6-2016

Genome-wide differentiation in closely related populations: The roles of selection and geographic isolation

Rebecca Safran  
*University of Colorado, Boulder, rebecca.safran@colorado.edu*

Elizabeth Scordato  
*University of Colorado, elizabeth.scordato@colorado.edu*

Matt Wilkins  
*University of Nebraska-Lincoln, m.r.wilkins06@gmail.com*

Joanna K. Hubbard  
*University of Nebraska-Lincoln, jkhubbard@unl.edu*

B. R. Jenkins  
*University of Colorado, Boulder*

*See next page for additional authors*

Follow this and additional works at: [http://digitalcommons.unl.edu/bioscifacpub](http://digitalcommons.unl.edu/bioscifacpub)

Part of the [Biology Commons](http://digitalcommons.unl.edu/bioscifacpub), and the [Genomics Commons](http://digitalcommons.unl.edu/bioscifacpub)

Safran, Rebecca; Scordato, Elizabeth; Wilkins, Matt; Hubbard, Joanna K.; Jenkins, B. R.; Albrecht, Tomáš; Flaxman, Samuel; Karaardiç, H.; Vortman, Yoni; Lotem, Arnon; Nosil, Patrik; Pap, P.; Shen, Sheng-Feng; Chan, S.-F.; Parchman, T. L.; and Kane, Nolan C., “Genome-wide differentiation in closely related populations: The roles of selection and geographic isolation” (2016). Faculty Publications in the Biological Sciences. 539.  
[http://digitalcommons.unl.edu/bioscifacpub/539](http://digitalcommons.unl.edu/bioscifacpub/539)

This Article is brought to you for free and open access by the Papers in the Biological Sciences at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Publications in the Biological Sciences by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.
Species are thought to arise primarily as populations diverge in spatial isolation and accumulate differences that preclude their reproduction upon secondary contact (Mayr 1963; Coyne & Orr 2004). This has been argued to be particularly true for birds (Price 2008). There is also long-standing interest in the extent to which population divergence can arise or increase in the context of incomplete spatial isolation, coupled with selection against immigrants or hybrids.
to offset homogenizing genetic exchange (‘divergence with gene flow’: Nosil et al. 2008; 2009). More generally, spatial isolation is a continuum, and divergence can arise as a consequence of spatial isolation and genetic drift (isolation by distance, IBD; Wright 1943; Slatkin 1993); spatially varying natural and sexual selection (isolation by adaptation, IBA, or a subset of IBA: isolation by environment, IBE; Rundle & Nosil 2005; Crispo et al. 2006; Nosil 2012, Schluter 2001, 2009; Shafer & Wolf 2013); or both (Lande 1980; Shafer & Wolf 2013; Wang & Bradburd 2014). While the contributions of spatial isolation and divergent selection to population differentiation have been research areas in evolutionary biology since the beginning of empirical population genetics (Fisher 1930), only more recently have we begun to characterize genomic divergence by simultaneous tests of IBA and IBD (e.g. Wang & Summers 2010; Bradburd et al. 2013; Wang et al. 2013; Wang & Bradburd 2014).

In comparison with accumulating evidence for the role of ecological adaptation (one form of IBA) and geographic isolation (IBD) in population divergence and speciation (Shafer & Wolf 2013; Wang & Bradburd 2014), there is little empirical data on the contribution of sexual selection to population genetic divergence (e.g. Parchman et al. 2013; Baldassarre et al. 2014; Morgans et al. 2014). Adaptation to the local environment may come in the form of sexual selection if competition over mates differs among closely related populations, or signal traits reflect variation in the environmental context in which these traits are developed or advertised (Ingleby et al. 2010; Maan & Seehausen 2011; Baldassarre et al. 2013; Safran et al. 2013; Matute 2014; Seehausen et al. 2014). The lack of such data represents a major gap in our understanding of speciation because divergence in sexual traits is a common form of phenotypic differentiation among populations that may strongly relate to reproductive isolation (e.g. West-Eberhard 1983; Panhuis et al. 2001; Irwin et al. 2001; Ritchie 2007; Curie Network 2012; Safran et al. 2012; Langerhans & Riesch 2013; Seddon et al. 2013). Moreover, previous work has shown that sexual signals can evolve more rapidly than traits related to ecological adaptation (e.g. Kingsolver et al. 2001; Svensson et al. 2006; Kingsolver & Pfennig 2007; Siepielski et al. 2011; Seddon et al. 2013). Collectively, these observations suggest that sexual signals could be an important feature of biological diversification, especially in the early, formative stages of speciation (Kraaijeveld et al. 2010). Here, a testable prediction for a role of sexual selection in divergence is that divergence in aspects of phenotype related to mate selection is positively associated with genetic divergence, either due to increased local adaptation, reduced immigration, or both. This is similar to evidence for ecological speciation, where divergence in traits related to ecological adaptation predicts population genetic differentiation (Shafer & Wolf 2013).

In its most general form, IBA predicts associations between genomic divergence and differentiation in traits related to both ecology and mating success. Note that ecological adaptations and mating adaptations are not mutually exclusive, and indeed, may even be related, as sexual signals necessarily evolve in an ecological context (Ingleby et al. 2010; Maan & Seehausen 2011; Baldassarre et al. 2013; Safran et al. 2013; Seehausen et al. 2014). Thus, important questions remain about the relative contributions and likely interactions of sexual selection, natural selection and geographic distance to genomic differentiation among closely related populations (e.g. van Doorn et al. 2009; Maan & Seehausen 2011; Wagner et al. 2012; Safran et al. 2013). For example, sexual selection has been shown to interact with ecological context in the early stages of population divergence and the formation of premating barriers to reproduction (Scordato et al. 2014), yet we know very little about the relative importance of each selective process in the accumulation of biologically relevant genomic differences among populations.

