This is a repository copy of *Genome-wide differentiation in closely related populations: the roles of selection and geographic isolation.*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/101847/

Version: Accepted Version

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

Safran RJ*[1], Scordato ESC[1], Wilkins MR[1,2], Hubbard, JK[1,2], Jenkins, BR[1], Albrecht T[3], Flaxman SM[1], Karaardıç H[4], Vortman Y[5], Lotem A[5], Nosil P[6], Pap P[7], Shen S[8], Chan, S-F[8], Parchman T[9], Kane NC[1]

* Corresponding author

[1] Department of Ecology and Evolutionary Biology, University of Colorado, Boulder, CO, USA 80309

[2] School of Biological Sciences, University of Nebraska-Lincoln, USA

[3] Department of Zoology, Charles University in Prague and Institute of Vertebrate Biology, Academy of Sciences of the Czech Republic, Czech Republic

[4] Elementary Science Education Department, Education Faculty, Alanya Alaaddin Keykubat University, Alanya, Turkey

[5] Department of Zoology, Tel-Aviv University, Israel

[6] Department of Animal and Plant Sciences, University of Sheffield, Sheffield, UK

[7] Department of Taxonomy and Ecology, Babeş-Bolyai University, Cluj-Napoca, Romania

[8] Biodiversity Research Center, Academia Sinica, Taiwan.

[9] Department of Biology, University of Nevada, Reno, NV, USA

Abstract

Population divergence in geographic isolation is due to a combination of factors. Natural and sexual selection may be important in shaping patterns of population differentiation, a pattern referred to as 'Isolation by Adaptation' (IBA). IBA can be complementary to the well-known pattern of 'Isolation by Distance' (IBD), in which the divergence of closely related populations (via any evolutionary process) is associated with geographic isolation. The barn swallow *Hirundo rustica* complex comprises six closely related subspecies, where divergent sexual selection is associated with phenotypic differentiation among allopatric populations. To investigate the relative contributions of selection and geographic distance to genome-wide differentiation, we compared genotypic and phenotypic variation from 350 barn swallows sampled across eight populations (28 pairwise comparisons) from four different subspecies. We report a draft whole genome sequence for *H. rustica*, to which we aligned a set of 9,493 single nucleotide polymorphisms (SNPs). Using statistical approaches to control for spatial autocorrelation of phenotypic variables and geographic distance, we find that divergence in traits related to migratory behavior and sexual signaling, as well as geographic distance together, explain over 70% of genome-wide divergence among populations. Controlling for IBD, we find 42% of genome-wide divergence is attributable to IBA through pairwise differences in traits related to migratory behavior and sexual signaling alone. By (i) combining these results with prior studies of how selection shapes morphological differentiation and (ii) accounting for spatial autocorrelation, we infer that morphological adaptation plays a large role in shaping population-level differentiation in this group of closely related populations.

Introduction

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 and 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 and Nosil 2005, Crispo et al 2006, Nosil et al. 2009, Schluter 2001, 2009, Shafer and Wolf 2013); or both (Lande 1980, Shafer and Wolf 2013, Wang and 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 and Summers 2010, Wang et al. 2013, Bradburd et al. 2013, Wang and Bradburd 2014).

In comparison to accumulating evidence for the role of ecological adaptation (one form of IBA) and geographic isolation (IBD) in population divergence and speciation (Shafer and Wolf 2013, Wang and Bradburd 2014), there is little empirical data on the contribution of sexual selection to population genetic divergence (e.g., Parchman et al

90   2013, Baldassarre et al. 2014, Morgans et al. 2014). Adaptation to the local

91   environment may come in the form of sexual selection if competition over mates differs

92   among closely related populations or signal traits reflect variation in the environmental

93   context in which these traits are developed or advertised (Ingleby et al. 2010, Maan and

94   Seehausen 2011, Baldassarre et al. 2013, Safran et al. 2013, Matute, 2014, Seehausen

95   et al. 2014). The lack of such data represents a major gap in our understanding of

96   speciation because divergence in sexual traits is a common form of phenotypic

97   differentiation among populations that may strongly relate to reproductive isolation (e.g.,

98   West-Eberhard 1983, Panhuis et al. 2001, Irwin et al. 2001, Ritchie 2007, Curie Network

99   2012, Safran et al. 2012, Langerhans and Riesch 2013, Seddon et al. 2013). Moreover,

100 previous work has shown that sexual signals can evolve more rapidly than traits related

101 to ecological adaptation (e.g., Kingsolver et al., 2001, Svensson et al. 2006, Kingsolver

102 and Pfennig 2007, Siepielski et al. 2011, Seddon et al. 2013). Collectively, these

103 observations suggest that sexual signals could be an important feature of biological

104 diversification, especially in the early, formative stages of speciation (Kraaijeveld et al.

105 2011). Here, a testable prediction for a role of sexual selection in divergence is that

106 divergence in aspects of phenotype related to mate selection is positively associated

107 with genetic divergence, either due to increased local adaptation, reduced immigration,

108 or both. This is similar to evidence for ecological speciation, where divergence in traits

109 related to ecological adaptation predicts population genetic differentiation (Shafer and

110 Wolf 2013).

111   Thus, in its most general form, IBA predicts associations between genomic

112 divergence and differentiation in traits related to both ecology and mating success. Note

113 that ecological adaptations and mating adaptations are not mutually exclusive, and

114 indeed, may even be related, as sexual signals necessarily evolve in an ecological

115 context (Ingleby et al. 2010, Maan and Seehausen 2011, Baldassarre et al. 2013,

116 Safran et al. 2013, Seehausen et al. 2014). Thus, important questions remain about the

117 relative contributions and likely interactions of sexual selection, natural selection, and

118 geographic distance to genomic differentiation among closely related populations (e.g.,

119 van Doorn et al. 2009, Maan and Seehausen 2011, Wagner et al. 2012, Safran et al.

120 2013). For example, sexual selection has been shown to interact with ecological context

121 in the early stages of population divergence and the formation of premating barriers to

122 reproduction (Scordato et al 2014), yet we know very little about the relative importance

123 of each selective process in the accumulation of biologically relevant genomic

124 differences among populations.

125   Closely related populations with variation in geographic distance are

126 advantageous for identifying spatial and trait-based predictors of genomic divergence.

127 To address the extent to which sexual and natural selection either interact or singly

128 influence population divergence, we analyzed patterns of genomic divergence as a

129 function of variation in aspects of phenotype known to be involved in mate selection and

130 migration, as well as relevant measures of environmental variation (using elevation data

131 and long-term climate databases), in a widespread, phenotypically divergent, yet young

132 group of subspecies: barn swallows (*Hirundo rustica*). Barn swallows include six sub-

133 species worldwide (Figure 1), with populations varying in: (1) geographic distance from

134 one another, (2) trait combinations with known importance in sexual signaling, (3)

135 breeding and non-breeding environments, and (4) migratory behavior (reviewed in

136   Scordato and Safran 2014). A recent phylogenetic reconstruction for this group of

137   closely related sub-species indicates that they are monophyletic with respect to other

138   members of the genus *Hirundo* (Dor et al. 2010). Additionally, recent mtDNA,

139   phylogeographic, and microsatellite analyses suggest this group formed rapidly

140   (between 100,000 and 27,000 years ago; Zink et al. 2006) and is not strongly

141   genetically differentiated, despite marked sexual signal and behavioral differentiation

142   among populations (Dor et al. 2010, Dor et al. 2012). Phenotypic and genomic

143   divergence among barn swallow populations is likely due to a combination of selection

144   and drift in the context of variable degrees of geographic isolation. For example,

145   populations in North America and Eurasia have likely diverged without gene flow,

146   whereas evidence suggests that populations in the Middle East have experienced

147   recent historical or ongoing gene flow with populations in Europe (Dor et al. 2010,

148   2012).

149   Previous research in barn swallows has demonstrated phenotypic divergence in

150   traits important to both sexual selection (morphological and behavioral sexual signals;

151   e.g., Safran and McGraw 2004, Safran et al. 2005, Vortman et al. 2011, Vortman et al.

152   2013, Scordato and Safran 2014) and natural selection (traits related to flight, foraging,

153   and migration; reviewed in Turner 2006). One of six well-characterized sub-species

154   within the larger *H. rustica* complex, the European barn swallow (*H.r. rustica*), has been

155   the subject of intense research activity over the last twenty years, with an emphasis on

156   sexual selection (Turner 2006). This work clearly demonstrates that females mate

157   preferentially with long-tailed males, and their offspring experience advantages over

158   those of shorter-tailed males (e.g., Møller 1994). By contrast, in two North American

159 populations of *H.r. erythrogaster,* males with darker plumage color have greater social

160 and extra-pair mating success (Safran et al. 2005, Safran et al. *in revision*). In the east-

161 Mediterranean distributed *H.r. transitiva,* males with a combination of darker ventral

162 plumage and longer streamer lengths are favored by females through social and extra-

163 pair mating decisions (Vortman et al. 2011, Vortman et al. 2013).

164   Both experimental and observational data suggest that tail streamer length and

165 ventral color are under varying degrees of sexual selection in different populations of

166 barn swallows, and previous work has defined characteristics of wing shape, such as

167 wing length, as being associated with migratory behavior (Von Ronn et al 2016). We

168 therefore use these traits to examine the roles of sexual and natural selection in

169 contributing to genome-wide divergence among populations of barn swallows. Further,

170 previous work has shown that variation in ventral color (Saino et al 2013, Hubbard et al.

171 2015) and tail streamer length have heritable components (Møller 1994), indicating that

172 they are subject to evolutionary change and thus relevant aspects of phenotype to

173 investigate in terms of their influence on genome-wide divergence.

174   Here, we genotyped thousands of SNPs (using genotyping-by-sequencing, GBS)

175 in 354 barn swallows from eight populations, representing four of six subspecies,

176 distributed across the Northern Hemisphere. Additionally, we report a draft whole

177 genome sequence for a male *Hirundo rustica erythrogaster*, which was constructed to

178 ensure that GBS sequences cleanly assembled to a reference and to reduce problems

179 with duplicates/paralogs. We use these population genomic data to address the extent

180 to which population genomic divergence (based on average genome-wide F_{ST}) is

181 associated with divergence in (i) sexual signals, (ii) wing length, (iii) features of climate,

182 and/or (iv) geographic distance. The first three factors all relate to IBA, while the fourth

183 relates to IBD. A specific objective of this study is to partition variance in pairwise

184 genomic divergence as a function of geographic distance (IBD); population differences

185 in ecology, i.e. climate, (IBA, influenced by natural selection); wing morphology related

186 to flight and migratory behavior (IBA, influenced by natural selection); and differences in

187 morphology related to sexual communication (IBA, influenced by sexual selection).

188

189 **METHODS**

190 **Field Data Collection**

191 We sampled individuals from eight locations representing four sub-species across the

192 barn swallow breeding range (Table 1). The following samples or measures were taken

193 from male barn swallows known to be breeding at each site: 1) a sample of ventral

194 feathers for objective color quantification, 2) length of wing and tail streamers (outer

195 rectrices; see Table 2 for a complete description), and 3) blood samples as a DNA

196 source from each bird (approx 50 µl, stored in 2% SDS lysis buffer). See Figure 1 and

197 Table 1 for sampling location and final sample size information.

198

199 **Genomics Methods**

200 ***Genome Assembly***. A draft genome was assembled, to provide a resource for future

201 work and to ensure that regions used for GBS cleanly assembled to a reference. A draft

202 genome assembled with moderate to high coverage enables the identification of single-

203 copy portions of the genome, reduces the challenges associated with distinguishing

204 close paralogs from alleles at the same locus, and (depending on quality and scaffold

205 sizes) can provide context of the genomic location of GBS loci. The DNA for our

206 genome assembly came from a male barn swallow with a well-known reproductive

207 history in our Boulder, Colorado study site (ID 2540-44680). This male was first

208 captured in 2008 and had five successive breeding seasons at the same location (from

209 2008 through 2012). To generate a draft reference genome we obtained sequences

210 from four lanes of Illumina HiSeq platform, two lanes of 101 base-pair reads from two

211 paired-end libraries and two lanes of 101 base reads from a mate-pair library from

212 Macrogen (www.macrogenusa.com). One paired-end library had an insert size of 176

213 bp, while a second paired-end library had an insert size of 454 bp. The mate-pair

214 average insert size was 1,458 bp. We obtained a total of 27 Gbp from paired-end library

215 one, and 17.4 Gbp from library two. From the two lanes of mate pair sequence we

216 obtained 43.4 Gbp. After cleaning to remove low quality reads (lower than a base

217 quality of 20 on either end of the read) and common contaminants, 129.4 million pairs of

218 reads remained in library one, 80.0 million pairs in library two, and 95.8 million pairs in

219 the mate pair library remained, for a total of 61.7 Gb of sequence. Based on an

220 estimated genome size of 1.3 Gb (Andrews et al. 2009) this is an average of 47x

221 coverage.

