Nuclear and mitochondrial sequences confirm complex colonization patterns and clear species boundaries for flightless weevils in the Galápagos archipelago

A. S. Sequeira\(^1\), S. M. Sijapati\(^1\), A. A. Lanteri\(^2\) and L. Roque Albelo\(^3\)

\(^1\)Department of Biological Sciences, Wellesley College, 106 Central Street, Wellesley, MA 02481, USA
\(^2\)Departamento Científico de Entomología, Museo de La Plata, 1900 La Plata, Argentina
\(^3\)Department of Terrestrial Invertebrates, Charles Darwin Research Station, Puerto Ayora, Galápagos, Ecuador

Nuclear sequence data were collected from endemic Galápagos species and an introduced close relative, and contrasted with mitochondrial DNA sequences, continuing investigation into the colonization history and modes of diversification in the weevil genus *Galapaganus*. The current combined phylogeography together with previously published penalized likelihood age estimates builds a complex picture of the archipelago’s colonization history. The present reconstruction relies on submerged platforms to explain the early divergence of the young southern Isabela endemics or the Española or San Cristobal populations. Diversity is later built through inter-island divergence starting on older islands and continuing on two simultaneous tracks towards younger islands. The amount of diversity generated through intra-island processes is skewed towards older islands, suggesting that island age significantly influences diversity. Phylogenetic concordance between nuclear and mitochondrial datasets and well-supported monophyletic species in mitochondrial derived topologies appear to reject the possibility of inter-species hybridization. These clear species boundaries might be related to the tight host associations of adult weevils in discrete ecological zones. If shared hosts facilitate hybridization, then host- or habitat-promoted divergences could prevent it, even in the case of species that share islands, since the altitudinal partitioning of habitats minimizes range overlap.

**Keywords:** colonization; diversification; species boundaries; phylogenetic concordance; weevils; Galápagos

1. INTRODUCTION

Oceanic islands are natural systems for studying the process of new species formation (Emerson & Kolm 2005). Among patchy populations, as island populations are assumed to be, gene flow is expected to decrease, whereas genetic differentiation should rise owing to genetic drift and differing selective pressures (Wright 1943; Kimura & Weiss 1964; Slatkin 1993), all factors favouring lineage splitting and speciation.

Accurate phylogenetic reconstructions of the relationships of island species groups allow examination of species formation within a temporal framework (Roderick & Gillespie 1998; Juan et al. 2000; Emerson 2002; Hormiga et al. 2003). The progression rule, as described for island colonization, is a stepwise progression down the island chain from the oldest to the youngest island accompanied by lineage bifurcation (Funk & Wagner 1995). Some form of this pattern has been observed for a variety of plant and animal radiations in several volcanic archipelagos including the Hawaiian Islands, Canary Islands and Galápagos (Funk & Wagner 1995; Juan et al. 2000; Parent & Crespi 2006). However, back-colonization towards a continent, within-island differentiation, adaptation and extinction, among other events, can confound a simple pattern of island emergence followed by colonization (Juan et al. 2000; Emerson 2002).

The weevil genus *Galapaganus* Lanteri (Coleoptera: Curculionidae) is a suitable system to study species radiations on islands. This genus contains 15 species and among the 13 that are flightless and fairly heavy bodied, 10 are endemic to the Galápagos (Lanteri 1992; table 1, figure 1). The sequence of events implied by the branching order of a recent comprehensive mitochondrial phylogeography of *Galapaganus* did not clearly display a pattern consistent with the progression rule, mostly owing to the basal position of the ancestor to the lineage that is endemic to the young island of Isabela (Sequeira et al. 2008). Even though a simple colonization model with formation of species on the oldest islands (Española or San Cristobal) and colonization of the nearest islands as they emerged could not be discounted, the results pointed towards a more complex pattern of colonization and speciation.

An intriguing pattern has emerged for the Galápagos, in which several endemic taxa display divergences older than the geological age estimates of the islands currently above water (Schmitz et al. 2007), in particular, iguanid...
By contrast, such explanations are not generally required to reconcile divergence times with the ages of the extant islands in other dynamic volcanic island systems such as the Canary and Hawaiian Islands (Fleischer et al. 1998; Emerson et al. 2000a,c; Gubitz et al. 2000; Price & Clague 2002; Hormiga et al. 2003; Emerson & Oromi 2005). Previous analyses of Galapaganus with the penalized likelihood (PL) age estimation method (Sanderson 2002) and the use of an extrinsic clock (employing published calibrations for invertebrate mitochondrial divergences; Sequeira et al. 2000) indicate that the age of island endemics and presumably that of the colonization of the archipelago by the ancestor of all island endemics (8.6–12 Ma) exceeds the geological estimates of the extant emerged islands (Cox 1983; Hickman & Lipps 1985). Explanations of this discordance could invoke geological evidence of older, drowned islands as potential colonization platforms for ancestors, effectively extending the time available for species diversification beyond the ages of the extant emerged islands (Christie et al. 1992; Werner et al. 1999).