Closely related populations with variation in geographic distance are advantageous for identifying spatial and trait-based predictors of genomic divergence. To address the extent to which sexual selection and natural selection either interact or singly influence population divergence, we analysed patterns of genomic divergence as a function of variation in aspects of phenotype known to be involved in mate selection and migration, as well as relevant measures of environmental variation (using elevation data and long-term climate databases), in a widespread, phenotypically divergent, yet young group of subspecies: barn swallows (Hirundo rustica). Barn swallows include six subspecies worldwide (Figure 1), with populations varying in (i) geographic distance from one another, (ii) trait combinations with known importance in sexual signalling, (iii) breeding and nonbreeding environments and (iv) migratory behavior (reviewed in Scordato & Safran 2014). A recent phylogenetic reconstruction for this group of closely related subspecies indicates that they are monophyletic with respect to other members of the genus Hirundo (Dor et al. 2010). Additionally, recent mtDNA, phylogeographic and microsatellite analyses suggest this group formed rapidly (between 100 000 and 27 000 years ago; Zink et al. 2006) and is not strongly genetically differentiated, despite marked sexual signal and behavioral differentiation among populations (Dor et al. 2010, 2012). Phenotypic and genomic divergence among barn swallow populations are likely due to a combination of selection and drift in the context of variable degrees of geographic isolation. For example, populations in North America and Eurasia have likely diverged without gene flow, whereas evidence suggests that populations in the Middle East have experienced recent historical or ongoing gene flow with populations in Europe (Dor et al. 2010, 2012).
Previous research in barn swallows has demonstrated phenotypic divergence in traits important to both sexual selection (morphological and behavioral sexual signals; e.g. Safran & McGraw 2004; Safran et al. 2005; Vortman et al. 2011, 2013; Scordato & Safran 2014) and natural selection (traits related to flight, foraging and migration; reviewed in Turner 2006). One of six well-characterized subspecies within the larger H. rustica complex, the European barn swallow (H. r. rustica), has been the subject of intense research activity over the last twenty years, with an emphasis on sexual selection (Turner 2006). This work clearly demonstrates that females mate preferentially with long-tailed males, and their offspring experience advantages over those of shorter-tailed males (e.g. Møller 1994). By contrast, in two North American populations of H. r. erythrogaster, males with darker plumage color have greater social and extrapair mating success (Safran et al. 2005; Safran et al., in Press).

In the east Mediterranean distributed H. r. transitiva, males with a combination of darker ventral plumage and longer streamer lengths are favoured by females through social and extra-pair mating decisions (Vortman et al. 2011, 2013). Both experimental and observational data suggest that tail streamer length and ventral color are under varying degrees of sexual selection in different populations of barn swallows, and previous work has defined characteristics of wing shape, such as wing length, as being associated with migratory behavior (von Rönn et al. 2016). We therefore use these traits to examine the roles of sexual and natural selection in contributing to genome-wide divergence among populations of barn swallows. Further, previous work has shown that variation in ventral color (Saino et al. 2013; Hubbard et al. 2015) and variation in tail streamer length have heritable components (Møller 1994), indicating that they are subject to evolutionary change and thus relevant aspects of phenotype to investigate in terms of their influence on genome-wide divergence.

Here, we genotyped thousands of SNPs (using genotyping- by-sequencing, GBS) in 354 barn swallows from eight
populations, representing four of the six subspecies, distributed across the Northern Hemisphere. Additionally, we report a draft whole-genome sequence for a male *Hirundo rustica erythrogaster*, which was constructed to ensure that GBS sequences cleanly assembled to a reference and to reduce problems with duplicates/paralogs. We use these population genomic data to examine the extent to which population genomic divergence (based on average genome-wide FST) is associated with divergence in (i) sexual signals, (ii) wing length, (iii) features of climate and/or (iv) geographic distance. The first three factors all relate to IBA, while the fourth relates to IBD. A specific objective of this study is to partition variance in pairwise genomic divergence as a function of geographic distance (IBD); population differences in ecology (IBA, influenced by natural selection); wing morphology related to flight and migratory behavior (IBA, influenced by natural selection); and differences in morphology related to sexual communication (IBA, influenced by sexual selection).

**Methods**

**Field data collection**

We sampled individuals from eight locations representing four subspecies across the barn swallow breeding range (Table 1). The following samples or measures were taken from male barn swallows known to be breeding at each site: (i) a sample of ventral feathers for objective color quantification, (ii) length of wing and tail streamers (outer rectrices; see Table 2 for a complete description) and (iii) blood samples as a DNA source from each bird (approx 50 ll, stored in 2% SDS lysis buffer). See Figure 1 and Table 1 for sampling location and final sample size information.

**Genomics methods**

**Genome assembly.** A draft genome assembled with moderate to high coverage enables the identification of single-copy portions of the genome, reduces the challenges associated with distinguishing close paralogs from alleles at the same locus, and, depending on quality and scaffold sizes, can identify the genomic location of GBS loci. The DNA for our genome assembly came from a male barn swallow with a well-known reproductive history in our Boulder, Colorado study site (ID 2540-44680). This male was first captured in 2008 and had five successive breeding seasons at the same location (from 2008 through 2012). To generate a draft reference genome, we obtained sequences from four lanes of Illumina HiSeq platform, two lanes of 101-base pair reads from two paired-end libraries and two lanes of 101-base pair reads from a mate-pair library from Macrogen (www.macrogenusa.com). One paired-end library had an insert size of 176 bp, while a second paired-end library had an insert size of 454 bp. The mate-pair average insert size was 1458 bp. We obtained a total of 27 Gbp from paired-end library one, and 17.4 Gbp from library two. From the two lanes of mate-pair sequence, we obtained 43.4 Gbp. After cleaning to remove low-quality reads (lower than a base quality of 20 on either end of the read) and common contaminants, 129.4 million pairs of reads remained in library one, 80.0 million pairs in library

**Table 1.** Sampling locations across four subspecies of *Hirundo rustica*. Acronyms are given for each location that match with locations on Figures 1 and 2.

<table>
<thead>
<tr>
<th>Sampling location (abbreviation for location on map)</th>
<th>Subspecies</th>
<th>Lat/long</th>
<th>Final sample size</th>
<th>Sampling dates: year (sample size)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boulder, Colorado, USA (CO)</td>
<td><em>Hirundo rustica erythrogaster</em></td>
<td>40.17, –105.10</td>
<td>144</td>
<td>2008 (50) 2009 (72) 2010 (22)</td>
</tr>
<tr>
<td>Czech Republic (CR)</td>
<td><em>Hirundo rustica rustica</em></td>
<td>49.06, 14.76</td>
<td>24</td>
<td>2010</td>
</tr>
<tr>
<td>Israel (IL)</td>
<td><em>Hirundo rustica transitiva</em></td>
<td>32.93, 35.54</td>
<td>45</td>
<td>2008 (3) 2009 (37) 2010 (5)</td>
</tr>
<tr>
<td>Ithaca, New York, USA (IA)</td>
<td><em>Hirundo rustica erythrogaster</em></td>
<td>42.44, –77.50</td>
<td>27</td>
<td>2002</td>
</tr>
<tr>
<td>Romania (RM)</td>
<td><em>Hirundo rustica rustica</em></td>
<td>46.75, 23.83</td>
<td>16</td>
<td>2010</td>
</tr>
<tr>
<td>Taiwan (TW)</td>
<td><em>Hirundo rustica gutturalis</em></td>
<td>25.09, 121.56</td>
<td>18</td>
<td>2010</td>
</tr>
<tr>
<td>Turkey (TR)</td>
<td><em>Hirundo rustica rustica</em></td>
<td>36.85, 31.16</td>
<td>50</td>
<td>2010 (50)</td>
</tr>
<tr>
<td>United Kingdom (UK)</td>
<td><em>Hirundo rustica rustica</em></td>
<td>50.50, –4.65</td>
<td>26</td>
<td>2009</td>
</tr>
<tr>
<td>Totals</td>
<td>8 populations</td>
<td>4 subspecies</td>
<td>350</td>
<td></td>
</tr>
</tbody>
</table>
two and 95.8 million pairs in the matepair library remained, for a total of 61.7 Gb of sequence. Based on an estimated genome size of 1.3 Gb (Andrews et al. 2009), this is an average of 47× coverage.

