1   Rediscovery of red wolf ghost alleles in a canid population along the American Gulf

2   Coast

3

4   Elizabeth Heppenheimer[1†], Kristin E. Brzeski[1,2*†], Ron Wooten[3], Will Waddell[4], Linda Y.

5   Rutledge[5], Michael J. Chamberlain[6], Daniel R. Stahler[7], Joseph W. Hinton[6], Bridgett M.

6   vonHoldt[1*]

7

8   [1]Princeton University, Ecology and Evolutionary Biology, Princeton, NJ 08544;

9   [2]Michigan Technological University, School of Forest Resources & Environmental

10   Science, Houghton, MI 49931; [3]Wildlife Biologist, Galveston, TX 77550; [4]Point

11   Defiance Zoo & Aquarium, Zoological & Environmental Education Department,

12   Tacoma, WA 98407; [5]Trent University, Environmental & Life Sciences, Peterborough,

13   Ontario, Canada K9L 0G2; [6]Warnell School of Forestry and Natural Resources,

14   University of Georgia, Athens, GA 30602; [7]Yellowstone Center for Resources, National

15   Park Service, Yellowstone National Park, WY 82190

16

17   †These authors contributed equally to this work

18   *Corresponding authors: *kbrzeski@mtu.edu* and *vonholdt@princeton.edu*

19

20   **Abstract**

21   Rediscovering species once thought to be extinct or on the edge of extinction is rare. Red

22   wolves have been extinct along the Gulf Coast region since 1980, with their last

23   populations found in coastal Louisiana and Texas. We report the rediscovery of red wolf

24   ghost alleles in a canid population on Galveston Island, Texas. We analyzed over 7,000

25   SNPs in 60 canid representatives from all legally recognized North American *Canis*

26   species and two phenotypically ambiguous canids from Galveston Island. We found

27   notably high Bayesian cluster assignments of the Galveston canids to captive red wolves

28   with extensive sharing of red wolf private alleles. Today, the only known extant wild red

29   wolves persist in a reintroduced population in North Carolina, which is dwindling

30   amongst political and taxonomic controversy. Our rediscovery of red wolf ancestry after

31   almost 40 years introduces both positive opportunities for additional conservation action

32   and difficult policy challenges.

33

Introduction

34

35   Red wolves (*Canis rufus*) once roamed across the southeastern United States but were

36   declared extinct in the wild by 1980 due to habitat loss, predator control programs,

37   disease, and interbreeding with encroaching coyotes (*Canis latrans*). In 1967, the U.S.

38   Fish & Wildlife Service (USFWS) listed red wolves as endangered under the U.S.

39   Endangered Species Preservation Act due to their rapid population decline in the

40   American south, and subsequently were among the first species listed on the 1973

41   Endangered Species Act (ESA), the landmark of the Unites States' environmental law

42   (*1*). On the brink of extinction, red wolf recovery was initiated through the trapping of

43   what was believed to be the last wild red wolves that lived along the Gulf Coast of

44   Louisiana and Texas in the 1970s (*1-5*). Individuals were selected as founders for the

45   captive breeding program based on morphology and behavioral traits considered to be

46   species informative (*6,7*). Over 240 canids were trapped from coastal Louisiana and

47   Texas between 1973 and 1977 (*6*). Forty individuals were selected for captive breeding,

48   of which 17 were deemed 100% wolf. However, only 14 wolves successfully produced

49   litters, which comprised the first generation of captive wolves from which all red wolves

50   in the recovery program descend.

51   Due to the highly successful captive breeding program, red wolves were restored

52   to the landscape in North Carolina less than a decade after becoming extinct in the wild

53   (*6*). This historic event represented the first attempt to reintroduce a wild-extinct species

54   in the United States and set a precedent for returning wild-extinct wildlife to the

55   landscape. The success of the red wolf recovery program was the foundation upon which

56   other wolf introductions were guided, including the gray wolf (*C. lupus*) reintroduction to

57   the northern Rocky Mountains in Yellowstone National Park, Wyoming, and central

58   Idaho, and the ongoing restoration efforts for the Mexican wolves (*C. lupus baileyi*) in

59   the southwest (*8,9*). Although successful by many measures (*7*), the North Carolina

60   experimental population (NCEP) of red wolves was reduced by the USFWS in response

61   to negative political pressure from the North Carolina Wildlife Resource Commission

62   and a minority of private landowners (*10*). Further, gunshot-related mortalities have

63   increased the probability that wolf packs deteriorate before the breeding season, which

64   facilitates the establishment of coyote-wolf breeding pairs (*11,12*). Consequently, the

65   NCEP has fewer than 40 surviving members (*13*) and are once again on the brink of

66   extinction in the wild.

67   Interbreeding between red wolves and coyotes is well documented and is viewed

68   as a threat to red wolf recovery (*14*). When historic populations of red wolves along the

69   Gulf Coast were surveyed, it was feared that these coastal populations were the last

70   remnants of pre-recovery wild wolves and were likely to quickly become genetically

71   extinct through introgressive swamping of coyote genetics (*15*). Yet, there continued to

72   be reports of red wolves in rural coastal areas of Louisiana and Texas since the 1970s

73   where coyotes were not considered part of the local fauna (*5,16*). Previous efforts to

74   detect surviving red wolves or their hybrids in the region proved unsuccessful (*17*).

75   However, the possibility remains that individuals with substantial red wolf ancestry have

76   naturally persisted in isolated areas of the Gulf Coast. For example, body measurements

77   of coyote-like canids in southwestern Louisiana were similar to those of confirmed red

78   wolf-coyote hybrids in the NCEP (*18*). These individuals would harbor ghost alleles of

4

79   the original red wolves, with these alleles likely lost in the contemporary red wolf

80   population during the extreme population bottleneck, drift, and inbreeding.

81   For red wolf ghost alleles to persist, a remnant Gulf Coast population would need

82   to be relatively isolated from frequent interbreeding with coyotes (*14*). Although red

83   wolves that co-occur with coyotes in the NCEP exhibit assortative mating patterns (*19*),

84   an island would promote low rates of genetic exchange by providing a geographic barrier

85   and could be an ideal location for which red wolf ghost alleles would persist under

86   limited hybridization. We report evidence that Galveston Island, TX may represent one

87   such location. The ancestral population from which all contemporary red wolves descend

88   were trapped from Jefferson, Chambers, southern Orange, and eastern Galveston counties

89   in Texas and Cameron and southern Calcasieu parishes in Louisiana (*16*) (Fig. 1). Given

90   Galveston Island's location and isolation from the mainland, it is a probable region to

91   harbor red wolf ghost alleles. Recent images captured of Galveston Island canids (Fig. 2)

92   piqued interest of local naturalists and two genetic samples were taken from roadkill

93   individuals. We conducted genomic analyses and found evidence of red wolf ancestry in

94   modern day Galveston Island canids.

95

96   **Results**

97   *Genome-wide SNP genotyping and diversity estimates*

98   We collected genomic and mitochondrial (mtDNA) sequence data for two canids

99   inhabiting Galveston Island, Texas (GI) of unknown taxonomic origin and 60 reference

100   North American canids: 29 coyotes from the American southeast (Alabama, Louisiana,

101   Oklahoma, Texas), 10 gray wolves from Yellowstone National Park, 10 eastern wolves

5

102 (*C. lycaon*) from Algonquin Provincial Park in Ontario, and 11 red wolves from the

103 Special Survival Plan captive breeding program (hereafter, red wolves) that collectively

104 represent all extant red wolf founders (Table S1). We used a modified restriction-

105 associated DNA sequencing (RADseq) protocol (*20*) to discover 7,068 genome-wide

106 polymorphic SNPs after filtering for a minimum of three reads per allele, removing sites

107 in high statistical linkage disequilibrium, sites with a minor allele frequency of 1%, and

108 excluded sites with more than 10 % missing data.

109   We observed a significant difference in heterozygosity across all pairwise

110 combinations of reference groups, where coyotes exhibited the highest genomic diversity

111 of all reference groups (H_E: coyotes=0.101, gray wolf=0.076, eastern wolf=0.087, red

112 wolf=0.061; Tables 1, S1). In contrast, red wolves displayed the lowest levels of genomic

113 diversity (H_E=0.061), a trend consistent with expected erosion of diversity due to a

114 declining effective population size, a subsequent bottleneck, and inbreeding (*21*). We

115 found the highest levels of genomic differentiation between red and gray wolves

116 (F_{ST}=0.136), with the lowest levels found between coyotes and wolf comparisons (F_{ST}:

117 red wolf-eastern wolf=0.093, gray wolf-eastern wolf=0.086, coyote-gray wolf=0.062,

118 coyote-red wolf=0.040, coyote-eastern wolf=0.042).

