Darwin’s finches represent a dynamic radiation of birds within the Galápagos Archipelago. Unlike classic island radiations dominated by island endemics and intuitive ‘conveyor belt’ colonization with little subsequent dispersal, species of Darwin’s finches have populations distributed across many islands and each island contains complex metacommunities of closely related birds. Understanding the role of metacommunity and structured population dynamics in speciation within this heterogeneous island system would provide insights into the roles of fragmentation and dispersal in evolution. In this study, a large multi-species dataset and a comparative ground finch dataset (two co-distributed lineages) were used to show how landscape features influence patterns of gene flow across the archipelago. Factors expected to regulate migration including distance and movement from large, central islands to small, peripheral islands were rejected in the multi-species dataset. Instead, the harsh northeast islands contributed individuals to the larger central islands. Successful immigration relies on three factors: arriving, surviving and reproducing, thus the dispersal towards the central islands may be either be due to more migrants orienting towards these land masses due to their large size and high elevation, or may reflect a higher likelihood of survival and successful reproduction due to the larger diversity of habitats and more environmentally stable ecosystems that these islands possess. Further, the overall directionality of migration was south-southwest against the dominant winds and currents. In comparing dispersal between the common cactus finch and medium ground finch, both species had similar migration rates but the cactus finch had approximately half the numbers of migrants due to lower effective populations sizes. Significant population structure in the cactus finch indicates potential for further speciation, while the medium ground finch maintains cohesive gene flow across islands. These patterns shed light on the macroevolutionary patterns that drive diversification and speciation within a radiation of highly-volant taxa.

Keywords: dispersal, Geospiza, metacommunity, migrate-n
Introduction

Island radiations are the foundation upon which much of evolutionary theory has been developed (Darwin 1859, Wallace 1871, MacArthur and Wilson 1967, Losos and Ricklefs 2009). The Hawaiian Archipelago, the Caribbean Islands, the Malay Archipelago and Melanesian Archipelagos and the Galápagos Archipelago are all examples of island systems that were instrumental in forming both classic and modern evolutionary theories (Grant 1999, Gillespie 2002, Losos et al. 2003, Lohman et al. 2011). Much of this understanding comes from comparing island landscape characteristics to speciation patterns and/or the process of species accumulation (MacArthur and Wilson 1967, Whittaker and Fernández-Palacios 2007, Losos and Ricklefs 2009, Warren et al. 2015). As models of the speciation process, most classic island radiations focus on ancestral and descendant populations, and typically connect island area and distance to speciation probability (e.g. Hawaiian Tetragonatha spiders, and Caribbean Anolis lizards, and many Galápagos species). However, these models miss one of the most critical stages in the speciation process for many species, the cessation of gene flow between semi-isolated populations within a larger structured population framework.

While island archipelagos are most famous for their ability to create new species, many of these systems also maintain disperse allopatic populations within species, maintained through ongoing gene flow (MacArthur and Wilson 1967, Whittaker and Fernández-Palacios 2007, Emerson and Gillespie 2008, Losos and Ricklefs 2009, Warren et al. 2015). Understanding these archipelago endemics within a comparative multi-species framework of potentially biased gene flow can provide unique insights into when species will diverge through isolation and local adaptation and when they will remain connected through the influx of alleles from other islands (Bolnick and Nosil 2007, Logan et al. 2016). Further, in comparative analyses of metacommunity dynamics of many co-distributed species, both species-specific and overall patterns may emerge to understand how the physical environmental landscape may shape evolutionary processes (Reid et al. 2006, Roberts 2006, Jones and Kennedy 2008).

Unlike most classic radiation examples in oceanic archipelagos, Darwin's finches in the Galápagos Archipelago do not follow the model of speciation through isolation of single island endemics (Grant 1999). The finches instead have many species co-distributed on multiple islands. Distance is less likely to rapidly sever gene flow in birds than in terrestrial fauna (Lomolino and Weiser 2001), therefore understanding the spatial relationship of available island habitats and dynamics of gene flow can provide a crucial missing piece in understanding the early stages of diversification.

Darwin's finches have approximately 18 distinct lineages with population structure within some species (Farrington et al. 2014, Lamichhaney et al. 2015). The finches can be categorized into a number of clades including the ground finches (Geospiza), the tree finches (Camarhynchus), the warbler finches (Certhidea), and three single-species lineages: the vegetarian finch Platyspiza crassirostris, the sharp beaked ground finch Geospiza difficilis (from Santiago, Fernandina and Pinta Islands) and the Cocos Island finch Pinaroloxias inornata. As most species are co-distributed on many islands throughout the archipelago (Grant 1999), Darwin's finches are an ideal system to study species specific and overall patterns of gene flow and divergence in a classical model system.

