1

2

3

4 # Genomic Insights into Isolation of the Threatened Florida Crested Caracara (*Caracara*
5 ***plancus***)

6

7 Natalie Payne[1*], John A. Erwin[2], Joan Morrison[3], James Dwyer[4], Melanie Culver[5, 6, 1]

8

9 [1] Genetics Graduate Interdisciplinary Program, University of Arizona, Tucson, AZ 85719, USA

10 [2] Florida International University College of Law, Miami, FL 33199, USA

11 [3] Department of Biology, Trinity College, 300 Summit Street, Hartford, CT 06106 USA

12 [4] EDM International, Inc., Fort Collins, CO 80525, USA

13 [5] U.S. Geological Survey, Arizona Cooperative Fish and Wildlife Research Unit, University of
14 Arizona, Tucson, AZ 85721, USA

15 [6] School of Natural Resources and the Environment, University of Arizona, Tucson, AZ 85721,
16 USA

17 *Correspondence: nataliermercer@gmail.com (N. Payne)

18 **Abstract**

We conducted a population genomic study of the crested caracara (*Caracara plancus*) using samples (*n* = 290) collected from individuals in Florida, Texas, and Arizona, USA. Crested caracaras are non-migratory raptors ranging from the southern tip of South America to the southern United States, including a federally protected relict population in Florida long thought to have been isolated since the last ice age. Our objectives were to evaluate genetic diversity and population structure of Florida's apparently isolated population and to evaluate taxonomic relationships of crested caracaras at the northern edge of their range. Using DNA purified from blood samples, we conducted double-digest restriction site associated DNA sequencing and sequenced the mitochondrial *ND2* gene. Analyses of population structure using over 9000 SNPs suggest that two major clusters are best supported, one cluster including only Florida individuals and the other cluster including Arizona and Texas individuals. Both SNPs and mitochondrial haplotypes reveal the Florida population to be highly differentiated genetically from Arizona and Texas populations, whereas Arizona and Texas populations are moderately differentiated from each other. The Florida population's mitochondrial haplotypes form a separate monophyletic group, while Arizona and Texas populations share mitochondrial haplotypes. Results of this study provide substantial genetic evidence that Florida's crested caracaras have experienced long-term isolation from caracaras in Arizona and Texas and, thus, represent a distinct evolutionary lineage possibly warranting distinction as an Evolutionarily Significant Unit (ESU) or subspecies. This study will inform conservation strategies focused on long-term survival of Florida's distinct, panmictic population.

Keywords: ddRADseq; Conservation Genomics; Falconidae; Genetic Isolation

Introduction

41 The crested caracara (*Caracara plancus*; hereafter “caracara”) is a medium-sized member of
42 Falconidae and the only extant species in its genus (Chesser *et al.* 2021; Gill *et al.* 2021). It is a
43 resident, non-migratory raptor that occurs throughout Central and South America, with
44 populations in the southern and northern portions of the range previously designated as separate
45 species (Chesser *et al*. 2021; Gill *et al*. 2021). The range of the “northern” crested caracara
46 (previously *C. cheriway*) includes northern South America and Central America and extends
47 northward to the southern U.S., with populations in Texas and Arizona and a relict population in
48 central Florida (Morrison and Dwyer 2023). While reliable estimates of population size are not
49 available, the Florida population has long been considered small, possibly around 500
50 individuals (Layne 1996). However, this number is based on limited data, and recent research on
51 communal roost use in this population revealed more than 300 non-breeding individuals together
52 at a single communal roost (Dwyer and Fraser 2018). Small, isolated populations are of
53 conservation concern because they are prone to loss of genetic diversity and inbreeding
54 (Woodruff 2001). Genetic homogeneity may interact with demographic and environmental
55 factors in small populations leading to an “extinction vortex”, a cycle which further reduces
56 population size and propels a population toward extinction (Gilpin and Soulé 1986).

57 Studies of intraspecific variation and population structure within the crested caracara throughout
58 the species’ range are lacking. The Florida population was considered a North American
59 subspecies (*Polyborus plancus audubonii*) when listed as threatened under the federal
60 Endangered Species Act (U.S. Fish and Wildlife Service 1987). However, all subspecies within
61 the genus *Caracara* have since been collapsed and are not currently considered taxonomically
62 distinct (Gill *et al.* 2021). Despite lack of taxonomic resolution within this species, a variety of
63 factors suggest the relict Florida population has experienced long-term isolation and

64 consequently may be limited in genetic diversity. Fossil records indicate crested caracaras have
65 inhabited Florida since the Pleistocene (Brodkorb 1959), and the population's isolation since
66 then likely has resulted from loss of grassland habitat along the Gulf Coast savanna corridor due
67 to interglacial climatic changes (Emslie 1998; Morgan and Emslie 2010) and subsequent
68 expansion of southern pine forests in northern Florida (Myers and Ewel 1990). This pattern
69 mirrors the isolation of other Plio-Pleistocene extra-limital populations of avian fauna in
70 peninsular Florida (Morgan and Emslie 2010), including the burrowing owl (*Athene cunicularia*
71 *floridana*) and Florida scrub-jay (*Aphelocoma coerulescens*). These two non-migratory
72 populations show strong genetic differentiation from their western counterparts (which are
73 recognized as different subspecies and species, respectively) and exhibit reduced genetic
74 diversity as a result of their long-term isolation (McDonald et al. 1999; Desmond et al. 2001;
75 Korfanta et al. 2005; Macías-Duarte et al. 2019).

76 Like the isolation of Florida scrub jays and burrowing owls, extensive telemetry data from
77 tagged caracaras and the absence of band recoveries of caracaras outside of Florida (Morrison
78 and Humphrey 2001; Dwyer *et al.* 2013) suggest there is no movement or genetic exchange
79 between Florida and other crested caracara populations. Understanding the present genetic
80 diversity and extent of isolation of Florida's caracara population is crucial to development of
81 informed management strategies needed for the conservation of this population.

82 We utilized double-digest restriction site associated DNA sequencing (ddRADseq) to develop a
83 baseline for genetic variation of the Florida population, including estimates of genetic diversity,
84 effective population size, and extent of inbreeding, and to compare these parameters among
85 caracara populations in Florida, Arizona, and Texas. Subsequently, we investigated the extent of
86 isolation of the Florida population and evaluated whether barriers to gene flow occur within the

87 Florida population. We also analyzed mitochondrial DNA haplotypes to further clarify taxonomy
88 of the Florida crested caracara in the context of other populations at the northern extent of the
89 species' range. To our knowledge, this is the first population genomic analysis of the crested
90 caracara in any portion of its range and the first to examine population isolation and molecular
91 genetic diversity of Florida's caracaras.

92 Materials and Methods

93 *Sample Collection and Extraction*

94 Blood samples were taken from the brachial vein of captured wild caracaras in Florida (FL; *n* =
95 273; Figure 1) and Arizona (AZ; *n* = 10). Seven blood samples were also obtained from captive
96 caracaras with known origins in Florida (*n* = 1) and Texas (TX; *n* = 6). DNA was extracted from
97 all samples (*n* = 290) using Qiagen's DNeasy Blood and Tissue Kit (Qiagen Ltd., Valencia, CA).

98 *ddRADseq Library Construction*

99 We prepared genomic libraries following the protocol of Peterson *et al.* (2012), using the
100 restriction enzyme pair *Nla*III and *Eco*RI and targeting a fragment size range of 254–326 base
101 pairs. Using the flex adapters from Peterson *et al.* (2012), with modifications to the P2 adapter to
102 include degenerate barcodes for PCR duplicate removal, we were able to pool up to 48
103 individuals. Size selection for fragments within the 254–326 base pair range (targeting 340–412
104 bp to account for the addition of adapter sequences) was done using a Pippin Prep (Sage
105 Scientific; Beverly, MA) with 2% agarose, 100-600 bp Pippin Gel Cassettes with Marker B.
106 Each pool of individuals was split into six PCR reactions, and each reaction was carried out for
107 12 cycles. Unique pool indices were added during PCR, and each library contained two pools

108 and 37–68 samples total. Libraries represented a total of 209 individuals, with samples from
109 Florida (*n* = 196), Arizona (*n* = 8), and Texas (*n* = 5). We also included 27 replicate pairs within
110 and between libraries to be used for calculating error rates. Libraries were sequenced on a
111 HiSeq4000 or HiSeq X Ten (150 bp paired-end) by Novogene (https://en.novogene.com/;
112 Beijing, China).

