Conserv Genet
DOI 10.1007/s10592-016-0863-0

RESEARCH ARTICLE

Genetic structure and viability selection in the golden eagle (*Aquila chrysaetos*), a vagile raptor with a Holarctic distribution

Jacqueline M. Doyle[1,2] · Todd E. Katzner[3] · Gary W. Roemer[4] · James W. Cain III[5] · Brian A. Millsap[6] · Carol L. McIntyre[7] · Sarah A. Sonsthagen[8] · Nadia B. Fernandez[1] · Maria Wheeler[9] · Zafer Bulut[10] · Peter H. Bloom[11] · J. Andrew DeWoody[1,12]

Received: 26 February 2016 / Accepted: 28 June 2016

Abstract Molecular markers can reveal interesting aspects of organismal ecology and evolution, especially when surveyed in rare or elusive species. Herein, we provide a preliminary assessment of golden eagle (*Aquila chrysaetos*) population structure in North America using novel single nucleotide polymorphisms (SNPs). These SNPs included one molecular sexing marker, two mitochondrial markers, 85 putatively neutral markers that were derived from noncoding regions within large intergenic intervals, and 74 putatively nonneutral markers found in or very near protein-coding genes. We genotyped 523 eagle samples at these 162 SNPs and quantified genotyping error rates and variability at each marker. Our samples corresponded to 344 individual golden eagles as assessed by unique multilocus genotypes. Observed heterozygosity of known adults was significantly higher than of chicks, as was the number of heterozygous loci, indicating that mean zygosity measured across all 159 autosomal markers was an indicator of fitness as it is associated with eagle survival to adulthood. Finally, we used chick samples of known provenance to test for population differentiation across portions of North America and found pronounced structure among geographic sampling sites. These data indicate that cryptic genetic population structure is likely widespread in the golden eagle gene pool, and that extensive field sampling and genotyping will be required to more clearly delineate management units within North America and elsewhere.

Keywords Allelic diversity · Effective population size · Repeatability · Fluidigm · Heterozygosity fitness correlation · Genetic mating system · Genetic monogamy

Electronic supplementary material The online version of this article (doi:10.1007/s10592-016-0863-0) contains supplementary material, which is available to authorized users.

✉ Jacqueline M. Doyle
jdoyle@towson.edu

[1] Department of Forestry and Natural Resources, Purdue University, 715 W. State Street, West Lafayette, IN 47907, USA

[2] Department of Biological Sciences, Towson University, 8000 York Road, Towson, MD 21252, USA

[3] U.S. Geological Survey, Forest and Rangeland Ecosystem Science Center, 970 Lusk Street, Boise, ID 83706, USA

[4] Department of Fish, Wildlife and Conservation Ecology, New Mexico State University, Las Cruces, NM 88003, USA

[5] U.S. Geological Survey, New Mexico Cooperative Fish and Wildlife Research Unit, Department of Fish, Wildlife and Conservation Ecology, New Mexico State University, Las Cruces, NM 88003, USA

[6] Division of Migratory Bird Management, U.S. Fish and Wildlife Service, 2105 Osuna NE, Albuquerque, NM 87113, USA

[7] National Park Service, 4175 Geist Road, Fairbanks, AK 99709, USA

[8] U.S. Geological Survey, Alaska Science Center, 4210 University Dr., Anchorage, AK 99508, USA

[9] Department of Biological Sciences, Duquesne University, 600 Forbes Ave., Pittsburgh, PA 15282, USA

[10] Department of Biochemistry, Selcuk University, Konya 42031, Turkey

[11] Bloom Biological Inc., 13611 Hewes Ave., Santa Ana, CA 92705, USA

[12] Department of Biological Sciences, Purdue University, 915 W. State Street, West Lafayette, IN 47907, USA

Published online: 11 July 2016

Introduction

The delineation of population genetic structure often has significant management implications and is important for conservation efforts (Ryder 1986; Fraser and Bernatchez 2001). Moritz (1994) argued that management units (i.e., populations that differ in allele frequencies) are of most relevance for population monitoring and demographic study. Fortunately, modern genetic markers can be used to identify management units but also to identify individuals, estimate population sizes (both census and effective), and to identify genomic signatures of natural selection. Such insights can profoundly improve our understanding of species' biology while leading to more effective conservation and management strategies (Avise 1994).

Golden eagles (*Aquila chrysaetos*) are sparsely distributed and usually difficult to sample (e.g., eyries are often found on sheer rock cliffs or in nest trees >20 m tall; Watson 2010). Consequently, their population genetic structure is poorly understood. Like other vagile predators with expansive natural distributions (e.g., some sharks; Ovenden et al. 2011), golden eagles may be genetically homogenous over broad geographic regions. However, natal philopatry or population bottlenecks associated with anthropogenic pressures may have sundered some eagle populations to the point where they should be considered distinct management units (Bourke et al. 2010; Millsap et al. 2014). For example, mitochondrial DNA (mtDNA) has revealed some genetic discontinuities in golden eagles (Nebel et al. 2015; also see Sonsthagen et al. 2012), but inferences of avian dispersal based on mtDNA are limited because of its maternal inheritance (Avise 2000). Bi-parentally inherited microsatellites provide some evidence of structure between island and mainland golden eagle populations (Sonsthagen et al. 2012; Ogden et al. 2015), but little evidence of continental structure (Wheeler 2014). Studies have recently shown, however, that single nucleotide polymorphisms (SNPs) and gene-associated markers have the capacity to reveal genetic structure undetected by mtDNA and/or microsatellites (Ruegg et al. 2014; Bekkevold et al. 2015; Malenfant et al. 2015), indicating they may be especially useful for studies of golden eagle biology.

Beyond population structure, SNP arrays have the potential to reveal differential selection among subspecies, populations, cohorts, or genomic regions (see the January 2016 issue of Molecular Ecology; e.g., Haasl and Payseur 2016). For example, SNP arrays have identified signals of selection associated with osmoregulation, salinity tolerance and growth genes in threespine stickleback (*Gasterosteus aculeatus*) populations (Ferchaud et al. 2014) as well as correlations between outlier genes and temperature and salinity parameters amongst Atlantic herring (*Clupea harengus*) populations (Limborg et al. 2012). Evidence of differential selection amongst populations may indicate local adaptation to environmental conditions, an important consideration when delineating management units (Funk et al. 2012; Shafer et al. 2015). Although golden eagles are found in a variety of habitats across the Northern Hemisphere (Watson 2010), it is not clear whether selection associated with different geographic regions should influence management decisions.

SNPs associated with functional genes may also shed light on selection among cohorts within the same population. There is considerable variance in golden eagle juvenile survivorship (McIntyre et al. 2006), and this variance could be due in part to heterozygosity-fitness correlations (HFCs; whereby more heterozygous individuals are more fit; see Clegg and Allard 1973; Mitton 1997) that are maintained by viability selection. Viability selection occurs when homozygosity has a negative effect on survival to adulthood, resulting in a more heterozygous adult population relative to juveniles (Clegg and Allard 1973; Cohas et al. 2009; Lampila et al. 2011). Studies of viability selection and HFCs have inconsistently identified significant effects (Chapman et al. 2009), but this inconsistency may stem from how genome-wide heterozygosity is calculated. Past studies have commonly used genetic diversity at microsatellite loci as an estimate of genome-wide heterozygosity, but SNPs sampled at high density will more accurately convey genome-wide heterozygosity (DeWoody and DeWoody 2005; Hoffman et al. 2014). The idea that SNPs provide a powerful platform to address natural selection is buttressed by theoretical work (Turelli and Ginzburg 1983) that indicates relatively uniform allele frequencies and the large number of loci surveyed should make SNP panels especially useful in detecting evidence of viability selection. The recent publication of the golden eagle genome sequence (Doyle et al. 2014) may provide mechanistic insights into such selective processes.

Herein, we validate and utilize a novel SNP assay for the golden eagle to address broad-scale patterns of population genetic structure and natural selection. Our goals were to: (1) conduct a preliminary test of the (null) panmixia hypothesis among samples collected across parts of North America, (2) test for viability selection and heterozygote advantage, and (3) identify candidate genes under selection amongst golden eagle populations.

Methods

Sample collection

Golden eagle feather, bone, or tissue samples were collected from Alaska, California, other western U.S. states

(Arizona, Colorado, Nebraska, New Mexico, Utah, Wyoming), and eight eastern U.S. states (Alabama, Georgia, North Carolina, New York, Pennsylvania, Tennessee, Virginia and West Virginia) (see Fig. 1). Samples were derived from regions with known migrants (Alaska, eastern U.S.) and regions that include year-round residents as well as migrants (western U.S.). Multiple independent replicates from 15 turbine-killed individuals (e.g., multiple feathers from a single carcass; Katzner et al. forthcoming) were included to assess genotyping error. To confirm Mendelian inheritance of our novel markers, we also included samples from 23 family groups that represent presumptive parent-offspring dyads. These family group samples consisted of adult feathers collected from nesting territories and blood sampled directly from the respective nestlings as in Rudnick et al. (2005). We also typed nine bald (*Haliaeetus leucocephalus*), 14 white-tailed (*H. albicilla*), and 14 imperial eagle (*Aquila heliaca*) samples to evaluate cross-species amplification of SNPs (Online Resources 1, 2).

SNP assay development

We aligned sequence reads from a single male golden eagle to the draft golden eagle genome assembly of Doyle et al. (2014) using BWA (Li and Durbin 2009). We then used Picard (http://broadinstitute.github.io/picard) to sort mapped reads and identify duplicates. We used GATK 3.2 (DePristo et al. 2011; Van der Auwera et al. 2013) to (1) identify and realign reads around insertions/deletions (indels), (2) call high-quality SNPs (Phred quality score of 50, no more than two alleles for nuclear SNPs) while masking indels, and (3) perform base recalibration. Base recalibration adjusts quality scores to reflect the fact that sequencing quality decreases with increasing numbers of cycles (Schunter et al. 2014). We subsequently used the recalibrated data set to identify SNPs with a minimum Phred quality score of 30 and a minimum depth of ten reads.

We used SNPdat 1.0.5 (Doran and Creevey 2013) to determine whether nuclear SNPs were present in exonic, intronic, or intergenic regions and BEDOPS 2.0 (Neph et al. 2012) to measure the distance between genes and SNPs. We subsequently identified markers in presumptive gene deserts (95 % percentile distance from known genes) as potentially "neutral" markers. SNPs present in the exons and introns of genes were annotated using BLAST® 2.2.3. We used IGV 2.3 (Robinson et al. 2011) to identify target SNPs with at least 60 nucleotides of high-quality flanking sequence upstream and downstream, GC content less than 65 %, and no other variable sites within 20 nucleotides.

Ultimately, we developed 95 autosomal nuclear markers from gene deserts and 94 autosomal nuclear markers from protein-coding genes. Over half (48) of the gene-associated markers were specifically targeted because of evidence for selection in other species (Table 1). We also developed two mitochondrial markers and one molecular-sexing marker for use with the Fluidigm® SNP Type ™ assay. The two mitochondrial DNA (mtDNA) markers were identified in preliminary population surveys via dideoxy sequencing. Briefly, we amplified a portion of the mtDNA control region using SpiCR31 and Pro primer sequences; PCR conditions and thermal profile were modified from Cadahía et al. (2009). To distinguish male from female golden eagles, we developed a novel molecular sexing marker (Online Resource 1) to assess nucleotide variation at a single site in the CHD1 gene using NCBI accessions AB096148 and AB096147. CHD1 was the first avian gene adapted for molecular sexing, as consistent polymorphisms between the Z- and W-chromosomes are used to distinguish males from females (Fridolfsson and Ellegren 1999).

SNP genotyping

We attempted to genotype >600 golden eagle samples at 192 candidate markers. Samples were genotyped using a Fluidigm® EP1™ Genotyping System and a specific target amplification step was incorporated to facilitate genotyping of low-quantity DNA samples. Individual SNP calls were visualized and edited using Fluidigm's® Genotyping Analysis Software; data from nuclear loci that did not produce obvious clusters of homozygotes and heterozygotes were excluded from further analyses. We also excluded data from poor quality DNA samples for which >3 % of the markers (i.e., n = 5) did not amplify (e.g. Ruegg et al. 2014).

Our data set included unknown replicate samples (e.g., multiple feathers from the same adult collected from a single eyrie). We used allelematch in R (Galpern et al. 2012) to group replicate genotypes into unique records that represent individual eagles. We allowed for as many as seven mismatches between replicates, based on preliminary analyses with the diagnostic function amUniqueProfile, which determines the number of mismatches that best allow replicate genotypes to sort into groups with minimal overlap (Galpern et al. 2012). All genotypes identified as replicates were subsequently confirmed visually.