Though island distributions for each species are well known, the levels of gene flow between island populations (within species) are not known despite anecdotal observations of occasional suspected immigrants (Boag and Grant 1984) or the founding of new populations (Grant and Grant 1995). Until a better understanding of inter-island migration for individual species and of archipelago-wide processes is established, our ability to understand evolution in this iconic system is impaired.

Landscape and climate variables throughout the Galápagos Archipelago are complex, with many possible variables contributing to shaping diversification and gene flow. The islands exhibit a variety of sizes, elevations, orientations and habitats that may all impact gene flow and genetic isolation. Distance (over water or large expanses of lava) may isolate populations, as might flying against wind currents which are predominantly north-northwest to west-northwest throughout the year (Trueman and d'Ozouville 2010). Additionally, the deep channel which separates the Northern Galápagos Volcanic Province Islands (Pinta, Marchena and Genovesa as well as Darwin and Wolf) and the fact that these islands were formed separately from the central islands may mean that many species experience a genetic break along this line (Sinton et al. 2018).

With these potentially influential landscape characteristics of the Galápagos Archipelago in mind, recent metapopulation and dispersal studies provide a number of hypotheses for overarching patterns of gene flow within this system. First, we expect that migration should be from older islands to younger islands (Wagner and Funk 1995, Cowie and Holland 2006, 2008, Parent et al. 2008, Poulakakis et al. 2012) – which would head west in the Galápagos Archipelago (Geist et al. 2014). Second, central islands should send migrants to the peripheral islands as assumed by peripheral speciation models (Mayr 1963, Servadio and Kirkpatrick 1997). Third, dispersal should be in line with dominant wind and water currents (Hoskin 2000, Roberts 2006, Poulakakis et al. 2012) and should be lower between the two regions that have never been connected (Northern Galápagos Volcanic Province Islands and Central Islands) (Geist et al. 2014). Finally, both larger and higher islands are generally expected to be sources of emigrants due to their greater range of habitats and lower climate stochasticity (Heinrichs et al. 2016).

In this study, we provide a first look into species-specific and archipelago-wide migration estimates in Darwin's finches and use them in a comparative framework to understand island archipelago characteristics that may drive paths of
gene flow critical for maintaining species integrity or promoting divergence and speciation. With a large metacommunity dataset (populations sampled across all 11 lineages with multi-island distributions excluding the isolated Darwin and Wolf islands), we test whether migration is limited by distance, directionality, aspects of island size or elevation (high elevation is linked to wetter and more drought-resistant habitat), and potential barriers such as crossing between the central and northeastern islands. In comparison to these large-scale processes that might be influenced by the diversity of ecology and niche found across this radiation, we then utilize a paired dataset of two closely related ground finch lineages with extremely broad distributions (Geospiza fortis and G. scandens: the medium ground finch and the common cactus finch) to tests for similarity of gene flow patterns throughout the archipelago.

Material and methods

Field collection

All birds were captured in mist nets over a multi-year period following Univ. of Cincinnati IACUC protocols (#03-06-17-01, 06-06-01-0). Brachial venipuncture was used to obtain a small drop of blood that was blotted onto EDTA-soaked filter paper and dried (Petren et al. 1999b, 2005). Birds were released after processing. 1190 birds are included in this study representing 11 species of Darwin’s finches (only species with ≥ 3 island populations were included: four ground finches, three tree finches, both warbler finches, and the sharp-beaked finch and vegetarian finch lineages). This dataset encompasses 68 populations from 13 of the largest islands in the Galápagos Archipelago (Table 1). Species names match Lamichhaney et al. (2015).

Additional specimens from Daphne Islet were included in preliminary analyses, but were removed from final analyses due to signals of interspecies introgression (see Results). Four species (G. fortis, G. scandens, C. parvulus, Geospiza olivacea) had separate populations in Isabela Island in the north and south regions in the microsatellite dataset, where previous estimates of genetic divergence showed extensive isolation due to vast expanses of uninhabitable lava landscapes (Parent et al. 2008). For these species, populations from the northern section of Isabela Island (Alcedo Volcano and above) were treated as separate populations from those in the southern section of Isabela (Sierra Negra Volcano) (Table 1). Only one population on Isabela Island was sampled for each species in the SNP dataset (all from Alcedo Volcano in the north). DNA was extracted from blood samples using standard methods (Petren et al. 1999a). All 1190 birds were included in the expanded microsatellite analysis of migration patterns, while a subset of the two most widely distributed species (G. fortis and G. scandens, 114 birds) were used for genomic assessment of species-specific migration patterns.

