Phenotypic differentiation is associated with divergent sexual selection among closely related barn swallow populations

M. R. WILKINS*†, H. KARAARDIÇ‡§, Y. VORTMAN¶**, T. L. PARCHMAN††,
T. ALBRECHT‡‡§§, A. PETRŽELKOVÁ‡‡§§, L. ÖZKAN‡¶¶, P. L. PAP***,
J. K. HUBBARD*†, A. K. HUND* & R. J. SAFRAN*

*Department of Ecology and Evolutionary Biology, University of Colorado, Boulder, CO, USA
†School of Biological Sciences, University of Nebraska-Lincoln, Lincoln, NE, USA
‡Department of Biology, Faculty of Science, Akdeniz University, Antalya, Turkey
§Elementary Science Education Department, Education Faculty, Alanya Alaaddin Keykubat University, Alanya, Turkey
¶Department of Zoology, Tel-Aviv University, Tel-Aviv, Israel
**Department of Animal Sciences, Hula Research Center, Tel-Hai College, Tel-Hai, Israel
††Department of Biology, University of Nevada, Reno, NV, USA
‡‡Institute of Vertebrate Biology, Czech Academy of Sciences, Brno, Czech Republic
§§Department of Zoology and Ecology, Charles University in Prague, Prague, Czech Republic
¶¶Department of Wildlife Ecology and Management, Faculty of Forestry, Düzce University, Düzce, Turkey
***Evolutionary Ecology Group, Hungarian Department of Biology and Ecology, Babeş-Bolyai University, Cluj-Napoca, Romania

Abstract

Sexual selection plays a key role in the diversification of numerous animal clades and may accelerate trait divergence during speciation. However, much of our understanding of this process comes from phylogenetic comparative studies, which rely on surrogate measures such as dimorphism that may not represent selection in wild populations. In this study, we assess sexual selection pressures for multiple male visual signals across four barn swallow (*Hirundo rustica*) populations. Our sample encompassed 2400 linear km and two described subspecies: European *H. r. rustica* (in the Czech Republic and Romania) and eastern Mediterranean *H. r. transitiva* (in Israel), as well as a potential area of contact (in Turkey). We demonstrate significant phenotypic differentiation in four sexual signalling axes, despite very low-level genomic divergence and no comparable divergence in an ecological trait. Moreover, the direction of phenotypic divergence is consistent with differences in sexual selection pressures among subspecies. Thus, *H. r. transitiva*, which have the darkest ventral plumage of any population, experience directional selection for darker plumage. Similarly, *H. r. rustica*, which have the longest tail feathers of any population, experience directional selection for elongated tail feathers and disruptive selection for ventral plumage saturation. These results suggest that sexual selection is the primary driver of phenotypic differentiation in this species. Our findings add to growing evidence of phenotypic divergence with gene flow. However, to our knowledge, this is the first study to relate direct measures of the strength and targets of sexual selection to phenotypic divergence among closely related wild populations.

Correspondence: Matthew R. Wilkins, 410 Manter Hall, 1104 T St, Lincoln, NE 68588-0118, USA.
Tel.: +256 655 6826; fax: +402 472 2083;
e-mail: m.r.wilkins06@gmail.com

Introduction

The role of sexual selection in population divergence and speciation has garnered tremendous research interest over the past several decades (West-Eberhard, 1983; Wilson *et al.*, 2000; Panhuis *et al.*, 2001; Ritchie, 2007;

Kraaijeveld *et al.*, 2011; Safran *et al.*, 2013). From this work, it is clear that sexual selection plays a key role in the diversification of numerous animal clades (See-hausen, 2000; Masta & Maddison, 2002; Stuart-Fox & Owens, 2003; Mendelson & Shaw, 2005; Boul *et al.*, 2007; Seddon *et al.*, 2008; Sauer & Hausdorf, 2009; Sullivan-Beckers & Cocroft, 2010). Sexual selection may also accelerate the rate of trait divergence relative to ecological selection during the process of speciation (Seddon *et al.*, 2013; Rowe *et al.*, 2015).

Much of our understanding of the role and relative importance of sexual selection in speciation comes from phylogenetic comparative studies, which reveal a correspondence between trait characteristics and rates of diversification (Seddon *et al.*, 2008, 2013; Huang & Rabosky, 2014), but cannot reveal the mechanistic basis of trait evolution itself – key for identifying a role of sexual selection in speciation (Safran *et al.*, 2013). To fully understand how differences in sexual signals and preferences affect phenotypic divergence and speciation, we need measures of the form of sexual selection in recently diverged wild populations. However, the clearest demonstrations to date of a role for sexual selection in signal and population divergence have been based on preference tests in a laboratory setting (Endler & Houde, 1995; Bosch *et al.*, 2000; Ryan *et al.*, 2007; Grace & Shaw, 2011; Williams *et al.*, 2013; Selz *et al.*, 2016) or among populations of unknown genetic relatedness (Rodríguez *et al.*, 2006; Sullivan-Beckers & Cocroft, 2010). Although there is some indication that sexual signals may evolve more rapidly than ecological traits (Arnegard *et al.*, 2010; Safran *et al.*, 2012), we also generally lack information on the relative influence of ecological and sexual selection in driving population divergence (Maan & Seehausen, 2011; Safran *et al.*, 2013). Thus, despite insight into variation in signals and mate preferences across diverse taxa (Rodríguez *et al.*, 2013), we still know little about how differences in sexual selection pressures relate to phenotypic and population divergence in nature.

We utilize the cosmopolitan barn swallow (*Hirundo rustica*) as a study system to investigate the relative importance of sexual selection for phenotypic divergence across closely related populations. Barn swallows comprise six closely related subspecies that breed across the Holarctic. Subspecies are characterized by pronounced phenotypic divergence in known signalling traits, including tail feather (streamer) length and ventral coloration, although all subspecies are ecologically similar—that is, aerial insectivores that build mud cup nests on human-constructed buildings and bridges (Scordato & Safran, 2014). In this study, we tested predictions based on the signatures of speciation by sexual selection outlined by Panhuis *et al.* (2001) and Safran *et al.* (2013). If sexual selection is a dominant driver of phenotypic divergence in this species, we expect the following: (i) greater divergence in sexual than

ecological traits and (ii) that the direction of sexual trait divergence corresponds to differences in the form of sexual selection among populations. We test these predictions by first examining phenotypic divergence in four putative sexual signalling axes describing ventral colour and tail feather length among four populations in Europe and Asia. We compare differentiation in these traits to divergence in a nonsignalling ecological trait (wing length) that is an important measure of flight and foraging performance (Marchetti *et al.*, 1995). We then assess whether trait divergence patterns are more consistent with population genetic structure based on genomewide restriction site-associated DNA sequencing (RADseq) data or differences in sexual selection pressures. In combination, our data and analyses enable us to relate direct measures of the strength and targets of sexual selection to phenotypic and genomic divergence among closely related wild populations.

Our study sites encompassed breeding populations of two (of six) currently described barn swallow subspecies (*Hirundo rustica rustica* and *H. r. transitiva*) with low levels of genetic differentiation (Dor *et al.*, 2010). Previous population genetic work utilizing neutral genetic and mitochondrial markers has indicated no evidence of population structure within the European subspecies *H. r. rustica* (Dor *et al.*, 2010, 2011; Santure *et al.*, 2010), which breed from Western Europe to parts of Asia (Turner, 2010). Current evidence suggests that low levels of genetic differentiation exist between widespread *H. r. rustica* and *H. r. transitiva*, which breed in a narrow range in the eastern Mediterranean. This pattern may result from recent gene flow between the two subspecies, but may also signify early population divergence as a result of a shift in the target of mate preferences (Dor *et al.*, 2011, Safran *et al.*, 2016a). Southern Turkey represents an area of potential contact between these subspecies; however, the degree of phenotypic and genetic overlap in this area was unknown prior to this study.

