Title: Low genomic divergence and high gene flow between locally adapted populations of the swamp sparrow

Authors: P. Deane-Coe,^{1,2*} R. Greenberg,^{3§} I. J. Lovette^{1,2} and R. G. Harrison.^{1§}

Affiliations: ¹Department of Ecology and Evolutionary Biology, Cornell University, Ithaca NY 14853 USA; ²Fuller Evolutionary Biology Program, Cornell Laboratory of Ornithology, Ithaca NY 14850 USA; ³Migratory Bird Center, Smithsonian Conservation Biology Institute, National Zoological Park, Washington DC 20008; [§]Posthumous

Corresponding author: *P. Deane-Coe (ped36@cornell.edu). 215 Tower Rd, Corson Hall, Department of Ecology and Evolutionary Biology, Cornell University, Ithaca NY 14853 USA.

Running head: Divergence and gene flow in swamp sparrows

Keywords: Local adaptation, adaptive divergence, natural selection, ecological speciation, hybridization

Abstract

Populations that have recently diverged across sharp environmental gradients provide an opportunity to study the mechanisms by which natural selection drives adaptive divergence. Inland and coastal populations of the North American swamp sparrow have become an emerging model system for studies of natural selection because they are morphologically and behaviourally distinct despite a very recent divergence time (<15,000 years), yet common garden experiments have demonstrated a genetic basis for their phenotypic differences. We characterized genomic patterns of variation within and between inland and coastal swamp sparrows via reduced representation sequencing in order to reconstruct the contributions of demography, gene flow and selection to this case of recent adaptive divergence. Compared to inland swamp sparrows, coastal swamp sparrows exhibited fewer polymorphic sites and reduced nucleotide diversity at those sites, indicating that a bottleneck and/or recent selective sweeps occurred in that population during coastal colonization and local adaptation. Estimates of genome-wide differentiation ($F_{ST}=0.02$) and sequence divergence ($\Phi_{ST}=0.05$) between inland and coastal populations were very low, consistent with postglacial divergence. A small number of SNPs were strongly differentiated (max $F_{ST}=0.8$) suggesting selection at linked sites. Swamp sparrows sampled from breeding sites at the habitat transition between freshwater and brackish marshes exhibited high levels of genetic admixture. Such evidence of active contemporary gene flow makes the evolution and maintenance of local adaptation in these two populations even more notable. We summarize several features of the swamp sparrow system that may facilitate the maintenance of adaptive diversity despite gene flow, including the presence of a magic trait.

Introduction

Birds offer many striking examples of local adaptation, particularly in terms of two evolutionarily labile traits: bill shape and plumage. Both traits are capable of substantial phenotypic change in response to natural selection, over both shallow and deep phylogenetic timescales (eg. Kusmierski et al. 1997; Lovette et al. 2002; Burns et al. 2002). In many cases, bill or plumage traits likely evolve via post-speciation selection, when allopatric taxa that are already reproductively isolated experience different forms of selection on these phenotypes depending on their environment (Dobzhansky 1940). Alternatively, when natural selection itself is responsible for driving genetic divergence between lineages during the evolution of local adaptation, we refer to the process as adaptive divergence. Since adaptive divergence may lead to the evolution of reproductive isolation (“ecological speciation”; Mayr 1963; Harrison 1991; Rice and Hostert 1993; Funk 1998; Schluter 2000, 2001; Coyne and Orr 2004; Nosil et al. 2005; Rundle and Nosil 2005; Nosil 2012), case studies of the early stages of adaptive divergence offer an opportunity to understand the mechanisms by which natural selection drives diversification over time. How does selection create and maintain locally adapted lineages before reproductive isolation has evolved between them?

Ecologically distinct populations of the North American swamp sparrow (*Melospiza georgiana*) have served as a natural laboratory for studies of adaptive divergence for decades. In the eastern USA there is both a widespread inland form of the swamp sparrow (*M. g. georgiana*), and a range-restricted coastal form (*M. g. nigrescens*) that breeds in brackish tidal marshes of the Delaware and Chesapeake Bays (Figure 1; Greenberg and Droege 1990; Beadell et al. 2003). “Coastal Plain” swamp sparrows were first recognized as distinct from inland swamp sparrows more than 60 years ago (Bond and Stewart 1951), and various phenotypes that distinguish the two forms have since been studied extensively. The coastal population differs from inland populations in at least ten traits (Table 1), including a deeper bill and melanic plumage (Figure 2; Greenberg and Droege 1990). Field studies have demonstrated that at least eight of the ten derived traits that distinguish coastal swamp sparrows are adaptive in coastal environments

(Table 1; Goldstein et al. 2004; Grenier and Greenberg 2005; Olsen et al. 2008, 2013; Peele et al. 2009; Greenberg et al. 2012), and common garden rearing experiments have confirmed that there is a strong genetic basis to most trait differences between the populations, including bill and plumage (Table 1; Ballentine and Greenberg 2010). Natural selection for deeper bills in coastal habitats has, in turn, biomechanically constrained the range of possible bandwidths at which coastal males can trill (Ballentine 2006), driving further population divergence through nonrandom mating. Inland swamp sparrow females prefer to mate with males that sing broad-bandwidth songs and maximize vocal performance, but coastal swamp sparrow females prefer the songs of coastal males, who compensate for a loss of bandwidth by increasing trill rate (Ballentine et al. 2013a, 2013b). Swamp sparrows are the third empirical example of bill shape as a “magic trait” (Gavrilets 2004): natural selection on the bills of Darwin’s finches and crossbills also imposes biomechanical constraints on song, facilitating divergence in this mating signal and driving the evolution of reproductive isolation (Podos 2001; Podos et al. 2004; Huber and Podos 2006; Huber et al. 2007; Smith and Benkman 2007).