113 *Bioinformatic Pipeline*

114 We received ddRAD sequencing data from Novogene as demultiplexed pools. We confirmed the
115 presence of RAD enzyme cut sites and per base and sequence quality of each pool using
116 FASTQC v. 0.11.9 (Andrews 2010). Prior to demultiplexing further, we removed PCR
117 duplicates based on the 5-nucleotide degenerate barcodes present at the 5' ends of the reverse
118 reads using program *clone_filter* in Stacks v. 2.60 (Catchen *et al.* 2011; Catchen *et al.* 2013;
119 Rochette *et al.* 2019). We trimmed an additional five bases from the 5' ends of the reverse reads
120 using Trimmomatic v. 0.39 (Bolger *et al.* 2014) to expose the *Eco*RI cut site at the 5' ends of the
121 reads. Prior to demultiplexing, we verified the removal of degenerate barcodes and additional
122 five bases from the reverse reads using FASTQC. To demultiplex each pool into individuals
123 based on in-line adapter barcodes at the 5' ends of forward reads and remove reads with adapter
124 contamination or missing RAD enzyme cut sites, we used the Stacks program *process_radtags*
125 (Catchen *et al.* 2011; Catchen *et al.* 2013; Rochette *et al.* 2019). We utilized R code developed
126 by Mastretta-Yanes *et al.* (2015) for calculating SNP error rates and 11 parameter combinations
127 used by Trumbo *et al.* (2019) to optimize parameters for the *denovo_map* pipeline using only
128 replicate pairs and one SNP per locus (--write_random_snp). Prior to utilizing selected
129 parameters to run *denovo_map* on the full dataset, reads from duplicated individuals were
130 combined.

131 To decrease computational time, components of the *denovo_map* wrapper in Stacks were carried
132 out individually on the full dataset (Catchen *et al.* n.d.; Rochette and Catchen 2017). For *ustacks*,
133 we divided the 209 samples into 11 sets of 19 samples, and we ran each set on a different
134 compute node with parameters -m 3, -M 4, and --max_locus_stacks 4. We then ran *cstacks* with
135 parameter -n 4 to assemble a catalog of loci using only a subset of 40 samples [all from Arizona
136 (*n* = 8) and Texas (*n* = 5), and the 27 samples from Florida with highest coverage], as
137 recommended in the Stacks manual (Catchen *et al*. n.d.) and Rochette and Catchen (2017). This
138 was done to reduce computational time and to enrich representation of loci found in our under-
139 sampled populations. For *sstacks*, we divided samples into 11 groups and ran them on separate
140 nodes, as was done for *ustacks*. We then performed the remainder of the pipeline (*tsv2bam*,
141 *gstacks*, and *populations*), using --write_random_snp to retain only one SNP per locus (since
142 multiple SNPs in the same locus are likely to exhibit linkage disequilibrium).

143 We sequentially filtered our results in PLINK v. 1.9 (Purcell *et al.* 2007), first removing loci
144 genotyped in less than 75% of individuals (--geno 0.25), then removing individuals with more
145 than 50% missing data (--mind 0.5), and then removing loci with a minor allele frequency
146 (MAF) less than 1% (--maf 0.01). We used the final ".map" file to create a list of loci to retain in
147 *populations*, and we reran *populations* using these loci and a new population map file excluding
148 the removed individuals. We utilized the resulting SNPs to estimate relatedness among
149 individuals using the R package *related* (Pew *et al*. 2015) with the dyadml estimator, allowing
150 for inbreeding. Using the output from *related*, we detected two instances of potentially
151 duplicated samples, possibly due to human error in the lab or field. We believe the samples were
152 potentially duplicated because, while possible, monozygotic twinning is thought to be a rare
153 occurrence in birds (Bassett *et al.* 1999), and to our knowledge there are no records of more

154 chicks than eggs in any nest. The relatedness values were 0.986 for one pair of samples and
155 0.978 for the other pair of samples (with a value of 1.0 expected for relatedness with "self"). To
156 resolve this, we used PLINK to remove the sample from each pair with more missing data,
157 determined with the R package *radiator* (Gosselin *et al.* 2020). We then repeated filtering of loci
158 below the 1% MAF threshold and reran *populations* using an updated list of loci and population
159 map file to obtain the population summary statistics (observed and expected heterozygosity,
160 nucleotide diversity, and the inbreeding coefficient for each population) and genotype files for
161 downstream analyses.

162 *ddRADseq Data Analyses*

163 We used the SNPs obtained by ddRADseq to calculate effective population size (N_e) in Florida
164 with N_eEstimator v. 2.1 (Do *et al.* 2014), using the linkage disequilibrium method accounting for
165 monogamy. We first estimated N_e by using the full SNP set. We additionally performed the
166 analysis with ten randomly selected subsets of 3000 loci, of those polymorphic within the Florida
167 population, and we evaluated estimates produced with and without a MAF cut-off (P_{crit} = 0.05
168 and 0.0). An additional correction based on haploid chromosome number (n = 46; Tagliarini *et*
169 *al.* 2007) was applied to the N_e and jackknifed 95% confidence interval (CI) estimates, using
170 equation 1a in Waples *et al.* (2016). This enabled us to account for the downward bias introduced
171 by using thousands of markers spread across relatively few chromosomes.

172 We also used the SNP dataset to assess population structure. To do so, we first used the R
173 package *adegenet* (Jombart 2008) to perform a Principal Components Analysis (PCA) and a
174 Discriminant Analysis of Principal Components (DAPC), which better discriminates between
175 groups by maximizing between-group variation while minimizing within-group variation. We

176 ran ADMIXTURE v. 1.3.0 (Alexander *et al.* 2009) to assess the most likely number of
177 populations from K = 1–10, with the value of K corresponding to the lowest cross-validation
178 error after 10 replicates being most likely, using both the full dataset and subset of 40 individuals
179 [all from Arizona (n = 8) and Texas (n = 5), and a subset of Florida samples selected randomly
180 (n = 27)], to reduce effects of unequal sample sizes. An additional filtering step was applied with
181 PLINK (--geno 0.25) prior to running ADMIXTURE with the subset due to some loci having
182 100% missing data, and 9099 variants were retained after filtering for the analysis. We
183 additionally ran ADMIXTURE on the Florida samples after pruning for relatedness with PLINK
184 (--rel-cutoff 0.2), followed by additional filtering (--geno 0.25 and --maf 0.01), retaining 131
185 individuals and 7035 variants for the analysis. ADMIXTURE results were visualized by plotting
186 the Q-matrices with *pophelper* (Francis 2017). Arlequin v. 3.5.2.2 (Excoffier and Lischer 2010)
187 was used to calculate pairwise F_{ST} between populations using 99,999 permutations and the
188 Arlequin project (.arp) file generated by PGDspider v. 2.1.1.5 (Lischer and Excoffier 2012).

189 We conducted a Mantel test to determine whether patterns of genetic variation may be better
190 explained by isolation by distance (IBD) than by panmixia or distinct subpopulations within
191 Florida. We performed the Mantel test using geographic Euclidean distances and genetic
192 distances (the inverse proportion of shared alleles, D_{PS}) between wild Florida individuals with
193 *adegenet* (Jombart 2008), and we assessed significance with 9,999 permutations.