As we developed a novel SNP assay, a number of analyses were undertaken to test the veracity of each marker. These analyses included (1) testing the repeatability of the entire assay using independent, replicate samples (Online Resources 1, 3); (2) conducting molecular sexing using both our novel CHD1 marker and a traditional PCR/gel method (Online Resource 1); and (3) comparing dideoxy mtDNA sequences with haplotypes derived from our novel SNP mtDNA markers (Online Resource 1). We also tested for linkage disequilibrium between markers (Online Resources 1, 4), compared genetic variation

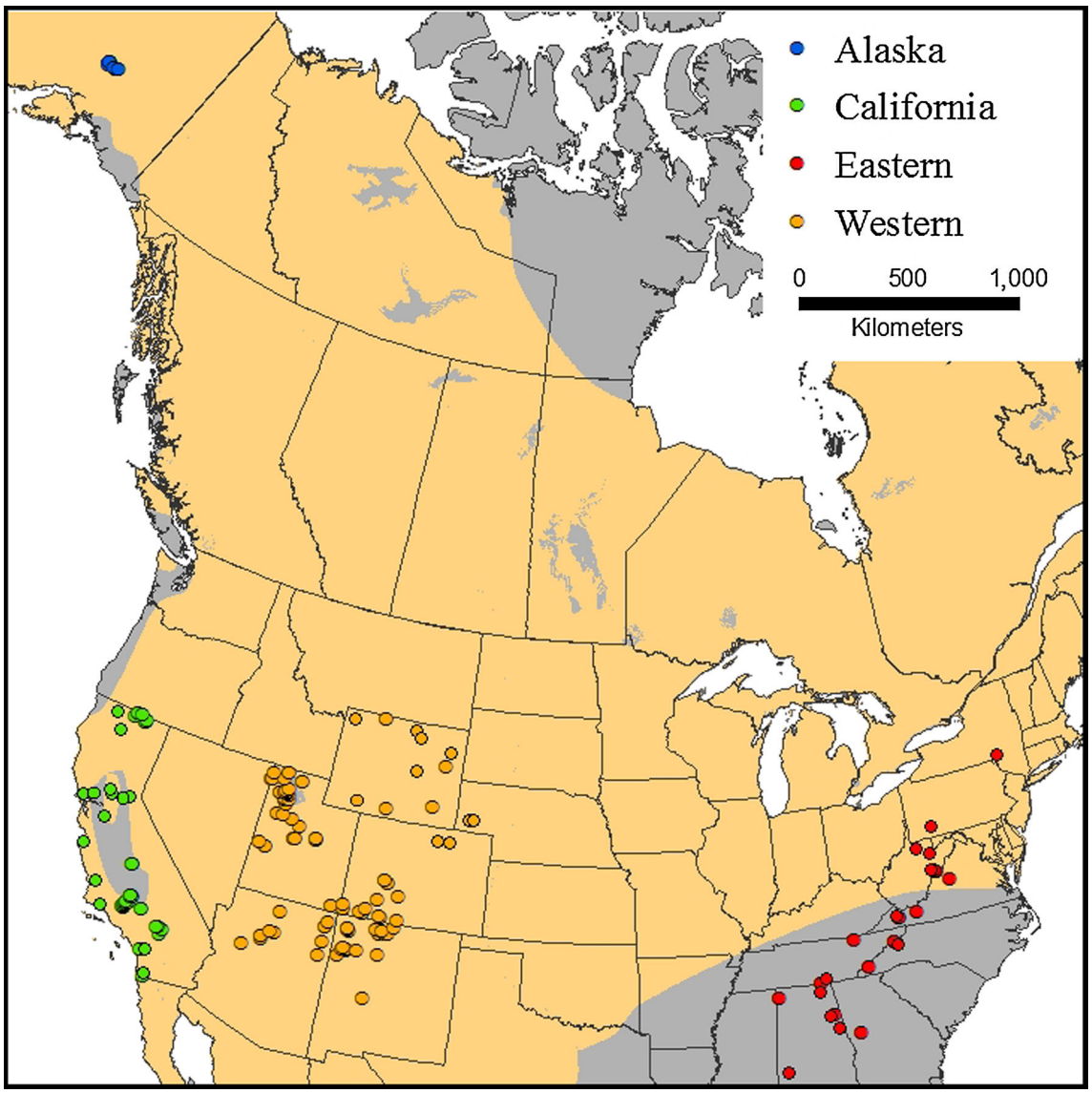


Fig. 1 Sampling locations including Alaska, California, other western states (Arizona, Colorado, Nebraska, New Mexico, Utah, Wyoming), and the eastern U.S. (i.e., Alabama, Georgia, North Carolina, New York, Pennsylvania, Tennessee, Virginia and West Virginia). When GPS coordinates for the sampling sites were not available, we used the centroid coordinates for the county where sampling occurred. The *yellow* shading indicates a portion of the IUCN range map. (Color figure online)

calculated from microsatellite and SNP markers (Online Resource 1), and conducted preliminary parentage analyses of golden eagle chicks (Online Resources 1, 5).

Heterozygosity and viability selection

GenAlEx 6.501 (Peakall and Smouse 2012) was used to calculate observed (H_O) and expected (H_E) heterozygosity. To test for viability selection (i.e., differential survival to adulthood), we compared the number of heterozygous loci in chicks to that of adult golden eagles. Chick samples were collected from eaglets still in the nest or those recently fledged, whereas adult samples were known breeders (identified via parentage analyses) or physically trapped birds that were aged according to molt limits (Jollie 1947; Bloom and Clark 2001). A much larger number of samples were derived from eagles of unknown age (hereafter referred to as non-chicks), but most were presumably preadults (i.e., post-fledgling juveniles and subadults less than 3 years of age) or adults. We compared the number of heterozygous loci in chicks ($n = 160$) to that of (1) known adults ($n = 30$) and (2) non-chicks (including both known and presumed adults as well as preadults; $n = 184$) using t-tests implemented in SigmaPlot 13.

Table 1 Description of 48 golden eagle SNPs associated with genes under selection in different species

Scaffold_position	Gene	Species	Citation
2285379_609976	BMP5	*Homo sapiens*	Voight et al. (2006)
2283606_765703	TLR3	*Gallus* sp.	Downing et al. (2010)
2285864_176661	IGF1R	*Gallus* sp.	Rubin et al. (2010)
2285559_384243	NASP	*Homo sapiens*, *Mus* sp.	Torgerson et al. (2002)
2284984_420012	TRIL	*Gallus* sp.	Downing et al. (2010)
2285293_608855	ZPBP	Mammal species	Swanson et al. (2001)
2285052_1961290	TLR5	*Gallus* sp.	Downing et al. (2010)
2285052_1961617	TLR5	*Gallus* sp.	Downing et al. (2010)
2285696_1028719	TBC1	*Gallus* sp.	Rubin et al. (2010)
2285438_337905	DRD4	*Parus major*	Mueller et al. (2013)
2285493_1252867	SOD1	*Chrysemys picta bellii*	Shaffer et al. (2013)
2284700_1489646	IFRD1	*Tursiops truncatus*	Nery et al. (2013)
2284756_267809	GRIN2A	*Taeniopygia guttata*	Nam et al. (2010)
2284841_422323	CENPJ	Mammal species	Evans et al. (2006)
2285862_268775	CDK5RAP2	Mammal species	Evans et al. (2006)
2283899_1490284	BMP4	*Falco peregrinus, Falco cherrug*	Zhan et al. (2013)
2284519_18585	Egeu-DAB1	*Agelaius phoeniceus*	Edwards et al. (1998)
2285364_777695	LRR1	*Geospiza magnirostris*	Rands et al. (2013)
2285348_755766	ANO10	*G. magnirostris*	Rands et al. (2013)
2285411_3331479	POU1F1	*G. magnirostris*	Rands et al. (2013)
2285451_264953	CCDC40	*G. magnirostris*	Rands et al. (2013)
2283951_456619	DTX3L	*G. magnirostris*	Rands et al. (2013)
2285967_60970	LRRC34	*G. magnirostris*	Rands et al. (2013)
2283459_581935	LRRC7	*Parus humilis*	Qu et al. (2013)
2284874_1176738	SPAM1	Mammal species	Kosiol et al. (2008)
2285287_2291260	MTOR	*Pseudopodoces humilis*	Cai et al. (2013)
2286215_568521	ZNFX1	*P. humilis*	Cai et al. (2013)
2286215_578202	ZNFX1	*P. humilis*	Cai et al. (2013)
2285088_1297347	ST6	*P. humilis*	Cai et al. (2013)
2285556_4721393	PIK3AP1	*Taeniopygia guttata*	Ekblom et al. (2010)
2285969_325201	GNS	*Alligator sinensis*	Wan et al. (2013)
2286096_174173	DHRS7C	*A. sinensis*	Wan et al. (2013)
2285189_455264	CREBBP	*A. sinensis*	Wan et al. (2013)
2285361_106037	TTN	*A. sinensis*	Wan et al. (2013)
2284110_1252933	GLB1L	*A. sinensis*	Wan et al. (2013)
2286471_3407124	OTOP1	*A. sinensis*	Wan et al. (2013)
2284988_747550	CNTN1	*A. sinensis*	Wan et al. (2013)
2285029_440974	SLC44A5	*A. sinensis*	Wan et al. (2013)
2284495_136729	GIGYF2	*A. sinensis*	Wan et al. (2013)
2280973_6559	PARP14	*A. sinensis*	Wan et al. (2013)
2284896_580802	DHRS7B	*A. sinensis*	Wan et al. (2013)
2283606_439166	SORBS2	*Taeniopygia guttata*	Warren et al. (2010)
2284858_2927263	MYH10	*T. guttata*	Warren et al. (2010)
2284858_2932994	MYH10	*T. guttata*	Warren et al. (2010)
2283532_528327	TRPV1	*T. guttata*	Warren et al. (2010)
2285862_3903664	CACNA1B	*T. guttata*	Warren et al. (2010)
2285761_825420	SP4	*T. guttata*	Warren et al. (2010)
2284393_111885	TMEM39B	*T. guttata*	Warren et al. (2010)

Putative gene functions are further described in Online Resource 6

Population structure

A pool of eaglets, subadults and adults (Table 2) was used to calculate N_e using the LD method (assuming monogamy) implemented with NeEstimator 2.01 (Do et al. 2014). We then tested the null hypothesis that golden eagles in the U.S. consist of a single panmictic population. Our initial test of the panmixia hypothesis used chicks that had not yet fledged and thus represented pre-dispersal individuals; thus these samples yielded natal genotypes of known geographic provenance (24 from Alaska, 4 from Arizona, 45 from Utah, 29 from California, 21 from Colorado, 4 from Nebraska, 20 from New Mexico, and 13 from Wyoming for a total of 160 eaglets). We tested panmixia using the Bayesian analysis implemented with STRUCTURE 2.3.4 (Pritchard et al. 2000) and Structure Harvester (Earl and vonHoldt 2012). Results were interpreted using mean likelihood values of K and ΔK. If ≥2 populations were identified during an initial STRUCTURE run, we followed Pritchard (2010) and ran STRUCTURE again using only individuals from each previously identified population (i.e. we ran STRUCTURE iteratively with subsets of the complete data set). We continued this process until no additional population structure could be identified.

We also tested panmixia by using 318 individuals (eaglets, pre-adults, and adults). This represents a less conservative approach (e.g., because samples collected from a given site could represent unknown dispersers) but allowed us to include samples from the eastern United States. These eastern North American samples derive from the breeding population in Northeastern Canada and represent a relatively discrete breeding population (Katzner et al. 2012; Morneau et al. 2015). In the case of feathers and blood samples collected from nesting territories, we retained adult but removed chick genotypes to prevent clustering algorithms from confusing family groups for population structure (Anderson and Dunham 2008). Otherwise, our initial analyses utilized all of the genotypes identified from all samples (Table 2). Once putative populations were identified, we estimated F-statistics with diveRsity (Keenan et al. 2013). To identify candidate genes under selection and to increase the likelihood of detecting population structure, we compared locus-specific pairwise F_{ST} values between both gene-associated and neutral nuclear markers. For mtDNA data, we used GenAlEx to test for genetic differentiation between putative populations by AMOVA and using 999 random permutations to test for significance (Peakall and Smouse 2012).

Results

SNP genotyping

We identified 1,239,990 candidate SNPs, of which 819,540 passed quality filtering criteria. Most of these (560,985) were in intergenic regions, but 212,118 were in genic regions (8473 in exons and 203,645 in introns). From these, we designed and tested 192 SNPs but ultimately excluded

Table 2 Number of samples and observed and expected heterozygosities for Alaskan, Californian, western and eastern golden eagles

	Individual sample size	Age	Observed heterozygosity	Expected heterozygosity
Alaska	24	Chicks	0.31 ± 0.02	0.31 ± 0.01
California	177	Adults, non-chicks, chicks	0.34 ± 0.01	0.35 ± 0.01
Western U.S.	113		0.32 ± 0.01	0.33 ± 0.01
New Mexico	20	Chicks		
Colorado	21	Chicks		
Nebraska	4	Chicks		
Utah	45	Chicks		
Wyoming	13	Chicks		
Arizona	10	Adults, non-chicks, chicks		
Eastern U.S.	30		0.28 ± 0.02	0.27 ± 0.01
Alabama	7	Adults, non-chicks		
Georgia	1	Adult		
New York	1	Non-chick		
North Carolina	4	Adults		
Pennsylvania	1	Adult		
Tennessee	2	Adults		
Virginia	11	Adults, non-chicks		
West Virginia	3	Adults		

30 markers that were monomorphic, clustered poorly, or were otherwise of low quality. These novel SNP loci and flanking sequences are described in Online Resources 6, 8. The 162 informative loci include 76 gene-associated nuclear markers (which may be targets of selection), 83 intergenic nuclear markers (which we assume evolve neutrally), 2 mitochondrial markers, and 1 nuclear molecular-sexing marker (Online Resources 6, 8). Ninety-nine low-quality samples (i.e., those where >5 SNPs did not amplify) were discarded from our analyses, meaning that the >600 samples we attempted to assay resulted in a total of 523 multilocus genotypes. After removing replicate genotypes (e.g., those that represent two or more feathers from the same adult), these 523 genotypes were assigned to 344 individual golden eagles sampled from 16 different states (Table 2, Online Resource 7).

Heterozygosity and viability selection

Average H_O and H_E at autosomal SNPs were 0.33 ± 0.01 and 0.34 ± 0.01, respectively, for all North American golden eagles sampled (Table 2). Twenty-one of 159 loci were not in Hardy–Weinberg equilibrium. Of these 21 SNPs, 13 were gene-associated and 8 were intergenic SNPs. Eighteen of these 21 SNP loci had F_{IS} values ranging from 0.12 to 0.40 (heterozygote deficiency), whereas only three of these loci had F_{IS} values ranging from −0.40 to −0.11 (heterozygote excess). Heterozygosity calculated with gene-associated markers ($H_O = 0.35 \pm 0.02$, $H_E = 0.36 \pm 0.02$) was not significantly different from heterozygosity calculated with intergenic neutral markers ($H_O = 0.31 \pm 0.02$, $H_E = 0.32 \pm 0.02$; Mann–Whitney Rank Sum Test: $U = 6563$, $p = 0.1$; Fig. 2).

