# CONSERVATION GENOMICS OF NORTH AMERICAN BALD (*HALIAEETUS LEUCOCEPHALUS*) AND GOLDEN (*AQUILA CHRYSAETOS*) EAGLES

By

MEGAN E. JUDKINS

Bachelor of Arts/Science in Natural Resources Ecology
and Management
Oklahoma State University
Stillwater, OK
2005

Master of Business Administration
Marylhurst University,
Marylhurst, Oregon 2012.

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
December, 2017

CONSERVATION GENOMICS OF NORTH AMERICAN BALD (*HALIAEETUS LEUCOCEPHALUS*) AND GOLDEN (*AQUILA CHRYSAETOS*) EAGLES

Dissertation Approved:

Dr. Ronald A. Van Den Bussche

Dissertation Adviser

Dr. Jim Lish

Dr. Meredith Hamilton

Dr. Andrew Doust

Outside Committee Member

ACKNOWLEDGEMENTS

There are a number of people that I must thank who made my dissertation possible and who supported me throughout this journey. To Victor Roubidoux who has supported all of my ideas in an effort to make the Grey Snow Eagle House the best it can be, thank you for allowing the bald and golden eagle conservation partnership between the Iowa Tribe of Oklahoma and the Van Den Bussche Lab to happen. Without your support, this project would have never been able to transform into the success that it is today. I also must express my deepest thanks and appreciation to Ron Van Den Bussche for taking a chance on this initially unfunded project that was fueled by the passion of three people who understand the importance of the conservation of eagles. I must also thank Ron for all the time, effort, and knowledge shared throughout the last six years that built me into the researcher and person I am today. To both Ron and Victor, I greatly appreciate that both of you trusted that I could handle both my management position at the aviary while I pursued this degree, so I did not have to choose between the two opportunities.

I must also thank the many individuals that provided me with knowledge and a willingness to collaborate. My committee members Dr. Andrew Doust, Dr. Meredith Hamilton, Dr. Jim Lish have provided a wealth of knowledge gained through courses taken, teaching assistantship opportunities, and conversations that cannot be replaced. Your kind words and support, even with just a hello in the hall, helped me continue

throughout my day. Furthermore, I am greatly appreciative of the overwhelming support from researchers, biologists, veterinarians, and wildlife rehabilitators from around the country who have provided their years of experience in eagle conversation as well as samples to my project. Without their help, my project would not have been able to have its range wide perspective.

Throughout my PhD, my family, friends, and husband have seen me through the good days and the bad. My family and friends always provided an ear to listen when things were not going quite as planned and cheered me on throughout all of my successes. To my parents, thank you for instilling in me an appreciation for knowledge and support as I have worked hard to achieve my dreams. To my husband, Tell, thank you for being my backbone throughout the last six years when balancing my job and dissertation left little spare time for other things. I cannot express how grateful I am to have had you by my side through this roller coaster ride. Finally, I must include my dogs, Shadow, Hunter, and Mattie, who willingly lost many hours of sleep staying up with me every night to ensure the day's goals were met and who excitedly got me up the next morning to begin the day on time, regardless of how sleep deprived we were.

Name: Megan E. Judkins

Date of Degree: DECEMBER 2017

Title of Study: CONSERVATION GENOMICS OF NORTH AMERICAN BALD (*HALIAEETUS LEUCOCEPHALUS*) AND GOLDEN (*AQUILA CHRYSAETOS*) EAGLES

Major Field: Zoology

Abstract: Although eagles play pivotal roles in most Native Americans, First Nations, and Mesoamericans cultures and represent many countries around the world as the national symbol, these ecologically important predators still face many anthropogenic stressors. Stressors currently affecting bald and golden eagle populations include electrocution, habitat loss and degradation, lead poisoning, wind turbines, and disturbances. Despite the number of stressors these species face, neither of these species have been thoroughly studied from a genetic standpoint causing concern about the long-term conservation of these ecologically and culturally important species. To begin addressing the lack of genetic information about these species, each chapter addressed critical information that is lacking for both species. In Chapter 1, the Holarctic phylogeographic patterns, genetic variation, and demographic history of golden eagles in North America were assessed. The results indicated that there are two genetic lineages of golden eagles with only one Holarctic haplotype, there is little to no modern day gene flow between Nearctic and Palearctic golden eagles, and that the current distribution of haplotypes in the Nearctic reveal a recent population expansion with moderate levels of gene flow. For Chapter 2 and 3, a custom 100K SNP array was designed and subsequently used to assess the levels of the partitioning of genetic variation, SNPs under putative selection, and to begin the development of biologically sound management units. The results of both of the chapters revealed significant levels of genetic structure within the population and a list of putative SNPs under selection was developed. Finally, in Chapter 4, a population of bald eagles in New Jersey was evaluated to determine the nest turnover rate and relatedness of hatchling individuals in the population. Results show the population has high rates of inbreeding, moderate levels of nest turnover rates, and the presence of half sibling and unrelated relationships between individuals raised in the same nest in the same year. The findings of my research provide a new baseline of data for the conservation of bald and golden eagles. This information will be critical in the development of conservation management plans.

TABLE OF CONTENTS

Chapter Page

Chapter Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

# CHAPTER I

# HOLARCTIC PHYLOGEOGRAPHY OF GOLDEN EAGLES (AQUILA CHRYSAETOS) AND EVALUATION OF ALTERNATIVE NORTH AMERICAN MANAGEMENT APPROACHES


## Abstract

The golden eagle (*Aquila chrysaetos*) is a long-lived bird of prey with a Holarctic distribution. This species has survived severe anthropogenic stressors that have reduced or eliminated populations in some parts of its range. Despite the ecological and cultural importance of the golden eagle few attempts have been made to determine the partitioning of genetic variation over large areas of the species' range. This study generated DNA sequence data from the mtDNA control region from 115 North American golden eagles and combined these data with the previously existing control region sequences from over 300 Nearctic, Palearctic and Mediterranean golden eagles to provide a clearer holistic picture of the Holarctic phylogeographic patterns of genetic variation in this species and the genetic variation and demographic history of golden eagles in North America. Results of our study support that there are two genetic lineages of golden eagles, one representing the Mediterranean and the other occurring throughout the Holarctic. The Holarctic linage reveals little to no modern day gene flow between

Nearctic and Palearctic golden eagles. Furthermore, the current distribution of haplotypes in the Nearctic shows a recent population expansion with moderate levels of gene flow.

## Introduction

Climate perturbations throughout history have had effects on the evolutionary history and phylogeography of many species. The Quaternary (~1.8 MYA to present) climate fluctuations have been well studied and shown to have impacted the current distributions and speciation events of many taxa in the Northern Hemisphere (Hewitt, 2000, Hewitt, 2004). Raptors throughout the Holarctic display varying patterns of genetic differentiation with much of this variation being attributed to how they were impacted during the Pleistocene.

When considering modern day phylogeography of raptors, an east to west differentiation of haplotypes is commonly observed. Cooper's hawks (*Accipiter cooperi*; Sonsthagen *et al.,* 2012b) and sharp-shinned hawks (*A. striatus*; Hull & Girman, 2005) in North America, as well as the white-tailed eagle (*Haliaeetus albicilla*; Hailer *et al.,* 2007), bearded vulture (*Gypaetus barbatus*; Godoy *et al.,* 2004), and the Eurasian black vulture (*Aegypius monachus*; Poulakakis *et al.,* 2008) in the Palearctic exhibit this trend. This contrasts with the northern goshawk (*A. gentilis*) in North America in which an east-west differentiation was not observed. Instead, the southwestern populations differed significantly from the other sampled locations (Bayard de Volo, 2008). When a species' range stretches across the Holarctic, typically two patterns of genetic structuring between the Palearctic and Nearctic exist. In the first pattern, American and Eurasian

lineages live in sympatry with a distinct American haplogroup resulting from recent gene flow across Beringia or the Bering Strait (Nebel *et al.,* 2015). Alternatively, due to a lack of gene flow across Beringia or the Bering Strait, speciation occurs and distinct species are found in both North America and Eurasia (Nebel *et al.*, 2015).

Golden eagles (*Aquila chrysaetos*) are the only eagle species found throughout the Holarctic. Golden eagles are able to survive in diverse environments by adapting to the available prey, landscape, and their choice of nesting sites. For example, golden eagles have adapted to nesting on the ground in tree-less steppe and desert habitat as opposed to cliffs and trees in other parts of the world, and their diet includes diverse prey such as ptarmigan, hares, tortoise, snakes, and lizards (Watson, 2010). Furthermore, golden eagles are highly mobile at different stages of their lives or during different times of the year. Juvenile eagles are nomadic for approximately 5 years until they become sexually mature, at which time it has been shown, through banding studies that North American individuals travel back to within approximately 46-175 km of their natal location to establish a breeding territory (Millsap *et al.,* 2014). North American golden eagles that breed in northern latitudes are obligatory migrants, as compared to their southern counterparts that do not migrate, and will move to more southern latitudes in autumn and return to their nesting areas in spring. Even though they are thought to be highly philopatric to their natal site (Millsap *et al.,* 2014), these movements may allow for gene flow throughout the range of the species through behaviors such as extra pair copulations or dispersal events.

Little is known about the phylogeographic patterns of genetic variation of this species. Nebel *et al.* (2015) provided insight into the phylogeographic structure of

Palearctic golden eagles through examination of 402 bp of the mitochondrial (mtDNA) control region, but the North American subspecies was poorly represented due to a lack of available samples. Nebel *et al.* (2015) identified 26 haplotypes, only one of which was restricted to North America, and they speculated that there were two historical refugia. Sonsthagen *et al.* (2012a) and Craig *et al.* (2016) also utilized this region of the mtDNA as well as microsatellite loci to analyze golden eagles in the far western portion of their North American range. Sonsthagen *et al. (2012a)* identified five novel haplotypes and high gene flow when they analyzed golden eagles in California and the Channel Islands, while Craig *et al.* (2016) identified the five hapolotypes previously detected by Sonsthagen *et al*. (2012a) and three novel haplotypes when analyzing samples from Alaska, Idaho, and California. Other genetic studies on golden eagles have utilized microsatellites, allozymes, and mitochondrial DNA, but these studies focused on very limited portions of the species range (Masuda *et al.,* 1998; Suchentrunk *et al.,* 1998; Wink *et al.* 2004*;* Bielikova *et al.,* 2010; Bourke *et al.*, 2010).

Throughout their range, golden eagles have faced many anthropogenic stressors including illegal shooting, windfarms, lead poisoning, habitat loss, organic pollutants, and game keepers driving the population into bottlenecks or causing local extinctions (Fielding *et al.,* 2006; Bourke *et al.*, 2010; Stauber *et al.*, 2010; Watson, 2010; Pagel *et al.*, 2013). For example, breeding populations of golden eagles have been extirpated east of the Mississippi River in the United States (Morneau *et al.*, 2015), in Ireland (Bourke *et al*., 2010), and in the Alpine foothills and lowlands of Germany, Austria, the Czech Republic, and Poland (Nebel *et al.,* 2015). These bottlenecks and the population fragmentation caused by regional and local extirpations can allow for inbreeding

depression and loss of genetic variation adding to the need of determining the current partitioning of genetic variation for management purposes.

Currently, in the United States there are two proposed models for developing Eagle Management Units (U.S. Fish and Wildlife Service, 2016). The first management proposal includes the migratory flyways found in North America (Atlantic, Mississippi, Central, and Pacific) with the Mississippi and Atlantic flyways being combined. The alternative proposal is based on the currently recognized Bird Conservation Regions (BCR) with support coming from dispersal from nesting site (U.S. Fish and Wildlife Service, 2016). Recent studies utilizing nuclear markers (Doyle *et al.*, 2015; Craig *et al.*, 2016; Van Den Bussche *et al.*, 2017), reveal similar groupings of golden eagles across studies, with some support leaning towards the Bird Conservation Region model. Generally, these studies support the genetic uniqueness of golden eagles in northern California, southern Idaho, and southern Oregon and another group consisting of golden eagles from western Canada and Alaska. They differ from the Bird Conservation Region model by grouping most of the central through eastern portion of the western golden eagle range into a single unit (Doyle et al., 2015; Van Den Bussche et al., 2017). Unfortunately, due to a lack of sufficient sampling across the North American range of golden eagles it is unknown if this pattern is an artifact of sampling or if these individuals should be grouped into a single management unit.

This study utilized a fragment of the mtDNA control region previously used by Nebel *et al.* (2015), Sonsthagan *et al.* (2012a), and Craig *et al.* (2016) to generate DNA sequence data from 115 North American golden eagles representing Alaska, Alberta, Arizona, California, Colorado, Idaho, Montana, Nebraska, New Mexico, Oklahoma,

Oregon, Wyoming, Yukon, and British Colombia. Our goals were to fill in geographic gaps across the range of North American golden eagles to provide a clearer picture of the Holarctic phylogeographic patterns of genetic variation of golden eagles. Additionally, we were interested in evaluating whether the mitochondrial sequence data provided insight into the most appropriate geographic structure for managing golden eagles by evaluating levels of genetic differentiation when the samples were partitioned by Migratory Flyways, Bird Conservation Regions, or Genetic Regions. Finally, we were interested in elucidating the demographic history of golden eagles in North America to answer questions regarding the impacts of population bottlenecks and subsequent expansions on North America golden eagles. This additional information will allow for a more complete understanding of the historical movements and current patterns of mitochondrial genetic variation for golden eagles throughout the Holarctic.

## Materials and Methods

### *DNA isolation and sequence development*

Working with U.S. Fish and Wildlife permitted biologists and wildlife rehabilitators we obtained blood samples from 124 North American golden eagles. When nests contained more than one individual, only one sample was included in the study to eliminate over representation of a haplotype. Individuals that were not collected as a hatchling had natal location determined by GPS transmitters or isotope data. Only two samples (one individual from Colorado and a second individual from Wyoming) may not have a true natal origin represented as they were collected during the month of December and natal location data was unavailable. Blood samples (~ 0.5 ml) were

collected, stored in lysis buffer (Longmire *et al.,* 1997), and subsequently shipped to our laboratory at Oklahoma State University where DNA was extracted following the protocol of Longmire *et al.* (1997). DNA quality was assessed by running an aliquot on a 1% agarose gel and quantified using a NanoDrop 3300 spectrophotometer (Thermo Scientific).

A 442 bp region of domain I and II of the hypervariable control region-1 (D-loop) was amplified using polymerase chain reaction (PCR) and primers GOEA_CR1L and GOEA_CR595H developed by Sonsthagen et al (2012). PCR reactions contained 6ul of GoTaq Flexi Buffer (Promega), 25mM $MgCl_2$, 1 uM of each primer, 10mg/mL BSA, 4mM dNTPs, 1 unit of GoTaq (Promega), and $ddH_2O$ to make a 30uL reaction. The thermoprofile was 95°C for 5 min, 35 cycles of 95°C for 1 min, 40°C for 1 min, 72°C for 1 min, and a final extension at 72°C for 7 min. PCR products were visualized by running on a 1% agarose gel and subsequently cleaned using the Promega Wizard Kit (Fitchburg, WI, USA). Cleaned product were sequenced using BigDye v3.1 chain terminators (Thermo Fisher), 5X sequencing buffer (Edge Bio), GOEA_CR1L primer, and $ddH_2O$. The sequencing thermoprofile consisted of 35 cycles of 96°C for 30 sec, 50°C for 35 sec, and 60°C for 4 min. Sequencing products were cleaned using either Performa DTR V3 96 well short plates or Performa DTR Gel Filtration cartridges (Edge Bio). The cleaned sequencing product (~20 uL) was added to 1.0 uL of HiDi followed by electrophoresis using an ABI 3130 Genetic Analyzer (Applied Biosystems). Base calling was performed using Sequencing Analysis 5.2 (Applied Biosystems). Sequences were visually inspected and those sequences that were not long enough or had ambiguous bases were resequenced using the alternative primer (GOEA_CR595H) and

de novo aligned in Geneious 7.1.9 to acquire the full 442 bp region. In addition to the sequences we generated, we also downloaded from GenBank the Old World golden eagle sequences generated by Nebel *et al.* (2015; Accession numbers KR259251 – KR259276), the southern California sequences generated by Sonsthagen *et al.* (2012a; Accession numbers JQ246417 – JQ246421), and the sequences from California, Idaho, and Oregon generated by Craig et al. (2016; Accession numbersKX687705-KX687711). All sequences were aligned using the ClustalX multiple sequence alignment in Geneious 7.1.9. The resulting alignment was imported into MacClade version 4.06 (Maddison and Maddison 2000) for visual inspection and to trim our generated sequences as well as those of Sonsthagen *et al.* (2012a) and Craig *et al.* (2016) to 402 bp so that we could directly compare these data with those of Nebel *et al.* (2015). Using the Redundant Taxa option in MacClade we grouped sequences into haplotypes and assigned names to the haplotypes. Further analyses were conducted utilizing the data of Craig *et al.* (2016) and Nebel *et al.* (2015) as the Sonsthagen *et al.* (2012a) haplotype frequencies and locations were not available. Furthermore, approximately half of the golden eagles represented in Sonsthagen *et al.* (2012a) were from the Channel Islands, a location that only acquired golden eagles in the mid-1990's when bald eagles (*Haliaeetus leucocephalus*) were extirpated from the island, meaning the natal locations of the eagles that moved into the area are unknown.

*Data analysis*

Relationships among control region haplotypes of golden eagles were evaluated through the construction of a minimum spanning network at the 95% confidence level

using TCS v2.1 (Clement *et al.*, 2000). As described below, analyses were conducted on: (1) the entire data set including the samples from our dataset, Craig *et al.* (2016), and Nebel *et al.* (2015); (2) data partitioned by geography into Nearctic, Palearctic, and Mediterranean; and, (3) with the Nearctic samples further broken down by the proposed Migratory Flyways (U.S. Fish and Wildlife Service, 2016), Bird Conservation Regions (U.S. Fish and Wildlife Service, 2016), or Genetic Regions suggested by nuclear markers in previous studies (Doyle *et al.*, 2015; Craig *et al.*, 2016; Van Den Bussche *et al.*, 2017).

To assess genetic diversity for each of the three data partitions, we calculated haplotype and nucleotide diversity in Arlequin 3.1 (Excoffier & Lischer 2010). We also calculated the allelic richness using Contrib 1.4 (Petit & Pons 1998) in which the population sample size was normalized using a rarefaction correction. The rarefaction value was set to the smallest population size. We performed neutrality tests, Tajima's D and Fu's Fs, in Arlequin 3.1 to confirm the selective neutrality for the region. Tajima's D and Fu's Fs were further used to determine demographic histories. We used an Analysis of Molecular Variance (AMOVA) and pairwise Φst using the Kimuira 2-parameter model (K80) of evolution in Arlequin 3.1 to assess the partitioning of genetic variation within and among the Nearctic, Palearctic, and Mediterranean sample localities. For the North American samples, we utilized a hierarchical AMOVA to test the three proposed conservation models (Migratory Flyways, Bird Conservation Regions, Genetic Regions). The hierarchical approach grouped samples by State or Canadian Province and then the States or Provinces were grouped according the Migratory Flyways, Bird Conservation Regions or Genetic Regions.

## Results

We were able to generate DNA sequence data from 115 North American golden eagles which combined with the samples from Nebel *et al.* (2015) and Craig *et al.* (2016) provided us with a total of 417 samples. The 402-bp region contained 27 polymorphic sites and resulted in 44 haplotypes. The 115 Nearctic samples that we generated DNA sequences from provided 19 haplotypes (Table 1). Of these 19 haplotypes, 11 were novel to this study, four previously described by Sonsthagen *et al.* (2012a) and Craig *et al.* (2016), three previously described by Craig *et al.* (2016) and one previously described by Sonsthagen *et al.* (2012a), Nebel *et al.* (2015) and Craig *et al.* (2016). Although Nebel *et al.* (2015) classified their haplotypes as Holarctic and Mediterranean, based on our increased sampling of North American golden eagles, only haplotype H5 is Holarctic. To provide greater clarity for understanding golden eagle phylogeographic patterns, we used the following haplotypic designations. The single Holarctic haplotype (H5 of Nebel et al., 2015) is represented as "H1", and all other Holarctic haplotypes of Nebel *et al.* (2015) are Palearctic and identified in this study by the letter "P". We retained the Mediterranean haplotypic designation of Nebel *et al.* (2015) and assigned it with the prefix "M". Finally, all haplotypes that are found only in North America are designated by the prefix "N". Representative sequences of 11 new haplotypes detected in this study and restricted to North America have been deposited in GenBank (Accession number XXXXXXXX – XXXXXXXX).

*Phylogeographic structure*

Using the computer program TCS we produced a haplotype network in which the 95% confidence limit to connection was 8 steps (Fig. 1). Within this network the Mediterranean haplotypes appear to be a distinct lineage with two connections to the North American – Holarctic – Palearctic lineage. Even though we detected 18 haplotypes restricted to North America, there is no resolution between the Nearctic and Palearctic haplotypes.

*Genetic diversity*

Haplotype and nucleotide diversities for the three lineages were: Palearctic (h= 0.8149, $\pi$= 0.01412), Nearctic (h= 0.7354, $\pi$= 0.0028), and Mediterranean (h= 0.5835, $\pi$= 0.0022) (Table 2). The allelic richness with the rarefaction correction for the Nearctic was 8.4, Palearctic was 7.0 and Mediterranean was 6.7 (Table 2). The Nearctic lineage was statistically significant for Tajima's D, Fu's F, raggedness and SSD, while the Palearctic and Mediterranean linage was not statistically significant for any neutrality measure or the mismatch analysis.

Within North America haplotype diversity did not vary greatly when samples were partitioned by Migratory Flyways, Bird Conservation Regions, or Genetic Regions with most groupings possessing haplotype diversities of approximately 0.74 (Table 2). Nucleotide diversity ranged from 0.0019 to 0.0039 depending on how the samples were partitioned (Table 2). Tajima's D was not statistically significant for any of the three North American groupings while Fu's F was statistically significant for all populations across groupings (Table 2).

*Genetic structure*

When the data were partitioned into Nearctic, Palearctic, and Mediterranean, the AMOVA revealed a high level of genetic structuring with 47.62% of the genetic variation partitioned among these three units. Furthermore, the pairwise Φst comparison revealed statistically significant genetic differentiation between all lineages with the Palearctic/Nearctic differentiation being 0.517, the Nearctic/Mediterranean differentiation being 0.9139, and the Palearctic/Mediterranean differentiation being 0.3.

Within North America, we evaluated the partitioning of genetic variation when individuals were partitioned in accordance with Migratory Flyways, Bird Conservation Regions, and Genetic Regions. Figure 3 illustrates the distribution of haplotypes when partitioned by these alternative management scenarios. When samples were grouped by Migratory Flyways, there was statistically significant differentiation among populations within Migratory Flyways ($\Phi_{SC}$ = 0.0336*)* and within populations ($\Phi_{ST}$ = 0.0624) however, there was not statistically significant differentiation between the two Migratory Flyways (Table 3). When the hierarchical model was removed and samples were grouped only according to Administrative Flyways, the pairwise $\Phi_{ST}$ was statistically significant (Table 4). The hierarchical AMOVA for the Bird Conservation Regions did not detect statistically significant population structure among populations, among regions or within populations (Table 3), but the pairwise $\Phi_{st}$ revealed significant genetic differentiation between golden eagles from Alaska, Alberta, British Columbia and Yukon when compared to individuals from California, Idaho, and Oregon (Table 4). Additionally, individuals from California, Idaho and Oregon were significantly differentiated from individuals from Montana and Wyoming ($\Phi_{st}$ = 0.12). Finally, low,

but statistically significant levels of genetic differentiation ($\Phi_{st}$ = 0.044) was also detected between individual from Montana and Wyoming when compared to the grouping of individuals from Arizona, Colorado, and New Mexico (Table 4). Combining the Nebraska and Oklahoma samples with samples from Arizona, Colorado and New Mexico to reflect the Genetic Region grouping, the hierarchical AMOVA did not reveal significant population differentiation at any level tested (Table 3), but similar to other analyses the pairwise $\Phi_{st}$ detected statistically significant pairwise genetic differentiation was detected between the grouping of individuals from Alaska, Alberta, British Columbia, and Yukon relative to individuals from California, Idaho, and Oregon (Table 4). In congruence with the analysis based on Bird Conservation Regions, this analysis also detected statistically significant genetic differentiation between individuals from California, Idaho, and Oregon when compared to individuals from Montana and Wyoming (Table 3 & 4).

*Population fluctuations*

A unimodal mismatch distribution was detected for the Mediterranean, Palearctic, and Nearctic lineages, indicating a more recent demographic expansion although the only lineage that had a significant p-value for raggedness was the Nearctic (Fig. 2, Table 2). Within the Nearctic groupings, raggedness and SSD were both statistically significant when the samples were grouped into Migratory Flyways. When grouped into Bird Conservation Regions, raggedness was only significant for the grouping of individuals from Alaska, Alberta, British Columbia and Yukon as well as the grouping of individuals from California, Idaho, and Oregon. When samples were partitioned in

accordance with the Genetic Regions raggedness was only significant for the grouping of our most northern individuals from Alaska, Alberta, British Columbia and Yukon. SSD was not significant for either Bird Conservation Regions or Genetic Regions adding support to the recent expansion hypothesis.

**Discussion**

When a species' range occupies the Holarctic, one of two phylogeographic patterns typically emerge. North American and Eurasian lineages occur in sympatry with a distinct North American haplogroup resulting from recent gene flow across Beringia or the Bering Strait with examples such as the raven (*Corvus corax*) (Omland *et al.* 2000) and the peregrine falcon (*Falco peregrinus*) (Bell *et al.* 2014). Alternatively, due to a lack of gene flow across Beringia or the Bering Strait, speciation occurs and distinct species are found in both North America and Eurasia with examples such as white-tailed eagle (*Haliaeetus albicilla*) and bald eagle (*Haliaeetus leucocephalus*) (Hailer *et al.* 2007) or three-toed woodpeckers (*Picoides tridactylus*) and American three-toed woodpeckers (*Picoides dorsalis*) (Zink *et al.* 2002). Our results support the conclusions of Nebel *et al.* (2015) in which there are two distinct mtDNA lineages: a Mediterranean lineage and a Holarctic lineage (Fig. 1). As also determined by Nebel *et al.* (2015), the Mediterranean lineage represents a monophyletic group with individuals characterized by these haplotypes being found only in Mediterranean locations. The Holarctic linage represents haplotypes found throughout the Palearctic and the Nearctic, with only one haplotype (H1) being truly Holarctic. In North America, H1 is the most numerous and most widespread haplotype making up 45% of individuals sequenced as opposed to the

Palearctic where it represents few (<1%) individuals with a strong location bias in Europe. Taking the widespread range of H1 into account as well as its central location in the haplotype network, suggests that this is an ancestral haplotype. The pattern of having a single shared haplotype between hemispheres was also found when using the mitochondrial control region in ravens (Omland *et al.*, 2006). The other haplotypes defined as Holarctic by Nebal *et al.*, (2015) are only found on a single continent, but our data do not recover reciprocal monophyly between the Palearctic and Nearctic haplotypes (Figs. 1, 2).

Despite not detecting reciprocal monophyly between the Nearctic and Palearctic haplotypes, pairwise genetic differentiation indicates little to no current gene flow between the Nearctic and Palearctic linages ($\Phi_{st}$ = 0.517). Additionally, pairwise comparison of $\Phi_{st}$ between the Nearctic and Mediterranean indicates nearly a complete lack of gene flow ($\Phi_{st}$ = 0.9139). Among these three large geographic regions, there does appear to be some gene flow occurring between the Palearctic and Mediterranean lineages with a pairwise $\Phi_{st}$ of 0.3. Our data also suggest that the Nearctic population of golden eagles has undergone a recent population expansion. When comparing the Nearctic to the Palearctic and Mediterranean lineages, both Tajima's D and Fu's F were statistically significant. This paired with the unimodal mismatch analysis that had significant values for raggedness support a recent population expansion.

The first mitochondrial analysis of North American golden eagles examined individuals representing the California Channel Islands and the California mainland (Sonsthagen et al. (2012a). Among the 42 individuals that they were able to obtain mitochondrial sequence data, Sonsthagen et al. (2012a) detected five haplotypes. More

recently, Craig et al. (2106) examined the same mitochondrial region of golden eagles from Alaska, California, and Idaho. Among the 49 individuals they examined, they detected eight haplotypes, three unique to birds in their study and the five haplotypes previously detected in southern California golden eagles by Sonsthagen et al. (2012a). Additionally, Craig *et al.* (2016) reported the possibility of two region specific haplotypes, in which one of the haplotypes they detected appeared to be restricted to California whereas a second haplotype appeared to be restricted to Idaho. To these previous mitochondrial studies of North American golden eagles, we add data from 115 additional individuals. Overall, we detected 19 haplotypes with 11 haplotypes being newly detected variants. Our measures of haplotype diversity are higher for the Nearctic samples than detected by Nebel *et al.* (2015) for Old World samples, but our measures of nucleotide and haplotype diversity are similar to the results obtained by Craig *et al.* (2016). With regards to the conclusion of Craig et al. (2016), our data do not support the possibility of their haplotypes being region specific as N14, which Craig et al. (2016) suggested was restricted to California, was also detected in Oregon in our analyses. Similarly, haplotype N10, which Craig et al. (2016) suggested was restricted to Idaho, was detected in samples from Oregon and New Mexico. Of the 11 new haplotypes that we identified, eight were region specific. Haplotype N5 and N17 were only found in Oregon, N6, N7, N9 were found only found in Alaska, N13 was only found in California, N16 was only found in Idaho, and N18 was only found in Colorado (Table 1). Because these sequences were only single occurrences, we suspect that with more thorough sampling, they may be found to be in other areas. The only Holarctic haplotype (H1) was found in every sampling location in the Nearctic except Oklahoma, where only

one individual was sampled (Table 1). Due to the wide geographic distribution of this haplotype and that locations bordering Oklahoma had this haplotype, we suspect that if the sampling efforts were increased in future studies, this haplotype would be found in Oklahoma as well.

Due to the uncertainty in how best to manage North American golden eagles (U.S. Fish and Wildlife, 2016), we partitioned our samples so that the proposed Migratory Flyways and Bird Conservation Regions could be tested as well as the results indicated in studies utilizing nuclear DNA (Genetic Regions). Although the hierarchical AMOVAs failed to reveal statistically significant levels of overall genetic differentiation for any of the potential groupings, the analysis of pairwise $\Phi_{st}$ did reveal several patterns of significant genetic differentiation (Table 4). Most notably, for both the Bird Conservation Regions and the Genetic Regions, statistically significant pairwise genetic differentiation was detected between the grouping of northern golden eagles (Alaska and Canada) when compared with golden eagles from California, Idaho, and Oregon. Moreover, we detected significant genetic differentiation between the California, Idaho, Oregon birds when compared with the birds from Montana and Wyoming. Both of these groupings are found in different Conservation Regions and Flyways.

Based on the samples that we analyzed, the primary difference between the Bird Conservation Regions and the Genetic Regions is that for the Bird Conservation Regions our samples representing Arizona, Colorado, and New Mexico are considered to belong to a group genetically differentiated from our samples from Nebraska and Oklahoma whereas for the Genetic Regions approach the samples from these five states are combined into a single unit. The additional partitioning of samples in the Bird

Conservation region detected significant genetic differentiation between samples from the Northern Plains (Montana and Wyoming) and the southern Rocky Mountain regions (Arizona, Colorado, and New Mexico).

These finding based on Bird Conservation Regions and Genetic Regions are similar to results reported by Doyle *et al.* (2016) in which the Alaska and California eagles were significantly different than the other western eagle population. This was also reported by Van Den Bussche et al. (2017), in which the Alaska and British Columbia eagles created a significant unit and the Idaho, California, and Oregon eagles created a significant unit. Finally, when partitioning the Nearctic samples into Migratory Flyways, Bird Conservation Regions, or Genetic Regions, the flyway subgroupings as well as the most northern golden eagles (Alaska and Canada) and the California, Idaho, and Oregon golden eagles appear to have experienced a recent population expansion (Table 2).