Reads were assembled using SOAPdenovo 2.04, using a k-mer length of 47, an edge coverage cut-off of 3, a k-mer frequency cut-off of 3 and an arcweight filter of 3. Repeats were resolved with reads. Otherwise, parameters were as defaults. After removing short scaffolds (below 1000 bp), we had a total of 100 153 scaffolds and a total assembly length of 1.1 Gb, 85% of the estimated 1.3 Gb genome size (Andrews et al. 2009). The average scaffold length was 11 010 bp, the longest scaffold was 732 517 bp, the N50 was 38 844 bp and N90 was 3718. Of this 1.1 Gb of assembled sequence, a total of 1.06 Gb could be conservatively mapped to the Ficedula albicollis genome (Ellegren et al. 2012), using blastn with a minimum e-value cut-off of 10−80 and a minimum sequence similarity of 80%.

The alignment of our Hirundo rustica genome assembly to the most closely related, well-annotated genome assembly for Collared Flycatcher Ficedula albicollis (Ellegren 2012) shows a high degree of sequence conservation. Over 91% of our single-copy, assembled sequence could be uniquely placed onto the genome. Thus, although we do not have a genetic or physical map placing our sequences onto chromosomes, we can provisionally place virtually all of our assembled genome using sequence similarity to existing avian genome sequences. Given the high degree of synteny found across passerines (Backström et al. 2008; Ellegren 2013; Kawakami et al. 2014), and among much more distant avian lineages (Derjusheva et al. 2004; Zhang et al. 2014), taking advantage of existing related genomes enabled us to infer which markers were likely to be linked on chromosomes. We also identified markers likely to be on autosomes and on the sex-determining Z chromosome. As the barn swallow populations we focus on in this paper exhibit only slight genome-wide differentiation (see Results), this draft genome is ideal for use as a reference for aligning the GBS SNPs described below.

Population genomics. We generated DNA sequence data from 354 individual barn swallows sampled across eight populations (see Table 1). We constructed reduced genomic complexity libraries for each individual using a restriction fragment-based procedure (Gompert et al. 2012; Parchman et al. 2012). We first digested genomic DNA with two restriction endonucleases (EcoRI and MseI) and ligated double-stranded adaptor oligonucleotides to the digested fragments. These oligonucleotides consisted of the priming sites for Illumina sequencing, followed by eight, nine or ten base pair barcode sequences that allow for the unique identification of sequences from each individual. This method has been successfully used to generate population genetic data for a large number of projects (Gompert et al. 2012, 2014; Nosil et al. 2012; Parchman et al. 2012, 2013; Mandeville et al. 2015); a full version of the protocol is available at dryad (doi: http://dx.doi.org/10.5061/dryad.17hr9.). We used 354 unique barcoded adaptors, which allowed us to sequence all individuals in one Illumina HiSeq sequencing lane. After the restriction and ligation reactions, we pooled all samples and used standard Illumina PCR primers to amplify the barcode-adapted fragments. We separated the amplification products on a 2% agarose gel and excised fragments between approximately 350 and 450 bp in length. We purified these fragments using Qiagen’s Qiaquick Gel Extraction Kit (Qiagen Inc.). Concentration and quality of the pooled library was evaluated on an Agilent BioAnalyzer qPCR. Sequencing was performed by the National Center for Genome Research (Santa Fe, NM, USA) with 100-base pair single-end reads, yielding 98 401 301 reads following standard contaminant filtering.

We used a custom Perl script (dryad doi: http://dx.doi.org/10.5061/dryad.17hr9.) to remove bases associated with the barcode and EcoRI cut site for all 98 401 301 sequences and to replace the sequence IDs with individual IDs for each DNA sample. This script also corrected barcodes with potential single or double base mismatches due to sequencing or oligonucleotide synthesis errors. Sequences were then split by individual for further analysis. Four of the individuals had fewer than 2000 reads and were disregarded, which left 350 individuals for further analysis. We used BWA v0.7.12 (Burrows–Wheeler Aligner; Li & Durbin 2009) to assemble data for each individual against the barn swallow genome assembly, with an edit distance of 4 and the remaining parameters set as default. This edit distance ensures that even with 1% sequencing errors typical of Illumina sequencing, distant alleles from divergent subspecies

### Table 2. Morphological and environmental traits measured in 350 individuals, across 8 sampling locations.

<table>
<thead>
<tr>
<th>Types of Traits</th>
<th>Traits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traits related to reproductive performance among subspecies of barn swallows</td>
<td>Tail streamer length, ventral color (throat, breast and vent, % brightness)</td>
</tr>
<tr>
<td>Trait related to flight behavior and migratory distance</td>
<td>Wing length (mm)</td>
</tr>
<tr>
<td>Features of the environment that affect food availability</td>
<td>Mean, variability and seasonal minima in breeding temperatures; elevation</td>
</tr>
</tbody>
</table>

---
will map to the genome, while nevertheless preventing non-specific alignments. Across all individuals, an average of 61.66% of reads assembled, with 97.5% of individuals having 57.9% or greater percentage of reads assembling. We used SAMTOOLS v 1.2 and BCFTOOLS v 1.2 (Li et al. 2009) to identify variant sites in the assembled sequences and obtain genotype likelihoods for variable sites. To identify variants, we required data for 60% or more individuals. This resulted in 67,773 single nucleotide variants. A complete list of parameters used for assembly and variant calling are available from the authors by request.

We used point estimates of allele frequencies in all 350 individuals to separate this set into 22,328 common variants (minimum minor allele frequency of 5% or greater) that were used for further analysis and 45,445 rare variants that were disregarded. We focused on comparatively common variants because they are more likely to occur in more than one location and contain information about population histories (Gompert et al. 2012). To avoid analysing SNPs with highly correlated allele frequencies, we randomly sampled a single SNP from each of the 100-bp regions in the assemblies (GBS loci are 100 bases in length). For the resulting final set of 9,493 single nucleotide polymorphisms (SNPs), the average coverage depth per individual per site was 1.5× (0.09–3.57, 2.5% and 97.5% quantiles). This kind of low coverage genomic data is appropriate for population-level inferences when analysed with models that incorporate uncertainty arising from variability in sequencing coverage across individuals and loci (e.g. Nielsen et al. 2011; Gompert et al. 2012; Buerkle and Gompert 2013).