194 *Mitochondrial DNA Sequencing*

195 We performed polymerase chain reaction (PCR) to amplify and sequence the mitochondrial *ND2*
196 region from caracaras from Florida (n = 271), Arizona (n = 10), and Texas (n = 6). The reverse
197 primer was H6313 (5'-ACTCTTRTTTAAGGCTTTGAAGGC-3'; Sorenson *et al.*, 1999), used

198 previously by Fuchs *et al.* (2012), and the forward primer was L5219 (5'-
199 CCCATACCCCGAAAATGATG-3'; Sorenson *et al.*, 1999). Reaction conditions were: 15 ng
200 DNA, 1 µL of each 10 µM primer, and 12.5 µL of HotStarTaq MasterMix (Qiagen) in a 25 µL
201 reaction volume. We used the following amplification conditions: enzyme activation at 95˚C for
202 15 min; 35 cycles of 94˚C for 45 sec, 55˚C for 45 sec, and 72˚C for 1 min; a final extension at
203 72˚C for 10 min. Products were visualized on a 1.5% agarose gel stained with ethidium bromide
204 and the resulting amplicon was sequenced at the University of Arizona Genetics Core on an
205 ABI3730 DNA Analyzer (Applied Biosystems, Thermo Fisher Scientific Inc., Waltham, MA,
206 USA). We sequenced the test reaction products in both forward and reverse directions. Due to
207 noise in chromatograms of the reverse reads, we proceeded with amplifying the region with
208 L5219 and H6313 and sequencing in the forward direction with L5219 and L1700Falco (5'-
209 GGACAATGRGAYMTYACCCA-3'; Fuchs *et al.* 2012) as an internal primer for the rest of the
210 samples. Two additional PCR reactions were conducted for cases of possible heteroplasmy (the
211 presence of two mitochondrial haplotypes in one individual), one using the original L5219-
212 H6313 primer pair and the other using L1700Falco-H6313. These templates were resequenced in
213 both forward and reverse directions. After a clean-up of the PCR amplicons, noise appeared
214 greatly reduced in most of the reverse reads.

215 We used Geneious Prime v. 2021.0.3 (https://www.geneious.com) to manually assemble and
216 trim reads for each individual. We assembled reads using MUSCLE v. 3.8.425 (Edgar 2004)
217 with a reference sequence (consensus of available *C. cheriway* sequences on GenBank:
218 MN231451, KX534743, KM876089) to help resolve ambiguous bases while visually inspecting
219 chromatograms. We identified individuals as heteroplasmic when both haplotypes observed
220 together were also observed separately in other individuals. A consensus of cleaned and trimmed

221 reads was obtained for each individual, and these sequences were trimmed to 717 bp and aligned
222 for use in downstream analyses.

223 *Mitochondrial DNA Data Analyses*

224 The mitochondrial haplotype alignment (excluding heteroplasmic individuals) was exported
225 from Geneious into DnaSP v. 6 (Rozas *et al.* 2017) to generate an Arlequin project file and
226 determine number of haplotypes (H), haplotype diversity (H_d), and nucleotide diversity (π) for
227 each population. Because we detected two potential instances of sample duplication with our
228 ddRADseq data, we excluded the haplotype of one individual from one of the duplicate pairs.
229 Samples from the other duplicate pair were already excluded from downstream analyses due to
230 heteroplasmy. Arlequin v. 3.5.2.2 was used to calculate pairwise F_{ST} estimates with 99,999
231 permutations, as with nuclear data. We visualized haplotypes in a median-joining haplotype
232 network (Bandelt *et al.* 1999) generated by PopART v. 1.7 (Leigh and Bryant 2015), using the
233 haplotype file produced by DnaSP.

234 We performed a Bayesian phylogenetic analysis using an alignment of each unique haplotype
235 identified in our data and on GenBank (sequences from both *C. cheriway* and *C. plancus*). We
236 included *Spiziapteryx circumcincta* (JF909882) and the extinct *C. creightoni* (MN231452) as
237 outgroups. We used JModelTest 2 (Darriba *et al.* 2012) to determine the best nuclear substitution
238 model (HKY+G) based on Bayesian Information Criterion (BIC), and we performed four
239 independent runs using 200,000 generations, sampling every 10 generations, and 25% burn-in
240 with MrBayes v. 3.2.7 (Ronquist *et al.* 2012). We confirmed convergence of all runs using
241 Tracer v. 1.7.2 (Rambaut *et al.* 2018). We generated a 50% majority-rule consensus tree with
242 MrBayes and visualized the tree using FigTree v. 1.4.4.

243 **Results**

244 We obtained a set of 191,963 SNPs in 209 individuals after the first run of *populations* in Stacks.
245 Applying the --geno 0.25 filter in PLINK removed 173,685 SNPs, and applying --mind 0.5
246 removed 10 individuals, all of which were from Florida. We removed an additional 8197 SNPs
247 with --maf 0.01, resulting in 10,081 SNPs and 199 individuals (FL = 186, AZ = 8, TX = 5)
248 retained in the final dataset. After removing one individual from each duplicate pair with PLINK
249 and re-filtering using an MAF threshold of 1%, we retained 9274 SNPs in 197 individuals (FL =
250 184, AZ = 8, TX = 5).

251 Results of genetic diversity obtained for caracara populations in Florida, Arizona, and Texas
252 from the final *populations* output of Stacks revealed that nucleotide diversity (π) and observed
253 and expected heterozygosity (H_O and H_E) were highest in Texas and lowest in Florida, with
254 Florida being the only population with observed heterozygosity lower than the expected
255 heterozygosity (Table 1). We did not detect high levels of inbreeding (F_{IS}), with estimates near
256 zero for all three populations (Table 1).

257 We first estimated effective population size (N_e) from the full set of SNP data. Without using a
258 MAF cut-off, we estimated the N_e of the Florida population to be 565.4 individuals (95% CI:
259 458.2, 671.2). Estimating N_e with a MAF cut-off of 0.05 revealed a similar estimate (N_e = 507.0),
260 although this value was outside the confidence interval (95% CI: 539.8, 894.4). Because large
261 numbers of SNP comparisons in LD-based N_e estimation can lead to inaccurate confidence
262 intervals (Jones *et al.* 2016), we also estimated N_e using ten sets of 3000 randomly selected SNPs
263 that were polymorphic within the Florida population. This method returned similar values of N_e
264 for the Florida population that fall within their confidence intervals, with estimates between 400

265 and 500 with an MAF cut-off of 0.05, and estimates over 500 with no MAF cut-off
266 (Supplementary Figure 1). We did not estimate N_e for the Arizona and Texas caracara
267 populations because of small sample sizes from those states (England *et al.* 2006).

268 The PCA revealed three population clusters corresponding to Florida, Arizona, and Texas
269 (Figure 2A). Each individual caracara was assigned correctly to the cluster matching its state of
270 origin. However, the DAPC supported two clusters (with the lowest BIC value; Figure 2B),
271 collapsing the Arizona and Texas clusters into one group and retaining Florida as a second
272 distinct group (Figure 2C). From the DAPC, the membership probability for each individual to
273 its appropriate group was 100% (Figure 2D).

274 Cross-validation (CV) errors from ADMIXTURE indicated that K = 2–4 were best supported
275 using the full dataset (Supplementary Figure 2). K = 4 had the lowest CV error (0.4139), but this
276 was nearly identical to that of K = 2 (0.4144) and K = 3 (0.4145). The ancestry clusters generated
277 by ADMIXTURE for K = 2 corresponded to the clusters from the DAPC; Texas shared more
278 ancestry with Arizona, while Florida individuals formed a second cluster (Supplementary Figure
279 3). At K = 2, the five Texas individuals showed some apparent admixture from Florida, with
280 minimal admixture within the Florida lineage. However, ADMIXTURE results using the subset
281 of 40 samples (all from Arizona and Texas and 27 from Florida) indicated that each state
282 population represents a distinct ancestral background at K = 3 (CV error = 0.576), as suggested
283 by PCA clusters (Figure 3). K = 2 was best supported with the subset (CV error = 0.502), with
284 the same pattern of ancestral assignments observed at K = 2 as with the full dataset, and higher
285 values of K, which were less supported by cross-validation (Supplementary Figure 4), suggested
286 possible substructure within Florida. However, relatedness values estimated with the R package
287 *related* suggested ADMIXTURE was subdividing the Florida ancestry based on close versus

288 distant familial relationships, indicating lack of major barriers to gene flow, rather than true
289 subpopulations reflected by K = 4. Running ADMIXTURE with only Florida samples after
290 pruning for relatedness showed K = 1 was best supported (Supplementary Figure 5). We also did
291 not detect any evidence of strong IBD within the Florida population. Instead, the Mantel test
292 revealed a weak but significant correlation between geographic and genetic distance among
293 Florida caracaras ($r = 0.127$, $p = 0.0001$; Supplementary Figure 6).