Metacommunity microsatellite dataset

Four multiplex PCR reactions were used to determine the genotype of each bird for 14 presumably unlinked microsatellite loci (Petren et al. 1999b). PCR products were analyzed on an Applied Biosystems 3730xl DNA Analyzer at the Cornell University Life Sciences Core Laboratories Center. Sample runs were genotyped in GENEMAPPER ver. 3.0 (Applied Biosystems) without knowledge of population origin to limit scoring bias. Errors within the data (i.e. input errors, allelic dropout, stutter and null alleles) were assessed in MICROCHECKER ver. 2.2 (Van Oosterhout et al. 2004).

Paired SNP dataset

A genomic SNP dataset was constructed through genotyping by sequencing (Elshire et al. 2011) at Cornell Univ. Genomics Facility outlined in Lawson and Petren (2017a) for G. fortis and G. scandens due to their inhabitation of nearly every island in the archipelago. After filtering for quality (> 5x representation of each allele, minimum minor allele frequency = 0.02, minimum quality score of 20, minimum site taxon coverage = 0.25), alignment to the G. fortis draft genome, and annotation to chromosome position using LiftOver to the zebra finch genome, 48 480 SNPs were recovered across Darwin’s finches (average depth of 103 reads) (Lawson and Petren 2017a). After further filtering for SNPs per species, and for various optimized datasets (see below for strategies to maximize individuals and SNPs for various analyses), ~20–30 000 SNPs were included in each analysis (see below for specifics). Only islands represented by both species were included (Table 2).

Population assignment in G. scandens and G. fortis

In order to determine the number of subpopulations that make up the total population (K) in the G. scandens and G. fortis SNP datasets, both Structure (Pritchard et al. 2000) and fastStructure (Raj et al. 2014) were used because they use different and potentially complementary algorithms. A downside of Structure is that with the most common method of determining K, known as the Evanno method (Evanno et al. 2005), a single panmictic population cannot be selected (K = 1). FastStructure, on the other hand, has been criticized for a greater potential to miss the true optimal K with its algorithm when population structure is weak (Raj et al. 2014). The SNP dataset was initially 60 G. scandens and 54 G. fortis. Individuals from Daphne and Albany Islets were removed for all final analyses based on preliminary data (see below). As Structure and fastStructure were both sensitive to missing information, but can accommodate a low degree of unknowns, a different filtering strategy was used for the ‘Structure SNP dataset’ than for the migrant-n dataset. This dataset was filtered to only include biallelic SNPs (minimum allele frequency of 0.05) and individuals with > 85% coverage using SNP extract in Tassel ver. 5 (Bradbury et al. 2007). This resulted in a working dataset of 46 G. scandens
Table 1. Sample sizes used for migration analysis according to species and island. (‘u’ = insufficient samples for analysis; ‘–’ = not recorded; ‘x’ = presumed extinct; ‘?’ = not detected in recent surveys). Isabela Island has two values reflecting the north and south areas of the island which have shown substantial genetic structure in previous studies. First value is from the northern part of the island (primarily Alcedo Volcano and above) and the second value is from the southern part of the island, all from Sierra Negra Volcano.