Known adult H_O and H_E averaged 0.35 ± 0.01 and 0.35 ± 0.01, respectively, across all autosomal markers. Non-chick H_O and H_E averaged 0.33 ± 0.01 and 0.34 ± 0.01 across all autosomal markers. Finally, chick H_O and H_E averaged 0.32 ± 0.01 and 0.33 ± 0.01, respectively, across all autosomal markers. Observed heterozygosity at gene-associated and putatively neutral markers, respectively, was 0.36 ± 0.02 and 0.33 ± 0.02 for adults, 0.35 ± 0.02 and 0.32 ± 0.02 for non-chicks, and 0.35 ± 0.02 and 0.31 ± 0.02 for chicks. To test the idea of viability selection (e.g., heterozygote advantage, whereby fitness decreases in concert with homozygosity), we compared the number of heterozygous loci in chicks to that of (1) known adults and (2) preadults and adults (i.e., non-chicks). Chicks had decreased numbers of heterozygous loci when compared to both known adults ($t_{1,188} = 2.77$, $p = 0.006$) and non-chicks ($t_{1,342} = 1.83$, $p = 0.07$). To visualize the comparisons between chicks, known adults and non-chicks, we generated 1000 bootstrap replicates where the 2.5th and 97.5th percentiles represent the 95 % confidence intervals for each age group (Fig. 3). We conducted an additional analysis comparing numbers of heterozygous loci between chicks and adults from a single geographic region (the Tehachapi mountain range in California) and found results similar to that of the full data set (data not shown).

Population structure

Our preliminary estimate of N_e, based on LD and a monogamous breeding system, was 398 ± 30. STRUCTURE analysis of 160 known-provenance individuals (i.e., eaglets) provide evidence that the samples we collected from across the western U.S. cluster into three genetically distinct populations (Fig. 4a) corresponding to Alaska, California and the other western states. Eaglets sampled from Alaska and California clustered together with particularly high affinity; the average probability of individuals being assigned to Alaskan and Californian clusters was 0.77 ± 0.15 and 0.71 ± 0.20 SD, respectively (Fig. 4a). More admixture was apparent within all other western birds, with individuals from the other six states sampled assigned to the cluster with an average probability of 0.50 ± 0.24 SD. We removed Alaskan and Californian samples from the data set and ran STRUCTURE once more using the remaining western samples, but detected no additional population structure. We additionally ran STRUCTURE after omitting loci out of Hardy–Weinberg equilibrium and again found evidence of three genetically distinct populations (data not shown).

We subsequently tested the panmixia hypothesis using a mixture of known and unknown provenance individuals (n = 318 golden eagles total). Sample sizes were identical to those shown in Table 2 with the exception of California where, after removing eaglet samples that represented members of family units (see methods), we considered 151 samples. The initial STRUCTURE analysis of panmixia rejected that hypothesis and provides evidence for genetic structure between Californian individuals and all others (Online Resource 1). We removed Californian samples from the data set and conducted an additional STRUCTURE analysis with the remaining samples. Without samples from California, all birds from the western states clustered separately from those sampled in the eastern states (Fig. 4b). A mtDNA AMOVA indicated that the majority of genotypic diversity was distributed within geographic sampling sites (97 %) rather than among sampling sites (3 %). This analysis indicated significant genetic differences among global populations ($\Phi_{PT} = 0.03$, $P = 0.03$), largely driven by differences between California and the eastern states ($\Phi_{PTP} = 0.10$, $P < 0.01$) and between California and the six other western states ($\Phi_{PTP} = 0.02$, $P = 0.02$).

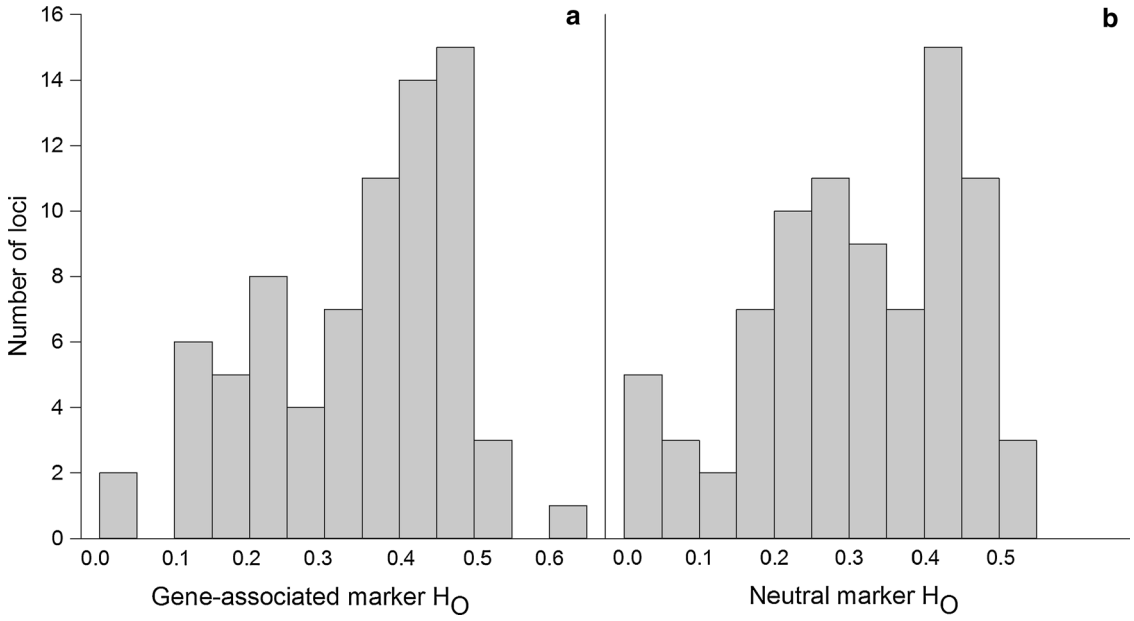


Fig. 2 Histogram showing observed heterozygosity on the x-axis and number of loci on the y-axis for **a** gene-associated loci and **b** putatively neutral loci

Given the results of our Bayesian analysis and the amount of genetic admixture present in the six western states, we calculated pairwise F_{ST} values for four presumptive populations: Alaska, California, eastern states and the six western states (Table 3). Significant population differentiation was evident in Alaska versus eastern, Alaska versus western and eastern versus western pairwise comparisons (i.e. 95 % confidence intervals did not overlap zero, Online Resource 1). There were no significant differences between F_{ST} values calculated with gene-associated and putatively neutral markers (Online Resource 1). A number of individual loci (both gene-associated and putatively neutral) had F_{ST} values ranging from 0.05 to 0.5, indicating moderate to substantial genetic differentiation (Fig. 5; Table 3). Genetic differentiation is expected of markers under different selective regimes in different environments, and Table 3 identifies markers with pairwise F_{ST} values greater than 0.10 in at least one pairwise population comparison. For example, F_{ST} values were significantly elevated among SNPs within the exons of genes implicated in cranial facial development (POU1F1), bone morphology (BMP4, BMP5), muscle development (AXIN1), immunity (HP, TRIL, TLR5), sperm competition (SPATA5, SPAG9), and oxidative stress (SOD1).

Discussion

Herein, we find evidence of population structure, differential selection, and viability selection in golden eagles using a powerful SNP genotyping panel. Our SNPs were identified using the genome sequence derived from a single individual (Doyle et al. 2014), in contrast to the more common approach of using reduced representation sequencing of multiple individuals as the raw material for calling SNPs (Senn et al. 2013; Ferchaud et al. 2014; Ruegg et al. 2014). The SNPtype assay performed well with low-copy DNA extracted from naturally shed feathers and degraded tissue. In our case, high quality genotypes were produced for 84 % of our field-collected samples (i.e., those that passed our filtering step). Furthermore, samples that passed our filtering step had a low error rate (0.4 %, Online Resource 1).

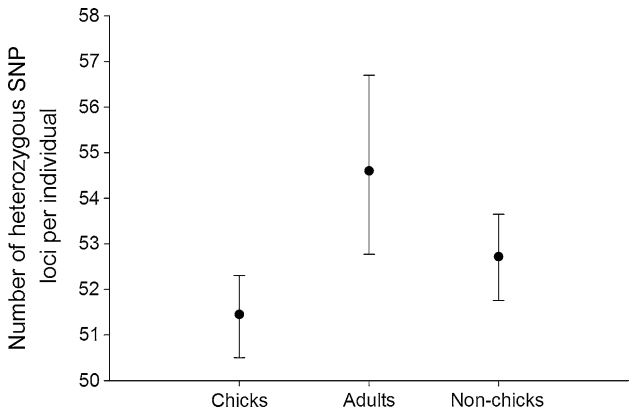


Fig. 3 Mean numbers of heterozygous loci (with 95 % CI) in chicks, adults and non-chicks. Chicks had decreased numbers of heterozygous loci when compared to both known adults ($t_{1,188} = 2.77$, $p = 0.006$) and non-chicks ($t_{1,342} = 1.83$, $p = 0.07$)

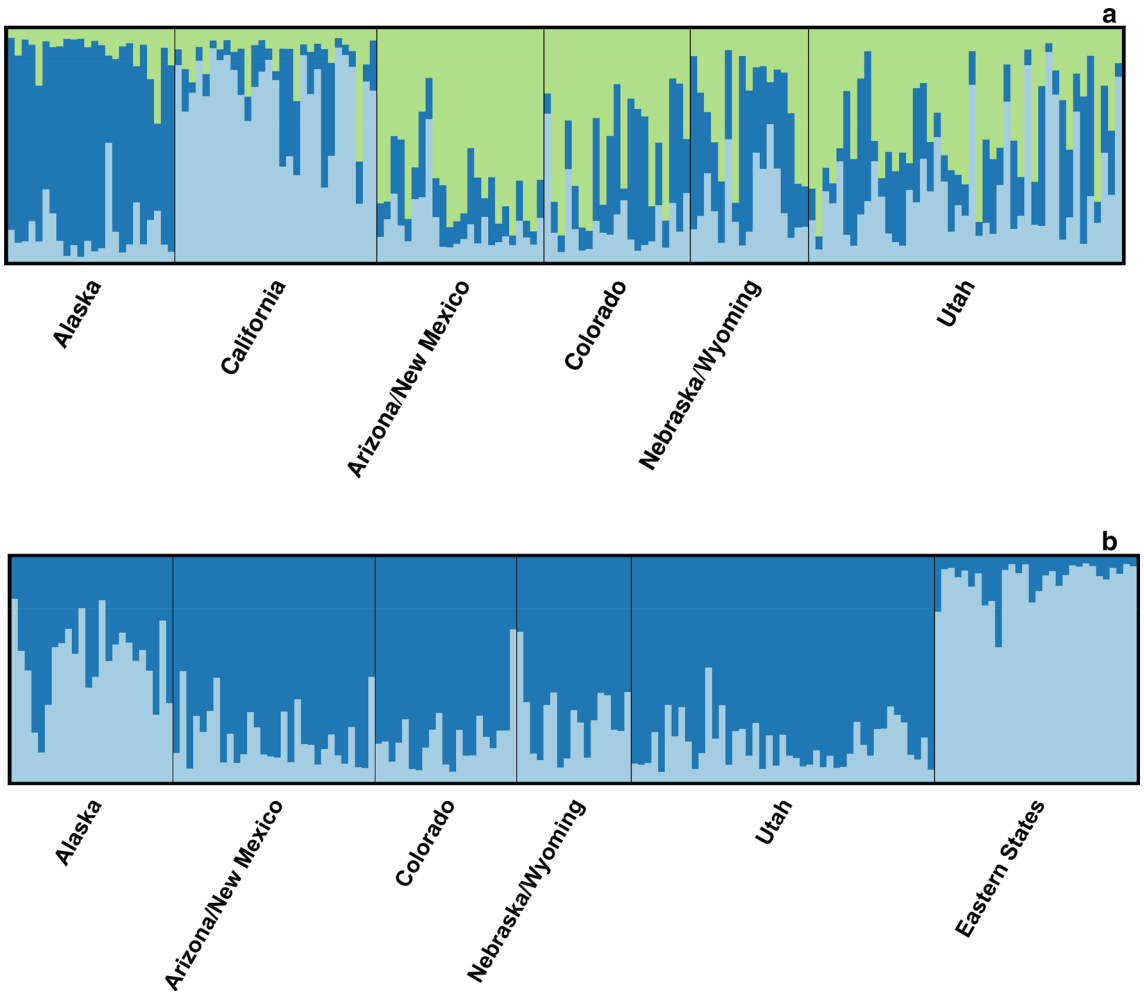


Fig. 4 STRUCTURE results for known and unknown-provenance eagles. **a** Results of STRUCTURE analysis for 160 known-provenance chicks sampled from Alaska, California, Arizona, New Mexico, Colorado, Nebraska, Wyoming, and Utah that were genotyped at 159 SNP loci. STRUCTURE results were CLUMPP-averaged across 10 runs when K is assumed to be equal to three. **b** Results of STRUCTURE analysis for a mix of 167 known and unknown provenance individuals sampled in Alaska, Colorado, Arizona/New Mexico, Utah, Nebraska/Wyoming and eastern states genotyped at 159 SNP loci. STRUCTURE results were CLUMPP-averaged across 10 runs when K is assumed to be equal to two

Heterozygosity and viability selection

Genome-wide genetic diversity is often correlated with fitness (Mitton and Pierce 1980; Chakraborty 1981; Chapman et al. 2009; Szulkin et al. 2010). By genotyping eagle chicks and adults at SNP loci, we tested for a heterozygosity-fitness correlation through the mechanism of viability selection and found that golden eaglets have fewer heterozygous loci than adults. This is consistent with the argument that homozygosity has a negative effect on survival to adulthood and results in a more heterozygous adult population relative to juveniles (see Clegg and Allard 1973; Cohas et al. 2009; Lampila et al. 2011). Both genome-wide and locus-specific effects have been proposed to explain relationships between heterozygosity and differential survival (Mitton and Pierce 1980; Chakraborty 1981; Chapman et al. 2009; Szulkin et al. 2010). If heterozygosity serves as a proxy for the inbreeding coefficient (Keller and Waller 2002), this suggests the reduced heterozygosity observed in golden eagle chicks relative to adults may be partly attributable to inbreeding (a genome-wide effect). Other eagles in the same genus (e.g., *A. heliaca*) mate randomly with respect to relatedness (Rudnick et al. 2005), indicating that inbreeding occasionally happens by chance alone. Thus, one possible mechanism underlying the reduced heterozygosity in golden eagle chicks would be viability selection against inbred individuals that express deleterious, recessive alleles.