222 Reads were assembled using SOAPdenovo 2.04, using a K-mer length of 47, an

223 edge coverage cutoff of 3, a k-mer frequency cutoff of 3, and an arcweight filter of 3.

224 Repeats were resolved with reads. Otherwise, parameters were as defaults. After

225 removing short scaffolds (below 1000 bp) we had a total of 100,153 scaffolds, and a

226 total assembly length of 1.1Gb, 85% of the estimated 1.3 Gb genome size (Andrews et

227 al. 2009). The average scaffold length was 11,010 bp, the longest scaffold was 732,517

228   bp, the N50 38,844bp, and N90 was 3,718. Of this 1.1 Gb of assembled sequence, a

229   total of 1.06 Gb could be conservatively mapped to the *Ficedula albicollis* genome

230   (Ellegren et al. 2012), using blastn with a minimum e-value cutoff of 10^{-80} and a

231   minimum sequence similarity of 80%.

232   This alignment of our *Hirundo rustica* genome assembly to the most closely-

233   related, well-annotated genome assembly for Collared Flycatcher *Ficedula albicollis*

234   (Ellegren 2012), shows a high degree of sequence conservation. Over 91% of our

235   single-copy, assembled sequence could be uniquely placed onto the genome. Thus,

236   although we do not have a genetic or physical map placing our sequences onto

237   chromosomes, we can provisionally place virtually all of our assembled genome using

238   sequence similarity to existing avian genome sequences. Given the high degree of

239   synteny found across passerines (Backström et al 2008, Ellegren 2013, Kawakami

240   2014), and among much more distant avian lineages (Derjusheva et al 2004, Zhang

241   2014), taking advantage of existing related genomes enabled us to infer which markers

242   were likely to be linked on chromosomes. We also identified markers likely to be on

243   autosomes and on the sex-determining Z chromosome. As the barn swallow

244   populations we focus on in this paper exhibit only slight genome wide differentiation

245   (see Results), this draft genome is ideal for use as a reference for aligning the GBS

246   SNPs described below.

247   ***Population Genomics.*** We generated DNA sequence data from 354 individual barn

248   swallows sampled across eight populations (see Table 1). We constructed reduced

249   genomic complexity libraries for each individual using a restriction fragment-based

250   procedure (Parchman et al., 2012; Gompert et al., 2012). We first digested genomic

251 DNA with two restriction endonucleases (EcoRI and MseI) and ligated double-stranded

252 adaptor oligonucleotides to the digested fragments. These oligonucleotides consisted of

253 the priming sites for Illumina sequencing, followed by eight, nine, or ten base pair

254 barcode sequences that allow for the unique identification of sequences from each

255 individual. This method has been successfully used to generate population genetic data

256 for a large number of projects (Gompert et al. 2012, 2014b, Parchman et al. 2012, 2013,

257 Nosil et al. 2012b, Mandeville et al. 2015); a full version of the protocol is available at

258 dryad (doi:XXX). We used 354 unique barcoded adaptors, which allowed us to

259 sequence all individuals in one Illumina Hiseq sequencing lane. After the restriction and

260 ligation reactions, we pooled all samples and used standard Illumina PCR primers to

261 amplify the barcode-adapted fragments. We separated the amplification products on a

262 2% agarose gel and excised fragments between approximately 350 and 450 bp in

263 length. We purified these fragments using Qiagen's Qiaquick Gel Extraction Kit (Qiagen

264 Inc.). Concentration and quality of the pooled library was evaluated on an Agilent

265 BioAnalyzer qPCR. Sequencing was performed by the National Center for Genome

266 Research (Santa Fe, NM, USA) with 100 base pair single end reads, and yielding

267 98,401,301 reads following standard contaminant filtering.

268   We used a custom Perl script (dryad doi:XXX) to remove bases associated with

269 the barcode and EcoRI cut site for all 98,401,301 sequences and to replace the

270 sequence IDs with individual IDs for each DNA sample. This script also corrected

271 barcodes with potential single or double base mismatches due to sequencing or

272 oligonucleotide synthesis errors. Sequences were then split by individual for further

273 analysis. Four of the individuals had fewer than 2,000 reads and were disregarded,

274   which left 350 individuals for further analysis. We used *bwa* v 0.7.12 (Burrows-Wheeler

275   Aligner; Li & Durbin 2009) to assemble data for each individual against the barn swallow

276   genome assembly, with an edit distance of 4 and the remaining parameters set as

277   default. This edit distance ensures that, even with 1% sequencing errors typical of

278   Illumina sequencing, distant alleles from divergent subspecies will map to the genome,

279   while nevertheless preventing non-specific alignments. Across all individuals an

280   average of 61.66% of reads assembled, with 97.5% of individuals having 57.9% or

281   greater percentage of reads assembling. We used *samtools* v 1.2 and *bcftools* v 1.2 (Li

282   et al., 2009) to identify variant sites in the assembled sequences and obtain genotype

283   likelihoods for variable sites. To identify variants, we required data for 60% or more

284   individuals. This resulted in 67,773 single nucleotide variants. A complete list of

285   parameters used for assembly and variant calling are available from the authors by

286   request.

287   We used point estimates of allele frequencies in all 350 individuals to separate

288   this set into 22,328 common variants (minimum minor allele frequency of 5% or greater)

289   that were used for further analysis and 45,445 rare variants that were disregarded. We

290   focused on comparatively common variants because they are more likely to occur in

291   more than one location and contain information about population histories (Gompert et

292   al. 2012). To avoid analyzing SNPs with highly correlated allele frequencies, we

293   randomly sampled a single SNP from each of the 100 bp regions in the assemblies

294   (GBS loci are 100 bases in length). For the resulting final set of 9493 single nucleotide

295   polymorphisms (SNPs), the average coverage depth per individual per site was 1.5×

296   (0.09-3.57, 2.5% and 97.5% quantiles). This kind of low coverage genomic data is

297   appropriate for population-level inferences when analyzed with models that incorporate

298   uncertainty arising from variability in sequencing coverage across individuals and loci

299   (e.g., Neilsen et al. 2011, Gompert et al. 2012, Gompert and Buerkle 2014).

300   ***Population genetic analyses****.* We used a Bayesian model to estimate population allele

301   frequencies for each of 9,493 variable nucleotides based on the point estimates of

302   genotypes from *bcftools* (described in Gompert et al., 2012). Genotypes and the

303   population allele frequencies were treated as unknown model parameters, and

304   genotype probabilities and allele frequencies were simultaneously estimated for each

305   sampling locality (separately from estimates for other localities). Importantly, this model

306   incorporates uncertainty from stochastic variation in sequencing coverage depth across

307   individuals and loci into the estimation process, and provides a sound approach for

308   estimating population level parameters for low coverage sequencing approaches.

309   Population allele frequencies serve as prior information, and genotype probabilities are

310   inferred for each locus in each individual. We obtained posterior probabilities for

311   parameters using Markov chain Monte Carlo (MCMC). Each analysis consisted of a

312   single chain iterated for 2,000 steps following a 1,000 step burn-in, with samples

313   retained every other step, yielding 1,000 samples from the posterior distributions. For

314   these data and this simple model, mixing and convergence were clearly evident in plots

315   of MCMC histories.

316   We obtained the mean genotype (scale of 0 to 2) from the posterior distributions

317   for each SNP in each individual. Furthermore, we calculated the mean genotype across

318   all individuals to center the genotypes at each locus before calculating the genetic

319   covariance between the genotype vectors for all pairs of individuals. We summarized

320   the genetic covariance matrix using principal component analysis (PCA). Because the

321   PCA will reflect all genetic covariances and will be affected by uneven sampling (e.g.,

322   144 samples from Colorado and 16 from Romania), we performed a second PCA in

323   which we randomly down-sampled all localities to the smallest sample size from any

324   locality (N=16). We also obtained mean allele frequencies from the posterior

325   distributions and transformed these to estimates of mean F_{ST} (Hudson 1992) and Nei's

326   D (Nei et al. 1983; Takezaki & Nei 1996) between all pairs of populations. F_{ST}

327   and Nei's D were highly correlated (r=0.996), so we only present results for F_{ST}. We

328   performed the PCA and calculated F_{ST} and Nei's D in R (using *prcomp* and custom

329   functions, R Core Team, 2015).

330

331   **Phenotypic and Environmental Variables**

332   Our goal was to compare divergence among traits known to be related to sexual

333   signaling (plumage color and streamer lengths) and migratory behavior (i.e., wing

334   length). We also compared aspects of environmental variability known to affect the

335   aerial insect populations that barn swallows prey upon. These variables include

336   elevation and several measures of temperature variation (e.g., Thomsen et al 2016).

337

338   ***Quantifying Feather Color.*** We collected 5-10 feathers from the throat and breast

339   (upper ventral region) and vent (lower ventral region, below the attachment of tail

340   streamers) and stored them in small envelopes in a dark, dry environment prior to

341   measurement (*following* Safran and McGraw 2004). We assessed the color of these

342   samples by measuring plumage brightness using an Ocean Optics USB4000

343   spectrometer (Dunedin, FL). Reflectance data were generated relative to a white

344   standard (Ocean Optics WS-1) and a dark standard (all light excluded), and spectra

345   were recorded with the SpectraSuite software package (version 2.0.125, Ocean Optics

346   Inc.). We find no evidence of UV reflectance in the ventral plumage color of barn

347   swallows (Safran and McGraw 2004) and thus used three traditional axes of color for

348   objective measurement of color variation. We used average brightness, which was

349   calculated from three separate measurements of the collected throat and breast

350   feathers, as a representative metric of overall ventral plumage color. Average

351   brightness is a good color metric, as all three traditional axes of color (hue, chroma, and

352   brightness) were previously found to be highly correlated across the ventral region of

353   individual barn swallows (McGraw et al. 2005, Safran and McGraw 2004, Hubbard

354   unpublished data), and brightness is the most variable dimension of color in this region

355   (Hubbard unpublished data). Lower brightness scores (% reflectance) indicate plumage

356   color that appears darker, redder, and more saturated when compared to feathers with

357   higher brightness scores.

358

359   ***Estimates of Phenotypic Divergence***. To determine pair-wise distance in phenotypes

360   among environmental and morphological traits, we used an unbiased effect size statistic

361   (ΔP; Safran et al. 2012) to calculate trait distance for each trait in pair-wise comparisons

362   among the 8 sampled populations. ΔP is calculated based upon a joint cumulative

363   distribution function (CDF) from all populations in the data set. Distances were

364   calculated for each pairwise comparison using the population median percentile in the

365   overall CDF. ΔP was developed specifically to analyze phenotypic distance among

366 closely related populations, as it easily accommodates simultaneous comparisons of

367 any number of traits across any number of populations, and is relatively insensitive to

368 unequal variances and sample sizes among populations (Safran et al. 2012). For all

369 analyses, we use the absolute value of pairwise distances.

370

371 ***Climate and Elevation Data***. Temperature and elevation data were obtained using the

372 CRUTEM database maintained by the Climatic Research Unit and available at

373 http://www.cru.uea.ac.uk/cru/data/crutem/ge/. Using this database, we downloaded the

374 last 50 years of temperature data from the 3 weather stations closest to each of our

375 sampling sites (see supplementary materials). From these data, we derived the mean,

376 minimum, and coefficient of variation in temperature during the breeding season.

377 Breeding seasons vary among our populations, and, accordingly, we used the following

378 months for climate data collection for each site: Colorado, USA: April-September; New

379 York, USA: May-August; UK: April-August; Israel, January-April; and Czech Republic,

380 Romania, Turkey and Taiwan, April- July. These measurements have been routinely

381 employed in studies of other avian taxa with widespread geographic ranges

382 (Rubenstein and Lovette 2007, Rubenstein 2007, Botero et al. 2009, Botero et al.

383 2014).

384

385 ***Estimates of Geographic Distance.*** Geographic distances between study sites were

386 calculated in the R package geosphere using the Haversine great circle distance

387 between points. This is the shortest "as the crow flies" distance, assuming a spherical

388 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 distance, we analyzed 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 distance, 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 = 10,000) 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 analyze 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 and Wolf 2013, Wang and Bradford 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,

412   depending on the analysis. For Mantel tests, we employed matrices of pairwise

413   differences for each population. For MRM and the variance partitioning analyses, these

414   matrices were transformed into vectors of pairwise differences for each variable and

415   each pair of populations. When the response variable is a single vector (as here, using

416   pairwise differences in mean F_{ST} between populations), variance partitioning is done by

417   partial regression.