The prevailing pattern of mitochondrial DNA (mtDNA)-derived time estimations supported colonization and within-island differentiation soon after island emergence for most of the islands (Sequeira et al. 2008). This pattern favours dispersal scenarios over vicariant events to explain the bulk of inter-island diversification. However, vicariance induced by lava flows was invoked to explain patterns of diversification rate variation within the younger and more volcanically active islands. The current comparison with a nuclear-based phylogeny seeks to confirm or propose alternative patterns of colonization to those derived solely from mtDNA.

Even though the emerged islands in the Galápagos archipelago are relatively geologically recent, their isolation from the South American continent (1000 km west of the coast of Ecuador) and their climatic and topographic diversity have promoted speciation in both plant and animal groups (Grant et al. 1996a; Peck 1996a; Sato et al. 2001; Tye et al. 2002; Nielsen 2004). In particular, there is a positive relationship between beetle diversity and island area, elevation and ecological complexity, but not island age (Peck 1994, 2005). Hence, it could be suggested that environmental diversity within each island in the Galápagos provides an ecological spectrum for species differentiation (Parent & Crespi 2006).

Patterns of species diversifications revealed by phylogenetic analyses can distinguish between inter- and
intra-island species formation (Emerson 2002; Jordal et al. 2006). Intra-island speciation can contribute to generation of island diversity, especially on islands above a certain size (Losos & Schluter 2000), while molecular phylogenetic studies of Canary Island arthropods revealed that intra-island speciation has had a much greater influence on species diversity on older islands than younger islands (Emerson & Oromi 2005). Intra-island speciation has been proposed as playing an important role in the generation of endemic diversity of Galapaganus land snails (Parent & Crespi 2006); it appears to occur with equal frequency on younger and older substrates. The previous Galapaganus phylogeny suggested that island age and not ecological complexity influenced speciation modes in Galapaganus, allowing more time for in situ speciation on older islands (Sequeira et al. 2008), in line with the pattern observed for Canary Island arthropods (Enghoff & Baez 1993; Emerson & Oromi 2005).

The island of Isabela provides a suitable test case for the relative contributions of island properties to diversity through intra-island diversification by exhibiting traits at the extremes of two of the variables in question (age and area) and comparable values for a third possible variable (habitat complexity). Sampling from several volcanoes from northern Isabela allows us to observe the relationships among populations/species within the island and to confirm or dispel the ideas of island area or age differentially promoting species formation or differentiation within this radiation.

When studying young species radiations, interpretations based solely on mtDNA phylogenies should be treated cautiously (Shaw 2002). Mitochondrial DNA is susceptible to selective sweeps and asymmetrical introgression that can severely mislead interpretation of the evolutionary history of a group (Ballard 2000; Shaw 2002; Machado & Hey 2003; Chan & Levin 2005). Analysis of multiple loci from mitochondrial and nuclear genomes is required to demarcate species boundaries confidently (Emerson 2002; Jordal et al. 2006; Linnen & Farrell 2007 but see Rubinoff & Holland 2005). Indeed, inter-species hybridization and introgression are prevalent in island beetles (Sota & Vogler 2001; Zhang et al. 2005; Jordal et al. 2006), flies (Pestano et al. 2003) and crickets (Shaw 2002).

The mtDNA-derived phylogeny of Galapaganus showed high correspondence with morphological species definitions, with high levels of resolution and support for monophyletic species, and lead us to believe that inter-species hybridization is not prevalent in this group (Sequeira et al. 2008). However, the exception was paraphyletic Galapaganus vandykei on Española and San Cristobal. Two phylogenetic patterns expected under prevalent mitochondrial introgression...
are non-monophyletic of species in the mitochondrial gene tree and discordance between mitochondrial and nuclear-based phylogenies. The possibility of hybridization of G. vandykei with Galapagos collaris and Galapaganus galapagoensis on San Cristobal is investigated here by assessing concordance with a nuclear-based phylogeny.

Nuclear sequence data and additional mtDNA sequences were added to those of Sequeira et al. (2008), continuing investigation into the colonization history and modes of diversification in Galapaganus. Well-resolved mitochondrial and nuclear phylogenies allow testing of the following explicit predictions: (i) if the colonization history of the archipelago does not follow the chronological ages of the extant islands owing to the use of sunken seamounts as colonization platforms, then the phylogeny/ies should not strictly follow the progression rule; (ii) if larger and topologically complex, though young islands, such as Isabela, are not promoting intra-island speciation, then weevil species/populations from Isabela will not be each other’s closest relatives in the estimated phylogeny; and (iii) if inter-species hybridization and mitochondrial introgression are not a widespread phenomenon in this radiation, then topologies of mitochondrial and nuclear-based phylogenies will be congruent.