**Conclusion**

Our study provides another example of intermediate divergence and the importance of studying the continuum of divergence for understanding geographical speciation, species limits, and conservation priorities (Omalnd *et al.*, 2006). This study also provides the first insight into the phylogeographic distribution and partitioning of mitochondrial genetic variation for North American golden eagles. To aid in determining more accurate historical events and to aid in determining proper management units, a more thorough sampling scheme should be carried out utilizing nuclear loci. Currently, single nucleotide polymorphisms (SNPs) have been used to

begin to elucidate the genetic structure of North American golden eagles, but with limited numbers of samples from throughout the range (Doyle *et al.,* 2016, Van Den Bussche *et al.,* 2017). Future studies that couple mitochondrial and nuclear loci can provided greater insight into the demography of golden eagle by revealing patterns of genetic variation at both maternally and biparentally inherited loci.

**Acknowledgments**

We would like to thank Bryan Bedrosian, Rob Domenech, Garth Herring, Gail Kratz, Brian Millsap, Gary Roemer, Victor Roubidoux, and Brian Smith for contributing golden eagle blood samples from their own research and rehabilitation efforts. Without their kindness and contribution, we would not have had the geographic coverage represented in this study. We would further like to thank the Iowa Tribe of Oklahoma for funding this study in collaboration with the Grey Snow Eagle House and the Ford Foundation for funding Megan Judkins' research assistantship while the work for this project was performed. We extend our appreciation to the other members of Megan Judkins dissertation committee, Drs. Meredith Hamilton, Michi Tobler, Andrew Doust, and Jim Lish whose review and feedback of an earlier version of this manuscript was critical. Finally, we thank the four anonymous reviewers of a previous version of this manuscript. Their comments were helpful in making this a much better manuscript.

**Literature Cited**


Bayard de Volo, Shelly. 2008. Genetic studies of Northern Goshawks (*Accipiter Gentilis)*: Genetic tagging and individual identification from feathers, and determining phylogeography, gene flow and population history for goshawks in North America. Unpublished D. Phil. Dissertation, Colorado State University.

Bell DA, Griffiths CS, Caballero IC, Hartley RR, Lawson RH. 2014. Genetic Evidence for Global Dispersal in the Peregrine Falcon ( *Falco peregrinus* ) and Affinity with the Taita Falcon ( *Falco fasciinucha)*. *Journal of Raptor Research* 48: 44–53.

Bielikova M, Ficek A, Valkova D, Turna J. 2010. Multiplex PCR amplification of 13 microsatellite loci for Aquila chrysaetos in forensic applications. *Biologia* 65: 1081–1088.

Bourke BP, Frantz AC, Lavers CP, Davison A, Dawson DA, Burke TA. 2010. Genetic signatures of population change in the British golden eagle (Aquila chrysaetos). *Conservation Genetics* 11: 1837–1846.

Clement M, Posada D, Crandall KA. 2000. TCS: A computer program to estimate gene genealogies. *Molecular Ecology* 9: 1657–1659.

Craig EH, Adams JR, Waits LP, Fuller MR, Whittington DM. 2016. Nuclear and mitochondrial DNA analyses of golden eagles (Aquila chrysaetos canadensis) from three areas in western North America; initial results and conservation implications. *PLoS ONE* 11: 1–15.

Doyle JM, Katzner TE, Roemer GW, Cain III JW, Millsap BA, McIntyre CL, Sonsthagen SA, Fernandez NB, Wheeler M, Bulut Z, Bloom PH, DeWoody A. 2016. Genetic structure and viability selection in the golden eagle (*Aquila chrysaetos*): a vagile raptor with a Holarctic distribution. *Conservation Genetics* 17: 1–16.

Excoffier L, Lischer HEL. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* 10: 564–567.

Fielding AH, Whitfield DP & McLeod DRA. 2006. Spatial association as an indicator of the potential for future interactions between wind energy developments and golden eagles Aquila chrysaetos in Scotland. *Biological Conservation* 131: 359–369.

Godoy JA, Negro JJ, Hiraldo F, Donazar JA. 2004. Phylogeography, genetic structure and diversity in the endangered bearded vulture (Gypaetus barbatus, L.) as revealed by mitochondrial DNA. *Molecular Ecology* 13: 371–390.

Hailer F, Helander B, Folkestad AO, Ganusevich SA, Garstad S, Hauff P, Koren C, Nygard T, Volke V, Vila C, Ellegren H. 2006. Bottlenecked but long-lived: high genetic diversity retained in white-tailed eagles upon recovery from population decline. *Biology letters* 2: 316–9.

Hewitt G. 2000. The genetic legacy of the Quaternary ice ages. *Nature* 405: 907–913.

Hewitt GM. 2004. Genetic consequences of climatic oscillations in the Quaternary. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 359: 183–195.

Hull JM, Girman DJ. 2005. Effects of Holocene climate change on the historical demography of migrating sharp-shinned hawks (Accipiter striatus velox) in North America. *Molecular Ecology* 14: 159–170.

Longmire, JL, Maltbie M, Baker RJ. 1997. Use of "lysis buffer" in DNA isolation and its implication for museum collections. *Occasional papers, The Museum, Texas Tech University*, 163:1 – 3.

Maddison, D. R. and W. P. Maddison, 2005. MacClade 4: Analysis of phylogeny and character evolution.

Masuda R, Noro M, Kurose N, Nishida-Umehara C, Takechi H, Yamazaki T, Kosuge M, Yoshida MC. 1998. Genetic characteristics of endangered Japanese golden eagles (Aquila chrysaetos japonica) based on mitochondrial DNA D-loop sequences and karyotypes.

Millsap BA, Harmata AR, Stahlecker DW, Mikesic DG. 2014. Natal Dispersal Distance of Bald and Golden Eagles Originating in the Coterminous United States as Inferred from Band Encounters. *Journal of Raptor Research* 48: 13–23.

Morneau F, Tremblay JA, Todd C, Chubbs TE, Maisonneuve C, Lemaitre J, Katzner T. 2015. Known Breeding Distribution and Abundance of Golden Eagles in Eastern North America. *Northeastern Naturalist* 22: 236–247.

Nebel C, Gamauf A, Haring E, Segelbacher G, Villers A, Zachos FE. 2015. Mitochondrial DNA analysis reveals Holarctic homogeneity and a distinct

Mediterranean lineage in the Golden eagle (Aquila chrysaetos). *Biological Journal of the Linnean Society* 116: 328–340.

Omland KE, Baker JM & Peters JL. 2006. Genetic signatures of intermediate divergence: Population history of Old and New World Holarctic ravens (Corvus corax). *Molecular Ecology* 15: 795–808.

Pagel JE, Kritz KJ, Millsap BA, Robert K. 2013. Bald Eagle and Golden Eagle Mortalities at Wind Energy Facilities in the Contiguous United States. *Journal of Raptor Research* 47: 311–315.

Petit, El Mousadik & Pons. 1998. Identifying populations for conservation on the basis of genetic markers. Conservation Biology 12, 844-855.

Poulakakis N, Antoniou A, Mantziou G, Parmakelis A, Skartsi T, Vasilakis D, Elorriaga J, De La Puente J, Gavashelishvili A, Ghasabyan M, Katzner T, McGrady M, Batbayar N, Fuller M, Natsagdorj T. 2008. Population structure, diversity, and phylogeography in the near-threatened Eurasian black vultures. *Natural History*: 859–872.

Sonsthagen SA., Coonan TJ, Latta BC, Sage GK, Talbot S. 2012a. Genetic diversity of a newly established population of golden eagles on the Channel Islands, California. *Biological Conservation* 146: 116–122.

Sonsthagen SA, Rosenfield RN, Bielefeldt J, Murphy, RK, Steward, AC, Stout WE, Driscoll TG, Bozek MA, Sloss BL, Talbot SL. 2012b. Genetic and morphological

divergence among Cooper's Hawk (*Accipiter cooperii*) populations breeding in North-Central and Western North America. *The Auk* 129: 427–437.

Stauber E, Finch N, Talcott PA, Gay JM. 2010. Lead poisoning of bald (Haliaeetus leucocephalus) and golden (Aquila chrysaetos) eagles in the U.S. inland Pacific Northwest region--an 18-year retrospective study: 1991-2008. *Journal of avian medicine and surgery* 24: 279–287.

Suchentrunk F, Haller H, Ratti P. 1999. Gene pool variability of a golden eagle (Aquila chrysaetos) population from the Swiss Alps. *Biological Conservation* 90: 151–155.

U.S. Fish and Wildlife Service. 2016. Bald and Golden Eagles: Population demographics and

estimation of sustainable take in the United States, 2016 update. Division of Migratory Bird Management, Washington D.C., USA.

Van Den Bussche RA, Judkins ME, Montague MJ, Warren WC. 2017. A resource of genome-wide single nucleotide polymorphisms (SNPs) for the conservation and management of golden eagles (*Aquila chrysaetos*). *Journal of Raptor Research* 51:368-377.

Watson J, Brockie K, and Watson D. 2010. The golden eagle. New Haven: Yale University Press. 2nd edition.

Wink, M, Clouet M, Goar LJ, Barrau C. 2004. Sequence variation

in the cytochrome b gene of subspecies of Golden Eagles Aquila chrysaetos. *Alauda* 72:153-157.

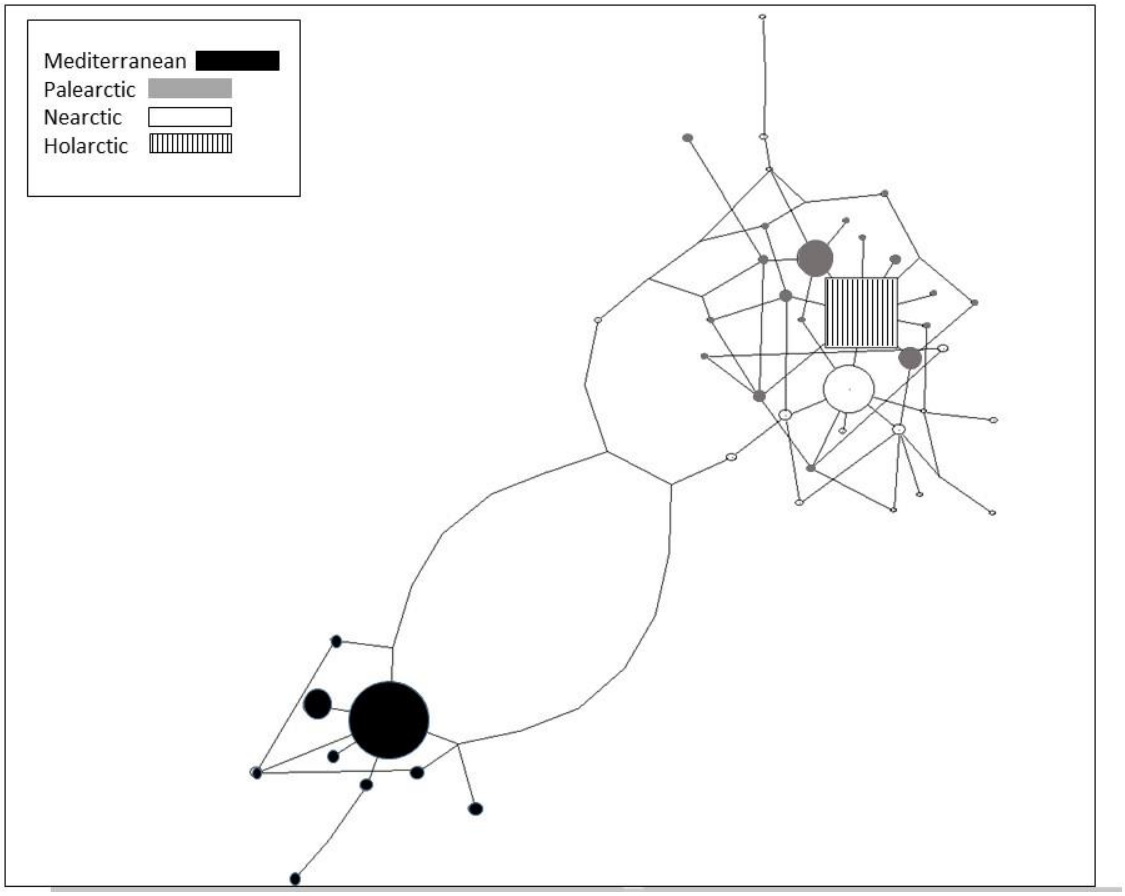
Zink RM, Barrowclough GF, Atwood JL, Blackwell-Rago RC. 2000. Genetics, taxonomy, and conservation of the threatened California Gnatcatcher. *Conservation Biology* 14: 1394–1405.

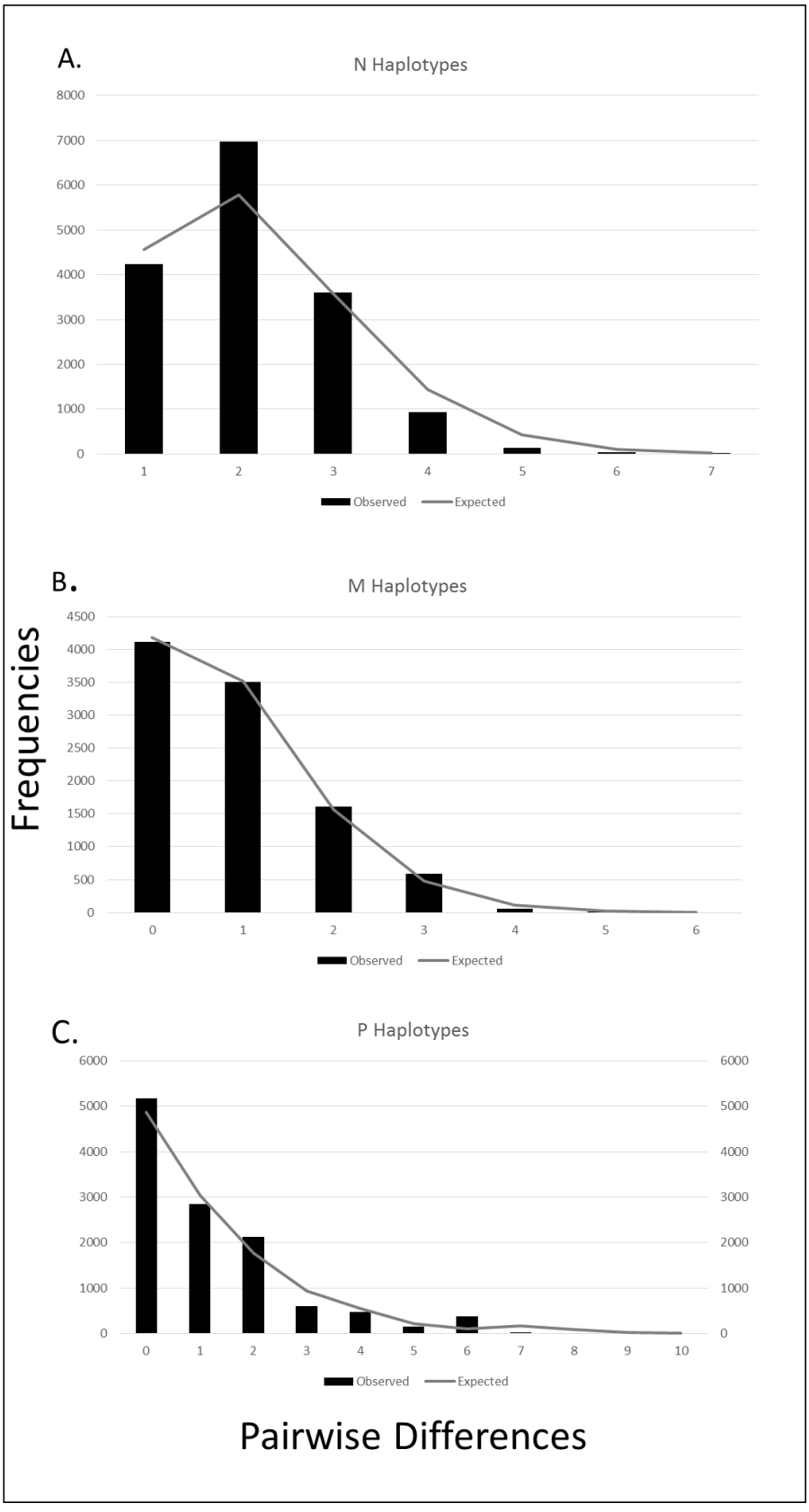
**Figure Legends**

Figure 1- Network diagram of all 44 control region haplotypes as determined by a minimum spanning network at the 95% confidence level using TCS v2.1. Colors represent lineages as determined in this study with yellow representing the single Holarctic haplotype, black representing the Mediterranean lineage, blue representing the Nearctic lineage, and orange representing the Palearctic lineage. The size of the shape is proportional to the number of individuals represented by that haplotype.

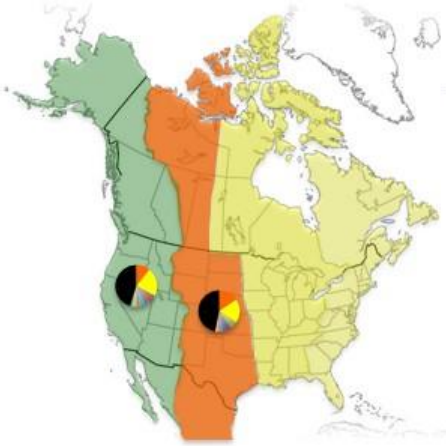
Figure 2- Mismatch distribution for each of the lineages: Mediterranean, Palearctic, and Nearctic. The vertical bars represent the observed frequencies while the line represents the expected frequencies under a model of sudden expansion.

Figure 3-Map of North American illustrating haplotype distributions when samples are partitioned according to Migratory Flyways, Bird Conservation Regions, and Genetic Regions. Pie graphs represent the haplotype frequency for the given Migratory Flyway, Bird Conservation Region, and Genetic Region. The colors in each map represent the different proposed management units for each model.

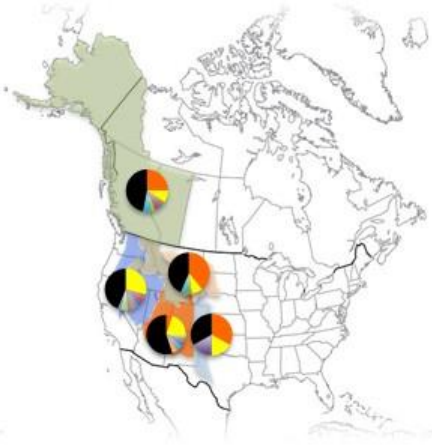

Mediterranean
Palearctic
Nearctic
Holarctic


A.
N Haplotypes
8000
7000
6000
5000
4000
3000
2000
1000
0
1
2
3
4
5
6
7
Observed
Expected
B.
M Haplotypes
Frequencies
4500
4000
3500
3000
2500
2000
1500
1000
500
0
0
1
2
3
4
5
6
Observed
Expected
C.
P Haplotypes
6000
5000
4000
3000
2000
1000
0
0
1
2
3
4
5
6
7
8
9
10
Observed
Expected
Pairwise Differences

Migratory Flyways



Bird Conservation Regions



Genetic Regions

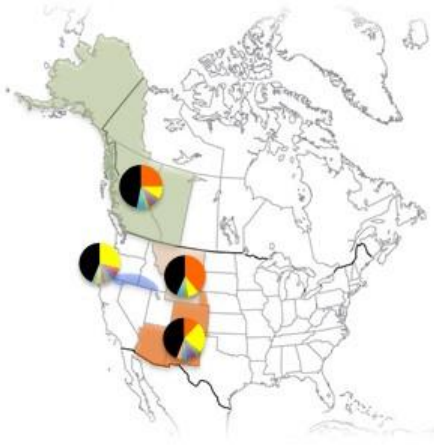


Table 1 Nearctic sample haplotype number with the total number of samples (n), total number of unique haplotypes, and the total number of each haplotype by state. States are classified according to Migratory Flyways (Pacific (P) or Central (C)), Bird Conservation Regions (CR1, 2, 3, 4, or 5), and Genetic Regions (GR1, 2, 3, or 4).

| | n | Total Haplotypes | N1 | N2 | N3 | N4 | N5 | A6 | N7 | N8 | N9 | N10 | N11 | N12 | N13 | N14 | N15 | N16 | N17 | N18 | H1 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Alaska (P, CR1, GR1) | 24 | 8 | 5 | 2 | | | | 1 | 1 | 1 | 1 | | 1 | | | | | | | | 12 |
| Alberta (P, CR1, GR1) | 2 | 2 | | | | | | | | | | | 1 | | | | | | | | 1 |
| British Columbia (P, CR1, GR1) | 4 | 3 | 2 | 1 | | | | | | | | | | | | | | | | | 1 |
| Yukon (P, CR1, GR1) | 2 | 2 | 1 | | | | | | | | | | | | | | | | | | 1 |
| California (P, CR2, GR2) | 33 | 6 | | 9 | | | | | | | | 2 | 2 | | 1 | | 3 | | | | 16 |
| Oregon (P, CR2, GR2) | 30 | 9 | | 12 | | | 1 | | | 2 | | 1 | 1 | 2 | | 1 | | | 1 | | 9 |
| Idaho (P, CR2, GR2) | 26 | 6 | | 5 | | | | | | 3 | | | | | | 1 | 2 | 1 | | | 14 |
| Wyoming (C, CR3, GR3) | 16 | 3 | 7 | 1 | | | | | | | | | | | | | | | | | 8 |
| Montana (C, CR3, GR3) | 7 | 5 | 2 | 1 | 1 | | | | | | | | 1 | | | | | | | | 2 |
| Nebraska (C, CR4, GR4) | 5 | 4 | 1 | 1 | | 1 | | | | | | | | | | | | | | | 2 |
| Oklahoma (C, CR4, GR4) | 1 | 1 | 1 | | | | | | | | | | | | | | | | | | |
| New Mexico (C, CR5, GR4) | 4 | 3 | | 1 | | | | | | | | 1 | | | | | | | | | 2 |
| Colorado (C, CR5, GR4) | 14 | 6 | 1 | 2 | 1 | | | | | | | | 1 | | | | | | | 1 | 8 |
| Arizona (C, CR5, GR4) | 2 | 2 | | 1 | | | | | | | | | | | | | | | | | 1 |
| Total | | | 20 | 36 | 2 | 1 | 1 | 1 | 1 | 6 | 1 | 4 | 7 | 2 | 1 | 2 | 5 | 1 | 1 | 1 | 77 |

Table 2 Statistical analyses of American golden eagles with samples grouped according to geographic lineages, Migratory Flyways, Bird Conservation Regions, and Genetic Regions. Statistics include number of samples (n) haplotype diversity (*h*), nucleotide diversity ($\pi$), allelic richness, Tajima's D and Fu's F, SSD and raggedness. Results that are bold indicate significance at the 0.05 level. State and Canadian Province abbreviations are as follows: Alberta (AB), Alaska (AK), Arizona (AZ), British Columbia (BC), California (CA), Colorado (CO), Idaho (ID), Montana (MT), Nebraska (NE), New Mexico (NM), Oklahoma (OK), Oregon (OR), Wyoming (WY), Yukon (YT).

| | n | *h* | $\pi$ | Allelic Richness | Tajima's D | Fu's Fs | SSD | Raggedness |
|---|---|---|---|---|---|---|---|---|
| **Lineages** | | | | | | | | |
| North American | 170 | 0.81 | 0.003 | 8.4 | **-1.48** | **13.62** | **0.01** | **0.12** |
| Palearctic | 217 | 0.74 | 0.014 | 7 | 1.71 | -1.41 | 0.08 | 0.7 |
| Mediterranean | 27 | 0.58 | 0.002 | 6.7 | -0.9 | -3.1 | 0.0001 | 0.54 |
| **Flyways** | | | | | | | | |
| Central | 99 | 0.73 | 0.004 | - | -0.75 | **-3.65** | **0.202** | **0.66** |
| Pacific | 80 | 0.72 | 0.003 | - | -0.95 | **-7.22** | **0.14** | **0.47** |
| **Conservation Regions** | | | | | | | | |
| AK-AB-BC-YT | 24 | 0.72 | 0.003 | 2.58 | -1.33 | **-3.42** | 0.02 | **0.17** |
| CA-ID-OR | 89 | 0.75 | 0.002 | 2.37 | -0.89 | **-5.67** | 0 | **0.09** |
| MT-WY | 40 | 0.71 | 0.002 | 1.94 | -1.3 | **-2.77** | 0.02 | 0.15 |
| NE-OK | 6 | 0.87 | 0.003 | 3 | -0.68 | **-0.99** | 0.01 | 0.13 |
| AZ-CO-NM | 20 | 0.68 | 0.002 | 2.29 | -1.33 | **3.61** | 0.01 | 0.09 |
| **Genetic Regions** | | | | | | | | |
| AK-AB-BC-YT | 24 | 0.72 | 0.003 | 6.13 | -1.33 | **3.42** | 0.02 | **0.17** |
| CA-ID-OR | 89 | 0.75 | 0.002 | 5.7 | -0.89 | **5.67** | 0.01 | 0.09 |
| MT-WY | 40 | 0.71 | 0.002 | 4 | -1.3 | **2.77** | 0.02 | 0.15 |
| NE-OK-AZ-CO-NM | 26 | 0.72 | 0.003 | 6.42 | -1.22 | **3.88** | 0.004 | 0.09 |

Table 1 Nearctic sample haplotype number with the total number of samples (n), total number of unique haplotypes, and the total number of each haplotype by state. States are classified according to Migratory Flyways (Pacific (P) or Central (C)), Bird Conservation Regions (CR1, 2, 3, 4, or 5), and Genetic Regions (GR1, 2, 3, or 4).

| | Migratory Flyways | | | | Bird Conservation Regions | | | | Genetic Regions | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Degrees of Freedom | Sum of Squares | Variance Components | Percentage of variation | Degrees of Freedom | Sum of Squares | Variance Components | Percentage of variation | Degrees of Freedom | Sum of Squares | Variance Components | Percentage of variation |
| Among Groups | 1 | 2.01 | 0.17 | 2.99 | 3 | 3.55 | 0.03 | 4.32 | 3 | 3.63 | 0.023 | 3.99 |
| Among Populations | 12 | 8.95 | 0.02 | 3.26 | 7 | 3.48 | -0.01 | -0.88 | 9 | 4.70 | -0.004 | -0.71 |
| Within Population | 156 | 84.20 | 0.54 | 93.75 | 137 | 76.13 | 0.56 | 96.55 | 141 | 79.33 | 0.56 | 96.72 |
| Fsc | **0.0336** | | | | -0.00917 | | | | -0.00737 | | | |
| Fst | **0.0624** | | | | 0.0345 | | | | 0.03278 | | | |
| Fct | 0.0299 | | | | 0.04325 | | | | 0.0399 | | | |

Table 4 Pairwise Φst values for Migratory Flyways, Bird Conservation Regions, and Genetic Regions. Number are bolded if statistically significant at 0.05.



| Migratory Flyways | | | Bird Conservation Regions | | | | | | Genetic Regions | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | West | East | | AK-AB-BC-Y | CA-ID-OR | MO-WY | NE-OK | AZ-CO-NM | | AK-AB-BC-Y | CA-ID-OR | MO-WY | NE-OK-AZ-CO-NM |
| Central | 0 | | AK-AB-BC-Y | 0 | | | | | AK-AB-BC-Y | 0 | | | |
| Pacific | **0.018** | 0 | CA-ID-OR | **0.09** | 0 | | | | CA-ID-OR | **0.09** | 0 | | |
| | | | MT-WY | -0.026 | **0.117** | 0 | | | MT-WY | -0.0262 | **0.117** | 0 | |
| | | | NE-OK | -0.062 | 0.068 | -0.0532 | 0 | | NE-OK-AZ-CO- | -0.0002 | 0.015 | 0.0178 | 0 |
| | | | AZ-CO-NM | 0.02 | -0.003 | **0.044** | 0.0072 | 0 | | | | | |

S1 Locality, age, provider, and haplotype information for the 115 samples sequenced for the Nearctic portion of the study.

| From | Age | State | Haplotype |
|---|---|---|---|
| Brain Smith | hatchling | Wyoming | H5 |
| Brian Smith | hatchling | California | H5 |
| Brian Smith | hatchling | California | H5 |
| Brian Smith | hatchling | California | N2 |
| Brian Smith | hatchling | California | N2 |
| Brian Smith | hatchling | Colorado | H5 |
| Brian Smith | hatchling | Colorado | H5 |
| Brian Smith | hatchling | Colorado | N1 |
| Brian Smith | adult | Colorado | N11 |
| Brian Smith | hatchling | Nebraska | H5 |
| Brian Smith | hatchling | Nebraska | H5 |
| Brian Smith | hatchling | Nebraska | N2 |
| Brian Smith | hatchling | Nebraska | N4 |
| Brian Smith | adult | Wyoming | H5 |
| Brian Smith | hatchling | Wyoming | H5 |
| Brian Smith | hatchling | Wyoming | H5 |
| Brian Smith | hatchling | Wyoming | H5 |
| Brian Smith | hatchling | Wyoming | H5 |
| Brian Smith | hatchling | Wyoming | H5 |
| Brian Smith | hatchling | Wyoming | N1 |
| Brian Smith | hatchling | Wyoming | N1 |
| Brian Smith | hatchling | Wyoming | N1 |
| Brian Smith | hatchling | Wyoming | N1 |
| Brian Smith | hatchling | Wyoming | N1 |
| Brian Smith | hatchling | Wyoming | N1 |
| Brian Smith | hatchling | Wyoming | N1 |
| Brian Smith | hatchling | Wyoming | N2 |
| Brian Smith | hatchling | Nebraska | N1 |
| Bryan Bedrosian | unknown | Alaska | N1 |
| Bryan Bedrosian | unknown | Alaska | N2 |
| Bryan Bedrosian | unknown | Montana | N1 |
| Bryan Bedrosian | unknown | Montana | N1 |
| Bryan Bedrosian | unknown | Montana | N3 |
| Bryan Bedrosian | unknown | Yukon | N1 |
| Garth Herring | hatchling | California | H5 |
| Garth Herring | hatchling | California | H5 |
| Garth Herring | hatchling | California | H5 |
| Garth Herring | hatchling | California | H5 |

| | | | |
|---|---|---|---|
| Garth Herring | hatchling | California | N13 |
| Garth Herring | hatchling | California | N2 |
| Garth Herring | hatchling | Idaho | H5 |
| Garth Herring | hatchling | Idaho | H5 |
| Garth Herring | hatchling | Idaho | H5 |
| Garth Herring | hatchling | Idaho | N15 |
| Garth Herring | hatchling | Idaho | N16 |
| Garth Herring | hatchling | Oregon | H5 |
| Garth Herring | hatchling | Oregon | H5 |
| Garth Herring | hatchling | Oregon | H5 |
| Garth Herring | hatchling | Oregon | H5 |
| Garth Herring | hatchling | Oregon | H5 |
| Garth Herring | hatchling | Oregon | H5 |
| Garth Herring | hatchling | Oregon | H5 |
| Garth Herring | hatchling | Oregon | H5 |
| Garth Herring | hatchling | Oregon | H5 |
| Garth Herring | hatchling | Oregon | H5 |
| Garth Herring | hatchling | Oregon | N10 |
| Garth Herring | hatchling | Oregon | N11 |
| Garth Herring | hatchling | Oregon | N12 |
| Garth Herring | hatchling | Oregon | N12 |
| Garth Herring | hatchling | Oregon | N14 |
| Garth Herring | hatchling | Oregon | N17 |
| Garth Herring | hatchling | Oregon | N2 |
| Garth Herring | hatchling | Oregon | N2 |
| Garth Herring | hatchling | Oregon | N2 |
| Garth Herring | hatchling | Oregon | N2 |
| Garth Herring | hatchling | Oregon | N2 |
| Garth Herring | hatchling | Oregon | N2 |
| Garth Herring | hatchling | Oregon | N2 |
| Garth Herring | hatchling | Oregon | N2 |
| Garth Herring | hatchling | Oregon | N2 |
| Garth Herring | hatchling | Oregon | N2 |
| Garth Herring | hatchling | Oregon | N2 |
| Garth Herring | hatchling | Oregon | N2 |
| Garth Herring | hatchling | Oregon | N5 |
| Garth Herring | hatchling | Oregon | N8 |
| Garth Herring | hatchling | Oregon | N8 |
| Gary Roemer | hatchling | Arizona | H5 |
| Gary Roemer | hatchling | Arizona | N2 |
| Gary Roemer | hatchling | Colorado | H5 |
| Gary Roemer | hatchling | Colorado | H5 |
| Gary Roemer | hatchling | Colorado | H5 |
| Gary Roemer | hatchling | Colorado | H5 |

| | | | |
|---|---|---|---|
| Gary Roemer | hatchling | Colorado | N18 |
| Gary Roemer | hatchling | Colorado | N2 |
| Gary Roemer | hatchling | Colorado | N2 |
| Gary Roemer | hatchling | New Mexico | H5 |
| Gary Roemer | hatchling | New Mexico | H5 |
| Gary Roemer | hatchling | New Mexico | N10 |
| Gary Roemer | hatchling | New Mexico | N2 |
| Victor Roubidoux | adult | Oklahoma | N1 |
| Gail Kratz | hatchling | Colorado | H5 |
| Gail Kratz | hatchling | Wyoming | H5 |
| Gail Kratz | hatchling | Colorado | N3 |
| Gail Kratz | hatchling | Colorado | H5 |
| Rob Domenech | adult | Alaska | H5 |
| Rob Domenech | subadult | Alaska | H5 |
| Rob Domenech | subadult | Alaska | H5 |
| Rob Domenech | subadult | Alaska | H5 |
| Rob Domenech | adult | Alaska | N1 |
| Rob Domenech | adult | Alaska | N1 |
| Rob Domenech | adult | Alaska | N1 |
| Rob Domenech | adult | Alaska | N1 |
| Rob Domenech | adult | Alaska | N2 |
| Rob Domenech | adult | Alaska | N6 |
| Rob Domenech | adult | Alaska | N7 |
| Rob Domenech | adult | Alberta | H5 |
| Rob Domenech | adult | Alberta | N11 |
| Rob Domenech | adult | British Columbia | H5 |
| Rob Domenech | adult | British Columbia | N1 |
| Rob Domenech | adult | British Columbia | N1 |
| Rob Domenech | adult | British Columbia | N2 |
| Rob Domenech | adult | Montana | H5 |
| Rob Domenech | adult | Montana | H5 |
| Rob Domenech | adult | Montana | N11 |
| Rob Domenech | adult | Montana | N2 |
| Rob Domenech | subadult | Yukon | H5 |

# CHAPTER II

## CHARACTERIZATION OF GENETIC STRUCTURE AND ADAPTATION USING A 37K SNP ARRAY FOR GOLDEN EAGLES (*AQUILA CHRYSAETOS)*


**ABSTRACT**

Golden eagles (*Aquila chrysaetos*) face many anthropogenic stressors throughout their Nearctic range. The presence of these stressors and their effects have triggered a recent push to develop more efficient conservation strategies to determine the current conservation status of golden eagles. The first step in any conservation management plan should be identification of genetically based conservation management units to aid in preserving genetic variation, determining gene flow, and avoiding inbreeding depression. Although previous studies have been conducted to determine genetically based management units for western golden eagles, a precise understanding of the size and location of these units is still unclear. In this study, we developed a custom 37K Affymetrix Axiom myDesign Single Nucleotide Polymorphism (SNP) array to further our previous attempts at developing management units and characterizing SNPs under putative selection. Using this array, we were able to successfully genotype 137 western golden eagles at 36,560 loci. Using STRUCTURE, Geneland, and a PCoA to evaluate population structure, a finer resolution for the genetic structure of western golden eagles

in the western U.S. is described. Finally, using a conservative approach to characterize SNPs under selection, we were able to identify 85 SNPs that are under putative selection.