294 From mitochondrial DNA data, we detected two haplotypes unique to Florida, two haplotypes
295 unique to Texas, and one haplotype shared by Arizona and Texas caracaras. Additionally, we
296 detected heteroplasmy (two mitochondrial haplotypes in a single individual) for both Florida
297 haplotypes in 32 samples, and heteroplasmy was detected irrespective of PCR batch. For samples
298 from each state, nucleotide diversity was low but increased with haplotype diversity and the
299 number of haplotypes (Table 1). Despite representation by the fewest samples ($n = 6$), the Texas
300 caracaras had the highest mitochondrial diversity, with three haplotypes represented and
301 haplotype diversity of 0.6. One of these haplotypes was shared with all ten Arizona individuals
302 ($H_d = 0$). Out of 238 non-heteroplasmic individuals, the sample of Florida's caracaras contained
303 only two unique haplotypes ($H_d = 0.148$).

304 Pairwise F_{ST} values calculated with nuclear SNP and mitochondrial DNA data revealed high,
305 significant differentiation between Florida and Arizona/Texas populations and moderate
306 differentiation between caracaras in Arizona and Texas (Table 2). The highest differentiation
307 was between caracaras in Florida and Arizona, with high differentiation at both the
308 mitochondrial level ($F_{ST} = 0.875$; $p < 0.0001$) and the nuclear level ($F_{ST} = 0.347$; $p < 0.0001$).
309 Texas caracaras were also highly differentiated from Florida's caracaras, with a mitochondrial
310 F_{ST} of 0.857 ($p < 0.0001$) and nuclear F_{ST} of 0.211 ($p < 0.0001$). The lowest differentiation was

311 observed between Arizona and Texas caracaras, with an F_{ST} of 0.183 at the nuclear level ($p <$
312 0.001) and F_{ST} of 0.091 at the mitochondrial level ($p = 0.12$).

313 The phylogenetic tree (Figure 4A) and haplotype network (Figure 4B) based on mitochondrial
314 *ND2* haplotypes revealed that the two Florida haplotypes formed a well-supported monophyletic
315 group (posterior probability = 0.95). The tree showed this "Florida subclade" nested within a
316 larger clade containing the shared Arizona/Texas haplotype and one other Texas haplotype. This
317 larger clade likely corresponds to northern crested caracara haplotypes (previously *C. cheriway*;
318 supported by the GenBank matches from Ecuador and North America); however, this clade was
319 not well-supported in our tree (posterior probability = 0.50). The tree revealed a second well-
320 supported clade (posterior probability = 0.97) primarily corresponding to haplotypes from
321 southern crested caracaras (supported by GenBank sequences from Argentina and Paraguay).
322 However, the southern crested caracara clade also contains one of the Texas haplotypes
323 (CRCA5), which in the haplotype network also presents as the most divergent of the five
324 haplotypes we identified in our samples, with seven substitutions separating it from the nearest
325 haplotype (CRCA3; Figure 4B).

326 **Discussion**

327 This study suggests that crested caracaras in Florida are genetically distinct from sampled
328 populations in Arizona and Texas. This finding is consistent with observations from banding and
329 telemetry data, which also indicate no movement of caracaras between Florida and the rest of the
330 U.S. These results support the hypothesis that crested caracaras have experienced long-term
331 isolation in Florida, and this history is reflected in the lower genetic diversity of this population.

332 Peninsular Florida has many other isolated populations and endemic species due to its unique
333 biogeography. For example, Florida populations of burrowing owls, snail kites (*Rostrhamus*
334 *sociabilis plumbeus*), and grasshopper sparrows (*Ammodramus savannarum floridanus*) are
335 imperiled and disjunct from more widely distributed subspecies in the Americas (Korfanta *et al.*
336 2005; Haas *et al.* 2009; Morgan and Emslie 2010; Bulgin *et al.* 2003). The relict Florida panther
337 (*Puma concolor coryi*) represents a more extreme case of genetic erosion (Roelke *et al.* 1993;
338 Culver *et al.* 2000), with such severe inbreeding depression and low population size that
339 supplementation with pumas from Texas was deemed necessary as a "genetic rescue" of the
340 population (Johnson *et al.* 2010). While the Florida panther's brush with extinction offers an
341 example of acute short-term crisis management for a small population without gene flow, long-
342 term habitat protections in Florida likely offer better long-term multi-species approaches to
343 preventing Florida's threatened and isolated fauna and flora populations from extinction.

344 *Population Structure and Differentiation*

345 The PCA, DAPC, ADMIXTURE, and F_{ST} results support two major clusters among caracara
346 populations occurring in the southern U.S. – one consisting of Arizona and Texas individuals and
347 the other consisting only of Florida individuals. These findings suggest that caracara populations
348 in Arizona and Texas have experienced more recent connectivity and moderate gene flow and
349 are highly differentiated from Florida's caracara population. This differentiation supports the
350 conclusion that the Florida population is closed and has been isolated for some time.
351 Furthermore, our population structure analyses indicate the lack of barriers to gene flow within
352 the Florida population. The lack of strong IBD further supports that Florida's caracaras exist as
353 one genetically panmictic population with potential for mating across the range in Florida, as

354 suggested by telemetry data showing that non-breeding caracaras range widely in Florida (Dwyer
355 *et al*. 2013).

356 Although the Texas population appeared to show admixture with Florida in some analyses, this
357 is more likely due to historical gene flow. The lack of mitochondrial haplotype sharing between
358 Texas and Florida individuals suggests modern gene flow does not occur between these
359 populations. Comparison with samples of other populations across the northern crested caracara
360 range, along with sampling of wild Texas caracaras, may be needed to clarify ancestral
361 relationships among populations.

362 *Genetic Diversity*

363 Our estimates of genetic diversity appear similarly low among caracara populations in Arizona,
364 Texas, and Florida, as represented by our samples, compared to estimates reported in a previous
365 study that developed 14 polymorphic microsatellite markers in 20 Florida caracaras (H_O = 0.41;
366 Vaughn *et al.* 2015); however, this does not indicate a reduction in genetic variation between the
367 time of the two studies. Rather, we believe this reflects differences in levels of polymorphism of
368 genetic marker types. Low heterozygosity estimates are often obtained from large SNP datasets,
369 which may have many low-frequency SNPs (Sunde *et al.* 2020; Zimmerman *et al.* 2020),
370 causing a lower magnitude of diversity relative to highly polymorphic markers. However,
371 genetic diversity and differentiation metrics estimated with SNPs and microsatellites tend to be
372 correlated, despite differences in magnitude (Zimmerman *et al.* 2020), and the heterozygosity
373 estimates reported here are within the range observed for other raptors using genomic data
374 (Doyle *et al.* 2019). The lower-than-expected heterozygosity in Florida's caracaras, and reduced
375 levels of diversity relative to other caracara populations, are likely due to isolation and relatively

376 small size of the Florida population. Given the current limited range of the species in Florida,
377 these values, along with the relatively small N_e, compared to proposed minimum N_e thresholds of
378 500 and 1000 for maintenance of genetic diversity (Franklin 1980; Frankham *et al*. 2014), may
379 be cause for concern regarding long-term population persistence. Habitat throughout the Florida
380 range continues to be lost (Smith *et al.* 2017), and a closed population will continue to lose
381 genetic diversity (Woodruff 2001).

382 *Mitochondrial Lineages*

383 Monophyly of haplotypes within the Florida population is apparent, in contrast to haplotype
384 sharing observed between Arizona and Texas populations. The absence of Arizona or Texas
385 haplotypes in the more than two hundred Florida caracaras sampled suggests female-mediated
386 gene flow does not or only rarely occurs between caracaras in Florida and these other states. We
387 also did not observe Florida haplotypes in the Arizona or Texas caracaras sampled, although we
388 were limited by the smaller number of samples from those states. These findings reflect the long-
389 term isolation of maternal lineages within the Florida population and further support the
390 conclusion of a closed population.