Table 3 SNPs with F_{ST} values greater than 0.10 in golden eagles from Alaska (AK), California (CA), and the eastern and western U.S

Locus ID	Ontology	Function	Pairwise FST values						
			Mean	AK/ CA	AK/ eastern	AK/ western	CA/ eastern	CA/ western	Eastern/ western
2286490_5809571			0.18	0.14	0.47	0.27	0.15	0.05	0.02
2285223_2776527			0.15	0.17	0.02	0.08	0.35	0.00	0.24
2285411_1155264			0.15	0.20	0.32	0.00	0.01	0.16	0.25
2283991_107866	LOC101750972[a]	Dystroglycan-like	0.13	0.10	0.38	0.03	0.11	0.01	0.18
2285411_3331479	POU1F1	POU class 1 homeobox 1	0.13	0.12	0.31	0.00	0.05	0.07	0.21
2285415_11368			0.12	0.10	0.06	0.07	0.28	0.00	0.24
2284289_123895	TPP2	Tripeptidyl peptidase II	0.11	0.00	0.28	0.00	0.21	0.00	0.17
2286534_2639559			0.11	0.06	0.29	0.02	0.09	0.01	0.18
2283459_581935	LRRC7	Leucine rich repeat containing 7	0.10	0.10	0.03	0.03	0.25	0.01	0.16
2285150_2303082	SPAG9	Sperm associated antigen 9	0.10	0.07	0.08	0.00	0.29	0.04	0.13
2283765_39193	AXIN1	Axin 1	0.09	0.13	NA	0.07	0.14	0.03	0.08
2285556_4721393	PIK3AP1	Phosphoinositide-3-kinase adaptor protein 1	0.09	0.05	0.27	0.06	0.08	0.00	0.07
2283606_765703	TLR3	Toll-like receptor	0.08	0.00	0.08	0.03	0.14	0.01	0.22
2283628_836853	DCHS1	Dachsous cadherin-related	0.08	0.18	0.00	0.00	0.10	0.18	0.01
2284393_111885	TMEM39B	Transmembrane protein 39B	0.08	0.04	0.00	0.16	0.06	0.03	0.19
2286215_578202	ZNFX1	Zinc finger	0.08	0.01	0.09	0.03	0.15	0.00	0.18
2283707_1417632	HP	Haptoglobin	0.07	0.08	0.00	0.04	0.18	0.00	0.12
2284301_376772	FBXO25	F-box protein 25	0.07	0.10	0.02	0.02	0.18	0.04	0.09
2284495_136729	GIGYF2	GRB10 interacting GYF protein 2	0.07	0.08	0.00	0.06	0.14	0.00	0.11
2284530_2060961			0.07	0.14	0.00	0.04	0.16	0.03	0.05
2285084_6689570			0.07	0.11	NA	0.03	0.11	0.05	0.03
2285245_2006110			0.07	0.04	0.17	0.17	0.04	0.02	0.01
2285246_2800379			0.07	0.01	0.04	0.03	0.00	0.11	0.21
2285510_1183382			0.07	0.00	0.10	0.00	0.11	0.01	0.18
2285660_2872308			0.07	0.07	0.03	0.06	0.13	0.00	0.11
2286243_1117656			0.07	0.11	NA	0.07	0.11	0.01	0.08
2283486_1788205			0.06	0.07	0.17	0.08	0.01	0.00	0.01
2283918_1095780			0.06	0.05	0.08	0.20	0.00	0.03	0.00
2283967_482740			0.06	0.16	0.00	0.10	0.08	0.00	0.04
2284530_1372527	ZNF407	Zinc finger protein 407	0.06	0.00	0.14	0.01	0.17	0.03	0.05
2284761_588233			0.06	0.06	0.05	0.16	0.00	0.03	0.04
2284984_420012	TRIL	TLR4 interactor with leucine-rich repeats	0.06	0.07	0.04	0.00	0.17	0.03	0.08
2285029_440974	SLC44A5	Solute carrier family 44	0.06	0.00	0.18	0.00	0.10	0.00	0.08
2285052_1961290	TLR5	Toll-like receptor 5	0.06	0.04	0.18	0.07	0.05	0.00	0.02
2285335_3668467			0.06	0.18	0.06	0.13	0.01	0.00	0.00
2285361_4224079	MDH1B	Malate dehydrogenase 1B	0.06	0.00	0.12	0.04	0.13	0.04	0.01
2285493_1252867	SOD1	Dismutase 1	0.06	0.04	0.05	0.03	0.11	0.00	0.10
2285595_365835			0.06	0.05	0.03	0.00	0.16	0.02	0.08
2285660_3724254			0.06	0.02	0.02	0.06	0.10	0.01	0.17
2285664_480103			0.06	0.11	0.03	0.11	0.06	0.00	0.05
2285841_858411			0.06	0.14	0.00	0.07	0.11	0.01	0.04
2285862_3903664	CACNA1B	Calcium channel	0.06	0.00	0.10	0.01	0.09	0.01	0.15
2285931_1993730			0.06	0.04	0.00	0.12	0.04	0.01	0.12
2285971_690363			0.06	0.00	0.09	0.00	0.12	0.00	0.14
2286345_950542			0.06	0.07	0.01	0.00	0.16	0.07	0.03

Table 3 continued

Locus ID	Ontology	Function	Pairwise FST values						
			Mean	AK/ CA	AK/ eastern	AK/ western	CA/ eastern	CA/ western	Eastern/ western
2286471_3407124	OTOP1	Otopetrin 1	0.06	0.20	0.02	0.07	0.06	0.02	0.00
2286509_389510			0.06	0.07	0.01	0.00	0.15	0.07	0.03
2283532_124899	SPECC1	Sperm antigen	0.05	0.00	0.07	0.00	0.15	0.01	0.07
2284110_1252933	GLB1L	Galactosidase	0.05	0.11	0.12	0.07	0.00	0.00	0.00
2284120_644297			0.05	0.00	0.14	0.00	0.09	0.00	0.08
2284827_653144			0.05	0.07	0.00	0.03	0.10	0.00	0.07
2284841_422323	CENPJ	Centromere protein J	0.05	0.13	0.01	0.07	0.06	0.02	0.01
2284950_3419210			0.05	0.07	NA	0.02	0.08	0.03	0.03
2284950_488816	ATP7A	ATPase, Cu ++ transporting	0.05	0.10	0.11	0.00	0.00	0.04	0.04
2285088_1297347	ST6GALNAC1	ST6 -*N*-acetylgalactosaminide alpha-2,6-sialyltransferase 1	0.05	0.04	0.03	0.01	0.13	0.01	0.09
2285696_1028719	TBC1D1	TBC1 domain family, member 1	0.05	0.06	0.11	0.00	0.00	0.04	0.09
2286308_3354983			0.05	0.03	0.00	0.08	0.06	0.02	0.11
2286471_850630			0.05	0.05	0.01	0.00	0.14	0.03	0.04
2283853_2441306			0.04	0.00	0.10	0.00	0.07	0.00	0.04
2283899_1490284	BMP4	Bone morphology protein	0.04	0.12	0.03	0.08	0.02	0.00	0.00
2285050_912631			0.04	0.00	0.06	0.00	0.08	0.00	0.13
2285323_569693			0.04	0.00	0.03	0.05	0.08	0.11	0.00
2285457_726468	SPATA5	Spermatogenesis associated 5	0.04	0.00	0.00	0.05	0.00	0.10	0.09
2285511_1233747			0.04	0.08	0.00	0.00	0.12	0.04	0.02
2285683_2302104			0.04	0.05	NA	0.02	0.06	0.01	0.03
2285967_60970	LRRC34	Leucine rich repeat containing 34	0.04	0.06	0.00	0.02	0.12	0.01	0.06
2286009_2270274			0.04	0.06	NA	0.03	0.06	0.00	0.04
2286035_1612375			0.04	0.05	0.10	0.06	0.00	0.00	0.00
2283606_439166	SORBS2	Sorbin and SH3 domain containing 2	0.03	0.00	0.05	0.00	0.05	0.01	0.10
2284874_331516	CADPS2	Ca ++-dependent secretion activator 2	0.03	0.00	0.10	0.01	0.03	0.00	0.02
Global			0.04	0.04	0.05	0.03	0.06	0.02	0.06

[a] Markers without ontologies or functions are located in intergenic regions

Alternatively, heterozygote superiority (also referred to as overdominance) could also explain our results if more heterozygous golden eagles have greater capacity to cope with ecological and environmental fluctuations (reviewed in Mitton 1997). In this case, zygosity-driven viability selection may be due not to genome-wide impacts such as inbreeding, but to locus-specific effects (e.g., selection on individual genes). For example, toll-like receptor genes, or TLRs, are part of the innate immune system and code for proteins that recognize microbial peptides as potential pathogens (Grueber et al. 2013, 2015). Our data indicate that *TLR5* is associated with SNPs that are F_{ST} outliers that could be targets of selection (Table 3). Clearly, genome-wide and locus specific effects are not mutually exclusive, and future tests of the relationship between heterozygosity, survival, and adult reproductive success will better elucidate the mechanism(s) by which heterozygosity influences fitness in golden eagles.

Population structure and effective population size

Golden eagles are reclusive, occupy large and remote habitats, are sparsely distributed on the landscape, and are generally difficult to sample in large numbers. Our geographic sampling was largely opportunistic and our preliminary N_e estimate of $\sim$400 obviously does not encompass the Asian or European gene pools. Our N_e estimate is much smaller than estimates from a partial survey of western U.S. golden eagle population size, which

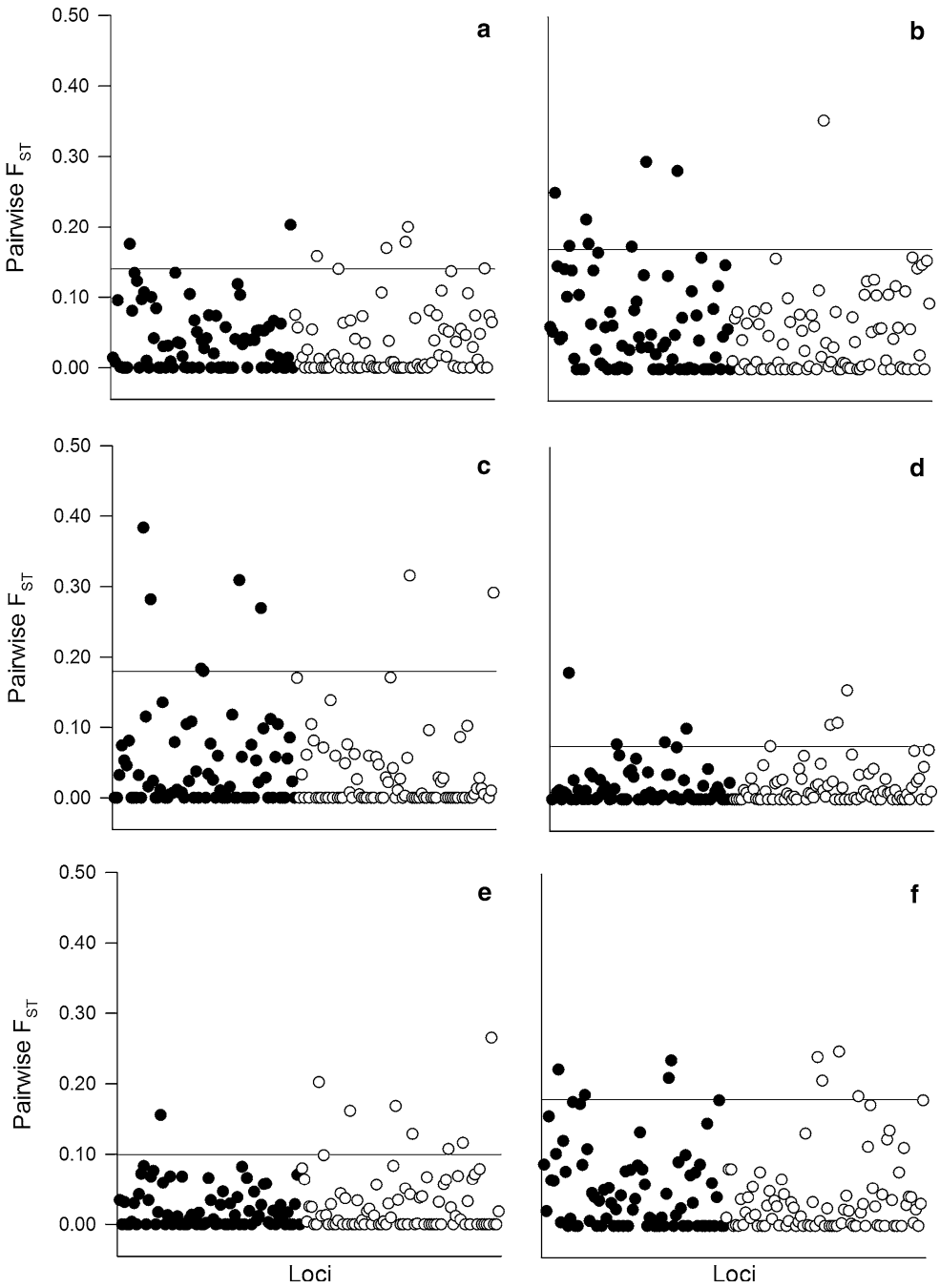


Fig. 5 Pairwise F_{ST} values for **a** Alaska versus California, **b** California versus eastern states, **c** Alaska versus eastern states, **d** California versus six western states, **e** Alaska versus six western states and **f** eastern versus western states at each SNP locus. SNP loci are ordered identically across the x-axis, with gene-associated markers in black and putatively neutral markers in white. Outliers found in the top 5 % of F_{ST} values represent markers that may be under selection, and are found above the horizontal reference line of each panel

Millsap et al. (2013) reports as between 25,450 and 41,700 (median ~32,000) individuals. Such a disparity between census and effective population size is not uncommon, as the two estimates often differ by at least an order of magnitude for a variety of reasons such as variance in family size (see Frankham 1995; Palstra and Ruzzante 2008). Palstra and Ruzzante's (2008) meta-analysis of 83 studies reported a median N_e estimate of 280, and show that 70 % of published estimates reported an $N_e < 500$. Our preliminary estimate of N_e in golden eagles falls within this general range, but more accurate estimates will require extensive geographic sampling to encompass more of the genetic variation contained across the Northern Hemisphere.