418

Results

Population Genomics

421   Principal component analysis of genetic covariances between an even sample of

422   individuals from each population (n = 16) revealed clear genomic differences between

423   some localities, but greater genomic similarity among the nearby sampling locations in

424   Czech Republic, Romania, Turkey, and Israel (Figure 2). The first two principal

425   components explained 52% of the genomic variation among individuals in each locality

426   (Figure 2). The separation of populations on PC1 is consistent with the phylogenetic

427   hypothesis (Dor et al. 2010) that separates eastern and western barn swallows, with

428   samples from Israel and Europe recoverable as a distinct lineage from samples in North

429   America (Colorado and New York) and southern Asia (Taiwan). Interestingly, while

430   Israel is considered a distinct subspecies based on phenotype, it cannot be

431   differentiated from European populations based on genomic covariance. PC2 explains a

432   relatively small amount (10.3%) of genetic variation, compared to PC1 (41.9%), and

433   further separates subspecies within the eastern and western clades. The PCA based on

434   all 350 individuals gave considerable weight in the first axis to distinction between the

435   large sample from North America (N=144 from Colorado and 27 from New York) and

436   samples from elsewhere, with PC1 accounting for 67.9% of the genetic variation.

437   Pairwise F_{ST} ranged from 0.024 to 0.073 (Table 3).

438

439   **Phenotypic and Environmental Variables**

440   ***Phenotypic and environmental divergence.*** Populations differed both in the extent of

441   phenotypic variation and environmental context (Tables 4 and 5). Mean percentiles

442   within a cumulative frequency distribution varied considerably among populations; the

443   distribution of variables measured from each population differed in terms of its

444   placement on the overall cumulative frequency distribution (Table 4). For example,

445   populations in the UK have the longest wing length, whereas populations in Taiwan

446   have the shortest. Populations in North America (Colorado and Ithaca) have the lightest

447   throat color yet the darkest breast color; thus, there is variation in the extent and

448   direction in which color patches differ among sub-species. The spread of percentile

449   values is indicative of the degree of pairwise divergence for each trait or climate

450   variable. For example, of all phenotypic traits, populations exhibited the most extreme

451   differences in wing length, with a range of percentiles from 9.02–91.16. Note that the

452   percentile values for the phenotypic measures are based on multiple individuals within

453   populations (Table 1), whereas the percentile values for the environmental measures

454   are based on one measure within each population. Analyses of variance reveal that all

455   of the phenotypic traits we measured in this study showed statistically significant

456   differences among populations (adjusted R-squared values range from 0.25 to 0.52;

457   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

ΔP) for various features of phenotype, including measures of ventral color, tail and wing

length, were positively correlated with genome-wide divergence among pairs of eight

geographically isolated populations (Mantel tests; Figure 4, 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, Table 6).

Geographic Distance, Environment, Phenotype. Several features of phenotypic

divergence also co-varied with geographic distance (Mantel tests; Figure 6, Table 6),

whereas environmental and climate features did not (Mantel tests; Figure 7, Table 6).

Thus, genome-wide divergence among closely related populations was associated with

both geographic and phenotypic trait distance, and the two are sometimes correlated

with each other. It is therefore necessary to adequately control for correlations between

phenotypic and geographic distance 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

481   regression (Wang 2013). We started with a maximal model that included the pairwise

482   distance matrix of F_{ST} values as the response variable and distance matrices (based on

483   ΔP values) for all nine phenotypic and ecological variables as predictors. We then used

484   backwards-stepwise model selection, sequentially deleting the least significant term and

485   rerunning the model, until coefficients for all remaining predictor variables were

486   significantly different from zero. The final multiple matrix regression model included two

487   aspects of phenotype that explained significant variation in pairwise F_{ST} while controlling

488   for spatial autocorrelation: wing length and throat color (full model $F_{2,17}$= 24.95, P =

489   0.002, r-squared = 0.757; parameter estimate and t-test: wing length coefficient = 0.82,

490   T=6.12, P = 0.002, throat color coefficient = 0.22, T = 1.79, P = 0.03, geographic

491   distance coefficient = 0.01, T = 0.09, P = 0.94).

492   We used the significant predictor variables from the multiple matrix regression in

493   variance partitioning and redundancy analyses to further examine the relative

494   contributions of wing length, throat color, and geographic distance in explaining

495   variation in genome-wide divergence among the populations in our data set. The

496   variance partitioning approach enabled us to test hypotheses about the relative

497   significance of IBA and IBD, taking into account correlations among phenotypic,

498   environmental, and geographic distance variables.

499   The best model fit from the variance partitioning analyses (Table 7) included the

500   effects of all three matrices (geographic distance, throat color distance, and wing length

501   distance). These models, in which various combinations of matrices are conditioned

502   upon one another, demonstrated that collectively 73% of genome-wide divergence is

503   attributable to three variables: wing length, throat coloration, and geographic distance

504   between populations (Figure 8). Further analyses enabled us to analyze the association

505   between each matrix (geographic distance, throat color, wing length) separately by

506   conditioning each variable on the others in all possible combinations (Table 7). For

507   example, when the matrix containing pairwise distance in wing length is conditioned on

508   the matrices containing pairwise differences in geographic distance and throat color, the

509   influence of wing length on its own accounted for 42% of pairwise genomic distance

510   among the populations in our sample (Figure 8, Table 7). When the matrix containing

511   pairwise distance in throat color is conditioned upon the matrices containing pairwise

512   differences in geographic distance and wing length, the influence of throat color on its

513   own explained 5% of pairwise genomic distance among the populations in our sample.

514   Finally, when conditioned upon the wing and color matrices, the effect of geographic

515   distance on its own did not explain additional variation in genomic distance among the

516   populations in our sample (Figure 8, Table 7). Interestingly, the only place where the

517   geographic distance matrix explained a significant amount of variation in genomic

518   divergence is when both geographic distance and wing length were considered side by

519   side and conditioned upon their correlation with the color matrix (Figure 8, Table 7).

520

Discussion

522   Geographic distance and phenotypic distance are strongly correlated among our study

523   populations. Thus, we applied two complementary methods of variance partitioning—

524   multiple matrix regression, and constrained redundancy analysis—which enabled us to

525   analyze the relative contributions of correlated matrices (geographic distance and

526   phenotypic distance) to genome-wide divergence. Overall, our results demonstrate clear

527   evidence that both IBA and IBD contribute to genome-wide divergence among these

528   closely related populations of barn swallows. When spatial autocorrelation between

529   phenotype and geographic distance are accounted for, our results suggest that

530   divergence in an ecological trait (wing length) and a sexual signaling trait (throat color)

531   play a larger role in population genetic divergence than does geographic distance.

532   Surprisingly, we found no evidence that elevation and temperature differences

533   were influential to genome-wide divergence. Although these features of the

534   environment, which are relevant for obligate aerial insectivores, did vary spatially, they

535   were neither associated with geographic distance nor genome-wide distance among the

536   populations in our sample. In a further investigation, we also found no evidence that

537   maximal temperatures or precipitation patterns at each location differed as a function of

538   geographic distance or were associated with genome-wide divergence (Safran,

539   unpublished data). Given the span of our sampling locations, ranging from Israel to

540   North America, these results either suggest that barn swallows occupy fairly similar

541   environments with respect to elevation and temperature or are not particularly sensitive

542   to these ecological variables. The latter explanation seems most likely as these

543   populations are very cosmopolitan in distribution both during the breeding and non-

544   breeding season and during long migratory trips where they likely inhabit a wide range

545   of environments.

546   ***Population Genomic Structure and Phenotypic Divergence.*** Our data demonstrate

547   genetic similarity of individuals within sampling localities, with greater differences

548   between populations that are separated by large geographic distances and genome-

549   wide clustering that generally corresponds to named subspecies. Principal component

550   analysis shows that the most genetically different populations along PC1 correspond to

551   a highly supported east – west split in the current phylogenetic hypothesis for barn

552   swallows (Zink et al. 2006, Dor et al. 2010). Populations in Asia (*H.r. gutturalis*) and

553   North America (*H.r. erythrogaster*) are more closely related to one another than either

554   are to populations in Europe (*H.r. rustica*) and the Middle East (*H.r transitiva*). PC2

555   further separates *H.r. gutturalis* (TW) from two samples from the North American *H.r.*

556   *erythrogaster* populations (IA and CO), and the UK population from other mainland

557   populations of *H.r. rustica* (TR, CR and RM). Samples from the Israeli subspecies (IL)

558   are clustered closely with other mainland European populations of barn swallows,

559   consistent with an unresolved relationship between this subspecies and *H.r. rustica* in

560   the current phylogeny (Dor et al. 2010). Despite being relatively genetically similar,

561   individuals from Israel and the continental European populations are fairly divergent in

562   phenotype, particularly with respect to ventral color.

563   Despite shallow genomic divergence, phenotypic differentiation is apparent in all

564   aspects of morphology we analyzed in this study, ranging from traits related to body

565   size and flight (wing and tail length) to ventral color. Phenotypic variation despite

566   shallow genomic divergence appears common among many taxa, including cichlid

567   fishes (e.g., Wagner et al. 2012) and particularly in birds (e.g. Parchman et al. 2006,

568   Rodrigues et al. 2014, Poelstra et al. 2014, Mason and Taylor 2015), where it is often

569   the case that a few genes are implicated in morphological variation against a fairly

570   homogenous genomic background (e.g., Poelstra et al. 2014, Kardos et al. 2015). In

571   other words, although a larger number of genes may be involved in generating plumage

572   coloration, studies to date have suggested that a large proportion of segregating trait

573 variation is due to variation in a small number of genetic loci. Collectively, these studies

574 suggest an important role of divergent selection on signaling traits in population

575 differentiation (Wagner et al. 2012, Poelstra et al. 2014), which might play a particularly

576 important role during the earliest stages of speciation (Kraaijeveld et al. 2010). Finer-

577 scale genomic analyses in a broader geographic context will enable us to test whether

578 phenotypic differentiation is a barrier to gene flow when different subspecies of barn

579 swallows are in secondary contact.

580 ***Isolation by Adaptation, Controlling for Isolation by Distance.*** Because trait

581 divergence is correlated with geographic distance among populations, we applied two

582 complementary statistical methods to tease apart the relative significance of geographic

583 and phenotypic distance in explaining genome-wide divergence. Both sets of analyses

584 reveal that wing length and, to a lesser extent, throat color, are most strongly associated

585 with genome-wide divergence, when geographic distance among populations is

586 accounted for statistically (Figure 8). The variance partitioning model enables us to

587 directly quantify the influence of correlations among geographic distance, wing length,

588 and throat color and to partition the contribution of each of these variables towards

589 differences in genome-wide divergence (Figure 8, Table 6). Thus, our study reveals an

590 important role of phenotypic divergence, and supports a model of IBA for explaining

591 genomic differentiation of geographically isolated populations, bearing in mind that

592 phenotypic divergence is also strongly correlated with geographic distance, and a model

593 of IBD, among populations (Figure 8).

594 IBD is one of the most common patterns in population genetic data (e.g., Jenkins

595 et al. 2010). Evidence is accumulating to suggest that IBA (which includes models of

596   Isolation by Environment, or IBE) is also well supported in a variety of empirical systems

597   among closely related populations (e.g., Edelaar et al. 2012, Lasky et al 2012, Lee and

598   Mitchell-Olds 2012, Shafer and Wolf 2013, Morgans et al. 2014, Wang and Bradburd

599   2014). A recent meta-analysis further revealed that the effect of geographic distance or

600   spatial autocorrelation (between phenotypes, environmental variables, and geographic

601   distance) is critical to control for in statistical tests of IBA but has rarely been done in

602   previous studies (Shafer and Wolf 2013). Simulation studies reveal that a failure to

603   account for spatial autocorrelation can lead to biased results (Shafer and Wolf 2013). In

604   our own study, separate analyses of IBA and IBD each revealed significant effects on

605   genome-wide divergence. Analyses that explicitly considered the correlations of

606   phenotypic and geographic distance were critical to separating out the relative

607   significance of each factor. Variance partitioning (Legendre and Fortin 2010) and

608   multiple matrix regression (Wang 2013) are excellent methods for dealing with spatial

609   autocorrelation, analyzing the relative significance of IBA and IBD, and are particularly

610   important alternatives to partial Mantel tests (Bradburd et al 2013, Wang 2013, Wang et

611   al. 2013), which are subject to false positive results (Diniz-Filho et al 2013).