Population genetic analyses. We used a Bayesian model to estimate population allele frequencies for each of 9493 variable nucleotides based on the point estimates of genotypes from BCFTOOLS (described in Gompert et al. 2012). Genotypes and the population allele frequencies were treated as unknown model parameters, and genotype probabilities and allele frequencies were simultaneously estimated for each sampling locality (separately from estimates for other localities). Importantly, this model incorporates uncertainty from stochastic variation in sequencing coverage depth across individuals and loci into the estimation process and provides a sound approach for estimating population-level parameters for low coverage sequencing approaches. Population allele frequencies serve as prior information, and genotype probabilities are inferred for each locus in each individual. We obtained posterior probabilities for parameters using Markov chain Monte Carlo (MCMC). Each analysis consisted of a single chain iterated for 2000 steps following a 1000-step burn-in, with samples retained every other step, yielding 1000 samples from the posterior distributions. For these data and this simple model, mixing and convergence were clearly evident in plots of MCMC histories.

We obtained the mean genotype (scale of 0–2) from the posterior distributions for each SNP in each individual. Furthermore, we calculated the mean genotype across all individuals to center the genotypes at each locus before calculating the genetic covariance between the genotype vectors for all pairs of individuals. We summarized the genetic covariance matrix using principal component analysis (PCA). Because the PCA will reflect all genetic covariances and will be affected by uneven sampling (e.g. 144 samples from Colorado and 16 from Romania), we performed a second PCA in which we randomly down-sampled all localities to the smallest sample size from any locality (N = 16). We also obtained mean allele frequencies from the posterior distributions and transformed these to estimates of mean F_{ST} (Hudson et al. 1992) and Nei’s D (Nei et al. 1983; Takezaki & Nei 1996) between all pairs of populations. F_{ST} and Nei’s D were highly correlated (r = 0.996), so we only present results for F_{ST}. We performed the PCA and calculated F_{ST} and Nei’s D in R (using pcomp and custom functions, R Core Team, 2015).

Phenotypic and environmental variables

Our goal was to compare divergence among traits known to be related to sexual signalling (plumage color and streamer lengths) and migratory behavior (i.e. wing length). We also compared aspects of environmental variability known to affect the aerial insect populations that barn swallows prey upon. These variables include elevation and several measures of temperature variation (e.g. Thomsen et al. 2016).

Quantifying feather color. We collected 5–10 feathers from the throat and breast (upper ventral region) and vent (lower ventral region, below the attachment of tail streamers) and stored them in small envelopes in a dark, dry environment prior to measurement (following Safran & McGraw 2004). We assessed the color of these samples by measuring plumage brightness using an Ocean Optics USB4000 spectrometer (Dunedin, FL). Reflectance data were generated relative to a white standard (Ocean Optics WS-1) and a dark standard (all light excluded), and spectra were recorded with the SPEC- TRASUITE software package (version 2.0.125, Ocean Optics Inc.). We find no evidence of UV reflectance in the ventral plumage color of barn swallows (Safran & McGraw 2004) and thus used three traditional axes of color for objective measurement of color variation. We used average brightness, which was calculated from three separate measurements of the collected throat and breast feathers, as a representative
metric of overall ventral plumage color. Average brightness is a good color metric, as all three traditional axes of color (hue, chroma and brightness) were previously found to be highly correlated across the ventral region of individual barn swallows (McGraw et al. 2005, Safran & McGraw 2004, J.K. Hubbard, unpublished data), and brightness is the most variable dimension of color in this region (J.K. Hubbard, unpublished data). Lower brightness scores (% reflectance) indicate plumage color that appears darker, redder and more saturated when compared to feathers with higher brightness scores.

Estimates of phenotypic divergence. To determine pairwise distance in phenotypes among environmental and morphological traits, we used an unbiased effect size statistic ($\Delta P$; Safran et al. 2012) to calculate trait distance for each trait in pairwise comparisons among the 8 sampled populations. $\Delta P$ is calculated based upon a joint cumulative distribution function (CDF) from all populations in the data set. Distances were calculated for each pairwise comparison using the population median percentile in the overall CDF. $\Delta P$ was developed specifically to analyse phenotypic distance among closely related populations, as it easily accommodates simultaneous comparisons of any number of traits across any number of populations, and is relatively insensitive to unequal variances and sample sizes among populations (Safran et al. 2012). For all analyses, we use the absolute value of pairwise distances.

Climate and elevation data. Temperature and elevation data were obtained using the CRUTEM database maintained by the Climatic Research Unit and available at http://www.cru.uea.ac.uk/cru/data/crutem/ge/. Using this database, we downloaded the last 50 years of temperature data from the three weather stations closest to each of our sampling sites (see Supporting information). From these data, we derived the mean, minimum and coefficient of variation in temperature during the breeding season. Breeding seasons vary among our populations, and, accordingly, we used the following months for climate data collection for each site: Colorado, USA: April–September; New York, USA: May–August; UK: April–August; Israel, January–April; and Czech Republic, Romania, Turkey and Taiwan, April–July. These measurements have been routinely employed in studies of other avian taxa with widespread geographic ranges (Rubenstein & Lovette 2007, 2007; Botero et al. 2009, 2014).

Estimates of geographic distance. Geographic distances between study sites were calculated in the R package “geosphere” using the Haversine great circle distance between points. This is the shortest “as the crow flies” distance, assuming a spherical earth.

Associations of geographic distance, environmental context and phenotype with population genomic divergence

To assess the degree to which genome-wide divergence is associated with geographic and phenotypic distances, we analysed correlations between pairwise trait distance and both pairwise geographic and genetic distances ($F_{ST}$), and assessed their significance (999 permutations) using Mantel tests using the R package “VEGAN.” Next, to quantify the relative strength of association among geographic, phenotypic and genomic variables, accounting for correlations between phenotypic and geographic distances, we used two complementary statistical approaches. First, we applied multiple matrix regression (Wang 2013). This performs multiple regression on distance matrices and uses permutation tests ($n = 10000$) to obtain $P$ and $R^2$ values using the MRM function in the R package “ecodist” (Wang 2013). In the second approach, we employed constrained redundancy analysis and variance partitioning to analyse the relative contributions of traits related to natural and sexual selection and environmental context in explaining pairwise genetic divergence (Legendre and Fortin 2010) using the R package “VEGAN” (Oksanen et al. 2015 v. 2.2). Redundancy analysis is a type of constrained ordination that quantifies how much variation in a set of variables is explained by a second set of variables, with the option of conditioning on a third set. This analysis is ideal for correlated matrices, as is often the case with the matrices related to phenotype, genotype, environmental variables and geographic distance among closely related populations (Shafer & Wolf 2013; Wang & Bradburd 2014). Using these methods, variance can then be partitioned between the constrained, conditioned and joint variable sets. Our data were structured in two different formats, depending on the analysis. For Mantel tests, we employed matrices of pairwise differences for each population. For MRM and the variance partitioning analyses, these matrices were transformed into vectors of pairwise differences for each variable and each pair of populations. When the response variable is a single vector (as here, using pairwise differences in mean $F_{ST}$ between populations), variance partitioning is done by partial regression.