2. MATERIAL AND METHODS

(a) Sampling

Samples were obtained from nine of the islands (figure 1, table 1). Nine of the 10 endemic Galapagos species and Galapaganus howdenae howdenae, widespread in coastal Ecuador and recently introduced in the agricultural area of Santa Cruz (Peck et al. 1998; Lanteri 2004; Causton et al. 2005), were included in the present study. Galapaganus howdenae howdenae is one of two members of the femoratus species group, while all endemics belong to the darwini species group (Lanteri 1992; Sequeira et al. 2008). Outgroups include species in the South American genus Naupactus, also in the tribe Naupactini of the subfamily Entiminae (table 1).

(b) DNA preparation, polymerase chain reaction amplification and sequencing

Collected specimens were preserved at −20°C in 100 per cent ethanol. DNA was isolated following Normark (1996) and Sequeira et al. (2008). The polymerase chain reaction (PCR) was used to amplify six nuclear gene regions: elongation factor 1 alpha (EF1-α); arginine kinase (ArgK); internal transcribed spacer 1 (ITS-1); and three anonymous loci (4b, 5, 6). Anonymous loci were developed for this study using a Topo Shotgun Subcloning Kit (Invitrogen, CA). In a subsample of populations from seven ingroup species, 15 anonymous regions were tested and the three containing larger levels of variation were selected to potentially maximize phylogenetic resolution. Blast searches of GenBank databases failed to return homologous sequences for the regions selected. Amplification reactions followed Sequeira et al. (2008) with primers for nuclear regions listed in table 2. A typical temperature profile for nuclear protein-coding regions consisted of 35 cycles with two annealing temperatures (four cycles at 47°C and 31 cycles at 50°C). A single annealing temperature of 55°C was used for known non-coding regions (ITS-1) and anonymous loci. The products were purified and sequenced using PCR primers following Sequeira et al. (2008).

Sequences were compiled, edited and aligned in Sequenser v. 4.6 (GeneCodes corporation, Ann Arbor, MI). Anonymous loci and ITS-1 alignments were performed in Clustal X (Thompson et al. 1997). Heterozygous sites (clear large peaks of equal size on all overlapping fragments) were found in very low frequencies and coded as Ns for phylogenetic analyses. Sequences for four mitochondrial gene regions (12S, partial sequence of cytochrome c oxidase subunits I and II and cytochrome b) were obtained for three additional populations (IS03, IS07, SA01; table 1), following Sequeira et al. (2008) and combined with data from that previous dataset.

A combined matrix of 4152 characters was compiled from all six nuclear regions, 655 for ITS-1, 919 for EF1-α, including one short mid-intron (128 bases), 828 for ArgK and 1747 for all three anonymous loci (645, 483 and 617 for loci 4b, 5 and 6, respectively) for 70 individuals from 24 populations of 10 Galapaganus species (table 1). Sequences have been submitted to GenBank under accession numbers EU748556–EU748797 and EU748853–EU748886, and EU748798–EU748852 for nuclear and mitochondrial sequences, respectively.

Table 2. Primers used for PCR amplification and sequencing of nuclear gene regions. (Primers for all anonymous loci were specifically designed for this study; primers for EF1-α, ArgK and ITS-1 were designed in the laboratory of Brian Farrell, Harvard University.)
(c) Phylogenetic analysis
Each nuclear partition was initially analysed separately using maximum parsimony (MP), maximum likelihood (ML) and Bayesian inference (BI), as detailed below. No major topological disagreements were observed among all six independent nuclear analyses. Although all genes are expected to have independent genealogical histories, to recover the underlying best estimate of the species history, evidence from all genes that share that history should be considered (Maddison 1997). Also, combining nuclear datasets ensures maximization of the signal present in each region compared with consensus approaches. Therefore, a concatenated nuclear dataset was compiled. Nuclear and mitochondrial datasets were not combined until significant discordance could be discounted (figure 2).

MP analysis was performed in PAUP* v. 4.0.b10 (Swofford 1998) including all nucleotide sites. Searches were performed using the heuristic search algorithm with the random addition sequences option, 1000 replicates and the tree bisection reconnection (TBR) branch-swapping algorithm. One thousand bootstrap replicates, each with 100 random addition sequences, were performed to assess the support for individual branches.