Although many of the adaptive traits in coastal swamp sparrows have a heritable genetic basis, coastal birds have not had much time to evolve them. Coastal tidal marshes did not exist in the northeastern or mid-atlantic regions during the Last Glacial Maximum (LGM) because northeastern coasts were under the ice sheet, and mid-atlantic coasts were hydrologically unstable due to glacial outflow from rivers and streams (Malamud-Roam et al. 2006). Tidal marsh habitats also require sediment accretion from rising sea levels in order to establish (Pethick 1984; Warren and Niering 1993). Coastal swamp sparrows have therefore likely colonized and adapted to this novel habitat rapidly, within the last 15,000 years. Consistent with a recent divergence, inland and coastal swamp sparrows have proven indistinguishable using allozymes and mitochondrial DNA (Balaban 1988; Greenberg et al. 1998). A recent study using microsatellite markers was the first to successfully resolve subtle genetic differentiation between coastal and inland swamp sparrows, and a genetic contact zone in northern New Jersey coincident with the ecotone between brackish and freshwater marshes (Greenberg et al. 2016).

To reconstruct the evolutionary processes responsible for adaptive divergence between inland and coastal swamp sparrows, we quantify genome-wide molecular variation for the species and extract patterns of historical demography, gene flow and selection. We also test for an association between phenotypic and genomic divergence, and characterize levels of admixture in the putative inland/coastal hybrid zone with genome-level sampling.

Figure 1. Breeding range of Swamp Sparrow subspecies (Greenberg and Droege 1990; Beadell et al. 2003). Freshwater (°) and brackish (•) sampling sites are shown with two-letter site codes (Table 2). Hatching denotes an area of very low breeding density in which the phenotypes of birds have not been studied. Dashed line denotes the approximate southern extent of the boreal subspecies (*ericrypta*), not included in this study.

Figure 2. Bill depth (A) and the extent of black crown plumage (B) are two of many traits that subtly distinguish inland (blue) and coastal (red) swamp sparrows. Morphological data include only birds from allopatric sites (N=29) and are from Greenberg and Droege (1990). Photos are of breeding males (inland: Hamilton County NY and Assunpink NJ, coastal: Fishing Bay MD).

Table 1. Proposed adaptive hypotheses for eight traits that distinguish coastal Swamp Sparrows from inland forms, and their empirical support from field studies. Trait differences indicated by an asterisk (*) persist when coastal and inland birds are reared in a common garden. Clutch size could not be measured in captivity (Ballentine and Greenberg 2010).

Trait	Hypotheses	Proposed selective agent	Proposed adaptive benefit	Support
↑ Bill depth*	<i>Trophic adaptation</i>	More abundant invertebrate prey in coastal habitat	Improves prey capture ^{1,2,3}	some
	<i>Abiotic adaptation</i>	Heat stress due to exposed, water-limited habitats	Convective radiative heat loss ⁴	Y
↑ Bill dimorphism	<i>Thermal budgets</i>	Territorial defense causes heat stress in males	Convective radiative heat loss ⁴	some
	<i>Intraspecific divergence</i>	High breeding densities, competition for resources	Sex-based diet specialization ⁵	N
	<i>Intrasexual selection</i>	High breeding densities, territorial competition	Competitive advantage ⁵	unclear
	<i>Female choice</i>	High breeding densities, competition for paternity	Females prefer larger bills ³	Y
↑ Plumage melanism*	<i>Predation adaptation</i>	Anoxia and iron sulphides cause darker sediments	Crypsis ¹	untested
	<i>Feather preservation</i>	Feather degrading bacteria thrive in saline habitats	Confers resistance to bacteria ⁶	Y
	<i>Intrasexual selection</i>	High breeding densities, territorial competition	Larger black plumage badges ⁷	Y
↓ Plumage rustiness*	<i>Predation adaptation</i>	Anoxia and iron sulphides cause darker sediments	Crypsis ¹	untested
	<i>Evolutionary constraint</i>	Selection for plumage melanism	None, females prefer rusty crowns ^{7,8}	Y
Delayed molt, extended breeding*	<i>Predation adaptation</i>	High predation on coast causes nest failure	Allows re-nesting ⁹	some
	<i>Abiotic adaptation</i>	Coastal habitats experience late spring flooding	Facilitates delayed breeding ¹	some
↓ Clutch size	<i>Predation adaptation</i>	Nest predation rates are higher in coastal habitats	Reduces provisioning, detectability ⁹	Y
	<i>Longevity trade-off</i>	Lifetime breeding effort favored over current	Improves annual survivorship ⁹	N
	<i>Abiotic adaptation</i>	Temperature extremes, egg loss during laying	Shorter laying avoids extreme events ⁹	Y
↓ Song bandwidth*	<i>Evolutionary constraint</i>	Larger bills constrain the bandwidth of male trills	None, broad bandwidth favored ^{10,11,12}	Y
↑ Song trill rate	<i>Acoustic adaptation</i>	Faster trills transmit better in open habitats	Improves signal transmission ^{13,14}	some
	<i>Female choice</i>	Faster songs are honest indicators of quality	Females prefer fast trills ^{10,12,15,16}	Y
	<i>Signal compensation</i>	Bill size constrains bandwidth, males trill faster	Females prefer challenging songs ^{10,11}	some