119

120 *Genetic clusters and assignment analysis*

121   An initial principal component analysis (PCA) of all canids revealed that spatial

122 clusters were largely concordant with taxonomic classifications (Fig. 1A) and consistent

123 with previous genome-wide analyses of wolf-like canids (*22*). Specifically, PC1

124 explained 7.0% of the total variation, polarized by gray wolves and coyotes, with eastern

6

125   and red wolves spatially intermediate, and the two GI canids proximal to the coyote

126   cluster. The second axis was informative for red wolves (PC2, 3.9% variation), a pattern

127   previously observed and likely attributed to the captive breeding bottleneck, inbreeding,

128   and subsequent drift. When we restricted our analysis to only reference red wolves,

129   reference coyotes, and the two GI canids, we found that coyotes and red wolves again

130   defined PC1 (6.0% variation) with an intermediate placement of the two GI canids, and

131   PC2 (4.4% variation) reflected geographic variation within coyotes (Fig. S1A).

132   We used a maximum likelihood framework in ADMIXTURE (*23*) to assess

133   genetic structure across all 62 canids and found the greatest support for three genetic

134   clusters (K=3, cv=0.35) composed of gray and eastern wolves, coyotes, and red wolves,

135   respectively (Figs. 1C, S1B). The two GI canids exhibited partial memberships only to

136   the red wolf and coyote clusters at K=3 (canid GI-1: $Q_{Red\ Wolf}$=0.60, Q_{Coyote}=0.40; canid

137   GI-2: $Q_{Red\ Wolf}$=0.60, Q_{Coyote}=0.40). Interestingly, two coyotes that were collected in

138   Louisiana also exhibited non-trivial assignment proportions to the red wolf genetic

139   cluster ($Q_{Red\ Wolf}$: LA-2=0.10; LA-3=0.11). A similar analysis of the GI canids with

140   reference coyotes and red wolves showed strong support for two genetic clusters, which

141   mirrored the patterns revealed by PCA (Fig. S1). Cluster memberships were similar as

142   before, where clusters corresponded to a red wolf and coyote group, with the two GI

143   canids assigned to each cluster (Fig. S1C). When we surveyed higher levels of

144   partitioning, K=3 revealed two distinct groups of coyotes corresponding to their historic

145   range of Oklahoma and Texas and to their southeastern expansion front across Louisiana

146   and Alabama. At this level of partitioning, the two GI canids retained non-trivial

147   assignments to red wolves (K=3 $Q_{Red\ Wolf}$: GI-1=0.27, GI-2=0.21) (Fig. S2). Interestingly,

148   the proportion of coyote ancestry for the two GI canids was attributed to the recently

149   expanded southeastern coyote population ($Q_{\text{Southeast Coyote}}$: GI-1=0.73; GI-2=0.79).

150   Although no definitive conclusion can be drawn from this data regarding the origin on the

151   red wolf allele sharing among the two GI canids, assignments to the southeastern coyote

152   cluster, rather than the historic range cluster, is consistent with interbreeding between the

153   wild red wolf and expanding coyote populations in the late 1970s.

154   We obtained the probability that each GI canid shared ancestry with reference

155   coyotes, red wolves, or both using the posterior probability assignment test in

156   STRUCTURE (*24*), where each GI canid was explicitly assigned to one or more of the

157   coyote and red wolf reference groups. We found similar results, where each GI canid had

158   partial assignments to the red wolf cluster ($Q_{\text{Red Wolf}}$: GI-1=0.33, GI-2=0.28) (Fig. S1C).

159

160   *Shared private alleles*

161   We explored the degree to which each of the GI canids shared private alleles with

162   the reference canid groups. This method is an informative approach to infer population

163   relationships (*25*), where an excess of shared private alleles may imply source population

164   or recent introgression (*26*). We surveyed 6,859 loci with non-missing data for canid GI-

165   1 and estimated the proportion of alleles carried that are private to each reference group

166   (Table 2). GI-1 shared the most private alleles with coyotes (S_{PA} n=184), followed by red

167   wolves (S_{PA} n=21), with few shared private alleles with either gray wolves (S_{PA} n=12) or

168   eastern wolves (S_{PA} n=12) (Table 2). This trend persisted following correction for

169   unequal sample size (Coyote S_{PAr}=0.0102; S_{PAr}=0.0059) (Table 2). There was minimal

170   sharing with other reference canid lineages (SPA_r: *C. lupus*=0.0035; *C. lycaon*=0.0045).

8

171   We surveyed 6,391 loci with non-missing data for canid GI-2 and found similar trends as

172   for GI-1, where after correcting for sample size, again the greatest private allele sharing

173   observed was between coyotes (S_{PA} n=138; S_{PAr}=0.0093) followed by red wolves

174   (SPA=14; S_{PAr}=0.0063) and the other reference canid lineages (*C. lupus*: S_{PA} n=10;

175   S_{PAr}=0.0036; *C. lycaon*: S_{PA} n=8; S_{PAr}=0.0039) (Table 2).

176   As expected, the greatest proportion of shared private alleles between GI-1 and

177   coyotes was observed for alleles with a frequency in coyotes of ≤70% (Proportion shared,

178   prop$_S$=1.00) (Fig. 2), and a low sharing ratio was observed for private alleles with low

179   frequency in coyotes (e.g. ≤10%; prop$_S$=0.07) (Fig. 2). That is, GI-1 tended to share

180   private alleles with coyotes that were relatively common across the coyote reference

181   population. Interestingly, the same trend was not observed for private allele sharing with

182   red wolves, where a relatively high proportion of allele sharing was observed for private

183   alleles with ≤10% frequency in the reference population (prop$_S$=0.18) (Fig. 2). However,

184   the greatest proportion of allele sharing between red wolves and GI-1 was observed for

185   private alleles at ≤30% frequency (prop$_S$=0.21).

186   Similar trends were observed for GI-2, where the highest private allele sharing

187   proportion with coyotes was observed for private alleles with ≤60% frequency

188   (prop$_S$=0.50) (Fig. 2) and the lowest proportion shared for alleles with ≤50% and ≤10%

189   in the reference coyotes (prop$_S$=0.00; 0.05) (Fig. 2). Additionally, when compared to red

190   wolves, GI-2 had a moderately high private allele sharing proportion with low frequency

191   private alleles (≤10%; prop$_S$=0.12) (Fig. 2). However, the highest sharing proportion for

192   GI-2 and red wolves was observed for private alleles with ≤80% frequency in the

193   reference population (prop$_S$=0.50) (Fig. 2).

9

194   As the GI canids predominantly shared private alleles with low to moderate

195   frequencies in the red wolf population, this likely reflects the loss of diversity in red

196   wolves due to the founding of the captive breeding population or subsequent inbreeding

197   (*21,27*). Founders of the captive breeding population likely did not represent all red wolf

198   diversity that existed on the landscape prior to the 1970s, when trapping to remove wild

199   red wolves and founder selection occurred.

200   We observed 30 shared private alleles between at least one GI canid and red

201   wolves. These markers were predominantly found in intergenic regions distributed across

202   21 chromosomes ($n_{intergenic}$=19; n_{exon}=3; n_{intron}=7; $n_{promoter}$=1) (Table S3). GI-1 was

203   heterozygous at 76% of the loci that contained red wolf private alleles (n=16 out of 21

204   total sites), whereas GI-2 was heterozygous at 71% of the loci (n=10 of 14 total sites)

205   (Table S3). We found five overlapping shared red wolf private alleles between the two GI

206   individuals; however, we estimated a high level of allele sharing (IBS=0.93), likely due

207   to originating from the same population and high probability of being related.

208   When the LD filter was removed, we retained 8,167 and 7,609 SNPs in GI-1 and

209   GI-2, respectively. GI-1 carried a total of 30 red wolf private alleles ($n_{homozygous}$=8) and

210   GI-2 carried 26 ($n_{homozygous}$=8) (Fig. S3). Although this provided a genome-level

211   perspective of red wolf allele sharing, the resolution was not sufficient to conclusively

212   identify contiguous shared private alleles in extended linkage disequilibrium due to

213   recent admixture.