<table>
<thead>
<tr>
<th>Species</th>
<th>Geospiza fortis</th>
<th>Geospiza fuliginosa</th>
<th>Geospiza magisteriis</th>
<th>Geospiza scandens</th>
<th>Geospiza difficilis</th>
<th>Camarhynchus parvulus</th>
<th>Camarhynchus psittacula</th>
<th>Camarhynchus pallida</th>
<th>Platsypisa crassirostris</th>
<th>Certhidea fusca</th>
<th>Certhidea olivacea</th>
</tr>
</thead>
<tbody>
<tr>
<td>Island</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Española</td>
<td>10</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>29</td>
</tr>
<tr>
<td>Fernandina</td>
<td>u</td>
<td>9</td>
<td>15</td>
<td>–</td>
<td>8</td>
<td>u</td>
<td>6</td>
<td>u</td>
<td>–</td>
<td>x</td>
<td>19</td>
</tr>
<tr>
<td>Floreana</td>
<td>10</td>
<td>9</td>
<td>x</td>
<td>5</td>
<td>x</td>
<td>22</td>
<td>u</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>x</td>
</tr>
<tr>
<td>Genovesa</td>
<td>–</td>
<td>18</td>
<td>48</td>
<td>–</td>
<td>722</td>
<td>10</td>
<td>18</td>
<td>18</td>
<td>23</td>
<td>–</td>
<td>26/16</td>
</tr>
<tr>
<td>Isabela</td>
<td>13/22</td>
<td>29/21</td>
<td>6</td>
<td>20</td>
<td>x</td>
<td>722</td>
<td>10</td>
<td>18</td>
<td>18</td>
<td>23/19</td>
<td>–</td>
</tr>
<tr>
<td>Marchena</td>
<td>u</td>
<td>12</td>
<td>10</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>8</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>19</td>
</tr>
<tr>
<td>Pinzón</td>
<td>13</td>
<td>21</td>
<td>15</td>
<td>13</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>u</td>
</tr>
<tr>
<td>Rábida</td>
<td>21</td>
<td>12</td>
<td>x</td>
<td>23</td>
<td>x</td>
<td>u</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>19/19</td>
<td>–</td>
</tr>
<tr>
<td>San Cristóbal</td>
<td>29</td>
<td>41</td>
<td>7</td>
<td>27</td>
<td>x</td>
<td>23</td>
<td>7</td>
<td>29</td>
<td>22</td>
<td>16</td>
<td>–</td>
</tr>
<tr>
<td>Santa Fe</td>
<td>25</td>
<td>14</td>
<td>23</td>
<td>20</td>
<td>24</td>
<td>u</td>
<td>u</td>
<td>6</td>
<td>–</td>
<td>–</td>
<td>31</td>
</tr>
<tr>
<td>Santiago</td>
<td>197</td>
<td>186</td>
<td>94</td>
<td>175</td>
<td>55</td>
<td>74</td>
<td>31</td>
<td>53</td>
<td>88</td>
<td>110</td>
<td>127</td>
</tr>
<tr>
<td>Total</td>
<td>197</td>
<td>186</td>
<td>94</td>
<td>175</td>
<td>55</td>
<td>74</td>
<td>31</td>
<td>53</td>
<td>88</td>
<td>110</td>
<td>127</td>
</tr>
</tbody>
</table>

Migration estimation

All migration analyses (metacommunity microsatellite data and paired SNP data) were run in Migrate-N v. 3.6.11 for the specifics of a SNP dataset. Datatype was set to Nucleotide and many of the same parameters as above, except modified for the specifics of a SNP dataset. Daunay was set to Nu.

Further filtering was then applied to ensure that all biallelic SNPs were truly biallelic across the entire dataset and Illuminus (0.02, maximum 0.98 to ensure that all sites, sites were further filtered to a minimum site heterozygosities and 40 G. fortis). Following this, filtering of individuals and group assignments using fastStructure's chooseK.py script for fastStructure models. We used the Bayesian clustering algorithm in Structure for the comparative SNP datasets for G. fortis, filtering strategies were needed to remove both individuals and assumed correlation of allele frequencies. We did not use the cluster assignments with the number of genetic clusters (K) and population structure analysis on the CIPRES science gateway (Miller et al. 2010), and thus required a different filtering strategy than the Structure models. In both cases, we implemented the admixture model for these nuclear markers (M of Nu. of individuals).

The final dataset consisted of 21,866 SNPs in G. fortis. The dataset was filtered for individuals and populations with high levels of missing data. Migrate-N validated and paired SNP data pairs were sampled from 3,339 SNP in G. fortis. The dataset was filtered for individuals and included the cluster assignments with the number of genetic clusters (K) and populations structure analysis on the CIPRES science gateway (Miller et al. 2010).
with a tratio of 2.0. Theta and migration priors were 0, 0.2, 0.02 and 0 1000 100 respectively as recommended by author suggestions. After trials for convergence and prior suitability were completed, long-inc was set at 100, long-sample was 20 000 and burnin was 100 000.