Our inferences about population structure are also preliminary, and we are cautious in our interpretations because inadequate geographic sampling can bias the delineation of management units (e.g. Beerli 2004; Palsbøll et al. 2007). Nevertheless, our SNP data provide the first snapshot of population genetic structure in North American golden

eagles and provide evidence of population differentiation among disparate geographic regions. Such differentiation is consistent with band encounter data that showed the median natal dispersal distance in golden eagles to be only ~50 km, and 80 % of natal dispersal being within 110 km (Millsap et al. 2014). More extensive sampling will be required to determine if population genetic structure exists on a global scale.

Eagles and other raptors are highly vagile, but geographic barriers can obstruct gene flow (e.g. Sonsthagen et al. 2012). For example, the Rocky and Sierra Nevada mountain ranges separate populations of genetically differentiated red-tailed hawks (*Buteo jamaicensis*, Hull et al. 2008). Distance can also reduce gene flow and differentiate gene pools at continental scales, as in white-tailed sea eagle populations from Greenland/Iceland compared to those of Amur/Japan (Hailer et al. 2007). Our own data indicate that the Sierra Nevada mountain range may act as a filter that reduces gene flow between California and other Western states, while both distance and geographic features likely decrease dispersal between Western and Eastern states. Sampling schemes that are able to more completely assay birds across their range and that can effectively differentiate migrants from non-migrant eagles would be of great consequence to future conservation of this species (Katzner et al. forthcoming).

The genetic distinctiveness of Alaskan eagles relative to eagles in six western states is interesting in light of the telemetry data of McIntyre et al. (2008). Those authors documented the migration of Alaskan golden eagles to Washington, Idaho, Wyoming, Colorado and New Mexico, but not California. Together, the telemetry and genetic data suggest that while Alaskan eagles are obligatory annual migrants, natal philopatry (Millsap et al. 2014) may limit gene flow with the populations we sampled in six western states. Extensive sampling from Canada (and perhaps Russia) will be required to more accurately define the limits of the nominal Alaskan gene pool. That said, golden eagles in Alaska and other parts of the far north combat extreme environmental challenges at the edge of the species' range. The F_{ST} outliers associated with some of our samples may be indicative of different life histories and/or represent local adaptation, as a number of these markers are associated with genes that are under selection in different avian species and potentially adaptive in golden eagles (e.g., *BMP4* and beak formation; Zhan et al. 2013).

Conclusions and conservation implications

Herein, we use a novel SNP assay to provide biological insights which could be relevant to the management of an elusive, iconic, and federally protected species. For example, our survey of SNP variation identified key loci that may be targets of natural selection as well as pronounced genetic structure among sampling sites. Thus, if all known genetic variability and adaptive potential is to be preserved, North American golden eagles may need to be conserved as distinct management units. Our discovery of viability selection may also have important conservation management implications. We observed viability selection in the form of increased average heterozygosity relative to age that may be due to genome-wide effects (such as inbreeding) or to heterozygote superiority at key loci. Although increased juvenile mortality associated with homozygosity may have marginal negative effects when populations are genetically diverse, threatened species often suffer from population-wide decreases in genetic variation (Spielman et al. 2004; Garner et al. 2005; Evans and Sheldon 2008). This could result in proportionally greater juvenile mortality in populations of conservation concern. Although our population genetic observations need further context in the form of additional geographic sampling and genomic analyses, they show significant promise for facilitating the conservation of golden eagles.

Acknowledgments The authors thank A. Capparella, J. Cooper, D. Driscoll, J. Fallon, D. Kramar, M. Kuishn, M. Lanzone, T. Miller, R. Murphy, K. O'Malley, J. Papp, K. Rogers, S. Slater, D. Stafford, D. Stahlecker, S. Thomas, L. Tran, S. Van Arsdae, and D. Wilst for their assistance collecting golden eagle samples. Special thanks to J. Willoughby for assistance generating Fig. 1, M. Sundaram for assistance with statistics, and to DeWoody lab members for comments on earlier drafts of the manuscript. The Nature Conservancy provided permission to use their lands. This work was supported by the U.S. Fish and Wildlife Service, the U.S. Bureau of Land Management (award numbers L12AC20102, L11PX02237, and L12AC2010), the California Department of Fish and Wildlife (Agreement #P1182024), and the Provost's Office at Purdue University (University Faculty Scholar program). Eagle tissue and feather samples were collected under appropriate scientific collecting permits. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government. The findings and conclusions in this article are those of the authors and do not necessarily represent the views of the U.S. Fish and Wildlife Service.

References

Anderson EC, Dunham KK (2008) The influence of family groups on inferences made with the program Structure. Mol Ecol Resour 8:1219–1229. doi:10.1111/j.1755-0998.2008.02355.x

Avise JC (1994) Molecular markers, natural history and evolution. Chapman and Hall, New York

Avise JC (2000) Phylogeography: The history and formation of species. Harvard University Press, Cambridge

Beerli P (2004) Effect of unsampled populations on the estimation of population sizes and migration rates between sampled populations. Mol Ecol 13:827–836. doi:10.1111/j.1365-294X.2004.02101.x

Bekkevold D, Helyar SJ, Limborg MT et al (2015) Gene-associated markers can assign origin in a weakly structured fish, Atlantic herring. ICES J Mar Sci 72:1790–1801

Bloom P, Clark W (2001) Molt and sequence of plumages of golden eagles and a technique for in-hand ageing. North Am Bird Bander 26:97–116

Bourke BP, Frantz AC, Lavers CP et al (2010) Genetic signatures of population change in the British golden eagle (*Aquila chrysaetos*). Conserv Genet 11:1837–1846. doi:10.1007/s10592-010-0076-x

Cadahía L, Pinsker W, Josénegro J et al (2009) Repeated sequence homogenization between the control and pseudo-control regions in the mitochondrial genomes of the subfamily aquilinae. J Exp Zool Part B 312:171–185. doi:10.1002/jez.b.21282

Cai Q, Qian X, Lang Y, Luo Y, Xu J, Pan S, Hui Y, Gou C, Cai Y, Hao M, Zhao J, Wang S, Wang Z, Zhang X, He R, Liu J, Luo L, Li Y, Wang J (2013) Genome sequence of ground tit Pseudopodoces humilis and its adaptation to high altitude. Genome Biol 14:R29

Chakraborty R (1981) The distribution of the number of heterozygous loci in an individual in natural populations. Genetics 98:461–466

Chapman JR, Nakagawa S, Coltman DW et al (2009) A quantitative review of heterozygosity-fitness correlations in animal populations. Mol Ecol 18:2746–2765. doi:10.1111/j.1365-294X.2009.04247.x

Clegg MT, Allard RW (1973) Viability versus fecundity selection in the slender wild oat, *Avena barbata* L. Science 181:667–668

Cohas A, Bonenfant C, Kempenaers B, Allainé D (2009) Age-specific effect of heterozygosity on survival in alpine marmots, *Marmota marmota*. Mol Ecol 18:1491–1503. doi:10.1111/j.1365-294X.2009.04116.x

DePristo MA, Banks E, Poplin R et al (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 43:491–498. doi:10.1038/ng.806

DeWoody YD, DeWoody JA (2005) On the estimation of genome-wide heterozygosity using molecular markers. J Hered 96:85–88. doi:10.1093/jhered/esi017

Do C, Waples RS, Peel D et al (2014) NeEstimator V2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. Mol Ecol Resour 14:209–214

Doran AG, Creevey CJ (2013) Snpdat: easy and rapid annotation of results from de novo snp discovery projects for model and non-model organisms. BMC Bioinformatics 14:45. doi:10.1186/1471-2105-14-45

Downing T, Lloyd AT, O'Farrelly C, Bradley DG (2010) The differential evolutionary dynamics of avian cytokine and TLR gene classes. J Immunol 184:6993–7000. doi:10.4049/jimmunol.0903092

Doyle JM, Katzner TE, Bloom PH et al (2014) The genome sequence of a widespread apex predator, the golden eagle (*Aquila chrysaetos*). PLoS ONE 9:20–22. doi:10.1371/journal.pone.0095599

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4:359–361. doi:10.1007/s12686-011-9548-7

Edwards SV, Gasper J, March M (1998) Genomics and polymorphism of Agph-DAB1, an Mhc class II B gene in red-winged blackbirds (*Agelaius phoeniceus*). Mol Biol Evol 15:236–250

Ekblom R, French L, Slate J, Burke T (2010) Evolutionary analysis and expression profiling of zebra finch immune genes. Genome Biol Evol 2:781–790. doi:10.1093/gbe/evq061

Evans S, Sheldon B (2008) Interspecific patterns of genetic diversity in birds: correlations with extinction risk. Conserv Biol 22:1016–1025

Evans PD, Vallender EJ, Lahn BT (2006) Molecular evolution of the brain size regulator genes CDK5RAP2 and CENPJ. Gene 375:75–79. doi:10.1016/j.gene.2006.02.019

Ferchaud A-L, Pedersen SH, Bekkevold D et al (2014) A low-density SNP array for analyzing differential selection in freshwater and marine populations of threespine stickleback (*Gasterosteus aculeatus*). BMC Genom 15:867. doi:10.1186/1471-2164-15-867

Frankham R (1995) Effective population size/adult population size ratios in wildlife: a review. Genet Res 66:95–107. doi:10.1017/S0016672308009695

Fraser DJ, Bernatchez L (2001) Adaptive evolutionary conservation: towards a unified concept for defining conservation units. Mol Ecol 10:2741–2752

Fridolfsson A-K, Ellegren H (1999) A simple and universal method for molecular sexing of non-ratite birds. J Avian Biol 30:116–121

Funk WC, McKay JK, Hohenlohe PA, Allendorf FW (2012) Harnessing genomics for delineating conservation units. Trends Ecol Evol 27:489–496

Galpern P, Manseau M, Hettinga P et al (2012) Allelematch: an R package for identifying unique multilocus genotypes where genotyping error and missing data may be present. Mol Ecol Resour 12:771–778. doi:10.1111/j.1755-0998.2012.03137.x

Garner A, Rachlow J, Hicks J (2005) Patterns of genetic diversity and its loss in mammalian populations. Conserv Biol 19:1215–1221

Grueber CE, Walls CP, Jamieson IG (2013) Genetic drift outweighs natural selection at toll-like receptor (TLR) immunity loci in a re-introduced population of a threatened species. Mol Ecol 22:4470–4482

Grueber CE, Knafler GJ, King TM, Senior AM, Grosser S, Robertson B, Weston KA, Brekke P, Harris CL, Jamieson IG (2015) Toll-like receptor diversity in 10 threatened bird species: relationship with microsatellite heterozygosity. Conserv Genet 16:595–611

Haasl R, Payseur BA (2016) Fifteen years of genomewide scans for selection: trends, lessons and unaddressed genetic sources of complication. Mol Ecol 25:5–23

Hailer F, Helander B, Folkestad AO et al (2007) Phylogeography of the white-tailed eagle, a generalist with large dispersal capacity. J Biogeogr 34:1193–1206. doi:10.1111/j.1365-2699.2007.01697.x

Hoffman JI, Simpson F, David P et al (2014) High-throughput sequencing reveals inbreeding depression in a natural population. Proc Natl Acad Sci USA 111:3775–3780. doi:10.1073/pnas.1318945111

Hull JM, Hull AC, Sacks BN et al (2008) Landscape characteristics influence morphological and genetic differentiation in a widespread raptor (*Buteo jamaicensis*). Mol Ecol 17:810–824. doi:10.1111/j.1365-294X.2007.03632.x

Jollie M (1947) Plumage changes in the Golden Eagle. Auk 64:549–576

Katzner T, Smith BW, Miller TA et al (2012) Status, biology and conservation priorities for North America's Eastern Golden Eagle (*Aquila chrysaetos*) population. Auk 129:168–176

Katzner TE, Nelson DM, Braham MA, et al (forthcoming) Golden eagle fatalities demonstrate the continental-scale consequences of local-scale renewable energy conservation

Keenan K, McGinnity P, Cross T et al (2013) DiveRsity: an R package for the estimation of population genetics parameters and their associated errors. Methods Ecol Evol 4:782–788

Keller L, Waller D (2002) Inbreeding effects in wild populations. Trends Ecol Evol 16:1099–1106

Kosiol C, Vinař T, da Fonseca RR et al (2008) Patterns of positive selection in six mammalian genomes. PLoS Genet 4(8):e1000144

Lampila S, Orell M, Kvist L (2011) Willow tit *Parus montanus* extrapair offspring are more heterozygous than their maternal half-siblings. J Avian Biol 42:355–362. doi:10.1111/j.1600-048X.2011.05349.x

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760. doi:10.1093/bioinformatics/btp324

Limborg MT, Helyar SJ, De Bruyn M et al (2012) Environmental selection on transcriptome-derived SNPs in a high gene flow marine fish, the Atlantic herring (*Clupea harengus*). Mol Ecol 21:2355–2356

Malenfant RM, Coltman DW, Davis CS (2015) Design of a 9 K illumina BeadChip for polar bears (*Ursus maritimus*) from RAD and transcriptome sequencing. Mol Ecol Resour 15:587–600. doi:10.1111/1755-0998.12327

