Sonia Cisneros, Paulina Couben and Ruth Boada. Freddy Cabrera and Wolfgang Buermann provided assistance in the field, and Johann Besserer from the Isabela Oceanographic Institute provided logistical assistance. Jenny Bollmer, Nicole Gottdenker and Jessica Rabenold collected many of the UMSL samples. A special thanks to Claudio Cruz and family in Floreana, Nicolas Palacios from Metropolitan Touring and our most grateful thanks to Roque and Pilar Sevilla for their hospitality. J. Pollinger, R. Jones and L. Journey helped with the laboratory work. The manuscript was improved by comments from R. Harrigan. This research was funded by the University of California Los Angeles Lida Scott Brown and National Science Foundation grant (IRCEB9977072) to TBS.

**References**


Goudet, J. 2001. *FSTAT* (version 2.9.3.): A Program to Windows to Estimate and Test Gene Diversity and Fixation Indices. Institut d’Écologie, Université de Lausanne, Lausanne, Switzerland.


Supporting information

Additional Supporting Information may be found in the online version of this article:

**Data S1** Specimen data and GenBank accession numbers for samples used in this study.

**Data S2** Microsatellite information.

**Data S3** The Bayesian Inference chronogram from BEAST based on the mtDNA combined dataset (ND2 and ATPase).

**Data S4** Phylogenetic network showing evolutionary relationships between haplotypes in yellow warbler samples over the geographic distribution.

**Data S5** Locus information from *aureola*.

**Data S6** Migration estimates (M) all islands (*π* = 10) using MIGRATE.

**Data S7** Factor loadings from the principal component analysis and percentage explained for male yellow warblers.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be reorganized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

Received 15 September 2011; revised 11 November 2011; accepted 27 November 2011