Archipelago characteristics and migration

GLM (general linear statistical modeling) was used as a heuristic tool to compare species and gauge the relative effects of different geographic factors in the R statistical software (R Development Core Team). Number of migrants (as estimated from Migrate-n) was log transformed to improve normality. Potential explanatory factors included in the model were: inter-island distance, source island elevation (all elevation estimates are max-elevation), receiving island elevation, delta elevation (difference between source and receiving island), source island area, receiving island area, crossing the gap between the main Galápagos islands and the separate northeast island (Marchena, Pinta and Genovesa) – treated as a categorical variable, and clade (warbler, tree, ground, vegetarian and G. difficilis). Data available in Supplementary material Appendix 1 Table A1. Relevant predictive variables to be included in the final model were determined by AIC through the R package MuMIn (tools for performing model selection and model averaging) 1.40.4 (Bartoń 2018) which compares subsets of the ‘global’ model. Analysis of variance (ANOVA) tests were also conducted in R to analyze the categorical variation of migration in terms of clades and species.

Directionality

The effects of directionality on migration were assessed independently using circular statistics in the software Oriana 4 (<www.kovcomp.co.uk/oriana>) using the angle of direction between the two closest points between two islands (Supplementary material Appendix 1 Table A2). Second order statistics were used to assess whether the number of migrants across species and directionality were correlated, and whether directionality of migration was random or in a dominant direction (Hotelling’s one sample test). As in the GLM analyses, all migration values were log transformed as this test is sensitive to skew. Though this dataset of islands does not allow movement in all directions from every island due to the inherent spatial structure (e.g. no islands are west of Fernandina), the non-linearity of this archipelago should limit biases in directionality. However, all findings should be considered within the framework of the natural structure of these islands.

Data deposition

SNP data along with species and population location are available in Dryad Digital Repository: <https://doi.org/10.5061/dryad.575md> (Lawson and Petren 2017b). Microsatellite files are available upon request due to the need to standardize datasets if they are to be combined.

Results

Metacommunity dataset

Migration estimation

In the multi-species microsatellite dataset, many potential predictive variables of spatial variation in migration had no significant relationship including: inter-island distance (p = 0.757), crossing the submarine trench that separates the northeastern islands from the central islands (p = 0.098), source or receiving island size (p = 0.766, 0.352 respectively), and source or receiving island elevation (p = 0.353, 0.145 respectively) (Supplementary material Appendix 1 Table A3). The best-fit model only included ‘delta island elevation’
was highly divergent from other than
2 – in the microsatellite Structure analy
2011.186° (~south-southwest, Fig. 1 bottom right corner), and
set (r = 0.144, p = 0.0003). Second order statistics of direc
Camarhynchus psittacula
directions. Warbler finches primarily dispersed
immigrants were Genovesa (p < 0.001), Española (p = 0.035),
and the vegetarian finch all had similar migration in
These patterns are consistent with findings for each source and receiving island show patterns that align
with the geophysical arrangement of islands, such as Pinzón Island sending many migrant almost due north and Santa Cruz sending the bulk of its migrants west (Supplementary material Appendix 1 Fig. A7, A8).

**Directionality**

There was an overall correlation of direction and migration number in Darwin's finches from the large, multispecies dataset (r = 0.144, p = 0.0003). Second order statistics of directionality showed an overall direction grand mean vector of 211.186° (~south-southwest, Fig. 1 bottom right corner), and that the distribution of migration weights was significantly non-random (Hotelling's test: F = 4.2, p = 0.016; Fig. 4). Specific results from each clade showed that tree finches, G. difficilis and the vegetarian finch all had similar migration in south and west directions. Ground finches migrated more evenly in all directions. Warbler finches primarily dispersed in northwest and southeast directions (Supplementary material Appendix 1 Fig. A6). Direction and weight of migration

**Migration estimation**

In the G. fortis and G. scandens SNP datasets, the number of migrants dispersing each generation was significantly higher in G. fortis than G. scandens (means 1.256, 0.078 respectively, t = 3.9886, df = 134.26, p = 0.0001. Welch two-sample t-test, Fig. 3). This is linked to higher theta values in G. fortis than G. scandens (means 0.164, 0.066 respectively, t = 3.635, df = 15.832, p = 0.0023), as unscaled migration estimates were actually higher in G. scandens than G. fortis (means 30.488, 38.280 respectively, t = -2.352, df = 111.39, p-value = 0.0204). Total immigrants and emigrants for each island showed Isabela as a major source of migrants for both species. Very few islands contributed immigrants in G. scandens, while G. fortis had a more even distribution from each island (Fig. 4, Supplementary material Appendix 1 Fig. A9).