McIntyre CL, Collopy MW, Lindberg MS (2006) Survival probability and mortality of migratory juvenile golden eagles from interior Alaska. J Wildl Manage 70:717–722

McIntyre CL, Douglas DC, Collopy MW (2008) Movements of golden eagles (*Aquila chrysaetos*) from interior Alaska during their first year of independence. Auk 125:214–224. doi:10.1525/auk.2008.125.1.214

Millsap G, Zimmerman G, Sauer J et al (2013) Golden eagle population trends in the western United States: 1968-2010. J Wildl Manage 77:1436–1448

Millsap B, Harmata A, Stahlecker D, Mikesic D (2014) Natal dispersal distance of Bald and Golden eagles originating in the Coterminous United States as inferred from band encounters. J Raptor Res 48:2014

Mitton JB (1997) Selection in natural populations. Oxford University Press, New York

Mitton JB, Pierce BA (1980) The distribution of individual heterozygosity in natural populations. Genetics 95:1043–1054

Moritz C (1994) Defining 'Evolutionary Significant Units' for conservation. Trends Ecol Evol 9:373–375

Morneau F, Tremblay JA, Todd C et al (2015) Known breeding distribution and abundance of golden eagles in Eastern North America. Northeastern Naturalist 22:236–247

Mueller JC, Korsten P, Hermannstaedter C et al (2013) Haplotype structure, adaptive history and associations with exploratory behaviour of the DRD4 gene region in four great tit (*Parus major*) populations. Mol Ecol 22:2797–2809. doi:10.1111/mec.12282

Nam K, Mugal C, Nabholz B et al (2010) Molecular evolution of genes in avian genomes. Genome Biol 11:R68. doi:10.1186/gb-2010-11-6-r68

Nebel C, Gamauf A, Haring E et al (2015) Mitochondrial DNA analysis reveals Holarctic homogeneity and a distinct Mediterranean lineage in the Golden eagle (*Aquila chrysaetos*). Biol J Linn Soc 116:328–340

Neph S, Kuehn MS, Reynolds AP et al (2012) BEDOPS: high-performance genomic feature operations. Bioinformatics 28:1919–1920. doi:10.1093/bioinformatics/bts277

Nery MF, Gonzalez DJ, Opazo JC (2013) How to make a dolphin: molecular signature of positive selection in Cetacean Genome. PLoS One 8(6):e65491

Ogden R, Heap E, Mcewing R et al (2015) Population structure and dispersal patterns in Scottish Golden Eagles *Aquila chrysaetos* revealed by molecular genetic analysis of territorial birds. Ibis (Lond 1859). doi:10.1111/ibi.12282

Ovenden JR, Morgan JAT, Street R et al (2011) Negligible evidence for regional genetic population structure for two shark species *Rhizoprionodon acutus* (Ruppell, 1837) and *Sphyrna lewini* (Griffith & Smith, 1834) with contrasting biology. Mar Biol 7:1497–1509

Palsbøll PJ, Bérubé M, Allendorf FW (2007) Identification of management units using population genetic data. Trends Ecol Evol 22:11–16. doi:10.1016/j.tree.2006.09.003

Palstra F, Ruzzante D (2008) Genetic estimates of contemporary effective population size: what can they tell us about the importance of genetic stochasticity for wild population persistence? Mol Ecol 17:3428–3447. doi:10.1111/j.1365-294X.2008.03842.x

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research–an update. Bioinformatics 28:2537–2539. doi:10.1093/bioinformatics/bts460

Pritchard JK (2010) Documentation for structure software: Version 2.3. http://pritchardlab.stanford.edu/structure_software/release_versions/v2.3.4/structure_doc.pdf. Accessed 27 May 2015

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Qu Y, Zhao H, Han N et al (2013) Ground tit genome reveals avian adaptation to living at high altitudes in the Tibetan plateau. Nat Commun 4:2071. doi:10.1038/ncomms3071

Rands CM, Darling A, Fujita M et al (2013) Insights into the evolution of Darwin's finches from comparative analysis of the *Geospiza magnirostris* genome sequence. BMC Genom 14:95. doi:10.1186/1471-2164-14-95

Robinson JT, Thorvaldsdóttir H, Winckler W et al (2011) Integrative genomics viewer. Nat Biotechnol 29:24–26

Rubin C-J, Zody MC, Eriksson J et al (2010) Whole-genome resequencing reveals loci under selection during chicken domestication. Nature 464:587–591. doi:10.1038/nature08832

Rudnick JA, Katzner TE, Bragin EA et al (2005) Using naturally shed feathers for individual identification, genetic parentage analyses, and population monitoring in an endangered Eastern imperial eagle (*Aquila heliaca*) population from Kazakhstan. Mol Ecol 14:2959–2967. doi:10.1111/j.1365-294X.2005.02641.x

Ruegg KC, Anderson EC, Paxton KL et al (2014) Mapping migration in a songbird using high-resolution genetic markers. Mol Ecol 23:5726–5739. doi:10.1111/mec.12977

Ryder OA (1986) Species conservation and systematics: the dilemma of subspecies. Trends Ecol Evol 1:9–10

Schunter C, Garza JC, Macpherson E, Pascual M (2014) SNP development from RNA-seq data in a nonmodel fish: how many individuals are needed for accurate allele frequency prediction? Mol Ecol Resour 14:157–165. doi:10.1111/1755-0998.12155

Senn H, Ogden R, Cezard T et al (2013) Reference-free SNP discovery for the Eurasian beaver from restriction site-associated DNA paired-end data. Mol Ecol 22:3141–3150. doi:10.1111/mec.12242

Shafer ABA, Wolf JBW, Alves PC et al (2015) Genomics and the challenging translation into conservation practice. Trends Ecol Evol 30:78–87

Shaffer HB, Minx P, Warren DE et al (2013) The western painted turtle genome, a model for the evolution of extreme physiological adaptations in a slowly evolving lineage. Genome Biol 14:R28. doi:10.1186/gb-2013-14-3-r28

Sonsthagen SA, Coonan TJ, Latta BC et al (2012) Genetic diversity of a newly established population of golden eagles on the channel islands, California. Biol Conserv 146:116–122. doi:10.1016/j.biocon.2011.11.031

Spielman D, Brook B, Frankham R (2004) Most species are not driven to extinction before genetic factors impact them. Proc Natl Acad Sci 101:15261–15264

Swanson W, Yang Z, Wolfner M, Aquadro C (2001) Positive Darwinian selection in the evolution of mammalian female reproductive proteins. Proc Natl Acad Sci USA 98:2509–2514

Szulkin M, Bierne N, David P (2010) Heterozygosity-fitness correlations: a time for reappraisal. Evolution 64:1202–1217. doi:10.1111/j.1558-5646.2010.00966.x

Torgerson D, Kulanthinal R, Singh R (2002) Mammalian sperm proteins are rapidly evolving: evidence of positive selection in functionally diverse genes. Mol Biol Evol 19:1973–1980

Turelli M, Ginzburg LR (1983) Should individual fitness increase with heterozygosity? Genetics 104:191–209

Van der Auwera GA, Carneiro MO, Hartl C et al (2013) From fastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. Curr Protoc Bioinform 43:11.10.1–11.10.33

Voight BF, Kudaravalli S, Wen X et al (2006) A map of recent positive selection in the human genome. PLoS Biol 4(3):e72

Wan Q-H, Pan S-K, Hu L et al (2013) Genome analysis and signature discovery for diving and sensory properties of the endangered Chinese alligator. Cell Res 23:1091–1105. doi:10.1038/cr.2013.104

Warren WC, Clayton DF, Ellegren H et al (2010) The genome of a songbird. Nature 464:757–762. doi:10.1038/nature08819

Watson J (2010) The golden eagle. Yale University Press, New Haven

Wheeler M (2014) The genetics of conservation translocations: a comparison of North American Golden eagles (*Aquila chrysaetos canadensis*) and Bald eagles (*Haliaeetus leucocephalus*),pp 1–220

Zhan X, Pan S, Wang J et al (2013) Peregrine and saker falcon genome sequences provide insights into evolution of a predatory lifestyle. Nat Genet 45:536–566. doi:10.1038/ng.2588

Supplementary Material for Genetic structure and viability selection in the Golden Eagle (*Aquila chrysaetos*), a vagile raptor with a holarctic distribution[a]

Jacqueline M. Doyle[*1], Todd E. Katzner[2], Gary W. Roemer[3], James W. Cain III[4], Brian Millsap[5], Carol McIntyre[6], Sarah A. Sonsthagen[7], Nadia Fernandez[1], Maria Wheeler[8], Zafer Bulut[9], Peter Bloom[10], and J. Andrew DeWoody[1,11]

[1] Department of Forestry and Natural Resources, Purdue University, 715 W. State Street, West Lafayette, IN 47907

[2] Forest and Rangeland Ecosystem Science Center, U.S. Geological Survey, 970 Lusk Street, Boise, ID 83706, USA

[3] Department of Fish, Wildlife and Conservation Ecology, New Mexico State University, Las Cruces, NM 88003, USA

[4] U.S. Geological Survey, New Mexico Cooperative Fish and Wildlife Research Unit, Department of Fish, Wildlife and Conservation Ecology, New Mexico State University, Las Cruces, NM 88003, USA

[5] Division of Migratory Bird Management, U.S. Fish and Wildlife Service, 2105 Osuna NE, Albuquerque, NM 87113, USA

[6] National Park Service, 4175 Geist Road, Fairbanks, AK 99709, USA

[7] USGS Alaska Science Center, 4210 University Dr., Anchorage, AK 99508, USA

[8] Department of Biological Sciences, Duquesne University, 600 Forbes Ave., Pittsburgh, PA 15282

[9] Department of Biochemistry, Selcuk University, Konya 42031, Turkey

[10] Bloom Biological Inc., 13611 Hewes Ave., Santa Ana, CA 92705, USA

[11] Department of Biological Sciences, Purdue University, 915 W. State Street, West Lafayette, IN 47907, USA

**Correspondent:* Jacqueline M. Doyle, Address: 715 W. State St., West Lafayette, IN 47907

Phone: 270-293-9486, Email: jmdoyle@purdue.edu

Repeatability

Three to seven replicate DNA samples (n = 78 in total) from 15 individual golden eagles were used to calculate error rates. Using all SNP calls generated at a site across replicate samples, the principle of plurality was used to determine a consensus sequence for each individual. The SNP typing error rate was calculated according to the equation $\frac{m}{d(s)}$, where *m* represents the total number of mismatches between each replicate sample and the consensus sequence across multiple samples from the same individual; *d* represents the total number of loci per replicate sample; and *s* represents the total number of replicate samples. Both incorrect SNP calls (errors of commission) and instances where no amplification occurred (errors of omission) contributed to *m*. Seventy-eight independent replicate samples from 15 golden eagles (Online Resource 3) were used to calculate an error rate (*e*) of 0.4%. Low-quality DNA samples generally had more incorrect SNP calls than higher-quality DNA samples (SI Figure 1).

Molecular sexing and mtDNA haplotyping

For validation purposes, a subset of samples was sexed with the traditional PCR/gel method (Fridolfsson and Ellegren 1999) and also by using our novel CHD1 marker. Similarly, mtDNA haplotypes were generated using both traditional dideoxy methods and novel SNP markers. We compared the concordance among assays in an attempt to validate our novel markers.

For validation purposes, 109 individual golden eagles were sexed using both our novel SNP assay and the traditional method using PCR and gel electrophoresis (Fridolfsson and Ellegren 1999). The results of the two methods (50 males and 59 females) were in complete concordance with one another (SI Figure 2). Fifty-four golden eagles were haplotyped using both our mtDNA SNP assays as well as traditional Sanger sequencing. Once again, the results of the two methods were in perfect concordance.

Over the last fifteen years, ornithologists have benefitted greatly from the development of highly conserved primers flanking CHD1 introns (Fridolfsson and Ellegren 1999). We developed a SNP assay from the CHD1 gene that reliably sexes golden, bald, imperial, and white-tailed sea eagles. The relevant region of the CHD1 gene is highly conserved among raptors (e.g., owls, vultures and hawks), which suggests that our SNP marker could be even more widely applied for raptor sexing.

Our SNP assay also includes mtDNA markers that can be used to assign haplotypes directly to individuals, or indirectly to sampling sites. The incorporation of mtDNA and nuclear markers into a single assay has a number of potential benefits for conservation geneticists. Mitochondrial and nuclear markers have different effective population sizes and confer different signals of genetic differentiation (Fahey et al. 2014). For example, our use of mitochondrial markers (and the calculation of Φ_{PT}) indicated significant differentiation between California and the Eastern and Western states, whereas nuclear markers (and the subsequent calculation of F_{ST}) indicated significant differentiation between Alaska and the Eastern and Western states. Thus a combination of the two marker types can be useful for resolving differences between regional gene pools as well as subspecies distinctions (Zink et al. 2013). This is of particular interest to management of golden eagles, given that 6 subspecies of golden eagles are thought to occupy its Holarctic range (Watson 2010) and that mtDNA has been used effectively to help delineate regional gene pools (Sonsthagen et al. 2012; Nebel et al. 2015).

Linkage disequilibrium

We deliberately minimized linkage disequilibrium (LD) by preferentially choosing single SNPs from a given scaffold or, when considering multiple SNPs from a single scaffold, choosing markers separated by intervals >300 kb. In a few cases, we included markers at physical distances <100 kb. To test for a relationship between physical distance and LD, we used SnpStats (Clayton 2014) to calculate r^2 (a measurement of pairwise LD, hereafter referred to as Δ^2), log-transformed the dependent variable to meet the assumption of normality and used linear regression to test for a relationship between physical distance and Δ^2.

Our 162 SNPs were distributed across 130 scaffolds. There were 35 instances where two markers were present on the same scaffold, allowing for measurement of physical distance between markers. Loci with very high (0.98) and very low (0.02) Δ^2 values were present when physical distances were less than 60 kb (SI Figure 3). There was a negative relationship between physical distance and log-transformed Δ^2 ($r^2 = 0.26$, $p = 0.002$). When all loci (including those that were the only markers present on a particular scaffold) were tested for pairwise LD we found only 8 cases (out of 13,041 pairwise comparisons) where $\Delta^2 \geq 0.2$, indicating moderate linkage disequilibrium between a few markers (Online Resource 4).