## Introduction

The development of arrays (also known as chips) for the high throughput analysis of hundreds to thousands of Single Nucleotide Polymorphisms (SNPs), distributed throughout the genome, for non-model species has significantly changed how species of conservation concern can be managed. Data generated from the increased number of loci can improve the accuracy of population genetic estimates, allow for 'real-time' migration tracking through the use of assignment tests, and provide new insights from previously unstudied areas of the genome (Nosil *et al.* 2005; Allendorf *et al.* 2010; Stapley *et al.* 2010; Hess *et al.* 2013; Lozier 2014; Pilot *et al.* 2014; Benestan *et al.* 2015; Cammen *et al.* 2015; Malenfant *et al.* 2015). Over the past decade there has been a rapid increase in the number of population genomic studies on non-model organisms that illustrate how advances in genomic technologies have revolutionized information that can be obtained and included in conservation based studies. Examples include the polar bear (*Ursus maritimus*, Malenfant *et al.* 2015), great tit ( *Parus major*, Van Bers *et al.* 2012), Atlantic salmon ( *Salmo salar*, Bourret *et al.* 2013), and grey wolf (*Canis lupis*, Schweizer *et al.* 2015).

In the United States, the breeding population of golden eagles (*Aquila chrysaetos*) east of the Mississippi River has been extirpated (Morneau *et al.* 2015) and although studies indicate that the western population as a whole is stable, some populations may be in decline (Kochert *et al.* 2002; Millsap *et al.* 2013). Currently, golden eagles face anthropogenic threats such as electrocutions, illegal shootings, lead toxicosis, collision

(i.e. vehicles, airplanes, wind turbines), and poisoning. To help combat the number of birds lost by these stressors, U.S. Fish and Wildlife Service is working to create a comprehensive management plan for golden eagles that includes determining conservation management units that are biologically relevant for the species. Currently, two scenarios are being considered. One is based on the Migratory Flyways found in North America (Atlantic, Mississippi, Central, and Pacific) with the Mississippi and Atlantic flyways being combined (U.S. FWS 2016). Support for managing golden eagles based on Migratory Flyways comes from the observation that it fits seasonal movement patterns (U.S. FWS 2016). The alternative scenario closely corresponds to the Bird Conservation Regions (U.S. FWS 2016) and is supported based on average natal dispersal distances (Millsap *et al*. 2014). Unfortunately, neither approach has been evaluated in light of genetic data and thus, it is unclear if either represent biologically relevant approaches for determining how best to manage western golden eagles.

In North America, golden eagles typically nest on the sides of cliffs, but can also be found nesting in Douglas firs, cottonwoods, and Ponderosa pines with an estimated nest density of 10-20 nesting pairs/1,000km$^2$ (Watson 2010). The difficulty in gaining access to nests combined with nests being sparsely distributed makes nest ascertainment difficult. In addition, golden eagles are protected in the United States under the Migratory Bird Treaty Act and the Bald and Golden Eagle Protection Act. Whereas these Acts protect eagles from disturbance and illegal take, they also make it exceedingly difficult to conduct research that requires biological samples.

Few genetic studies exist for North American golden eagles. Using nine microsatellite loci and sequencing a portion of the mtDNA control region for 71 golden

eagles from the Channel Islands and adjacent California mainland, Sonsthagen *et al.* (2012) documented population structure along with gene flow between the two regions; however, the number of genetic clusters they detected varied depending upon the analysis used. Craig *et al.* (2016) also used the mtDNA control region and microsatellites to evaluate the level of differentiation between golden eagles in Alaksa, California, and Idaho. Their study revealed high levels of genetic diversity among western golden eagles from each of these areas. Due to low samples sizes from Alaska (6 individuals), they did not include these individuals in analyses of population differenetiation. When comparing golden eagles from California and Idaho, Craig *et al*. (2016) did not detect any significant genetic differentiation.

Doyle *et al.* (2016) utilized 162 SNPs to perform a population genetics analysis of golden eagles from across North America. Their analyses revealed three genetic clusters, (1) Alaska, (2) California, and (3) golden eagles from Arizona, Colorado, Nebraska, New Mexico, Utah, and Wyoming. Van Den Bussche *et al.* (2017) utilized a suite of 30,006 SNPs for their population genetics analysis and similar to Doyle *et al.* (2016) detected three genetic groupings, (1) Alaska and British Columbia, (2) northern California, southern Oregon, southern Idaho and, (3) a cluster consisting of golden eagles from Arizona, Colorado, Nebraska, New Mexico, Oklahoma, and Wyoming.

Most recently, Judkins & Van Den Bussche (In Press) added to the mitochondrial findings of Sonsthagen *et al.* (2012) and Craig *et al.* (2016) by sequencing the control region of 115 western golden eagles to evaluate if the mitochondrial geneome supported either the proposed Migratory Flyway or Bird Conservation Region models. Although the results of their analyses did not provide support for either of the proposed management

scenarios, in support of the conclusions by Doyle et al. (2016) and Van Den Bussche et al. (2017) they did detect statistically significant genetic differentiation between their most northern samples from Alaska and Canada relative to their samples from northern California, southern Oregon, and southern Idaho. Judkins and Van Den Bussche (In Press) also found statistically significant genetic differentiation between western golden eagles from northern California, southern Oregon and southern Idaho when compared with birds from Montana and Wyoming. Thus, the mitochondrial results support many of the conclusions derived from the nuclear SNP data (Doyle et al., 201; Van Den Bussche et al. 2017).

Although Van Den Bussche *et al.* (2017) utililized a large number of SNPs, the SNPs examined were choosen from the longest scaffolds and subsequentyly filtered, thereby only representing 38.2% of the assembly with no knowledge of the location of the SNPs (coding or noncoding). Furthermore, in both the Doyle *et al.* (2016) and Van Den Bussche *et al.* (2017) studies, STRUCTURE was the only program used to determine the genetic structure of the population. It has been well documented that STRUCTURE has difficulty elucidating genetic structure when $F_{st}$ is low (Latch *et al.* 2006), when there is a pattern of isolation by distance (Pritchard *et al.* 2000; Schwartz & McKelvey 2009), or when sample sizes are small (Evanno *et al.* 2005). Therefore, when any of the previously mentioned situations are present, it is best to use additional programs that rely on different assumptions to ascertain genetic structure. Additionally, STRUCTURE does not allow for spatial input as compared to other Bayesian programs such as Geneland (Guillot *et al.* 2005) or Baps (Corander *et al.* 2003). Even though these recently published studies provided insight into the partitioning of genetic variation

within and among particular population segments of western golden eagles, they still leave information gaps that are critical for conservation efforts.

In this study, we describe the development of an Axiom myDesign custom SNP array for golden eagles that contains over 37,000 loci. These SNPs are distributed throughout the golden eagle genome and are representative of both intergenic and genic regions. We chose to develop this SNP chip because medium to high density genome-wide SNP arrays have been shown to be powerful tools to disentangle the relative roles of natural selection, genetic drift, and gene flow in observed patterns of genetic variation for other species (Luikart *et al.* 2003; Nielsen 2005; Stinchcombe & Hoekstra 2008; Malenfant *et al*. 2015). We therefore believed this would be a valuable approach for the generation of data relevant to the management and conservation of western golden eagles. We investigated the performance of this 37K SNP chip by performing preliminary population genomic assessments of 138 western golden eagles and compared our results to recently published studies of golden eagles that utilized SNPs (Doyle *et al.* 2016; Van Den Bussche *et al.* 2017), microsatellite loci (Craig *et al*. 2016), and mtDNA (Craig *et al.* 2016, Judkins and Van Den Bussche, In Press). We demonstrated that the custom SNP chip has high genotyping success, provides considerable power for assessing population genetic parameters, and identified several SNPs (and associated candidate genes) under selection. This study represents the first step in the application of a medium density SNP chip to the large-scale population genomics of an apex predator, the golden eagle.

## Materials and Methods

*SNP Chip Development*

SNP isolation and identification was conducted as described in Van Den Bussche *et al.* (2017) which provided a starting point of about 1.8 million SNPs for development of a golden eagle Axiom myDesign custom array. To screen for SNPs suitable for chip development we utilized several options in the PLINK software package (Purcell *et al.* 2007). Filtering options included a minor allele frequency (MAF) of 0.05, a minimal genotype frequency of 0.3, and a Hardy-Weinberg equilibrium (HWE) of 0.001. Loci passing this initial filtering were further thinned by examining all SNPs within a 10KB sliding window and retaining only a single SNP in each 10 KB window using operations in VCFTools (Danecek *et al.* 2011). All SNPs remaining after these filtering steps, as well as 160 nuclear SNPs, including a SNP for gender determination, originally identified by Doyle et al. (2016) were sent to Affymetrix (Santa Clara, CA) for additional quality assessment to ensure that they were appropriate for the Axiom myDesign custom array platform. All SNPs remaining after our filtering and filtering by Affymetrix were compared to the annotated golden eagle genome (http://www.ncbi.nlm.nih.gov/nuccore/NW_011950869.1) to determine location in the genome and severity of the SNP using the program SNPEff (Cingolani *et al.* 2012). We then chose SNPs that were in genes that other studies found to be ecologically relevant (Van Bers *et al.* 2012; Malenfant *et al.* 2015), SNPs that were upstream and downstream of genes, and SNPs that were in intergenic regions and not necessarily closely associated with any genes for population genetic analyses.

To evaluate the accuracy of the gender determination SNP identified by Doyle *et al.* (2016), we performed standard PCR-based molecular sexing of 57 (41%) individuals that were genotyped using the chip. For the PCR-based gender determination, we followed the protocol of Ito *et al.* (2003) using a 30 µl PCR containing 1 unit of *Taq* DNA polymerase, 0.16 mM of each deoxynucleoside triphosphate, 6 µl of 10x buffer, 2 mM of $MgCl_2$, 10.0 mg/ml of bovine serum albumin, 1.0 µL of 2 µM NP primer 0.5 µL of 2 µM MP primer, 0.5 µL of 2 µM P2 primer, and 1 µL of DNA. The thermal profile for the sexing PCR began with a 1:00 minute denaturation at 93°C, followed by 35 cycles of 93°C for 0:10, 52°C for 0:35, and 68°C 0:30, and a final elongation at 72°C for 7 minutes. Final PCR products were run on a 3% agarose gel to determine if one (male) or two (female) bands were present. If the reaction did not amplify using the primers MP and P2, primers developed by Banhos *et al.* (2008) were substituted using the above conditions. The substituted primers, CHD1Wr and CHD1Zr, were developed to amplify a 100bp shorter region of the CHD gene in order to cope with degraded DNA.

To determine genotyping error of our SNP chip, 15 arbitrarily chosen individuals, representing 11% of the individuals genotyped, were genotyped twice. The genotypes of the duplicate runs were compared to identify discrepancies between runs. We evaluated the following two types of potential genotyping error: the error rate for a no call at a SNP in one run versus a base call for the same SNP in the duplicate run and errors in which different bases were called for the same SNP in independent runs. These two error rates were calculated independently and then combined for an overall error rate.

*Sample Selection*

Blood samples from 138 western golden eagles (Fig 1, S1) were obtained by working with U.S. Fish and Wildlife biologists and licensed rehabilitators. If we received samples of multiple hatchlings/fledglings from the same nest, only one individual was used in all analyses to avoid violating any program assumptions due to highly related individuals. For this study, we included 124 hatchling/fledgling western golden eagles with known natal locations from Arizona (n = 3), California (n = 8), Colorado (n = 26), Idaho (n = 5), Montana (n = 4), Nebraska (n = 10), New Mexico (n = 19), Oklahoma (n = 1), Oregon (n = 28), Texas (n = 4), Utah (n = 3), and Wyoming (n = 13). The remaining 14 western golden eagles included an adult from Colorado collected during the non-breeding season, an adult from Wyoming collected during the non-breeding season, and 12 juvenile individuals (Alaska=7, British Columbia=3) that were caught during their southern migration into the United States. Subsequent isotope analyses were performed to determine their natal location (Domenech *et al. 2015*). The last two western golden eagles were fitted with GPS transmitters to determine the summer nesting location. Studies have determined that the typical median natal dispersal of golden eagles is 46.4 km (Millsap *et al*. 2014). Therefore, assuming the current nesting locations provides a close approximation to an individual's natal location, these individuals represent Alaska (n = 2). Blood aliquots (0.5 ml) from all 138 individuals were stored in lysis buffer and sent to our lab at Oklahoma State University where DNA was extracted using the protocol described by Longmire *et al.* (1997). DNA quality was assessed by running an aliquot on a 1% agarose gel and quantified using a NanoDrop 3300 spectrophotometer (Thermo Scientific).

The Axiom myDesign custom array we designed was used to genotype 138 western golden eagles at Eurofins (River Falls, WI). After receiving the data from Eurofins, we scored all individuals using options in the Axiom Analysis Suite v2.0.0.35. All scored SNPs were filtered to remove any loci that were poorly clustered, had a minor allele frequency less than 0.01, or were monomorphic.

*Population Structure Assessment*

To determine population structure we used three approaches, STRUCTURE v2.3.4 (Pritchard *et al.* 2000), Geneland (Guillot *et al.* 2005) and a Principle Coordinates Analysis (PCoA). Although both STRUCTURE and Geneland utilize a Bayesian algorithm, due to different algorithms and assumptions, results from each program and each run within a single program can vary (Hobbs *et al.* 2011). Furthermore, Geneland allows for the incorporation of specific geographic coordinates of sampled individuals in the analysis *a priori* to aid in determining population structure.

The computer program STRUCTURE v2.3.4 was used to probabilistically cluster all individuals based on their multilocus genotypes without prior locality information. The Bayesian clustering approach calculates the posterior probabilities of K clusters while minimizing linkage disequilibrium and maximizing clusters of individuals meeting the assumptions of HWE. STRUCTURE was run using a model that allowed for admixture and correlated allele frequencies for 10,000 burn-in iterations followed by 50,000 Markov Chain Monte Carlo (MCMC) iterations. Clusters (K) ranging from 1 to 10 were analyzed using eight independent runs per cluster scenario. Structure Harvester (Earl & VonHoldt 2012) was subsequently used to evaluate the output following the Evanno method (Evanno *et al.*, 2005) to determine the most appropriate K. Clumpp

(Jakobsson & Rosenberg 2007) was then used to determine the assignment probability of each individual to each cluster with the final results being visualized with the program Destruct (Rosenberg 2004).

Geneland was run in the program R 3.3.2. Due to the large number of SNPs analyzed, we created 10 datasets of 1,000 non-overlapping SNPs and each dataset was analyzed three times for a total of 30 runs. Each run was conducted using the correlated frequency model with geographical coordinates (latitude and longitude) of nest as a location prior. The model was allowed to evaluate a maximum of eight populations with the number of iterations=100,000. Once K and the individual's cluster membership was determined for each run, the results of the 30 runs were used to select the most probable K.

Finally, we performed a principal components analysis (PCoA) to cluster individuals based on genetic distances. The PCoA was run in Adegenet (Jombart 2008) using program R 3.3.2. PCoA is a distance-based model which utilizes the standardized covariance matrix of population pairwise genetic distance. The PCoA output was compared to the STRUCTURE and Geneland results to aid in determining the most appropriate number of clusters using a non-Bayesian approach.

*Relatedness and Population Genetics*

Once we determined the most appropriate number of genetic clusters, Genepop 4.4.3 (Raymond & Rousset 1995) was used to ensure that SNPs were in HWE within each cluster and tested for significance using a sequential Bonferroni correction. Arlequin v3.5 was used to determine pairwise population $F_{st}$ values using a p-value < 0.05 to

indicate significance. The kinship option in KING (Manichaikul *et al.* 2010) was used to determine relatedness among all individuals. Finally, the effective population size was calculated using the linkage disequilibrium option in NeEstimator v2.01 (Do *et al.* 2014) using a critical value of 0.05 and a monogamous mating system

*Outlier Detection*

Three programs were used to scan for outlier loci: Lositan (Antao *et al.* 2008), Arlequin v3.5, and BayeScan v2.1 (Foll & Gaggiotti 2008). As there are issues pertaining to which loci are under selection and which are false positives (Narum & Hess 2011), we used a conservative approach and only considered a SNP to be an outlier if it was found to be an outlier in all three methods.

FDIST2 (Beaumont & Nichols 1996) as implemented in Lositan was run under the infinite alleles model using 100,000 simulations, a neutral and forced mean $F_{st}$, a 0.99 CI criterion, and a FDR cut-off of 0.1. As Lositan is known to produce different results in different runs (Blanco-Bercial & Bucklin 2016), these parameters were run three times and loci that were identified in all three analyses were considered to be under selection. We also tested for loci potentially under selection using the program Arlequin 3.5.2.2 using 50,000 simulations and 1,000 demes. SNPs were determined to be under selection if they had a p-value < 0.01. Finally, we ran BayeScan for 100,000 iterations using 20 pilot runs and an additional 50,000 iterations for burn-in. We used a q-value of 0.1 as our cut off value for determining SNPs under selection.

SNPs that were found to be under selection in all three programs were then analyzed in SNPEff and associated genes determined using the reference golden eagle

genome and NCBI databases. Gene functions and annotations were determined using Gene Ontogeny (GO) in Panther utilizing the chicken (*Gallus gallus*) genome (Thomas *et al.* 2003; Mi *et al.* 2013) and a subsequent literature review to determine specific gene function in Aves.

## Results

### *SNP Chip Development, Sample Scoring, and Error Rate*

After all quality controls and filtering were completed, the Axiom myDesign custom array consisted of 37,562 SNPs. Of these, 157 nuclear SNPs were obtained from Doyle *et al.* (2016), 4,719 SNPs were in genic regions, and 32,686 SNPs were in intergenic regions of the genome. The 138 western golden eagle blood samples (Fig 1,S1) were sent to Eurofins for genotyping using this custom array. DNA from one of these samples (a sample from Wyoming) did not pass quality control after being scored. This individual was removed from subsequent analyses resulting in 137 western golden eagles (69 females and 66 males) for all population analyses. Analysis of loci from these 137 individuals revealed 142 monomorphic SNPs, 220 SNPs with a MAF lower than 0.01, and 628 SNPs that were poorly clustered leaving 36,560 SNPs for downstream analyses.

Error rates for the gender determination SNP were determined to be very low; we were able to successfully PCR amplify and determine gender on all 57 individuals. When gender was determined based on the SNP identified by Doyle *et al*. (2016), the gender of 55 individuals was identified using the SNP chip, while two of the samples were recorded as no calls for the sexing SNP on the chip. Of the 55 samples that were

analyzed in both methods, no discrepancies were determined between the two methods resulting in 100% concordance (27 female, 28 male).

Genotyping error for the overall array was assessed using 15 arbitrarily chosen DNA samples (11%) that were genotyped twice. Errors where a base was called for a SNP in one run and the same SNP had a non-distinguishable base in the second run occurred at an average rate of 0.586% and a median rate of 0.456%. Errors in which a different base was called for a specific SNP in each run occurred at an average rate of 0.229% and a median rate of 0.192%. Thus, based on our sample of 137 western golden eagles, our mean overall error rate was 0.82% and the median error rate was 0.644%.

*Population Structure*

The genetic structuring programs determined different optimum number of clusters (K). The three clusters (Fig 2) determined from STRUCTURE represent a clear unit that separates the Alaska and British Columbia golden eagles and a unit representing northern California, southern Oregon and southern Idaho. The third cluster contains golden eagles from the remaining eight states and shows a pattern more similar to a cline of genetic variation as opposed to a discrete genetic unit. Geneland was unable to determine a single optimum for the best K with the 30 runs consisting of two runs representing three clusters, 12 runs representing four clusters, 12 runs representing five clusters, and four runs representing six clusters. The major difference between the four and five-unit cluster was that although the Colorado samples were separated among two different clusters in both analyses, the four-unit cluster placed the majority of the Colorado samples with Wyoming, Nebraska, and Montana, whereas the five-unit cluster

placed many of these Colorado samples in their own cluster (Fig 3A and 3B). Furthermore, the four-unit cluster was unable to assign one individual into any of the clusters and the five-unit cluster was unable to assign three individuals into any of the clusters. Although these Geneland results differed from the STRUCTURE results, they were not discordant with STRUCTURE, rather they provided greater resolution than STRUCTURE provided. The PCoA determined that four-unit cluster was the most probable (Fig 4) with the results closely resembling the Geneland results for the four-unit cluster. The two adult birds with unknown natal locations from Wyoming and Colorado are displayed with an asterisk.

*Population Genetics and Relatedness*

When considering either the four or five-unit clusters determined by Geneland, none of the loci were out of HWE. Also, pairwise $F_{st}$ values for both the four and five-unit clusters revealed statistically significant levels of genetic differentiation for all pairwise population comparisons, providing support for both models (Table 1). When some of the Colorado samples were separated from the WY/NE/MT to make the five-unit cluster, the new Colorado group possessed low, but statistically significant genetic differentiation ($F_{ST}$ = 0.012) when compared with its original grouping. However, the Colorado group shows very low, but statistically significant, genetic differentiation ($F_{ST}$ = 0.009) when compared to the larger CO/UT/TX/NM/AZ/OK/WY that contains the other Colorado individuals (Table 1).

Among the 9,316 pairwise kinship relationship values for these 137 western golden eagles, 98% (9,173) of the comparisons revealed unrelated individuals. Of the 143 comparisons that revealed some level of kinship, 10 (0.1%) were equal to siblings, 1

was equivalent to half-sibs, 26 (2.8%) were first cousins, 106 (1.14%) approximated second cousins (Table 3). Of the samples that showed a relationship equal to or greater than second cousins, 62 pairings were from different states and 81 were from the same state (Table 3). For the pairs that were from different state, 46 were pairings in which the multiple states were from the same genetic cluster as defined by the four-unit cluster Geneland analysis. Nine of the pairings showed Oregon individuals that had first or second cousin relationships with individuals from the CO/UT/TX/NM/AZ/OK/WY or the WY/NE/MT/CO cluster. Finally, seven of these pairings consisted of individuals from the two eastern most clusters.

The effective population size was calculated for the four-unit cluster. AK/BC had the largest effective population size while the WY/NE/MT/CO cluster had the smallest. The effective population sizes are as follows: WY/NE/MT/CO was 146 ±1, AK/BC 1643 ± 113, CA/ID/OR 268 ± 2, and WY/NE/CO/MO 146 ±1 (Table 2).

*SNPs Under Selection*

Each of the programs used for identifying SNPs under putative selection, Lositan, Arlequin and Bayescan, was run using the four-unit cluster as determined by Geneland. Lositan revealed 838 SNPs potentially under selection for the four-unit cluster ($F_{st}$ 0.058-0.28). Using options in Arlequin, 988 ($F_{st}$ 0.091218-0.291) were found. Finally, Bayescan revealed 111 SNPs for four clusters ($F_{st}$ 0.059585-0.13845) as potentially being under selection. Taking a conservative approach and only considering loci that were suggestive of being under selection in all three analyses, we were left with 85 SNPs.

Of the 85 SNPs considered to be under selection there was one synonymous variant, two missense variant, three upstream gene variants, one intron variant, one 5' UTR variant, one 3' UTR variant, and 75 located in intergenic regions. These 85 SNPs were associated with 120 genes and regions. In Panther, the molecular function and biological process were analyzed with the complete results in Supplementary Table 2 and the GO Slim, a broader overview of the ontology, in Figure 5.

## Discussion

The purpose of this study was to create an Axiom myDesign custom array to further investigate the use of SNPs in determining genetically based conservation units for golden eagles. By utilizing a large SNP dataset we were able to incorporate modern genomic approaches at the whole genome level. Results indicate the design was successful as 97% of the selected SNPs and 99% of individuals were successfully scored. Moreover, the error rate for calling genotypes was low (mean error rate = 0.82%; median error rate = 0.644%). Utilizing the SNPs included in the array, we detected a finer scale genetic structure across the western United States than previously described in other studies (Doyle *et al.* 2016, Van Den Bussche *et al.* 2017). A dataset of SNPs under putative selection was also identified.

### *Population Structure*

Previous studies on western golden eagles (Doyle *et al.* 2016, Van Den Bussche *et al.* 2017) using only STRUCTURE indicated that there were three putative clusters for western golden eagles in the western United States which closely resembled our

STRUCTURE results in which the Alaska/Canadian samples grouped in a cluster, the Idaho, Oregon, California areas grouped as a cluster, and all other samples grouped together with considerable genomic admixture. To provide clarity for the number of clusters and to evaluate the high degree of admixture in golden eagles from the central Rocky Mountain and Plains states, we performed additional clustering analyses and standard population genetic statistics. These analyses revealed additional structuring in this portion of the western golden eagle range and importantly, these additional clusters did not contradict the STRUCTURE results, rather they provided greater resolution with regards to population structuring.

The U.S. Fish and Wildlife Service is currently in a decision-making process for determining how to manage golden eagles utilizing an adaptive management approach (Katzner *et al.* 2013). This approach is based on the best scientific data available but also allows for changes based on newer data, if it is relevant. To date, in addition to this study, there are three other studies that utilize nuclear loci for the management of western golden eagles (Craig et al. 2016; Doyle *et al*. 2016, Van Den Bussche et *al.* 2017). Although none of these studies drew specific management conclusions, when the data from this study are interpreted in light of these previous studies, some patterns emerge.

All of the analyses we performed supported our samples from Alaska and British Columbia as representing a discrete genetic entity (Figs. 2-4; Table 1). Whereas it is important to note that these Alaskan and British Columbia individuals were assigned to these areas by isotope data or GPS tracking data as opposed to having a known natal location, our analyses supports their genetic uniqueness as compared to golden eagles found in the contiguous United States. This northern grouping, albeit with some

differences in geographical representation (Doyle *et al.* [2016] did not examine individuals from Canada), was also identified in the studies by Doyle *et al*. (2016) and Van Den Bussche *et al.* (2017). The genetic distinction of this group is also supported through the analysis of about 170 western golden eagles at the mitochondrial control region (Judkins and Van Den Bussche, In Press). The adult eagle that was sampled in Wyoming appears to fit more closely with the British Columbia samples in both the STRUCTURE and the PCoA analyses (Figs. 2 and 4, respectively). As this bird was caught in Wyoming in December, it was probably a migratory individual that had a natal location farther north. When this individual was removed from the Wyoming sample set in the PCoA (not shown), the ellipse for the Wyoming sample narrowed along the Y axis making a more precise ellipse for the remaining Wyoming samples.

The second group we detected in all of our analyses consists of individuals from northern California, southern Oregon, and southern Idaho (Figs. 2 – 4; Table 1). In addition to being a well-supported group in our study, this grouping of golden eagles receives support from multiple studies, albeit with some caveats. The genetic break in California is unclear based on these studies. Craig *et al.* (2016) examined individuals from central California, southern Oregon and southern Idaho and detected no significant genetic differentiation among these samples. Van Den Bussche *et al.* (2017) examined individuals from northern California, southern Oregon and southern Idaho and these individuals formed a well-defined genetic entity, as was also found in this study. How much of the California range of golden eagles to include in this group is uncertain. Doyle *et al.* (2016) examined golden eagles throughout most of the north-south length of California and found strong support for those individuals belonging to a single genetic

unit. However, they did not include birds from the very northern portion of California as was included in the study by Van Den Bussche *et al.* (2017) and this study nor did they include any individuals from southern Oregon or southern Idaho. Clearly, to better refine the number of discrete genetic entities in this portion of the golden eagle range, we need to sample individuals from throughout the California distribution of nesting golden eagles.

Whereas these genetic entities appear to be well supported based on the available data, the number of discrete genetic entities in the remainder of the range of western golden eagles is uncertain. We included representation of golden eagles from the western states of Montana, Wyoming, Nebraska, Colorado, Utah, Oklahoma, Texas, New Mexico, and Arizona. STRUCTURE revealed that these individuals grouped together, but they clearly represent considerable admixture of genomes (Fig. 2). This same pattern was detected by Doyle *et al.* (2016) and Van Den Bussche *et al.* (2017). However, analyzing these data with Geneland and conducting a PCoA, provides further insight. The PCoA partitions these individuals into two groups with individuals from Montana, Colorado, Wyoming, and Nebraska forming one group and individuals from Utah, Colorado, Oklahoma, New Mexico, Texas, and Arizona forming a second group (Fig 4). This same four-unit cluster was also revealed as an acceptable possibility based on our analysis of these data using Geneland (Fig 3). Examination of the PCoA reveals that our samples of Colorado golden eagles possess the largest spread across both the X and Y-axis and suggests that additional sampling of individuals throughout the range of western golden eagles will be necessary to refine these units (Fig. 4). Alternatively, it may be that for some unknown reason, individuals from the California, Idaho, Oregon cluster as well

as individuals from the Montana, Wyoming, Nebraska cluster have a higher propensity for dispersal from their natal area to Colorado than many other birds. Furthermore, the adult golden eagle that was captured in Colorado clustered better with the Idaho, Oregon, California cluster. This individual could represent a migratory individual as the sample was collected in December. It could also be a dispersing individual thereby adding further support to some of the higher kinship values calculated in KING between the Idaho, California, Oregon cluster and the eagles in the eastern clusters. When this sample was removed from the PCoA (not shown), the ellipse for the Colorado individuals narrowed along the X-axis, helping to remove some of the broadness in the Colorado cluster.