Previous research across 22 other European and North African populations of *H. r. rustica*, not sampled here, demonstrated consistent directional selection for elongated tail feathers (streamers) (Møller *et al.*, 2006). In contrast, correlational (Vortman *et al.*, 2011) and experimental (Vortman *et al.*, 2013) work suggests that *H. r. transitiva* females in Israel select males based on the combination of two sexual traits: elongated tail feathers and dark melanin-based feather plumage. Despite recent evidence of sexual dichromatism in ventral plumage pigmentation (i.e. darker males) within *H. r. rustica* (Saino *et al.*, 2013), the relationship between ventral plumage colour and reproductive success has not been tested in European populations and precisely how selection varies along these phenotypic axes within and across subspecies remains unclear.

In contrast to sexual signalling divergence, there is no known divergence in feeding ecology within the

subspecies complex. However, there are differences in migratory behaviour, as *H. r. rustica* are long-distance migrants to sub-Saharan Africa, whereas *H. r. transitiva* are considered residents in the eastern Mediterranean (Vortman *et al.*, 2011). Thus, divergence in sexual signals as a result of differential targets of female preference may explain apparent morphological and genetic differences among these subspecies. Female preferences may also be important for maintaining population structure between these subspecies if divergent signals are linked to locally adapted phenotypes (social or migratory) and lead to assortative mating based on these phenotypes (van Doorn *et al.*, 2009). Thus, the barn swallow complex offers a tractable system to investigate how the targets of female preference affect phenotypic divergence at a continental scale.

Materials and methods

Field methods

We studied barn swallows at breeding sites near Ami'ad, Israel (Lat 32.92888, Lon 35.54073), in 2009; Boğazkent, Turkey (Lat 36.85771, Lon 31.1606), in 2010 and 2011; Cojocna, Romania (Lat 46.75319, Lon 23.83464), in 2010; and Lužnice, Czech Republic (Lat 49.0692, Lon 14.7112), in 2013 (shown in Fig. 1a). Based on previous genetic analyses (Dor *et al.*, 2010; Santure *et al.*, 2010), our populations in the Czech Republic and Romania were expected to be genetically undifferentiated *H. r. rustica*, whereas the Israeli population is known to be *H. r. transitiva* (Dor *et al.*, 2011). No previous genetic sampling has characterized Turkish barn swallows. Therefore, given Turkey's importance as a migratory flyway (Leshem & Yom-Tov, 1998; Tottrup *et al.*, 2008) and its proximity to Israel, we were unsure of this population's subspecies affiliation. Thus, an additional goal of this study was to better understand the placement of this population in the context of other members of the *Hirundo rustica* subspecies complex.

At each location, barn swallows were captured during the early breeding season, at which time we took morphological measures (wing and streamer length), as well as blood and feather samples. In addition, unique combinations of a colour band and/or coloured ink markings on the white spots of tail feathers were applied to later link each individual to an active nest. In Israel (2009), Turkey (2010–2011) and the Czech Republic (2013), we further determined the date of breeding onset (date the first egg was laid) for banded individuals by closely monitoring each nest found within study areas. Fitness data were not collected in Romania in 2010, and we thus only present morphological divergence data for this population. We present 2 years of data in Turkey because the nature of selection in this potential contact zone was unknown and might be more subject to changes in the direction or

targets of selection, based on the provenance (and mating preferences) of immigrants.

Colour measurements

Feather samples from four ventral patches (throat, breast, belly, vent) were taped to a standard white card background so that they overlapped as they do on the body of a bird. The colour of each patch was measured using a spectrometer (USB 4000, Ocean Optics), pulsed xenon light (PX-2, Ocean Optics) and SpectraSuite software (v2.0.151). The probe was held at 90 degrees to the feather surface at a distance such that a 2.5 mm diameter of the surface was illuminated and measured. Each measurement was an average of 20 scans of the spectrometer, and each sample was measured three times, lifting the probe between measurements, and averaged. From the generated spectra, we used the 'TCS' function in the package {pavo} (Maia *et al.* 2013) using R v3.1.3 (R Core Team, 2015) to calculate four colour measures for each patch within the tetrahedral colour space (Endler & Mielke, 2005; Stoddard & Prum, 2008). These colour metrics describe the relative stimulation of the four bird cone channels (u, s, m and l), relative to the achromatic origin. Thus, each colour can be described by a vector and represented by θ (horizontal angle), Φ (vertical angle) and r (vector length), with θ and Φ reflecting visible and ultraviolet hues. As barn swallow ventral pigmentation does not reflect in the UV spectrum (Safran & McGraw, 2004; Vortman *et al.*, 2011; Saino *et al.*, 2013), we only used θ as a measure of chromatic colour. Additionally, because the colour space is a tetrahedron, rather than a sphere, the maximum value of r varies by hue. As such, we use r_A (achieved chroma, r/r_{max}) as a relative measure of saturation. Last, we calculated brilliance, or the averaged reflectance values for each measured wavelength between 300 and 700 nm. Collectively, our colour metrics, θ, r_A and brilliance describe the colour of each ventral plumage patch, quantified using the average avian UV visual model, as defined in {pavo}.

Phenotypic variables

To reduce the dimensionality of our phenotypic measures, and account for strong correlations between some traits, we performed principal components analysis (PCA) on 13 colour and morphological measures, including males from all populations (raw trait means are shown in Table S1). We extracted four PCs with eigenvalues near one (PC4 had an eigenvalue of 0.96, but loaded highly for streamer length, which has known sexual signalling importance); varimax rotation was applied to increase interpretability of loadings, using the 'principal' function in the R package {psych} (Revelle, 2015). There were no significant correlations between traits for any population*year, after
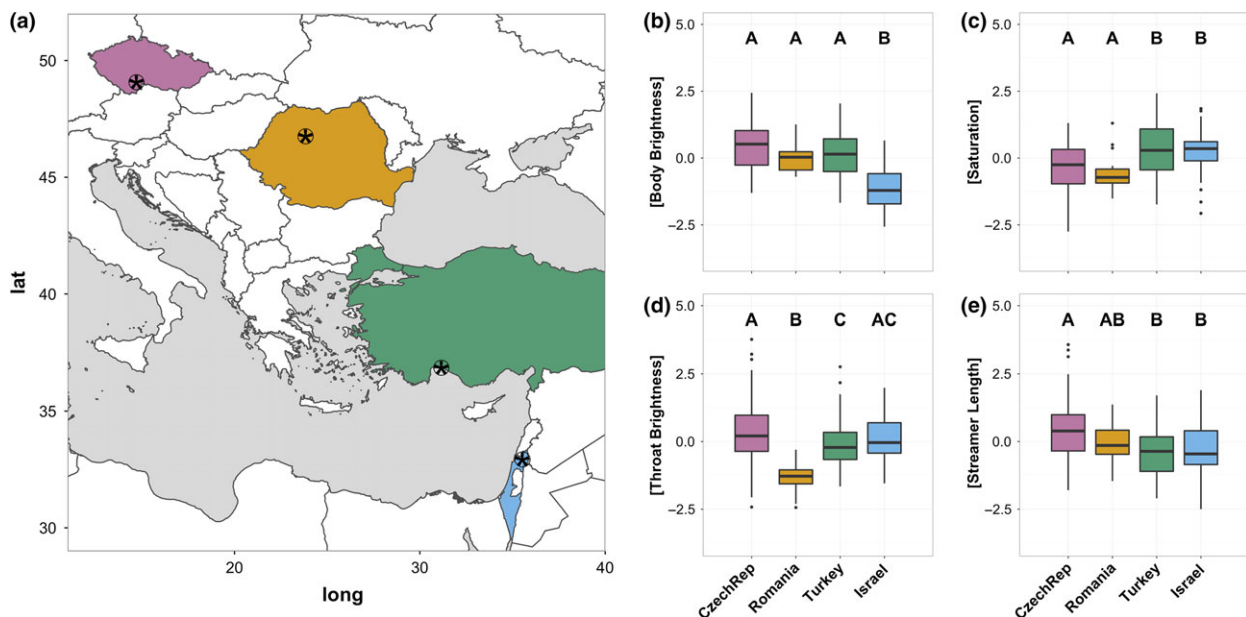
Fig. 1 Study sites, inter- and intrapopulation phenotypic variation for male barn swallows. (a) Sampling sites (west to east) in the Czech Republic, Romania, Turkey and Israel. Panels show population differences in: (b) [Body Brightness], (c) [Saturation], (d) [Throat Brightness] and (e) [Streamer Length]. For Turkey, only data for first capture were included, to avoid pseudoreplication. Different letters above boxplots denote significant differences ($P < 0.05$) between population means, according to Tukey's honestly significant difference tests.