612   ***Barn swallows and divergent selection.*** Our results support an important role of IBA

613   in genome-wide divergence. In particular, differences in wing length, the most divergent

614   phenotypic trait among populations, explained a significant amount of variation in

615   among-population genomic divergence; throat color is also associated with genomic

616   divergence, but to a lesser extent. From these results and our understanding of the

617   function of these traits, we infer an influence of both natural selection and sexual

618   selection in genomic divergence among these eight populations of barn swallows.

619   These results are intuitive for several reasons. First, wing length and shape are traits

620   associated with migratory behavior (Marchetti et al 1995, Lockwood et al 1998). Barn

621   swallows vary in migratory distance, and there is evidence of migratory divides

622   throughout their range (e.g., Irwin and Irwin 2005, von Rönn et al. 2016). All four

623   representatives of the *H.r. rustica* subspecies (UK, Czech Republic, Romania, and

624   Turkey) have the longest wing lengths among populations sampled, whereas individuals

625   from distantly-related Taiwan (*H.r. gutturalis*) have the shortest. Interestingly, the sole

626   non-migratory population in our study, Israel, has intermediate wing lengths that overlap

627   in variation with individuals from Turkey and Romania and is genomically

628   indistinguishable from *H. r. rustica*, based on pairwise F_{ST}. In other migratory bird

629   populations, variation in migratory behavior can influence both the evolution of wing

630   shape and genomic differentiation (e.g., Ruegg 2007, 2008, Delmore et al. 2012, 2014,

631   2015, Rolshausen et al 2009, von Rönn et al. 2016). Thus, variation in migratory routes

632   and behaviors (e.g., timing of arrival to breeding grounds, which may affect mate

633   selection) is proposed as a potentially influential contributor to population divergence

634   (Irwin and Irwin 2005, Rolshausen et al 2009). Further study is required to better

635   understand the role of migratory behavior in divergence among barn swallow

636   populations, as well as the extent to which these seasonal movement patterns influence

637   trait evolution.

638   Our analyses of IBA also revealed that population-level differences in color

639   contribute to genomic differentiation. Plumage color in various ventral regions (throat,

640   breast, vent) varies both within and among closely related populations of barn swallows

641   (Scordato and Safran 2014). Both throat and breast coloration are variable among the

642   populations in this study, but in different ways. For example, males in our North

643   American populations (Colorado, New York) have the darkest breast color, yet the

644   lightest throat color compared to other populations. Ventral coloration has been the

645   focus of several correlational and experimental studies in barn swallows. Together,

646   these studies indicate that melanin-based ventral color varies among populations

647   (Scordato and Safran 2014), is heritable (Hubbard et al. 2015), and relates to social

648   mate selection and paternity allocation (Safran and McGraw 2004, Vortman et al 2011).

649   Thus, patterns of ventral color differentiation are likely under divergent sexual selection,

650   although this hypothesis requires further testing.

651   Patterns of IBA may be underlain by divergent selection on traits that directly

652   impede gene flow or through the build up of an association between genome-wide

653   divergence and trait divergence in allopatric populations. Whereas models of IBA have

654   typically been applied to scenarios where gene flow is possible between populations,

655   most of the eight populations in our study are geographically isolated from one another,

656   with the exception of those within continental Europe (Czech Republic and Romania). It

657   is possible that gene flow occurs among populations sampled from western and central

658   Europe and Israel, though a formal analysis is needed to make direct inferences about

659   recent historical or ongoing gene flow in these regions. Nevertheless, divergence in the

660   absence of gene flow is an important aspect of avian diversification (e.g., Price 2008)

661   and thus it is highly relevant to ask questions about associations between phenotypic

662   and genomic divergence in order to better understand patterns and consequences of

663   trait evolution in isolated, yet closely related populations. Still, ongoing studies of IBA in

664   locations where divergent populations come into sympatry will be critical for inferring

665   whether trait divergence that may have evolved in isolation imposes a barrier to gene

666   flow.

667

Acknowledgements

679   References

680   Andrews CB, Mackenzie SA, and Gregory TR (2009) Genome size and wing parameters in
681   passerine birds. *Proceedings of the Royal Society of London B,* **276**:55-61.

682   Backström N, Karaiskou N, Leder EH, Gustafsson L, Primmer CR, Qvarnström A, Ellegren H
683   (2008) A gene-based genetic linkage map of the collared flycatcher (Ficedula albicollis) reveals
684   extensive synteny and gene-order conservation during 100 million years of avian
685   evolution. *Genetics* **179**: 1479-1495

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

689   Baldassarre DT, White TA, Karubian J, Webster MS (2014). Genomic and morphological
690   analysis of a semipermeable avian hybrid zone suggests asymmetrical introgression of a sexual
691   signal. *Evolution*, **68**: 2644–2657.

692   Botero CA, Boogert N, Lovette IJ, Vehrencamp SL. (2009) Climatic patterns predict the
693   elaboration of song displays in mockingbirds. *Current Biology* **19**:1-5.

694   Botero CA, Dor R, McCain CM, Safran RJ (2014) Environmental harshness is positively
695   correlated with intraspecific divergence in mammals and birds. *Molecular Ecology* **23**: 259-268.

696   Bradburd GS, Ralph PL, Coop GM (2013) Disentangling the effects of geographic and
697   ecological isolation on genetic differentiation*. Evolution*, **67**, 3258–3273.

698   Buerkle CA, Gompert Z (2013) Population genomics based on low coverage sequencing: how
699   low should we go? *Molecular Ecology* **22**: 3028–3035

700   Coyne JA and Orr HA (2004) Speciation. Sunderland: Sinauer Associates, Inc.

701   Crispo E, Bentzen P, Reznick DN, Kinnison MT, Hendry AP (2006) The relative influence of
702   natural selection and geography on gene flow in guppies. *Molecular Ecology*, **15**, 49–62.

703   Curie Network (2012). What do we need to know about speciation? *Trends in Ecology &*
704   *Evolution*, **27**, 27-39.

705   Delmore KE, Fox JW, Irwin DE. (2012) Dramatic intraspecific differences in migratory routes,
706   stopover and wintering sites revealed using light-level geolocators. *Proceedings of the Royal*
707   *Society B: Biological Sciences* **279**: 4582-4589

708   Delmore, KE, Hübner S, Kane NC, Schuster R, Andrew RL, Câmara F, Guigo R, Irwin DE
709   (2015) Genomic analysis of a migratory divide reveals candidate genes for migration and
710   implicates selective sweeps in generating islands of differentiation. *Molecular Ecology* **24**: 1873-
711   1888.

712   Delmore KE, Irwin DE (2014) Hybrid songbirds employ intermediate routes in a migratory
713   divide. *Ecology Letters* **17**: 1211-121

714   Derjusheva, Svetlana, et al. (2004) High chromosome conservation detected by comparative
715   chromosome painting in chicken, pigeon and passerine birds. *Chromosome Research* **12.7** 715-
716   723

717   Diniz-Filho, JAF, Soares TN, Lima JS, Dobrovolski R, Landeiro VL, de Campos Telles MP,
718   Rangel T, Bini L.M. (2013) Mantel test in population genetics *Genet Mol Biol*. **36**: 475–485.

719   Dor R, Safran RJ, Sheldon FH, Winkler DW, Lovette IJ (2010) Phylogeny of the genus *Hirundo*
720   and the Barn Swallow subspecies complex. *Molecular Phylogenetics and Evolution* **56**: 409-418

721 Dor R, Safran RJ, Vortman Y, Lotem A, McGowen A, Evans MR, Lovette IJ (2012) Population
722 genetics and morphological comparisons of migratory European *(Hirundo rustica rustica*) and
723 sedentary East-Mediterranean *(H. r. transitiva*) Barn Swallows. *Journal of Heredity*. **103**: 55-63.

724 Edelaar P, Alonso D, Lagerveld S, Senar JC, Bjorklund M € (2012) Population differentiation
725 and restricted gene flow in Spanish crossbills: not isolation-by-distance but isolation-by ecology.
726 *Journal of Evolutionary Biology*, **25**, 417–430.

727 Ellegren H1, Smeds L, Burri R, Olason PI, Backström N, Kawakami T, Künstner A, Mäkinen
728 H, Nadachowska-Brzyska K, Qvarnström A, Uebbing S, Wolf JB. (2012) The genomic
729 landscape of species divergence in *Ficedula* flycatchers. *Nature.* **491**(7426):756-60. doi:
730 10.1038/nature11584.

731 Ellegren H. (2013). The evolutionary genomics of birds. *Annual Review of Ecology, Evolution,*
732 *and Systematics*, **44,** 239-259.

733 Fisher RA (1930) The Genetical Theory of Natural Selection, Clarendon Press, Oxford

734 Gompert Z, Lucas LK, Nice CC, Fordyce JA, Forister ML, A BC (2012) Genomic regions with a
735 history of divergent selection affect fitness of hybrids between two butterfly species. *Evolution*,
736 **66**, 2167–2181.

737 Gompert Z, Lucas LK, Buerkle CA, Forister ML, Fordyce JA, Nice CC (2014) Admixture and the
738 organization of genetic diversity in a butterfly species complex revealed through common and
739 rare genetic variants. *Molecular Ecology*, **23**, 4555–4573.

740 Hubbard JK, Jenkins BR, Safran RJ. (2015) Quantitative genetics of plumage color: lifetime
741 effects of early nest environment on a colorful sexual signal. *Ecology and Evolution*. **5**: 3436-
742 3449

743 Hudson RR, Slatkin M, Maddison WP (1992) Estimation of levels of gene flow from DNA
744 sequence data. *Genetics* **132**, 583–589

745 Ingleby FC, Hunt J, Hosken DJ (2010) The role of genotype by-environment interactions in
746 sexual selection. *Journal of Evolutionary Biology*, **23**, 2031–2045.

747 Irwin, DE, Bensch S, Price TD (2001) Speciation in a ring. *Nature* **409**: 333-337

748 Irwin, DE, Irwin JH (2005) Siberian migratory divides: the role of seasonal migration in
749 speciation. Pages 27-40 in *Birds of Two Worlds: The Ecology and Evolution of Migration*, edited
750 by R. Greenberg and P. P. Marra. Johns Hopkins University Press

751 Jenkins, DG, Carey, M. Czerniewska, J, Fletcher, J, Hether T, Jones A, Knight S, Knox J, Long
752 T, Mannino M, McGuire M, Riffle A, Segelsky S, Shappell L, Sterner A, Strickler T and Turs R.
753 (2010) A meta-analysis of isolation by distance: relic or reference standard for landscape
754 genetics? *Ecography* **33**: 315-320.

755 Kardos M, Husby A, McFarlane E, Qvanstrom A, Ellegren H. (2015) Whole-genome
756 resequencing of extreme phenotypes in collared flycatchers highlights the difficulty of detecting
757 quantitative trait loci in natural populations *Molecular Ecology Resources* doi: 10.1111/1755-
758 0998.12498
759
760 Kingsolver JG, Hoekstra HE, Hoekstra JM, Berrigan D, Vignieri SN, Hill CE, Hoang A, Gibert P,
761 Beerli P (2001) The strength of phenotypic selection in natural populations. *The American*
762 *Naturalist* **157**:245-261.

763 Kingsolver JG, Pfennig DW (2007) Patterns and power of phenotypic selection in nature.
764 *Bioscience* **57**:561-571.

Kraaijeveld K, Kraaijeveld-Smit FJL, Maan ME (2010). Sexual selection and speciation: the comparative evidence revisited. *Biological Reviews*, **86**, 367–377.