391 We detected a high amount of heteroplasmy at a single site for both haplotypes identified within
392 the Florida population. Heteroplasmy for both these haplotypes has been detected previously
393 (GenBank: MN231451, a caracara from Florida), and heteroplasmy due to paternal leakage has
394 been observed in other avian species (Kvist *et al.* 2003; Gandolfi *et al.* 2017). We hypothesize
395 that heteroplasmy in Florida caracaras may be due to paternal leakage, initially, and subsequent
396 inheritance from heteroplasmic mothers. In a small, isolated population, heteroplasmy
397 originating from one or few instances of paternal leakage could quickly rise in frequency as

398 heteroplasmic females breed (Parakatselaki and Ladoukakis 2021). Mechanisms allowing for
399 increased incidences of heteroplasmy may be evolutionarily advantageous for small populations
400 through increasing genetic diversity and adaptive potential and buffering against genetic
401 diversity loss associated with population bottlenecks (Tikochinski *et al.* 2020). However, the
402 heteroplasmic mutation detected is synonymous, so positive selection for heteroplasmy at this
403 site may be less likely.

404 The highly divergent haplotype (CRCA5) possessed by a single caracara from Texas calls into
405 question this bird's origin. We did not include this individual in a ddRAD library, so its ancestry
406 proportions are unknown. The phylogenetic tree places this haplotype in a clade with other
407 haplotypes from southern crested caracaras. Based on telemetry data of breeding adult caracaras
408 (Morrison 2003) and non-breeding caracaras (Dwyer *et al.* 2013), and the absence of band
409 recoveries outside of Florida, the crested caracara is considered non-migratory. Populations of
410 the formerly designated northern and southern crested caracaras have limited range overlap in
411 northern South America, so gene flow between Arizona or Texas and the more southern
412 populations of crested caracaras is unlikely. However, recent data from telemetered caracaras in
413 Arizona indicate that some individuals may travel longer distances than previously thought, with
414 movements recorded between Arizona and Central Mexico (Morrison, unpublished data),
415 perhaps enhancing gene flow among southern U.S. and central American populations. This is
416 consistent with our findings that mitochondrial haplotypes are shared among Arizona, Texas,
417 Mexico, Nicaragua, and Ecuador (based on origins of sequences on GenBank), although post-
418 Pleistocene expansion could also explain haplotype sharing across North and Mesoamerican
419 populations. The history of the Texas individual in captivity is unknown, but it is possible its
420 origin was somewhere in the southern part of the caracara's range, and it may have been

421 transported into Texas. Alternatively, this divergent haplotype may be present in the Texas
422 population due to incomplete lineage sorting. Obtaining more samples from wild Texas caracaras
423 for future analyses could help resolve occurrence of this haplotype there.

424 *Crested Caracara Systematics*

425 There have been multiple changes to the taxonomic designation of crested caracaras over time.
426 Northern and southern crested caracaras were considered one species (Hellmayr and Conover
427 1949) based on minimal morphological differences, but Dove and Banks (1999) concluded the
428 northern crested caracara (*C. cheriway*) and southern crested caracara (*C. plancus*) were distinct
429 species based on variation in plumage and morphometric characteristics as well as limited range
430 overlap in northern South America (Morrison and Dwyer 2023). Recently, the American
431 Ornithological Society (Chesser *et al.* 2021) and International Ornithologists' Union (Gill *et al.*
432 2021) again merged the two species as *Caracara plancus* based on low genetic divergence
433 (Fuchs *et al.* 2012). These authors used seven autosomal markers and a 2.4 kb mitochondrial
434 region spanning multiple genes in their phylogenetic study of the subfamily Polyborinae.
435 However, as the authors noted, the information is limited for broad taxonomic inferences
436 because only one *C. plancus* individual and five *C. cheriway* individuals were included.

437 Our analyses of mitochondrial DNA suggest two major subclades of the crested caracara, likely
438 corresponding to what were previously considered northern and southern crested caracaras,
439 based on locations of our sequenced samples and reference sequences from GenBank. The
440 overall crested caracara clade and northern crested caracara subclade were not well-supported in
441 our tree, although other studies have found support for these clades (Fuchs *et al.* 2012; Oswald *et*
442 *al.* 2019). We believe this is likely due to identifying few polymorphic sites among all crested

443 caracara haplotypes we analyzed, resulting in lower resolution for phylogenetic analysis. Fuchs
444 *et al.* (2012) also found southern and northern crested caracaras formed well-supported separate
445 monophyletic groups with both mitochondrial and nuclear DNA. Together these results may
446 suggest reproductive isolation between northern and southern crested caracaras, a result
447 supported by the limited number of intermediate morphological characters and narrow range of
448 overlap reported by Dove and Banks (1999). While our analyses show that the Florida
449 population is genetically distinct from other U.S. populations, we were unable to sample the
450 species' range outside the U.S., and much of the crested caracara's intraspecific variation is left
451 to be described. Clarifying crested caracara taxonomy utilizing nuclear and mitochondrial
452 genomic data from a large number of individuals sampled across the species' range, including
453 the contact zone of putative species *C. plancus* and *C. cheriway*, is a high-priority objective of
454 future research.

455 *Conservation Implications*

456 Our results indicate that the Florida population of crested caracaras represents a distinct
457 evolutionary lineage, having experienced long-term isolation from other caracara populations.
458 The monophyly of Florida haplotypes in our mitochondrial *ND2* phylogeny, along with
459 population structure results indicating a high degree of isolation between Florida and allopatric
460 populations in Arizona and Texas, provide support for consideration of Florida's crested caracara
461 population as an Evolutionarily Significant Unit (Ryder 1986; Moritz 1994) and possible
462 designation as a separate subspecies (O'Brien and Mayr 1991). Irrespective of potential changes
463 in taxonomy, this population is consistent with the definition of a Distinct Population Segment
464 (DPS) under the ESA. Conservation and management actions which prioritize this isolated

465 population's unique genetic diversity and evolutionary potential will help ensure the long-term
466 persistence of this population.

467 **Acknowledgements**

468 We thank Erin Hahn for completing many DNA extractions. We are grateful to Daryl Trumbo and the
469 Funk lab for their advice in ddRADseq experimental design. We thank Karla Vargas, Eldridge Wisely,
470 Hans-Werner Herrman, and Leigh Combrink for providing analytical advice and support during
471 manuscript preparation. We thank Robert Fitak for critical review of this manuscript for USGS, and we
472 are grateful to Robert Fleischer, Jérôme Fuchs, and one anonymous reviewer, whose recommendations
473 improved the final manuscript. We also thank Alicia Mastretta-Yanes for assistance with using R code for
474 error calculations. We are grateful to the Worobey lab, especially Tom Watts and Grace Quirk, for
475 allowing us to use the Pippin Prep DNA size selection system. Thanks to Doug Jenness and Rich Glinski
476 for their assistance obtaining blood samples from Arizona's caracaras. We thank personnel at the
477 following facilities for providing the samples from Texas caracaras: John Ball Zoo, Caldwell Zoo,
478 Cameron Park Zoo, and Gladys Porter Zoo.

479 Vertebrates were captured and blood samples were taken by Joan L. Morrison and James F. Dwyer under
480 U.S. Fish and Wildlife Service Threatened and Endangered Species Section 10(a)(1)(A) Recovery Permit
481 No. TE106708-3, Federal Master Bird Banding Permit No. 22800, and State of Arizona (Game and Fish
482 Department) Scientific Activity License No. SP651715. A separate Florida state (Fish and Wildlife
483 Commission) permit for capturing and handling crested caracaras was not required.