ML topologies were first constructed in PHYML (Guindon et al. 2004) and further optimized in PAUP* using the model selected by ModelTest (Posada & Crandall 1998) for all partitions together (general time reversible, GTR). Bayesian-inference analysis was performed with MrBayes v. 3.1.1 (Huelsenbeck & Ronquist 2001). The model selected for each gene region incorporated in Bayesian searches was chosen according to the Akaike information criterion in MrModeltest v. 2.2 (Nylander et al. 2004). For the analysis of the concatenated nuclear sequences, a site-specific model was applied in which coding regions were analysed under a Hasegawa, Kishino, Yano (HKY) model, estimating the proportion of invariable sites, and the gamma distribution (HKY + I + G) and a more complex GTR + I + G model was applied for nuclear non-coding regions (ITS-1, EF1-α intron and anonymous loci). Mitochondrial regions were analysed under a site-specific model in which first and second codon positions and 12S sequences were analysed under a HKY + I + G model, while the more complex GTR + I + G model was applied for third codon positions. The combined mitochondrial and nuclear dataset was treated as six different partitions incorporating the two models used in the independent analyses. Each search in MrBayes started from a random tree and was continued for 2 000 000 generations with four simultaneous chains, each sampled every 100 generations. Majority-rule consensus trees were constructed after all trees below the stationarity level were discarded.

(d) Colonization scenarios
Alternative scenarios of island colonization sequence were inferred from the topology of the MP and BI phylogenetic trees of the combined dataset and the relative geographical setting of the islands following methods described by Thorpe et al. (1994) and previously used for other island radiations (Thorpe et al. 1994; Parent & Crespi 2006). This process is based on the supposition that an island is more likely to be colonized from a close island than from a distant one. However, because this method is sensitive to taxon sampling, the general direction of colonization inferred was used in conjunction with MP reconstructions of island of origin for colonizers as ancestral character states performed in MacClade (Maddison & Maddison 1993). Equally parsimonious sources of colonists were incorporated into the scenarios proposed as alternatives (figure 3).

(e) Bayesian and Shimodaira-Hasegawa tests of island assemblage, morphospecies monophyly and topological congruence
To test for island assemblage and morphospecies monophyly, we applied hypothesis-specific tests. Likelihood ratio tests were performed comparing the unconstrained ML tree with trees that were constrained for each of the hypotheses to be tested (one-tailed Shimodaira-Hasegawa (SH) log-likelihood test (Shimodaira & Hasegawa 1999) as implemented in PAUP*). We tested the island assemblages of San Cristobal, Santa Cruz and Isabela for monophyly. In addition we tested for the single origin of endemics inhabiting several islands (G. vandykei). Each of the constrained searches was performed with the same settings as the unconstrained searches.

Parametric bootstrapping (PB) was used to evaluate the effect of constraining the monophyly of the three islands and one morphospecies not inferred as monophyletic in our analyses. Sequences were simulated on a constrained tree for a given hypothesis constructed with PAUP*, using maximum likelihood distances with parameter estimates derived from the ModelTest analysis. Simulated sequences (100 datasets) were generated in Seq-Gen (Rambaut & Grassly 1997) using the same model of sequence evolution and parameter estimates used to construct the hypothetical tree. The resulting distribution of differences was then compared with the MP tree length differences for the empirical constrained and non-constrained trees (Ruedi et al. 1998; Emerson et al. 2000b). All tree searches were performed in PAUP* with 1000 random addition sequences and TBR with no max-trees limit. Bayesian posterior probabilities (BP) were also determined for island groupings not supported by our analyses.

SH tests were also used as a means of evaluating congruence between nuclear and mitochondrial datasets and among all nuclear datasets. Topological constraints were constructed from topologies obtained from analyses with each gene region and tested against all other datasets.

3. RESULTS
(a) Phylogenetic analyses and colonization patterns
(i) Independent analyses of the nuclear and mitochondrial datasets
The EF1-α exon and ArgK regions contained no insertions or deletions and were unambiguously aligned with verifiable open-reading frames. The short intron region within EF1-α and ITS-1 displayed most of the length differences between the dawini and femoratus species group sequences and the outgroup Naupectus species, but not within either of the species groups. Most of the variations in the anonymous loci were in the form of substitutions with few length differences.