Linkage disequilibrium (nonindependence) and heterozygote deficiency (potential null alleles) are present in our marker panel but occurred at only a small proportion of the loci (8% and 11%, respectively). Given the overwhelming statistical power associated with our SNP panel, future workers may rely on a subset of our 162 SNPs (e.g., all neutral or all gene-associated markers). The 96 SNPs recommended in Online Resource 8 are among the most polymorphic in our samples and cluster cleanly (i.e., require little manual editing).

Comparison of microsatellite- and SNP-based genetic variation

A subset of our samples underwent microsatellite genotyping in addition to SNP genotyping in order to compare genetic variation associated with the microsatellite and SNP assays. Individuals were genotyped at 9 microsatellite loci consisting of Hal10, Hal13 and IEAAAG13 (triplex), Aa11, Aa15, Aa36, IEAAAG04 (tetraplex), IEAAAG14 and Aa27 (Martinez-Cruz et al. 2002; Busch et al. 2005; Hailer et al. 2005; Katzner et al. 2015). Seventy-one golden eagle samples were genotyped with both microsatellite and SNP markers. These individuals were heterozygous at 0-8 of the 9 microsatellite markers ($H_O = 0.51 \pm 0.06$, $H_E = 0.51 \pm 0.07$) and 36-72 of the 159 SNP markers. There was no significant relationship between the number of heterozygous microsatellite loci and heterozygous SNP loci per individual (linear regression: $r^2 = 0.01$, $p < 0.05$, SI Figure 4). Microsatellite H_o and H_e were calculated with GenAlEx 6.501 (Peakall and Smouse 2012).

Population structure

STRUCTURE analysis using only known providence individuals is described in detail in the main text of the manuscript. We used STRUCTURE and both known and unknown providence individuals (n = 318 golden eagles total) in an additional test of the panmixia hypothesis. The initial STRUCTURE analysis of panmixia rejected the hypothesis and provides evidence for genetic structure between Californian individuals and all others (SI Figure 5). We removed Californian samples from the data set and conducted an additional STRUCTURE analysis with the remaining samples. Without samples from California, all birds from the Western states clustered separately from those sampled in the Eastern states (Figure 4b).

Given the results of our Bayesian analysis and the amount of genetic admixture present in Western states, we calculated pairwise F_{ST} values for four presumptive populations: Alaska, California, Eastern states and Western states (Table 3). Significant population differentiation was evident in Alaska vs. Eastern, Alaska vs. Western and Eastern vs. Western pairwise comparisons (i.e. 95% confidence intervals did not overlap zero, SI Table 1). There were no significant differences between F_{ST} values calculated with gene-associated and putatively neutral markers (SI Table 1).

Paternity analyses

Some of our samples from California, where we collected breeding eagle feathers noninvasively from nesting territories and blood directly from juveniles, were amenable to parentage analyses. Samples from putative family groups were assigned to individual eagles using allelematch (Galpern et al. 2012). Parentage assignment was conducted on the California samples using Cervus 3.0.7 (Kalinowski et al. 2007); all golden eagles genotyped from California were considered candidate parents. An adult male or female with a critical delta-value (95% confidence) greater than 2.0 and 2.1, respectively, was considered the most probable biological parent. GenAlEx 6.501 was used to calculate probability of identity (P_{ID}), probability of identity when related individuals are present (PI_{SIB}), and the probability of exclusion (PE) for the entire data set.

Our suite of novel SNPs provide strong discriminatory power in that $P_{ID,}$ PI_{SIB} and PE were estimated as 3.7×10^{-48}, 5.8×10^{-25}, and 0.999, respectively. Overall, we genotyped 23 chicks from 14 different nests and used these data to confirm the Mendelian inheritance of all 159 nuclear autosomal SNPs, the maternal inheritance of both mtDNA SNPs, and the sex-linked (ZW) inheritance of the sexing marker. We identified chicks and both parents at 7 such nests in California (see SI Table 2, Online Resource 5). In an additional 7 cases, we identified chicks as well as either a male or female parent (Online Resource 5). We conducted genetic parentage analyses for 23 chicks using all non-chick golden eagles genotyped in California included as candidate parents. In the case of each chick, the male and female adults genotyped at the nest were always identified as the most probabilistic parent(s).

Although behavioral observations suggest golden eagles are primarily monogamous (Cramp and Simmons 1980), no test of the genetic mating system has been performed. In fact, Watson (2010) reviewed several reports of trios of golden eagles incubating, brooding and feeding chicks at a single nest (e.g. Bergo 1988), and argued that these might be episodes of polygamy or polygyny. A plausible alternative hypothesis is that one member of each trio is the juvenile offspring of the mated pair, helping to care for its siblings (e.g. European bee-eaters, Jones et al. 1991). Our own data indicate that each of 23 chicks considered could be assigned to the adult male and/or female present at their nest. Although no additional adults were genetically tagged at any of the nests, future sampling may reveal whether "helpers" assist parents and care for siblings.

Cross-species utility

Bald, imperial and white-tailed eagle SNP calls (both nuclear and mtDNA) were visualized and edited as in golden eagles. We quantified the number of polymorphic and monomorphic SNPs for each species and compared interspecific results from the traditional generic sexing assay (Fridolfsson and Ellegren 1999) to our novel eagle SNP sexing assay. We typed 42 samples from bald, imperial and white-tailed eagles but the vast majority of our loci were monomorphic (SI Table 3, Online Resource 2). Three to eight markers amplified and were polymorphic in each of these species. Approximately 15

155 markers exhibit fixed differences between our samples of imperial (genus *Aquila*) and bald/white-tailed
156 eagles (genus *Haliaeetus*). In all three species, our novel SNP sexing assay worked well and
157 corresponded perfectly to the traditional assay (Fridolfsson and Ellegren 1999).
158 When markers designed for one species are applied to non-target species, polymorphism
159 decreases as divergence time between the source and target species increases (Primmer et al. 1996;
160 VonHoldt et al. 2011; Miller et al. 2012). Not surprisingly, most of our SNP loci designed for golden
161 eagles were monomorphic in samples of bald, imperial and white-tailed sea eagles. However, there were
162 fixed differences between *A. heliaca* and the two *Haliaeetus* species at 16 loci. These markers may
163 ultimately be useful for species identification when sampling feathers from mixed-roost assemblages
164 (e.g., Rudnick et al. 2007).
165

Literature cited

Bergo G (1988) Trios is the Golden Eagle *Aquila chrysaetos* (L.). Fauna Nor Ser C, Cindus 11:40–44.

Clayton D (2014) snpStats: SnpMatrix and XSnpMatrix classes and methods.

Cramp S, Simmons K (1980) The birds of the Palearctic, 2nd edn. University Press, Oxford

Fahey AL, Ricklefs RE, Dewoody JA (2014) DNA-based approaches for evaluating historical demography in terrestrial vertebrates. Biol J Linn Soc 112:367–386. doi: 10.1111/bij.12259

Fridolfsson A-K, Ellegren H (1999) A simple and universal method for molecular sexing of non-ratite birds. J Avian Biol 30:116–121.

Galpern P, Manseau M, Hettinga P, et al (2012) Allelematch: An R package for identifying unique multilocus genotypes where genotyping error and missing data may be present. Mol Ecol Resour 12:771–778. doi: 10.1111/j.1755-0998.2012.03137.x

Jones C, Lessells C, Krebs J (1991) Helpers-at-the-nest in European Bee-eaters (*Merops apiaster*): a genetic analysis. In: DNA fingerprinting: approaches and applications. pp 169–192

Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Mol Ecol 16:1099–1106. doi: 10.1111/j.1365-294X.2007.03089.x

Miller JM, Kijas JW, Heaton MP, et al (2012) Consistent divergence times and allele sharing measured from cross-species application of SNP chips developed for three domestic species. Mol Ecol Resour 12:1145–1150. doi: 10.1111/1755-0998.12017

Nebel C, Gamauf A, Haring E, et al (2015) Mitochondrial DNA analysis reveals Holarctic homogeneity and a distinct Mediterranean lineage in the Golden eagle (*Aquila chrysaetos*). Biol J Linn Soc 116:328-340

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research--an update. Bioinformatics 28:2537–9. doi: 10.1093/bioinformatics/bts460

Primmer CR, Moller AP, Ellegren H (1996) A wide-range survey of cross-species microsatellite amplification in birds. Mol Ecol 5:365–378.

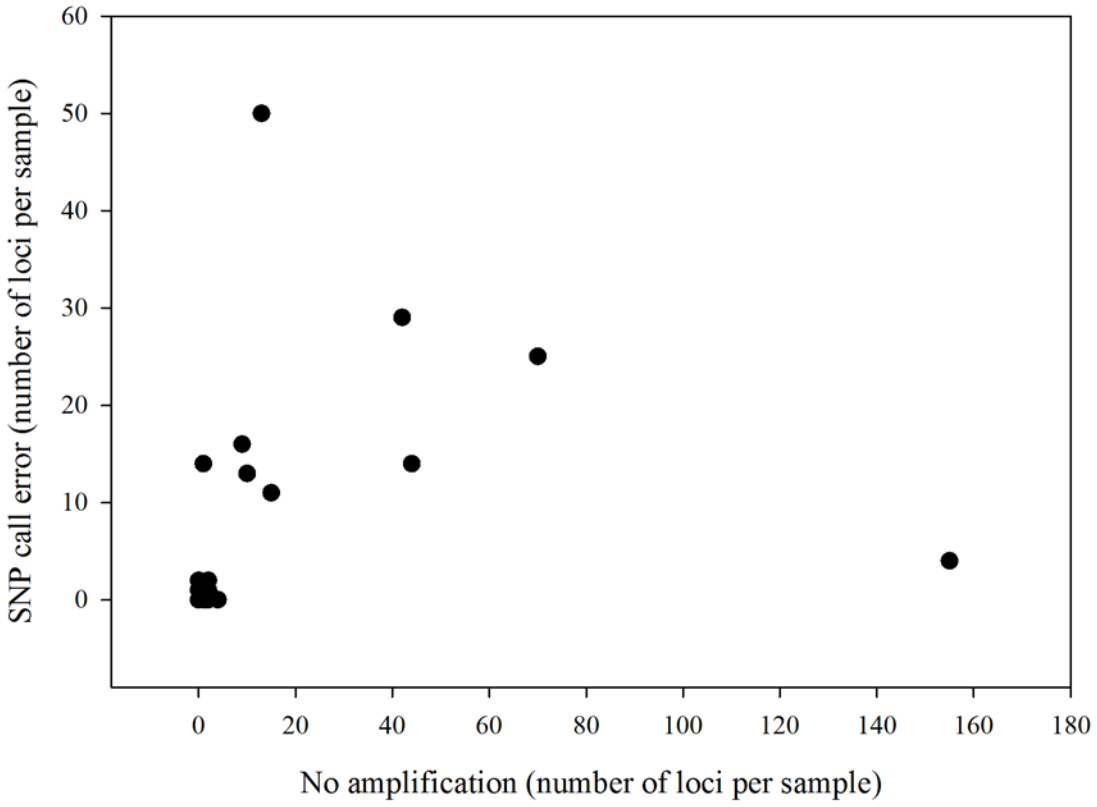
Rudnick JA, Katzner TE, Bragin EA, DeWoody JA (2007) Species identification of birds through genetic analysis of naturally shed feathers. Mol Ecol Notes 7:757–762. doi: 10.1111/j.1471-8286.2007.01796.x

Sonsthagen SA, Coonan TJ, Latta BC, et al (2012) Genetic diversity of a newly established population of golden eagles on the Channel Islands, California. Biol Conserv 146:116–122. doi: 10.1016/j.biocon.2011.11.031

VonHoldt BM, Pollinger JP, Earl DA, et al (2011) A genome-wide perspective on the evolutionary history of enigmatic wolf-like canids. Genome Res 21:1294–1305. doi: 10.1101/gr.116301.110

Watson J (2010) The Golden Eagle. Yale University Press, New Haven

Zink RM, Groth JG, Vázquez-Miranda H, Barrowclough GF (2013) Phylogeography of the California Gnatcatcher (*Polioptila californica*) using multilocus DNA sequences and ecological niche modeling. Auk 130:449–458. doi: 10.1525/auk.2013.12241



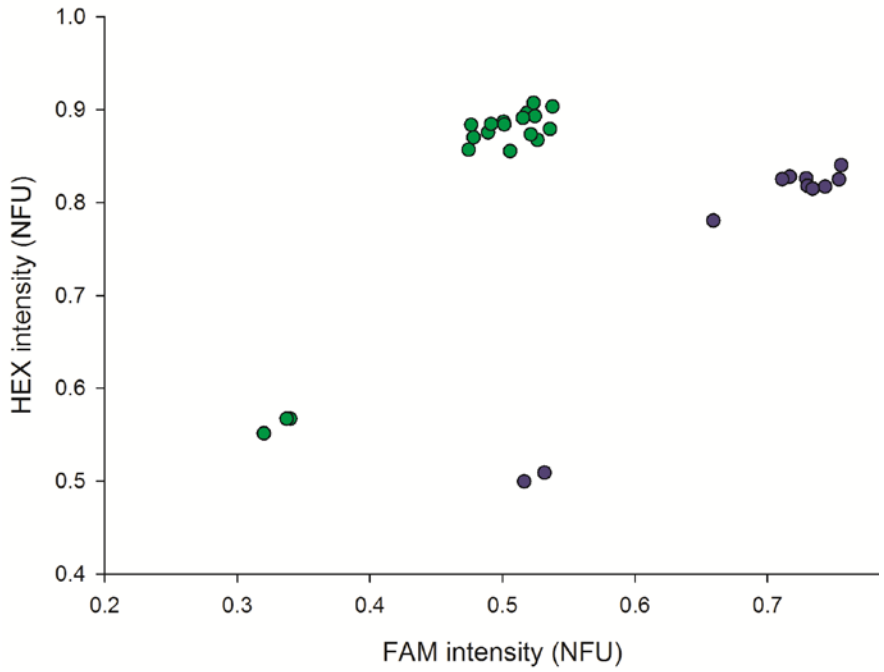
206

207 SI Figure 1: Nine low-quality DNA samples (i.e. those with a large number of loci that did not amplify)

208 generally had more incorrect SNP calls than higher-quality DNA samples. Seventy-nine high quality