484 **Funding**

485 This work was funded by the Fish and Wildlife Foundation of Florida.

486 **DISCLAIMERS**

487 Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement
488 by the U.S. Government.
489
490 A names review of scientific and common names, place names, Federally recognized Tribes, and geologic
491 names was completed for this manuscript.
492
493
494 Although this code has been processed successfully on a computer system at the U.S. Geological Survey
495 (USGS), no warranty expressed or implied is made regarding the display or utility of the data for other
496 purposes, nor on all computer systems, nor shall the act of distribution constitute any such warranty. The
497 USGS or the U.S. Government shall not be held liable for improper or incorrect use of the data described
498 and/or contained herein.
499

500 **Data Availability**

501 Sample locations, genotypes, and ddRADseq reads will be deposited in Dryad. Mitochondrial haplotypes
502 are available on GenBank (accessions OR340981– OR340985).

503 ## References

504 Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated
505 individuals. *Genome Res* 2009;19:1655–1664.

506 Andrews S. FastQC: A Quality Control Tool for High Throughput Sequence Data [Online].
507 http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (2010).

508 Bandelt HJ, Forster P, Röhl A. Median-joining networks for inferring intraspecific
509 phylogenies. *Mol Biol Evol* 1999;16:37–48.

510 Bassett SM, Potter MA, Fordham RA, Johnston E V. Genetically identical avian twins. *J*
511 *Zool* 1999;247:475–478.

512 Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence
513 data. *Bioinformatics* 2014;30:2114–2120.

514 Brodkorb P. Pleistocene avifauna of Arredondo, Florida. *Bulletin of the Florida State*
515 *Museum, Biological Sciences* 1959;4:269-291.

516 Bulgin NL, Gibbs HL, Vickery P, Baker AJ. Ancestral polymorphisms in genetic markers
517 obscure detection of evolutionarily distinct populations in the endangered Florida
518 grasshopper sparrow (*Ammodramus savannarum floridanus*). *Mol Ecol* 2003;12:831–
519 844.

520 Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA. Stacks: an analysis tool set
521 for population genomics. *Mol Ecol* 2013;22:3124–3140.

522 Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH. Stacks: Building and
523 genotyping loci de novo from short-read sequences. *G3 Genes, Genomes, Genet*
524 2011;1:171–182.

525 Catchen J, Rochette N, Cresko WA, Hohenlohe PA, Amores A, *et al.* Stacks: Stacks
526 Manual. https://catchenlab.life.illinois.edu/stacks/manual/ (accessed 22 February
527 2021).

528 Chesser RT, Billerman SM, Burns KJ, Cicero C, Dunn JL, *et al.* Sixty-second Supplement
529 to the American Ornithological Society's Check-list of North American Birds.
530 *Ornithology* 2021;138:1–18.

531 Culver M, Johnson WE, Pecon-Slattery J, O'Brien SJ. Genomic ancestry of the American
532 puma (*Puma concolor*). *J Hered* 2000;91:186–197.

533 Darriba D, Taboada GL, Doallo R, Posada D. JModelTest 2: More models, new heuristics
534 and parallel computing. *Nature Methods* 2012;9:772.

535 Desmond MJ, Parsons TJ, Powers TO, Savidge JA. An initial examination of mitochondrial
536 dna structure in Burrowing Owl populations. *J Raptor Res* 2001;35:274–281.

537 Do C, Waples RS, Peel D, Macbeth GM, Tillett BJ, *et al.* NeEstimator v2: Re-
538 implementation of software for the estimation of contemporary effective population
539 size (Ne) from genetic data. *Mol Ecol Resour* 2014;14:209–214.

540 Dove CJ, Banks RC. A taxonomic study of crested caracaras (Falconidae). *Wilson Bull*
541 1999;111:330–339.

542 Doyle JM, Willoughby JR, Bell DA, Bloom PH, Bragin EA, *et al.* Elevated heterozygosity
543 in adults relative to juveniles provides evidence of viability selection on eagles and
544 falcons. *J Hered* 2019;110:696–706.

545 Dwyer JF, Fraser JD, Morrison JL. Evolution of communal roosting: A social refuge–
546 territory prospecting hypothesis. *J Raptor Res* 2018;52:407–419.

547 Dwyer JF, Fraser JD, Morrison JL. Range sizes and habitat use of non-breeding Crested
548 Caracaras in Florida. *J F Ornithol* 2013;84:223–233.

549 Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high
550 throughput. *Nucleic Acids Res* 2004;32:1792–1797.

551 Emslie SD. Avian Community, Climate, and Sea-Level Changes in the Plio-Pleistocene of
552 the Florida Peninsula. *Ornithol Monogr* 1998;50:1–113.

553 England PR, Cornuet JM, Berthier P, Tallmon DA, Luikart G. Estimating effective
554 population size from linkage disequilibrium: Severe bias in small samples. *Conserv*
555 *Genet* 2006;7:303–308.

556 Excoffier L, Lischer HEL. Arlequin suite ver 3.5: A new series of programs to perform
557 population genetics analyses under Linux and Windows. *Mol Ecol Resour*
558 2010;10:564–567.

559 Francis RM. pophelper: an R package and web app to analyse and visualize population
560 structure. *Mol Ecol Resour* 2017;17:27–32.

561 Frankham R, Bradshaw CJA, Brook BW. Genetics in conservation management: Revised
562 recommendations for the 50/500 rules, Red List criteria and population viability
563 analyses. *Biol Conserv* 2014;170:56–63.

564 Franklin IR. Evolutionary changes in small populations. In: Soule ME, Wilcox BA
565 (editors). Conservation biology: an evolutionary-ecological prospective. Sunderland,
566 Mass.: Sinauer Associates; 1980. pp. 135–149.

567 Fuchs J, Johnson JA, Mindell DP. Molecular systematics of the caracaras and allies
568 (Falconidae: Polyborinae) inferred from mitochondrial and nuclear sequence data. *Ibis*
569 2012;154:520–532.

570 Gandolfi A, Crestanello B, Fagotti A, Simoncelli F, Chiesa S, et al. New Evidences of
571 Mitochondrial DNA Heteroplasmy by Putative Paternal Leakage between the Rock
572 Partridge (*Alectoris graeca*) and the Chukar Partridge (*Alectoris chukar*). *PLoS One*
573 2017;12:e0170507.

574 Gill F, Donsker D, Rasmussen P. IOC World Bird List (v. 11.1) Proposed Splits/Lumps.
575 https://www.worldbirdnames.org/new/updates/proposed-splits/ (2021, accessed 22
576 February 2021).

577 Gilpin ME, Soulé ME. Minimum viable populations: processes of species extinction. In:
578 Soulé ME (editor). Conservation biology: the science of scarcity and diversity.
579 Sunderland, Mass.: Sinauer Associates; 1986. pp. 19–34.

580 Gosselin T, Lamothe M, Devloo-Delva F, Grewe P. radiator: RADseq Data Exploration,
581 Manipulation and Visualization using R. 2020.
582 https://thierrygosselin.github.io/radiator/

583 Haas SE, Kimball RT, Martin J, Kitchens WM. Genetic divergence among Snail Kite
584 subspecies: implications for the conservation of the endangered Florida Snail Kite
585 *Rostrhamus sociabilis*. *Ibis* 2009;151:181–185.

586 Hellmayr CE, Conover B. Catalogue of Birds of the Americas and the Adjacent Islands.
587 *Zool Ser, F Museum Hist* 1949;13, Part 1:1–358.

588 Johnson W, Onorato D, Roelke M, Land E, Cunningham M, et al. Genetic restoration of the
589 Florida panther. *Science* 2010;329:1641–1645.

590 Jombart T. Adegenet: A R package for the multivariate analysis of genetic markers.
591 *Bioinformatics* 2008;24:1403–1405.

592 Jones AT, Ovenden JR, Wang YG. Improved confidence intervals for the linkage
593 disequilibrium method for estimating effective population size. *Heredity*
594 2016;117:217–223.

595 Korfanta NM, McDonald DB, Glenn TC. Burrowing owl (*Athene cunicularia*) population
596 genetics: A comparison of North American forms and migratory habits. *Auk*
597 2005;122:464–478.

598 Kvist L, Martens J, Nazarenko AA, Orell M. Paternal Leakage of Mitochondrial DNA in
599 the Great Tit (Parus major). *Mol Biol Evol* 2003;20:243–247.