The effective population size was calculated for the four-unit cluster (Table 2). The estimated effective population sizes are large enough that inbreeding should not be a major problem at this time. The effective population size did not greatly increase or decrease when taking the sample size or the geographic size of each unit into account. Rather, effective population size seems to correspond to historical persecution pressure with the highest effective population sizes being in Alaska and Canada and lowest in the more agriculture areas of the United States where extirpation programs were not uncommon (Watson 2010)

*SNPs Under Selection*

Given that there is slightly more support for a four-unit cluster model (Geneland, PCoA, and pairwise differentiation) coupled with our lack of sufficient sampling across the range of western golden eagles, we took a conservative approach for determining SNPs under putative selection. Interestingly, even though we used SNPs from Doyle *et*

*al.* (2016), only one of their proposed SNPs, BMP4, was found to be under selection in our final dataset. The difference between our findings and those of Doyle *et al.* (2016) is likely due to the stringent requirements we invoked to consider a SNP to be under selection.

SNPs that were deemed to be under selection served a role in a variety of biological pathways and molecular function (Fig. 5). Within chickens, BMP4 and CDKN3 were found to have roles in beak formation (Zhan *et al.* 2013). SLC8A1 and FSHR both were found in chickens to have a role in egg shell and egg formation respectively (Jonchère *et al.* 2012; Xu *et al.* 2017). In chickens, FSHR plays a role in lipid biosynthesis around the abdominal area. Other pathways included inner ear formation (BMP4) (Gerlach *et al.* 2000), cardiac muscle (MEF2C) (Takebayashi-Suzuki *et al.* 2001), and various gland development (BMP4, ISL1) (Proszkowiec-Weglarz *et al.* 2011; Neves *et al.* 2012). As this study only aimed to identify a stringently selected group of SNPs under selection, additional studies determining the effects of the SNPs in the pathways are definitely warranted.

*Implications for Future Studies*

While more sampling needs to be conducted, none of the cluster results clearly add support to either of the currently proposed models (Migratory Flyway or Bird Conservation Region) for determining management units despite the fact these regions were suggested based on the documented movements of golden eagles. The Migratory Flyway model was proposed to reflect seasonal movement of golden eagles as determined by band returns during their migration whereas the Bird Conservation

Regions model was proposed because of nest proximity and dispersal (U.S. FWS 2016). As none of the genetic data fit these models, the genetic data may reflect more gene flow than currently recognized between western golden eagles. The KING analysis revealed 14 pairings of individuals that were related at a second cousin or greater level (>0.01) that were separated across geographic space and were from different genetic clusters. Studies have shown that it can take as few as one to as high as 10 immigrants into a population to minimize the loss of polymorphism and heterozygosity while maintaining differing allele frequencies between subpopulation (Mills & Allendorf 1996). As our samples size represents less than 1% of the estimated population of golden eagle in the United States and we are already observing related individuals across large geographic distances, there may be much more gene flow between western golden eagles populations than previously thought. By adding more geographic spread in the future, more a more accurate representation of the gene flow can be assessed.

By utilizing more markers and more thorough analyses, genetically based management units for western golden eagles are becoming clearer. While the samples utilized in this study only provide a preliminary insight into the development of these management units, the results strongly suggest that there is more substructure to the population than previously thought. Prior to more research being conducted, caution is warranted when managing utilizing Migratory Flyways and Bird Conservation Units as these proposed units may not truly reflect genetically based units.

**Acknowledgements**

We would like to thank Bryan Bedrosian, Rob Domenech, Garth Herring, Gail Kratz, Bill Murphy, Victor Roubidoux, and Brian Smith for contributing golden eagle blood samples from their research and rehabilitation efforts. We would also like to thank the Iowa Tribe of Oklahoma, Shakopee Mdewakanton Sioux, Epply Foundation, and United States Fish and Wildlife Service for providing critical funds to this study as well as the Ford Foundation for funding Megan Judkins' research assistantship while the project was being conducted. Finally, all of the other members of Megan's dissertation committee, Drs. Meredith Hamilton, Andrew Doust, and Jim Lish are sincerely thanked for their efforts, guidance and comments on previous versions of this manuscript.

**Literature Cited**

Allendorf FW, Hohenlohe PA, Luikart G (2010) Genomics and the future of conservation genetics. *Nature reviews. Genetics*, 11, 697–709.

Antao T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G (2008) LOSITAN: A workbench to detect molecular adaptation based on a Fst-outlier method. *BMC Bioinformatics*, 9, 323.

Banhos A, Hrbek T, Gravena W, Sanaiotti T, Farias IP (2008) Genomic resources for the conservation and management of the harpy eagle (Harpia harpyja, Falconiformes, Accipitridae). *Genetics and Molecular Biology*, 31, 146–154.

Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society B: Biological Sciences*, 1619–1626.

Benestan L, Gosselin T, Perrier C, Sainte-Marie B, Rochette R, Bernatchex L (2015) RAD genotyping reveals fine-scale genetic structuring and provides powerful population assignment in a widely distributed marine species, the American lobster (Homarus americanus). *Molecular Ecology*, 24, 3299–315.

Blanco-Bercial L, Bucklin A (2016) New view of population genetics of zooplankton: RAD-seq analysis reveals population structure of the North Atlantic planktonic copepod *Centropages typicus*. *Molecular Ecology*, 25, 1566-1580.

Bourret V, Kent MP, Primmer CR, Vasemagi A, Karlsson S, Hindar K, McGinnity P, Verspoor E, Bernatchez L, Lien S (2013) SNP-array reveals genome-wide patterns of geographical and potential adaptive divergence across the natural range of Atlantic salmon (*Salmo salar*). *Molecular Ecology*, 22, 532–551.

Cammen KM, Schultz TF, Rosel PE, Wells RS, Read AJ (2015) Genomewide investigation of adaptation to harmful algal blooms in common bottlenose dolphins ( *Tursiops truncatus* ). *Molecular Ecology*, 24, 4697–4710.

Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Ruden DM, Lu X (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118 ; iso-2; iso-3. *Fly*, 6, 1–13.

Corander J, Waldmann P, Sillanpää MJ (2003) Bayesian analysis of genetic differentiation between populations. *Genetics*, 163, 367–374.

Craig EH, Adams JR, Waits LP, Fuller MR, Whittington DM (2016) Nuclear and mitochondrial DNA analyses of golden eagles (Aquila chrysaetos canadensis) from three areas in western North America; initial results and conservation implications. *PLoS ONE*, 11, 1–15.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R, 1000 Genomes Project Analysis Group (2011) The variant call format and VCFtools. *Bioinformatics*, 27, 2156–2158.

Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J. & Ovenden, J. R. (2014). NeEstimator V2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Molecular Ecology Resources*, 14, 209-214.

Domenech R, Pitzkathy T, Smith G (2015) Estimating natal origins of migratory juvenile golden eagles using stable hydrogen isotopes. *Journal of Raptor Research* , 49, 308–

315.

Doyle JM, Katzner TE, Roemer GW, Cain III JW, Millsap BA, McIntyre CL, Sonsthagen SA, Fernandez NB, Wheeler M, Bulut Z, Bloom PH, DeWoody JA (2016) Genetic structure and viability selection in the golden eagle (Aquila chrysaetos), a vagile raptor with a Holarctic distribution. *Conservation Genetics*, 17, 1307-1322.

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4, 359–361.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, 14, 2611–2620.

Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, 180, 977–993.

Gerlach LM, Hutson MR, Germiller JA, Nguyen-Luu D, Victor JC, Barald KF (2000) Addition of the BMP4 antagonist, noggin, disrupts avian inner ear development. *Development*, 127, 45–54.

Guillot G, Mortier F, Estoup A (2005) GENELAND: A computer package for landscape genetics. *Molecular Ecology Notes*, 5, 712–715.

Hess JE, Campbell NR, Close DA, Docker MF, Narum SR (2013) Population genomics of Pacific lamprey: Adaptive variation in a highly dispersive species. *Molecular Ecology*, 22, 2898–2916.

Hobbs GI, Chadwick EA, Bruford MW, Slater FM (2011) Bayesian clustering techniques

and progressive partitioning to identify population structuring within a recovering otter population in the UK. *Journal of Applied Ecology*, 48, 1206–1217.

Ito H, Sudo-Yamaji A, Abe M, Murase T, Tsubota T (2003) Sex identification by alternative polymerase chain reaction methods in Falconiformes. *Zoological Science*, 20, 339–344.

Jakobsson M, Rosenberg NA (2007) CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, 23, 1801–1806.

Jombart T (2008) Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics,* 24, 1403–1405.

Jonchère V, Brionne A, Gautron J, Nys Y (2012) Identification of uterine ion transporters for mineralisation precursors of the avian eggshell, *BMC Physiology*, 12:1-17.

Judkins ME, Van Den Bussche RA (In Press) Holarctic phylogeography of golden eagles (Aquila chrysaetos) and evaluation of alternative North American management approaches. *Biological Journal of the Linnean Society*.

Katzner T, Smith BW, Miller TA, Brandes D, Cooper J, Lanzone M, Brauning D, Farmer C, Harding S, Kramer DE, Koppie C, Maisonneuve C, Martell M, Mojica EK, Todd C, Tremblay JA, Wheeler M, Brinker DF, CHubbs TE, Gubler, O'malley K, Mehus S, Poerter B, Brooke RP, Watts BD, Bildstein KL (2013) Status, biology, and conservation priorities for North America's eastern golden eagle (*Aquila chrysaetos*) population. *The American Ornithologists' Union*, 129 (1):169-176.

Kochert MN, Steenhof K, Kociiert M (2002) Golden Eagles in the US and Canada: Status, trends, and conservation challenges. *Journal of Raptor Research*, 36, 32–40.

Latch EK, Dharmarajan G, Glaubitz JC, Rhodes OE (2006) Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. *Conservation Genetics*, 7, 295–302.

Longmire JL, Maltbie M, Baker RJ (1997) Use of "lysis buffer" in DNA isolation and its implication for museum collections. *Occasional Papers of the Museum of Texas Tech University*, 163, 1–3.

Lozier JD (2014) Revisiting comparisons of genetic diversity in stable and declining species : assessing genome-wide polymorphism in North American bumble bees using RAD sequencing. *Molecular Ecology*, 23, 788–801.

Luikart G, England PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews Genetics*, 4, 981–994.

Malenfant RM, Coltman DW, Davis CS (2015) Design of a 9K illumina BeadChip for polar bears ( *Ursus maritimus* ) from RAD and transcriptome sequencing. *Molecular Ecology Resources*, 15, 587–600.

Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM (2010) Robust relationship inference in genome-wide association studies. *Bioinformatics*, 26, 2867–2873.

Mi H, Muruganujan A, Thomas PD (2013) PANTHER in 2013: Modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Research*, 41, 377–386.

Mills LS, Allendorf FW (1996) The One-Migrant-per-Generation Rule in Conservation

and Management. *Conservation Biology*, 10, 1509–1518.

Millsap BA, Zimmerman GS, Sauer JR, Nielson RM, Otto M, Bjerre E, Murphy R (2013) Golden eagle population trends in the western United States: 1968-2010. *Journal of Wildlife Management*, 77, 1436–1448.

Millsap BA, Harmata AR, Stahlecker DW, Mikesic DG (2014) Natal Dispersal Distance of Bald and Golden Eagles Originating in the Coterminous United States as Inferred from Band Encounters. *Journal of Raptor Research,* 48, 13–23.

Morneau F, Tremblay JA, Todd C, Chubbs TE (2015) Known Breeding Distribution and Abundance of Golden Eagles in Eastern North America. *Northeastern Naturalist*, 22, 236–247.

Narum SR, Hess JE (2011) Comparison of FST outlier tests for SNP loci under selection. *Molecular Ecology Resources*, 11, 184–194.

Neves H, Dupin E, Parreira L, Le Douarin NM (2012) Modulation of Bmp4 signalling in the epithelial-mesenchymal interactions that take place in early thymus and parathyroid development in avian embryos. *Developmental Biology*, 361, 208–219.

Nielsen R (2005) Molecular Signatures of Natural Selection. *Annual Review of Genetics*, 39, 197–218.

Nosil P, Vines TH, Funk DJ (2005) Perspective: Reproductive isolation caused by natural selection against immigrants from divergent habitats. *Evolution: International Journal of Organic Evolution*, 59, 705–719.

Pilot M, Greco C, VonHoldt BM, Jedrzejewska B, Randi E, Jedrzejewski W, Sidorovich VE, Ostrander EA, WAyne RK (2014) Genome-wide signatures of population bottlenecks and diversifying selection in European wolves. *Heredity,* 112, 428–42.

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

Proszkowiec-Weglarz M, Higgins SE, Porter TE (2011) Changes in gene expression during pituitary morphogenesis and organogenesis in the chick embryo. *Endocrinology*, 152, 989–1000.

Purcell S, Neale B, Todd-brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, De Bakker PIW, Daly MJ, Sham PC (2007) PLINK : A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *American Journal of Human Genetics* , 81, 559–575.

Raymond M, Rousset F (1995) Genepop (Version-1.2) - Population-Genetics Software for Exact Tests and Ecumenicism. *Journal of Heredity*, 86, 248–249.

Rosenberg NA (2004) DISTRUCT: A program for the graphical display of population structure. *Molecular Ecology Notes*, 4, 137–138.

Schwartz MK, McKelvey KS (2009) Why sampling scheme matters: the effect of sampling scheme on landscape genetic results. *Conservation Genetics*, 10, 441–452.

Schweizer RM, Vonholdt BM, Harrigan R, Knowles JC, Musiani M, Coltman D, Novembre J, Wayne RK. (2015) Genetic subdivision and candidate genes under selection in North American gray wolves. *Molecular Ecology*, 25, 380-402.

Sonsthagen SA, Coonan TJ, Latta BC, Sage GK, Talbot SL (2012) Genetic diversity of a newly established population of golden eagles on the Channel Islands, California. *Biological Conservation*, 146, 116–122.

Stapley J, Reger J, Feulner PGD, Smadja C, Galindo J, Ekblom R, Bennison C, Ball AD, Beckerman AP, Slate, J (2010) Adaptation genomics: the next generation. *Trends in*

*Ecology & Evolution*, 25, 705–712.

Stinchcombe JR, Hoekstra HE (2008) Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity*, 100, 158–170.

Takebayashi-Suzuki K, Pauliks LB, Eltsefon Y, Mikawa T (2001) Purkinje fibers of the avian heart express a myogenic transcription factor program distinct from cardiac and skeletal muscle. *Developmental Biology*, 234, 390–401.

Thomas PD, Campbell MJ, Kejariwal A, Mi H, Karlak B (2003) PANTHER: A Library of Protein Families and Subfamilies Indexed by Function. *Genome Research*, 13, 2129–2141.

U.S. Fish and Wildlife Service. 2016. Bald and Golden Eagles: Population demographics and

estimation of sustainable take in the United States, 2016 update. Division of Migratory Bird Management, Washington D.C., USA.

Van Bers NEM, Santure AW, Van Oers K, Cauwer ID, Dibbits BW, Mateman C, CRooijmans, RPMA, Sheldon BC, Visser ME, GRoenen MAM, Slate J (2012) The design and cross-population application of a genome-wide SNP chip for the great tit Parus major. *Molecular Ecology Resources*, 12, 753–770.

Van Den Bussche RA, Judkins ME, Montague MJ, Warren WC (2017) A Resource of Genome-Wide Single Nucleotide Polymorphisms ( Snps ) for the Conservation and Management of Golden Eagles. *Journal of Raptor Research*, 51, 368–377.

Watson, J. 2010. The Golden Eagle. Second Edition. Yale University Press, New Haven.

Xu J, Gao X, Li X, Ye Q, Jebessa E, Abdalla BA, Nei Q (2017) Molecular characterization, expression profile of the FSHR gene and its association with egg production traits in muscovy duck. *Journal of Genetics*, 96, 341–352.

Zhan X, Pan S, Wang J, Dixon A, He J, Muller MG, Ni P, Liu Y, Hou H, Chen Y, Xia J, Luo Q, Xu P, Chen Y, Liao S, Cao, Gao, Wang Z, Yue Z, Li G, Yin Y, Fox NC, Wang J, Bruford MW (2013) Peregrine and saker falcon genome sequences provide insights into evolution of a predatory lifestyle. *Nature Genetics*, 45, 563–6.

**Figure Legends**

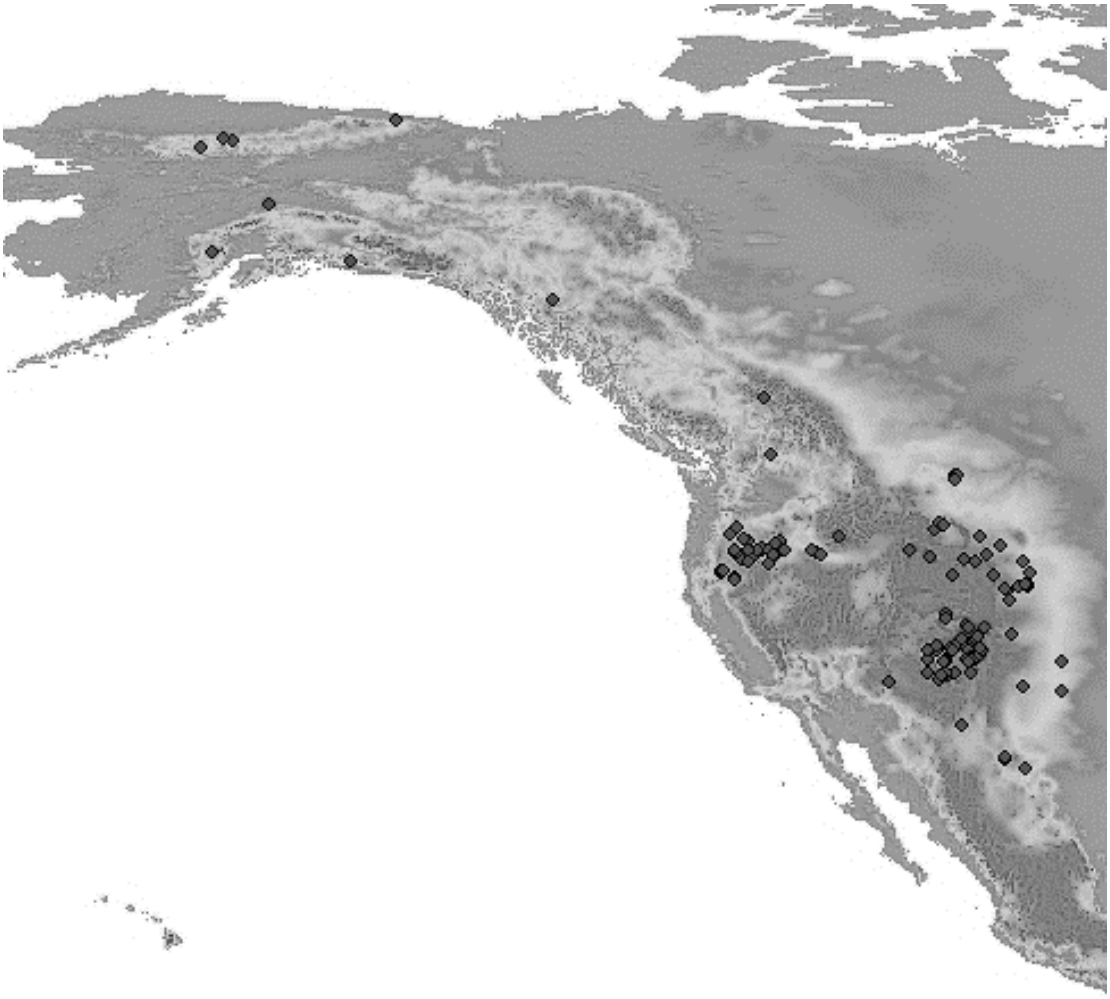
Figure 1 Distribution of collection locations of golden eagle samples analyzed for this study. Locations shows represent the actual hatch site of juveniles that were collected for the study (Arizona, California, Colorado, Idaho, Montana, Nebraska, New Mexico, Oklahoma, Oregon, Texas, Utah) or the approximate natal location using isotopes (Alaska, British Columbia).

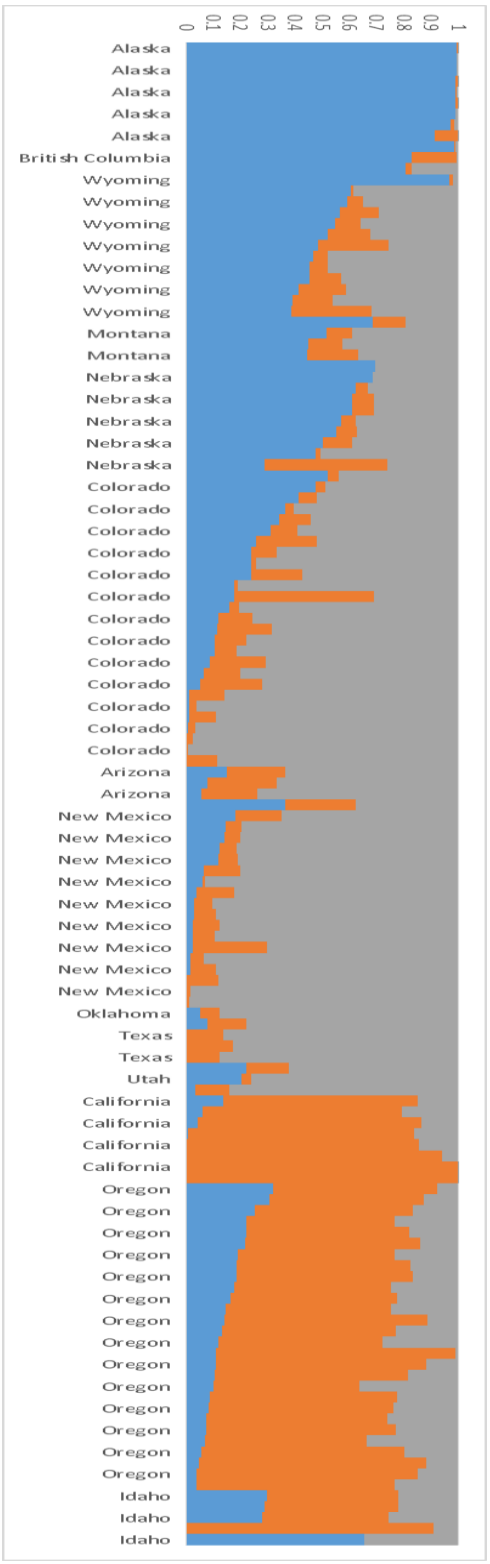
Figure 2 Results of STRUCTURE for 36,560 single nucleotide polymorphisms and 137 individuals using CLUMPP to average 8 runs of K equaling 3. Blue represents the Alaska-British cluster, orange represents the Idaho-California-Oregon cluster, and grey represents the Arizona-Colorado-Montana-Nebraska-New Mexico-Oklahoma-Texas-Utah cluster.

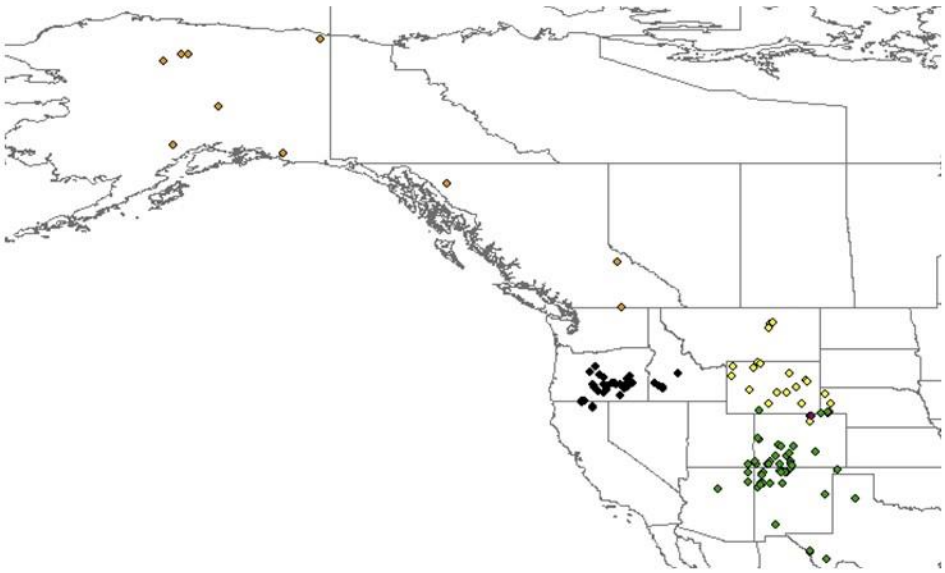
Figure 3 Distribution of K equals four and five from Geneland. A) The K equals 4 results are shown with yellow, black, green, and orange symbols representing the four unique clusters. Pink was used to represent samples that could not be placed within a cluster. B) The K equals 5 results are shown with yellow, black, orange, green, and blue symbols representing the five unique clusters. Pink was used to represent samples that could not be placed within a cluster.

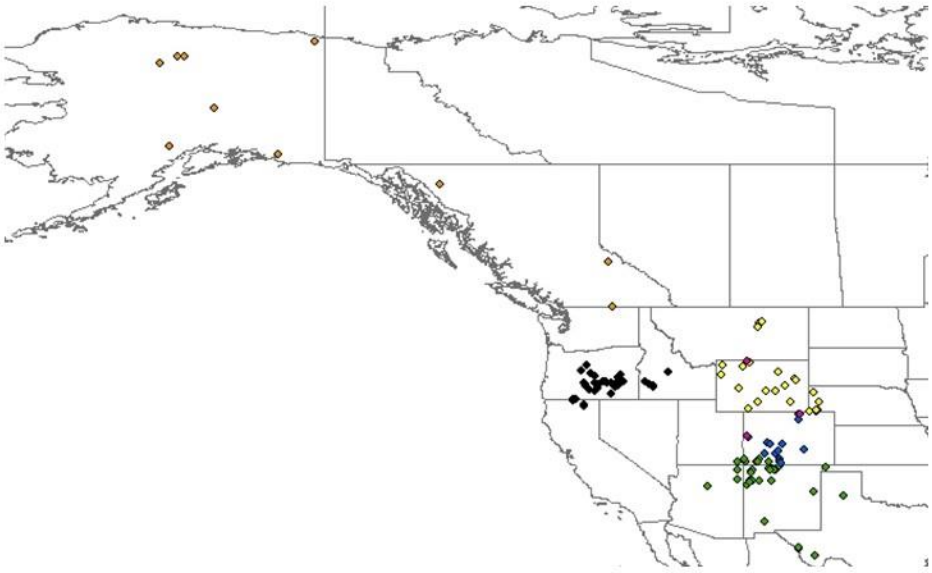
Figure 4 Results of the Principal coordinates analysis (PCoA) run using Adagent in program R using 36,560 SNPs with colors representing the state from which the sample

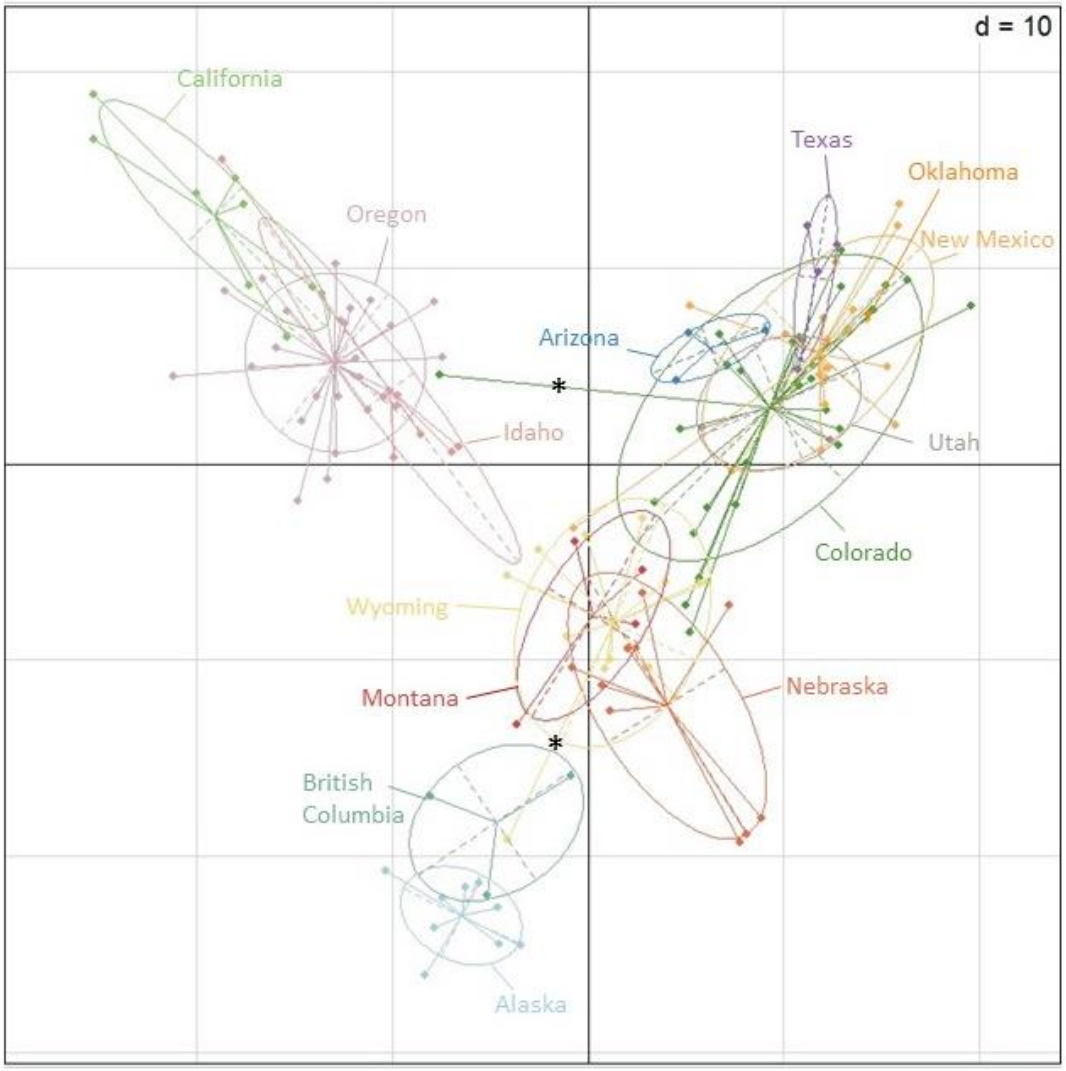
was obtained. The black stars represent adult birds with unknown natal locations that were placed in a cluster other than their sampled location.