accounting for false-discovery rate (Table S2). Fitting all populations on the same set of axes afforded us consistent units to compare divergence and selection pressures among populations. The number of unique males with complete morphological and colour measures for each population was as follows: Israel = 57; Turkey 2010 = 61; Turkey 2011 = 58; Romania = 16; and Czech Republic = 84. We assessed interpopulation variation in phenotypic axes using ANOVA. We additionally Z-transformed (i.e. standardized to a mean of zero and unit variance) right wing length measures to be on the same scale as PC scores and compared divergence across populations in a trait that is expected to have a similar ecological function across populations and lacks a role in sexual signalling. For traits that showed significant interpopulation variation, we calculated pairwise differences using Tukey's HSD *post hoc* tests.

Population genetic analysis

Our sample sizes for genetic analyses were as follows: Israel, 2009 = 37; Turkey, 2010 = 57; Romania, 2010 = 16; Czech Republic, 2010 = 24. DNA was extracted from blood or feather samples, and single nucleotide polymorphism (SNP) data were generated using a genotyping-by-sequencing (GBS) approach (Gompert *et al.*, 2012; Parchman *et al.*, 2012). Two restriction enzymes, EcoRI and MseI, were used to digest genomic DNA, and custom oligonucleotide adaptors were ligated to digested fragments. EcoRI adaptors

were built with unique eight to ten base pair DNA barcode sequences and an Illumina adaptor; the MseI adaptors contained the opposite Illumina adaptor. Uniquely barcoded ligation products from all individuals were pooled and PCR-amplified using standard Illumina primers. Libraries were size selected for a region between 350 and 450 and sequenced on a single Illumina HiSeq lane at the National Center for Genome Resources (Santa Fe, NM).

After quality and contaminant filtering, we used a Perl script to recognize barcodes assigned to each individual bird, to correct errors in barcode sequences and to remove sequences containing portions of the Illumina adaptors. We then aligned all reads for all individuals against the barn swallow genome assembly (see Safran *et al.* 2016) using the aln and samse algorithms in 'bwa' v 0.7.8 (Burrows-Wheeler Aligner; Li & Durbin, 2009), using an edit distance of 4 and the remaining parameters set as default. An average of 61.66% of reads assembled per individual, with 97.5% of individuals having > 57.9% of reads assembling. We identified single nucleotide variants and estimated genotype likelihoods using 'samtools' v. 1.19 and 'bcftools' v. 1.19 (Li *et al.*, 2009). We only considered single nucleotide variants when 60% of the individuals had at least one read at the locus. We removed variable sites with more than one alternate allele and loci with minor allele frequencies < 5%. For assembled contigs containing more than one SNP, we randomly selected a single SNP to increase the independence of loci used in subsequent

analyses. The GBS data generated for these populations were part of a larger, worldwide population genetic analysis of barn swallows (Safran *et al.* 2016), where a more detailed description of assembly and variant calling can be found.

We used a hierarchical Bayesian model that incorporates uncertainty in sequencing coverage and error across loci and individuals to estimate allele frequencies and genotype probabilities simultaneously for all individuals based on estimated genotype likelihoods (Gompert *et al.*, 2012). This model treats population allele frequencies as priors and simultaneously allows the estimation of allele frequencies and genotype probabilities while incorporating uncertainty arising from variation in sequence coverage. We obtained posterior estimates of genotype probabilities by running 2000 MCMC steps after a 1000 step burn-in and thinning every other step. Mixing and convergence of MCMC steps were clearly evident upon examination of the MCMC histories. The estimated genotype probabilities were then converted to composite genotype values, where an individual's genotype ranged from 0 to 2 at a locus. To further summarize genetic differentiation among populations, we calculated pairwise F_{ST} (Hudson *et al.*, 1992) using code written in R. We assessed the significance of F_{ST} estimates using a permutation-based approach. We further assessed the relationship between pairwise F_{ST} and geographic distance using a Mantel test, using 'mantel.rtest' in the R package {ade4} (Dray & Dufour, 2007).

Sexual selection across populations

Standardized directional (β') and quadratic (γ') selection gradients were calculated for each trait, for each population*year following (Lande and Arnold 1983). All phenotypic factors were scaled and centred by subtracting from each population mean and dividing by the standard deviation. Relative fitness was calculated by dividing by mean fitness for each population and used as the response variable. Directional β' estimates were calculated as the coefficients of multivariate linear regression, including all four putative sexual signalling axes. Quadratic γ' estimates were calculated as the coefficients of the second-degree polynomial term in a model including all four signalling axes, four quadratic terms and the six cross-product combinations. Quadratic estimates and standard errors were doubled, following Stinchcombe *et al.* (2008).

Our measure of fitness was clutch initiation (the Julian day the first egg was laid in a focal male's nest), as this fitness component correlates strongly with fledging success in males' nests in Turkey (2010: Spearman's $\rho = -0.572$, $n = 45$, $P = 7.70\text{e-}6$; 2011: $\rho = -0.443$, $n = 38$, $P = 0.002$) and other populations (Møller, 1994; Safran & McGraw, 2004) and is unaffected by experimental protocols which differed across sites.

Moreover, in a previous study, sexual selection differentials calculated from female clutch initiation and fledging success were strongly correlated across 22 European and North African populations (Møller *et al.*, 2006), reinforcing the suitability of this reproductive metric as a surrogate of fitness. However, as the directionality of this variable is opposite of most fitness measures, we will refer to it as RelCI (relative clutch initiation), rather than 'w' typically used to denote relative fitness. For interpreting significance of selection gradients, we corrected *P*-values for false-discovery rate, as this is a superior method for controlling analysis-wide type I error when performing multiple comparisons (Benjamini *et al.*, 2001; Nakagawa, 2004). Our sample sizes for selection gradients were as follows: Israel = 29; Turkey 2010 = 52; Turkey 2011 = 50; and Czech Republic = 59.