The concatenated nuclear dataset contained 435 parsimony informative characters, 320 derived from non-protein coding regions (139 ITS-1, 24 EF1-α intron and 157 anonymous loci) and 115 from coding regions (67 EF1-α and 48 ArgK). Common resolved and well-supported topological features observed across the three independently analysed nuclear partitions (coding: EF1-α exon, ArgK; non-coding: ITS-1, EF1-α intron; anonymous: all three anonymous loci) are the monophyly of the genus, of the endemic members of the dawini species group, of the Galapaganus
...Galapaganus caroli from Santiago. Unique topological features (resolved in one of the partitions and not in others) include the monophyly of Galapaganus caroli endemic to Floreana (anonymous loci), Galapaganus darwini from Culpepper (Darwin) and Wenman (Wolf), and the basal position of the G. vandykei individuals from Española (ITS-1 and EF1-α intron). Despite the...
particulars of the individual nuclear topologies, SH tests performed among the three partitions and the tests performed for each of the three datasets on the ML constrained topologies produced by the two alternative datasets yielded no significant differences in any case (e.g. coding dataset constrained by topologies of non-coding and anonymous loci datasets), six tests in total ($-\ln L$ differences range: 6.35–22.67; $p > 0.05$). The MP topology for the nuclear concatenated dataset (figure 2b) is the strict consensus derived from 35 MP trees ($L = 1239$). ML and BI topologies do not differ greatly from the MP consensus tree with the (a) MP and (b) Bayesian topologies. Dotted arrows indicate equally parsimonious scenarios in each case. Numbers below the branches in the topology correspond to the arrows in the inset for two alternative colonization scenarios: MP: normal font, BI: italics. IC indicates initial colonization and asterisks indicate intra-island speciation.

Figure 3. MP strict consensus tree derived from a concatenated mitochondrial and nuclear dataset. Topological differences from the Bayesian majority rule consensus tree are indicated by dashed grey lines. Numbers above and below branches as in figure 2.

Phil. Trans. R. Soc. B (2008)
exception of the additional monophyly of *G. galapagoensis* and the sister group relationship of *G. collaris* and *G. vandykei* in San Cristobal (figure 2b). Resolved and supported groupings in the nuclear–combined phylogeny include the monophyly of the genus and of the endemic *darwini*, of *G. conwayensis* from northern Isabela, of *G. vandykei* from the northern islands and of *G. blairi* from Santiago. A salient aspect of the nuclear–combined topology is the lack of resolution between species other than the basal position of the ancestor to *G. vandykei* from Española within the island endemics, albeit without support in the form of bootstrap or BPs. Lack of resolution is probably the result of the small amount of variation present in the nuclear data, rather than of homoplasy caused by a rapid radiation. Despite the relative absence of between species resolution and support within the *darwini* group, the existing monophyly of several species and island groupings (e.g. Santiago, Culpepper, Wenman) is confirmed in this topology and argues against significant conflict with patterns revealed by mitochondrial data (figure 2a).

The nuclear and mitochondrial MP- and BI-derived phylogenies are compared side by side in figure 2 (ML topologies were again very similar to the BI majority rule consensus). The mitochondrial-derived phylogeny confirmed results from the previous analysis, including many resolved and well-supported interspecies relationships, monophyly of all but one of the morphologically defined species (*G. vandykei* is paraphyletic) and the basal position of the ancestor to *Galapaganus williamsi* from southern Isabela. The present improved dataset and analyses provide increased bootstrap support for the intriguing basal position of the ancestor of the endemic species of the young island of Isabela. Also, the species and populations from the older island of San Cristobal now formed a paraphyletic assemblage instead of the polyphyletic one in the previous estimate.

(ii) Combined analyses

Given the lack of resolution at intermediate levels in the nuclear topology, the main topological difference from the mitochondrial dataset (SH tests) is the basal position of the ancestor to *G. vandykei* from Española. The mitochondrial topology is more resolved and displays the ancestor to *G. williamsi* from southern Isabela at the base of the topology. Differences were not significant in either of the comparisons tested (mitochondrial dataset/nuclear tree, Diff − ln L = 38.74, p > 0.05; nuclear dataset/mitochondrial tree, Diff − ln L = 29.07, p > 0.05). Both datasets therefore fail to reject the alternative species proposed as the first island endemic.

The combined dataset from mitochondrial and nuclear sequences contained 1104 informative characters. Figure 3 shows the MP strict consensus of 240 MP trees (*L* = 3892). The combined topology resembles the one resulting from the mitochondrial data (figure 2a) except for the position of *G. vandykei* from Española, here derived from the ancestral divergence that gave rise to *G. williamsi* from southern Isabela and subtending the rest of the *darwini* species group. The Bayesian majority rule consensus topology derived from the combined analysis is almost identical to the MP topology. The one intriguing difference is the basal position of the ancestor to the *G. vandykei* population from Española, exchanging that position with the southern Isabela endemics in the MP topology. Despite the low posterior probability values supporting these relationships, the branching order suggests an alternative colonization scenario.