214   Overall, the GI canids carried 21 alleles that were absent from all reference

215   populations. These GI canid-specific alleles were distributed throughout the genome (n,

216   intergenic=16; exon=1; intron=4; promoter=0) (Fig. S3, Table S4B). GI-1 carried 14 such

10

217   private alleles and was homozygous for 50% of loci, where GI-2 carried 16 private alleles

218   and was homozygous for 56%. There were nine overlapping private alleles between the

219   two GI canids, which again likely reflect a high probability of relatedness (Table S4B). It

220   is possible that these alleles represent at least some of the genomic diversity in the

221   historical wild red wolf population that was lost as the result of selecting 14 captive

222   breeding founders from the wild, but this is speculation in the absence of documented

223   historical red wolf samples.

224

225   *Matrilineal haplotyping*

226   Canid mtDNA haplotypes are well established and form two clear gene tree

227   clades: a clade composed of Eurasian-evolved gray wolves and domestic dogs, and a

228   clade consisting of North American canids (coyotes, eastern wolves, and red wolves; Fig.

229   1D) (*28,29*). To further understand admixture in the GI canids and possible parental

230   mtDNA contributions, we amplified the mtDNA control region that was previously found

231   to contain a unique red wolf haplotype in contemporary wolves (*30*) and unique

232   haplotypes in ancient canids of the American southeast (*31*). Both GI canids carried

233   mtDNA haplotypes identical to previously published coyote haplotypes from the Great

234   Plains states (GI-1: haplotype la77; accession JN982588; *32*) and Texas (GI-2: la143;

235   accession FM209386; *17*). We reconstructed the two well supported mtDNA gene tree

236   clades that are commonly identified with a high posterior probability (Prob=0.98) and

237   found that the two GI canids clearly grouped with North American canids, specifically

238   coyotes, although nodal support within the two clades was generally low (Figs. 1D, S4).

239   For instance, the posterior probability was low (Prob<0.5) for most nodes within the

240   North American clade, especially within coyote haplotypes, which show very little

241   phylogenetic structuring across their range (*32*) (Figs. 1D, S4).

242

243   **Discussion**

244   We rediscovered ghost red wolf alleles present on the Gulf Coast landscape 40

245   years after red wolves were believed to be extinct in the region. Through interbreeding

246   with coyotes, endangered and extinct red wolf genetic variation has persisted and could

247   represent a reservoir of previously lost red wolf ancestry. This unprecedented discovery

248   opens new avenues for contemporary red wolf conservation and management, where

249   ghost alleles could be re-introduced into the current captive and experimental

250   populations. These admixed individuals consequently can be of great conservation value

251   but there is no ESA policy providing protection for admixed individuals that serve as

252   reservoirs for extinct genetic variation. An 'intercross policy' was introduced in 1996 to

253   assist prioritizing protection efforts but this policy was never fully adopted (*33*). Several

254   commentaries have encouraged an updated implementation of the ESA and Species

255   Status Assessments, especially as admixed genomes are increasingly being described and

256   viewed as a source of potentially beneficial genetic variation in the face of rapid climate

257   change (e.g. *34*). Although red wolves represent one of the greatest species recovery

258   stories in ESA history, debates regarding historical and on-going interbreeding with

259   coyotes highlight the ESA's short-comings associated with admixed individuals and the

260   difficulty in setting management objectives given our evolving understanding of admixed

261   genomes across wild populations (*35*).

262   In this study, our initial impetus to evaluate admixed individuals was due to the

263   phenotype displayed by canids on Galveston Island, which appear to overlap with the

264   canonical red wolf (Fig. 2). Given these suspected admixed canids, roadkill individuals

265   were sampled for genomic testing. Our survey of their genomes revealed a surprising

266   amount of allele sharing with the captive breeding population of red wolves. This shared

267   variation could be the consequence of two potential scenarios: 1) surviving ancestral

268   polymorphisms from the shared common ancestor of coyotes and red wolves that have

269   drifted to a high frequency in the captive breeding red wolf population and in a small

270   portion of Gulf Coast coyotes; or 2) wild coyotes in the Gulf Coast region are a reservoir

271   of red wolf ghost alleles that have persisted into the 21st century. Neither of these

272   potential explanations require adherence to a specific species concept. For instance,

273   incomplete lineage sorting from a shared common ancestor could occur whether red

274   wolves are a subspecies of the gray wolf, conspecific with Eastern wolves, or an

275   independent lineage with a possible ancient hybrid origin (*36*) (Fig. S5). Similarly,

276   interbreeding with the ancestral red wolf population would have resulted in the

277   introgression of red wolf alleles and associated phenotypes into Gulf Coast coyotes under

278   each species concept. Our findings of admixture and composition of private alleles are

279   most consistent with the second scenario, where the Galveston Island canids are admixed

280   coyotes carrying red wolves ghost alleles. Further, Galveston Island is found within the

281   historic red wolf range from where the original founders for the captive and reintroduced

282   populations were captured in the 1970s (Fig. S6). On an island, this canid population

283   likely experienced long-term reductions in gene flow with the southeastern coyote

284   population and a greater probability of retaining unique red wolf alleles. In further

285   support that coyotes of the American Gulf Coast likely serve as a ghost allele reservoir of

286   red wolf ancestry, we also identified two coyotes with red wolf admixture from

287   Louisiana's Gulf Coast, a second geographic region in which trapping efforts were

288   conducted to build a captive red wolf population (*16*). These findings provide substantial

289   support that ancestral red wolf genetic variation persists as ghost alleles in the regional

290   coyotes of the southeastern United States.

291   Our discovery warrants further genetic surveys of coyote populations in Louisiana

292   and Texas to establish the level and extent to which remnant red wolf alleles are found

293   exclusively in admixed coyotes. There are potentially admixed coyotes in the region that

294   exhibit higher levels of red wolf ancestry. With these surveys in place, conservation

295   efforts then face the opportunity to consider the role of remnant genetic variation in the

296   future of the red wolf. The NCEP of red wolves are a listable entity under the ESA in

297   need of proactive conservation (*36*). However, in the age of an extinction crisis,

298   innovative mechanisms to preserve and utilize adaptive potential are in great demand.

299   Today, every federally recognized red wolf individual is a descendant from 14 founders,

300   of which only 12 are genetically represented. These founders were removed from a single

301   geographic location in the 1970s and vastly underrepresent the original genomic diversity

302   present in southeastern wolves (*5*). Our discovery of red wolf ghost alleles in

303   southeastern coyotes demonstrate the ability to uncover ancestral variation and establish a

304   new component of biodiversity conservation. A minority of conservation priorities have

305   considered a 'de-introgression' strategy in which admixed individuals are bred in a

306   specific design to recover the extinct genotype (*37*).

307   The "tree of life" approach to conservation is under challenge, as a new paradigm

308   has been proposed to include admixed genomes (*35,38*). Historical and contemporary red

309   wolves face anthropogenically-mediated hybridization, but introgression is also likely a

310   natural process in the evolution of *Canis* lineages. As an important evolutionary process,

311   introgression could protect adaptive potential and maintain processes that sustain

312   ecosystems. As a result, incorporating admixed entities into conservation policy and here,

313   red wolf restoration, may be the next step in broader biodiversity conservation. Another

314   pivotal next step in red wolf restoration is the identification of a new re-introduction site

315   for a wild population of red wolves. Our discovery of red wolf ghost alleles indicates

316   there are geographic regions that can harbor endangered genetic variation and may guide

317   future efforts for red wolf reintroduction. The foundation upon which that effort will be

318   built rests exclusively on describing large-scale geographic patterns of red wolf ghost

319   alleles in the American southeast.

320

321   **Materials and Methods**

322   *Genomic DNA preparation and sample selection*

323   We obtained tissue samples from two roadkill canids of unknown taxonomic

324   affiliation on Galveston Island, TX (GI canids hereafter). We extracted genomic DNA

325   using a DNeasy Blood and Tissue Kit (Qiagen) following the manufacturer's instructions.