Kawakami, T , Smeds, L , Backström, N , Husby, A , Qvarnström, A , Mugal, C F & Ellegren, H (2014) A high-density linkage map enables a second-generation collared flycatcher genome assembly and reveals the patterns of avian recombination rate variation and chromosomal evolution *Molecular ecology*, 23, 4035-4058

Lande, R (1980) Sexual Dimorphism, Sexual Selection, and Adaptation in Polygenic Characters *Evolution* **34***,* 292-305

Langerhans BR, Riesch R (2013) Speciation by selection: a framework for understanding ecology's role in speciation Current Zoology **59**, 31-52

Lasky JR, Des Marais DL, McKay JK et al (2012) Characterizing genomic variation of Arabidopsis thaliana: the roles of geography and climate. *Molecular Ecology*, **21**, 5512–5529

Lee C-R, Mitchell-Olds T (2011) Quantifying effects of environmental and geographical factors on patterns of genetic differentiation. *Molecular Ecology*, **20**, 4631–4642

Legendre and Fortin (2010) Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. *Molecular Ecology Resources* **10**: 831–844 doi: 10 1111/j 1755-0998 2010 02866 x

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, et al (2009) The Sequence Alignment/Map format and SAMtools *Bioinformatics*, **25**, 2078–2079

Lockwood R, Swaddle JP, Rayner, JM (1998) Avian wingtip shape reconsidered: wingtip shape indices and morphological adaptations to migration *Journal of Avian Biology* **29**, 273-292

Maan, M E and Seehausen, O (2011) Ecology, sexual selection and speciation *Ecology Letters* **14**, 591–602

Marchetti, K, Price TD, Richman, A (1995) Correlates of wing morphology with foraging behaviour and migration distance in the genus Phylloscopus *Journal of Avian Biology* **26**, 177-181

Mandeville, E G, Parchman, T L, McDonald, D B and Buerkle, C A, 2015. Highly variable reproductive isolation among pairs of Catostomus species. *Molecular ecology*, **24**, 1856-1872.

Mason NA, Taylor SA (2015) Differentially expressed genes match morphology and plumage despite largely homogeneous genomes in a Holarctic songbird *Molecular Ecology* **24**: 3009 - 3025

Matute DR (2014) The magnitude of behavioral isolation in *Drosophila* is affected by characteristics of the mating community. *Ecology and Evolution.* **4**: 2945–2956

Mayr E (1963) Animal Species and Evolution. Harvard University Press, Cambridge, MA

Møller, AP (1994) Sexual Selection and the Barn Swallow. Oxford: Oxford University Press

Morgans, CL, Cooke GM and TJ Ord (2014) How populations differentiate despite gene flow: sexual and natural selection drive phenotypic divergence within a land fish, the Pacific leaping blenny *BMC Evolutionary Biology* **14**:97

Nei M, Maruyama T, Wu C (1983) Models of evolution of reproductive isolation. *Genetics*, **103**, 557–579

Nielsen R, Paul JS, Albrechtsen A, Song YS. (2011) Genotype and SNP calling from next-generation sequencing data." *Nature Reviews Genetics*, **12**, 443-451

Nosil, P (2012) *Ecological speciation* Oxford University Press, Oxford

Nosil, P, Gompert,Z. Farkas TA, Comeault AA, Feder JL, Buerkle CA, Parchman TL (2012b) Genomic consequences of multiple speciation processes in a stick insect *Proceedings of the Royal Society of London B* **279**: 5058-5065

Nosil P, Vines TH, Funk DJ (2005) Reproductive isolation caused by natural selection against immigrants from divergent habitats *Evolution*, **59**, 705–719

Nosil P, Egan SP, Funk DJ (2008) Heterogeneous genomic differentiation between walking-stick ecotypes: "isolation by adaptation" and multiple roles for divergent selection. *Evolution*, **62**, 316–336

Oksanen J F, Blanchet G Kindt R, Legendre P, Minchin PR , O'Hara R B , et al (2015) vegan: Community Ecology Package R Package Version 2 2-1 Available online at: http://CRAN R-project org/package=vegan

Panhuis TM, Butlin R, Zuk MR, Tregenza T (2001) Sexual selection and speciation *Trends in Ecology and Evolution*, **16**, 364-371

Parchman TL, Gompert Z, Mudge J et al (2012) Genome-wide association genetics of an adaptive trait in lodgepole pine *Molecular Ecology*, **21**, 2991–3005

Parchman, TL , Z Gompert, G Zhang, M J Braun, R Brumfield, D B McDonald, J A C Uy, E D Jarvis, B A Schlinger, and C A Buerkle 2013 The genomic consequences of adaptive divergence and reproductive isolation between species of manakins *Molecular Ecology* **22**:3304-3317

Parchman, TL, Benkman C, Britch SC (2006) Patterns of genetic variation in the adaptive radiation of North American crossbills (Aves: Loxia) *Molecular Ecology* **15**:1873-1887

Poelstra,JWN, Vijay CM, Bossu H, Lantz B, Ryll I, Müller V, Baglione P, Unneberg M Wikelski MG, Grabher J, Wolf JB (2014) The Genomic Landscape Underlying Phenotypic Integrity In The Face Of Gene Flow In Crows. *Science* **20** 1410-1414

Price TD (2008) *Speciation in Birds*, Roberts and Co, Greenwood Village, Colorado

R Core Team (2015) R: A language and environment for statistical computing R Foundation for Statistical Computing, Vienna, Austria URL http://www R-project org/

Rodrigues P, Lopes RJ , Reis S, Resendes R, Ramos, JA and ã o da Cunha R T (2014) Genetic diversity and morphological variation of the common chaffinch *Fringilla coelebs* in the Azores *Journal of Avian Biology* **45**: 167–178

Rolshausen G, Segelbacher G, Hobson KA, Schaefer HM (2009) A Recent Migratory Divide Facilitates the Contemporary Evolution of Reproductive Isolation and Phenotypic Divergence in Sympatry *Current Biology*, **19**: 2097-2101

Rubenstein DR and IJ Lovette (2007) Temporal environmental variability drives the evolution of cooperative breeding in birds *Current Biology* **17**:1414-1419

849   Ruegg K (2007) Divergence between subspecies groups of Swainson's thrush (*Catharus*
850   *ustulatus ustulatus* and *C.u. swainsoni*) *Ornithological Monographs*, **63**: 67- 77

851   Ruegg K (2008) Genetic, morphological and ecological characterization of a hybrid zone that
852   spans a migratory divide. *Evolution*, **62**: 452-466

853   Ritchie MG (2007) Sexual selection and speciation. *Annual Reviews of Ecology and*
854   *Systematics*, **38**, 79-102

855   Rundle H, Nosil P (2005) Ecological speciation *Ecology Letters* **8**, 336-352

856   Safran, RJ, McGraw KJ (2004) Plumage Coloration, Not Length Or Symmetry of Tail-Streamers,
857   Is A Sexually Selected Trait in North American Barn Swallows. *Behavioral Ecology* **15**: 455-461

858   Safran RJ. Neuman CR, McGraw KJ, Lovette IJ (2005) Dynamic Paternity Allocation As A
859   Function of Male Plumage Color in Barn Swallows. *Science* **309**: 2210-2212

860   Safran RJ , Flaxman SM. Kopp M, Irwin DE, Briggs D, Evans MR, Funk WC, Gray DA, Hebets E
861   A, Seddon N, Scordato ESC, Symes LB, Tobias J A, Toews DPL, Uy, JAC (2012) A robust
862   new metric of phenotypic distance to estimate and compare multiple trait differences among
863   populations *Current Zoology* **58**: 423-436

864   Safran RJ, Scordato ESC, Symes LB, Rodrıguez RL, Mendelson TC (2013) Contributions of
865   natural and sexual selection to the evolution of premating reproductive isolation: a research
866   agenda *Trends in Ecology & Evolution*, **28**, 643–650

867   Saino, N, Romano M, Rubolini D, Teplitsky C, Ambrosini R, Caprioli M, Canova L, Wakamatsu
868   K (2013) Sexual dimorphism in melanin pigmentation, feather coloration and Its heritability in the
869   barn swallow (Hirundo rustica). *PLoS One* 8:e58024.

870   Scordato, ESC, Symes, LB, Mendelson, TC, and Safran RJ (2014) The role of ecology in
871   speciation by sexual selection: a systematic empirical review *Journal of Heredity* **14**: 782–794

872   Scordato ESC, Safran RJ (2014) Geographic variation in sexual selection and implications for
873   speciation in the barn swallow *Avian Research* 5:8

874   Seddon N, Botero, CAB, Dunn PO, MacGregor H, Rubenstein DR, Tobias JA, Uy JAC, Weir J,
875   Whittingham LW, Safran, RJ (2013) Sexual selection accelerates the evolution of phenotypic
876   divergence and reproductive isolation in birds. *Proceedings of the Royal Society of London B*
877   *Biological Sciences* **280:** 20131065

878   Seehausen O, Butlin RK, Keller I, Wagner CE, Boughman JW, Hohenlohe PA *et al* (2014)
879   Genomics and the origin of species *Nature Reviews, Genetics* **15**: 176–192

880   Siepielski, AM, DiBattista JD, Evans J, Carlson, SM (2011) Differences in the temporal
881   dynamics of phenotypic selection among fitness components in the wild *Proceedings of the*
882   *Royal Society of London B Biological Sciences* **278**:1572-1580

883   Schluter D (2001) Ecology and the Origin of Species.*Trends in Ecology and Evolution* **16**, 372-
884   399

885   Schluter D (2009) Evidence for ecological speciation and its alternative. *Science* **323**, 737-741

886   Shafer ABA, Wolf JBW (2013) Widespread evidence for incipient ecological speciation: a meta-
887   analysis of isolation-by-ecology *Ecology Letters*, **16**, 940–950

888   Slatkin M (1993) Isolation by distance in equilibrium and nonequilibrium populations Evolution,
889   47, 264–279

890 Svensson, E , Eroukhmanoff, F, Friberg M (2006) Effects of natural and sexual selection on
891 adaptive population divergence and premating isolation in a damselfly *Evolution* **60**:1242-1253

892 Takezaki N, Nei M (1996) Genetic distances and reconstruction of phylogenetic trees from
893 microsatellite DNA Genetics, **144**, 389–399

894 Thomsen PF, Jørgensen PS, Bruun HH, Pedersen J, Riis-Nielsen T, Jonko K, Słowińska I,
895 Rahbek C, Karsholt O (2016), Resource specialists lead local insect community turnover
896 associated with temperature – analysis of an 18-year full-seasonal record of moths and beetles.
897 *J Anim Ecol* 85: 251–261.
898

899 Turner AK (2006) *The Barn Swallow* London: T & AD Poyser

900 van Doorn, G S *et al* (2009) On the origin of species by natural and sexual selection *Science*
901 **326**, 1704–7

902 von Rönn, JAC, Shafer ABA, Wolf, JBW (2016) Disruptive selection without evolution across a
903 migratory divide *In press*, Molecular Ecology

904 Vortman Y, Lotem A, Dor R, Lovette IJ, Safran RJ (2011) The sexual signals of the East-
905 Mediterranean barn swallow (*Hirundo rustica transitiva*): evidence for geographic variation in
906 patterns of signal-based reproductive performance *Behavioral Ecology* **22**:1344–1352

907 Vortman Y, Dor R, Lovette IJ, Lotem A, Safran RJ (2013) Multiple signals and behavioral
908 reproductive isolation in a diverging population. *American Naturalist* **182**:514-523

909 Wagner, CE, Harmon LJ. Seehausen, O (2012) Ecological opportunity and sexual selection
910 together predict adaptive radiation *Nature* **487**, 366–369

911 Wang IJ, Summers K (2010) Genetic structure is correlated with phenotypic divergence rather
912 than geographic isolation in the highly polymorphic strawberry poison-dart frog. *Molecular*
913 *Ecology*, **19**, 447–458

914 Wang IJ (2013) Examining the full effects of landscape heterogeneity on spatial genetic
915 variation: a multiple matrix regression approach for quantifying geographic and ecological
916 isolation *Evolution*, **16**, 175–182

917 Wang IJ, Glor RE, Losos JB (2013) Quantifying the roles of ecology and geography in spatial
918 genetic divergence. *Ecology Letters*, **16**, 175–182

919 Wang, IJ and Bradburd GS (2014) Isolation by environment. *Molecular Ecology* **23**, 5649-5662

920 West-Eberhard, M J (1983) Sexual selection, social competition, and speciation *Quarterly*
921 *Review of Biology* **58**, 155–183

922 Wright S (1943) Isolation by distance. *Genetics*, **28**, 114–138

923 Zhang G, Li C, Li Q, Li B, Larkin DM, Lee C, Ödeen, A (2014) Comparative genomics reveals
924 insights into avian genome evolution and adaptation *Science*, **346**, 1311-1320

925 Zink RM, Pavlova A Rohwer and Drovetski SV (2006) Barn swallows before barns: population
926 histories and intercontinental colonization *Proceedings of the Royal Society of London B*:
927 **273**:1245-1251

928

929

930

931 *Data accessibility*

932 Genome assembly: XXX

933 Phenotype and environmental data: Dryad XXX

934 Protocol and models: Dryad XXX

935

936 *Authors' contributions:*

937 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.,

938 collected data in the field, N.C and E.S.C.S. completed the draft assembly of the genome, B.R.J.