The arrows in the inset in figure 3 represent the two alternative colonization scenarios constructed following Thorpe et al. (1994) and MP optimizations that match relationships between the species and populations of endemics derived from both phylogeny estimates. These rather complex scenarios have three common characteristics as follows: the position close to the original colonization of the archipelago of an ancestor of the endemic species of the young island of Isabela; two parallel colonization tracks from Española or San Cristobal, one involving the colonization of Pinta and northern Isabela from Santa Cruz; and the other featuring a stepwise pattern through San Cristobal and Santiago towards the young northernmost islands. Previously published estimates of divergence times and geological estimates of island ages impose some constraints on the interpretation of these patterns, especially for the location of the original colonization.

(b) Intra-island speciation versus multiple colonizations

The combined mitochondrial MP and BI topologies (figure 3) displayed two instances of intra-island speciation on San Cristobal. The highland species *G. collaris* appears to have originated from lowland populations of *G. vandykei* within San Cristobal. *Galapaganus galapagoensis*, the third endemic on this older island, derives from the ancestor of the other island endemics, *G. collaris*/*G. vandykei*, in a paraphyletic grouping that later gave rise to endemics on the northern islands of Santiago, Culpepper (Darwin) and Wenman (Wolf).

Previous analyses supported the origin of Santa Cruz endemics as the result of multiple colonizations (Sequeira et al. 2008), albeit that monophyly of these lineages could not be rejected. The current estimate proposes Santa Cruz as a source of *G. conwayensis* colonists for Pinta and northern Isabela and through divergence of endemics.

Isabela is the third island with multiple endemic species. The topology reported here clearly supports one colonization of southern Isabela from an early ancestor and a later colonization of northern Isabela from Santa Cruz. Results of SH, PB and Bayesian monophyly tests of island assemblages are summarized in table 3. All three tests significantly reject the monophyly of Isabela endemics and favour multiple colonizations to explain the origin of *Galapaganus* diversity in this young and large island for the combined and mitochondrial datasets (*p* < 0.01), but not for the nuclear dataset. However, despite the high support of both mitochondrial and combined datasets for the paraphyly of the San Cristobal assemblage, none of the tests applied could significantly reject the monophyly of the lineages in the oldest island of the archipelago. Interpretation of tests for a single origin of Santa Cruz endemics is less straightforward; SH
and PB tests fail to reject a single origin for *Galapaganus ashlocki* and the *G. conwayensis* populations from Santa Cruz, supporting intra-island speciation for this older island. However, PB tests for the mitochondrial dataset reject island monophyly in favour of multiple colonizations.

**Table 3. Results of island and species monophyly tests performed on the combined (C), mitochondrial (M) and nuclear datasets (N). (The SH column gives the results of the Shimodaira–Hasegawa test (likelihood score differences and significance). The PB column gives the MP tree length score differences between constrained and non-constrained topologies for those monophyletic island and morphospecies groupings not present in the analysis and the probability values under parametric bootstrapping. The post-probability column lists the posterior probabilities of those same groupings. Asterisks indicate significant differences between hypothesis trees (constrained) and empirical topologies, suggesting that the hypothesis of monophyly of islands or morphospecies is rejected in favour of the alternative hypothesis of multiple colonizations or species non-monophyly.)**

<table>
<thead>
<tr>
<th>island or morpho species tested</th>
<th>dataset</th>
<th>SH</th>
<th>PB</th>
<th>post-probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isabela</td>
<td>C</td>
<td>72.44</td>
<td>38, (p &lt; 0.01)^*</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>109.76</td>
<td>27, (p &lt; 0.01)^*</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>22.43</td>
<td>5, (p &gt; 0.05)</td>
<td>0.26</td>
</tr>
<tr>
<td>San Cristobal</td>
<td>C</td>
<td>16.38</td>
<td>6, (p &gt; 0.05)</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>20.44</td>
<td>4, (p &gt; 0.05)</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>15.42</td>
<td>7, (p &gt; 0.05)</td>
<td>0.33</td>
</tr>
<tr>
<td>Santa Cruz</td>
<td>C</td>
<td>20.54</td>
<td>9, (p &gt; 0.05)</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>18.62</td>
<td>15, (p &lt; 0.05)^*</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>29.17</td>
<td>7, (p &gt; 0.05)</td>
<td>0.33</td>
</tr>
<tr>
<td>G. vandykei</td>
<td>C</td>
<td>30.17</td>
<td>10, (p &gt; 0.05)</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>37.84</td>
<td>9, (p &gt; 0.05)</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>11.42</td>
<td>5, (p &gt; 0.05)</td>
<td>0.13</td>
</tr>
</tbody>
</table>

and PB tests fail to reject a single origin for *Galapaganus ashlocki* and the *G. conwayensis* populations from Santa Cruz, supporting intra-island speciation for this older island. However, PB tests for the mitochondrial dataset reject island monophyly in favour of multiple colonizations.