0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Alaska
Alaska
Alaska
Alaska
Alaska
British Columbia
Wyoming
Wyoming
Wyoming
Wyoming
Wyoming
Wyoming
Wyoming
Montana
Montana
Nebraska
Nebraska
Nebraska
Nebraska
Nebraska
Colorado
Colorado
Colorado
Colorado
Colorado
Colorado
Colorado
Colorado
Colorado
Colorado
Colorado
Colorado
Colorado
Arizona
Arizona
New Mexico
New Mexico
New Mexico
New Mexico
New Mexico
New Mexico
New Mexico
New Mexico
New Mexico
Oklahoma
Texas
Texas
Utah
California
California
California
California
Oregon
Oregon
Oregon
Oregon
Oregon
Oregon
Oregon
Oregon
Oregon
Oregon
Oregon
Oregon
Oregon
Idaho
Idaho
Idaho






d = 10
California
Oregon
Texas
Oklahoma
New Mexico
Arizona
Idaho
Utah
Colorado
Wyoming
Montana
Nebraska
British
Columbia
Alaska

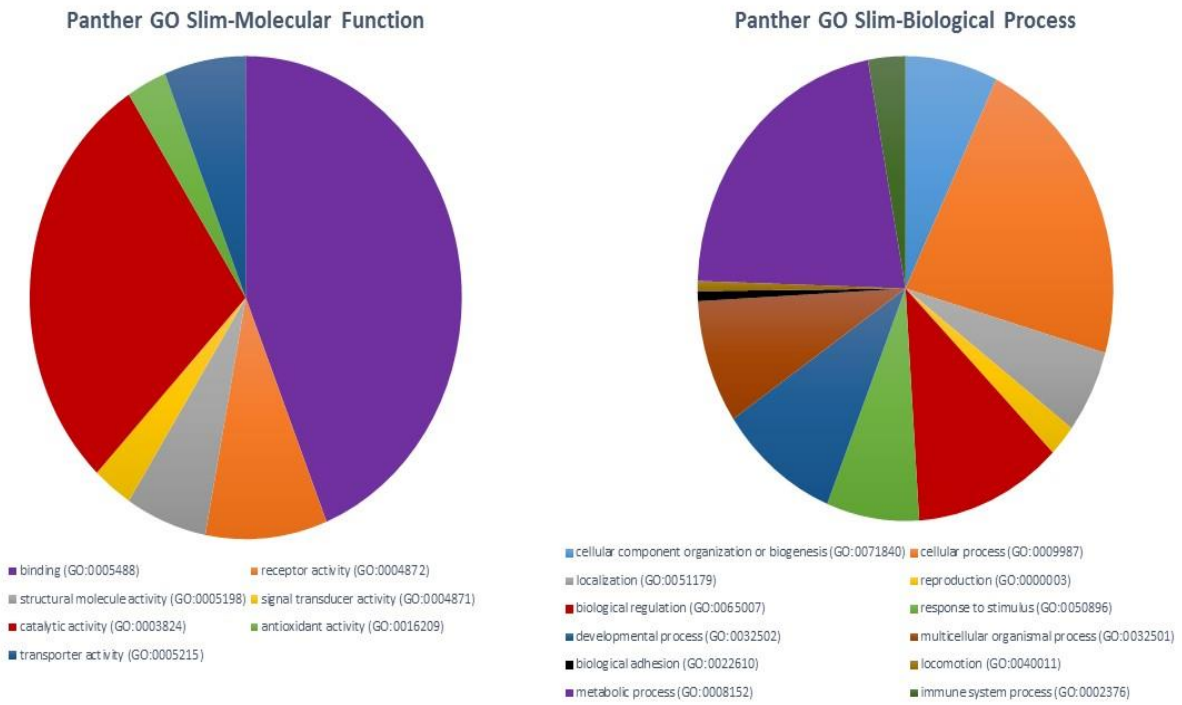

Panther GO Slim-Molecular Function
binding (GO:0005488)
receptor activity (GO:0004872)
structural molecule activity (GO:0005198)
signal transducer activity (GO:0004871)
catalytic activity (GO:0003824)
antioxidant activity (GO:0016209)
transporter activity (GO:0005215)
Panther GO Slim-Biological Process
cellular component organization or biogenesis (GO:0071840)
cellular process (GO:0009987)
localization (GO:0051179)
reproduction (GO:0000003)
biological regulation (GO:0065007)
response to stimulus (GO:0050896)
developmental process (GO:0032502)
multicellular organismal process (GO:0032501)
biological adhesion (GO:0022610)
locomotion (GO:0040011)
metabolic process (GO:0008152)
immune system process (GO:0002376)

Table 1 Pairwise $F_{st}$ values for K equals four and five clusters as reported in Geneland. Bold values indicate significant values of P > 0.05. Postal code abbreviations are used for the states and provinces.

| A. | UT/NE/CO/TX/NM/AZ/OK | AK/BC | CA/OR/ID | WY/NE/CO/MO |
|---|---|---|---|---|
| UT/NE/CO/TX/NM/AZ/OK | 0 | | | |
| AK/BC | **0.03068** | 0 | | |
| CA/OR/ID | **0.01754** | **0.03022** | 0 | |
| WY/NE/CO/MO | **0.0098** | **0.01983** | **0.01617** | 0 |

| B. | CO/UT/TX/NM/AZ/OK/WY | AK/BC | CA/OR/ID | WY/NE/MO | CO |
|---|---|---|---|---|---|
| CO/UT/TX/NM/AZ/OK/WY | 0 | | | | |
| AK/BC | **0.03424** | 0 | | | |
| CA/OR/ID | **0.02008** | **0.03022** | 0 | | |
| WY/NE/MO | **0.01491** | **0.02146** | **0.01777** | 0 | |
| CO | **0.00934** | **0.03307** | **0.02** | **0.01211** | 0 |

Table 2 Effective population size for the four-unit cluster using NeEstimator v 2.01. Calculations were based on a monogamous system using the linkage disequilibrium setting.

| Population | Sample Size | Ne | 95% CI | |
|---|---|---|---|---|
| CO/UT/TX/NM/AZ/OK/WY | 60 | 242.7 | 242.3 | 243.1 |
| AK/BC | 12 | 1643.7 | 1530.1 | 1775.4 |
| CA/ID/OR | 41 | 268.3 | 267.6 | 269 |
| WY/NE/CO/MO | 23 | 146.5 | 146 | 146.9 |

Table 3 KING results indicating the relationship values for all golden eagles used in the study. Same State Pairings indicate a pairing in which the samples were from the same state. Different State Pairings indicate a pairing in which samples were obtained from different states.

| Kinship Value | Same State Pairing | Different State Pairing |
|---|---|---|
| 0.18-0.3 (Siblings) | 9 | 1 |
| 0.09-0.179 (1/2 siblings) | 1 | 0 |
| 0.03-0.089 (1st cousins) | 19 | 7 |
| 0.01-0.029 (2nd cousins) | 52 | 54 |
| <0.099 | 1042 | 8131 |

Table S1 Band number, provider, and the state of collection for samples used in this study.

| Sample Number | Provider | State |
|---|---|---|
| 0799-00753 | Brian Smith | Wyoming |
| 709-03870 | Brian Smith | Nebraska |
| 799-00737 | Brian Smith | Colorado |
| 799-00745 | Brian Smith | Colorado |
| 0679-02601 | Brian Smith | California |
| 0679-02649 | Brian Smith | California |
| 0709-00754 | Brian Smith | Nebraska |
| 0709-02969 | Brian Smith | Nebraska |
| 0709-02973 | Brian Smith | Colorado |
| 0709-03858 | Brian Smith | Wyoming |
| 0709-03859 | Brian Smith | Colorado |
| 0709-03861 | Brian Smith | Nebraska |
| 0709-03865 | Brian Smith | Wyoming |
| 0709-03866 | Brian Smith | Wyoming |
| 0709-04382 | Brian Smith | Colorado |
| 0799-00096 | Brian Smith | Nebraska |
| 0799-00543 | Brian Smith | Wyoming |
| 0799-00731 | Brian Smith | Wyoming |
| 0799-00732 | Brian Smith | Wyoming |
| 0799-00733 | Brian Smith | Wyoming |
| 0799-00734 | Brian Smith | Wyoming |
| 0799-00735 | Brian Smith | Wyoming |
| 0799-00751 | Brian Smith | Colorado |
| 0799-00752 | Brian Smith | Wyoming |
| 0799-00755 | Brian Smith | Wyoming |
| 709-02983 | Brian Smith | Colorado |
| 0799-00656 | Brian Smith | Nebraska |
| 719-00025 | Bryan Bedrosian | Alaska |
| 719-00830 | Bryan Bedrosian | Alaska |
| 799-01003 | Bryan Bedrosian | Wyoming |
| 799-01004 | Bryan Bedrosian | Wyoming |
| 799-01005 | Bryan Bedrosian | Montana |
| 799-01006 | Bryan Bedrosian | Montana |
| 799-01008 | Bryan Bedrosian | Montana |
| 799-01009 | Bryan Bedrosian | Montana |
| 709-2976 | Dale Stahlecker | Colorado |
| 709-3879 | Dale Stahlecker | Colorado |

| | | |
|---|---|---|
| 709-3880 | Dale Stahlecker | Colorado |
| 709-3881 | Dale Stahlecker | Colorado |
| 709-3883 | Dale Stahlecker | Nebraska |
| 709-3884 | Dale Stahlecker | Nebraska |
| 709-3886 | Dale Stahlecker | Colorado |
| 709-750 | Dale Stahlecker | Colorado |
| 799-1251 | Dale Stahlecker | Colorado |
| 799-1253 | Dale Stahlecker | Nebraska |
| 799-1254 | Dale Stahlecker | Nebraska |
| 0709-01687 | Garth Herring | Oregon |
| 0709-01689 | Garth Herring | Oregon |
| 0709-02012 | Garth Herring | Oregon |
| 0709-02013 | Garth Herring | Oregon |
| 0709-02018 | Garth Herring | Oregon |
| 0709-02021 | Garth Herring | Oregon |
| 0709-02023 | Garth Herring | Oregon |
| 0709-02030 | Garth Herring | Oregon |
| 0709-02034 | Garth Herring | Oregon |
| 0709-02051 | Garth Herring | Oregon |
| 0709-02275 | Garth Herring | Oregon |
| 0709-02276 | Garth Herring | Oregon |
| 0709-02314 | Garth Herring | Oregon |
| 0709-04007 | Garth Herring | Idaho |
| 0709-04008 | Garth Herring | Idaho |
| 0709-04151 | Garth Herring | Oregon |
| 0709-04168 | Garth Herring | Oregon |
| 0709-04169 | Garth Herring | Oregon |
| 0709-04175 | Garth Herring | Oregon |
| 0709-04176 | Garth Herring | Oregon |
| 0709-04177 | Garth Herring | Oregon |
| 0709-04178 | Garth Herring | Oregon |
| 0709-04181 | Garth Herring | Oregon |
| 0709-04192 | Garth Herring | Oregon |
| 0799-00153 | Garth Herring | Oregon |
| 0799-00662 | Garth Herring | Idaho |
| 0799-00712 | Garth Herring | Oregon |
| 0799-00714 | Garth Herring | Oregon |
| 629-33469 | Garth Herring | California |
| 629-33471 | Garth Herring | California |
| 629-33472 | Garth Herring | California |

| | | |
|---|---|---|
| 629-33473 | Garth Herring | California |
| 629-33474 | Garth Herring | California |
| 629-33480 | Garth Herring | California |
| 709-02037 | Garth Herring | Oregon |
| 709-04003 | Garth Herring | Idaho |
| 709-04010 | Garth Herring | Idaho |
| 709-04191 | Garth Herring | Oregon |
| GE-KRRE-B1 | Garth Herring | Oregon |
| 0629-50758 | Gary Roemer | Utah |
| 0709-02965 | Gary Roemer | Utah |
| 0709-02971 | Gary Roemer | Utah |
| 0709-03874 | Gary Roemer | Texas |
| 0709-03875 | Gary Roemer | Texas |
| 0709-03876 | Gary Roemer | Texas |
| 0799-00741 | Gary Roemer | Texas |
| 0799-00746 | Gary Roemer | New Mexico |
| 623 | Gary Roemer | New Mexico |
| 629-50751 | Gary Roemer | Arizona |
| 629-50753 | Gary Roemer | New Mexico |
| 629-50755 | Gary Roemer | New Mexico |
| 629-50756 | Gary Roemer | New Mexico |
| 629-50757 | Gary Roemer | Arizona |
| 629-50759 | Gary Roemer | Arizona |
| 629-50760 | Gary Roemer | New Mexico |
| 629-50761 | Gary Roemer | New Mexico |
| 629-50762 | Gary Roemer | New Mexico |
| 629-50764 | Gary Roemer | New Mexico |
| 629-50765 | Gary Roemer | New Mexico |
| 629-50766 | Gary Roemer | Colorado |
| 629-50769 | Gary Roemer | Colorado |
| 629-50770 | Gary Roemer | Colorado |
| 629-50772 | Gary Roemer | Colorado |
| 629-50773 | Gary Roemer | New Mexico |
| 629-50774 | Gary Roemer | New Mexico |
| 629-50786 | Gary Roemer | New Mexico |
| 629-50788 | Gary Roemer | Colorado |
| 629-50790 | Gary Roemer | Colorado |
| 629-50796 | Gary Roemer | New Mexico |
| 679-02283 | Gary Roemer | New Mexico |
| 679-02284 | Gary Roemer | New Mexico |

| | | |
|---|---|---|
| 709-02951 | Gary Roemer | Colorado |
| 709-02954 | Gary Roemer | Colorado |
| 709-02957 | Gary Roemer | Colorado |
| 709-03852 | Gary Roemer | Colorado |
| 709-03854 | Gary Roemer | Colorado |
| 709-03855 | Gary Roemer | Colorado |
| 799-00098 | Gary Roemer | New Mexico |
| 799-00100 | Gary Roemer | Colorado |
| 629-50752 | Gary Rroemer | New Mexico |
| E47 | Grey Snow Eagle House | Oklahoma |
| 709-02953 | Gary Roemer | New Mexico |
| 0629-51877 | Rob Domenech | British Columbia |
| 0709-00357 | Rob Domenech | Alaska |
| 0709-01852 | Rob Domenech | British Columbia |
| 0709-01873 | Rob Domenech | Alaska |
| 0709-02483 | Rob Domenech | Alaska |
| 0719-00347 | Rob Domenech | Alaska |
| 0719-00349 | Rob Domenech | Alaska |
| 0719-00367 | Rob Domenech | Alaska |
| 0719-00384 | Rob Domenech | Alaska |
| 0719-02458 | Rob Domenech | British Columbia |

Table S2 GO molecular functions and biological process for the genes that were associated with SNPs under selection.

| Gene/Region | GO Database Molecular Function | GO Database Biological Process |
|---|---|---|
| ATXN1 | protein C-terminus binding(GO:0008022);poly(U) RNA binding(GO:0008266);poly(G) binding(GO:0034046);identical protein binding(GO:0042802);protein self-association(GO:0043621) | negative regulation of transcription, DNA-templated(GO:0045892);nuclear export(GO:0051168) |
| BASP1 | | multicellular organism development(GO:0007275);negative regulation of transcription, DNA-templated(GO:0045892) |
| BMP4 | heparin binding(GO:0008201);co-receptor binding(GO:0039706);chemoattractant activity(GO:0042056);cytokine activity(GO:0005125);transforming growth factor beta receptor binding(GO:0005160);growth factor activity(GO:0008083);BMP receptor binding(GO:0070700) | activation of MAPKK activity(GO:0000186);anterior/posterior axis specification(GO:0009948);aortic valve morphogenesis(GO:0003180);beak morphogenesis(GO:0071729);blood vessel endothelial cell proliferation involved in sprouting angiogenesis(GO:0002043);BMP signaling pathway involved in heart development(GO:0061312);BMP signaling pathway involved in heart induction(GO:0003130);BMP signaling pathway involved in nephric duct formation(GO:0071893);BMP signaling pathway involved in renal system segmentation(GO:0061151);BMP signaling pathway involved in ureter morphogenesis(GO:0061149);BMP signaling pathway(GO:0030509);branching involved in prostate gland morphogenesis(GO:0060442);branching involved in ureteric bud morphogenesis(GO:0001658);bronchus development(GO:0060433);bud dilation involved in lung branching(GO:0060503);bud elongation involved in lung branching(GO:0060449);cardiac muscle cell differentiation(GO:0055007);cartilage development(GO:0051216);cell development(GO:0048468);cell proliferation(GO:0008283);chondrocyte differentiation(GO:0002062);cloacal septation(GO:0060197);common-partner SMAD protein phosphorylation(GO:0007182);coronary vasculature development(GO:0060976);cranial |

| | | |
|---|---|---|
| | | suture morphogenesis(GO:0060363);deltoid tuberosity development(GO:0035993);dorsal/ventral neural tube patterning(GO:0021904);embryonic cranial skeleton morphogenesis(GO:0048701);embryonic digit morphogenesis(GO:0042733);embryonic hindlimb morphogenesis(GO:0035116);embryonic skeletal joint morphogenesis(GO:0060272);embryonic skeletal system morphogenesis(GO:0048704);endocardial cushion development(GO:0003197);endochondral ossification(GO:0001958);endoderm development(GO:0007492);epithelial cell proliferation involved in lung morphogenesis(GO:0060502);epithelial-mesenchymal cell signaling(GO:0060684);erythrocyte differentiation(GO:0030218);germ cell development(GO:0007281);glomerular capillary formation(GO:0072104);glomerular visceral epithelial cell development(GO:0072015);inner ear receptor cell differentiation(GO:0060113);intermediate mesodermal cell differentiation(GO:0048392);lens induction in camera-type eye(GO:0060235);lung alveolus development(GO:0048286);lymphoid progenitor cell differentiation(GO:0002320);macrophage differentiation(GO:0030225);membranous septum morphogenesis(GO:0003149);mesenchymal cell differentiation involved in kidney development(GO:0072161);mesenchymal cell proliferation involved in ureter development(GO:0072198);mesenchymal cell proliferation involved in ureteric bud development(GO:0072138);mesenchymal to epithelial transition involved in metanephros morphogenesis(GO:0003337);mesodermal cell fate determination(GO:0007500);metanephric collecting duct development(GO:0072205);monocyte differentiation(GO:0030224);negative regulation of apoptotic process(GO:0043066);negative regulation of branch elongation involved in ureteric bud branching by BMP signaling pathway(GO:0072097);negative regulation of cell proliferation involved in heart morphogenesis(GO:2000137);negative regulation of chondrocyte differentiation(GO:0032331);negative regulation of epithelial cell proliferation(GO:0050680);negative regulation of extrinsic apoptotic signaling pathway(GO:2001237);negative regulation of glomerular mesangial cell |

| | | |
|---|---|---|
| | | proliferation(GO:0072125);negative regulation of immature T cell proliferation in thymus(GO:0033088);negative regulation of MAP kinase activity(GO:0043407);negative regulation of mesenchymal cell proliferation involved in ureter development(GO:0072200);negative regulation of metanephric comma-shaped body morphogenesis(GO:2000007);negative regulation of metanephric S-shaped body morphogenesis(GO:2000005);negative regulation of mitotic nuclear division(GO:0045839);negative regulation of myoblast differentiation(GO:0045662);negative regulation of prostatic bud formation(GO:0060686);negative regulation of striated muscle tissue development(GO:0045843);negative regulation of thymocyte apoptotic process(GO:0070244);negative regulation of transcription from RNA polymerase II promoter(GO:0000122);negative regulation of transcription, DNA-templated(GO:0045892);nephric duct formation(GO:0072179);neural tube closure(GO:0001843);neuron fate commitment(GO:0048663);ossification(GO:0001503);osteoblast differentiation(GO:0001649);outflow tract septum morphogenesis(GO:0003148);pharyngeal arch artery morphogenesis(GO:0061626);pituitary gland development(GO:0021983);polarity specification of dorsal/ventral axis(GO:0009951);positive chemotaxis(GO:0050918);positive regulation of apoptotic process(GO:0043065);positive regulation of BMP signaling pathway(GO:0030513);positive regulation of bone mineralization(GO:0030501);positive regulation of branching involved in lung morphogenesis(GO:0061047);positive regulation of cardiac muscle fiber development(GO:0055020);positive regulation of cartilage development(GO:0061036);positive regulation of cell proliferation involved in outflow tract morphogenesis(GO:1901964);positive regulation of collagen biosynthetic process(GO:0032967);positive regulation of DNA-dependent DNA replication(GO:2000105);positive regulation of endothelial cell differentiation(GO:0045603);positive regulation of endothelial cell migration(GO:0010595);positive regulation of endothelial cell proliferation(GO:0001938);positive regulation of epidermal cell |

| | | |
|---|---|---|
| | | differentiation(GO:0045606);positive regulation of ERK1 and ERK2 cascade(GO:0070374);positive regulation of kidney development(GO:0090184);positive regulation of neuron differentiation(GO:0045666);positive regulation of osteoblast differentiation(GO:0045669);positive regulation of pathway-restricted SMAD protein phosphorylation(GO:0010862);positive regulation of pri-miRNA transcription from RNA polymerase II promoter(GO:1902895);positive regulation of production of miRNAs involved in gene silencing by miRNA(GO:1903800);positive regulation of progesterone secretion(GO:2000872);positive regulation of protein binding(GO:0032092);positive regulation of SMAD protein import into nucleus(GO:0060391);positive regulation of smooth muscle cell proliferation(GO:0048661);positive regulation of transcription, DNA-templated(GO:0045893);post-embryonic development(GO:0009791);protein localization to nucleus(GO:0034504);pulmonary artery endothelial tube morphogenesis(GO:0061155);pulmonary valve morphogenesis(GO:0003184);regulation of branching involved in prostate gland morphogenesis(GO:0060687);regulation of cell fate commitment(GO:0010453);regulation of MAPK cascade(GO:0043408);regulation of myotome development(GO:2000290);regulation of odontogenesis of dentin-containing tooth(GO:0042487);regulation of smooth muscle cell differentiation(GO:0051150);renal system process(GO:0003014);secondary heart field specification(GO:0003139);SMAD protein signal transduction(GO:0060395);smoothened signaling pathway(GO:0007224);specification of animal organ position(GO:0010159);specification of ureteric bud anterior/posterior symmetry by BMP signaling pathway(GO:0072101);steroid hormone mediated signaling pathway(GO:0043401);telencephalon regionalization(GO:0021978);tendon cell differentiation(GO:0035990);trachea formation(GO:0060440);type B pancreatic cell development(GO:0003323);ureter epithelial cell differentiation(GO:0072192);ureter smooth muscle cell differentiation(GO:0072193);ureteric bud morphogenesis(GO:0060675); |

| | | |
|---|---|---|
| CDH10 | calcium ion binding(GO:0005509) | homophilic cell adhesion via plasma membrane adhesion molecules(GO:0007156) |
| CDH12 | calcium ion binding(GO:0005509) | homophilic cell adhesion via plasma membrane adhesion molecules(GO:0007156) |
| CDKN3 | protein serine/threonine phosphatase activity(GO:0004722);protein tyrosine phosphatase activity(GO:0004725);protein tyrosine/serine/threonine phosphatase activity(GO:0008138) | G1/S transition of mitotic cell cycle(GO:0000082);cell cycle arrest(GO:0007050);peptidyl-tyrosine dephosphorylation(GO:0035335) |
| CELF2 | nucleotide binding(GO:0000166);translation repressor activity, nucleic acid binding(GO:0000900);nucleic acid binding(GO:0003676);RNA binding(GO:0003723);BRE binding(GO:0042835) | alternative mRNA splicing, via spliceosome(GO:0000380);negative regulation of translation(GO:0017148) |
| COLEC11 | mannose binding(GO:0005537) | multicellular organism development(GO:0007275);developmental process(GO:0032502) |
| CRELD2 | calcium ion binding(GO:0005509) | |
| DNAH5 | ATP binding(GO:0005524);ATP-dependent microtubule motor activity, minus-end-directed(GO:0008569);dynein light chain binding(GO:0045503);dynein intermediate chain binding(GO:0045505);dynein light intermediate chain binding(GO:0051959) | cilium movement(GO:0003341);determination of left/right symmetry(GO:0007368);heart development(GO:0007507);lateral ventricle development(GO:0021670);flagellated sperm motility(GO:0030317);outer dynein arm assembly(GO:0036158) |
| DTWD2 | | positive regulation of protein targeting to mitochondrion(GO:1903955) |
| ELAVL2 | RNA binding(GO:0003723) | |
| ESM1 | hepatocyte growth factor receptor binding(GO:0005171);integrin binding(GO:0005178);insulin-like growth factor binding(GO:0005520) | angiogenesis(GO:0001525);regulation of cell growth(GO:0001558);sprouting angiogenesis(GO:0002040);positive regulation of cell proliferation(GO:0008284);positive regulation of hepatocyte growth factor receptor signaling pathway(GO:1902204) |

| | | |
|---|---|---|
| ESRP2 | mRNA binding(GO:0003729);RNA binding(GO:0003723) | mRNA processing(GO:0006397);RNA splicing(GO:0008380);regulation of RNA splicing(GO:0043484) |
| FSHR | follicle-stimulating hormone receptor activity(GO:0004963);G-protein coupled peptide receptor activity(GO:0008528) | activation of adenylate cyclase activity(GO:0007190);adenylate cyclase-activating G-protein coupled receptor signaling pathway(GO:0007189);adenylate cyclase-inhibiting G-protein coupled receptor signaling pathway(GO:0007193);basement membrane organization(GO:0071711);cellular response to follicle-stimulating hormone stimulus(GO:0071372);cellular water homeostasis(GO:0009992);follicle-stimulating hormone signaling pathway(GO:0042699);hormone-mediated signaling pathway(GO:0009755);locomotory behavior(GO:0007626);negative regulation of bone resorption(GO:0045779);neuron projection development(GO:0031175);phospholipase C-activating G-protein coupled receptor signaling pathway(GO:0007200);positive regulation of adenylate cyclase activity(GO:0045762);positive regulation of ERK1 and ERK2 cascade(GO:0070374);positive regulation of intracellular estrogen receptor signaling pathway(GO:0033148);positive regulation of luteinizing hormone secretion(GO:0033686);positive regulation of phosphatidylinositol 3-kinase signaling(GO:0014068);primary ovarian follicle growth(GO:0001545);regulation of acetylcholine metabolic process(GO:0060408);regulation of chromosome organization(GO:0033044);regulation of hormone metabolic process(GO:0032350);regulation of MAPK cascade(GO:0043408);regulation of osteoclast differentiation(GO:0045670);regulation of platelet-derived growth factor receptor signaling pathway(GO:0010640);regulation of protein kinase A signaling(GO:0010738);regulation of systemic arterial blood pressure(GO:0003073);Sertoli cell development(GO:0060009);Sertoli cell proliferation(GO:0060011);spermatogenesis, exchange of chromosomal proteins(GO:0035093);uterus development(GO:0060065); |
| FSTL4 | calcium ion binding(GO:0005509) | |

| | | |
|---|---|---|
| GABRG2 | GABA-A receptor activity(GO:0004890);extracellular ligand-gated ion channel activity(GO:0005230);chloride channel activity(GO:0005254) | gamma-aminobutyric acid signaling pathway(GO:0007214);chemical synaptic transmission(GO:0007268);post-embryonic development(GO:0009791);adult behavior(GO:0030534);cellular response to histamine(GO:0071420);chloride transmembrane transport(GO:1902476) |
| GATA3 | RNA polymerase II regulatory region sequence-specific DNA binding(GO:0000977);RNA polymerase II core promoter sequence-specific DNA binding(GO:0000979);core promoter proximal region sequence-specific DNA binding(GO:0000987);transcriptional activator activity, RNA polymerase II core promoter proximal region sequence-specific binding(GO:0001077);transcriptional repressor activity, RNA polymerase II core promoter proximal region sequence-specific binding(GO:0001078);RNA polymerase II transcription factor binding(GO:0001085);enhancer sequence-specific DNA binding(GO:0001158);transcriptional activator activity, RNA polymerase II transcription regulatory region sequence-specific binding(GO:0001228);chromatin binding(GO:0003682);transcription coactivator activity(GO:0003713);interleukin-2 receptor binding(GO:0005134);protein binding(GO:0005515);zinc ion binding(GO:0008270);transcription regulatory region DNA binding(GO:0044212);protein dimerization | anatomical structure formation involved in morphogenesis(GO:0048646);aortic valve morphogenesis(GO:0003180);auditory receptor cell differentiation(GO:0042491);axon guidance(GO:0007411);canonical Wnt signaling pathway involved in metanephric kidney development(GO:0061290);cardiac right ventricle morphogenesis(GO:0003215);cell fate determination(GO:0001709);cell maturation(GO:0048469);cellular response to BMP stimulus(GO:0071773);cellular response to interferon-alpha(GO:0035457);cellular response to interleukin-4(GO:0071353);cellular response to tumor necrosis factor(GO:0071356);chromatin remodeling(GO:0006338);developmental growth(GO:0048589);ear development(GO:0043583);embryonic hemopoiesis(GO:0035162);erythrocyte differentiation(GO:0030218);humoral immune response(GO:0006959);inner ear morphogenesis(GO:0042472);interferon-gamma secretion(GO:0072643);interleukin-4 secretion(GO:0072602);kidney development(GO:0001822);lens development in camera-type eye(GO:0002088);lymphocyte migration(GO:0072676);male gonad development(GO:0008584);mast cell differentiation(GO:0060374);mesenchymal to epithelial transition(GO:0060231);mesonephros development(GO:0001823);negative regulation of cell cycle(GO:0045786);negative regulation of cell motility(GO:2000146);negative regulation of cell proliferation involved in mesonephros development(GO:2000607);negative regulation of cell proliferation(GO:0008285);negative regulation of DNA demethylation(GO:1901536);negative regulation of endothelial cell apoptotic process(GO:2000352);negative regulation of fat cell differentiation(GO:0045599);negative regulation of fibroblast growth factor receptor signaling pathway involved in ureteric bud formation(GO:2000703);negative regulation of glial cell-derived neurotrophic factor receptor signaling pathway involved in ureteric bud formation(GO:2000734);negative regulation of |

| | | |
|---|---|---|
| | activity(GO:0046983);E-box binding(GO:0070888);HMG box domain binding(GO:0071837) | inflammatory response(GO:0050728);negative regulation of interferon-gamma production(GO:0032689);negative regulation of interleukin-2 production(GO:0032703);negative regulation of mammary gland epithelial cell proliferation(GO:0033600);negative regulation of transcription from RNA polymerase II promoter(GO:0000122);negative regulation of transcription, DNA-templated(GO:0045892);nephric duct formation(GO:0072179);nephric duct morphogenesis(GO:0072178);neuron migration(GO:0001764);norepinephrine biosynthetic process(GO:0042421);otic vesicle development(GO:0071599);parathyroid gland development(GO:0060017);parathyroid hormone secretion(GO:0035898);pharyngeal system development(GO:0060037);phosphatidylinositol 3-kinase signaling(GO:0014065);positive regulation of endothelial cell migration(GO:0010595);positive regulation of histone H3-K14 acetylation(GO:0071442);positive regulation of histone H3-K9 acetylation(GO:2000617);positive regulation of interleukin-13 secretion(GO:2000667);positive regulation of interleukin-4 production(GO:0032753);positive regulation of interleukin-5 secretion(GO:2000664);positive regulation of protein kinase B signaling(GO:0051897);positive regulation of signal transduction(GO:0009967);positive regulation of T cell differentiation(GO:0045582);positive regulation of T-helper 2 cell cytokine production(GO:2000553);positive regulation of thyroid hormone generation(GO:2000611);positive regulation of transcription from RNA polymerase II promoter(GO:0045944);positive regulation of transcription regulatory region DNA binding(GO:2000679);positive regulation of transcription, DNA-templated(GO:0045893);positive regulation of ureteric bud formation(GO:0072107);post-embryonic development(GO:0009791);pro-T cell differentiation(GO:0002572);regulation of CD4-positive, alpha-beta T cell differentiation(GO:0043370);regulation of cellular response to X-ray(GO:2000683);regulation of cytokine biosynthetic process(GO:0042035);regulation of establishment of cell polarity(GO:2000114);regulation of histone H3-K27 methylation(GO:0061085);regulation of histone H3-K4 methylation(GO:0051569);regulation of nephron |

| | | |
|---|---|---|
| | | tubule epithelial cell differentiation(GO:0072182);regulation of neuron apoptotic process(GO:0043523);regulation of neuron projection development(GO:0010975);response to estrogen(GO:0043627);response to virus(GO:0009615);signal transduction(GO:0007165);sympathetic nervous system development(GO:0048485);T cell receptor signaling pathway(GO:0050852);T-helper 2 cell differentiation(GO:0045064);thymic T cell selection(GO:0045061);thymus development(GO:0048538);TOR signaling(GO:0031929);transcription from RNA polymerase II promoter(GO:0006366);ureter maturation(GO:0035799);ureteric bud formation(GO:0060676);uterus development(GO:0060065);ventricular septum development(GO:0003281) |
| IGFBP5 | fibronectin binding(GO:0001968);insulin-like growth factor I binding(GO:0031994);insulin-like growth factor II binding(GO:0031995) | negative regulation of transcription from RNA polymerase II promoter(GO:0000122);neural crest cell migration(GO:0001755);secondary heart field specification(GO:0003139);outflow tract septum morphogenesis(GO:0003148);endocardial cushion morphogenesis(GO:0003203);cardiac right ventricle morphogenesis(GO:0003215);regulation of secondary heart field cardioblast proliferation(GO:0003266);transcription, DNA-templated(GO:0006351);positive regulation of cell proliferation(GO:0008284);positive regulation of vascular endothelial growth factor production(GO:0010575);spinal cord motor neuron cell fate specification(GO:0021520);visceral motor neuron differentiation(GO:0021524);trigeminal nerve development(GO:0021559);pituitary gland development(GO:0021983);pancreas development(GO:0031016);retinal ganglion cell axon guidance(GO:0031290);positive regulation of granulocyte macrophage colony-stimulating factor production(GO:0032725);positive regulation of interferon-gamma production(GO:0032729);positive regulation of interleukin-1 alpha production(GO:0032730);positive regulation of interleukin-1 beta production(GO:0032731);positive regulation of interleukin-12 production(GO:0032735);positive regulation of interleukin-6 production(GO:0032755);positive regulation of tumor necrosis factor production(GO:0032760);positive regulation of histone acetylation(GO:0035066);positive regulation of tyrosine phosphorylation of STAT protein(GO:0042531);positive regulation of DNA |