326   We selected reference samples that collectively represented all possible wild canid

327   evolutionary lineages in North America that could have contributed to the ancestry and

328   genetic variation of the two GI canids (Table S1). The reference tissue samples were

329   obtained via helicopter darting and post-mortality sampling by the US National Park

15

330   Service (*C. lupus*), through the Ontario Ministry of Natural Resources archives (*C.*

331   *lycaon*), or the Point Defiance Zoo and Aquarium archives (*C. rufus*). Reference coyote

332   samples were obtained through the Sam Noble Oklahoma Museum of Natural History,

333   samples collected from animal captured for radio collaring, or donations from hunters

334   and trappers. We selected captive red wolves that captured genetic diversity of the 12

335   original founders and who are genetically represented in extant wolves that reproduced in

336   the captive breeding program. These captive red wolves are considered the reference red

337   wolves in all analyses.

338

339   *RADseq library preparation and bioinformatic processing of sequence data*

340   We used a modified version of the RADseq protocol by (*20*). We digested DNA

341   with *Sbf1*, ligated a unique barcode adapter to the fragments, and pooled between 96 and

342   153 samples. Each pool was subsequently sheared to 400 bp in a Covaris LE220 at

343   Princeton University's Lewis Siegler Institute Genomics Core Facility. We recovered

344   ligated fragments using a streptavidin bead binding assay and prepared genomic libraries

345   for Illumina HiSeq sequencing following either the standard TruSeq protocol for the

346   NEBNext Ultra or NEBNext UltraII DNA Library Prep Kit (New England Biolabs). We

347   conducted a size selection step using Agencourt AMPure XP beads (Beckman Coulter) to

348   retain fragments 300-400 bp in size. We also used AMPure XP beads for library

349   purification. We standardized genomic libraries to 10nM for 2X150nt sequencing on an

350   Illumina HiSeq 2500 platform.

351   We filtered raw paired-end sequence data to retain reads that contained one of the

352   96 possible barcodes and the expected restriction enzyme cut-site using a custom perl

16

353   script (flip_trim_sbfI_170601.pl, see Supporting Information). We discovered variant

354   sites following the STACKS v 1.42 pipeline (*39*). Reads were de-multiplexed using

355   *Process_Radtags*, allowing a mismatch of two to rescue barcodes. We discarded reads

356   with an uncalled base or with an average quality score (≤10) within a sliding window

357   equivalent to 15% of the total read length. We removed PCR duplicates using

358   *Clone_Filter* with default parameters. All samples were mapped to the Canfam3.1

359   assembly of dog genome (*40*) with STAMPY v 1.0.31 (*41*). We filtered mapped reads in

360   Samtools v 0.1.18 (*42*) to retain those with MAPQ >96 and exported as a bam file.

361   Variant calling was completed in STACKS following a standard pipeline for reference

362   mapped data (i.e. *pstacks, cstacks, sstacks, populations*) (*39*). We required a minimum

363   stack depth of 3 reads (-m) in *pstacks* and allowed a maximum per locus missingness of

364   10% in *populations*. Further, to reduce biases resulting from linked markers, we enabled

365   the *–write-single-snp* flag in *populations* and filtered for statistical linkage disequilibrium

366   (LD) across sites using the *--indep-pairwise* 50 5 0.5 flag in Plink v1.90b3i (*43*). We

367   conducted a final filtering to retain sites that also had a minimum minor allele of 1%.

368   Standard metrics of genomic diversity (observed heterozygosity, H_O; private

369   allele count, $N_{PA,}$ Pairwise F_{ST}) for all reference groups were calculated using functions in

370   STACKS. We evaluated significant differences in genome-wide heterozygosity pairwise

371   estimates with a pairwise Wilcoxon rank sum test implemented in *R* with a false

372   discovery rate correction for multiple testing (FDR<0.05). Allelic richness (A_r) and

373   private allelic richness (PA_r) were calculated using a rarefaction approach implemented

374   in ADZE (*25*) with a maximum tolerance of 10% missing data and maximum

17

375   standardized sample size (g) set to the smallest n for the samples considered (20; eastern

376   and gray wolves).

377

378   *Mitochondrial DNA sequencing and haplotyping*

379   Mitochondrial sequence was amplified from genomic DNA with primers for

380   control region (Thr-L 15926: 5′-CAATTCCCCGGTCTTGTAAACC-3′; DL-H 16340:

381   5′-CCTGAAGTAGGAACCAGATG-3) and thermocycling conditions following (*28*).

382   Amplified products were bi-directionally sequenced using a service provided by

383   GeneWiz (New Jersey). Each sample was sequenced in duplicate to confirm ambiguous

384   sites. Sequences were view, corrected, and aligned with Geneious v6.16 software (*44*).

385

386   *Population Structure Analyses*

387   To visualize genetic clustering patterns of the filtered SNP dataset, we conducted

388   a Principal Component Analysis (PCA) in *flashPCA* (*45*). Additional PCAs were

389   completed on a subset of the total sample size for specific comparisons (i.e. inclusion of

390   only reference coyotes, reference red wolves, and the two GI individuals). We

391   implemented a maximum-likelihood analysis to infer population structure using the

392   program ADMIXTURE v1.3. (*23*). We evaluated between 1 and 10 genetic partitions

393   (*K*), evaluated the fit of each partition using the cross-validation flag, and considered the

394   best fit number of partitions to have the lowest cross-validation score. We first considered

395   the entire dataset, with subsequent analyses conducted on subsets of the total sample size

396   for specific comparisons. For example, we analyzed only the reference coyotes, reference

397   red wolves, and the two GI canids using the aforementioned parameters.

18

398   Although this maximum-likelihood cluster analysis is useful for evaluating

399 specific levels of data partitioning, it is not an explicit ancestry analysis. Using a

400 Bayesian framework, we conducted a posterior probability assignment test in Structure v.

401 2.3.4 (*24*) that included all reference coyotes and red wolf individuals as the training set

402 of samples. We then assigned each of the GI canids to one or more of these reference

403 groups (*K*=2) using 10,000 repetitions following a burn-in of 2,500.

404

405 *Private Allele Sharing Analyses*

406   We evaluated the degree of sharing of private alleles among the GI canids and all

407 possible reference groups, considering each GI canid separately. To avoid spurious

408 identification of private alleles due to the presence of any missing data, we restricted

409 analyses to loci that were 100% genotyped in each GI canid and identified alleles private

410 to each reference group in STACKS. We then determined the number of shared private

411 alleles between the GI canid and the reference groups. Further, we calculated shared

412 private allelic richness with each reference group using a rarefaction approach in ADZE

413 with a tolerance of 15% missing data and a maximum standardized sample size (g) of

414 two.

415   We estimated the frequency of each shared private allele in the corresponding

416 reference coyote or red wolf population. This frequency distribution was binned in

417 frequency intervals of 10%. In other words, the number of shared private alleles for each

418 GI canid was divided by the total number of private alleles and binned in 10% frequency

419 intervals based on the allele's frequency in the corresponding reference population.

19

420   We determined the genomic coordinates of all shared red wolf private alleles with

421   each of the GI canids. We annotated each site as intergenic, within an intron or exon, or

422   within a putatively regulatory region (within 2 kb of a transcription start site) using a

423   custom python script and the Ensembl gene database (chr_site.py; see Supporting

424   Information) (*46*). To further evaluate the fine scale distribution of shared red wolf

425   private alleles across the genomes of GI-1 and GI-2, we removed the LD filter and

426   recalculated shared private alleles with the red wolf reference group as described above.

427   We evaluated the alleles found only in GI-1 and GI-2 to determine whether either GI

428   canid harbored any unique genomic diversity absent from the reference groups. We

429   calculated the identity by state (IBS) between the two GI canids in Plink v1.90b3i using

430   the --ibs-matrix argument.