Results

Population genomics

Principal component analysis of genetic covariances between an even sample of individuals from each population ($n = 16$) revealed clear genomic differences between some localities, but greater genomic similarity among the nearby sampling
locations in Czech Republic, Romania, Turkey and Israel (Figure 2). The first two principal components explained 52% of the genomic variation among individuals in each locality (Figure 2). The separation of populations on PC1 is consistent with the phylogenetic hypothesis (Dor et al. 2010) that separates eastern and western barn swallows, with samples from Israel and Europe recoverable as a distinct lineage from samples in North America (Colorado and New York) and southeastern Asia (Taiwan). Interestingly, while Israel is considered a distinct subspecies based on phenotype, it cannot be differentiated from European populations based on genomic covariance. PC2 explains a relatively small amount (10.3%) of genetic variation compared to PC1 (41.9%), and further separates subspecies within the eastern and western clades. The PCA based on all 350 individuals gave considerable weight in the first axis to distinction between the large sample from North America (N = 144 from Colorado and 27 from New York) and samples from elsewhere, with PC1 accounting for 67.9% of the genetic variation. Pairwise $F_{ST}$ ranged from 0.024 to 0.073 (Table 3).

**Phenotypic and environmental variables**

Phenotypic and environmental divergence. Populations differed in the extent of both phenotypic variation and environmental context (Tables 4 and 5). Mean percentiles within a cumulative frequency distribution varied considerably among populations; the distribution of variables measured from each population differed in terms of its placement on the overall cumulative frequency distribution (Table 4). For example, populations in the UK have the longest wing length, whereas populations in Taiwan have the shortest. Populations in North America (Colorado and Ithaca) have the lightest throat color yet the darkest breast color; thus, there is variation in the extent and direction in which color patches differ among subspecies. The spread of percentile values is indicative of the degree of pairwise divergence for each trait or climate variable. For example, of all phenotypic traits, populations exhibited the most extreme differences in wing length, with a range of percentiles from 9.02 to 91.16. Note that the percentile values for the phenotypic measures are based on multiple

![Figure 2](image-url). Statistical summary of population genetic structure based on principal component axes one (PC1), and two (PC2) derived from genotype probabilities. Acronyms for sampling sites are as follows: CO, Colorado, USA; CR, Czech Republic; IA, New York, USA; IL, Israel; RM, Romania; TR, Turkey; TW, Taiwan; UK, United Kingdom.
individuals within populations (Table 1), whereas the percentile values for the environmental measures are based on one measure within each population. Analyses of variance reveal that all of the phenotypic traits we measured in this study showed statistically significant differences among populations (adjusted $R^2$-squared values range from 0.25 to 0.52; Table 5).

**Associations of geographic distance, environmental context and phenotype with population genomic divergence**

**Geographic distance.** Geographic distance between pairs of populations predicted genome-wide divergence, consistent with a model of isolation by distance ($r = 0.628$, Mantel $P$ value < 0.008, Figure 3, Table 3).

**Phenotypic divergence and environmental context.** Distance matrices (based on $\Delta P$) for various features of phenotype, including measures of ventral color, tail and wing length, were positively correlated with genomewide divergence among pairs of eight geographically isolated populations (Mantel tests; Figure 4 and Table 6). None of the pairwise differences in environmental variables (Mantel tests; elevation and various measures of breeding season temperature) were significantly associated with genome-wide divergence among our study populations (Figure 5 and Table 6).

**Geographic distance, environment, phenotype.** Several features of phenotypic divergence also covaried with geographic distance (Mantel tests; Figure 6 and Table 6), whereas environmental and climate features did not (Mantel tests; Figure 7 and Table 6). Thus, genome-wide divergence among closely related populations was associated with both geographic and phenotypic trait distances, and the two are sometimes correlated with each other. It is therefore necessary to adequately control for correlations between phenotypic and geographic distances in order to infer the relative significance of IBA and IBD. We did this using two complementary approaches.

First, to investigate the associations of specific traits with pairwise, genome-wide $F_{ST}$, while accounting for correlations among variables in our model (e.g. phenotypes correlated with one another and with geographic distance), we used multiple matrix regression (Wang 2013). We started with a maximal model that included the pairwise distance matrix of $F_{ST}$ values as the response variable and distance matrices (based on $\Delta P$ values) for all nine phenotypic and ecological

---

**Table 3.** Average pairwise $F_{ST}$ (upper diagonal in italicized text) and geographic distance (km; lower diagonal).

<table>
<thead>
<tr>
<th>Sampling location</th>
<th>Colorado</th>
<th>Czech Republic</th>
<th>New York</th>
<th>Israel</th>
<th>Romania</th>
<th>Turkey</th>
<th>Taiwan</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colorado</td>
<td>NA</td>
<td>0.049</td>
<td>0.024</td>
<td>0.054</td>
<td>0.053</td>
<td>0.046</td>
<td>0.046</td>
<td>0.059</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>853.948</td>
<td>NA</td>
<td>0.059</td>
<td>0.038</td>
<td>0.041</td>
<td>0.030</td>
<td>0.062</td>
<td>0.048</td>
</tr>
<tr>
<td>New York</td>
<td>2397.257</td>
<td>6726.714</td>
<td>NA</td>
<td>0.062</td>
<td>0.063</td>
<td>0.055</td>
<td>0.055</td>
<td>0.067</td>
</tr>
<tr>
<td>Israel</td>
<td>10959.901</td>
<td>2454.926</td>
<td>9169.197</td>
<td>NA</td>
<td>0.044</td>
<td>0.031</td>
<td>0.065</td>
<td>0.051</td>
</tr>
<tr>
<td>Romania</td>
<td>9131.466</td>
<td>700.843</td>
<td>7399.893</td>
<td>1828.448</td>
<td>NA</td>
<td>0.037</td>
<td>0.067</td>
<td>0.053</td>
</tr>
<tr>
<td>Turkey</td>
<td>10381.422</td>
<td>1864.741</td>
<td>8578.322</td>
<td>591.36</td>
<td>1255.353</td>
<td>NA</td>
<td>0.059</td>
<td>0.045</td>
</tr>
<tr>
<td>Taiwan</td>
<td>11321.578</td>
<td>9083.006</td>
<td>12305.315</td>
<td>8199.932</td>
<td>8582.409</td>
<td>8425.255</td>
<td>NA</td>
<td>0.073</td>
</tr>
<tr>
<td>UK</td>
<td>7337.099</td>
<td>1443.954</td>
<td>5361.685</td>
<td>3818.056</td>
<td>2144.723</td>
<td>3230.591</td>
<td>10145.824</td>
<td>NA</td>
</tr>
</tbody>
</table>

**Table 4.** Percentiles for male morphological traits across all populations used for calculating DP differences. Lower wing and tail streamer percentiles indicate shorter wings and streamers, respectively. Higher color percentiles indicate lighter color. Greater elevation percentiles are indicative of higher elevations at breeding locations. Greater percentiles for variation, minima and average temperatures are indicative of more variable, hotter climates measured over the period of data collection (1965–2015).