**Population structure of G. fortis and G. scandens**

Removal of the Daphne population – The Daphne population of G. scandens was highly divergent from other populations in both the SNP and microsatellite datasets (Supplementary material Appendix 1 Fig. A1). The Daphne population of G. fortis shared alleles with individuals from Pinta and Marchena, but not other populations. As this population is thus somewhat distinct, but has documented suspected introgression with other species (Grant et al. 2004, Farrington et al. 2014), the Daphne populations from both species were removed from further analyses.

Geospiza fortis – in the microsatellite Structure analysis, G. fortis had a best fit of K = 3 (Supplementary material Appendix 1 Fig. A2). As Structure cannot evaluate K = 1 in the Evanno method (described above), a panmictic population (K = 1) could not be evaluated. K = 3 was also the best-supported model for G. fortis in the SNP dataset evaluated using Structure, however inspection of the groupings and the Ln Likelihood plot suggest K = 1 is likely the correct model (Supplementary material Appendix 1 Fig. A2), as supported by the findings from fastStructure on the SNP dataset which identified panmictic populations (K = 1) as the best fit for G. fortis.

Geospiza scandens – in the microsatellite dataset, G. scandens had the best fit at K = 4 (Supplementary material Appendix 1 Fig. A2). In the SNP dataset, Structure found G. scandens to have ΔK peaks at K = 3 and K = 5 in the SNP dataset (Supplementary material Appendix 1 Fig. A2), however fastStructure identified a panmictic populations (K = 1) as the best fit for G. scandens on the SNP dataset implying that the inability to evaluate K = 1 in Structure may have lead to an inflated number of subdivisions identified than experienced in nature.
The model that best predicts migration weights for the *G. scandens* SNP datasets had no significant predictor variables of archipelago characteristics (Supplementary material Appendix 1 Table A4). The *G. fortis* SNP dataset had aspects of island variation as significant predictors (larger source islands, \( p = 0.004 \) and higher source islands, \( p = 0.005 \)). If the specific source and receiving islands are also included as variables in the potential model, ‘source island’ is also significant in the *G. fortis* model, with all islands significant as either high or low sources. Isabela, San Cristóbal, Santa Cruz had positive relationships (produced more emigrants; all \( p \leq 0.001 \)); Marchena, Pinta, Rábida and Santa Fe had negative relationships (produced fewer emigrants; \( p \leq 0.001, 0.004, < 0.001, < 0.001 \)). The significance of migration rates for Santiago and Floreana could not be calculated due to collinearity of the full dataset (Supplementary material Appendix 1 Table

Figure 1. Directionality and sum migration between islands across all species. Top 14 panels: the sum of all migrants per generation (Nm) across species from each source island to all potential sink islands displayed as arrow thickness. Arrows represent a range of migrants from 0.123 (Genovesa to Española) to 34.66 (San Cristóbal to Santa Fe) per generation. Bottom corner: overall directionality of migration weight is 211.186°.
In *G. scandens*, the source island also is included in the best fit model, with Marchena, Pinta, San Cristóbal, Santa Fe, Santa Cruz and Santiago all having low emigration rates (all p values < 0.001; Floreana significance not estimated as above).

**Directionality**

The two ground finch datasets showed similar directionality (*Geospiza fortis* grand mean = 172.6°, *G. scandens* grand mean = 252.5°) for both species migrating towards south-southwestern islands. In *G. fortis*, the correlation of number of migrants and direction was significant (r = 0.279, p = 0.005), while directionality was primarily in one direction as opposed to random, but not significant (F = 2.49, p = 0.09). In *G. scandens*, the correlation of migrants and direction was significant (r = 0.441, p < 0.001) and migration had a strong signal of directionality (F = 4.679, p = 0.012) (Supplementary material Appendix 1 Table A2).

**Discussion**

The endemic fauna of the Galápagos Archipelago exhibit a range of patterns of isolation within the archipelago. Many species, such as the Galápagos mockingbird *Mimus parvulus*, exhibit small radiations of lineages characterized by lack of gene flow between island populations (Arbogast et al. 2006, Parent et al. 2008). Darwin’s finches, however, are entirely different, with a relatively large radiation of lineages characterized by sympathy and metacommunity structure. Understanding this radiation, in particular, sheds light on evolutionary processes linked to isolation and gene flow in a structured landscape leading to rapid speciation.