431

432   *Mitochondrial DNA sequence analysis*

433   We compared the GI canids mtDNA haplotypes to mtDNA control region

434   sequences available on GenBank the represented possible *Canis* ancestors from multiple

435   evolutionary lineages: domestic dogs, gray wolves, eastern wolves, red wolves, and

436   coyotes (Table S2). Using consensus sequences of 234bp, we estimated gene trees using

437   Bayesian methods implemented in BEAST v1.8.4 with red fox (*Vulpes Vulpes*) as an

438   outgroup (*47*). In BEAST, we use a constant size coalescent tree prior, an uncorrelated

439   lognormal relaxed molecular clock, and used a random starting tree. We conducted two

440   independent Markov Chain Monte Carlo (MCMC) analyses for 25 million steps,

441   sampling every 2,500 steps, and combined tree estimates from each run with

442   LogCombiner v1.8.4 with a 10% burnin. Convergence on the posterior distribution was

443   determined based on viewing the log files in Tracerv1.6, where convergence is attained

444   when the effective sample size of a parameter (i.e. the number of effectively independent

445   draws from the posterior distribution) is at least 200; all parameters in our analyses had

446   effective sample size values greater than 400. To visualize our estimated gene trees, we

447   first calculated the maximum clade credibility in TreeAnnotator v1.8.4 and upload the

448   most likely tree in the Interactive Tree of Life v3.6.3 online platform (*48*). Full tree

449   results are presented in Figure S5 with posterior probabilities over 0.90 reported;

450   truncated results that show the clearly detected phylogenetic relationships among North

451   American canids and GI canids are presented in Figure 2D.

452

453   **Acknowledgments**

463

464   **Data and materials availability**

465   All data needed to evaluate the conclusions in this paper can be obtained from the

466 National Center for Biotechnology Information Sequence Read Archive (SUB4398163).

467 Additional data can be requested from the corresponding authors.

468
469

Literature Cited

1. H. McCarley, C. J. Carley, *Recent changes in the distribution and status of wild red wolves Canis rufus.* (Endangered Species Report No. 4. U.S. Fish and Wildlife Service, Albuquerque, NM, 1979).

2. J.L. Paradiso, Recent records of red wolves from the Gulf Coast of Texas. *Southwest. Nat.* **10**, 318-319 (1965).

3. D.H Pimlott, P. W. Joslin, The status and distribution of the red wolf. *T. N. Am Wildl. Nat. Res.* **33**, 373–389 (1968).

4. R.M Nowak, Report on the red wolf. *Defenders of Wildlife News* **45**, 82-94 (1970).

5. Carley, C. J. *Activities and findings of the Red Wolf Recovery Program from late 1973 to July 1, 1975*. U.S. Fish and Wildlife Service. Albuquerque, NM, 1975).

6. M. K. Phillips, W.T. Parker, Red wolf recovery: a progress report. *Conserv. Biol.* **2**, 139-141 (1988).

7. P.W. Hedrick, J. Fredrickson, Captive breeding and the reintroduction of Mexican and red wolves. *Mol. Ecol.* **17**, 344–350 (2008).

8. S.H. Fritts, E.E. Bangs, J.A. Fontaine, M.R. Johnson, M.K. Phillips, E.D. Koch, J.R. Gunson, Planning and implementing a reintroduction of wolves to Yellowstone National Park and central Idaho. *Restor. Ecol.* **5**, 7-27 (1997).

9. P.W. Hedrick, P.S. Miller, E. Geffen, R.K Wayne, Genetic evaluation of the three captive Mexican wolf lineages. *Zoo. Biol.* **16**, 47-69 (1997).

10. *Endangered and Threatened Wildlife and Plants; Proposed Replacement of the Regulations for the Nonessential Experimental Population of Red Wolves in Northeastern North Carolina*. (Department of the Interior Fish and Wildlife Service Federal Register Vol. 83, No. 125, 2018).

11. J.H. Bohling, LP Waits, Factors influencing red wolf–coyote hybridization in eastern North Carolina, USA. *Biol. Cons.* **184**, 108-116 (2015).

12. J.W. Hinton, K.E Brzeski, D.R. Rabon, M.J. Chamberlain, Effects of anthropogenic mortality on critically endangered red wolf *Canis rufus* breeding pairs: implications for red wolf recovery. *Oryx* **51**, 174-181. (2017).

13. E. Weller, *Red Wolf (Canis rufus)*. (United States Fish and Wildlife Service 5-Year Review, retrieved from https://www.fws.gov/southeast/wildlife/mammal/red-wolf/#, 2018).

14. R.J. Fredrickson, P.W. Hedrick, Dynamics of hybridization and introgression in red wolves and coyotes. *Conserv. Biol.* **20**, 1272-1283 (2006).

15. H. McCarley, The taxonomic status of wild *Canis* (Canidae) in the south central United States. *Southwest. Nat.* **7**, 227-235 (1962).

16. G.A. Riley, R. T. McBride, A survey of the red wolf (*Canis rufus*). *U. S. Fish and Wildlife Service, Special Scientific Report-Wildlife* **162**, 1-15 (1972).

17. F. Hailer, J.A. Leonard, Hybridization among three native North American *Canis* species in a region of natural sympatry. *PLoS One* **3**, e3333 (2008).

18. J.W. Hinton, M.J. Chamberlain, Morphometrics of Canis taxa in eastern North Carolina. *J. Mammal.* **95**, 855-861 (2014).

19. J.W. Hinton, J.L. Gittleman, F.T. van Manen, M.J. Chamberlain, Size-assortative choice and mate availability influences hybridization between red wolves (*Canis rufus*) and coyotes (*Canis latrans*). *Ecol. Evol.* **8**, 3927-3940 (2018).

515   20. O.A. Ali, S.M. O'Rourke, S.J. Amish, M.H. Meek, G. Luikart, C. Jeffres, M.R.
516   Miller, RAD Capture (Rapture): Flexible and efficient sequence-based genotyping.
517   *Genetics* **202**(2), 389-400 (2015).
518   21. K.E Brzeski, D. R. Rabon Jr, M.J. Chamberlain, L.P. Waits, S.S. Taylor, Inbreeding
519   and inbreeding depression in endangered red wolves (*Canis rufus*). *Mol. Ecol.*
520   **23**,4241-4255 (2014).
521   22. B.M. vonHoldt, J.P. Pollinger, D.A. Earl, J.C. Knowles, A.R. Boyko, H. Parker, E.
522   Geffen, M. Pilot, W. Jedrzejewski, B. Jedrzejewska, V. Sidorovich, C. Greco, E.
523   Randi, M. Musiani, R. Kays, C.D. Bustamante, E.A. Ostrander, J. Novembre, R.K.
524   Wayne, A genome-wide perspective on the evolutionary history of enigmatic wolf-
525   like canids. *Genome Res.* **21**, 1294-1305 (2011).
526   23. D.H. Alexander, J. Novembre, K. Lange, Fast model-based estimation of ancestry in
527   unrelated individuals. *Genome Res*. **19**, 1655-1664 (2009).
528   24. J.K. Pritchard, M. Stephens, P. Donnelly, Inference of population structure using
529   multilocus genotype data. *Genetics* **155**, 945–959 (2000).
530   25. Z.A. Szpiech, M. Jakobsson, N.A. Rosenberg, ADZE: A rarefaction approach for
531   counting alleles private to combinations of populations. *Bioinformatics* **24**, 2498-
532   2504 (2008).
533   26. M.K. Gonder, S. Locatelli, L. Ghobrial, M.W. Mitchell, J.T. Kujawski, F.J.
534   Lankester, C.B Stewart, S.A. Tishkoff, Evidence from Cameroon reveals differences
535   in the genetic structure and histories of chimpanzee populations. *P. Natl. A. Sci.USA*
536   **108**, 4766-4771 (2011).
537   27. J.L. Simonis, R.B. Harrison, S.T. Long, D.R. Rabon Jr, W.T. Waddell, L.J. Faust,
538   Managed movement increases metapopulation viability of the endangered red wolf. *J.*
539   *Wildlife Manage.* **82**(3), 573-582 (2018).
540   28. C. Vilà, I.R.Amorim, J.A. Leonard, D. Posada, J. Castroviejo, F. Petrucci-Fonseca,
541   K.A. Crandall, H. Ellegren, R.K. Wayne, Mitochondrial DNA phylogeography and
542   population history of the gray wolf *Canis lupus*. *Mol. Ecol.* **8**, 2089-2103 (1999).
543   29. L.Y. Rutledge, C.J. Garroway, K.M. Loveless, B.R. Patterson, Genetic differentiation
544   of eastern wolves in Algonquin Park despite bridging gene flow between coyotes and
545   gray wolves. *Heredity* **105**(6), 520 (2010).
546   30. J.A. Leonard, R.K. Wayne, J. Wheeler, R. Valadez, S. Guillén, C. Vilà, C. DNA
547   evidence for Old World origin of New World dogs. *Science* **298**, 1613-1616 (2002).
548   31. K.E. Brzeski, M.B. DeBiasse, D.R. Rabon Jr., M.K. Chamberlain, S.S. Taylor,
549   Mitochondrial DNA variation in southeastern Pre-Columbian canids. *J Hered.* **107**(3),
550   287-293 (2016).
551   32. S. Koblmüller, R.K. Wayne, J.A. Leonard, Impact of Quarternary climate changes
552   and interspecific competition on the demographic history of a highly mobile
553   generalist carnivore, the coyote. *Biol. Letters* **8**, 644-647 (2012).
554   33. S.M. Haig, F.W Allendorf, in *The Endangered Species Act at Thirty: Volume 2:*
555   *conserving biodiversity in human-dominated landscapes*, D.D. Scott, F.W. Davis,
556   Eds. (Island Press, Washington, DC, 2006), pp.150-163.
557   34. M. Becker, N. Gruenheit, M. Steel, C. Voelckel, O. Deusch, P.B. Heenan, P.A.
558   McLenachan, O. Kardailsky, J.W. Leigh, P.J. Lockhard, Hybridization may facilitate
559   in situ survival of endemic species through periods of climate change. *Nat. Clim.*
560   *Change* **3**, 1039-1043 (2013).