<table>
<thead>
<tr>
<th>Trait</th>
<th>Colorado</th>
<th>Czech Republic</th>
<th>Israel</th>
<th>New York</th>
<th>Romania</th>
<th>Taiwan</th>
<th>Turkey</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wing length</td>
<td>24.08</td>
<td>70.73</td>
<td>61.00</td>
<td>30.22</td>
<td>61.87</td>
<td>9.02</td>
<td>61.00</td>
<td>91.16</td>
</tr>
<tr>
<td>Streamer length</td>
<td>21.82</td>
<td>72.23</td>
<td>60.22</td>
<td>15.05</td>
<td>79.79</td>
<td>28.66</td>
<td>67.16</td>
<td>60.22</td>
</tr>
<tr>
<td>Throat color</td>
<td>77.74</td>
<td>59.42</td>
<td>39.96</td>
<td>80.77</td>
<td>18.35</td>
<td>28.26</td>
<td>47.12</td>
<td>44.62</td>
</tr>
<tr>
<td>Breast color</td>
<td>21.03</td>
<td>58.63</td>
<td>27.61</td>
<td>13.09</td>
<td>58.72</td>
<td>80.77</td>
<td>55.41</td>
<td>70.39</td>
</tr>
<tr>
<td>Vent color</td>
<td>18.10</td>
<td>27.01</td>
<td>16.02</td>
<td>20.32</td>
<td>21.45</td>
<td>32.99</td>
<td>25.32</td>
<td>30.14</td>
</tr>
<tr>
<td>Elevation</td>
<td>100</td>
<td>62.5</td>
<td>75</td>
<td>37.5</td>
<td>50</td>
<td>12.5</td>
<td>87.5</td>
<td>25</td>
</tr>
<tr>
<td>Var temp</td>
<td>37.5</td>
<td>75</td>
<td>87.5</td>
<td>62.5</td>
<td>50</td>
<td>25</td>
<td>100</td>
<td>12.5</td>
</tr>
<tr>
<td>Min temp</td>
<td>62.5</td>
<td>25</td>
<td>12.5</td>
<td>5 87.5</td>
<td>50</td>
<td>100</td>
<td>87.5</td>
<td>37.5</td>
</tr>
<tr>
<td>Mean temp</td>
<td>62.5</td>
<td>25</td>
<td>12.5</td>
<td>75</td>
<td>50</td>
<td>100</td>
<td>87.5</td>
<td>37.5</td>
</tr>
</tbody>
</table>
variables as predictors. We then used backwards stepwise model selection, sequentially deleting the least significant term and rerunning the model, until coefficients for all remaining predictor variables were significantly different from zero. The final multiple matrix regression model included two aspects of phenotype that explained significant variation in pairwise $F_{ST}$ while controlling for spatial autocorrelation: wing length and throat color (full model $F_{2,17} = 24.95, P = 0.002, R$-squared = 0.757; parameter estimate and $t$-test: wing length coefficient = 0.82, $T = 6.12, P = 0.002$, throat color coefficient = 0.22, $T = 1.79, P = 0.03$, geographic distance coefficient = 0.01, $T = 0.09, P = 0.94$).

We used the significant predictor variables from the multiple matrix regression in variance partitioning and redundancy analyses to further examine the relative contributions of wing length, throat color and geographic distance in explaining variation in genomewide divergence among the populations in our data set. The variance partitioning approach enabled us to test hypotheses about the relative significance of IBA and IBD, taking into account correlations among phenotypic, environmental and geographic distance variables.

The best model fit from the variance partitioning analyses (Table 7) included the effects of all three matrices (geographic distance, throat color distance and wing length distance). These models, in which various combinations of matrices are conditioned upon one another, demonstrated that collectively 73% of genomewide divergence is attributable to three variables: wing length, throat coloration and geographic distance between populations (Figure 8). Further analyses enabled us to analyze the association between each matrix (geographic distance, throat color, wing length) separately by conditioning each variable on the others in all possible combinations (Table 7). For example, when the matrix containing pairwise distance in wing length is conditioned on the matrices containing pairwise differences in geographic distance and throat color, the influence of wing length on its

<table>
<thead>
<tr>
<th>Trait</th>
<th>$F$</th>
<th>$P$</th>
<th>Adj $R$-squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wing length</td>
<td>54.11</td>
<td>$&lt;$0.001</td>
<td>0.52</td>
</tr>
<tr>
<td>Streamer length</td>
<td>54.27</td>
<td>$&lt;$0.001</td>
<td>0.52</td>
</tr>
<tr>
<td>Throat color</td>
<td>18.84</td>
<td>$&lt;$0.001</td>
<td>0.25</td>
</tr>
<tr>
<td>Breast color</td>
<td>28.16</td>
<td>$&lt;$0.001</td>
<td>0.37</td>
</tr>
<tr>
<td>Vent color</td>
<td>24.97</td>
<td>$&lt;$0.001</td>
<td>0.34</td>
</tr>
</tbody>
</table>

Figure 3. Average pairwise genetic distance ($F_{ST}$) as a function of geographic distance (km) among pairwise comparisons of eight barn swallow populations, consistent with a pattern of isolation by distance.
own accounted for 42% of pairwise genomic distance among the populations in our sample (Figure 8 and Table 7). When the matrix containing pairwise distance in throat color is conditioned upon the matrices containing pairwise differences in geographic distance and wing length, the influence of throat color on its own explained 5% of pairwise genomic distance among the populations in our sample. Finally, when conditioned upon the wing and color matrices, the effect of geographic distance on its own did not explain additional variation in genomic distance among the populations in our sample (Figure 8 and Table 7). Interestingly, the only place where the geographic distance matrix explained a significant amount of variation in genomic divergence is when both geographic distance and wing length were considered side by side and conditioned upon their correlation with the color matrix (Figure 8 and Table 7).