939 conducted most of the lab work, R.J.S., E.S.C.S., S.M.F., T. P. analyzed the data with input from

940 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

941 final approval for publication.

942

943 **Table 1.** Sampling locations across four subspecies of *Hirundo rustica*. Acronyms are

944 given for each location that match with locations on Figures 1 and 2.

945

946

Sampling location [abbreviation for location on map]	subspecies	Lat/Long	Final Sample size	Sampling dates: year (sample size)
Boulder, Colorado, USA [CO]	*Hirundo rustica erythrogaster*	40.17, -105.10	144	2008 (50) 2009 (72) 2010 (22)
Czech Republic [CR]	*Hirundo rustica rustica*	49.06, 14.76	24	2010
Israel [IL]	*Hirundo rustica transitiva*	32.93, 35.54	45	2008 (3) 2009 (37) 2010 (5)
Ithaca, New York, USA [IA]	*Hirundo rustica erythrogaster*	42.44, -77.50	27	2002
Romania [RM]	*Hirundo rustica rustica*	46.75, 23.83	16	2010
Taiwan [TW]	*Hirundo rustica gutturalis*	25.09, 121.56	18	2010
Turkey [TR]	*Hirundo rustica rustica*	36.85, 31.16	50	2010 (50)
United Kingdom [UK]	*Hirundo rustica rustica*	50.50, -4.65	26	2009
Totals				
8 populations	*4 subspecies*		350	

947
948
949

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

Types of Traits	Traits
Traits related to reproductive performance among sub-species of barn swallows	Tail streamer length, ventral color (throat breast and vent, % brightness)
Trait related to flight behavior and migratory distance	Wing length (mm)
Features of the environment that affect food availability	Mean, variability, and seasonal minima in breeding temperatures; elevation

952

953   Table 3. Average pairwise F_{ST} (upper diagonal) and geographic distance (km; lower

954   diagonal).

955

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

956

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

	Population Sampled								
Trait	Colorado	Czech Republic	Israel	New York	Romania	Taiwan	Turkey	UK	*Range in percentiles)*
Wing length	24.08	70.73	61.00	30.22	61.87	9.02	61.00	91.16	9.02-91.16
Streamer length	21.82	72.23	60.22	15.05	79.79	28.66	67.16	60.22	15.05-79.79
Throat Color	77.74	59.42	39.96	80.77	18.35	28.26	47.12	44.62	18.35-80.77
Breast color	21.03	58.63	27.61	21.99	58.72	80.77	55.41	70.39	21.03-80.77
Vent color	18.10	27.01	16.02	20.32	21.45	32.99	25.32	30.14	16.02-30.14
Elevation	100	62.5	75	37.5	50	12.5	87.5	25	12.5-100
Var temp	37.5	75	87.5	62.5	50	25	100	12.5	12.5-100
Min temp	62.5	25	12.5	87.5	50	100	87.5	37.5	12.5-100
Mean temp	62.5	25	12.5	75	50	100	87.5	37.5	12.5-100

963

964 Table 5. Results of ANOVAs to show geographic variation in morphological traits among
965 eight closely related populations of barn swallows.

Trait	F	P	Adj R-squared
Wing length	$54.11_{7,353}$	< 0.001	0.52
Streamer length	$54.27_{7,352}$	< 0.001	0.52
Throat Color	$18.84_{7,346}$	< 0.001	0.25
Breast color	$28.16_{7,345}$	< 0.001	0.37
Vent color	$24.97_{7,345}$	< 0.001	0.34

966

967

968

969

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.

975

Variable	Association with genome wide F_{ST}			Association with geographic distance		
	Mantel r	*CI*	*P value*	*Mantel r*	*CI*	*P value*
Wing length	0.85	0.79-0.91	**0.002**	0.60	0.34-0.76	**0.010**
Tail length	0.67	0.60-0.81	**0.012**	0.59	0.38-0.85	**0.013**
Throat color	0.29	-0.08-0.66	0.091	0.48	0.27-0.78	**0.030**
Breast color	0.34	0.16-0.53	**0.023**	0.46	0.24-0.74	**0.031**
Vent Color	0.40	0.14-0.59	**0.037**	0.30	-0.11-0.54	0.094
Elevation	-0.04	-0.27-0.38	0.593	0.17	-0.07-0.35	0.202
Min breeding temp	0.07	-0.18-0.40	0.348	-0.01	-0.28-0.38	0.483
Mean breeding temp	0.20	-0.18-0.64	0.145	0.13	-0.14-0.54	0.274
Var breeding temp	0.07	-0.11-0.29	0.346	-0.05	-0.30-0.27	0.560

976 **Table 7.** Quantitative tests of variance partitioning (Figure 7): RDA analyses of the effects of
977 selection and geographic distance on genome-wide divergence. WING = pairwise differences in
978 wing length, COLOR = pairwise differences in throat color, DIST = pairwise differences in
979 geographic distance. Note that the matrix title before the symbol '|' indicates the variance due
980 to that matrix that contributed to genome-wide divergence, conditioned upon the matrices listed
981 after the '|' symbol.

982

| Model	F	*P*	Adjusted R-squared
WING, COLOR, DIST	$WING_{1,24} = 78.36$	< 0.001	0.73
	$COLOR_{1,24} = 7.03$	0.010	
	$DIST_{1,24} = 0.03$	0.878	
WING|COLOR+DIST	$WING_{1,24} = 44.17$	< 0.001	0.42
COLOR|WING+DIST	$COLOR_{1,24} = 6.03$	0.018	0.05
DIST|WING+COLOR	$ECO_{1,24} = 0.03$	0.878	0.00
WING+COLOR|DIST	$WING_{1,24} = 38.66$	< 0.001	0.40
	$COLOR_{1,24} = 6.03$	0.021	
DIST+COLOR|WING	$DIST_{1,24} = 1.04$	0.321	0.05
	$COLOR_{1,24} = 6.03$	0.024	
DIST+WING|COLOR	$DIST_{1,24} = 32.15$	< 0.001	0.70
	$WING_{1,24} = 44.17$	< 0.001	

983

984

985

986

Figure 1. *Hirundo rustica* complex breeding distribution map indicating eight sampling locations. Cartoons of male phenotypes are shown for each subspecies (with permission by artist Hilary Burn). 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.

992

993
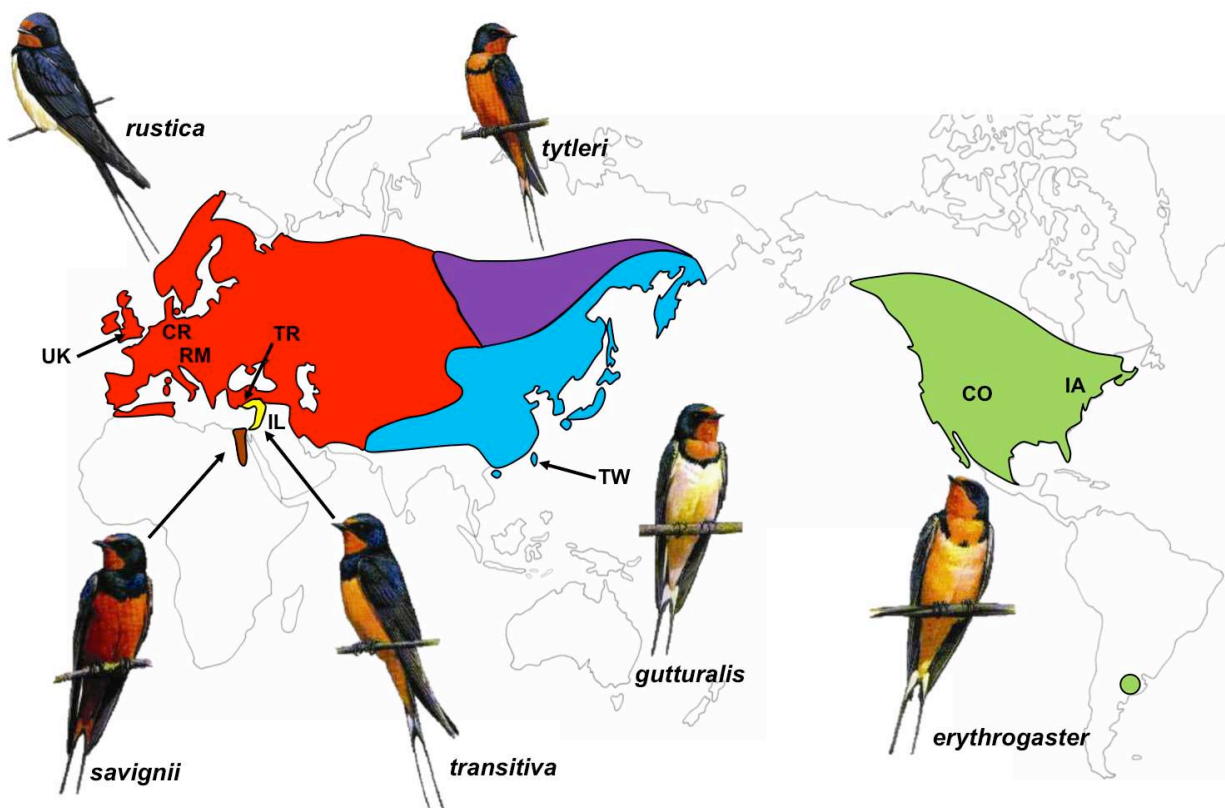
Figure 2. Statistical summary of population genetic structure based on principal component axes one (PC1), and two (PC2) derived from genotype probabilities. Acronyms for sampling sites are as follows: CO = Colorado, USA; CR = Czech Republic; IA = New York, USA; IL = Israel; RM = Romania; TR = Turkey; TW = Taiwan; UK = United Kingdom.
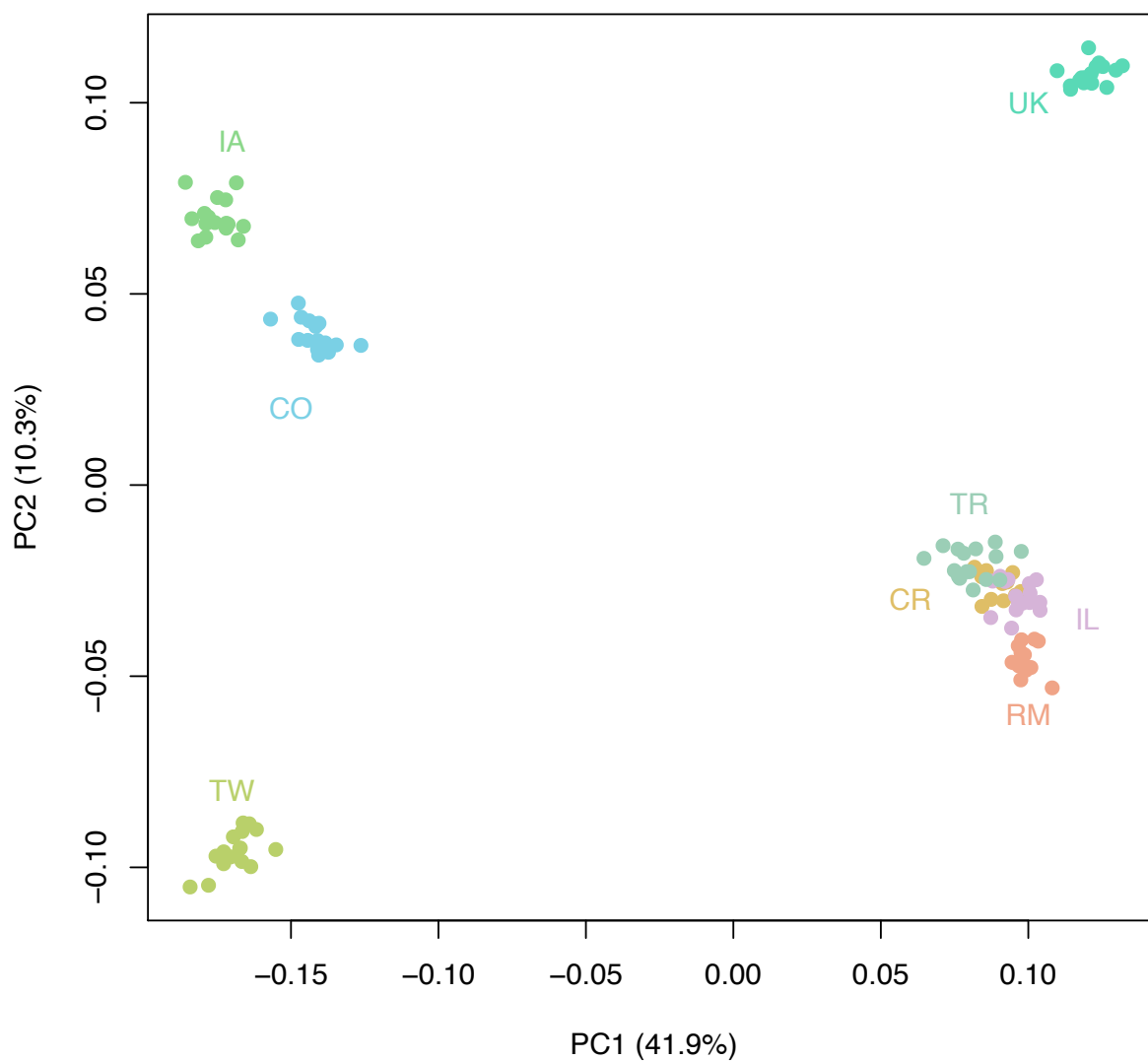
999