| | | |
|---|---|---|
| | | binding(GO:0043388);negative regulation of neuron apoptotic process(GO:0043524);positive regulation of cell differentiation(GO:0045597);negative regulation of neuron differentiation(GO:0045665);positive regulation of angiogenesis(GO:0045766);negative regulation of transcription, DNA-templated(GO:0045892);positive regulation of transcription from RNA polymerase II promoter(GO:0045944);sensory system development(GO:0048880);peripheral nervous system neuron axonogenesis(GO:0048936);negative regulation of inflammatory response(GO:0050728);ventricular cardiac muscle tissue morphogenesis(GO:0055010);pharyngeal system development(GO:0060037);cardiac muscle cell myoblast differentiation(GO:0060379);innervation(GO:0060384);atrial septum morphogenesis(GO:0060413);cardiac cell fate determination(GO:0060913);positive regulation of granulocyte colony-stimulating factor production(GO:0071657);negative regulation of canonical Wnt signaling pathway(GO:0090090); |
| ISL1 | RNA polymerase II activating transcription factor binding(GO:0001102);RRNA polymerase II transcription coactivator activity(GO:0001105);enhancer sequence-specific DNA binding(GO:0001158);bHLH transcription factor binding(GO:0043425);metal ion binding(GO:0046872);promoter-specific chromatin binding(GO:1990841) | negative regulation of transcription from RNA polymerase II promoter(GO:0000122);neural crest cell migration(GO:0001755);secondary heart field specification(GO:0003139);outflow tract septum morphogenesis(GO:0003148);endocardial cushion morphogenesis(GO:0003203);cardiac right ventricle morphogenesis(GO:0003215);regulation of secondary heart field cardioblast proliferation(GO:0003266);transcription, DNA-templated(GO:0006351);positive regulation of cell proliferation(GO:0008284);positive regulation of vascular endothelial growth factor production(GO:0010575);spinal cord motor neuron cell fate specification(GO:0021520);visceral motor neuron differentiation(GO:0021524);trigeminal nerve development(GO:0021559);pituitary gland development(GO:0021983);pancreas development(GO:0031016);retinal ganglion cell axon guidance(GO:0031290);positive regulation of granulocyte macrophage colony-stimulating factor production(GO:0032725);positive regulation of interferon-gamma production(GO:0032729);positive regulation of interleukin-1 alpha production(GO:0032730);positive regulation of interleukin-1 beta production(GO:0032731);positive regulation of interleukin-12 production(GO:0032735);positive regulation of interleukin-6 |

| | |
|---|---|
| | production(GO:0032755);positive regulation of tumor necrosis factor production(GO:0032760);positive regulation of histone acetylation(GO:0035066);positive regulation of tyrosine phosphorylation of STAT protein(GO:0042531);positive regulation of DNA binding(GO:0043388);negative regulation of neuron apoptotic process(GO:0043524);positive regulation of cell differentiation(GO:0045597);negative regulation of neuron differentiation(GO:0045665);positive regulation of angiogenesis(GO:0045766);negative regulation of transcription, DNA-templated(GO:0045892);positive regulation of transcription from RNA polymerase II promoter(GO:0045944);sensory system development(GO:0048880);peripheral nervous system neuron axonogenesis(GO:0048936);negative regulation of inflammatory response(GO:0050728);ventricular cardiac muscle tissue morphogenesis(GO:0055010);pharyngeal system development(GO:0060037);cardiac muscle cell myoblast differentiation(GO:0060379);innervation(GO:0060384);atrial septum morphogenesis(GO:0060413);cardiac cell fate determination(GO:0060913);cardiac cell fate determination(GO:0060913);positive regulation of granulocyte colony-stimulating factor production(GO:0071657);negative regulation of canonical Wnt signaling pathway(GO:0090090) |

| | | |
|---|---|---|
| ITGA1 | protein phosphatase binding(GO:0019903);collagen binding involved in cell-matrix adhesion(GO:0098639) | cell-matrix adhesion(GO:0007160);integrin-mediated signaling pathway(GO:0007229);negative regulation of cell proliferation(GO:0008285);neutrophil chemotaxis(GO:0030593);positive regulation of phosphoprotein phosphatase activity(GO:0032516);negative regulation of epidermal growth factor receptor signaling pathway(GO:0042059);cellular extravasation(GO:0045123);cell adhesion(GO:0007155);integrin-mediated signaling pathway(GO:0007229);negative regulation of cell proliferation(GO:0008285);positive regulation of phosphoprotein phosphatase activity(GO:0032516);negative regulation of epidermal growth factor receptor signaling pathway(GO:0042059);cell adhesion(GO:0007155);integrin-mediated signaling pathway(GO:0007229) |
| KHDRBS3 | RNA binding(GO:0003723) | |
| KLF12 | nucleic acid binding(GO:0003676) | |
| KLHL1 | ubiquitin-protein transferase activity(GO:0004842) | adult walking behavior(GO:0007628);dendrite development(GO:0016358);protein ubiquitination(GO:0016567);cerebellar Purkinje cell layer development(GO:0021680) |
| LMO4 | enhancer sequence-specific DNA binding(GO:0001158);transcription factor binding(GO:0008134);metal ion binding(GO:0046872) | neural tube closure(GO:0001843);ventricular septum development(GO:0003281);transcription from RNA polymerase II promoter(GO:0006366);ventral spinal cord interneuron differentiation(GO:0021514);spinal cord motor neuron differentiation(GO:0021522);spinal cord association neuron differentiation(GO:0021527);regulation of cell migration(GO:0030334);negative regulation of protein complex assembly(GO:0031333);positive regulation of kinase activity(GO:0033674);regulation of cell fate specification(GO:0042659);positive regulation of transcription from RNA polymerase II promoter(GO:0045944);thymus development(GO:0048538);regulation of cell activation(GO:0050865) |

| | | |
|---|---|---|
| LRRC4C | protein kinase inhibitor activity(GO:0004860) | negative regulation of protein kinase activity(GO:0006469);cytokine-mediated signaling pathway(GO:0019221);negative regulation of JAK-STAT cascade(GO:0046426);regulation of axonogenesis(GO:0050770) |
| MELK | protein serine/threonine kinase activity(GO:0004674);non-membrane spanning protein tyrosine kinase activity(GO:0004715);ATP binding(GO:0005524) | protein phosphorylation(GO:0006468);establishment of cell polarity(GO:0030010);microtubule cytoskeleton organization(GO:0000226);protein phosphorylation(GO:0006468);peptidyl-tyrosine phosphorylation(GO:0018108);establishment of cell polarity(GO:0030010);intracellular signal transduction(GO:0035556) |
| MICU2 | calcium ion binding(GO:0005509);protein heterodimerization activity(GO:0046982) | mitochondrial calcium uptake(GO:0036444);positive regulation of mitochondrial calcium ion concentration(GO:0051561);negative regulation of mitochondrial calcium ion concentration(GO:0051562) |
| MMP17 | metalloendopeptidase activity(GO:0004222);zinc ion binding(GO:0008270) | proteolysis(GO:0006508) |
| MRPS30 | structural constituent of ribosome(GO:0003735) | translation(GO:0006412) |
| MRPS5 | RNA binding(GO:0003723);structural constituent of ribosome(GO:0003735) | translation(GO:0006412) |
| MSH3 | Y-form DNA binding(GO:0000403);heteroduplex DNA loop binding(GO:0000404);double-strand/single-strand DNA junction binding(GO:0000406);damaged DNA binding(GO:0003684);single-stranded DNA binding(GO:0003697);ATP binding(GO:0005524);DNA-dependent ATPase activity(GO:0008094);enzyme binding(GO:0019899);single guanine insertion binding(GO:0032142);dinucleotide repeat insertion binding(GO:0032181) | mismatch repair(GO:0006298);somatic recombination of immunoglobulin gene segments(GO:0016447);maintenance of DNA repeat elements(GO:0043570);negative regulation of DNA recombination(GO:0045910);positive regulation of helicase activity(GO:0051096) |

| | | |
|---|---|---|
| NNT | NAD(P)+ transhydrogenase activity(GO:0008746);NAD (P)+ transhydrogenase (AB-specific) activity(GO:0008750);NAD P binding(GO:0050661);NAD binding(GO:0051287) | oxidation-reduction process(GO:0055114);NADPH regeneration(GO:0006740);proton transport(GO:0015992);oxidation-reduction process(GO:0055114) |
| NPY2R | peptide YY receptor activity(GO:0001601) | outflow tract morphogenesis(GO:0003151);cardiac left ventricle morphogenesis(GO:0003214);adenylate cyclase-inhibiting G-protein coupled receptor signaling pathway(GO:0007193);neuropeptide signaling pathway(GO:0007218);chemical synaptic transmission(GO:0007268);feeding behavior(GO:0007631) |
| NRXN1 | metal ion binding(GO:0046872) | cell adhesion(GO:0007155) |
| NRXN1 | | angiogenesis(GO:0001525);cell adhesion(GO:0007155) |
| PAX5 | DNA binding(GO:0003677) | transcription, DNA-templated(GO:0006351);regulation of transcription, DNA-templated(GO:0006355);multicellular organism development(GO:0007275) |
| POLQ | DNA binding(GO:0003677);chro matin binding(GO:0003682);DNA -directed DNA polymerase activity(GO:0003887);ATP binding(GO:0005524);singl e-stranded DNA-dependent ATPase activity(GO:0043142);5'-deoxyribose-5-phosphate lyase activity(GO:0051575) | DNA-dependent DNA replication(GO:0006261);base-excision repair(GO:0006284);double-strand break repair(GO:0006302);cellular response to DNA damage stimulus(GO:0006974);somatic hypermutation of immunoglobulin genes(GO:0016446);protein homooligomerization(GO:0051260);DNA biosynthetic process(GO:0071897);double-strand break repair via alternative nonhomologous end joining(GO:0097681);negative regulation of double-strand break repair via homologous recombination(GO:2000042) |
| PPP4R2 | protein phosphatase regulator activity(GO:0019888);prote in binding, bridging(GO:0030674) | regulation of double-strand break repair via homologous recombination(GO:0010569);regulation of catalytic activity(GO:0050790) |
| PSAT1 | O-phospho-L-serine:2-oxoglutarate aminotransferase activity(GO:0004648) | L-serine biosynthetic process(GO:0006564) |

| | | |
|---|---|---|
| PXDN | extracellular matrix structural constituent(GO:0005201);heme binding(GO:0020037);peroxidase activity(GO:0004601);heme binding(GO:0020037) | response to oxidative stress(GO:0006979);extracellular matrix organization(GO:0030198);hydrogen peroxide catabolic process(GO:0042744);oxidation-reduction process(GO:0055114);cellular oxidant detoxification(GO:0098869);response to oxidative stress(GO:0006979);oxidation-reduction process(GO:0055114);cellular oxidant detoxification(GO:0098869) |
| RAPGEF6 | guanyl-nucleotide exchange factor activity(GO:0005085); | positive regulation of GTPase activity(GO:0043547);small GTPase mediated signal transduction(GO:0007264);positive regulation of GTPase activity(GO:0043547) |
| RASEF | GTPase activity(GO:0003924);calcium ion binding(GO:0005509);GTP binding(GO:0005525) | |
| RASGRF2 | Rho guanyl-nucleotide exchange factor activity(GO:0005089) | small GTPase mediated signal transduction(GO:0007264);regulation of Rho protein signal transduction(GO:0035023);positive regulation of GTPase activity(GO:0043547) |
| SEMA6A | transmembrane signaling receptor activity(GO:0004888);semaphorin receptor binding(GO:0030215) | neuron migration(GO:0001764);axon guidance(GO:0007411);centrosome localization(GO:0051642);semaphorin-plexin signaling pathway(GO:0071526);positive regulation of neuron migration(GO:2001224) |
| SFSWAP | RNA binding(GO:0003723) | alternative mRNA splicing, via spliceosome(GO:0000380);mRNA 5'-splice site recognition(GO:0000395);RNA processing(GO:0006396) |
| SGCE | | muscle system process(GO:0003012);membrane organization(GO:0061024) |
| SKIV2L2 | RNA binding(GO:0003723);ATP-dependent RNA helicase activity(GO:0004004);ATP binding(GO:0005524) | maturation of 5.8S rRNA(GO:0000460);RNA catabolic process(GO:0006401) |
| SLC30A8 | zinc ion transmembrane transporter activity(GO:0005385) | cellular zinc ion homeostasis(GO:0006882);zinc II ion transmembrane transport(GO:0071577) |

| | | |
|---|---|---|
| SLC39A10 | zinc ion transmembrane transporter activity(GO:0005385) | negative regulation of B cell apoptotic process(GO:0002903);cellular zinc ion homeostasis(GO:0006882);positive regulation of B cell proliferation(GO:0030890);positive regulation of B cell receptor signaling pathway(GO:0050861);zinc II ion transmembrane import(GO:0071578);positive regulation of protein tyrosine phosphatase activity(GO:1903615) |
| SLC8A1 | calcium:sodium antiporter activity(GO:0005432);ankyrin binding(GO:0030506) | cellular sodium ion homeostasis(GO:0006883);positive regulation of bone mineralization(GO:0030501);response to muscle stretch(GO:0035994);calcium ion import(GO:0070509);calcium ion transmembrane transport(GO:0070588);sodium ion import(GO:0097369);positive regulation of the force of heart contraction(GO:0098735);calcium ion transport(GO:0006816);cell communication(GO:0007154);sodium ion transmembrane transport(GO:0035725) |
| SORCS3 | | learning(GO:0007612);memory(GO:0007613);regulation of long term synaptic depression(GO:1900452) |
| SOX11 | DNA binding(GO:0003677);transcription factor activity, sequence-specific DNA binding(GO:0003700) | transcription, DNA-templated(GO:0006351);regulation of transcription, DNA-templated(GO:0006355);nervous system development(GO:0007399);cell differentiation(GO:0030154);positive regulation of canonical Wnt signaling pathway(GO:0090263) |
| SPTLC1 | serine C-palmitoyltransferase activity(GO:0004758);pyridoxal phosphate binding(GO:0030170) | sphingomyelin biosynthetic process(GO:0006686);sphinganine biosynthetic process(GO:0046511);sphingosine biosynthetic process(GO:0046512);ceramide biosynthetic process(GO:0046513);positive regulation of lipophagy(GO:1904504) |
| STX11 | SNARE binding(GO:0000149);SNAP receptor activity(GO:0005484) | intracellular protein transport(GO:0006886);synaptic vesicle fusion to presynaptic active zone membrane(GO:0031629);vesicle docking(GO:0048278) |

| | | |
|---|---|---|
| TARS | threonine-tRNA ligase activity(GO:0004829);ATP binding(GO:0005524) | threonyl-tRNA aminoacylation(GO:0006435) |
| TENM1 | protein binding(GO:0005515);protein homodimerization activity(GO:0042803);protein heterodimerization activity(GO:0046982);cell adhesion molecule binding(GO:0050839);catalytic activity(GO:0003824) | cell morphogenesis(GO:0000902);transcription, DNA-templated(GO:0006351);regulation of transcription from RNA polymerase III promoter(GO:0006359);heterophilic cell-cell adhesion via plasma membrane cell adhesion molecules(GO:0007157);neuropeptide signaling pathway(GO:0007218);positive regulation of actin filament polymerization(GO:0030838);positive regulation of peptidyl-serine phosphorylation(GO:0033138);positive regulation of MAP kinase activity(GO:0043406);neuron development(GO:0048666);positive regulation of filopodium assembly(GO:0051491);positive regulation of intracellular protein transport(GO:0090316);signal transduction(GO:0007165) |
| TLE1 | RNA polymerase II transcription corepressor activity(GO:0001106);transcription factor binding(GO:0008134);identical protein binding(GO:0042802) | negative regulation of transcription from RNA polymerase II promoter(GO:0000122);regulation of transcription, DNA-templated(GO:0006355);positive regulation of gene expression(GO:0010628);negative regulation of I-kappaB kinase/NF-kappaB signaling(GO:0043124);negative regulation of anoikis(GO:2000811) |
| TLE4 | | cell fate determination(GO:0001709);regionalization(GO:0003002);regulation of transcription, DNA-templated(GO:0006355);establishment of tissue polarity(GO:0007164);positive regulation of gene expression(GO:0010628);negative regulation of gene expression(GO:0010629);midbrain development(GO:0030901);formation of anatomical boundary(GO:0048859) |
| TMEM167A | | protein secretion(GO:0009306);intracellular transport(GO:0046907) |
| TPO | cytokine activity(GO:0005125);hormone activity(GO:0005179) | cell proliferation(GO:0008283) |

| | | |
|---|---|---|
| TPO | iodide peroxidase activity(GO:0004447);peroxidase activity(GO:0004601);calcium ion binding(GO:0005509);heme binding(GO:0020037) | thyroid hormone generation(GO:0006590);response to oxidative stress(GO:0006979);embryonic hemopoiesis(GO:0035162);oxidation-reduction process(GO:0055114);cellular oxidant detoxification(GO:0098869) |
| TSHZ3 | transcriptional repressor activity, RNA polymerase II core promoter proximal region sequence-specific binding(GO:0001078);DNA binding(GO:0003677);chromatin binding(GO:0003682) | negative regulation of transcription from RNA polymerase II promoter(GO:0000122);regulation of transcription from RNA polymerase II promoter(GO:0006357);multicellular organism development(GO:0007275) |
| VRK2 | protein kinase activity(GO:0004672);ATP binding(GO:0005524);protein kinase activity(GO:0004672) | protein phosphorylation(GO:0006468) |
| ZNF507 | nucleic acid binding(GO:0003676) | |
| ADGRG6 | | |
| ATP6AP1L | | |
| CAAP1 | | |
| CCDC85A | | |
| CDC42SE2 | | |
| CETN3 | | |
| DACH1 | | |
| DIAPH2 | | |
| EMB | | |
| EMCN | | |
| FAM135B | | |
| FAM19A5 | | |
| FOXA1 | | |
| HIVEP2 | | |
| LOC105399097 | | |
| LOC105399204 | | |
| LOC105399208 | | |
| LOC105399597 | | |
| LOC105399814 | | |
| LOC105399815 | | |

| |
|---|
| LOC105400347 |
| LOC105400396 |
| LOC105400551 |
| LOC105400624 |
| LOC105401319 |
| LOC105401627 |
| LOC105402490 |
| LOC105402630 |
| LOC105403851 |
| LOC105403868 |
| LOC105404008 |
| LOC105404061 |
| LOC105404364 |
| LOC105406972 |
| LOC105407148 |
| LOC105408456 |
| LOC105408723 |
| LOC105408724 |
| LOC105411528 |
| LOC105411531 |
| LOC105411696 |
| LOC105411756 |
| MCTP2 |
| MEF2C |
| MIPOL1 |
| NAV2 |
| NPR3 |
| PPP1R9A |
| PPRC1 |

| |
|---|
| RAD21 |
| SORCS1 |
| SOX11 |
| STMND1 |
| TBC1D4 |
| TMEFF2 |
| TNS1 |
| TWIST2 |
| VSTM2B |

# CHAPTER III

## A 50K SNP ARRAY REVELAS HIGH LEVELS OF GENETIC STRUCTURE FOR BALD EAGLES (*HALIAEETUS LEUCOCEPHALUS)* IN NORTH AMERICA


### Abstract

Bald eagles underwent a severe population bottleneck in the mid-1900s due to the use of Dichlorodiphenyltrichloroethane (DDT) as a pesticide. After DDT's ban in 1972, the population began to recover with the increase being attributed to reintroduction programs. Although the bald eagle population has increased, populations still face stressors such as electrocution, lead poisoning, wind turbines, illegal shooting, and climate change, therefore there is a push to develop informed management plans to aid in their future conservation. The effect of the bottleneck and subsequent reintroductions have not been studied from a genetic perspective, making management decisions for the population difficult. This study presents the first range-wide genomic analyses of the bald eagle population using a custom 50K Affymetrix Axiom myDesign single nucleotide polymorphism (SNP) array. Despite the bottleneck and reintroductions which could have led to a greatly admixed population, we found highly differentiated and supported genetic clusters of bald eagles. Using STRUCTURE, PCoA, and population genetic statistics, we were able to identify seven unique groups of bald eagles throughout North America.

Finally, we were able to identify 379 SNPs under putative selection which may play a number of roles in the biological processes of bald eagles.

## Introduction

A large portion of the bald eagle (*Haliaeetus leucocephalus*) range, primarily east of the Mississippi River, underwent a severe population bottleneck caused by the use of Dichlorodiphenyltrichloroethane (DDT). From the 1940s until it was banned in 1972, DDT was heavily used in the U.S. as an insecticide (Bowerman *et al.* 1995; McEwan and Hirth 2012) . Fish contaminated with DDT were consumed by bald eagles and the DDT quickly metabolized into Dichlorodiphenyldichloroethylene (DDE), a stable and toxic chemical that caused egg shell-thinning, resulting in egg breakage and an increase in unfertile embryos (Newton 1979). Furthermore, DDE bioaccumulated in adipose tissue thereby allowing the chemical to be metabolized long after contact or ingestion (Newton 1979). DDT and other anthropogenic factors reduced the bald eagle population in the contiguous 48 states to ~ 417 nesting pairs, nearly causing the extirpation of the U.S. (excluding Alaska) population of bald eagles (Sprunt *et al.* 1963).

A population bottleneck has the ability to decrease the genetic variation in the gene pool with the resulting decrease in genetic variation being dependent upon the size of the population reduction and the length (in terms of generations) that the population remains reduced in size (Nei *et al.* 1975; Maruyama *et al.* 1985; Brown *et al.* 2007). The population that remains after the bottleneck can be impacted both demographically as well as genetically. Demographic impacts can lead to inbreeding; since the population has been reduced to small numbers, remaining individuals may breed to closely related

individuals. Inbreeding has been shown to lower genetic diversity and cause the loss of alleles. Moreover, even without inbreeding, the fraction of the population that survived the bottleneck likely does not have full representation of the genetic variation present prior to the bottleneck. Thus, these demographic and genetic factors work synergistically to reduce standing genetic variation in the population, thereby decreasing fitness and limiting the adaptive potential of the species. Although many species have undergone bottlenecks from natural or anthropogenic factors, the outcome of the resulting recoveries have been mixed with regards to impact on genetic variation. For example, a decline in genetic variation has been documented in species such as the Scandinavian wolf (*Canis lupis*, Seddon *et al.* 2005), Mauritius Kesterel (*Falco punctatus*, Ewing *et al.* 2008), Seychelles paradise flycatcher (*Terpsiphone corvina*, Bristol *et al.* 2013), and bearded vulture (*Gypaetus barbatus, L.*, Godoy *et al.* 2004). In contrast, population bottlenecks did not result in a large loss of genetic variation in the Eurasian black vultures (*Aegypius monachus*, Poulakakis *et al.* 2008), peregrine falcons (*Falco peregrinus*, Brown *et al.* 2007), or the Spanish imperial eagle (*Aquila adalberti*, Martinez-Cruz *et al.* 2004).

For conservation programs in which the cause of the population decline can be determined and either corrected or mitigated, reintroduction or translocation programs are becoming a common management tool (Armstrong & Seddon 2008) but with varied levels of success. Reintroductions and translocations have increased levels of genetic variation and reduced inbreeding in black bears (*Ursus americanus*; Van Den Bussche *et al.* 2009) and Florida panthers (*Felis concolor*; Johnson *et al.* 2010). However, for species such as the South Island Saddleback (*Philestrunus carunculatus carunculaturs*), Stewart Island Robin (*Petroica australis rakiura*), North Island Robin (*Petrocia*

*longipes*), the Takahe (*Porphyrio hochstetteri*) and the Mauritius kestrel, the reintroduction programs were not nearly as successful (Ewing *et al.* 2008; Jamieson 2010). The recovery of the U.S. population of bald eagles has been attributed in part to the large reintroduction/translocation program conducted from the 1970s through the early 2000s (Table 1). Unfortunately, because so few genetic studies have been conducted on bald eagles, not only is it unclear what impact the large population bottleneck had on standing levels of genetic variation but it is also unclear as to how many of these reintroduction/translocation events were successful. If the reintroductions were successful, it is unclear what impact they had not only on standing genetic variation but also the partitioning of genetic variation across the range of bald eagles.

Morizot *et al.* (1985) conducted the first biochemical genetic analysis of bald eagles from Alaska, Washington, Oregon, and Arizona by examining five presumptive allozymic loci and detected trends of a north-to south cline. Knight *et al.* (2013) used 32 allozymic loci to evaluate levels of genetic differentiation between hatchling eagles from Colorado and Ontario. They found virtually no allozymic genetic differentiation between the two groups despite the geographic distance. Unfortunately, neither of these studies were designed to assess the impact of the DDT-induced bottleneck or resulting translocations on intra and interpopulation genetic variation. As there are so few published genetic studies on bald eagles, studies on population genetic variation in white-tailed sea eagle (*H. albicilla*) may provide some insight into the current genetic state of bald eagles. White-tailed sea eagles are the sister species to bald eagles and their European ecological equivalent. White-tailed sea eagles also underwent a population decline due to DDT and the result of this bottleneck is better documented in the literature

than the bald eagle bottleneck. For example, Hailer *et al.* (2007) found highly supported phylogeographic structure throughout the Palearctic for white-tailed sea eagles when using mtDNA. Honnen *et al.* (2010) examined 100 European white-tailed sea eagles using microsatellites and mtDNA and documented a potential cline of genetic variation from north-west to south-east as well as high levels of genetic variation. Thus, although this area suffered high mortality rates from DDT, there is no genetic signature of the bottleneck.

It is believed that the bald eagle population is currently in exponential growth (B. Millsap pers. comm.). Unfortunately, there are several anthropogenic stressors that may cause harm to bald eagles especially when taking specific populations into consideration. For example, lead poisoning, wind farm placement, collision with powerlines, and environmental toxicants are all contributing to bald eagle mortality (Bowerman *et al.* 1995; Millsap *et al.* 2004; Watts *et al.* 2008; Stauber *et al.* 2010; Pagel *et al.* 2013; Mojica *et al.* 2016). To help combat these stressors and develop an impactful management plan, the U.S. Fish and Wildlife Service (USFWS) is working to determine biologically relevant management units for bald eagles (U.S. Fish and Wildlife 2016). Currently, USFWS is considering two management scenarios for implementation (U. S. Fish and Wildlife Service, 2016). The first is based on the North American Migratory Flyways (Atlantic, Mississippi, Central, and Pacific) with support for this management scenario coming from the observations of migratory movement for the species. The second is based on the average natal dispersal distances (Millsap *et al.* 2014) which closely correspond to the current USFWS regions. Unfortunately, the development of these two approaches did not consider genetic data, nor has genetic data been used thus

far to evaluate the validity of these scenarios. Therefore, it is unknown if either of these scenarios represent biologically relevant management units.

Based on behavioral, ecological, and morphological data, some researchers have proposed recognition of three subspecies of bald eagles: northern (*H. l. alascanus*), southern (*H. l. leucocephalus*) and Sonoran desert (Morizot *et al* 1985, Simmons *et al* 1998, K Jacobson pers. comm., B. Millsap pers. comm.). The northern and southern subspecies are primarily based on different migratory behavior and sizes between the two geographic areas. The northern subspecies migrate into the southern United States when food supplies diminish in the winter, whereas individuals ascribed to the southern subspecies do not exhibit this migratory behavior to the same degree (Simons *et al.* 1988). Furthermore, individuals ascribed to the northern subspecies are larger than their southern counterpart. In contrast to the northern and southern subspecies, the Sonoran desert bald eagle is adapted to the Sonoran desert environment which has required these individuals to change their diet and nesting sites relative to bald eagles throughout the remainder of North America. Finally, Sonoran desert bald eagles are even smaller than other southern bald eagles (K. Jackobson pers. comm.). While there does appear to be behavioral, ecological, and morphological reasons to argue for the three subspecies, no range-wide genetic study designed to test for these designations have been conducted.

This study represents the first range-wide population genomic analysis of bald eagles. To begin to address the questions outlined above related to genetic signatures of population bottlenecks and reintroductions, subspecies, management units, and adaptation, we created a 50K Axiom myDesign custom SNP array. We describe the design of the array's development and its use in evaluating the genomic structure of 171

bald eagles from throughout their range. We further use this array to determine the impacts of the bottleneck and translocation programs. Addressing these questions provides information necessary, and currently unavailable, for the development of biologically relevant management and conservation plans for the species.

## Materials and Methods

### *SNP Calling and Array Development*

To begin development for the SNP chip, two sequencing methods were utilized. The first was restriction-site associated sequencing (RAD-tag), which utilizes short fragments of DNA that are adjacent to restriction enzyme recognition sites. By only focusing on areas around the tags on either side of a restriction site, high-densities of SNPs can be obtained when individuals are multiplexed and barcoded (Pujolar *et al.* 2013). The second method, low coverage-whole genome sequencing, was utilized to increase the number of potential SNPs for the SNP array by including regions not associated with restriction sites. For both protocols, DNA was isolated using the protocol described in Longmire *et al.* (1997).

The RAD-tag sequencing approach utilized 200 bald eagle DNA samples (12ng/ul) (Alaska 13, Alberta 2, Arizona 4, Arkansas 2, California 14, Delaware 2, Florida 5, Illinois 1, Iowa 3, Kentucky 2, Louisiana 4, Maryland 31, Michigan 2, Minnesota 11, Nebraska 5, New Jersy 41, North Carolina 1, Nunavut 1, Northwest Territories 2, Oklahoma 18, Oregon 3, Pennsylvania 1, Saskatchewan 4, Utah 2, Wisconsin 22, Wyoming 4) that were sent to the Glenn Lab in Athens, GA where a double digestion protocol was performed (Hoffberg *et al.* 2016). All sequence libraries

for 3-RADseq were prepared using standard protocols and sequenced with 150 pair end reads on the Illumina Next-Seq platform. Illumina sequencing data was filtered for quality using standard Illumina protocols. RAD-tag data was processed using the RAD-tag bioinformatics suite in Stacks (Li 2011) where standard runtime settings using the recommended protocol with process_radtags with NdeI and Hind1 double restriction sites were utilized. The ustacks portion of the Stacks protocol was used to construct alleles and call SNPs using a maximum likelihood framework (Hohenlohe *et al.* 2010).

The whole genome sequencing utilized 28 bald eagle DNA samples subsampled from the 200 samples that were used in the RAD-tag sequencing (200ng) (Arkansas 2, Iowa 2, Michigan 2, Nebraska 5, Oklahoma 10, Oregon 3, Utah 1, Wisconsin 3) that were shipped to the McDonnell Genome Institute at Washington University, St. Louis, MO for DNA sequencing. All genomic libraries for whole genome sequencing were sequenced using the HiSeq 2500 platform. Illumina sequencing data was quality filtered using standard Illumina protocols.

All generated read data from both protocols were aligned to the bald eagle genome (www.ncbi.nlm.nih.gov/nuccore/NW_010972436.1) using the short read alignment program Bowtie2 (Langmead & Salzberg 2013). Standard seed and read alignment metrics were used as recommend by whole genome alignment protocols. SNPs were integrated from all 228 sequenced samples using the mpileup function of alignment modification software Samtools (Li 2011). A combination of Bcftool and Samtools were used to further filter all SNPs using a minimum coverage depth of 10x and a minimal phred quality value of 18. The initial VCF output was generated and formatted in

VCFTools (Danecek *et al.* 2011). Finally, VCFtools was also used to separate SNP calls from other variants such as repeat or indels.