1. Greenberg and Droege 1990 2. Grenier and Greenberg 2005 3. Olsen et al. 2013 4. Greenberg et al. 2012 5. Greenberg and Olsen 2010 6. Peele et al. 2009 7. Olsen et al. 2008a 8. Olsen et al. 2010 9. Olsen et al. 2008b 10. Ballentine 2006 11. Ballentine et al. 2013a 12. Liu et al. 2008 13. Morton 1975 14. Wiley and Richards 1978 15. Ballentine 2009 16. Ballentine et al. 2013b

Methods

Sample collection and phenotypic measurements. We compiled 92 DNA samples from male and female swamp sparrows from 15 sites distributed across the breeding range of inland and coastal forms in the northeastern US, including samples from the putative contact zone in northern New Jersey (Figure 1; Tables 2, S1). Most were blood samples collected from birds banded in the 2001 breeding season (N=66), and the remainder were blood samples from banded birds or tissue samples from voucher specimens, collected by P. Deane-Coe (P.D.) or other collectors from the Cornell University Museum of Vertebrates during the 2007-2014 breeding seasons (N=26). We collected four voucher specimens and six blood samples from a previously unstudied coastal population at Fishing Bay Wildlife Management Area (MD) on the eastern shore of the Chesapeake Bay, representing the most southern breeding location sampled to date (Table S1). We also took a complete set of high resolution photographs for all birds banded in 2012-2014, using a Canon DSLR camera against a standardized color reference target (Xrite Digital SG Colorchecker), and collected standard morphological data for all banded birds.

Library preparation and sequencing. We extracted DNA using the DNeasy kit (Qiagen), quantified concentrations on a Qubit (Invitrogen, Carlsbad, CA, USA), and balanced concentrations across samples using either dilution or vacuum filtration. We digested 175ng of each sample with enzymes SBf1 and Msp1. We followed standard ddRAD protocols for digestion, Illumina TruSeq primer ligation and clean-up as described by Peterson et al. (2012). We chose a combination of eight index groups and 12 barcodes, each 5 or 6 bases in length (Table S2). Prior to barcode ligation, we selected fragments from 400-600bp in size using the Pippin protocol (Sage Science, MA, USA). Barcode sequences were semi-randomized by systematically distributing individuals from each sampling location across index groups and barcode identities (Table S2). Within index groups, we ligated barcodes with either 14 or 16 cycles of PCR and visually inspected the products of these reactions on a gel to confirm that amplification error had not distorted the size distribution of fragments. We balanced DNA

concentration to 2nM across index groups following quantification on a Bioanalyzer, and pooled index groups. We used two lanes of an Illumina HiSeq in Rapid Run mode for sequencing (yielding over 200 million 150bp reads), and we performed quality filtering and individual genotyping using a custom modification of the STACKS pipeline (Catchen et al. 2011). In brief, we executed `process_radtags` commands to demultiplex reads and used fastX tools for custom processing. We clipped reads with insert sizes short enough to contain adapter sequence, and trimmed all reads to 145bp. The first eight base pairs of each read did not meet quality standards due to low complexity across index sequences, and were also trimmed. We determined that trimming the sequence containing the enzyme cut site did not negatively affect the performance of the STACKS assembly by performing multiple parallel processing runs. In fact, this step improved the quality of the assembly by removing sections of low quality sequence and preserving the high quality remainder of each read. We caution, however, that this modified assembly approach may not work as efficiently in comparisons across a deeper divergence time than in our system, since successful assembly of reads from multiple individuals will rely on sequence similarity downstream of the cut site.

SNP genotyping. We quality-filtered reads, retaining those with Illumina quality scores above 24, and the resulting 137bp fragments were assembled into a catalogue of 72,246 unique loci according to a minimum depth of five reads per stack, two locus mismatches allowed within individuals, and four locus mismatches allowed in the catalog. We dropped two individuals from the northern part of the inland range (TCm4, TCu1; both sampled far from the hybrid zone) prior to SNP discovery because they possessed unusual and divergent haplotypes at several loci, representing either sequencing error or gene flow from the unsampled boreal subspecies *M. g. ericrypta*. We grouped remaining individuals (N=90) into three populations for SNP discovery, according to whether they originated from the allopatric range of either the inland (sites HC, GE, TC, ER, GC) or coastal (WB, BH, PM, FB) populations, or from sites near the contact zone between the two (HA, NE, ME, GS, CH, EH). After extensive exploration of the impact of

missing data on the accuracy of population genetic estimates from our data, we conservatively required that a given SNP be present in at least 50% of the individuals in each of the three populations. A total of 29,058 single-nucleotide sites on 4,256 stacks that met these criteria were polymorphic across allopatric inland, allopatric coastal, and sites near the contact zone. We selected the first SNP from each remaining locus to define a panel of independent SNPs for subsequent analyses of genome-wide variation and divergence (N=4,238 after additional SNP quality filtering). We performed an Analysis of Molecular Variance (AMOVA; Meirmans 2006) in the program ARLEQUIN (Excoffier and Lischer 2010) to assess the distribution of total genetic variation across different hierarchical categories of population structure. We extracted a range of different population genetic metrics and locus-specific allele frequencies from the output of STACKS, including F_{ST} (a relative measure of differentiation in allele frequencies; Wright 1965; Weir and Cockerham 1984) and Φ_{ST} (a relative measure of mutational distance between populations; Meirmans and Hedrick 2010). We also tested for F_{ST} outliers using Bayescan (Foll and Gaggiotti 2008).