561   35. R.N Jackiw, G. Mandil, H.A. Hager, A framework to guide the conservation of
562   species hybrids based on ethical and ecological considerations. *Conserv. Biol.* **29**,
563   1040-1051 (2015).
564   36. R. Waples, R. Kays, R.J. Fredrickson, K. Pacifici, L.S. Mills, Is the red wolf a listable
565   unit under the US Endangered Species Act? *J. Hered.* **109**(5), 585-597 (2018).
566   37. C. Amador, B.J. Hayes, H.D. Daetwyler, Genomic selection for recovery of original
567   genetic background from hybrids of endangered and common breeds. *Evol. Appl.* **7**,
568   227-237 (2014).
569   38. F.A. Allendorf, R.F. Leary, P. Spruell, J.K. Wenburg, The problems with hybrids:
570   setting conservation guidelines. *Trends Ecol. Evol.* **16**, 613-622 (2001).
571   39. J. Catchen, P.A. Hohenlohe, S. Bassham, A. Amores, W.A. Cresko, Stacks: an
572   analysis tool set for population genomics. *Mol. Ecol.* **22**, 3124-3140 (2013).
573   40. K. Lindblad-Toh, C. M. Wade, T. S. Mikkelsen, E. K. Karlsson, D. B. Jaffe, M.
574   Kamal, M. Clamp, J. L. Chang, E. J. Kulbokas III, M. C. Zody, E. Mauceli, X. Xie,
575   M. Breen, R. K. Wayne, E. A. Ostrander, C. P. Ponting, F. Galibert, D. R. Smith, P. J.
576   deJong, E. Kirkness, P. Alvarez, T. Biagi, W. Brockman, J. Butler, C.-W. Chin, A.
577   Cook, J. Cuff, M.J. Daly, D. DeCaprio, S. Gnerre, M. Grabherr, M. Kellis, M. Kleber,
578   C. Bardeleben, L. Goodstadt, A. Heger, C. Hitte, L. Kim, K.-P. Koepfli, H.G. Parker,
579   J.P. Pollinger, S.M.J. Searle, N.B. Sutter, R. Thomas, C. Webber, J. Baldwin, A.
580   Abebe, A. Abouelleil, L. Aftuck, M. Ait-zahra, T. Aldredge, N. Allen, P. An, S.
581   Anderson, C. Antoine, H. Arachchi, A. Aslam, L. Ayotte, P. Bachantsang, A. Barry,
582   T. Bayul, M. Benamara, A. Berlin, D. Bessette, B. Blitshteyn, T. Bloom, J. Blye, L.
583   Boguslavskiy, C. Bonnet, B. Boukhgalter, A. Brown, P. Cahill, N. Calixte, J.
584   Camarata, Y. Cheshatsang, J. Chu, M. Citroen, A. Collymore, P. Cooke, T. Dawoe,
585   R. Daza, K. Decktor, S. DeGray, N. Dhargay, K. Dooley, K. Dooley, P. Dorje, K.
586   Dorjee, L. Dorris, N. Duffey, A. Dupes, O. Egbiremolen, R. Elong, J. Falk, A. Farina,
587   S. Faro, D. Ferguson, P. Ferreira, S. Fisher, M. FitzGerald, K. Foley, C. Foley, A.
588   Franke, D. Friedrich, D. Gage, M. Garber, G. Gearin, G. Giannoukos, T. Goode, A.
589   Goyette, J. Graham, E. Grandbois, K. Gyaltsen, N. Hafez, D. Hagopian, B. Hagos, J.
590   Hall, C. Healy, R. Hegarty, T. Honan, A. Horn, N. Houde, L. Hughes, L. Hunnicutt,
591   M. Husby, B. Jester, C. Jones, A. Kamat, B. Kanga, C. Kells, D. Khazanovich, A. C.
592   Kieu, P. LeVine, J. Liu, X. Liu, Y. Lokyitsang, T. Lokyitsang, A. Lui, J. Macdonald,
593   J. Major, R. Marabella, K. Maru, C. Matthews, S. McDonough, T. Mehta, J. Meldrim,
594   A. Melnikov, L. Meneus, A. Mihalev, T. Mihova, K. Miller, R. Mittelman, V.
595   Mlenga, L. Mulrain, G. Munson, A. Navidi, J. Naylor, T. Nguyen, N. Nguyen, C.
596   Nguyen, T. Nguyen, R. Nicol, N. Norbu, C. Norbu, N. Novod, T. Nyima, P. Olandt,
597   B. O'Neill, K. O'Neill, S. Osman, L. Oyono, C. Patti, D. Perrin, P. Phunkhang, F. Pierre,
598   M. Priest, A. Rachupka, S. Raghuraman, R. Rameau, V. Ray, C. Raymond, F. Rege,
599   C. Rise, J. Rogers, P. Rogov, J. Sahalie, S. Settipalli, T. Sharpe, T. Shea, M. Sheehan,
600   N. Sherpa, J. Shi, D. Shih, J. Sloan, C. Smith, T. Sparrow, J. Stalker, N. Stange-
601   Thomann, S. Stavropoulos, C. Stone, S. Stone, S. Sykes, P. Tchuinga, P. Tenzing, S.
602   Tesfaye, D. Thoulutsang, Y. Thoulutsang, K. Topham, I. Topping, T. Tsamla, H.
603   Vassiliev, V. Venkataraman, A. Vo, T. Wangchuk, T. Wangdi, M. Weiand, J.
604   Wilkinson, A. Wilson, S. Yadav, S. Yang, X. Yang, G. Young, Q. Yu, J. Zainoun, L.
605   Zembek, A. Zimmer; for Broad Sequencing Platform members, E. S. Lander,