Cessation of gene flow lies at the core of our understanding of the process of speciation (Slatkin 1987). Reduced migration within a structured population, therefore, may be the first step towards new lineages arising within a radiation. The warbler finches, for example, may be poised for further speciation due to low median migration rates and significant population genetic structure within both *Certhidea* lineages (Farrington and Petren 2011, Farrington et al. 2014). Counter-intuitively, these lineages show little morphological variation despite deep genetic divergences within lineages (Tonnis et al. 2005, Farrington and Petren 2011) and the first divergence from the rest of the radiation approximately one million years ago (Lamichhaney et al. 2015). If larger taxon cycle forces are acting within this system, we would expect younger lineages to have widespread and well-connected populations and older lineages to exhibit highly structured populations with limited dispersal (Wilson 1961, Jønsson et al. 2014). The basal warbler finches certainly fit this pattern, as do the youngest radiations: the tree finch and ground finch lineages (Lamichhaney et al. 2015). Tree finches and ground finches show patterns of many moderately-connected islands with small quartile boxes and few populations highly connected or entirely disconnected populations (Fig. 2). This contrasts with groups such as the warbler finches, which have many populations with no estimated dispersal and other populations with high dispersal (Fig. 2).
When both overall finch metacommunity patterns and species-specific differences are considered, the interaction between large-scale landscape effects on gene flow and species-specific variation linked to ecology or chance can be better interpreted than by studying each species in isolation. In the larger multi-species dataset, there is a non-random migration bias towards the south-southwest, primarily from low islands to high islands. Furthermore, distance and area were

![Figure 3](image3.png)

Figure 3. Migration estimates from *G. fortis* and *G. scandens* SNP datasets. Box-and-whisker plots show the median (horizontal line), upper and lower quartiles (box) and maximum and minimum values (excluding outliers represented by a dot). Data points for each migration estimate are shown in translucent circles. Migration rates (M) were similar, but the species have vastly different effective population sizes (Theta), yielding significantly different numbers of migrants per generation (Nm).

![Figure 4](image4.png)

Figure 4. Estimates of migrants per generation for *G. fortis* and *G. scandens* from the SNP datasets. Arrow thickness is proportional to the number of migrants per generation (Nm). The medium ground finch *G. fortis* has fairly consistent numbers of migrants between islands, while the common cactus finch *G. scandens* has relatively low migration except for a few well connected islands.
not strongly correlated to the number of migrants between islands. Due to limitations with fewer loci in microsatellite datasets, the species-specific resolution may not be as precise as in the SNP dataset, but this pattern is counter to previous assumptions about migration in island systems. Though our initial hypotheses from other island systems suggested: 1) migration should be from older islands to younger islands – straight west in the Galápagos Archipelago; 2) central islands should send migrants to the peripheral islands as assumed by peripheral speciation models; and 3) tall and large islands would be sources of emigrants due to their greater range of habitats and lower climate stochasticity, we found that none of these fit our data. Additionally, the direction of migration is nearly opposite of the predominant north-northwest winds throughout the year (Trueman and d’Ozouville 2010).

Comparative patterns of population structure within a heterogeneous island system may reveal underlying evolutionary processes responsible for adaptive radiations. Within the metacommunity, this study found a major influx of individuals from the small and harsh peripheral islands to larger central islands with stable climates. While potentially counterintuitive, this pattern may be particularly relevant to species that occupy particularly stochastic and ephemeral landscapes. The Galápagos Islands are famous for their intense selective pressures related to their oscillating environmental conditions (Grant and Grant 2006). While studies on highly stochastic and ephemeral metapopulations are still rather recent, recent work has highlighted the importance of peripheral populations for maintaining crucial traits in boom-bust cycles (Altermatt and Ebert 2010). Insights gained from comparative studies in ephemeral systems will be critical for future conservation work as environments become less stable.

The observed patterns of structured gene flow for this metacommunity may be due to either intrinsic or extrinsic factors. An example of intrinsic factors affecting speciation could be extreme selective pressure on peripheral islands, which leads to beak morphologies tightly tied to diet partitioning in those islands. Beak traits tend to converge on the central islands during high rain years (Grant 1999, Grant and Grant 2011, 2014), making birds less adapted to divergent beak shape optima. When peripheral birds colonize central islands after periods of admixing, they are likely to outcompete their ‘softer’ conspecifics as seen in other structured populations with variable selection intensity across populations (Petren et al. 2005, Altermatt and Ebert 2010).