Characterizing genome-wide patterns of variation and divergence. To determine whether genome-wide SNP variation reflected evolutionary distinctness of swamp sparrows from freshwater and coastal marshes, we performed k-means clustering on the results of a Principle Coordinates Analysis (PCA) using Discriminant Analysis of Principle Components (DAPC) (Adegenet; Jombart et al. 2010), and performed bayesian assignment tests using STRUCTURE (Pritchard et al. 2000). We also performed k-means clustering manually in R, to confirm the results of the DAPC algorithm. To resolve the location and extent of admixture, we estimated the probability of assignment for all individuals in the dataset to the significant genetic clusters determined in the previous step, comparing the assignments generated by both DAPC and STRUCTURE.

Testing morphological associations. We tested for significant associations between phenotypic scores for adaptive traits (bill depth, length, width, volume and extent of black crown plumage) and discriminant scores based on genome-wide SNP variation using Spearman Rank correlation (correlation coefficient = ρ). For consistency, only birds sampled in 2001 and measured by R. Greenberg were used in these analyses. We performed analyses of covariance (ANCOVA) to test for significant associations between morphology and discriminant score for any traits that significantly scaled with weight. We also performed T-tests to assess the degree of morphological differentiation between distinct genetic clusters inferred from DAPC.

Table 2. Locations and dates for Swamp Sparrows sampled in this study (N=92).

Code	Sampling locations	N	Date
HC	Hamilton County, NY	3	2013
TC	Tompkins County, NY	12	2008-2014
GE	Genesee County, NY	1	2007
ER	Erie National Wildlife Refuge, PA	8	2001
GC	Garrett County, MD	8	2001
HA	Hawley, PA	3	2001
NE	Newton, NJ	6	2001
ME	Meadowlands, NJ	6	2001
GS	Great Swamp National Wildlife Refuge, NJ	10	2001
CH	Cheesequake State Park, NJ	5	2001
EH	Egg Harbor, NJ	1	2001
WB	Woodland Beach, DE	2	2001
BH	Bombay Hook, DE	8	2001
PM	Port Mahon, DE	9	2001
FB	Fishing Bay Wildlife Management Area, MD	10	2014

Results

Genome-wide differentiation and divergence. Out of the 29,058 SNPs in the complete dataset, only 11% (3,126 SNPs) showed significant allele frequency differences (corrected $F_{ST} > 0$) between coastal and inland swamp sparrows. Our panel of 4,238 independent SNPs also reflected overall similarity between coastal and inland genomes (median $F_{ST}=0.00$, mean $F_{ST}=0.02$), and sequence divergence was similarly low (median $\Phi_{ST}=0.03$, mean $\Phi_{ST}=0.05$). There was a long tail to the distributions of both F_{ST} and Φ_{ST} (max $F_{ST}=0.80$; max $\Phi_{ST}=0.77$), representing SNPs with large allele frequency differences (F_{ST}), and loci with large mutational distances between inland and coastal haplotypes (Φ_{ST}), but Bayescan did not identify any of these as significant outliers. An AMOVA including only allopatric coastal or inland sites indicated that most of the total genetic variance is partitioned across individuals (76%) or across individuals within sites (19%), providing further indication that most variable sites within the genome are not structured between coastal and inland swamp sparrows. Higher variation within groups than between groups is commonly observed in investigations of recently diverged taxa (e.g., Campagna et al. 2015). A very low proportion of the variance (1.5%) could be attributed to structure between breeding sites within each population, indicating little to no role for isolation by distance or other geographic or demographic processes operating within coastal or inland ranges. Consistent with the presence of a long tail to the distributions of F_{ST} and Φ_{ST} , however, a small subset of genome-wide variation (3%) was partitioned between coastal and inland habitats.

The program STRUCTURE did not resolve strong support for two distinct genetic groups (best $K=1$), representing genetic variation within swamp sparrows as a gradual cline in allele frequencies across many individuals of intermediate assignment when $K=2$ was enforced (Figure 3). In contrast, according to DAPC, total genomic variation within our samples was best represented by a single discriminant axis from a scaled PCA, and the lowest Bayesian Information Criterion (BIC) scores supported two genetic clusters corresponding to birds sampled from inland freshwater vs. coastal brackish marshes (Figure 3, S1). A conventional PCA

with manually applied k-means clustering also robustly supported the genetic distinctness of swamp sparrows from coastal vs. inland habitats.

Molecular variation within each population. As a group, birds sampled from coastal brackish tidal marshes possessed fewer polymorphisms across the genome than those from inland freshwater marshes (2.8% vs. 3.7% of sites), and had fewer private alleles (2,367 vs. 6,473). Average nucleotide diversity was also subtly but significantly lower in coastal birds compared to inland birds (inland $\pi = 0.11 \pm 0.14$, coastal $\pi = 0.10 \pm 0.15$; $p = 3.2 \times 10^{-7}$), however, despite a smaller population size coastal birds exhibited more negative values of F_{IS} than inland birds (inland $F_{IS} = -2.0$, coastal $F_{IS} = -3.2$; $p = 6.3 \times 10^{-29}$), suggesting that inbreeding is not responsible for this loss of variation.