600 Layne JN. Audubon's Crested Caracara. In: Rogers, Jr. JA, Kale II HW, Smith HT (editors).
601 *Rare and endangered biota of Florida. Vol V: Birds*. Gainesville: University of Florida
602 Press; 1996. pp. 197–210.

603 Leigh JW, Bryant D. Popart: full-feature software for haplotype network construction.
604 *Methods Ecol Evol* 2015;6:1110–1116.

605 Lischer HEL, Excoffier L. PGDSpider: an automated data conversion tool for connecting
606 population genetics and genomics programs. *Bioinformatics* 2012;28:298–299.

607 Macías-Duarte A, Conway CJ, Holroyd GL, Valdez-Gómez HE, Culver M. Genetic
608 Variation among Island and Continental Populations of Burrowing Owl (*Athene*
609 *cunicularia*) Subspecies in North America. *J Raptor Res* 2019;53:127–133.

610 Mastretta-Yanes A, Arrigo N, Alvarez N, Jorgensen TH, Piñero D, *et al.* Restriction site-
611 associated DNA sequencing, genotyping error estimation and de novo assembly
612 optimization for population genetic inference. *Mol Ecol Resour* 2015;15:28–41.

613 McDonald DB, Potts WK, Fitzpatrick JW, Woolfenden GE. Contrasting genetic structures
614 in sister species of North American scrub-jays. *Proc R Soc B Biol Sci* 1999;266:1117–
615 1125.

616 Morgan GS, Emslie SD. Tropical and western influences in vertebrate faunas from the
617 Pliocene and Pleistocene of Florida. *Quat Int* 2010;217:143–158.

618 Moritz C. Defining 'Evolutionarily Significant Units'. *Trends Ecol Evol* 1994;9:373–375.

619 Morrison JL. Age-Specific Survival of Florida's Crested Caracaras. *J F Ornithol*
620 2003;74:321–330.

621 Morrison JL, Humphrey SR. Conservation Value of Private Lands for Crested Caracaras in
622 Florida. *Conserv Biol* 2001;15:675–684.

623 Morrison JL, Dwyer JF. Crested Caracara (*Caracara plancus*), version 1.1. In: Sly N
624 (editor). Birds of the World. Ithaca, NY, USA: Cornell Lab of Ornithology; 2023.

625 Myers RL, Ewel JJ. Ecosystems of Florida. Orlando, FL, USA: University of Central
626 Florida Press; 1990.

627 O'Brien SJ, Mayr E. Bureaucratic mischief: Recognizing endangered species and
628 subspecies. *Science* 1991;251:1187–1188.

629 Oswald JA, Allen JM, Witt KE, Folk RA, Albury NA, *et al.* Ancient DNA from a 2,500-
630 year-old Caribbean fossil places an extinct bird (Caracara creightoni) in a phylogenetic
631 context. *Mol Phylogenet Evol* 2019;140:106576.

632 Parakatselaki ME, Ladoukakis ED. mtDNA Heteroplasmy: Origin, Detection, Significance,
633 and Evolutionary Consequences. *Life* 2021;11: 633.

634 Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE. Double Digest RADseq: An
635 Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-
636 Model Species. *PLoS One* 2012;7:e37135.

637 Pew J, Muir PH, Wang J, Frasier TR. related: an R package for analysing pairwise
638 relatedness from codominant molecular markers. *Mol Ecol Resour* 2015;15:557–561.

639 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, *et al.* PLINK: A tool set for
640 whole-genome association and population-based linkage analyses. *Am J Hum Genet*
641 2007;81:559–575.

642 Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. Posterior Summarization in
643 Bayesian Phylogenetics Using Tracer 1.7. *Syst Biol* 2018;67:901–904.

644 Rochette NC, Catchen JM. Deriving genotypes from RAD-seq short-read data using Stacks.
645 *Nat Protoc* 2017;12:2640–2659.

646 Rochette NC, Rivera-Colón AG, Catchen JM. Stacks 2: Analytical methods for paired-end
647 sequencing improve RADseq-based population genomics. *Mol Ecol* 2019;28:4737–
648 4754.

649 Roelke ME, Martenson JS, O'Brien SJ. The consequences of demographic reduction and
650 genetic depletion in the endangered Florida panther. *Curr Biol* 1993;3:340–350.

651 Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, *et al.* MrBayes 3.2:
652 efficient Bayesian phylogenetic inference and model choice across a large model
653 space. *Syst Biol* 2012;61:539–542.

654 Rozas J, Ferrer-Mata A, Sanchez-DelBarrio JC, Guirao-Rico S, Librado P, *et al.* DnaSP 6:
655 DNA sequence polymorphism analysis of large data sets. *Mol Biol Evol*
656 2017;34:3299–3302.

657 Ryder OA. Species conservation and systematics: the dilemma of subspecies. *Trends Ecol*
658 *Evol* 1986;1:9–10.

659 Smith JA, Dwyer JF, Fraser JD, Morrison JL. A habitat model to aid the conservation of
660 crested caracaras. *J Wildl Manage* 2017;81:712–719.

661 Sorenson MD, Ast JC, Dimcheff DE, Yuri T, Mindell DP. Primers for a PCR-Based
662 Approach to Mitochondrial Genome Sequencing in Birds and Other Vertebrates. *Mol*
663 *Phylogenet Evol* 1999;12:105–114.

664 Sunde J, Yıldırım Y, Tibblin P, Forsman A. Comparing the Performance of Microsatellites
665 and RADseq in Population Genetic Studies: Analysis of Data for Pike (Esox lucius)
666 and a Synthesis of Previous Studies. *Front Genet* 2020;11:218.

667 Tagliarini MM, Nagamachi CY, Pieczarka JC, Correa De Oliveira EH. Description of two
668 new karyotypes and cytotaxonomic considerations on Falconiformes. *Rev Bras Ornitol*
669 2007;15:261–266.

670 Tikochinski Y, Carreras C, Tikochinski G, Vilaça ST. Population-specific signatures of
671 intra-individual mitochondrial DNA heteroplasmy and their potential evolutionary
672 advantages. *Sci Rep* 2020;10:211.

673 Trumbo DR, Salerno PE, Logan KA, Alldredge MW, Gagne RB, *et al.* Urbanization
674 impacts apex predator gene flow but not genetic diversity across an urban-rural divide.
675 *Mol Ecol* 2019;28:4926–4940.

676 U.S. Fish and Wildlife Service. Endangered and Threatened Wildlife and Plants; Threatened
677 Status for the Florida Population of Audubon's Crested Caracara. *Federal Register*
678 1987;52:25229–25232.

679 Vaughn EE, Dwyer JF, Morrison JL, Culver M. Development and characterization of
680 polymorphic microsatellite markers for the crested caracara, Caracara cheriway.
681 *Conserv Genet Resour* 2015;7:557–559.

682 Waples RK, Larson WA, Waples RS. Estimating contemporary effective population size in
683 non-model species using linkage disequilibrium across thousands of loci. *Heredity*
684 2016;117:233–240.

685 Woodruff DS. Populations, Species, and Conservation Genetics. In: Levin SA (editor).
686 *Encyclopedia of Biodiversity*. New York: Elsevier; 2001. pp. 811–829.

687 Zimmerman SJ, Aldridge CL, Oyler-Mccance SJ. An empirical comparison of population
688 genetic analyses using microsatellite and SNP data for a species of conservation
689 concern. *BMC Genomics* 2020;21:1–16.