25

606   Genome sequence, comparative analysis and haplotype structure of the domestic dog.
607   Nature **438**, 803-819 (2005).
608 41. G. Lunter, M. Goodson, Stampy: A statistical algorithm for sensitive and fast
609   mapping of Illumina sequence reads. *Genome Res.* **21**, 936–939 (2011).
610 42. H. Li, A statistical framework for SNP calling, mutation discovery, association
611   mapping and population genetic parameter estimation from sequencing data.
612   *Bioinformatics* **27**(21), 2987-2993 (2011).
613 43. S. Purcell, B. Neale, K. Todd-Brown, L. Thomas, M.A.R. Ferreira, D. Bender, J.
614   Maller, P. Sklar, P.I.W de Bakker, M.J. Daly, P.C. Sham, PLINK: a tool set for
615   whole-genome association and population-based linkage analysis. *Am. J. Hum Genet,*
616   **81**, 559-579 (2007).
617 44. M. Kearse, R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton,
618   A. Cooper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, P. Meintjes, A.
619   Drummond, Geneious Basic: an integrated and extendable desktop software platform
620   for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647-1649
621   (2012).
622 45. G. Abraham, M. Inouye, Fast principal component analysis of large-scale genome-
623   wide data. *PLoS ONE* **9**, 1-5 (2014).
624 46. D.R. Zerbino, P. Achuthan, W. Akanni, M.R. Amode, D. Barrell, J. Bhai, K. Billis, C.
625   Cummins, A. Gall, C.G. Girón, L. Gil, L. Gordon, L. Haggerty, E. Haskell, T.
626   Hourlier, O.G. Izuogu, S.H. Janacek, T. Juettemann, J.K. To, M.R. Laird, I. Lavidas,
627   Z. Liu, J.E. Loveland, T. Maurel, W. McLaren, B. Moore, J. Mudge, D.N. Murphy,
628   V. Newman, M. Nuhn, D. Ogeh, C.K. Ong, A. Parker, M. Patricio, H.S. Riat, H.
629   Schuilenburg, D. Sheppard, H. Sparrow, K. Taylor, A. Thormann, A. Vullo, B. Walts,
630   A. Zadissa, A. Frankish, S.E. Hunt, M. Kostadima, N. Langridge, F.J. Martin, M.
631   Muffato, E. Perry, M. Ruffier, D.M. Staines, S.J. Trevanion, B.L. Aken, F.
632   Cunningham, A. Yates, P. Flicek, Ensembl 2018. *Nucleic Acids Res.* **46**(D1), D754–
633   D761 (2018).
634 47. A.J. Drummond, M.A. Suchard, D. Xie, A. Rambaut, Bayesian phylogenetics with
635   BEAUTi and the BEAST 1.7. *Mol. Biol. Evol.* **29**(8), 1969-1973 (2012).
636 48. I. Letunic, P. Bork, Interactive Tree Of Life (iTOL) v3: an online tool for the display
637   and annotation of phylogenetic and other trees. *Nucleic Acids Res.* **44**(W1), W242-
638   W245 (2016).
639
640

641 **Figure 1.** Analyses of genome wide SNP and mtDNA data across all legally recognized
642 wild *Canis* species and two canids from Galveston Island. A) Map of area. Site 1 is
643 Sabine National Wildlife Refuge, site 2 is McFadden National Wildlife Refuge, and site 3
644 is Brazoria National Wildlife Refuge. Cluster patterns were assessed across 7,068 SNPs
645 with a B) principal component analysis and C) admixture analysis of *K*=2-4 partitions. D)
646 Clade membership was determined by reconstruction of the Bayesian haplotype tree with
647 the highest posterior probability (Prob=0.98) from 234bp of mtDNA sequence data from
648 the control region; taxonomic designation of eastern wolf is based on assigned clade and
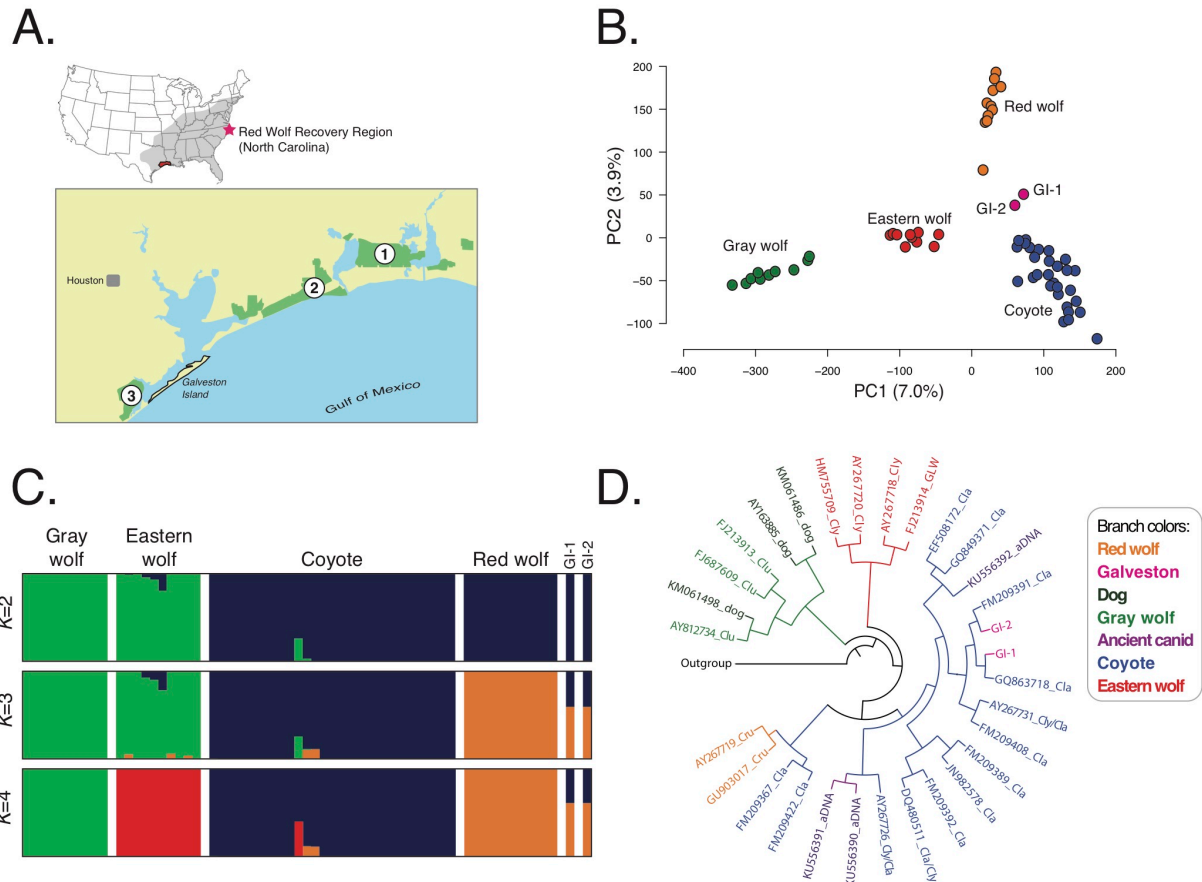649 sample location, not necessarily field identification.

650
651

Figure 2. Photographic comparison of coyotes, Galveston Island (GI) canids, and red wolves. Photo credit and location as follows: A) Pack of GI canids, Galveston Island, TX, credit: R. Wooten. B) Western coyote, Intermountain West, United States, credit: Wikimedia commons, Rich Keen/DPRA. GI canid laying on airport runway, Galveston Island, TX, credit: R. Wooten. Captive female red wolf, Alligator River National Wildlife Refuge, NC, credit: R. Nordsven, USFWS. C) Western coyote, Joshua Tree National Park, CA, credit: Wikimedia commons, Michael Vamstad/NPS. Head-shots of GI canids, Galveston Island, TX, credit: R. Wooten. Wild juvenile male red wolf prior to release, Albemarle Peninsula, NC, credit R. Nordsven, USFWS.
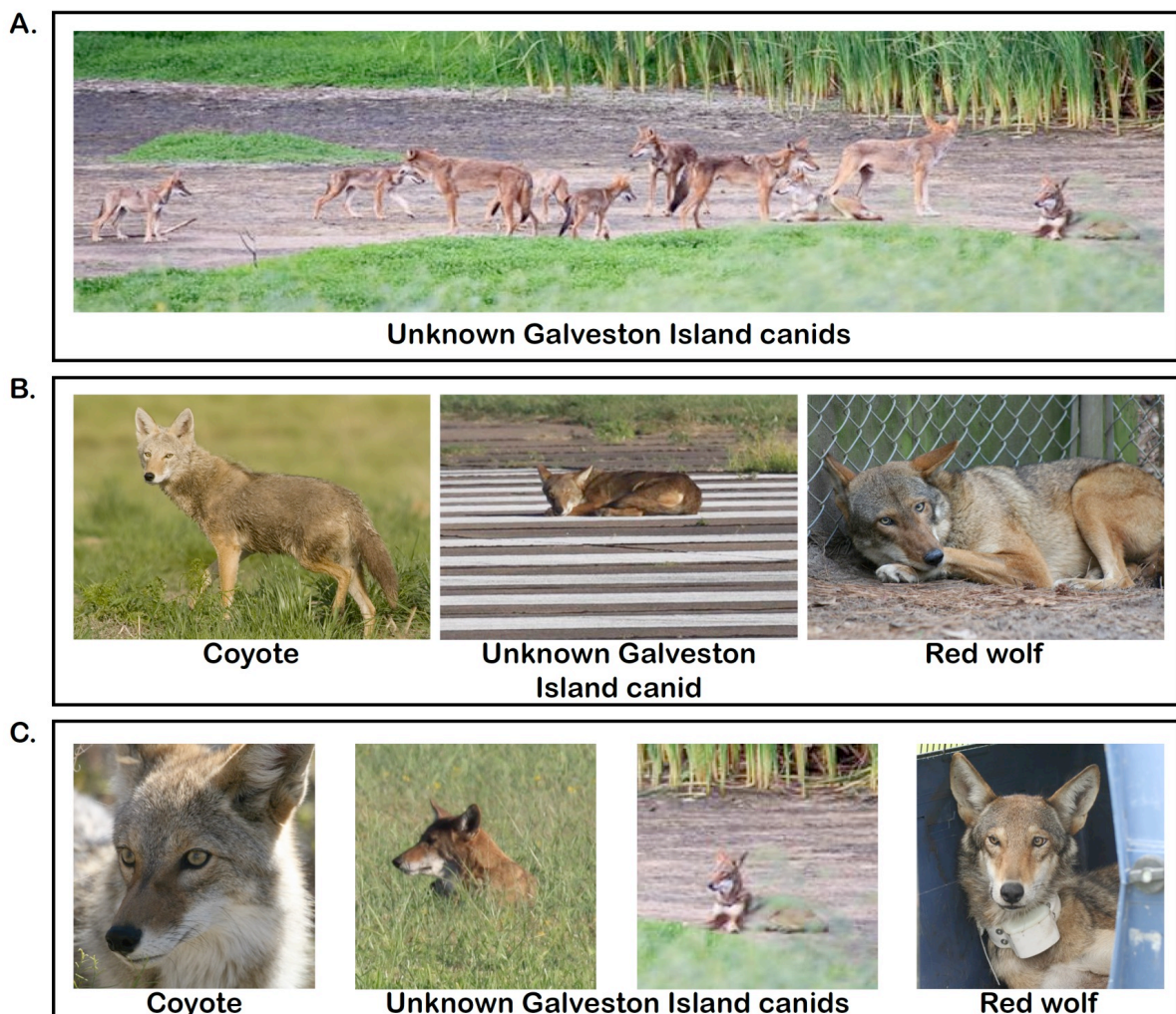
663   **Figure 3.** Sharing of red wolf and coyote private alleles with the two Galveston Island,
664   Texas canids, A) GI-1 and B) GI-2, and their frequencies in their respective reference
665   populations.