Location and extent of admixture. The two methods used to infer individual assignment probabilities (STRUCTURE and DAPC) exhibited discordance in terms of the extent of admixture, but both did exhibit concordance with respect to the location of the zone of admixture, which occurs in and around the ecotonal transition between freshwater and brackish habitats in northern NJ (Figure 4). Birds from two sites (ME and CH) had highly heterogeneous individual assignments to one cluster or the other according to DAPC, consistent with the wide range of discriminant axis values attributed to individuals from these northern brackish habitats (Figure 3, 4). Heterogeneity across individuals at these sites was also observed from the assignments calculated in STRUCTURE.

Figure 3. Discriminant analysis of 22 principal components, via K-means clustering and model selection, identified two clusters of genetically related swamp sparrows from a panel of 4,238 independent SNPs. A single axis of variation best distinguished the clusters. The x-axis represents the placement of individuals (short vertical lines) along this discriminant axis. The y-axis is a scaled metric representing the smoothed density of individuals corresponding to each discriminant value. Assignment probabilities consistently placed swamp sparrows from inland freshwater sites (blue) and coastal brackish sites (red) in separate clusters, with the exception of birds from two brackish sites in northern NJ (black). Individuals from these sites exhibited highly heterogeneous assignments to one cluster or the other. Brackets represent the range of genetic variation sampled from each site. Refer to Table 2 for site codes.

Figure 4. Assignment probabilities of individual swamp sparrows to one of two genetic clusters, computed by the programs STRUCTURE (top) and DAPC (bottom). Refer to Figure 1 and Table 2 for two-letter site codes. Black line indicates the ecotone between freshwater and brackish habitats.

Associations between phenotype scores and genomic assignment. Bill length, bill width and wing chord measures were not significantly correlated with the placement of individuals on the discriminant axis (Figure S1). Correlations between DAPC discriminant axis score and morphological score were significant for the following traits: Bill depth ($q=0.46$, $p<0.001$), bill volume ($q=0.35$, $p<0.05$), extent of black crown plumage ($q=0.29$, $p<0.05$), tarsus ($q=0.20$, $p<0.05$) and weight ($q=0.33$, $p<0.05$). Bill depth and volume did not significantly scale with tarsus length, a common proxy for overall size; however both metrics did scale with weight (Figure S2; bill depth, $q=0.41$, $p<0.01$; bill volume, $q=0.49$, $p<0.001$). We conducted analyses of covariance (ANCOVA) to test for an association between phenotype and discriminant axis score when the confounding effect of weight was removed, and both relationships remained strong and significant ($p<0.001$). Since variation in bill depth influences bill volume, and bill width and length were not significant, we only show the data for bill depth and black crown plumage (Figure 5). Both bill depth and black crown plumage also significantly differed between the two distinct genetic clusters inferred from DAPC (corresponding to inland and coastal sites).

Figure 5. Morphological traits distinguishing inland and coastal swamp sparrows, and their association with genomic discrimination scores (left) and cluster assignment (right). Both bill depth (A) and the extent of black crown plumage (B) positively correlate with DAPC score. Birds assigned to the coastal cluster had significantly deeper bills and more extensive black crown plumage than those assigned to the inland cluster ($p < 0.001$ and $p < 0.05$, respectively).

Discussion

Recent adaptive divergence between inland and coastal swamp sparrows. In our comparisons, genome-wide patterns of variation supported previous evidence that inland and coastal swamp sparrows are very closely related, but have diverged very recently and are now on distinct evolutionary trajectories. We found evidence of population genetic structure corresponding to inland and coastal subspecies, but it was sufficiently low to be consistent with divergence during the most recent 15,000 years, and it was only detected by the more sensitive of the two methods of genetic cluster assignment (DAPC inferred two significant clusters but STRUCTURE did not). Across all swamp sparrows sampled, degrees of morphological variation did correlate with degrees of genome-wide differentiation, and both bill depth and black crown plumage were significantly associated with the genomic cluster to which birds were assigned.

Divergent SNPs despite a highly similar genomic background. Inland and coastal swamp sparrow genomes are highly similar, but a long tail in the F_{ST} distribution also suggests selection in regions of the genome linked to these sites. Bayescan detected no significant outliers, however, which is not surprising given the subtle phenotypic divergence in question: Outlier tests scan for sites with reciprocally fixed or nearly fixed haplotype frequencies, and the adaptive and heritable phenotypes that differentiate inland and coastal swamp sparrows are not fixed in either population. Rather, bill depth and plumage melanism in coastal birds is an exaggerated subset of the total trait variation among inland birds (Figure 2; Greenberg and Droege 1990). Detection of functional variation responsible for adaptive phenotypes will require much more extensive genomic sampling, both to increase the proportion of the genome scanned and to enhance our ability to detect significant deviations in allele frequencies due to selection by more tightly estimating the background distribution of F_{ST} across the rest of the genome.

Evidence of a founder effect during coastal colonization. Compared to inland swamp sparrows, coastal swamp sparrows exhibited fewer polymorphic sites, reduced nucleotide diversity at those

sites, and fewer private alleles. This pattern of reduced molecular variation is not likely due to post-colonization genetic drift in the smaller coastal population since average F_{IS} is negative in coastal birds, indicating less inbreeding. In fact, coastal populations appear to exhibit even less inbreeding than the large inland population (more negative F_{IS}). Reduced population genetic variation on the coast without evidence of contemporary inbreeding is therefore consistent with an alternative demographic scenario: a recent population bottleneck during coastal colonization. If coastal colonists originated from the same ancestral source population as contemporary inland populations, these founders may have carried only a subset of ancestral variation with them.