Extrinsically, birds moving from a low island to a high island are likely to encounter conditions more conducive for survival and reproduction while birds accustomed to consistently moist habitat on high islands may fare poorly on drier, drought prone, low islands. Alternatively, the small peripheral islands may have larger boom-bust cycles linked to the more extreme seasonality on these islands (Grant 1999, Grant et al. 2000, Grant and Grant 2002). Periodic crowding followed by periods of low resources and limited options for within-island dispersal may increase overwater dispersal towards the larger and more environmentally stable central islands. Other factors may affect migration bias but are difficult to assess including the greater visibility of larger islands on the horizon compared to smaller islands and whether harsher conditions may cause birds from smaller islands to migrate more frequently in search of more favorable conditions. This visual draw possibility is supported by the significant interaction term between distance and ‘receiving island elevation’ where ‘close’ islands had no relationship between the elevation of the receiving island and the number of migrants, but middle distance and far islands had increasing effects on receiving island height and the number of migrants moving towards them (Supplementary material Appendix 1 Fig. A4).

Looking beyond the overall metapopulation of Darwin’s finches, even closely-related species that are co-distributed on the same extensive set of islands (G. scandens and G. fortis) can have markedly different numbers of successful dispersers flying between islands leading to divergent evolutionary trajectories. Despite having nearly identical migration rates (as Nm accounts for the number of successful individuals that go on to contribute to the gene pool of their destination and not simply the number of birds that disperse), species-specific differences in effective island population sizes between these lineages lead to far less gene flow in G. scandens than its congener (Fig. 3). This genetic result is corroborated by morphology, as G. scandens has much higher morphological variation between populations than seen in populations of G. fortis (Lack 1940, Grant 1999). Lower gene flow, coupled with small effective population sizes and heterogeneous environmental conditions and competition environments on each island may be leading to enhanced speciation potential in G. scandens and also greater vulnerability to local extinctions. Contributing to the effects of smaller effective population sizes and observed low abundance in ground finch communities in the field (Abbott et al. 1977, Dvorak et al. 2012), G. scandens has a narrow dietary niche (cactus seeds) which limits dispersal success to finding cactus forests on each island to inhabit. As these are often small and marginal habitats on many islands, migration estimates may reflect a lower success rate instead of simply fewer individuals dispersing. These variations highlight the need for further in-depth research on lineages within this radiation to determine the role of species-specific ecology in the overall dispersal patterns across the archipelago. Further, conservation considerations for these species can be enhanced by a better understanding of population genetic diversity and gene flow within these structured populations. Species with lower effective population sizes in their populations and less genetic connectivity between populations, such as G. scandens, can be seen as needing additional conservation protection for divergent forms.

Through a comparative framework of interisland gene flow, new insights are possible concerning overarching patterns of population connectivity within structured landscapes such as the Galápagos Archipelago. However, even in a young radiation of forms such as Darwin’s finches, this approach has limitations. Species-specific ecological differences, such as niche specialists and generalists may have entirely different
potential areas of occupation across the archipelago, making each island a mosaic of potential areas instead of a single unit as this model considers. Further, species with very little gene flow contribute relatively less data to overall patterns (such as directionality) than species with significant gene flow around the archipelago. Species- or clade-specific analyses of directionality (Supplementary material Appendix 1 Fig. A6) highlight that some species move in all directions (e.g. ground and tree finches) while others have one or two main directions of dispersal (e.g. vegetarian finch and *G. difficilis*). In the future, further studies with additional information on ecology, niche breadth and dispersal ability will enhance the understanding of divergence and evolutionary potential in complex landscapes such as these. Adaptive ecological radiations are unique in their ability to inform on complex evolutionary processes. By clarifying evolutionary patterns and processes in these systems, we move closer to understanding the interplay between landscape and ecology in driving evolutionary processes.

**Acknowledgements** – We thank the Charles Darwin Research Station and the Charles Darwin Foundation for logistical support. We also thank the many researchers and museums that have collaborated to provide samples over this long-running analysis of this system. Two anonymous reviewers greatly improved this manuscript. Support was provided by the Univ. of Cincinnati and the McMicken College of Arts and Sciences.

**Permits** – We thank the Galápagos National Park Service for permission to conduct this research.

**References**


Grant, P. R. and Grant, B. R. 2014. 40 Years of evolution: Darwin's finches on Daphne Major Island. – Princeton Univ. Press.


Supplementary material (available online as Appendix ecog-04511 at <www.ecography.org/appendix/ecog-04511>). Appendix 1.