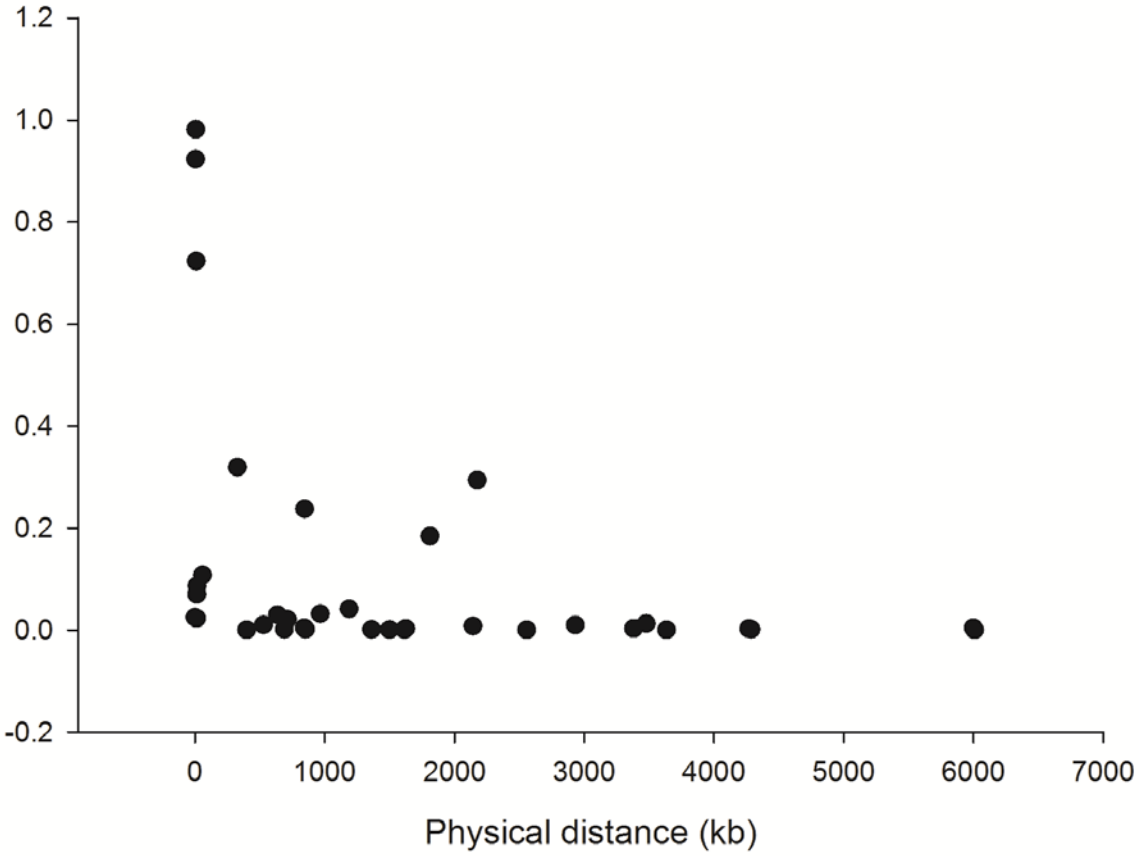
209 samples had fewer than five loci that did not amplify or were in error.

210

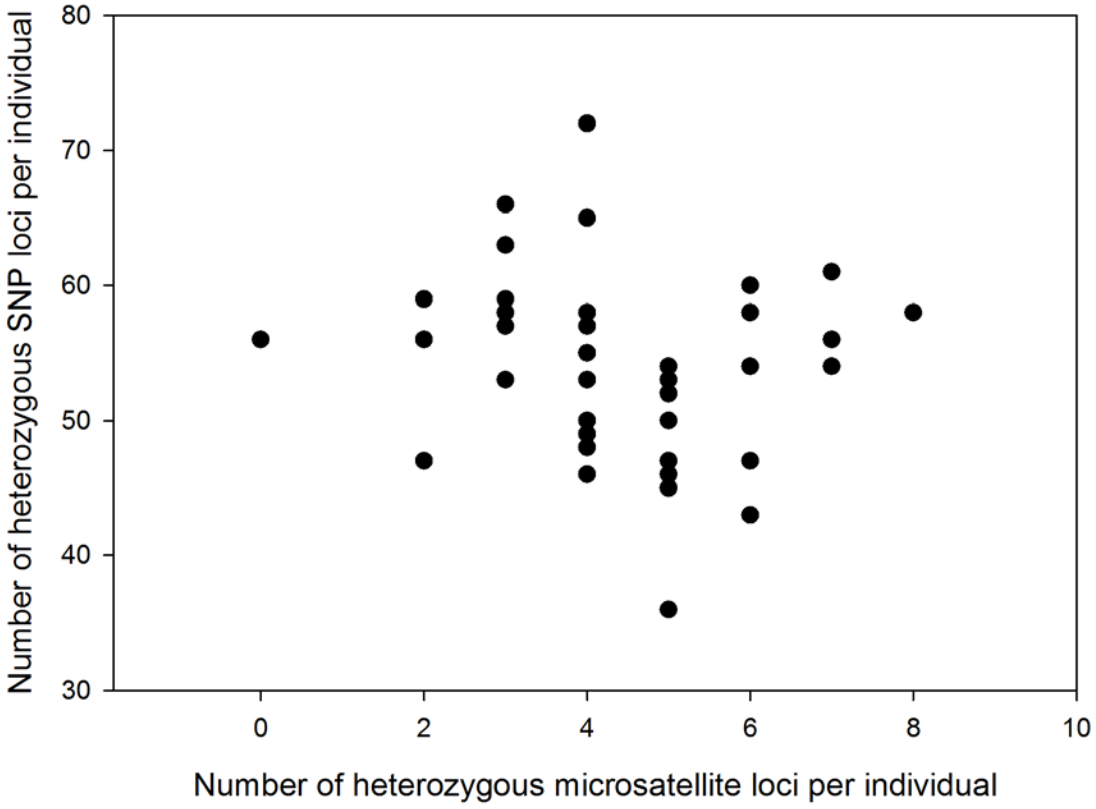


211

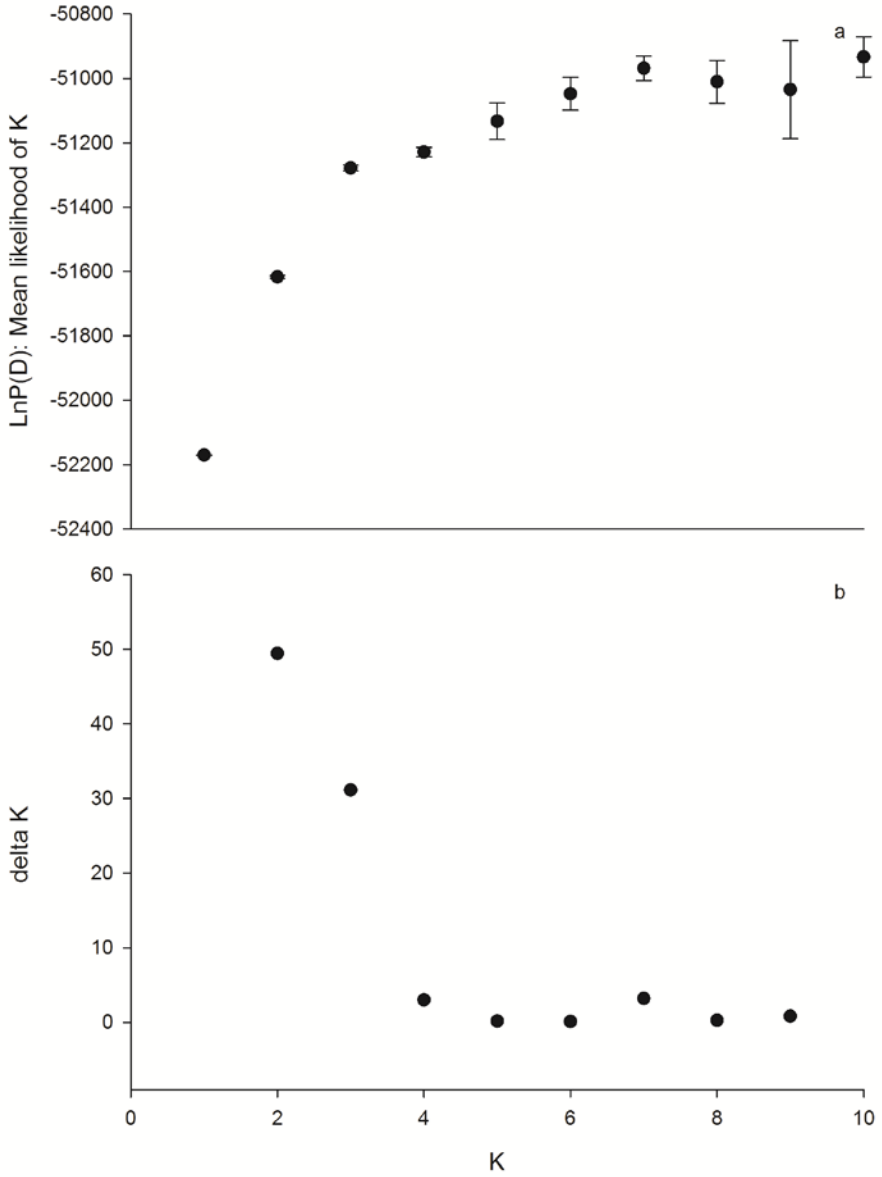
212 SI Figure 2: Males (green) and females (blue) can be distinguished from one another based on HEX and
213 FAM intensity measured in normalized fluorescence units (NFU). Males and females are distinguished
214 using the Fluidigm SNP genotyping analysis software clustering algorithm.



SI Figure 3: Linkage disequilibrium (Δ^2) decreases as physical distance increases.



SI Figure 4: There is no significant relationship between the number of heterozygous microsatellite loci and heterozygous SNP loci per individual (linear regression: $r^2 = 0.02$, $p < 0.05$).



SI Figure 5: Results of a STRUCTURE analysis of a mix of 318 known and unknown provenance individuals sampled in Alaska, California, Colorado, Arizona/New Mexico, Utah, Nebraska/Wyoming and Eastern states genotyped at 159 SNP loci. Panel a shows mean likelihood values (±SD) for 10 STRUCTURE runs (each with 10 iterations). Panel b shows delta K values 2 through 9. Viewed together, these results indicate that K = 2.

SI Table 1: Mean F_{ST} values for gene-associated and putatively neutral markers for each pairwise comparison. We tested for differences between F_{ST} values calculated with gene-associated and neutral markers for each pairwise comparison using a Mann U Whitney rank sum test.

Pairwise comparison	Global F_{ST}	95% CI	Marker type	Mean F_{ST} ± SE	Mann U Whitney
Alaska/California	0.038	-0.708, 0.153	Gene-associated	0.06 ± 0.01	U = 747, p = 0.50
			Neutral	0.06 ± 0.01	
Alaska/Eastern	0.052	0.041, 0.066	Gene-associated	0.05 ± 0.01	U = 613, p = 0.05
			Neutral	0.08 ± 0.01	
Alaska/Western	0.027	0.018, 0.039	Gene-associated	0.03 ± 0.01	U = 742, p = 0.46
			Neutral	0.05 ± 0.01	
California/Eastern	0.062	-0.659, 0.174	Gene-associated	0.01 ± 0.01	U = 714, p = 0.27
			Neutral	0.01 ± 0.01	
California/Western	0.019	-0.579, 0.113	Gene-associated	0.03 ± 0.01	U = 805, p = 0.90
			Neutral	0.03 ± 0.01	
Eastern/Western	0.055	0.047, 0.065	Gene-associated	0.02 ± 0.01	U = 727, p = 0.37
			Neutral	0.02 ± 0.01	

SI Table 2: An exemplar family from the Buette Valley Steepe Island nest identified by genotyping feather and blood samples. The resident male and female were identified from 7 and 2 feathers, respectively. 2258884_9766, 2280973_6559 and 2283459_581935 are three representative SNP loci of the 159 typed for each sample.

Sex	Family member	Sample type	2258884_9766a	2258884_9766b	2280973_6559a	2280973_6559b	2283459_581935a	2283459_581935b
male	adult male	feather	C	T	A	C	T	T
male	adult male	feather	C	T	A	C	T	T
male	adult male	feather	C	T	A	C	T	T
male	adult male	feather	C	T	A	C	T	T
male	adult male	feather	C	T	A	C	T	T
male	adult male	feather	C	T	A	C	T	T
male	adult male	feather	C	T	A	C	T	T
female	adult female	feather	T	T	A	C	C	C
female	adult female	feather	T	T	A	C	C	C
female	chick	blood	T	T	C	C	C	T
male	chick	blood	C	T	C	C	C	T

SI Table 3: Number of SNP markers (designed for *A. chrysaetos*) that were invalid (because they did not amplify or clustered poorly), monomorphic or polymorphic in *A. helica*, *H. leucocephalus*, and *H. albicilla*. Also shown are the numbers of markers with differentially fixed alleles in *A. helica* vs. *H. leucocephalus*, *A. helica* vs. *H. albicilla* and *H. leucocephalus* vs. *H. albicilla*.

					Alleles differentially fixed between species		
Species	Sample size	Invalid	Monomorphic	Polymorphic	x *A. helica*	x *H. leucocephalus*	x *H. albicilla*
A. helica	14	3	157	3	-	16	15
H. leucocephalus	9	12	147	4	16	-	0
H. albicilla	14	12	143	8	15	0	-