Gene flow at the inland/coastal ecotone. Admixture estimates for individuals from across the northeastern US support earlier suggestions that there is a zone of contact between inland and coastal swamp sparrows in northern New Jersey, where inland freshwater marshes transition to brackish coastal marshes. Depending on the assignment algorithm used, this transition between genetically distinct groups is inferred to be either abrupt (DAPC) or gradual (STRUCTURE), reflecting different sensitivities to the subtle signal of divergence present in the data. Finer-scale geographic sampling of sparrows from within this region is ongoing and will facilitate the use of more computationally sophisticated methods to infer ancestry (Introgress, Gompert and Buerkle 2010; BGC, Gompert and Buerkle 2012; HapMix, Price et al. 2009; RASPBerry, Wegmann et al. 2011).

Local adaptation despite gene flow. In the present day, adaptive diversity appears to be maintained by selection despite gene flow since we detected highly admixed individuals at the ecotone between habitats in northern New Jersey. Two aspects of the swamp sparrow system may explain the maintenance of locally adaptive phenotypes despite active gene flow at that contact zone. First, since adaptive trait variation in coastal birds is a subset of the variation present inland (Figure 2), inland populations do harbor some coastally adaptive standing variation. Gene flow from the large inland population may therefore seed both adaptive and non-

adaptive variants into the coastal gene pool, and if selection were strong, differential survival would mean that only the adaptive variants are maintained on the coast (eg. differential habitat performance; Harrison 1986, 1990). This would be an example of immigrant inviability maintaining adaptive diversity (Nosil et al. 2005). Second, there is robust evidence that divergent natural selection on bill size drives positive assortative mating according to habitat in swamp sparrows (Balletine 2006; Balletine et al. 2013a, 2013b). This process could maintain local adaptation by constraining gene flow to a narrow band of ecotonal or transitional habitat where divergent selection on bill shape is not strong or is not present. Future research in this hybrid zone that includes explicit tests of either mechanism will provide valuable insight into the mechanisms of adaptive evolution.

Acknowledgements

The authors thank J. Searle, R. Fleischer and three anonymous reviewers for comments that improved this manuscript. The Smithsonian Conservation Biology Institute's Center for Conservation and Evolutionary Genetics (R. Fleischer), and Cornell University Museum of Vertebrates (C. Dardia, V. Rohwer) contributed samples collected by N. Perlut, N. Mason and G. Seeholzer, among others. Blackwater National Wildlife Refuge (M. Whitbeck) and Fishing Bay Wildlife Management Area (J. Moulis) provided logistical support, and A. Bessler, K. Bostwick and K. Deane-Coe provided field assistance. S. Bogdanowicz and L. Campagna guided ddRAD sequencing and assembly. This research was supported by the Athena Fund at the Laboratory of Ornithology, Cornell Vertebrate Genomics, the Frank M. Chapman Fund at the American Museum of Natural History and Andrew W. Mellon Student Research fund. Fellowship support for P.D. was provided by Cornell University, the Cornell Lab of Ornithology and the Natural Sciences and Engineering Research Council of Canada (PGSD3 405451-11). Sampling was conducted under appropriate state and federal permits and all procedures conformed to protocols approved by Cornell's Institutional Animal Care and Use Committee (IACUC). Authors Russ Greenberg and Rick Harrison passed away before the final version of this manuscript, but both were heavily involved in discussion of the ecological (R.G.) and evolutionary (R.H.) aspects of this system. They left an indelible mark on their respective fields, and are deeply missed.

References

- Balaban, E. 1988. Cultural and genetic variation in Swamp Sparrows (*Melospiza georgiana*) 1. Song variation, genetic variation, and their relationship. *Behaviour*. 105:250–291.
- Ballentine, B. 2006. Morphological adaptation influences the evolution of a mating signal. *Evolution*. 60:1936-1944.
- Ballentine, B. 2009. The ability to perform physically challenging songs predicts age and size in male swamp sparrows. *Animal Behaviour*. 77:973-978.
- Ballentine, B., and R. Greenberg. 2010. Common garden experiment reveals genetic control of phenotypic divergence between swamp sparrow subspecies that lack divergence in neutral genotypes. *PLoS One*. 5:e10229.
- Ballentine, B., B. Horton, E. T. Brown and R. Greenberg. 2013a. Divergent selection on bill morphology contributes to nonrandom mating between swamp sparrow subspecies. *Animal Behaviour*. 86:467-473.
- Ballentine, B., K. W. Gkoo and R. Greenberg. 2013b. Mechanisms of song divergence between swamp sparrow subspecies. *Behaviour*. 150:1165-1181.
- Beadell, J., R. Greenberg, S. Droege and J. A. Royle. 2003. Distribution, abundance, and habitat affinities of the Coastal Plain Swamp Sparrow. *Wilson Bulletin*. 115:38-44.
- Bond, G. M., and R. E. Stewart. 1951. A new Swamp Sparrow from the Maryland Coastal Plain. *Wilson Bulletin*. 63:38-40.
- Burns, K. J., S. J. Hackett, and N. K. Klein. 2002. Phylogenetic relationships and morphological diversity in Darwin's finches and their relatives. *Evolution*. 56(6):1240-52.
- Campagna, L., I. Gronau, L. F. Silveira, A. Siepel and I. J. Lovette. 2015. Distinguishing noise from signal in patterns of genomic divergence in a highly polymorphic avian radiation. *Molecular Ecology*. 24(16):4238-4251.
- Catchen, J. M., A. Amores, P. Hohenlohe, W. Cresko and J. H. Postlethwait. 2011. Stacks: building and genotyping loci de novo from short-read sequences. *Genes Genomes and Genetics*. 1:171-182.
- Coyne, J. A., and H. A. Orr. 2004. *Speciation*. Sinauer Associates, Sunderland, MA.
- Dobzhansky, T. Speciation as a stage in evolutionary divergence. 1940. *The American Naturalist*. 74(753):312-21.