Because our analysis of selection gradients revealed divergent targets of sexual selection among populations, we visualized fitness surfaces for Israel, the Czech Republic and Turkey for two different years. To facilitate interpretation, we used Z-transformed raw data for the traits which had the highest PC loadings on the divergent trait axes and mapped fitness onto these axes as contour plots using the R package {fields} (Nychka *et al.*, 2015).

Results

Dimension reduction of phenotypic variables

Our principal components analysis of 13 colour and morphological traits for four populations of male barn swallows resulted in four rotated components with eigenvalues near one. These four components, describing 77% of the cumulative variance, have been renamed according to loadings (shown in Table 1). Higher values of [Body Brightness] indicate lighter colour and yellower hue for belly, breast and vent; higher [Throat Brightness] values indicate lighter colour and yellower hue for the throat patch; higher values of [Saturation] reflect greater colour intensity in all ventral patches, particularly in the throat and vent; and higher [Streamer Length] values correspond to longer tail streamers.

Phenotypic variation among populations

Figure 1b demonstrates variation in the four putative sexual signalling axes. Populations did not vary in wing length (ANOVA, $F_{3,327} = 1.643$, $P = 0.179$), although at least one population differed from the other three for all four potential sexual signalling traits (Fig. 1b–d). [Body Brightness] was significantly lower in Israel than the other populations by 1.57–1.17 SD, whereas Saturation was higher in Israel (by 0.61–0.78 SD) and Turkey (by 0.72–0.88 SD) than in Romania or the Czech

Table 1 Loadings for varimax-rotated principal components describing barn swallow phenotypes. The axis of maximal loading is highlighted in grey for each trait. Eigenvalues are for rotated axes.

	[Body Brightness]	[Throat Brightness]	[Saturation]	[Streamer Length]
Maximum tail streamer length	0.069	−0.051	−0.023	0.943
Throat theta	0.087	0.940	0.194	−0.020
Throat r achieved	0.068	0.020	0.875	0.093
Throat brilliance	0.116	0.869	−0.339	−0.032
Breast theta	0.873	−0.018	0.093	−0.041
Breast r achieved	−0.692	−0.031	0.408	−0.194
Breast brilliance	0.814	0.029	−0.209	0.109
Belly theta	0.833	0.016	0.083	−0.144
Belly r achieved	−0.611	−0.145	0.530	−0.094
Belly brilliance	0.712	0.218	−0.341	0.094
Vent theta	0.864	0.169	−0.024	0.088
Vent r achieved	−0.443	−0.094	0.657	−0.261
Vent brilliance	0.767	0.137	−0.317	0.265
Eigenvalue	5.039	2.074	1.769	1.136
Proportion variance	0.388	0.16	0.136	0.087
Cumulative variance	0.388	0.547	0.683	0.771

Republic. [Throat Brightness] showed the greatest variation among populations, with Romania having the darkest by 1.18–1.67 SD and nearby Czech Republic having the lightest throat colour by 0.25–1.67 SD. Given that we did not predict such a striking difference in colour among European populations and to account for the smaller sample size in Romania ($n = 16$ vs. $n = 84$ in Czech Republic), we subsampled 16 random males from the Czech Republic and performed a two-sample Wilcoxon test, repeating this 1000 times. For this procedure, all $P \leq 0.003$, indicating the robustness of throat colour differences between these two European populations. Additionally, [Streamer Length] was significantly (0.73–0.78 SD) greater in the Czech Republic than Turkey or Israel, whereas Romania was intermediate to these three populations (difference range: 0.27–0.46 SD; Fig. 1d).

Sexual selection

Table 2 shows standardized linear (β') and quadratic (γ') selection gradients for the four putative sexual signals across three populations: Czech Republic, Turkey (for 2010 and 2011) and Israel. Two linear selection gradients were significant for putative sexual traits, even after controlling for false-discovery rate: [Streamer Length] in the Czech Republic and [Body Brightness] in Israel. The negative gradient for the Czech Republic indicates that males with longer streamers bred earlier and had higher reproductive success. In contrast, in

Israel, darker males bred earlier. In the Czech Republic, there was also significant negative quadratic selection for [Saturation], potentially indicating disruptive selection on ventral plumage colour intensity (i.e. high and low colour intensities bred earlier; note the directionality is reverse from typical quadratic selection gradients because lower clutch initiation dates are beneficial for fitness). However, as noted by Kingsolver & Diamond (2011), the sign of quadratic selection gradients does not directly relate to the pattern of selection, as the proximity of most population trait means to fitness peaks should bias γ towards stabilizing selection. There was negative directional and positive quadratic selection for [Streamer Length] in Turkey in 2011; however, neither of these relationships was significant after controlling for false-discovery rate.

Age is a potential confounding factor with any comparative population study. Because ventral darkness and streamer length are known to increase with age (Lifjeld *et al.*, 2011; Vortman *et al.*, 2015), we reran our analyses on a subset of individuals that we could confidently assign as yearling males (those which we had either banded as nestlings or were unbanded males breeding at a site where we banded adults the previous year). This reduced our sample size to 40 males in the Czech Republic and 23 males in Turkey 2011. We performed a pooled PCA (with similar loadings to the full analysis) and calculated directional selection gradients, shown in Table S3. These results were highly correlated with values from the full data set (Spearman's $\rho = 0.762$, $P = 0.037$), indicating that patterns of divergent selection are not due to age-related differences in our population samples.

To further visualize differences in selection pressures among populations, we made contour plots of bivariate fitness surfaces for the traits loading highest on [Streamer Length] (streamer length) and [Body Brightness] (breast θ) across the four population*years for which we had reproductive data. From these plots (Fig. 2), there are striking differences between the fitness surface of Israel and the other populations. Surfaces for the Czech Republic and both years in Turkey are mostly flat (green hues, and contour lines near one, indicating average fitness values across most of the surface), with slight variation in fitness favouring longer streamers. In stark contrast, there is a pronounced peak in the lower right of the Israel fitness surface, indicating strong benefits of early breeding for males with the greatest combination of dark breast plumage and long streamers in that population.

Population genetics

Initial variant calling utilizing samtools and bcftools resulted in 67 773 single nucleotide variants. After discarding loci with minor allele frequencies < 0.05, and randomly retaining one single nucleotide variant per

Table 2 Standardized directional (*β′*) and quadratic (*γ′*) selection differentials for the four putative sexual signal axes across population*years (using relative clutch initiation date as a fitness metric). Bolded values are significant after correcting for false-discovery rate.