1000

1001

1002

1003 **Figure 3.** Average pairwise genetic distance (F_{ST}) as a function of geographic distance
1004 (km) among pairwise comparisons of eight barn swallow populations, consistent with a
1005 pattern of isolation by distance
1006
1007
1008
1009
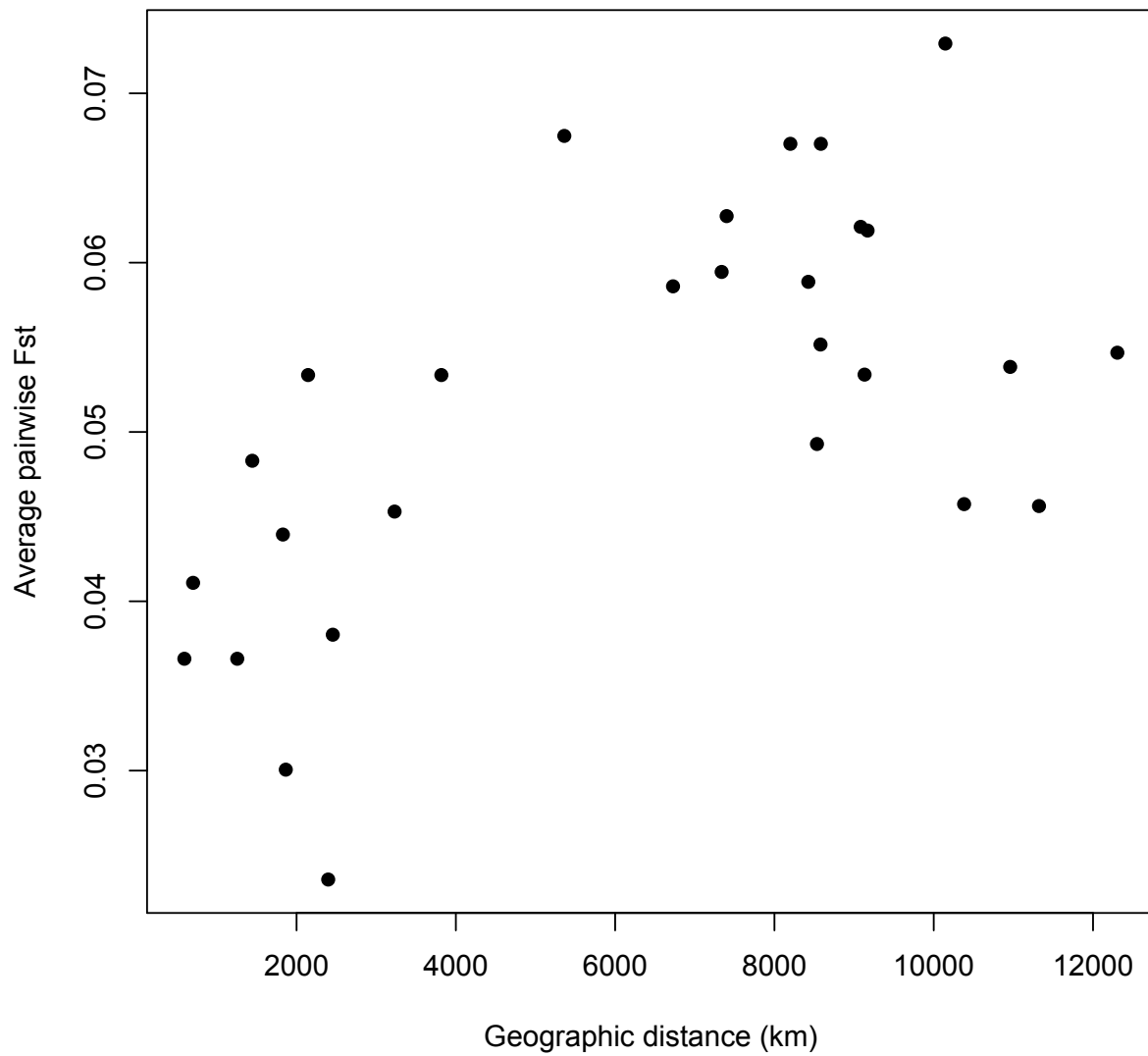
1010

1011
1012
1013 **Figure 4.** Average pairwise genetic distance (F_{ST}) is strongly correlated with phenotypic
1014 distance among pairwise comparisons of eight barn swallow populations, consistent
1015 with a pattern of IBA. Table 6 shows results of mantel tests between Fst and these four
1016 phenotype measures; all are significantly associated with the exception of throat color.
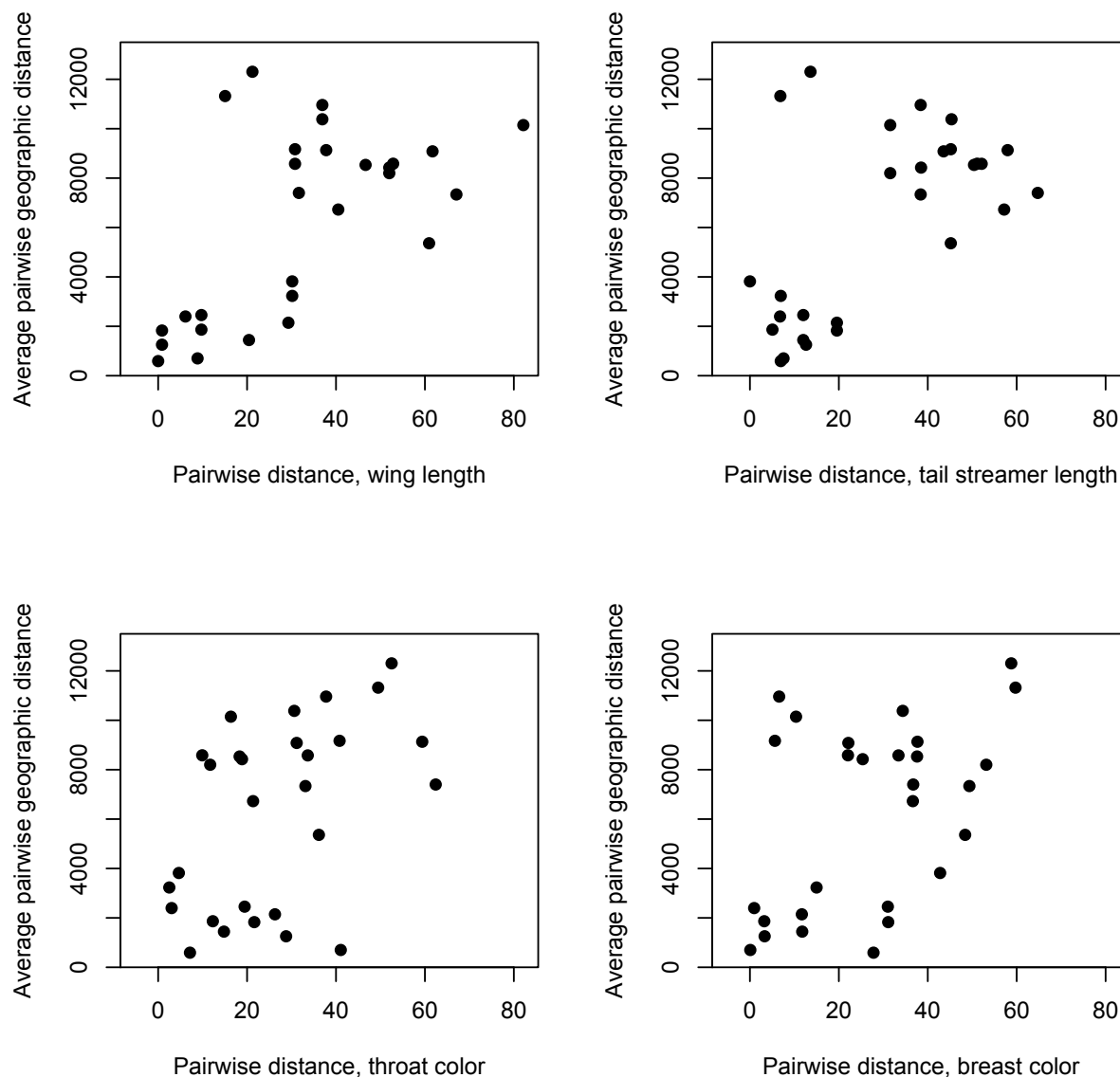1017
1018
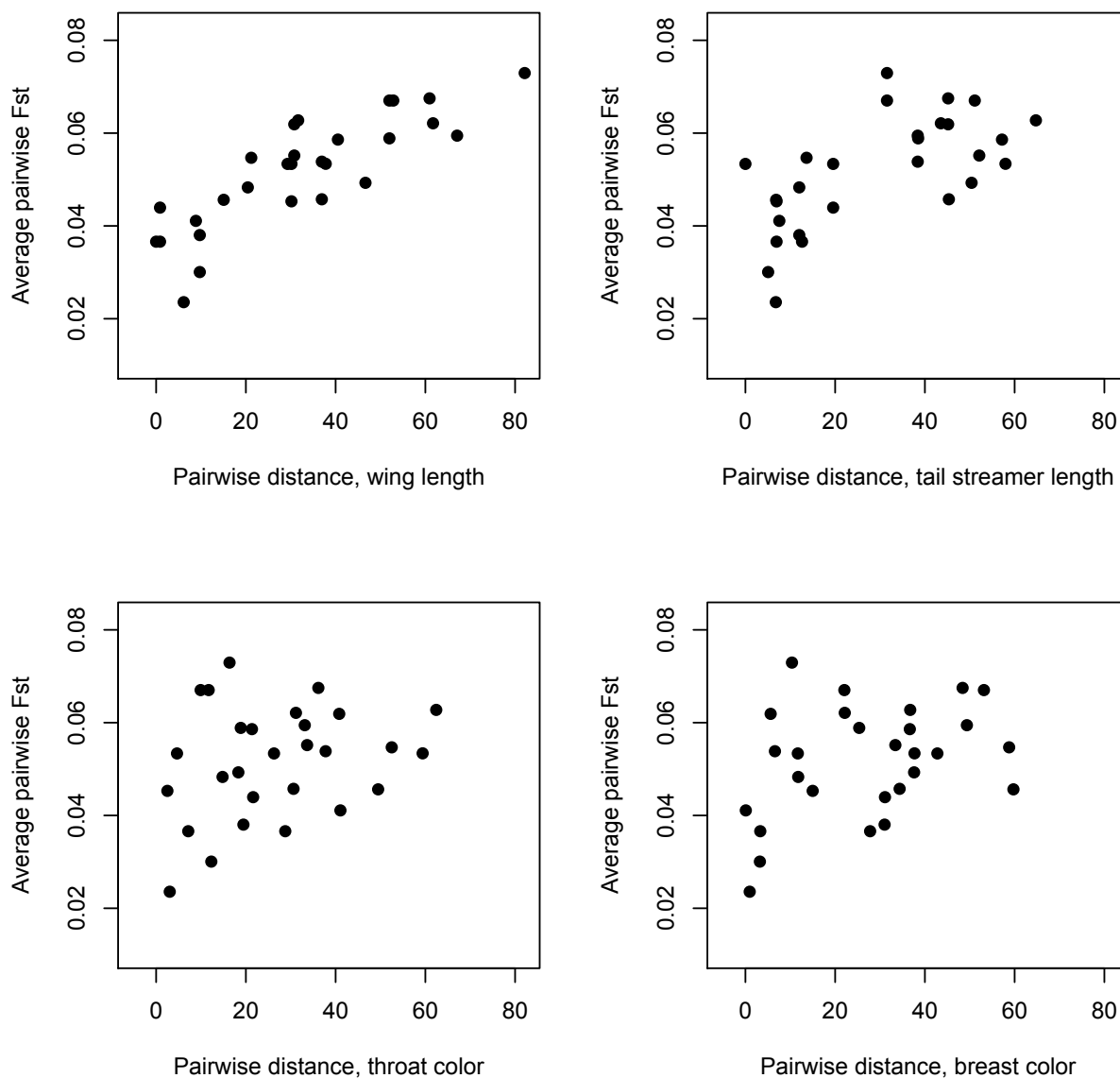
1019
1020
1021
1022
1023