- Excoffier, L., and H. E. L. Lischer. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetic analyses under Linux and Windows. *Molecular Ecology Resources*. 10:564-567.
- Excoffier, L., P. E. Smouse and J. M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*. 131(2):479-491.
- Foll, M., and O. Gaggiotti. 2008. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics*. 180(2):977-93.
- Funk, D. J. 1998. Isolating a role for natural selection in speciation: host adaptation and sexual isolation in *Neochlamisus bebbianae* leaf beetles. *Evolution*. 52(6):1744-1759.
- Goldstein, G., K. R. Flory, B. A. Browne, S. Majid, J. M. Ichida and E. H. Burt Jr. 2004. Bacterial degradation of black and white feathers. *The Auk*. 121:656-659.
- Gompert, Z., and C. A. Buerkle. 2010. INTROGRESS: a software package for mapping components of isolation in hybrids. *Molecular Ecology Resources*. 10:378-384.
- Gompert, Z., and C. A. Buerkle. 2012. bgc: Software for Bayesian estimation of genomic clines. *Molecular Ecology Resources*. 12(6):1168-76.
- Greenberg, R., and S. Droege. 1990. Adaptation to tidal marshes in breeding populations of the Swamp Sparrow. *The Condor*. 92:393-404.
- Greenberg, R., and J. E. Maldonado. 2006. Diversity and endemism in tidal-marsh vertebrates. *Studies in Avian Biology*. 32:32-53.
- Greenberg, R. and B. Olsen. 2010. Bill size and dimorphism in tidal-marsh sparrows: island-like processes in a continental habitat. *Ecology*. 91:2428-2436.
- Greenberg, R., P. J. Cordero, S. Droege and R. C. Fleischer. 1998. Morphological adaptation with no mitochondrial DNA differentiation in the coastal plain Swamp Sparrow. *The Auk*. 115:706-712.
- Greenberg, R., R. Danner, B. Olsen and D. Luther. 2012. High summer temperature explains bill size variation in salt marsh sparrows. *Ecography*. 35:146-152.
- Greenberg, R., K. M. Cammen, A. G. Wilson, B. J. Olsen, B. Ballentine, N. C. McInerney and R. C. Fleischer. Geographic population structure and subspecific boundaries in a tidal marsh sparrow. *Conservation Genetics* 17(3):603-613.
- Grenier, J. L., and R. Greenberg. 2005. A biogeographic pattern in sparrow bill morphology: parallel adaptation to tidal marshes. *Evolution*. 59:1588-1595.

- Groppe J., J. Greenwald, E. Wiater, J. Rodriguez-Leon, A. N. Economides, W. Kwiatkowski, M. Affolter, W. W. Vale, J. C. Izpisua Belmonte and S. Choe. 2002. Structural basis of BMP signalling inhibition by the cystine knot protein Noggin. *Nature*. 420(6916):636–42.
- Harrison, R. G.. 1990. Hybrid zones: windows on evolutionary process. Pp. 69-128 in D. Futuyma and J. Antonovics, eds. *Oxford surveys in evolutionary biology*. Oxford Univ. Press, New York, NY.
- Harrison, R. G. 1991. Molecular changes at speciation. *Annual Review of Ecology, Evolution and Systematics*. 22:281-308.
- Harrison, R. G. 2012. The language of speciation. *Evolution*. 66(12):3643-3657.
- Harrison, R. G., and E. L. Larson. 2014. Hybridization, introgression, and the nature of species boundaries. *J Heredity*. 105:795-809.
- Huber, S. K., and J. Podos. 2006. Beak morphology and song features covary in a population of Darwin's finches (*Geospiza fortis*). *Biological Journal of the Linnean Society*. 88(3):489-498.
- Jombart, T., S. Devillard and F. Balloux. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*. 11(1):94.
- Kane, N. C., M. G. King, M. S. Barker, A. Raduski, S. Karrenberg, Y. Yatabe, S. J. Knapp and L. H. Rieseberg. 2009. Genetic analyses indicate highly porous genomes and high levels of gene flow between divergent *Helianthus* species. *Evolution*. 63:2061-2075.
- Kusmierski, R., G. Borgia, A. Uy, and R. H. Crozier. 1997. Labile evolution of display traits in bowerbirds indicates reduced effects of phylogenetic constraint. *Proceedings of the Royal Society of London B: Biological Sciences*. 264(1380):307-13.
- Liu, I. A., B. Lohr, B. Olsen and R. Greenberg. 2008. Macrogeographic vocal variation in subspecies of swamp sparrow. *The Condor*. 110:102-109.
- Lovette, I. J., E. Bermingham, and R. E. Ricklefs. 2002. Clade-specific morphological diversification and adaptive radiation in Hawaiian songbirds. *Proceedings of the Royal Society of London B: Biological Sciences*. 269(1486):37-42.
- Malamud-Roam, K. P., F. P. Malamud-Roam, E. B. Watson, J. P. Collins and B. L. Ingram. 2006. The Quaternary geography and biogeography of tidal saltmarshes. *Studies in Avian Biology*. 32:11-31.
- Marshall, J. T. 1948. Ecologic Races of Song Sparrows in the San Francisco Bay Region: Part II. Geographic Variation. *The Condor*. 50(6):233-256.