Country	Trait	N	*β′*	SE	*t* Value	*P* value	Adj. *P* value	*γ′*	SE	*t* Value	*P* value	Adj. *P* value
Czech Rep	[Body Brightness]	59	0.011	0.013	0.837	0.406	0.502	0.031	0.018	1.685	0.099	0.198
Czech Rep	[Throat Brightness]		−0.020	0.014	−1.492	0.142	0.283	0.029	0.022	1.302	0.200	0.200
Czech Rep	[Saturation]		0.009	0.014	0.676	0.502	0.502	**−0.094**	**0.028**	**−3.362**	**0.002**	**0.006**
Czech Rep	[Streamer Length]		**−0.035**	**0.013**	**−2.679**	**0.010**	**0.039**	0.021	0.016	1.333	0.190	0.200
Turkey10	[Body Brightness]	52	−0.005	0.022	−0.231	0.818	0.837	0.044	0.039	1.113	0.273	0.546
Turkey10	[Throat Brightness]		0.006	0.027	0.207	0.837	0.837	0.016	0.075	0.216	0.830	0.830
Turkey10	[Saturation]		−0.027	0.021	−1.267	0.211	0.423	0.085	0.048	1.783	0.083	0.331
Turkey10	[Streamer Length]		−0.037	0.023	−1.584	0.120	0.423	−0.016	0.048	−0.334	0.740	0.830
Turkey11	[Body Brightness]	50	−0.021	0.016	−1.279	0.207	0.415	0.013	0.027	0.464	0.646	0.646
Turkey11	[Throat Brightness]		0.008	0.016	0.514	0.610	0.701	−0.022	0.034	−0.641	0.526	0.646
Turkey11	[Saturation]		0.007	0.018	0.386	0.701	0.701	−0.034	0.038	−0.893	0.378	0.646
Turkey11	[Streamer Length]		−0.034	0.016	−2.076	0.044	0.175	0.078	0.033	2.367	0.024	0.094
Israel	[Body Brightness]	29	**0.177**	**0.065**	**2.721**	**0.012**	**0.048**	−0.206	0.234	−0.880	0.394	0.963
Israel	[Throat Brightness]		0.070	0.061	1.141	0.265	0.530	0.153	0.212	0.723	0.481	0.963
Israel	[Saturation]		0.038	0.057	0.672	0.508	0.678	−0.003	0.109	−0.030	0.977	0.977
Israel	[Streamer Length]		0.018	0.053	0.333	0.742	0.742	−0.020	0.128	−0.155	0.879	0.977
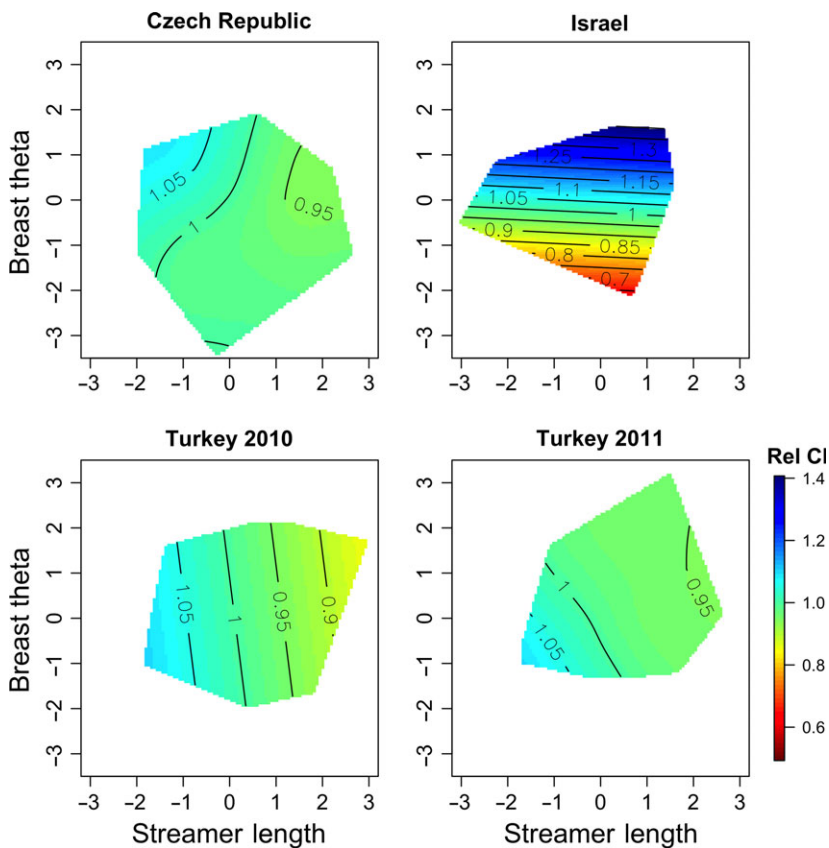
Fig. 2 Contour plots showing fitness landscapes for the three populations in which we measured breeding onset. To facilitate interpretation, *x*- and *y*-axes are the centred, scaled traits with highest loadings on [Streamer Length] and [Body Brightness], respectively. Colours and contour lines reflect Rel CI, that is clutch initiation relative to the population mean (our surrogate of fitness), interpolated using thin plate splines. Red hues are associated with earlier than average breeding and higher overall fitness, whereas cooler (bluer) colours reflect worse reproductive performance. There is a striking difference in the fitness landscape of the Israeli population, with much greater variation in breeding onset and a pronounced peak, favouring males with darker breast colour (lower breast theta) and longer streamers. The other population*years show much weaker selection, favouring longer streamers, with little influence of breast colour.

GBS contig, we retained a final set of 9493 SNPs. For this final set of loci, the average coverage depth per site per individual was 1.5×. Although this coverage is low, this type of data is appropriate for population-level inferences when analysed with models that incorporate uncertainty arising from variability in sequencing coverage (e.g. Neilsen *et al.*, 2011; Gompert *et al.*, 2012), such as the model we used to estimate genotype

probabilities (Gompert *et al.*, 2012; Buerkle & Gompert, 2013). GBS data were part of a larger population genetic study of worldwide barn swallow populations, and further details of assembly, variant calling and associated parameter setting and data can be found in Safran *et al.* 2016a. Table 3 demonstrates average F_{ST} values (upper diagonal) for each population pair based on 9493 loci. All values are significant (Table S4), but very close to zero, indicating extremely little genetic differentiation among our sample populations, despite being separated by as much as 2400 km. The highest F_{ST} value of 0.044 was between Romania and Israel, although this value is close to the comparison between Romania and nearby Czech Republic (F_{ST} = 0.041). Moreover, there is no relationship between F_{ST} and geographic distance (Mantel test: Pearson's r = 0.139, permutations = 1000, P = 0.373), further suggesting that geography has no detectable effect on genetic differentiation for these four populations. Thus, there is evidence of significant phenotypic differentiation, without a signature of isolation by distance.

Discussion

Consistent with predictions for a role of sexual selection in speciation (see Panhuis *et al.* (2001), Safran *et al.* (2013), we found the following: (i) greater divergence in putative sexual signalling than a nonsignalling ecological trait and (ii) the direction of divergence among subspecies is consistent with differences in fitness surfaces, despite very shallow genomewide divergence. Based on these results and previous work in Europe and Israel (Møller, 1988, 1994; Vortman *et al.*, 2011, 2013; Safran *et al.*, 2016b), we infer that divergence in mating preferences among populations has driven phenotypic divergence. Although migratory behaviour differs between long-distance migrating *H. r. rustica* and resident *H. r. transitiva*, wing length— an important migratory trait (Marchetti *et al.*, 1995)— did not vary across populations. Our results add to growing evidence that shifts in mating preferences may lead to phenotypic divergence with little or no differentiation in ecological traits in the early stages of speciation (Irwin *et al.*, 2001; Mendelson & Shaw, 2005; Arnegard *et al.*, 2010; Safran *et al.*, 2012).

Table 3 Genomic divergence and geographic distance between barn swallow populations. Lower diagonal: geographic distance in kilometre, calculated with the haversine method; upper diagonal: mean F_{ST} values.

	Czech Rep	Romania	Turkey	Israel
Czech Rep	–	0.041	0.030	0.038
Romania	727	–	0.037	0.044
Turkey	1901	1257	–	0.031
Israel	2492	1831	592	–

Population divergence in phenotype

It is not surprising that males from Israel had significantly lower [Body Brightness] (i.e. darker ventral plumage) than any other population (Fig. 1b), as this is consistent with previous comparative work (Dor *et al.*, 2011) and subspecies descriptions (Turner, 2010; Turner & Christie, 2012). [Streamer Length] follows the expected pattern from previous work as well, with significantly longer streamers within *H. r. rustica* males in the Czech Republic, compared to *H. r. transitiva* males in Israel. The most variable phenotypic trait was [Throat Brightness], which showed significant differences between three of the four populations, with Israel falling between Turkey and the Czech Republic. Curiously, males from Romania had the darkest throat colour of any population, whereas males from the Czech Republic had the lightest. Thus, geographically proximate male Romanian swallows had much darker throats than their Czech counterparts, but they did not differ along any other PC axis. Although we do not have reproductive data for Romania, it is possible that this difference results from selection for darker throat colour, as has been shown for populations in the United States and Japan (Safran & McGraw, 2004; Hasegawa *et al.*, 2010).