1024 **Figure 5.** Average pairwise genetic distance (F_{ST}) as a function of environmental
1025 distance among pairwise comparisons of eight barn swallow populations. Table 6 shows
1026 results of mantel tests between F_{ST} and environmental distance; that there are no
1027 statistically significant associations between genetic distance and these environmental
1028 variables.
1029
1030
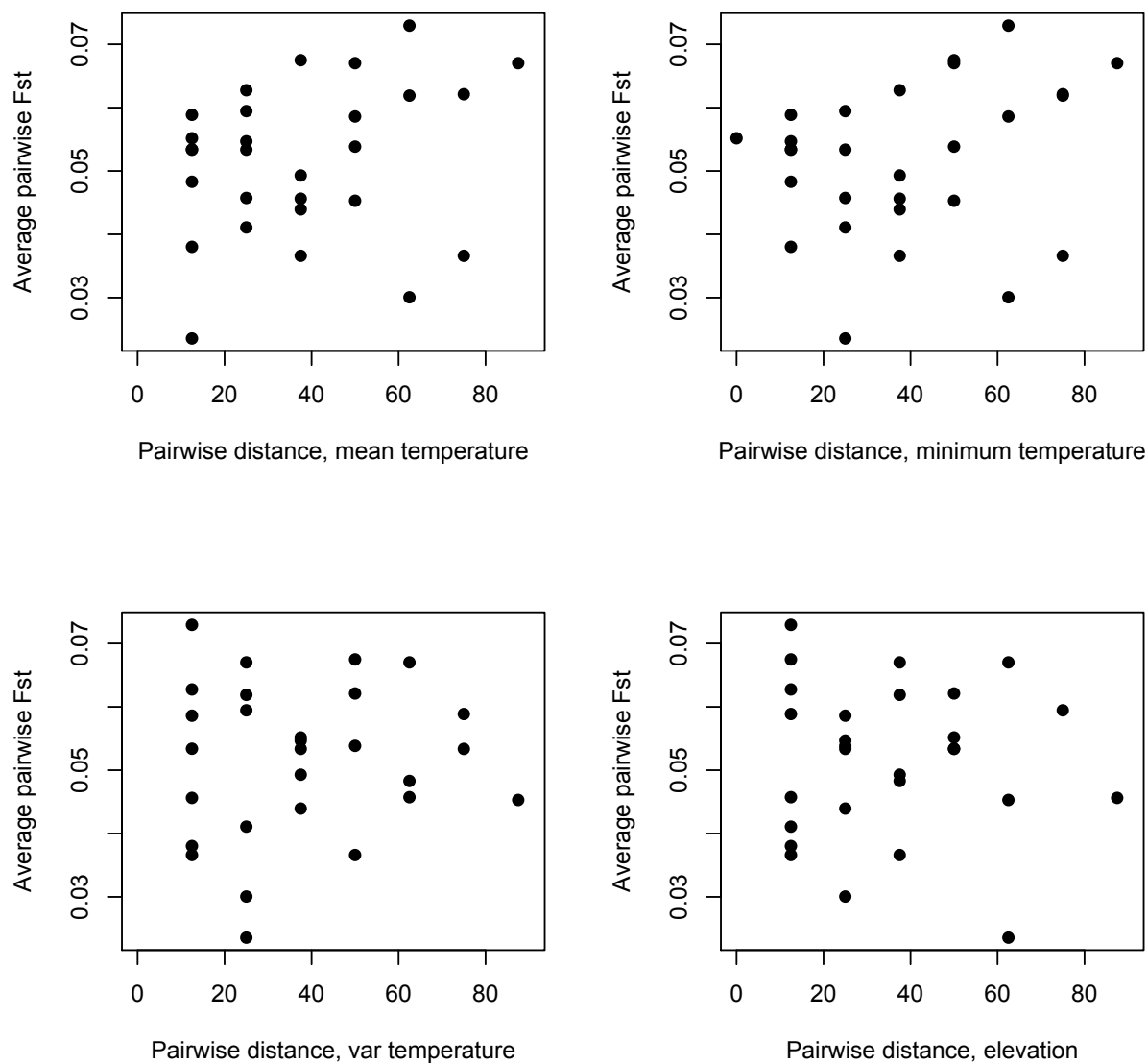
1031
1032
1033
1034

1035 **Figure 6.** Average pairwise geographic distance as a function of phenotypic distance
1036 among pairwise comparisons of eight barn swallow populations indicate spatial
1037 autocorrelation between morphological and geographic distance. Table 6 shows results
1038 of mantel tests between geographic distance and phenotype distance; all of these
1039 associations are statistically significant.
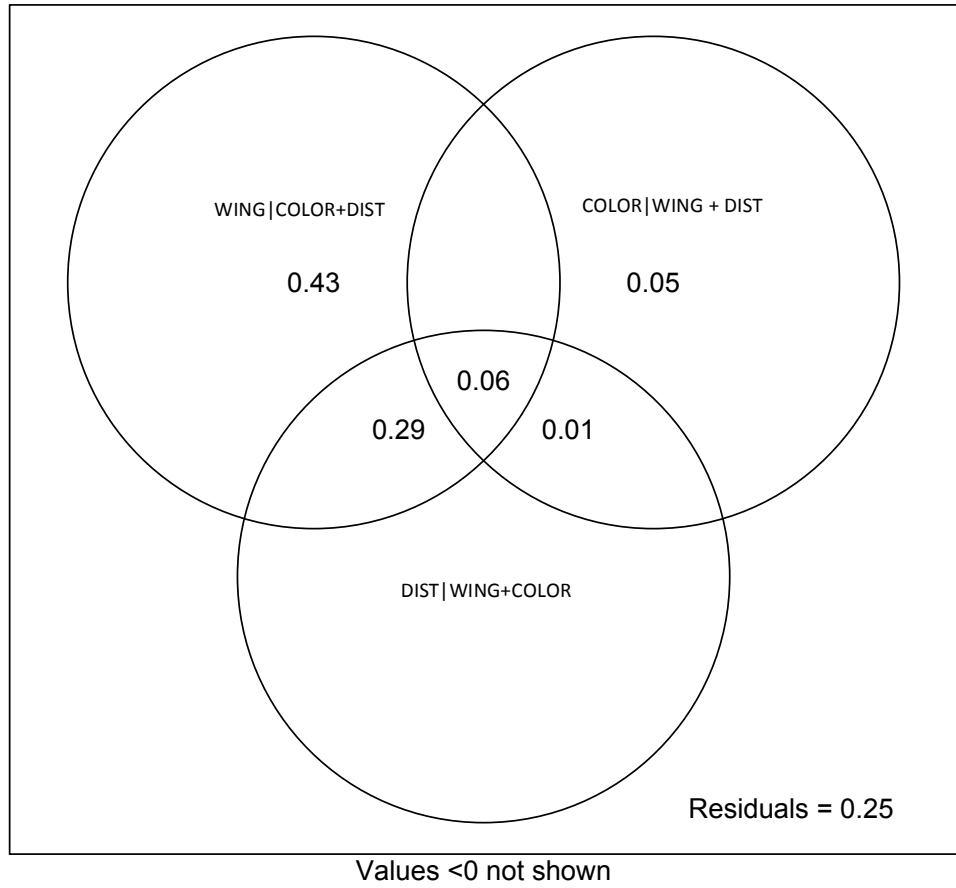1040

1041

1042   **Figure 7.** Average pairwise geographic distance as a function of environmental distance
1043   among pairwise comparisons of eight barn swallow populations. Table 6 shows results
1044   of Mantel tests between geographic distance and environmental distance; none of these
1045   associations are statistically significant.
1046
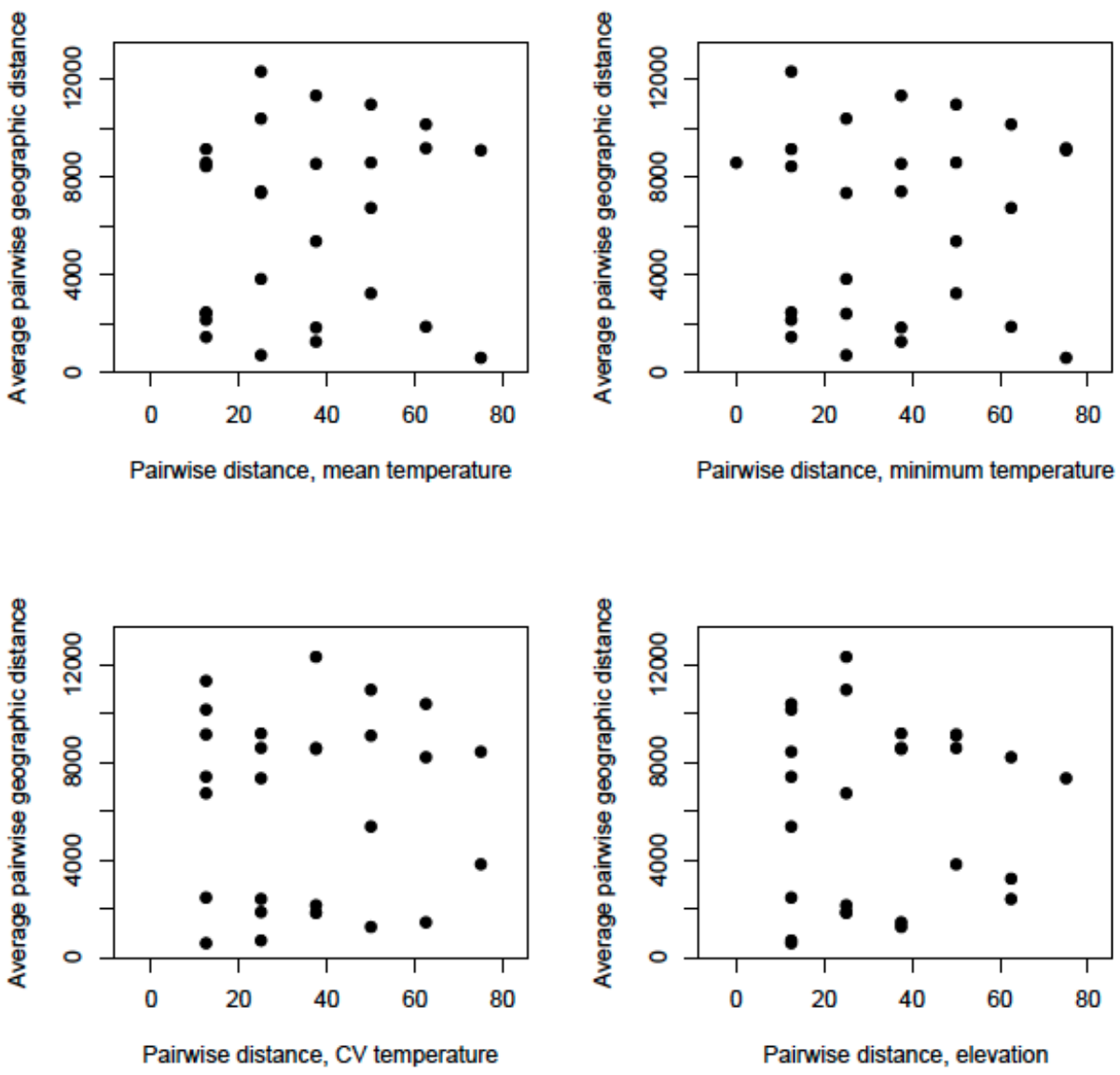
1047   **Figure 8.** Variance partitioning plot to represent the extent to which natural selection (WING),
1048   sexual selection (COLOR), and geographic distance (DIST) explain variance in pairwise,
1049   genome-wide differences in F_{ST}. Note that the matrix title before the symbol '|' indicates the
1050   variance due to that matrix that contributed to genome-wide divergence, conditioned upon the
1051   matrices listed after the '|' symbol.
1052
1053
1054

Values <0 not shown

1055