- Mayr, E. 1963. Animal species and evolution. Belknap Press of Harvard University Press, Cambridge, MA.
- Meirmans, P. G. 2006. Using the AMOVA framework to estimate a standardized genetic differentiation measure. *Evolution*. 60(11):2399-2402.
- Meirmans, P. G., and P. W. Hedrick. 2010. Assessing population structure: F_{ST} and related measures. *Molecular Ecology Resources*. 11(1):5-18.
- Morton, E. S. 1975. Ecological sources of selection on avian sounds. *The American Naturalist*. 109:17-34.
- Nosil, P. 2012. Ecological speciation. Oxford Univ. Press, Oxford, U.K.
- Nosil, P., T. H. Vines and D. J. Funk. 2005. Reproductive isolation caused by natural selection against immigrants from divergent habitats. *Evolution*. 59(4):705-719.
- Olsen, B., R. Greenberg, R. C. Fleischer and J. R. Walters. 2008a.. Extrapair paternity in the the swamp sparrow, *Melospiza georgiana*: male access or female preference? *Behavioural Ecology and Sociobiology*. 63:285-294.
- Olsen, B., R. Greenberg, I. A. Liu, J. M. Fetch and J. R. Walters. 2010. Interactions between sexual and natural selection on the evolution of a plumage badge. *Evolutionary Ecology*. 24:731-748.
- Olsen, B., J. M. Felch, R. Greenberg and J. Walters. 2008b. Causes of reduced clutch size in a tidal marsh endemic. *Oecologia*. 158:421-435.
- Olsen, B. J., R. Greenberg, J. R. Walters and R. C. Fleischer. 2013. Sexual dimorphism in a feeding apparatus is driven by mate choice and not niche partitioning. *Behavioural Ecology*. 00:1-12.
- Olson N. D., S. P. Lund, R. E. Colman, J. T. Foster, J. W. Sahl, J. M. Schupp, P. Keim, J. B. Morrow, M. L. Salit and J. M. Zook. 2015. Best practices for evaluating single nucleotide variant calling methods for microbial genomics. *Frontiers in Genetics*. 6:235.
- Peele, A. M., E. H. Burt Jr, M. R. Schroeder and R. S. Greenberg. 2009. Dark color of the Coastal Plain Swamp Sparrow (*Melospiza georgiana nigrescens*) may be an evolutionary response to occurrence and abundance of salt-tolerant feather-degrading bacilli in its plumage. *The Auk*. 126:531-535.
- Peterson, B. K., J. N. Weber, E. H. Kay, H. S. Fisher and H. E. Hoekstra. 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One*. 7:e37135.

- Pethick, J. S. 1984. An introduction to coastal geomorphology. Dept. of Geography, Univ. of Hull, U.K.
- Podos, J. 2001. Correlated evolution of morphology and vocal signal structure in Darwin's finches. *Nature*. 409(6817):185-188.
- Podos, J., S. K. Huber and B. Taft. 2004. Bird song: the interface of evolution and mechanism. *Annual Review of Ecology Evolution and Systematics*. 35:55-87.
- Poelstra, J. W., N. Vijay, M. P. Hoepfner and J. B. Wolf. 2015. Transcriptomics of colour patterning and coloration shifts in crows. *Molecular Ecology*. 24(18):4617-4628.
- Price, A. L., A. Tandon, N. Patterson, K. C. Barnes, N. Rafaels, I. Ruczinski, T. H. Beaty, R. Mathias, D. Reich and S. Myers. 2009. Sensitive detection of chromosomal segments of distinct ancestry in admixed populations. *PLoS Genetics*. 5(6):e1000519.
- Pritchard, J. K., M. Stephens and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics*. 155:945-959.
- Rice, W. R., and E. E. Hostert. 1993. Laboratory experiments on speciation: what have we learned in 40 years? *Evolution*. 47(6):1637-1653.
- Rundle, H. D., and P. Nosil. 2005. Ecological speciation. *Ecology Letters*. 8(3):336-352.
- Schluter, D. 2000. The ecology of adaptive radiation. Oxford Univ. Press, Oxford, U.K.
- Schluter, D. 2001. Ecology and the origin of species. *Trends in Ecology and Evolution*. 16(7):372-380.
- Smith, J. W, and C. W. Benkman. 2007. A coevolutionary arms race causes ecological speciation in crossbills. *The American Naturalist*. 169(4):455-65.
- Warren, R. S., and W. A. Niering. 1993. Vegetation change on a Northeast tidal marsh: interaction of sea-level rise and marsh accretion. *Ecology*. 74(1):96-103.
- Weir, B. S., and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. *Evolution*. 38(6):1358-1370.
- Wegmann, D., D. E. Kessner, K. R. Veeramah, R. A. Mathias, D. L. Nicolae, L. R. Yanek, Y. V. Sun, D. G. Torgerson, N. Rafaels, T. Mosley and L. C. Becker. 2011. Recombination rates in admixed individuals identified by ancestry-based inference. *Nature Genetics*. 43(9):847-853.
- Wiley, R. H., and D. G. Richards. 1978. Physical constraints on acoustic communication in atmosphere: implications for evolution of animal vocalizations. *Behavioural Ecology and Sociobiology*. 3:69-94.

Wright, S. 1965. The interpretation of population structure by F-statistic with special regard to systems of mating. *Evolution*. 19(3): 395-420.