Given the lack of any external barrier from the known *H. r. rustica* breeding range, we predicted our Turkish population would be more phenotypically similar to widely distributed European populations than *H. r. transitiva*, which migrate only very short distances. Indeed, males from Turkey had high [Body Brightness], similar to the two European populations (Fig. 1b); however, streamers were significantly shorter, and saturation higher than these populations, and were more similar to the Israel population (Fig. 1). Additionally, [ThroatBrightness] was significantly different and intermediate to the divergent colours of Romania and the Czech Republic and was not different from Israel. These results seem to indicate that the Turkish population is either composed of assortatively breeding *H. r. transitiva* and *H. r. rustica*, or an admixture of these.

Selection and phenotypic divergence

A great deal of research has established that longer tail streamers are sexually selected within *H. r. rustica* throughout Western Europe (Møller, 1994; Møller *et al.*, 2006). Although reproductive benefits of male colour have not previously been considered within *H. r. rustica*, recent evidence of sexual dichromatism in throat pigmentation within an Italian population (Saino *et al.*, 2013) suggests that throat colour may be a target of selection in that population. Here, we directly compare estimates of selection for colour and streamer length among barn swallow populations. Consistent with previous work in other European populations

(Møller *et al.*, 2006), we found males in the Czech Republic with longer streamers bred significantly earlier than males with short streamers (Table 2). Additionally, we found evidence of disruptive selection for [Saturation], but no directional selection on any colour axis. As shown in the fitness landscape in Fig. 2, it is clear that there is a small, but significant advantage of earlier breeding for males in the Czech Republic with longer streamers, with no benefits of being darker (contour lines are oriented more or less vertically). These selection patterns are consistent with phenotypic divergence, as the Czech Republic males had the longest streamers and among the lightest and least saturated ventral colour measures of any population (Fig. 1, Table S1).

In contrast, for our Turkish population, of unknown subspecies, no significant selection gradients were detected after correcting for false-discovery rate. Fitness surfaces for both 2010 and 2011 (Fig. 2) are qualitatively similar to the Czech Republic, with a small benefit of longer streamers, and no effect of colour on breeding onset. When paired with the phenotypic patterns, wherein Turkish males had long streamers and light colour similar to European populations, but high saturation similar to the Israeli population, these results suggest that sexual selection in the Turkish population: (i) has more variable targets (i.e. patterns of female mate preference may vary based on the proportion of rustica- or transitiva-like males), or (ii) primarily involve traits not considered in this analysis, such as song. Further work should assess the importance of song in mate choice within this population and reveal whether there is assortative mating based on phenotype.

Finally, consistent with previous correlational results (Vortman *et al.*, 2011), we found significant selection for darker *H. r. transitiva* males from Israel. Indeed, the fitness surface for Israel (Fig. 2) is much steeper, and oriented perpendicularly to the other populations, with the darkest males breeding significantly earlier, with no effect of streamer length. This is consistent with phenotypic divergence patterns, as males in Israel were much darker than any other population, with shorter streamers. However, previous studies in Israel have shown reproductive benefits associated with long tail streamers (Vortman *et al.*, 2011), and experimental manipulation of colour and streamer length has shown that the combination of increased streamer length and ventral darkness is preferred by females (Vortman *et al.*, 2013). This inconsistency may result from our use of clutch initiation as a surrogate for reproductive performance. However, given strong correlations between reproductive metrics, we propose another interpretation. Although ventral colour and streamer length are uncorrelated in this population (Vortman *et al.*, 2011, 2015; Table S2), the fitness surface for Israel (Fig. 2) shows a disjoint in the distributions of these two traits. That is, males with short tail streamers tended to have average ventral colour. The most successful males (red area on the contour plot) had the darkest breast plumage and also above average streamers (to the right of standardized mean zero in Fig. 2). Although there was no consistent relationship between streamer length and breeding onset, the most successful males had a combination of the darkest plumage and average to above average streamers, concordant with previous work. As both of these traits may indicate age (Lifjeld *et al.*, 2011; Vortman *et al.*, 2015), further work is required to determine what the signalling value of colour and streamer length are to females, while controlling for breeding experience.

Table 3 demonstrates extremely low genetic differentiation across populations, consistent with previous population genetic studies of barn swallow (Santure *et al.*, 2010; Dor *et al.*, 2011). This suggests that the observed phenotypic divergence has occurred in the context of recent historical or ongoing gene flow. We currently know little about the genetic architecture underlying sexually selected phenotypes (Wilkinson *et al.*, 2015), and deeper genetic sampling is required to assess the relationship between phenotype and genotype. For example, it is possible that localized genomic divergence may be occurring at loci responsible for different signalling traits among populations, in spite of pronounced gene flow throughout the rest of the genome (Wu, 2001). Moreover, phenotypic plasticity and genotype-by-environment interactions (Ingleby *et al.*, 2010) may provide additional or alternative explanations for phenotypic divergence in spite of genetic homogeneity.

Conclusions

In this study, we demonstrate significant phenotypic variation in four putative sexual signalling trait axes, in the absence of comparable divergence in an ecological trait, among four barn swallow populations. Our comparison of sexual selection gradients demonstrates that divergence in the strength and targets of selection for dark ventral colour and streamer elongation correspond to the direction of population-level differentiation in these traits. Moreover, despite significant pairwise differences in sexual phenotypes, there was little evidence of population genetic structure. Collectively, this work contributes to a small number of field studies that explore how local variation in sexual selection pressures relates to phenotypic divergence with gene flow at a large scale. Based on our results, we suggest that signal and population divergence are facilitated not only by the strength and direction (i.e. sign) of selection (e.g. Rodríguez *et al.*, 2013), but also the targets (i.e. phenotypic axes) of selection. As more studies characterize variation in selective targets, we may find that cases of consistent selection for a single trait, such as in the European barn swallow (Møller, 1994; Møller *et al.*, 2006), are unusual. Future challenges will

include determining the mechanisms through which mate preferences shift and characterizing the conditions under which such shifts lead to phenotypic and population divergence.

Acknowledgments

Conflict of interest

None.

References

Arnegard, M.E., McIntyre, P.B., Harmon, L.J., Zelditch, M.L., Crampton, W.G.R., Davis, J.K. *et al.* 2010. Sexual signal evolution outpaces ecological divergence during electric fish species radiation. *Am. Nat.* **176**: 335–356.

Benjamini, Y., Drai, D., Elmer, G., Kafkafi, N. & Golani, I. 2001. Controlling the false discovery rate in behavior genetics research. *Behav. Brain Res.* **125**: 279–284.

Bosch, J., Rand, A. & Ryan, M. 2000. Signal variation and call preferences for whine frequency in the túngara frog, *Physalaemus pustulosus*. *Behav. Ecol. Sociobiol.* **49**: 62–66.

Boul, K.E., Funk, W.C., Darst, C.R., Cannatella, D.C. & Ryan, M.J. 2007. Sexual selection drives speciation in an Amazonian frog. *Proc. Biol. Sci.* **274**: 399–406.