Discussion

Geographic distance and phenotypic distance are strongly correlated among our study populations. Thus, we applied two complementary methods of variance partitioning – multiple matrix regression, and constrained redundancy analysis – which enabled us to analyse the relative contributions of correlated matrices (geographic distance and phenotypic distance) to genome-wide divergence. Overall, our results demonstrate clear evidence that both IBA and IBD contribute to genome-wide divergence among these closely related populations of barn swallows. When spatial autocorrelation between phenotype and geographic distance are accounted for, our results suggest that divergence in an ecological trait (wing length) and a sexual signalling trait (throat color) play a larger role in population genetic divergence than does geographic distance.

Surprisingly, we found no evidence that elevation and temperature differences were influential to genomewide divergence. Although these features of the environment, which are relevant for obligate aerial insectivores, did vary spatially, they were neither associated with geographic distance nor genome-wide divergence among the populations in our sample. In a further investigation, we also found no evidence that maximal temperatures or precipitation patterns at each location differed as a function of geographic distance or were associated with genome-wide divergence (R.J. Safran, unpublished data). Given the span of our sampling locations, ranging from Israel to North America, these results either suggest that barn swallows occupy fairly similar environments with respect to elevation and temperature or are not particularly sensitive to these ecological variables. The latter

Figure 4. Average pairwise genetic distance ($F_{ST}$) is strongly correlated with phenotypic distance among pairwise comparisons of eight barn swallow populations, consistent with a pattern of IBA. Table 6 shows results of Mantel tests between $F_{ST}$ and these four phenotype measures; all are significantly associated with the exception of throat color.
explanation seems most likely as these populations are very cosmopolitan in distribution during both the breeding and nonbreeding season and during long migratory trips where they likely inhabit a wide range of environments.

**Population genomic structure and phenotypic divergence**

Our data demonstrate genetic similarity of individuals within sampling localities, with greater differences between populations that are separated by large geographic distances, and genome-wide clustering that generally corresponds to named subspecies. Principal component analysis shows that the most genetically differentiated populations along PC1 correspond to a highly supported east–west split in the current phylogenetic hypothesis for barn swallows (Zink et al. 2006; Dor et al. 2010). Populations in Asia (H. r. gutturalis) and North America (H. r. erythrogaster) are more closely related to one another than either are to populations in Europe.

**Table 6.** Results of Mantel tests. Tests of isolation by adaptation and phenotypic-genetic distance correlations using traits related to mate selection (color, tail length) and migratory behavior (wing length) and environmental traits including elevation and various metrics of temperature during the breeding season. To control for multiple testing, we have indicated the confidence intervals around the Mantel test coefficient.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Association with genome-wide $F_{ST}$</th>
<th>Association with geographic distance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mantel $r$  CI</td>
<td>$P$ value</td>
</tr>
<tr>
<td>Wing length</td>
<td>0.85  0.79–0.91</td>
<td>0.002</td>
</tr>
<tr>
<td>Tail length</td>
<td>0.67  0.60–0.81</td>
<td>0.012</td>
</tr>
<tr>
<td>Throat color</td>
<td>0.29  −0.08 to 0.66</td>
<td>0.091</td>
</tr>
<tr>
<td>Breast color</td>
<td>0.34  0.16–0.53</td>
<td>0.023</td>
</tr>
<tr>
<td>Vent color</td>
<td>0.40  0.14–0.59</td>
<td>0.037</td>
</tr>
<tr>
<td>Elevation</td>
<td>−0.04 −0.27 to 0.38</td>
<td>0.593</td>
</tr>
<tr>
<td>Min breeding temp</td>
<td>0.07  −0.18 to 0.40</td>
<td>0.348</td>
</tr>
<tr>
<td>Mean breeding temp</td>
<td>0.20  −0.18 to 0.64</td>
<td>0.145</td>
</tr>
<tr>
<td>Var breeding temp</td>
<td>0.07  −0.11 to 0.29</td>
<td>0.346</td>
</tr>
</tbody>
</table>

$P$-values highlighted in boldface are statistically significant at $P < 0.05$.

**Figure 5.** Average pairwise genetic distance ($F_{ST}$) as a function of environmental distance among pairwise comparisons of eight barn swallow populations. Table 6 shows results of Mantel tests between $F_{ST}$ and environmental distance; that there are no statistically significant associations between genetic distance and these environmental variables.
(H. r. rustica) and the Middle East (H. r. transitiva). PC2 further separates H. r. gutturalis (TW) from two samples from the North American H. r. erythrogaster populations (IA and CO), and the UK population from other mainland populations of H. r. rustica (TR, CR and RM). Samples from the Israeli subspecies (IL) are clustered closely with other mainland European populations of barn swallows, consistent with an unresolved relationship between this subspecies and H. r. rustica in the current phylogeny (Dor et al. 2010). Despite being relatively genetically similar, individuals from Israel and the continental European populations are fairly divergent in phenotype, particularly with respect to ventral color.

Despite shallow genomic divergence, phenotypic differentiation is apparent in all aspects of morphology we analysed in this study, ranging from traits related to body size and flight (wing and tail length) to ventral color. Phenotypic variation despite shallow genomic divergence appears common among many taxa, including cichlid fishes (e.g. Wagner et al. 2012) and particularly in birds (e.g. Parchman et al. 2006; Poelstra et al. 2014; Rodrigues et al. 2014; Mason & Taylor 2015), where it is often the case that a few genes are implicated in morphological variation against a fairly homogenous genomic background (e.g. Poelstra et al. 2014; Kardos et al. 2016). In other words, although a larger number of genes may be involved in generating plumage coloration, studies to date have suggested that a large proportion of segregating trait variation is due to variation in a small number of genetic loci. Collectively, these studies suggest an important role of divergent selection on signalling traits in population differentiation (Wagner et al. 2012; Poelstra et al. 2014), which might play a particularly important role during the earliest stages of speciation (Kraaijeveld et al. 2010). Finer-scale genomic analyses in a broader geographic context will enable us to test whether phenotypic differentiation is a barrier to gene flow when different subspecies of barn swallows are in secondary contact.

**Isolation by adaptation, controlling for isolation by distance**

Because trait divergence is correlated with geographic distance among populations, we applied two complementary statistical methods to tease apart the relative significance of geographic and phenotypic distances in explaining genomewide divergence. Both sets of analyses reveal that wing length and, to a lesser extent, throat color, are most strongly associated with genomewide divergence, when geographic distance
among populations is accounted for statistically (Figure 8). The variance partitioning model enables us to directly quantify the influence of correlations among geographic distance, wing length and throat color and to partition the contribution of each of these variables towards differences in genome-wide divergence (Figure 8 and Table 6). Thus, our study reveals an important role of phenotypic divergence and supports a model of IBA for explaining genomic differentiation of geographically isolated populations, bearing in mind that phenotypic divergence is also strongly correlated with geographic distance, and a model of IBD, among populations (Figure 8).