690

691

692

693

	ddRAD SNP Data					Mitochondrial Data			
Population	N	H_O	H_E	π	F_{IS}	N	H	H_d	π
Florida	184	0.23212	0.23433	0.00082	0.01512	238	2	0.148	0.00021
Arizona	8	0.25440	0.23712	0.00089	0.00497	10	1	0	0
Texas	5	0.27954	0.25669	0.00102	0.02504	6	3	0.6	0.00418

694

695 **Table 1.** Summary statistics for populations of the crested caracara in Florida, Arizona, and
696 Texas (N = individuals sampled). Statistics calculated using SNPs from ddRADseq were the
697 observed heterozygosity (H_O), expected heterozygosity (H_E), and the inbreeding coefficient (F_{IS}),
698 and nucleotide diversity (π) was calculated using variant and invariant sites. The number of
699 haplotypes (H), haplotype diversity (H_d), and nucleotide diversity (π) were calculated with
700 mitochondrial haplotypes.
701

Below Diagonal: mtDNA F_{ST}	Above Diagonal: ddRAD F_{ST}			
		FL	TX	AZ
	FL		0.21137**	0.34673**
	TX	0.85713**		0.18337*
	AZ	0.87544**	0.09091	

702 **Table 2.** Pairwise values of genetic differentiation (F_{ST}) calculated among caracara populations
703 in Florida, Texas, and Arizona. Florida's caracara population appears highly differentiated from
704 populations in both Texas and Arizona, while populations in Texas and Arizona are moderately
705 differentiated. Significance is denoted by asterisks: * $p < 0.001$, ** $p < 0.0001$.

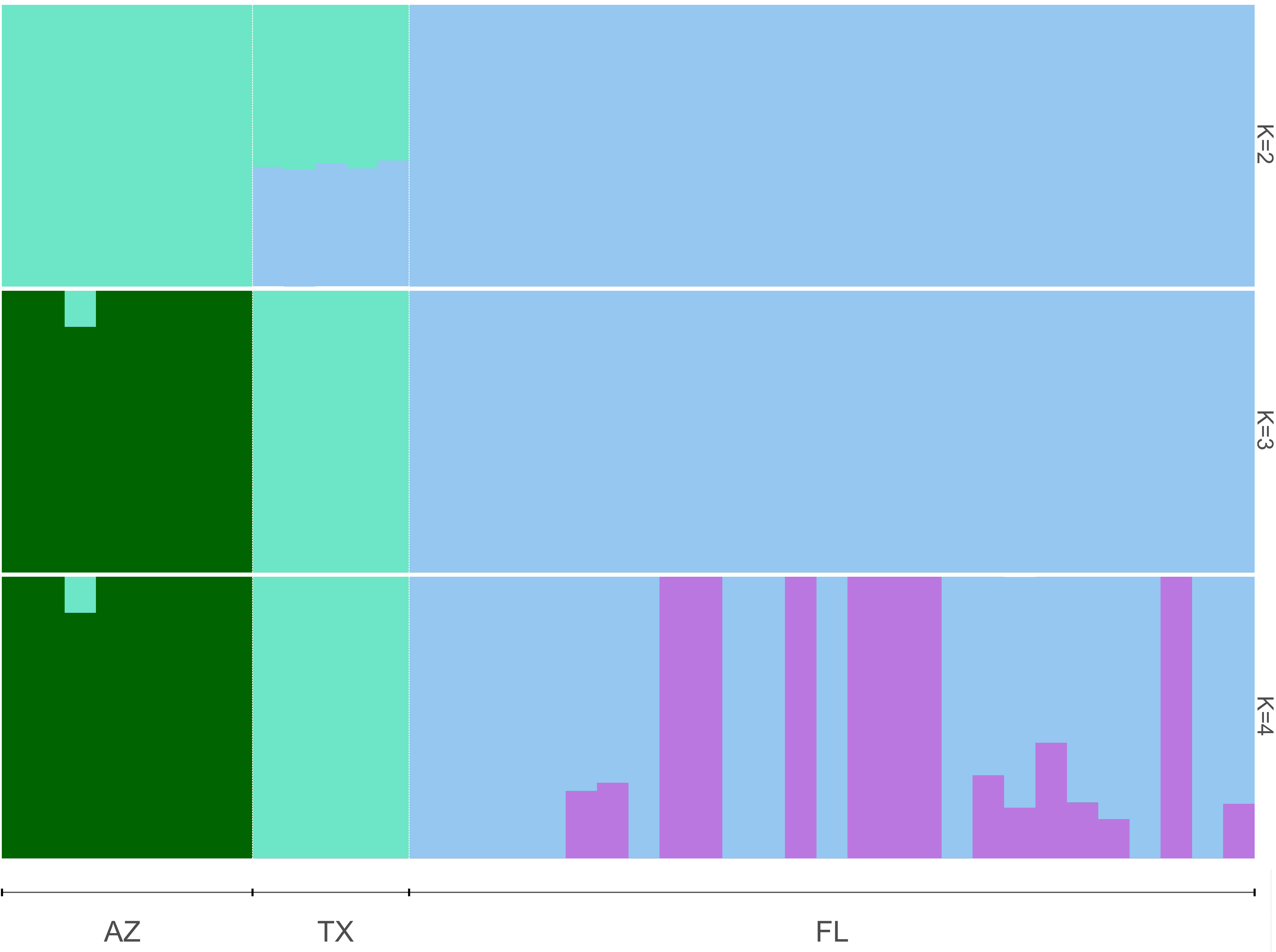
706

707 **Figure 1.** Map of sampling sites of wild caracaras in Florida (green). Sites for samples included
708 in ddRAD libraries or with the 717-bp mitochondrial haplotype sequenced are shown, colored
709 based on sampling year.
710

711 **Figure 2.** Results of Principal Components Analysis (PCA) and Discriminant Analysis of
712 Principal Components (DAPC). (A) The PCA shows three clusters, corresponding to state
713 populations (Arizona, Texas, and Florida). Principal component 1 is on the x-axis (7.28% of
714 variation), and principal component 2 is on the y-axis (1.59% of variation). (B) Prior to
715 performing the DAPC, the find.clusters command in *adegenet* showed that K = 2 had the lowest
716 Bayesian Information Criterion (BIC) value and was therefore most supported. (C) The DAPC
717 scatter plot showing the results of performing the DAPC with K = 2. One cluster was composed
718 of only Florida samples (light blue), and the other was composed of Arizona and Texas samples
719 (light green). (D) The composition plot shows the group assignment probability for each
720 individual to the DAPC clusters. Each individual had a membership probability of 1.0 for its
721 assigned cluster. Together, these results suggest two major groups, corresponding to Florida and
722 Arizona/Texas, with Arizona and Texas being more recently connected populations.

723 **Figure 3.** Results of ancestry analyses using ADMIXTURE for K = 2 – 4, using the subset of 40
724 individuals (AZ = 8, TX = 5, FL = 27 randomly selected). K = 2 was best supported, followed
725 sequentially by K = 3 and K = 4. The plot for K = 2 shows Florida individuals have ancestry
726 from one cluster and Arizona and Texas individuals predominantly have ancestry from a second
727 cluster. At K = 3, Texas individuals form a separate cluster, and at K = 4, ADMIXTURE
728 identified close relatives in the Florida population, rather than subpopulations.

729 **Figure 4.** Mitochondrial phylogenetic tree and haplotype network. (A) 50% majority-rule
730 consensus tree with posterior probabilities generated by MrBayes for the five haplotypes
731 identified in our samples and additional reference and outgroup sequences from GenBank. The
732 Florida haplotypes form a distinct monophyletic group (highlighted in blue). This Florida
733 subclade is within a clade that appears to correspond to northern crested caracaras (previously *C.*
734 *cheriway*); however, this clade is not well-supported in our tree. The tree additionally shows a
735 well-supported clade corresponding primarily to southern crested caracaras (highlighted in
736 orange). Information on *C. plancus* reference sequences obtained from GenBank, including
737 geographic origin and voucher number, is included in the table. (B) Median-joining network
738 showing the five haplotypes identified in our samples. Circle size corresponds to the number of
739 sequenced individuals with the haplotype, and hash marks represent the number of genetic
740 changes between haplotypes. Haplotypes 1 and 2 (CRCA1 and CRCA2) were found only in
741 Florida, haplotype 3 (CRCA3) was found in both Arizona and Texas, and haplotypes 4 and 5
742 (CRCA4 and CRCA5) were found only in Texas.
743



Cluster 1
Cluster 2
Cluster 3
Cluster 4
K=2
K=3
K=4
AZ
TX
FL