To screen for highly variable SNPS suitable for chip production, PLINK (Purcell *et al.* 2007) was used to filter for a minor allele frequency of 0.05 (MAF), a minimal genotype frequency of 0.3, and a Hardy-Weinberg equilibrium (HWE) of 0.001. Loci that passed this set of criteria were thinned by examining all SNPs within a 10KB sliding window and retaining only a single SNP in each 10 KB window using operations in VCFTools. All SNPs that remained after this final filtering step were sent to Affymetrix for additional filtering to ensure they were appropriate for the array. These filtering steps included estimation of binding specificity and strength. The SNPs remaining after Affymetrix's quality control steps were annotated to the bald eagle genome from NCBI using SNPEff (Cingolani *et al.* 2012) to determine the severity of the SNP. We then prioritized SNPs that were in ecologically relevant genes, genes that were located in proximity to known genes, and SNPS in intergenic regions for the population genetic analyses.

*Sample Selection*

Blood samples (0.5 ml) from 171 bald eagle were obtained from permitted rehabilitators, veterinarians, and biologists and stored in lysis buffer (Longmire *et al.* 1997). Samples from Arizona (n = 28), Florida (n = 18), Illinois (n = 1), Iowa (n = 3), Maryland (n =10), Minnesota (n = 10), New Jersey (n = 24), Oklahoma (n = 19), Virginia (n = 15), Wisconsin (n = 22), and Wyoming (n = 5) were obtained from hatchling individuals with known natal locations. When samples for multiple hatchlings from the

same nest were received, only one individual was included in for analysis. Samples from Nunavut (n = 2), Ontario (n = 4), and Saskatchewan (n = 1) represent individuals that were fitted with GPS transmitters during their southern migration into the United States. Studies have determined that the typical median natal dispersal of bald eagles is 46-175 km (Millsap *et al*, 2014) and therefore we are assuming that the nesting location provides a close approximation to the natal location. Finally, samples from Alaska (n = 13) were adult birds that entered rehabilitation care during the summer breeding months, therefore these samples are assumed to be nesting individuals from Alaska.

Aliquots of whole blood were sent to our laboratory at Oklahoma State University where DNA was extracted following a standard protocol (Longmire *et al.* 1997). Extracted DNA was run on a 1% agarose gel to assess quality and then quantified using a NanoDrop 3300 spectrophotometer (Thermo Scientific). An aliquot of DNA was sent to Eurofins (River Falls, WI) where it was further assessed for quality and concentration for genotyping using our custom Axiom myDesign Bald Eagle SNP array. After receiving the raw data from Eurofins, we scored the samples using options in the Axiom Analysis Suite v2.0.0.35. All scored SNPs were filtered so that poorly clustered SNPs, SNPs with minor allele frequencies less than 0.01, and monomorphic SNPs were removed prior to downstream analyses.

Fifteen arbitrarily chosen individuals, representing 8.8% of the genotyped samples, were genotyped twice to calculate a genotyping error of the array. Each SNP call between the two independent runs was analyzed and two types of genotyping errors were identified. The first occurred when a SNP was a no call in one run while a base was called in the duplicate run. The second type of error occurred when two different bases

were called for the same SNP in the independent runs. The two error rates were calculated separately and subsequently combined for an overall error rate of the array.

*Population Structure Assessment*

STRUCTURE v 2.3.4 (Pritchard *et al.* 2000) uses a Bayesian clustering approach to calculate the posterior probabilities for the correct K while maximizing the cluster of individuals meeting the assumptions of Hardy-Weinberg equilibrium (HWE) and minimizing linkage disequilibrium. STRUCTURE was run using the admixture model and correlated allele frequencies for 10,000 burn-in iterations and 50,000 Markov Chain Monte Carlo (MCMC) iterations. The number of clusters (K) was analyzed from 1 to 15 using eight independent runs per cluster. Structure Harvester (Earl & vonHoldt 2012) was used to visualize the STRUCTURE output and to evaluate this output using the Evanno method (Evanno *et al.* 2005), thereby allowing the most appropriate K to be determined. Clumpp (Jakobsson & Rosenberg 2007) was subsequently used to determine the assignment probability of each individual to each cluster. The final results were visualized using Destruct (Rosenberg 2004).

As Bayesian clustering programs, such as STRUCUTRE, rely on a specific set of assumptions, problems can arise if these assumptions are violated. Factors that can cause these violations and can greatly impact the clustering results include: populations displaying patterns of isolation by distance (Ruiz-Gonzalez *et al.* 2015), small sample size (Evanno *et al.* 2005), and population subdivisions characterized by low $F_{st}$ (Latch *et al.* 2006). Therefore, to provide further clarity into the population structure, a Principal Coordinates Analysis (PCoA) was run in Adegenet (Jombart 2008) using program R 3.3.2.

*Relatedness and Population Genetics*

To determine the level of kinship between individuals in the dataset, we utilized the kinship option in the program KING (Manichaikul *et al.* 2010). Genepop 4.4.3 (Raymond & Rousset 1995) was used to test for HWE for each SNP within each cluster. Significance was determined using a sequential Bonferroni correction. Genepop 4.4.3 was also used to calculate the inbreeding coefficient for each cluster. Pairwise population $F_{st}$ values were calculated in Arlequin v3.5 (Excoffier & Lischer 2010) using a p-value of $< 0.05$ to indicate significance. Finally, the effective population size was calculated using the linkage disequilibrium option in NeEstimator v2.01 (Do *et al.* 2014) using a critical value of 0.05 and a monogamous mating system.

*Outlier Detection*

To detect SNPs under selection, three methods were implemented: Lositan (Antao *et al.* 2008), Arlequin v3.5, and BayeScan v2.1 (Foll & Gaggiotti 2008). Three methods were used as there are known issues with false positives in outlier detection analyses (Narum *et al.* 2013). To help combat these problems, we considered a SNP to be under putative selection if it was found to be an outlier in all three programs.

Lositan was run using the infinite alleles model with 100,000 simulations, a forced and neutral mean $F_{st}$, a FDR cut-off of 0.1, and a 0.99 CI criterion. Since Lositan can produce different results for the same dataset in different runs, Lositan was run three times and loci that were identified in all three runs were considered to be under selection. Arlequin was run using 50,000 simulations and 1,000 demes. SNPs with a p-value $< 0.01$ were considered to be under selection. Finally, BayeScan was run for 100,000 iterations

utilizing 20 pilot runs and 50,000 iterations for burn-in. A q-value of 0.1 was used as a cut of value for determining SNPs under selection.

SNPs that were identified as being under selection in all three programs were then analyzed in SNPEff. Gene functions and annotations were identified using Gene Ontogeny (GO) as implemented in Panther using the chicken (*Gallus gallus*) genome (Thomas *et al.* 2003; Mi *et al.* 2013) as a reference. A literature review was also conducted to determine the specific gene function in Aves.

## Results

### *SNP Chip Development and Sample Scoring*

The RAD-tag sequencing resulted in an average of 13,770,170 reads per individual with an average of 2,065,525,548 base pairs per individual. On average there were 37,000 stacks per individual prior to filtering. After filtering for HWE and MAF this was reduced to 728 stacks per individual. A total of 168,243 SNPs were generated for bald eagles utilizing the RAD-tag method. The whole genome sequencing resulted in 134,250,617 reads with an average of 2,397,332 reads per individual. Prior to filtering the whole genome sequencing approach generated 4,222,395 SNPs for bald eagles. After filtering for HWE, minimal genotype frequency, and HWE this number was reduced to 2,089,831.

The combined RAD-tag and whole sequencing SNP dataset was filtered for a quality score of >15 and a coverage of >20 resulting in 1,894,471 SNPs which were sent to Affymetrix for review. It was determined that 1,417,390 SNPs were highly recommended for the array by tiling the individual SNP utilizing the forward strand

individually, the reverse strand individually, or both strands. These highly recommended SNPs were thinned using a 10kb sliding window in which only one SNP was allowed in the window.

The final Axiom myDesign custom array consisted of 50,789 SNPs after all quality controls and filter analyses were conducted. These SNPs consisted of 17,105 genic regions and 33,684 intergenic regions. Blood samples for 171 bald eagles were genotyped using the Axiom Analysis Suite software where 2 samples (one Florida, one Arizona) did not pass the quality controls after being scored. These individuals were removed from all subsequent analyses. Genotypes of the remaining 169 samples revealed 1,358 monomorphic SNPs, 2,472 SNPs with a MAF lower than 0.01, and 273 SNPs that were poorly clustered. After these SNPs were removed 45,952 SNPs remained for downstream population genetic analyses.

The error rate for the array was determined by genotyping 15 arbitrarily chosen DNA samples (8.8%) twice. Genotyping errors where a base call in one run was different than the base call in the duplicate run occurred at an average rate of 0.29% and a median rate of 0.29%. Genotyping errors that occurred when a base was called in one run but a no call was called in the subsequent run occurred on average 0.64% and at a median rate of 0.61%. Therefore, the overall error rate for the 169 eagles was on average 0.93% with a median error rate of 0.90%.

*Population Structure*

STRUCTURE determined that the most appropriate number of clusters was five (Fig. 2). Two of the clusters described by STRUCTURE separated Alaska and Arizona

into their own distinct clusters. A third cluster containing individuals from Wyoming, Nunavut, Ontario, Oklahoma, Minnesota, Wisconsin, Illinois, and Iowa represented the upper Midwest. One individual in Arizona aligned more closely with this cluster than the Arizona cluster. The final two clusters were distributed mainly throughout the east coast representing Maryland, Florida, New Jersey, and Virginia. Individuals that fit more closely with east coast clusters were also found in Oklahoma and Arizona. The PCoA (Fig. 3) supported the overall findings of STRUCTURE. Alaska and Arizona were separated from the remainder of the samples in distinct groupings in the two lower quadrants. The other sampling locations are spread across the top two quadrants with the east coast samples found primarily in one quadrant and the upper Midwest samples being found the other. The Oklahoma samples are found as an intermediate group between the upper Midwest samples and the samples from the east coast, as also represented in the results from STRUCTURE.

Evaluating the results of STRUCTURE in light of the PCoA and also considering that observation that STRUCTURE has difficulty in situations with small sample sizes, subsequent population genetic analyses were conducted using seven groups for the following reasons. Based on both the PCoA and STRUCTURE, individuals from Alaska appear to represent a distinct genetic group with mean level of group membership for individuals belonging to this group being 97.75% (Table 2). The 27 individuals from Arizona comprise a second group with mean level of membership for this group comprising 10.07%, 76.47%, and 11.64% of Genomes 1, 3, and 5, respectively (Table 2). Although STRUCTURE appears to group the five individuals from Wyoming with individuals from the upper Midwest and Canada, closer evaluation of population means

reveals that the Wyoming individuals possess a different genetic characteristic of genomes than individuals from the upper Midwest and Canada and therefore were considered as a distinct genetic group (Table 2). We also consider the Oklahoma population as a discrete genetic entity for further analyses due to genomic characteristics reflecting both the translocations from Florida (Table 1) as well as gene flow from the upper Midwest (Table 2). Finally, we separate Florida from the remainder of the east coast samples. Our rationale is due to differences in genomic characteristics of these populations (Table 2) coupled with the fact that Florida samples were used for reintroductions along the east coast.

*Population Genetics and Kinship*

When grouping the samples into seven groups represented by Alaska, Arizona, east coast (Maryland, New Jersey, Virginia), Florida, Oklahoma, upper Midwest (Wisconsin, Minnesota, Nunavut, Ontario, Saskatchewan, Iowa, and Illinois), and Wyoming, all pairwise $F_{st}$ values were statistically significant (Table 3). Alaska has particularly high levels of genetic differentiation when compared to all other groups (0.11-0.2). Arizona also shows high levels of differentiation especially when compared to Wyoming (0.11) and Alaska (0.20). The observation that pairwise comparisons among these seven groups revealed moderate to high levels of statistically significant genetic differentiation, provides support for separating Wyoming from the upper Midwestern samples, Florida from the remaining east coast samples, and keeping the Oklahoma samples as a unique group. Based on these seven groups, none of the loci were out of HWE when using a Bonferroni correction. Finally, the inbreeding coefficients for the

seven populations ranged from 0.26-0.32 with Alaska and the upper mid-west being the lowest and highest, respectively (Table 4).

Of the 14,197 pairwise kinship values for the 169 bald eagles, 97% were determined to be unrelated (13,778). Of the remaining 419 relationships that showed some level of relatedness, 16 (3.8%) were found to be equivalent to full siblings, 28 (6.68%) were equivalent to half siblings, 125 (29.8%) were equal to first cousins, and 249 (59.4%) were equivalent to second cousins (Table 5). Of the samples that showed second cousin or greater relationships (>0.01), 224 were located within the same state while 194 were located in different states. For the related comparisons that had individuals from different states, 111 were from different genetic groups when the samples were arranged into the seven groupings.

The effective population size was calculated for six clusters. Wyoming was not included as the sample size was too small for accurate analysis. The results showed an effective population size between 52-2,876 individuals with Arizona and Florida possessing the lowest effective population sizes and Alaska maintaining the highest effective population size (Table 6).

*SNPs Under Putative Selection*

Lositan, Bayescan, and Arlequin were run using the seven populations as defined by the pairwise $F_{st}$. Lositan identified 1,440 SNPs that were identified in each of the three runs conducted ($F_{st}$ 0.78-0.03). Arlequin identified 1,900 SNPs under putative selection ($F_{st}$ 0-0.82). Finally, Bayescan identified 711 SNPs under putative selection ($F_{st}$ 0.18-0.34). Taking a conservative approach and only considering those SNPs identified by all

three programs resulted in 379 SNPs under putative selection. Of these SNPs, 283 were located in intergenic regions, two were 3' UTR variants, one was a 5' UTR variant, 13 were downstream gene variants, 64 were intron variants, three were splice region variants, three were synonymous variants, and nine were upstream gene variants. It was determined that 247 SNPs in the intergenic region were associated with genes or gene regions. The coding SNPs were associated with 84 genes or gene regions that were further analyzed using Panther and a literature review. A complete list of biological processes and molecular function as determined by Panther can be found in Supplementary Table 2 while the Panther GO slims can be found in Figure 4.

**Discussion**

This study utilized a combination of RAD-tag sequencing and full genome sequencing to develop a custom 50K Axiom myDesign array to provide the first range-wide genomic analysis of bald eagles. The array proved to be successful as 99% of the eagle samples analyzed and 92% of the SNPs included on the array were successfully scored with low genotyping error (0.9%). Due to the large population bottleneck experienced by bald eagles coupled with the translocations and the fact that limited genetic work has been conducted on bald eagles, we were uncertain as to levels of genetic variation and population differentiation that we would be able to detect. Thus, our sampling scheme was designed to maximize the chances of detecting genetic differentiation. We included samples from regions of the continent not impacted by the bottleneck (Alaska, Arizona, Canada, Florida, Maryland, Minnesota, Virginia, Wisconsin, and Wyoming) as well as samples from areas impacted by the bottleneck and

subsequent translocations/reintroductions (New Jersey, Oklahoma). Despite a large bottleneck occurring within the bald eagle population, the results of this study clearly show that genetic structure exists as well as genetic footprints from the reintroduction efforts. Finally, a preliminary list of SNPs under putative selection were identified.

*Population Structure*

Considering the results of the STRUCTURE, PCoA, and population genetic analyses, seven highly supported distinct groups of bald eagles were identified. When coupling the genetic results with the reintroduction and life histories of eagles within these seven regions, a greater clarity for the genetic patterns emerge. Four of the identified groups were source populations that were not as significantly impacted by DDT and displayed very little admixture. The first unique group consists of bald eagles from Alaska and this is the most genetically homogeneous group showing little admixture from other groups (Table 2). Furthermore, the Alaska group represents a discrete cluster on the PCoA and the highest pairwise $F_{st}$ values when compared to the other six groups (0.11-0.20). Although our sample size for the Alaskan population is low, this grouping had the highest effective population size of the seven groups adding further support for Alaskan bald eagles not being significantly impacted by the DDT-induced bottleneck (Table 6). While the inbreeding coefficient for Alaska was high, it was the lowest of any of the other clusters (Table 4).

The second group represents the Arizona sampling location with most of the samples being from the Sonoran Desert population. Genetic composition of this group of eagles is 76.47% Genome 3, which has highest representation in the Arizona birds, ~10%

Genome 1, which is shared with eastern North America and Oklahoma, and 11.64% Genome 5 which is characteristic of the upper Midwest (Table 2). All but three of these eagles grouped tightly within a single quadrant of the PCoA (Fig. 2). As the Sonoran Desert population has been extensively studied, the histories of these three outlier eagles are known. The individual that was placed most closely with the upper Midwest group has a breeding location that is in the Arizona portion of the four corners region 125 miles away from the Sonoran Desert population (K. Jacobson pers. comm.). The two other individuals that are characterized by higher amounts of admixture with the upper Midwest cluster were offspring from a male banded in Texas that dispersed into the area and began breeding with a Sonoran Desert female (K. Jacobson pers. comm.). Therefore, the genetic placement of these eagles make sense and the admixture observed in the population can be attributed to random dispersing individuals entering the area. Removing these three individuals from the population calculation (Table 2) increases representation of Genome 3 for the remaining Arizona bald eagles to 84.39% (+/- 13.12). Although this grouping of bald eagles was not impacted by DDT, it has one of the lowest effective population sizes and high levels of inbreeding among the seven groups we detected (Table 4, 6). The low effective population size and high levels of inbreeding provides additional support that Arizona bald eagles receive little introgression from other populations of bald eagles and represent a distinct genetic entity. Future studies clearly need to add bald eagles from Texas, California, Colorado, New Mexico, Texas, and Utah to gain a clearer understanding of the genetic characteristics and partitioning of genetic variation among southwestern bald eagles.

Birds from Wyoming represent an area of the range that was not significantly impacted by the DDT-induced bottleneck. Although sample sizes for this population are low, they do share genetic characteristics with bald eagles from Alaska as well as those from the upper Midwest (Table 2). They also show high levels of inbreeding when compared to the other clusters (Table 4). Future studies should increase sample sizes from this area to determine the genetic uniqueness of this population as well as to better understand what role these individuals may have played in the recovery of bald eagles.

The upper Midwest region consists of birds from Wisconsin, Minnesota, Nunavut, Ontario, Saskatchewan, Iowa, and Illinois with 83.55% of the population genome being characterized by Genome 5 (Table 2). Within this population there appears to be low, but approximately equal contributions of admixture from the other four genomes detected by STRUCTURE (Table 2). Finally, closer examination of individuals from Nunavut, Ontario, and Saskatchewan indicate that these Canadian bald eagles may represent a separate genetic entity but due to logistical and permitting difficulties, we were not able to obtain sufficient samples to evaluate their genetic distinction. Bald eagles represented by this group were minimally impacted by the DDT-induced bottleneck and this is reflective in this group possessing the second highest effective population size (Table 6), although this group did exhibit the highest level of inbreeding when compared to the other seven clusters (Table 4). Future studies should include bald eagles from across Canada to provide more insight into the genetic structure across North America. Including additional samples from across Canada would make it possible to determine partitioning of genetic variation not only between Canada and the upper Midwestern states but also across Canada. Such an analysis would also provide better estimates of the

effective population size for this area of bald eagle range and could provide explanation to the high inbreeding coefficient.

Finally, although no reintroductions into Florida were performed, these samples were made up of two of the genomes identified by STRUCTURE (Fig. 2). All Florida individuals were characterized as possessing almost solely one of these two clusters or were characterized as possessing close to half of each genome. Furthermore, the pairwise $F_{st}$ showed low, but statistically significant levels of population differentiation between Florida and the other east coast samples (Table 3). Further examination of the individuals revealed no correlations between geography or geographic distance between the sampling locations of these individuals. Finally, even though bald eagles from Florida were used as a source for translocations along the east coast, the samples from Florida possessed the lowest effective population size and exhibited high levels of inbreeding (Table 4, 6). Additional sampling is needed from Florida and other southeastern areas to determine what is driving the diverse genetic variation, but low effective population size and high inbreeding coefficient in Florida.

The remaining two groups were populations in which reintroduction attempts were conducted. The east coast group consisting of New Jersey, Maryland, and Virginia was subject to reintroduction mainly in the New Jersey area. Interestingly, the two primary genotypes characterizing these individuals are the same genotypes that are found in Florida. Although there were reintroductions from the upper Midwest and Alaska into these areas (Table 1) there is little genetic signature left of these individuals in the hatchlings sampled in this study (Table 2, Fig 1). This may be a sign that there were selection pressures against these translocated individuals for currently unknown reasons.

The inbreeding coefficient was high compared to the other clusters of bald eagles, but that is to be expected for areas where the reintroductions occurred (Table 4). Finally, the effective population size for these east coast bald eagles is substantially higher than their source population but still considerably lower than areas of the country that were not significantly impacted by DDT (Table 6).

Results for Oklahoma revealed a diverse grouping of eagles that included individuals that are more similar to the east coast and Florida groups as well as the upper Midwest group. As eagles were reintroduced into Oklahoma from Florida, it would be assumed that the east coast genotypes are from these events. Prior to 1940 and the creation of manmade lakes, Oklahoma did not have a population of bald eagles. As manmade lakes were built creating viable habitat, individuals from the north moved into the new habitat (Lish and Sherrod 2007). As several of the Oklahoma eagles clustered closely with the upper Midwest genotypes, we assume that these northern individuals dispersed into the state after the DDT bottleneck as they did when the population was first established. Even though Oklahoma bald eagles possess high levels of genetic diversity, this population is characterized by one of the lower effective population sizes and higher inbreeding coefficients (Table 4.6).

Throughout Oklahoma, Florida, and the east coast we see patterns from the genomes represented by Genome 1 and 2 (Table 2). We believe that Genome 1 represents the historical Florida samples that were reintroduced into various parts of the United States and the Genome 2 represents Canadian individuals whose natal locations are not represented in this study that because of reintroductions have allowed for gene flow or individuals to disperse throughout the east coast. This is supported by several

observations about the dataset. First, samples representing Florida in this dataset are primarily Genome 1(55.55%), with Genome 2 only making up 28.49% (Table 2). Next, Oklahoma has several individuals that are represented by a majority of Genome 1, thereby revealing the genetic footprints of the reintroductions from Florida. Although, Oklahoma has trace amounts of genome 2 present, it is also present in all other groups. Furthermore, New Jersey, an area that had reintroductions from Nova Scotia and Manitoba, is primarily Genome 2 as shown in the STRUCTURE analyses (Figure 1). Finally, Virginia and Maryland, two populations that were not heavily involved in reintroduction practices, are primarily Genome 1. More samples from Canada and throughout the eastern United States are warranted to confirm this hypothesis.

*SNPs Under Selection*

Taking a conservative approach to help eliminate problems associated with false positives (Foll & Gaggiotti 2008), we identified 379 SNPs under putative selection and found regions associated with these SNPs play a role in a variety of biological pathways and molecular functions (Fig 4). Nie *et al.* (2016) found that PAM, RNF38, SLCO4C1, ST8SIA4, AND TRAF5 were all associated with a chicken's earlobe color and may be associated with egg color. Since these genes are believed to have an effect on egg formation, a problem also associated with DDT, the genes associated with these SNPs may be revealing adaptive differences in eagles that were and were not affected by DDT. Furthermore, there are noticeable phenotypic differences in bald eagles head coloration, specifically around where the chicken earlobe is located. While all bald eagles have a white head by the age of five, the amount of white varies for some individuals as some

individuals having a more “dirty” appearance on the side of the head with some residual browns and some individuals have a completely pure white head. As these genes seem to have effects on the coloration present in chickens in this location, the differences in bald eagles may also be under selection. TRAF5 and TBCD were both associated with exposure to avian influenza in chickens, H9N2 (Degen *et al.* 2006; Reemers *et al.* 2009) and H5N1 (Balasubramaniam *et al.* 2012), respectively. As the different strains of avian influenza are only in certain parts of the country, these SNPs may be driven by exposure to the virus. ADIPOR2 was found to be tied with fasting, in which a significant decrease in gene regulation was documented in chickens when a fast over 48 hours occurred (Maddineni *et al.* 2017). BAMBI was documented as having an important role in face morphogenesis (Higashihori *et al.* 2008) and tip formation in the wing (Casanova *et al.* 2012). ACSS1, SCD5, SLC44A1 were found to have roles in metabolism (Castro *et al.* 2012; Du *et al.* 2017; Zhang *et al.* 2017). Finally, FBN2, NODAL, TAL2, and TRABD2A had various roles in development (Burke *et al.* 2000; Schlange *et al.* 2002; Ferran *et al.* 2009; Merchán *et al.* 2011; Reis *et al.* 2014). As this study only aimed to develop a preliminary suite of SNPs under putative selection and their functions, additional studies are warranted to identify the effects of the SNPs in their various pathways.

*Conservation Implications and Future Research*

The USFWS is currently working to create and refine a comprehensive adaptive management plan for bald eagles with their proposed management units revolving around either the USFWS regions or the migratory flyways (U.S. Fish and Wildlife 2016).

Results of our study do not support either of these two management scenarios. As the King analyses revealed that there were 111 related pairwise comparisons there had individuals from different groups, part of the reason that these scenarios may not be reflected in the genetic data could be due to the bottleneck and reintroductions. As there is no information about the genetic structure prior to the bottleneck and reintroductions, these events may have caused deviations from a historical population structure that would have supported one of these scenarios.

When considering the potential northern, southern, and Sonoran Desert subspecies of bald eagles, our results were able to provide some insights. Our data highly supports that the Sonoran Desert eagles are genetically different from all other bald eagles in our sample set. We suggest that a more thorough sampling scheme of the southwestern states be conducted so differentiation between the Sonoran desert eagles and other bald eagles in the southwest can be evaluated. In evaluating the proposed northern and southern subspecies, we were not able to identify units that would separate the population into a solely northern and a southern subspecies. While there are clear differences in the genetics of eagles over geographic distance, there is no geographic boundary that would divide our sample set into a northern and southern group. Rather we identified five unique clusters of bald eagles that can be further broke down into seven unique groups that correspond to different areas of the country, especially when taking the reintroduction efforts into account.

Overall, for three of the six genetic units the effective population sizes are large enough to avoid inbreeding depression, although the current inbreeding coefficients for all units are high. One factor that could be driving the high overall inbreeding coefficients

is the low dispersal distances from natal locations observed for bald eagles (Millsap *et al.* 2014) in addition to the population bottleneck. More analyses should be conducted as the most inbreed population in this study was the upper mid-west an area that did not suffer an extreme population decline. Generally, the groups with the larger effective population sizes represent populations that were not as impacted by the bottleneck. Interestingly, while Florida was a source population and not heavily impacted by the bottleneck, this group had the lowest effective population size. This could have been caused by unknown effects of taking individuals for the reintroductions or undocumented population declines as a result of human encroachment. Arizona also showed a low effective population size, however this is to be expected with the small census size for the population. However, additional sampling needs to be conducted in the southwestern U.S. as well as states surrounding Oklahoma to better determine the genetic uniqueness of the Arizona and Oklahoma populations as well as to obtain better estimates of the effective population sizes for these areas.

This study provides the first insight into the population structure of bald eagles. Despite the population bottleneck, we identified clear genetic structure and selection occurring within the population. As the substructure identified in our results does not support the currently proposed management scenarios, caution is warranted for managing eagles in these manors until more samples can be utilized to identify the exact boundaries for these newly identified units. Furthermore, the source location should also be considered for any future translocation attempts as to help preserve the genetic structure observed in this study and to help ensure survival of the individuals as our data suggests there may have been selection against individuals translocated in previous attempts.

**Acknowledgements**

We would like to thank, Bryan Bedrosian, Brian Millsap, Erica Miller, Kenneth Jacobson, Russ Horton, Todd Katzner, Victor Roubidoux, and Victoria Vosberg for contributing the bald eagle samples used in this study. Without their generosity, the breadth of this study would not have been possible. We would also like to the Iowa Tribe of Oklahoma and Eppley Foundation for providing the necessary funding for this project as well as the Ford Foundation for funding Megan Judkins' research fellowship while the project was being conducted. A special thanks to all of the members of Megan's dissertation committee, Drs. Meredith Hamilton, Andrew Doust, and Jim Lish as their efforts on previous versions of the manuscript were invaluable.

**Literature Cited**

Antao T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G (2008) LOSITAN: A workbench to detect molecular adaptation based on a Fst-outlier method. *BMC Bioinformatics*, 9, 323.

Armstrong DP, Seddon PJ (2008) Directions in reintroduction biology. *Trends in Ecology and Evolution*, 23, 20–25.

Balasubramaniam VRMT, Wai TH, Omar AR, Othman I, Hassan SS (2012) Cellular transcripts of chicken brain tissues in response to H5N1 and Newcastle disease virus infection, *Virology Journal*, 9, 1-12.

Bowerman WW, Giesy JP, Best DA, Kramer VJ (1995) A review of factors affecting productivity of bald eagles in the great lakes region: Implications for recovery. *Environmental Health Perspectives*, 103, 51–59.

Bristol RM, Tucker R, Dawson DA, Horsburgh G, Prys-Jones RP, Frantz AC, Krupa A, Shah NJ, Burke T, Groombridge JJ (2013) Comparison of historical bottleneck effects and genetic consequences of re-introduction in a critically endangered island passerine. *Molecular Ecology*, 22, 4644–4662.

Brown JW, Van Coeverden De Groot PJ, Birt TP, Seutin G, Boag, PT, Friesen VL (2007) Appraisal of the consequences of the DDT-induced bottleneck on the level and geographic distribution of neutral genetic variation in Canadian peregrine falcons, Falco peregrinus. *Molecular Ecology*, 16, 327–343.

Burke RD, Wang D, Mark S (2000) Distribution of fibrillin I in extracellular matrix and epithelia during early development of avian embryos. *Anatomy and Embryology*, 201, 317–326.

Casanova J, Badia-Careaga C, Uribe V, Sanz-Ezquerro J (2012) Bambi and Sp8 expression mark digit tips and their absence shows that chick wing digits 2 and 3 are truncated, 7, 1–11.

Castro LFC, Lopes-marques M, Wilson JM, Roch E, Reis-Henriques MA, Santos MM, Cunha I (2012) A novel Acetyl-CoA synthetase short-chain subfamily member 1 ( Acss1 ) gene indicates a dynamic history of paralogue retention and loss in vertebrates, *Gene*, 497, 249–255.

Cingolani P, Platts A, Wang LL *et al.* (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118 ; iso-2; iso-3, *Fly*, 6, 1–13.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R, 1000 Genomes Project Analysis Group (2011) The variant call format and VCFtools.,*Bioinformatics*, 27, 2156–2158.

Degen WGJ, Smith J, Simmelink B, Glass EJ, Burt DW, Schijns VEJC (2006) Molecular immunophenotyping of lungs and spleens in naive and vaccinated chickens early after pulmonary avian influenza A (H9N2) virus infection, *Vaccine*, 24, 6096–6109.

Do C, Waples RS, Peel D, Macbeth GM, Tillett BJ, Ovenden JR (2014) NeEstimator v2: Re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data, *Molecular Ecology Resources*, 14, 209–214.

Du YF, Ding QL, Li YM, Fang WR (2017) Identification of differentially expressed genes and pathways for myofiber characteristics in soleus muscles between chicken breeds differing in meat quality, *Animal Biotechnology*, 28, 83–93.

Earl DA, VonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4, 359–361.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, 14, 2611–2620.

Ewing SR, Nager RG, Nicoll MAC, Aumjaud A, Jones CG, Keller LF (2008) Inbreeding and loss of genetic variation in a reintroduced population of Mauritius Kestrel, *Conservation Biology*, 22, 395–404.

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10, 564–567.

Ferran JL, de Oliveira ED, Merchán P, Sandoval JE, Sanchez-Arrones L, Martinez-de-La-Torre M, Puelles L (2009) Genoarchitectonic profile of developing nuclear groups in the chicken pretectum. *Journal of Comparative Neurology*, 517, 405–451.

Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective, *Genetics*, 180, 977–993.