IBD is one of the most common patterns in population genetic data (e.g. Jenkins et al. 2010). Evidence is accumulating to suggest that IBA (which includes models of isolation by environment, or IBE) is also well supported in a variety of empirical systems among closely related populations (e.g. Lee & Mitchell-Olds 2011; Edelaar et al. 2012; Lasky et al. 2012; Shafer & Wolf 2013; Morgans et al. 2014; Wang & Bradburd 2014). A recent meta-analysis further revealed that the effect of geographic distance or spatial autocorrelation (between phenotypes, environmental variables and geographic distance) is critical to control for in statistical tests of IBA but has rarely been done in previous studies (Shafer & Wolf 2013). In our own study, separate analyses of IBA and IBD each revealed significant effects on genome-wide divergence. Analyses that explicitly considered the correlations of phenotypic and geographic distances were critical to separating out the relative significance of each factor. Variance partitioning (Legendre and Fortin 2010) and multiple matrix regression (Wang 2013) are excellent methods for dealing with spatial autocorrelation, analysing the relative significance of IBA and IBD, and are particularly important alternatives to partial Mantel tests (Bradburd et al. 2013; Wang 2013; Wang et al. 2013), which are subject to false-positive results (Diniz-Filho et al. 2013).

Barn swallows and divergent selection

Our results support an important role of IBA in genome-wide divergence. In particular, differences in wing length, the most divergent phenotypic trait among populations, explained a significant amount of variation in among-population genomic divergence; throat color is also associated with genomic divergence, but to a lesser extent. From these results and our understanding of the function of these traits, we infer an influence of both natural selection and sexual selection in genomic divergence among these eight populations of barn swallows. These results are intuitive for several reasons.
First, wing length and shape are traits associated with migratory behavior (Marchetti et al. 1995; Lockwood et al. 1998). Barn swallows vary in migratory distance, and there is evidence of migratory divides throughout their range (e.g. Irwin & Irwin 2005; von Rönn et al. 2016). All four representatives of the *H. rustica* subspecies (UK, Czech Republic, Romania and Turkey) have the longest wing lengths among populations sampled, whereas individuals from distantly related Taiwan (*H. gutturalis*) have the shortest. Interestingly, the sole nonmigratory population in our study, Israel, has...
intermediate wing lengths that overlap in variation with individuals from Turkey and Romania and is genomically indistinguishable from *H. r. rustica*, based on pairwise $F_{ST}$. In other migratory bird populations, variation in migratory behavior can influence both the evolution of wing shape and genomic differentiation (e.g. Ruegg 2007, 2008; Rolshausen *et al.* 2009; Delmore *et al.* 2012, 2015; Delmore & Irwin 2014; von Rönn *et al.* 2016). Thus, variation in migratory routes and behaviors (e.g. timing of arrival to breeding grounds, which may affect mate selection) is proposed as a potentially influential contributor to population divergence (Irwin & Irwin 2005; Rolshausen *et al.* 2009). Further study is required to better understand the role of migratory behavior in divergence among barn swallow populations, as well as the extent to which these seasonal movement patterns influence trait evolution.

Our analyses of IBA also revealed that population-level differences in color contribute to genomic differentiation. Plumage color in various ventral regions (throat, breast, vent) varies both within and among closely related populations of barn swallows (Scordato & Safran 2014). Both throat coloration and breast coloration are variable among the populations in this study, but in different ways. For example, males in our North American populations (Colorado, New York) have the darkest breast color, yet the lightest throat color compared with other populations. Ventral coloration has been the focus of several correlational and experimental studies in barn swallows. Together, these studies indicate that melanin-based ventral color varies among populations (Scordato & Safran 2014), is heritable (Hubbard *et al.* 2015) and relates to social mate selection and paternity allocation (Safran & McGraw 2004; Vortman *et al.* 2011). Thus, patterns of ventral color differentiation are likely under divergent sexual selection, although this hypothesis requires further testing.

Patterns of IBA may be underlain by divergent selection on traits that directly impede gene flow, or through the build up of an association between genome-wide divergence and trait divergence in allopatric populations. Whereas models of IBA have typically been applied to scenarios where gene flow is possible between populations, most of the eight populations in our study are geographically isolated from one another, with the exception of those within continental Europe (Czech Republic and Romania). It is possible that gene flow occurs among populations sampled from western and central Europe and Israel, although a formal analysis is needed to make direct inferences about recent historical or ongoing gene flow in these regions. Nevertheless, divergence in the absence of gene flow is an important aspect of avian diversification (e.g. Price 2008), and thus it is highly relevant to ask questions about associations between phenotypic and genomic divergence in order to better understand patterns and consequences of trait evolution in isolated yet closely related populations. Still, ongoing studies of IBA in locations where divergent populations come into sympathy will be critical for inferring whether trait divergence that may have evolved in isolation imposes a barrier to gene flow.

Acknowledgments — We thank C. Alex Buerkle for collaboration on all aspects of this project and the Safran, Kane, and Flaxman laboratory groups at the University of Colorado for feedback and advice. We thanks Matthew Evans for collaboration on barn swallows in the UK. Funding for this project was provided by the National Science Foundation (DEB-CAREER–1149942) to R.J.S. M.R.W was supported by an NSF graduate research fellowship and P.N by a Royal Society of London University Research Fellowship. Work in the Czech Republic was funded by the Ministry of Education, Youth and Sports of the Czech Republic (project LH14045). All field and laboratory protocols were compliant with University of Colorado, and federal animal care and import/export protocols. We are grateful for the excellent reviews provided by Dr. Anna Santure and three anonymous reviewers.

References


Baldassarre DT, Thomassen HA, Karubian J, Webster MS (2013) The role of ecological variation in driving divergence of sexual and non-sexual traits in the red-backed fairy-wren (*Malurus melanopechus*). *BMC Evolutionary Biology*, 13, 75.


**Author contributions**

R.J.S. designed the study, R.J.S., M.R.W, J.K.H, B.R.J., T.A., H.K., Y.V., P.P., S.S. and S.C., collected data in the field, N.C and E.S.C.S. completed the draft assembly of the genome, B.R.J. conducted most of the laboratory work, R.J.S., E.S.C.S., S.M.F. and T.P. analysed the data with input from P.N., R.J.S. wrote the manuscript with input from E.S.C.S., P.N., T.P and N.K. All authors gave final approval for publication.

**Data accessibility**


**Supporting information** follows the References.

**Figure S1** Statistical summary of population genetic structure based on principal component axes one (PC1), and two (PC2) [Panel A], three (PC3), and four (PC4) [Panel B] based on genotype probabilities.