Buerkle, C.A. & Gompert, Z. 2013. Population genomics based on low coverage sequencing: how low should we go? *Mol. Ecol.* **22**: 3028–3035.

van Doorn, G.S., Edelaar, P. & Weissing, F.J. 2009. On the origin of species by natural and sexual selection. *Science* **326**: 1704–1707.

Dor, R., Safran, R.J., Sheldon, F.H., Winkler, D.W. & Lovette, I.J. 2010. Phylogeny of the genus *Hirundo* and the barn swallow subspecies complex. *Mol. Phylogenet. Evol.* **56**: 409–418.

Dor, R., Safran, R.J., Vortman, Y., Lotem, A., McGowan, A., Evans, M.R. *et al.* 2011. Population genetics and morphological comparisons of migratory European (*Hirundo rustica rustica*) and sedentary East-Mediterranean (*Hirundo rustica transitiva*) barn swallows. *J. Hered.* **103**: 55–63.

Dray, S. & Dufour, A.B. 2007. The ade4 package: implementing the duality diagram for ecologists. *J. Stat. Softw.* **22**: 1–20.

Endler, J.A. & Houde, A.E. 1995. Geographic variation in female preferences for male traits in *Poecilia reticulata*. *Evolution* **49**: 456.

Endler, J.A. & Mielke, P. 2005. Comparing entire colour patterns as birds see them. *Biol. J. Linn. Soc.* **86**: 405–431.

Gompert, Z., Lucas, L.L.K., Nice, C.C., Fordyce, J.A., Forister, M.L. & Buerkle, C.A. 2012. Genomic regions with a history of divergent selection affect fitness of hybrids between two butterfly species. *Evolution* **66**: 2167–2181.

Grace, J.L. & Shaw, K.L. 2011. Coevolution of male mating signal and female preference during early lineage divergence of the Hawaiian cricket, *Laupala cerasina*. *Evolution* **65**: 2184–2196.

Hasegawa, M., Arai, E., Watanabe, M. & Nakamura, M. 2010. Mating advantage of multiple male ornaments in the barn swallow *Hirundo rustica* gutturalis. *Ornithol. Sci.* **9**: 141–148.

Huang, H. & Rabosky, D.L. 2014. Sexual selection and diversification: reexamining the correlation between dichromatism and speciation rate in birds. *Am. Nat.* **184**: E101–E114.

Hudson, R.R., Slatkin, M. & Maddison, W.P. 1992. Estimation of levels of gene flow from DNA sequence data. *Genetics* **132**: 583–589.

Ingleby, F.C., Hunt, J. & Hosken, D.J. 2010. The role of genotype-by-environment interactions in sexual selection. *J. Evol. Biol.* **23**: 2031–2045.

Irwin, D.E., Bensch, S. & Price, T.D. 2001. Speciation in a ring. *Nature* **409**: 333–337.

Kingsolver, J.G. & Diamond, S.E. 2011. Phenotypic selection in natural populations: what limits directional selection? *Am. Nat.* **177**: 346–357.

Kraaijeveld, K., Kraaijeveld-Smit, F.J.L. & Maan, M.E. 2011. Sexual selection and speciation: the comparative evidence revisited. *Biol. Rev. Camb. Philos. Soc.* **86**: 367–377.

Lande, R. & Arnold, S. 1983. The measurement of selection on correlated characters. *Evolution (N. Y)* **37**: 1210–1226.

Leshem, Y. & Yom-Tov, Y. 1998. Routes of migrating soaring birds. *Ibis* **140**: 41–52.

Li, H. & Durbin, R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**: 1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N. *et al.* 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**: 2078–2079.

Lifjeld, J.T., Kleven, O., Jacobsen, F., McGraw, K.J., Safran, R.J. & Robertson, R.J. 2011. Age before beauty? Relationships between fertilization success and age-dependent ornaments in barn swallows. *Behav. Ecol. Sociobiol.* **65**: 1687–1697.

Maan, M.E. & Seehausen, O. 2011. Ecology, sexual selection and speciation. *Ecol. Lett.* **14**: 591–602.

Maia, R., Eliason, C.M., Bitton, P.-P., Doucet, S.M. & Shawkey, M.D. 2013. pavo: an R package for the analysis,

visualization and organization of spectral data. *Methods Ecol. Evol.* **4**: 906–913.

Marchetti, K., Price, T. & Richman, A. 1995. Correlates of wing morphology with foraging behaviour and migration distance in the genus *Phylloscopus*. *J. Avian Biol.* **26**: 177–181.

Masta, S.E. & Maddison, W.P. 2002. Sexual selection driving diversification in jumping spiders. *Proc. Natl. Acad. Sci. USA* **99**: 4442–4447.

Mendelson, T.C. & Shaw, K.L. 2005. Sexual behaviour: rapid speciation in an arthropod. *Nature* **433**: 375–376.

Møller, A.P. 1988. Female choice selects for male sexual tail ornaments in the monogamous swallow. *Nature* **332**: 640–642.

Møller, A.P. 1994. *Sexual Selection and the Barn Swallow*. Oxford University Press, New York.

Møller, A.P., Chabi, Y., Cuervo, J.J., De Lope, F., Kilpimaa, J., Kose, M. *et al.* 2006. An analysis of continent-wide patterns of sexual selection in a passerine bird. *Evolution* **60**: 856–868.

Nakagawa, S. 2004. A farewell to Bonferroni: the problems of low statistical power and publication bias. *Behav. Ecol.* **15**: 1044–1045.

Neilsen, P.M., Noll, J.E., Suetani, R.J., Schulz, R.B., Al-ejeh, F., Evdokiou, A. *et al.* 2011. Mutant p53 uses p63 as a molecular chaperone to alter gene expression and induce a pro-invasive secretome Paul. *Oncotarget* **2**: 1203–1217.

Nychka, D., Furrer, R., Paige, J. & Sain, S. 2015. *Fields: Tools for Spatial Data*. R package version 8.4-1. University Corporation for Atmospheric Research, Boulder, CO.

Panhuis, T.M., Butlin, R., Zuk, M. & Tregenza, T. 2001. Sexual selection and speciation. *Trends Ecol. Evol.* **16**: 364–371.

Parchman, T.L., Gompert, Z., Mudge, J., Schilkey, F.D., Benkman, C.W. & Buerkle, C.A. 2012. Genome-wide association genetics of an adaptive trait in lodgepole pine. *Mol. Ecol.* **21**: 2991–3005.

R Core Team. 2015. *R: A Language and Environment for Statistical Computing*. R Core Team, Vienna, Austria.

Revelle, W. 2015. *psych: Procedures for Personality and Psychological Research*. Northwestern University, Evanston, IL.

Ritchie, M.G. 2007. Sexual Selection and Speciation. *Annu. Rev. Ecol. Evol. Syst.* **38**: 79–102.

Rodríguez, R.L., Ramaswamy, K. & Cocroft, R.B. 2006. Evidence that female preferences have shaped male signal evolution in a clade of specialized plant-feeding insects. *Proc. R. Soc. B Biol. Sci.* **273**: 2585–2593.

Rodríguez, R.L., Boughman, J.W., Gray, D.A., Hebets, E.A., Höbel, G. & Symes, L.B. 2013. Diversification under sexual selection: the relative roles of mate preference strength and the degree of divergence in mate preferences. *Ecol. Lett.* **16**: 964–974.