666
667
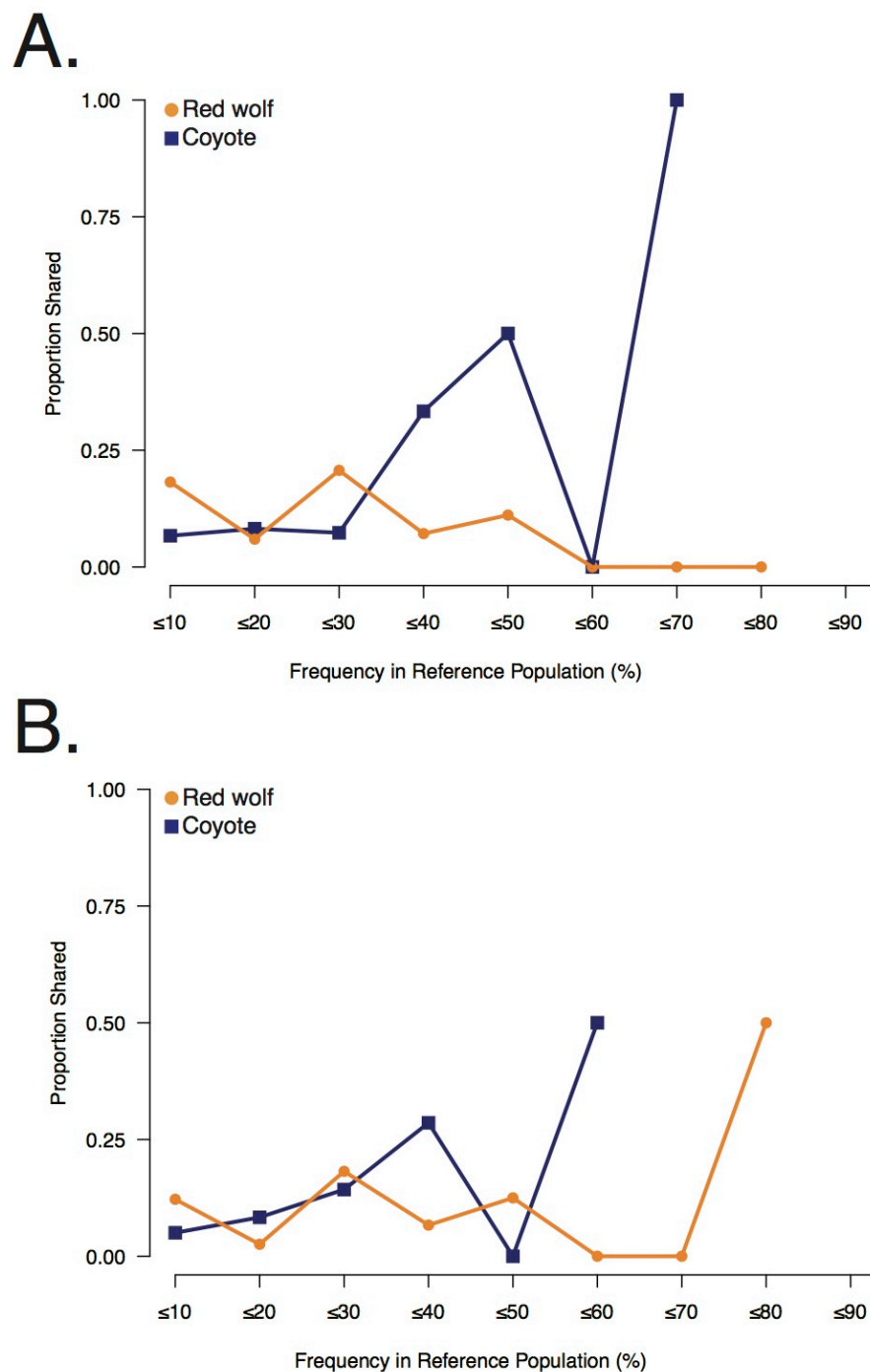
668 **Table 1. Diversity statistics for each reference group of *Canis* in North America.**
669 Summary statistics for each *Canis* reference population derived from 7,047 genome-wide
670 polymorphic SNPs. (Abbreviations: n, sample size; H_O, observed heterozygosity; H_E,
671 Expected heterozygosity; A_R, Allelic Richness; PA_R, Private Allelic Richness)

Group	Sampling Location[*]	n	H_O	H_E	A_R[*]	N_{PA}	PA_R[**]
Coyote	Southeastern USA	29	0.085	0.101	1.52	2,686	0.28
Gray wolf	Yellowstone National Park	10	0.072	0.076	1.27	368	0.10
Eastern wolf	Algonquin Provincial Park	10	0.079	0.087	1.36	332	0.11
Red wolf (captive)	Point Defiance Zoo & Aquarium	11	0.051	0.061	1.17	191	0.04

672 *For details regarding which U.S. states are in the sampling region, please see Table S1.
673 **Allelic richness and private allelic richness are reported for minimum sample size (g)
674 of 18, the maximum obtainable g for eastern wolves & gray wolves given the sample size
675 and tolerance threshold.
676
677

678 **Table 2. Private allele sharing between reference groups and each Galveston Island,**
679 **TX canid.** Summary statistics for GI-1 and GI-2 were calculated over 6,859 and 6,391
680 genome-wide SNPs, respectively, reflecting the number of loci with non-missing data for
681 each individual. (Abbreviations: N_{PA}, number of private alleles; percentage; S_{PA}, shared
682 private alleles; S_{Par}, shared private allelic richness)

Reference Group	GI-1				GI-2			
	N_{PA}	S_{PA} (count)	S_{PA} (%)	S_{PAr}	N_{PA}	S_{PA} (count)	S_{PA} (%)	S_{PAr}
Coyote	2,632	184	6.99%	0.0102	2,439	138	5.66%	0.0093
Gray wolf	362	12	3.31%	0.0035	335	10	2.99%	0.0036
Eastern wolf	329	12	3.65%	0.0045	303	8	2.64%	0.0039
Red wolf	188	21	11.17%	0.0059	171	14	8.19%	0.0063

683
684
685

31