**4. DISCUSSION**

**(a) Colonization patterns**

The combined mitochondrial and nuclear dataset can be interpreted as indicating a single colonization from the South American continent, originating from within the range of *G. h. howdenae*, other species of the femoratus species group and wingless members of the darwini species group (Lanteri 1992, 2004). The current phylogenetic estimate does not include continental congeners other than *G. h. howdenae*, precluding more precise identification of the continental source of island colonists other than coastal Ecuador and Peru (Lanteri 2004; Peck 2005). The assumption of monophyly of all island endemics and a single colonization of the archipelago can only be effectively tested by including all close continental relatives (Emerson 2002). Therefore, alternative scenarios proposing multiple colonizations by already divergent wingless lineages or back-colonization of the continent by wingless weevils cannot be completely discounted until samples of *Galapaganus propinquus*, *Galapaganus squashosus* and *Galapaganus lacertosus*, the continental darwini group species, become available. However, both these scenarios are unlikely, given the direction of the prevailing oceanic currents (Wyrtki 1966; Graham 1975), which would aid in transport of colonists from the continent (Wright 1983; Peck & Kukalova-Peck 1990; Lopez et al. 1992).

Previous PL age estimations based on a phylogenetic estimate derived from mitochondrial sequence data have dated the colonization of the archipelago by a *Galapaganus* ancestor (8.6–12 Ma) earlier than the geological age estimates of the islands (Sequeira et al. 2008). Discord between molecular and geological dating in Galápagos is not exclusive to this weevil radiation (Wright 1983; Lopez et al. 1992); however, it is not generalized across all radiations in the archipelago, since the colonization and divergence times of endemic finches (Sato et al. 2001), hawks (Bollmer et al. 2006) and microlepidoptera (Schmitz et al. 2007) fall within the age range of the emerged islands as in other volcanic island systems (e.g. Hawaii; Price & Clague 2002).

Because most estimates of the timing of within-island diversification in this group fall within the range of the geological estimates of each island’s emergence, Sequeira et al. (2008) favoured dispersal explanations for the generation of *Galapaganus* diversity. However, in a few instances (e.g. *G. conwayensis* on Santa Cruz and Pinta), more recent vicariant events caused by sea-level changes could have contributed to the current distributions.

The current combined phylogeny, together with the previous PL age estimate, builds a complex picture of the colonization history of the archipelago by *Galapaganus*. Considering the time of the initial colonization, as derived from previous studies, the present reconstruction would need to rely on current submerged platforms for the original colonization to explain the early divergence of the southern Isabela endemic or even the
Española or San Cristóbal populations, since the maximum age of all islands is much less than the time of the initial colonization (Isabela: 0.7 Ma; Española: 5.6 Ma; San Cristóbal: 6.3 Ma). Diversity would later be built mostly through inter-island divergence starting on Española or San Cristóbal and continuing on two simultaneous tracks across the archipelago towards the younger islands. Although not strictly progressing down an island chain in a single linear pattern, similar age-related patterns starting on the older island of Española have been observed for bulimulid snails (Parent & Crespi 2006) and flightless Stomion beetles (Finston & Peck 1997). Despite the complexity of the patterns suggested, there are aspects that appear as a progression towards the younger islands. One scenario involves a stepwise colonization pattern towards the northernmost islands from G. blairi of the central island of Santiago.

(b) Intra-island speciation, island area and island age

With the current sampling of Galapaganus species, and despite the availability of samples from three islands that harbour multiple endemic species (San Cristóbal, Santa Cruz, Isabela), in situ or intra-island speciation is reconstructed in the phylogeny as having occurred only on San Cristóbal and less clearly on Santa Cruz. Topology tests cannot reject an intra-island origin for Santa Cruz endemics. Intra-island speciation seems to be more prevalent in older than in younger islands in Galápagos, as in arthropod groups in the Canary Islands (Emerson & Oromi 2005).

Multiple lines of evidence reject in situ speciation to explain Galapaganus diversity in the large, young and topologically complex island of Isabela. The phylogenetic topology suggests that two separate colonizations occurred and topology tests significantly reject the monophyly of a combined northern and southern Isabela assemblage. This is the only case of multiple species sharing an island in which both the mitochondrial and combined data significantly support multiple colonizations versus in situ differentiation as a plausible explanation for the origin of the island’s Galapaganus fauna. Isabela was formed by six major volcanoes now connected by extensive lava flows with little or no vegetation. The most important of these putative barriers is a 12 km wide lava field (the Perry Isthmus), which separates northern and southern Isabela. Despite the fact that Galapaganus weevils are tightly associated with their specific host plants, lava flows between volcanoes do not appear to be effective barriers within the northern area (Darwin and Alcedo volcanoes) and further analyses should be performed before considering each volcano or even the northern and southern regions as separate islands with regard to weevil biogeography.