Rowe, M., Albrecht, T., Cramer, E.R.A., Johnsen, A., Laskemoen, T., Weir, J.T. *et al.* 2015. Postcopulatory sexual selection is associated with accelerated evolution of sperm morphology. *Evolution* **69**: 1044–1052.

Ryan, M.J., Bernal, X.E. & Rand, A.S. 2007. Patterns of mating call preferences in túngara frogs, Physalaemus pustulosus. *J. Evol. Biol.* **20**: 2235–2247.

Safran, R.J. & McGraw, K.J. 2004. Plumage coloration, not length or symmetry of tail-streamers, is a sexually selected trait in North American barn swallows. *Behav. Ecol.* **15**: 455–461.

Safran, R., Flaxman, S., Kopp, M., Irwin, D.E., Briggs, D., Evans, M.R. *et al.* 2012. A robust new metric of phenotypic distance to estimate and compare multiple trait differences among populations. *Curr. Zool.* **58**: 426–439.

Safran, R.J., Scordato, E.S.C., Symes, L.B., Rodríguez, R.L. & Mendelson, T.C. 2013. Contributions of natural and sexual selection to the evolution of premating reproductive isolation: a research agenda. *Trends Ecol. Evol.* **28**: 643–650.

Safran, R.J., Scordato, E.S.C., Wilkins, M.R., Hubbard, J.K., Jenkins, B.R., Albrecht, T., et al. 2016a. Genome-wide differentiation in closely related populations: the roles of selection and geographic isolation. *Mol. Ecol.* **25**: 3769–4090.

Safran, R.J., Vortman, Y., Jenkins, B.R., Hubbard, J.K., Wilkins, M.R., Bradley, R.J., et al. 2016b. The maintenance of phenotypic divergence through sexual selection: An experimental study in barn swallows Hirundo rustica. *Evolution.* **70**: 2074–2084.

Saino, N., Romano, M., Rubolini, D., Teplitsky, C., Ambrosini, R., Caprioli, M. *et al.* 2013. Sexual dimorphism in melanin pigmentation, feather coloration and its heritability in the barn swallow (*Hirundo rustica*). *PLoS One* **8**: e58024.

Santure, A.W., Ewen, J.G., Sicard, D., Roff, D.A. & Møller, A.P. 2010. Population structure in the barn swallow, *Hirundo rustica*: a comparison between neutral DNA markers and quantitative traits. *Biol. J. Linn. Soc.* **99**: 306–314.

Sauer, J. & Hausdorf, B. 2009. Sexual selection is involved in speciation in a land snail radiation on crete. *Evolution* **63**: 2535–2546.

Scordato, E.S.C. & Safran, R.J. 2014. Geographic variation in sexual selection and implications for speciation in the barn swallow. *Avian Res.* **5**: 1–13.

Seddon, N., Merrill, R.M. & Tobias, J. 2008. Sexually selected traits predict patterns of species richness in a diverse clade of suboscine birds. *Am. Nat.* **171**: 620–631.

Seddon, N., Botero, C.A., Tobias, J.A., Dunn, P.O., Macgregor, H.E.A., Rubenstein, D.R. *et al.* 2013. Sexual selection accelerates signal evolution during speciation in birds. *Proc. R. Soc. B Biol. Sci.* **280**: 20131065.

Seehausen, O. 2000. Explosive speciation rates and unusual species richness in haplochromine cichlid fishes: effects of sexual selection. *Adv. Ecol. Res.* **31**: 237–274.

Selz, O.M., Thommen, R., Pierotti, M.E.R., Anaya-Rojas, J.M. & Seehausen, O. 2016. Differences in male coloration are predicted by divergent sexual selection between populations of a cichlid fish. *Proc. R. Soc. B Biol. Sci.* **283**: 20160172.

Stinchcombe, J.R., Agrawal, A.F., Hohenlohe, P.A., Arnold, S.J. & Blows, M.W. 2008. Estimating nonlinear selection gradients using quadratic regression coefficients: double or nothing? *Evolution* **62**: 2435–2440.

Stoddard, M.C. & Prum, R.O. 2008. Evolution of avian plumage color in a tetrahedral color space: a phylogenetic analysis of new world buntings. *Am. Nat.* **171**: 755–776.

Stuart-Fox, D. & Owens, I.P.F. 2003. Species richness in agamid lizards: chance, body size, sexual selection or ecology? *J. Evol. Biol.* **16**: 659–669.

Sullivan-Beckers, L. & Cocroft, R.B. 2010. The importance of female choice, male-male competition, and signal

transmission as causes of selection on male mating signals. *Evolution* **64**: 3158–3171.

Tottrup, A.P., Thorup, K., Rainio, K., Yosef, R., Lehikoinen, E. & Rahbek, C. 2008. Avian migrants adjust migration in response to environmental conditions en route. *Biol. Lett.* **4**: 685–688.

Turner, A. 2010. *The Barn Swallow*. T & AD Poyser, London, UK.

Turner, A. & Christie, D.A. 2012. Barn Swallow (*Hirundo rustica*). In: *Handbook of the Birds of the World Alive*. (J. del Hoyo, A. Elliott, J. Sargatal, D.A. Christie & E. de Juana, eds.) R package version 8.4-1. Lynx Edicions, Barcelona. (retrieved from http://www.hbw.com/node/57729 on 8 October 2014).

Vortman, Y., Lotem, A., Dor, R., Lovette, I.J. & Safran, R.J. 2011. The sexual signals of the East-Mediterranean barn swallow: a different swallow tale. *Behav. Ecol.* **22**: 1344–1352.

Vortman, Y., Lotem, A., Dor, R., Lovette, I. & Safran, R.J. 2013. Multiple sexual signals and behavioral reproductive isolation in a diverging population. *Am. Nat.* **182**: 514–523.

Vortman, Y., Safran, R.J., Reiner Brodetzki, T., Dor, R. & Lotem, A. 2015. Expression of multiple sexual signals by fathers and sons in the east-mediterranean barn swallow: are advertising strategies heritable? *PLoS One* **10**: e0118054.

West-Eberhard, M.J. 1983. Sexual selection, social competition, and speciation. *Q. Rev. Biol.* **58**: 155–183.

Wilkinson, G.S., Breden, F., Mank, J.E., Ritchie, M.G., Higginson, A.D., Radwan, J. *et al.* 2015. The locus of sexual selection: moving sexual selection studies into the post-genomics era. *J. Evol. Biol.* **28**: 739–755.

Williams, T.H., Gumm, J.M. & Mendelson, T.C. 2013. Sexual selection acting on a speciation trait in darters (Percidae: Etheostoma). *Behav. Ecol.* **24**: 1407–1414.

Wilson, A.B., Noack-Kunnmann, K. & Meyer, A. 2000. Incipient speciation in sympatric Nicaraguan crater lake cichlid fishes: sexual selection versus ecological diversification. *Proc. Biol. Sci.* **267**: 2133–2141.

Wu, C.I. 2001. The genic view of the process of speciation. *J. Evol. Biol.* **14**: 851–865.

Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article:

Table S1 Raw means and standard deviations for the 18 measured traits used in this study.

Table S2 Correlations between phenotypic axes across populations.

Table S3 Directional selection gradients for males that could confidently be considered yearlings (which we banded as nestlings the previous year or were new to a site where we had banded adults the previous year).

Table S4 We compared observed F_{ST} (Hudson *et al.*, 1992) estimates among populations to a null distribution of F_{ST} values (based on 100 permutations of population membership).

Figure S1 Fitness LOESS curves for the ecological trait (right wing length) and four putative sexual signaling axes (PCs) for all four population*years.