The importance of intra-island speciation to species richness in island systems has been explicitly linked to island area and habitat diversity (Cowie 1995; Losos & Schluter 2000); both are strong predictors of species richness in Galápagos land snails (Parent & Crespi 2006). Isabela is the largest island (more than 4000 km²) and could be considered highly complex ecologically, or at least equally as complex as the older islands, since each volcano contains all ecological/vegetation zones found on Santa Cruz and San Cristóbal (Wiggins & Porter 1971; Peck 1991, 1996b, 2005). However despite its area, Isabela does not contribute to Galapaganus species diversity through intra-island speciation but by being colonized multiple times. Alternatively, the island’s large area and central location may make it the target of multiple colonizations. As for Canary Island arthropods (Emerson & Oromi 2005), intra-island divergences in Galapaganus are more prevalent on older but smaller islands such as San Cristóbal and Santa Cruz, suggesting that island age has had a greater influence on species diversity than island area by allowing more time for in situ speciation in older islands (Cowie 1995). While most of the Galapaganus species diversity is the product of inter-island divergence, the amount of intra-island diversification is skewed towards older islands. Island age thus influences diversity more than island area in this radiation.

(c) No evidence of inter-species hybridization and mitochondrial introgression

In some island radiations, behavioural changes that may accompany frequent founder events can facilitate inter-species hybridization (Roderick & Gillespie 1998). In some of these instances patterns of variation inferred from mtDNA have proved misleading about species relationships (DeSalle et al. 1997; Sota & Vogler 2001; Shaw 2002 but see Rubino & Holland 2005). Studies contrasting nuclear- and mitochondrial-derived topologies have provided evidence of pervasive mitochondrial introgression, though not exclusively, in island systems (Shaw 2002; Alves et al. 2003; Jordal et al. 2006; Linnen & Farrell 2007). However, both phylogenetic signals that could be interpreted as possible indicators of inter-species hybridization, i.e. non-monophyly of species assemblages in mitochondrial phylogenies and mitochondrial/nuclear topological incongruence, were not significantly supported by the present analyses of Galapaganus.

The apparent lack of inter-species mitochondrial gene flow confirms that mitochondrial genes are reliable markers for species delimitations and phylogenetic reconstructions of colonization timing and history in this group. These clear species boundaries might be related to the tight association of weevils with their hosts in discrete ecological zones, even as adults. Each species feeds on plants restricted to one or two contiguous ecological zones (Lanteri 1992; Peck 2005). When two or more species occupy an island, one is usually confined to the moist highlands, while its counterpart inhabits the lowlands (Lanteri 1992). Adults of species that do not share islands with other Galapaganus species could be considered habitat generalists since they feed on a variety of hosts across all ecological zones. A previous ML optimization of habitat preferences in Galapaganus reconstructed multiple shifts towards the more restricted highland preferences on different islands and suggested that habitat shifts could be linked to species formation (Sequeira et al. 2008). If sharing hosts facilitates hybridization, then host- or habitat-promoted divergences could prevent it, even in the case of species that share islands, given that the altitudinal partitioning of the habitats will minimize range overlap. In
Galapag anus, as in other island radiations, potential range overlap is avoided by the nature of the island system in which inter-island colonizations underlie the bulk of the species divergences (Grant & Grant 1992; Finston et al. 1997; Roderick & Gillespie 1998; Sato et al. 1999; Shaw 2002). Our hypothesis of absence of inter-species hybridization and mitochondrial introgression is supported by these analyses. Potential future lines of inquiry will focus on morphological or behavioural evidence of reproductive isolation that could accompany ecological preferences, and these clear species boundaries, especially for species pairs that share islands, and that in the presence of major climatic shifts (i.e. El Niño events) could experience some range overlap.

Scientific research permits for the collecting trips were obtained from the Parque Nacional Galápagos through the Department of Terrestrial Invertebrates at the Charles Darwin Research Station (Santa Cruz). We gratefully acknowledge S. Cárdenas, L. Cruz (and the crew from ‘El Pirata’), A. Mieles, A. M. Ortega, J. Rosado and station volunteers for invaluable assistance in the field and S. Peck for providing specimens. Field logistical support was provided by the Department of Terrestrial Invertebrates at the Charles Darwin Research Station.

REFERENCES


Rubino, D. & Holland, B. S. 2005 Between the two extremes: mitochondrial DNA is neither the panacea nor the nemesis of phylogenetic and systematic inference. Syst. Biol. 54, 952–961. (doi:10.1038/sj/sysbio.8200234674)