Godoy JA, Negro JJ, Hiraldo F, Donázar JA (2004) Phylogeography, genetic structure and diversity in the endangered bearded vulture as revealed by mitochondrial DNA. *Molecular Ecology*, 13, 371–390.

Hailer F, Helander B, Folkestad AO, Ganusevich SA, Garstad S, Hauff P, Koren C, Masterov VB, Nygard T, Rudnick JA, Shiraki S, Skarphedinsson K, Volke V, Wille

F, Vila C (2007) Phylogeography of the white-tailed eagle, a generalist with large dispersal capacity. *Journal of Biogeography*, 34, 1193–1206.

Higashihori N, Song Y, Richman JM (2008) Expression and regulation of the decoy bone morphogenetic protein receptor BAMBI in the developing avian face, *Developmental Dynamics*, 237, 1500–1508.

Hoffberg SL, Kieran TJ, Catchen JM, Devault A, Faircloth BC, Mauricio R, Glenn TC (2016) RADcap: Sequence capture of dual-digest RADseq libraries with identifiable duplicates and reduced missing data, *Molecular Ecology Resources,*16, 1264-1278.

Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson EA, Cresko WA (2010) Population genomics of parallel adaptation in threespine stickleback using sequenced RAD Tags, *PLoS Genetics*, 6, 1-23.

Honnen AC, Hailer F, Kenntner N, Literak I, Dubska L, Zachos, FE (2010) Mitochondrial DNA and nuclear microsatellites reveal high diversity and genetic structure in an avian top predator, the white-tailed sea eagle, in central Europe. *Biological Journal of the Linnean Society*, 99, 727–737.

Jakobsson M, Rosenberg NA (2007) CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, 23, 1801–1806.

Jamieson IG (2011) Founder effects, inbreeding, and loss of genetic diversity in four avian reintroduction programs, *Conservation Biology*, 25:115 - 123

Johnson JA, Talbot SL, Sage GK, Burnham KK, Brown JW, Maechtle TL, Seegar WS, Yates MA, Anderson B, Mindell DP (2010) The use of genetics for the management of a recovering population: temporal assessment of migratory

peregrine falcons in North America. *PLoS ONE*, 5, 1-15.

Jombart T (2008) Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24, 1403–1405.

Knight RL, Craig GR, Smith MH, Grier JW, Mclean G (2013) Genetic variation and nesting bald eagles, *Cooper Ornithological Society*, 97, 282–283.

Langmead B, Salzberg S (2013) Fast gapped-read alignment with Bowtie2. *Nature methods*, 9, 357–359.

Latch EK, Dharmarajan G, Glaubitz JC, Rhodes OE (2006) Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation, *Conservation Genetics*, 7, 295–302.

Li H (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, 27, 2987–2993.

Lish JW, Sherrod SK (2007) A History of Bald Eagle Nesting Activity in Oklahoma, *Proceedings of the Oklahoma Academy of Science*, 20, 15–20.

Longmire JL, Maltbie M, Baker RJ (1997) Use of "lysis buffer" in DNA isolation and its implication for museum collections, *Occasional Papers of the Museum of Texas Tech University*, 163, 1–3.

Maddineni S, Metzger S, Oco O, Iii GH, Ramachandran R (2017) Adiponectin gene is expressed in multiple tissues in the chicken: Food deprivation influences adiponectin messenger ribonucleic acid expression, *Endocrinology*, 146, 4250–4256.

Manichaikul A, Mychaleckyj JC, Rich SS, Kaly K, Sale M, Chen WM (2010) Robust

relationship inference in genome-wide association studies, *Bioinformatics*, 26, 2867–2873.

Martinez-Cruz B, Godoy JA, Negro JJ (2004) Population genetics after fragmentation: the case of the endangered Spanish imperial eagle, *Molecular Ecology*, 13, 2243–2255.

Maruyama T, Fuerst PA (1985) Population bottlenecks and nonequilibrium models in population genetics. II. Number of alleles in a small population that was formed by a recent bottleneck, *Genetics*, 111, 675-689.

McEwan L, Hirth DH (2012) Southern bald eagle productivity and nest site selection. *The Journal of Wildlife Management*, 43, 585–594.

Merchán P, Bardet SM, Puelles L, Ferran JL (2011) Comparison of pretectal genoarchitectonic pattern between quail and chicken embryos. *Frontiers in Neuroanatomy*, 5, 1–19.

Mi H, Muruganujan A, Thomas PD (2013) PANTHER in 2013: Modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Research*, 41, 377–386.

Millsap B, Breen T, McConnell E, Steffer T, Phillips L, Douglass N, Taylor S (2004) Comparative fecundity and survival of bald eagles fledged from suburban and rural natal areas in Florida. *Journal of Wildlife Management*, 68, 1018–1031.

Millsap BA, Harmata AR, Stahlecker DW, Mikesic DG (2014) Natal Dispersal Distance of Bald and Golden Eagles Originating in the Coterminous United States as Inferred from Band Encounters, *Journal of Raptor Research*, 48, 13–23.

Mojica EK, Watts BD, Turrin CL (2016) Utilization Probability Map for Migrating Bald

Eagles in Northeastern North America: A Tool for Siting Wind Energy Facilities and Other Flight Hazards, *Plos One*, 11, 1-11.

Morizot DC, Anthony RG, Grubb TG, Hoffman SW, Schmidt ME, Ferrell RE (1985) Clinal genetic variation at enzyme loci in bald eagles (Haliaeetus leucocephalus) from the western United States, *Biochemical genetics*, 23, 337–45.

Narum SR, Buerkle CA, Davey JW, Miller MR, Hohenlohe PA (2013) Genotyping-by-sequencing in ecological and conservation genomics, *Molecular Ecology*, 22, 2841–2847.

Nei M, Maruyama T, Chakraborty R (1975) The Bottleneck Effect and Genetic Variability in Populations, *Evolution*, 29, 1–10.

Newton, I (1979) Population ecology ofraptors. Buteo Books, Vermillion, SD

Nie C, Zhang Z, Zheng J, Sun H, Ning Z, Xu G, Yang N, Qu L (2016) Genome-wide association study revealed genomic regions related to white/red earlobe color trait in the Rhode Island Red chickens, *BMC Genetics*, 17, 115.

Pagel JE, Kritz KJ, Millsap BA, Murphy RK, Kershner EL, Covington S (2013) Bald eagle and golden eagle mortalities at wind energy facilities in the contiguous United States. *Journal of Raptor Research*, 47, 311–315.

Poulakakis N, Antoniou A, Mantziou G, Parmakelis A, Skartsi T, Vasilakis D, Elorriaga J, De La Puente J, Gavashelishvili A, Ghasabyan M, Katzner T, McGrady M, Batbayar N, Fuller M, Natsagdorj T (2008) Population structure, diversity, and phylogeography in the near-threatened Eurasian black vultures, *Natural History*, 95, 859–872.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using

multilocus genotype data, *Genetics*, 155, 945–959.

Pujolar JM, Jacobsen MW, Frydenberg J, Larsen PF, Maes GE, Zane L, Jian JB, Cheng L, Hansen MM (2013) A resource of genome-wide single-nucleotide polymorphisms generated by RAD tag sequencing in the critically endangered European eel. *Molecular Ecology Resources*, 13, 706–714.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, BAkker PIW, Daly MJ, Sham PC (2007) Report Plink: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses, 81, 559–575.

Raymond M, Rousset F (1995) Genepop (Version-1.2) - Population-Genetics Software for Exact Tests and Ecumenicism, *Journal of Heredity*, 86, 248–249.

Reemers SS, Van Haarlem DA, Groot Koerkramp MJ, Vervelde L (2009) Differential gene-expression and host-response profiles against avian influenza virus within the chicken lung due to anatomy and airflow, *Journal of General Virology*, 90, 2134–2146.

*Reis AH*, Macdonald BT, Feistel K, Brito JM, Amado NG, Xu C, Abreu JG, He X (2014) Expression and evolution of the Tiki1 and Tiki2 genes in vertebrates, *International Journal of Developmental Biology*, 58, 355–362.

Rosenberg NA (2004) DISTRUCT: A program for the graphical display of population structure, *Molecular Ecology Notes*, 4, 137–138.

Ruiz-Gonzalez A, Cushman SA, Madeira MJ, Randi E, Gómez-Moliner BJ (2015) Isolation by distance, resistance and/or clusters? Lessons learned from a forest-dwelling carnivore inhabiting a heterogeneous landscape, *Molecular Ecology*, 24, 5110-5129.

Schlange T, Arnold H, Brand T (2002) BMP2 is a positive regulator of Nodal signaling during left-right axis formation in the chicken embryo, *Development*, 129, 31421–3429.

Seddon JM, Parker HG, Ostrander EA, Ellegren H (2005) SNPs in ecological and conservation studies: A test in the Scandinavian wolf population, *Molecular Ecology*, 14, 503–511.

Simons T, Sherrod SK, Collopy MW, Jenkins MA (1988) Restoring the Bald Eagle. *American Scientist*, 76, 253–260.

Sprunt A ,Robertson WB, Postupalsky S ,Hensel RJ, Knoder CF, Ligas FJ (1973) Comparative productivity of six bald eagle populations, Tram. North American Wildlife Conference,38:96-10

Stauber E, Finch N, Talcott PA, Gay JM (2010) Lead poisoning of bald (Haliaeetus leucocephalus) and golden (Aquila chrysaetos) eagles in the U.S. inland Pacific northwest region--An 18-year retrospective study: 1991-2008, *Journal of avian medicine and surgery*, 24, 279–287.

Thomas PD, Campbell MJ, Kejariwal A, Mi H, Karlak B (2003) PANTHER: A Library of Protein Families and Subfamilies Indexed by Function, *Genome Research*, 13, 2129–2141.

U.S. Fish and Wildlife Service. 2016. Bald and Golden Eagles: Population demographics and

estimation of sustainable take in the United States, 2016 update. Division of Migratory Bird Management, Washington D.C., USA.

Van Den Bussche RA, Onorato DP, Hellgren EC, Gardner-Santana LC, McKinney BR,

Villalobos JD, Lack JB (2009). Mitochondrial DNA phylogeography ob black bears (*Ursus americanus*) in central and southern North America: conservation implications. *Journal of Mammalogy*, 90:1075 - 1082.

Watts BD, Therres GD, Byrd MA (2008) Recovery of the Chesapeake Bay Bald Eagle Nesting Population. *The Journal of Wildlife Management*, 72, 152–158.

Zhang M, Yan F-B, Li F *et al.* (2017) Genome-wide DNA methylation profiles reveal novel candidate genes associated with meat quality at different age stages in hens, *Scientific Reports*, 7, 1-8.

**Figure Legends**

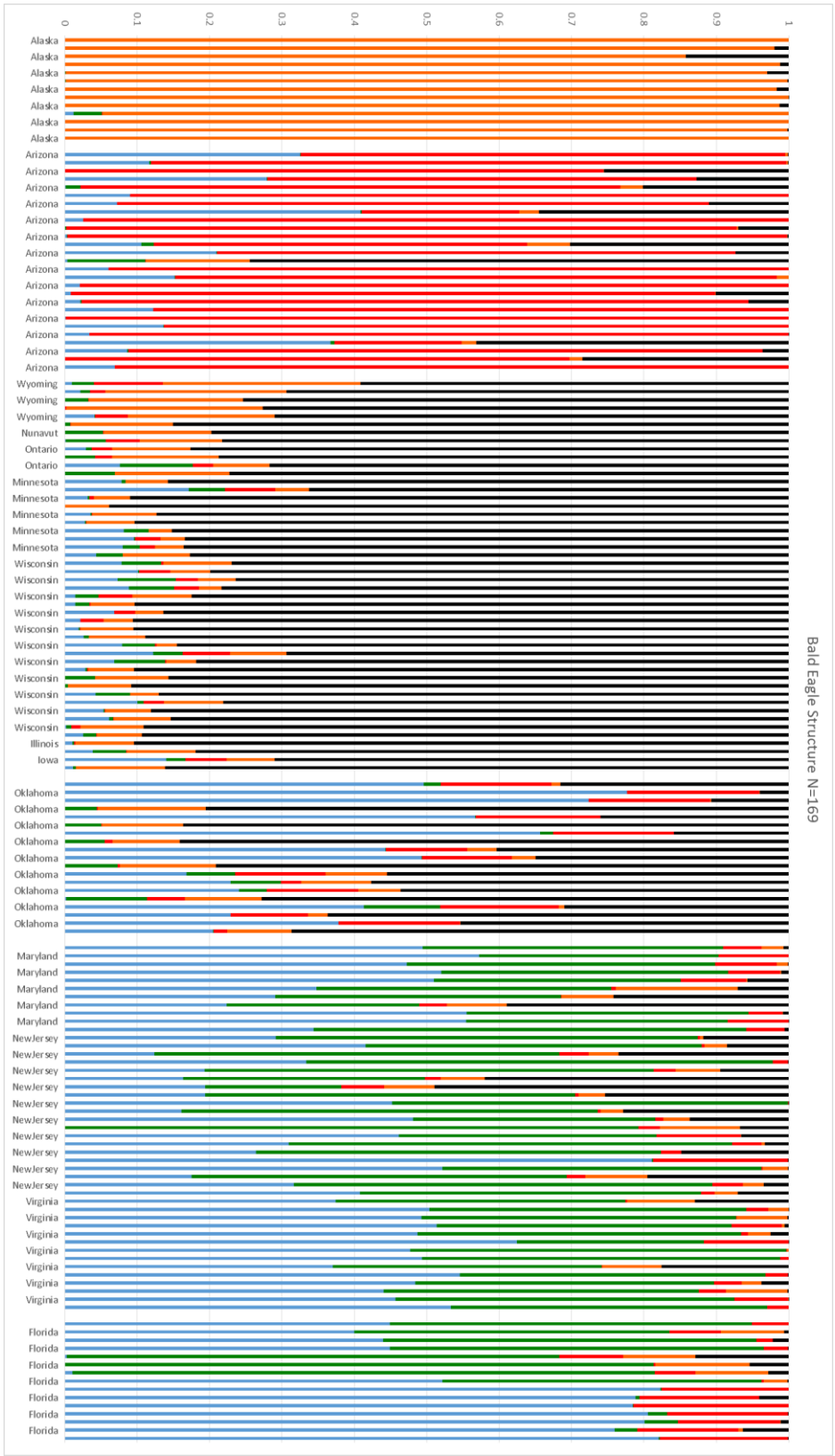
Figure 1 Map displaying the sample numbers for the states and provinces used in this study.

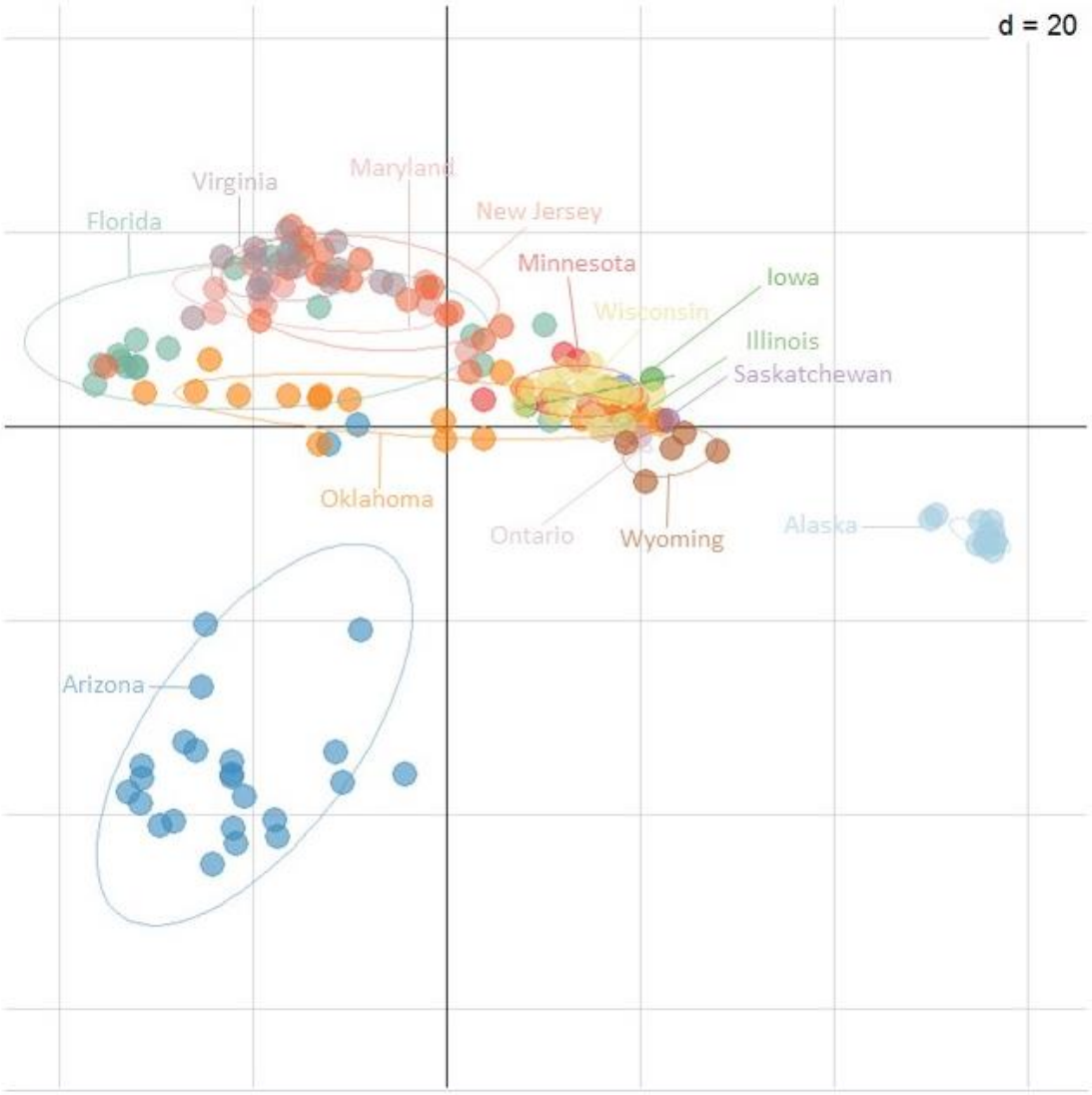
Figure 2 STRUCTURE results for 45,952 single nucleotide polymorphisms and 169 eagles. Figure displays the results of CLUMPP in which 8 runs of K equaled 5. The orange represents the Alaska cluster, the red represents the Arizona cluster, the black represents the upper Midwest cluster, and the green and the blue represent the east coast clusters.

Figure 3 Principle coordinates analysis run in Adagent in program R using 45,952 single nucleotide polymorphisms and 169 individuals. Colors represent the state or province believed to be the natal location of the sampled individual. Because of the high amount of clustering in the upper right quadrant, Nunavut is not able to be seen.

Figure 4 Panther GO Slims for the 84 genes associated with the single nucleotide polymorphisms under selection in coding regions. Genes are categorized by Molecular Function and Biological Process.

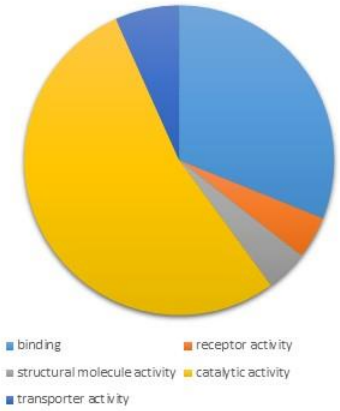

13
2
1
4
10
22
5
3
10
1
24
15
28
19
18


0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Alaska
Alaska
Alaska
Alaska
Alaska
Alaska
Alaska
Arizona
Arizona
Arizona
Arizona
Arizona
Arizona
Arizona
Arizona
Arizona
Arizona
Arizona
Arizona
Arizona
Arizona
Wyoming
Wyoming
Wyoming
Nunavut
Ontario
Ontario
Minnesota
Minnesota
Minnesota
Minnesota
Minnesota
Wisconsin
Wisconsin
Wisconsin
Wisconsin
Wisconsin
Wisconsin
Wisconsin
Wisconsin
Wisconsin
Wisconsin
Wisconsin
Illinois
Iowa
Oklahoma
Oklahoma
Oklahoma
Oklahoma
Oklahoma
Oklahoma
Oklahoma
Oklahoma
Oklahoma
Oklahoma
Maryland
Maryland
Maryland
Maryland
Maryland
NewJersey
NewJersey
NewJersey
NewJersey
NewJersey
NewJersey
NewJersey
NewJersey
NewJersey
Virginia
Virginia
Virginia
Virginia
Virginia
Virginia
Virginia
Florida
Florida
Florida
Florida
Florida
Florida
Florida
Bald Eagle Structure N=169


d = 20
Virginia
Maryland
Florida
New Jersey
Minnesota
Iowa
Wisconsin
Illinois
Saskatchewan
Oklahoma
Ontario
Wyoming
Alaska
Arizona

Panther GO Slim for Molecular Function and Biological Processes

A.




B.

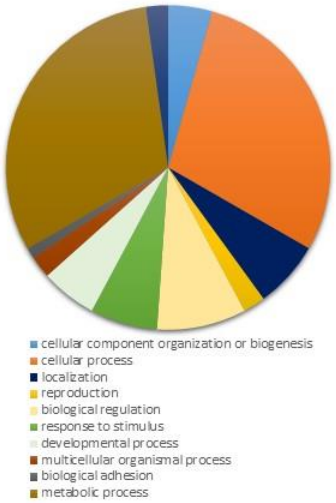


Table 1 Dates, release sites, natal location, and number of eagles translocated across North America after the DDT bottleneck.

| Date | Release site | From | Number of birds | Citation |
|---|---|---|---|---|
| 1985-1991 | Alabama | Unknown | 91 | Center for Biological Diversity, 2007 |
| Unknown | Alabama | Florida | 275* | Simons *et al.* 1988 |
| 1982 | Arkansas | Wisconsin, Minnesota | - | Center for Biological Diversity, 2007 |
| 1986-1994,1999-2000 | Big Sur, California | Unknown | 70 | Ventana Wildlife Society, 2012 |
| Unknown | California | Alaska | - | Environment Alaska |
| 2002-2006 | Channel Island | Unknown | - | Center for Biological Diversity, 2007 |
| 1995-1998 | District of Columbia | Wisconsin | 4 | Center for Biological Diversity, 2007 |
| Unknown | Georgia | Florida | 275* | Simons et al. 1988 |
| 1979-1995 | Georgia | Unknown | 89 | Georgia Department of Natural Resources |
| 1985-1989 | Lake Monroe, Indiana | Wisconsin, Alaska | 73 | Center for Biological Diversity, 2007 |
| 1970-1980 | Maine | Minnesota, Wisconsin | - | Center for Biological Diversity, 2007 |
| 1981-1990 | Mingo National Wildlife Refuge, Schell-Osage Conservation area, Missouri | Minnesota, Wisconsin | 74 | Center for Biological Diversity, 2007 |
| Unknown | Mississippi | Florida | 275* | Simons et al. 1988 |
| Unknown | Missouri | Alaska | - | Environment Alaska |
| 1983 | New Jersey | Nova Scotia, Manitoba | 60 | E. Miller and K. Clark Pers. Comm 2017 |
| Unknown | New York | Alaska, Great Lakes | 198 | Center for Biological Diversity, 2007 |
| 2002-2006 | New York | Wisconsin | 20 | Center for Biological Diversity, 2007 |
| 1983-1988 | North Carolina | Unknown | 29 | North Carolina Wildlife Resources Commission, 2005 |
| Unknown | North Carolina | Florida | 275* | Simons et al. 1988 |
| Unknown | North Carolina | Alaska | - | Environment Alaska |
| 1984-1990 | Oklahoma | Florida | 90* | Center for Biological Diversity, 2007 |
| 1983-1989 | Pennsylvania | Saskatchewan | 92 | Center for Biological Diversity, 2007 |
| 1982-1988 | Quabbin Reservoir, Massachusetts | Manitoba, Michigan, Nova Scotia | 41 | Center for Biological Diversity, 2007 |

| | | | | |
|---|---|---|---|---|
| 1979-1993 | Sapelo Island, Georgia | Patuxent Wildlife Research Center | 89 | Center for Biological Diversity, 2007 |
| 1980 | Tennessee | Unknown | 300 | Center for Biological Diversity, 2007 |
| 2007 | Tennessee | San Francisco Zoo | 6 | Center for Biological Diversity, 2007 |
| Unknown | Tennessee | Alaska | - | Environment Alaska |
| 2004-2006 | Vermont | Unknown | 26 | Center for Biological Diversity, 2007 |

.

Table 2. Summary results of the STRUCTURE reflecting the percentage of the population belonging to each of the genomes along with standard deviation, in parentheses.

| Location | Sample Size | Genome 1 | Genome 2 | Genome 3 | Genome 4 | Genome 5 |
|---|---|---|---|---|---|---|
| Alaska | 13 | 0.00% (0.33%) | 0.32% (1.09%) | 0.00% (0.02%) | 97.75% (3.91%) | 1.83% (3.84%) |
| Arizona | 27 | 10.07% (11.90%) | 0.61% (2.09%) | 76.47% (26.14%) | 1.21% (2.97%) | 11.64% (17.82%) |
| Wyoming | 5 | 1.45% (1.72%) | 1.55% (1.51%) | 3.29% (3.93%) | 24.20% (3.26%) | 69.52% (6.21%) |
| Midwest | 43 | 4.76% (4.25%) | 2.61% (2.66%) | 1.52% (2.03%) | 7.57% (3.33%) | 83.55% (6.45%) |
| Oklahoma | 19 | 31.68% (25.85%) | 3.47% (3.74%) | 9.92% (6.88%) | 5.51% (5.07%) | 49.43% (25.46%) |
| East Coast | 45 | 39.89% (15.80%) | 44.35% (13.19%) | 3.93% (3.94%) | 3.27% (3.85%) | 8.56% (12.18%) |
| Florida | 17 | 55.55% (30.62%) | 28.49% (31.25%) | 11.08% (7.30%) | 2.70% (4.58%) | 2.18% (3.42%) |

Table 3 Pairwise $F_{st}$ values for the seven groups identified in this study. Bold values indicate statistical significance at $P > 0.05$.

| | Central | Wyoming | East Coast | Florida | Alaska | Oklahoma | Arizona |
|---|---|---|---|---|---|---|---|
| Central | 0 | | | | | | |
| Wyoming | **0.03** | 0 | | | | | |
| East | **0.05** | **0.08** | 0 | | | | |
| Florida | **0.06** | **0.10** | **0.01** | 0 | | | |
| Alaska | **0.11** | **0.12** | **0.16** | **0.20** | 0 | | |
| Oklahoma | **0.02** | **0.05** | **0.03** | **0.03** | **0.14** | 0 | |
| Arizona | **0.08** | **0.11** | **0.06** | **0.06** | **0.20** | **0.05** | 0 |

Table 4 Genepop 4.4.2 calculations for the inbreeding coefficient ($F_{IS}$) of the seven units of bald eagles.

| Cluster | $F_{is}$ |
|---|---|
| Arizona | 0.286 |
| Upper Mid-west | 0.3218 |
| East Coast | 0.3038 |
| Oklahoma | 0.3083 |
| Florida | 0.2807 |
| Wyoming | 0.3099 |
| Alaska | 0.2615 |

Table 5 King values showing the kingship values between the 169 bald eagles included in this study. Pairings are broken down to include siblings, half siblings, first cousins, and second cousins. Pairings are further broken down in which both samples were from the same state and pairs in which the individuals in the pair were from different states.

| Kinship Value | Same State Pairing | Different State Pairing |
|---|---|---|
| 0.18-0.3 (Siblings) | 11 | 5 |
| 0.09-0.179 (1/2 siblings) | 26 | 2 |
| 0.03-0.089 (1st cousins) | 88 | 37 |
| 0.01-0.029 (2nd cousins) | 99 | 150 |
| <0.099 | 1148 | 12630 |

Table 6 Effective population size for six out of the seven population defined in the study using NeEstimator v 2.01. Calculations were based on the linkage disequilibrium setting using a monogamous system. Wyoming was not included due to small sample size

| Population | Sample Size | Ne | CI | |
|---|---|---|---|---|
| Arizona | 27 | 54.1 | 54.1 | 54.2 |
| Upper Mid-west | 43 | 1111.9 | 1107.5 | 1116.3 |
| East Coast | 45 | 225.8 | 225.6 | 226 |
| Oklahoma | 19 | 78.1 | 78 | 78.1 |
| Florida | 17 | 52.1 | 52.1 | 52.2 |
| Alaska | 13 | 2876.5 | 2731 | 3038.2 |

Supplementary Table 1 State, band number, and sampling type of the bald eagles included in the study.

| State | Band Number | Type |
|---|---|---|
| Alaska | VicBE1436 | Rehabilitation |
| Alaska | VicBE1434 | Rehabilitation |
| Alaska | VicBE1440 | Rehabilitation |
| Alaska | VicBE1441 | Rehabilitation |
| Alaska | VicBE1442 | Rehabilitation |
| Alaska | VicBE1443 | Rehabilitation |
| Alaska | VicBE1444 | Rehabilitation |
| Alaska | VicBE1445 | Rehabilitation |
| Alaska | VicBE1446 | Rehabilitation |
| Alaska | VicBE1447 | Rehabilitation |
| Alaska | VicBE1448 | Rehabilitation |
| Alaska | VicBE1450 | Rehabilitation |
| Alaska | VicBE1451 | Rehabilitation |
| Arizona | 00J05 | Hatchling |
| Arizona | 00J23 | Hatchling |
| Arizona | 01J10 | Hatchling |
| Arizona | 02J01 | Hatchling |
| Arizona | 02J04 | Hatchling |
| Arizona | 02J05.03 | Hatchling |
| Arizona | 02J11 | Hatchling |
| Arizona | 07J15 | Hatchling |
| Arizona | 07J18 | Hatchling |
| Arizona | 07J22 | Hatchling |
| Arizona | 07J23 | Hatchling |
| Arizona | 07J26 | Hatchling |
| Arizona | 10J26 | Hatchling |
| Arizona | 15J28 | Hatchling |
| Arizona | 96J14 | Hatchling |
| Arizona | 99J04 | Hatchling |
| Arizona | 99J06 | Hatchling |
| Arizona | 99J07 | Hatchling |
| Arizona | 99J09 | Hatchling |
| Arizona | 99J11 | Hatchling |
| Arizona | 99J13 | Hatchling |
| Arizona | 99J15 | Hatchling |
| Arizona | 99J17 | Hatchling |
| Arizona | 99J19 | Hatchling |
| Arizona | 99J21 | Hatchling |
| Arizona | 99J23 | Hatchling |
| Arizona | 99J24 | Hatchling |

| | | |
|---|---|---|
| Florida | 077-15 | Rehabilitation |
| Florida | 088-15 | Rehabilitation |
| Florida | 130-15 | Rehabilitation |
| Florida | 135-15 | Rehabilitation |
| Florida | 168-15 | Rehabilitation |
| Florida | 187-15 | Rehabilitation |
| Florida | 190-15 | Rehabilitation |
| Florida | 191-15 | Rehabilitation |
| Florida | 195-15 | Rehabilitation |
| Florida | 201-15 | Rehabilitation |
| Florida | 209-15 | Rehabilitation |
| Florida | 214-15 | Rehabilitation |
| Florida | 412-15 | Rehabilitation |
| Florida | 562-15 | Rehabilitation |
| Florida | E15 | Rehabilitation |
| Florida | E64 | Rehabilitation |
| Florida | JessE462 | Rehabilitation |
| Illinois | 0709-04925 | Hatchling |
| Iowa | 0709-04920 | Hatchling |
| Iowa | 0709-04930 | Hatchling |
| Iowa | 0709-04948 | Hatchling |
| Maryland | 0679-01275 | Hatchling |
| Maryland | 0679-01277 | Hatchling |
| Maryland | 0679-01278 | Hatchling |
| Maryland | 0679-01284 | Hatchling |
| Maryland | 0679-01290 | Hatchling |
| Maryland | 0679-01291 | Hatchling |
| Maryland | 0679-01292 | Hatchling |
| Maryland | 0679-01297 | Hatchling |
| Maryland | 0679-01343 | Hatchling |
| Maryland | 0679-01360 | Hatchling |
| Minnesota | 0629-48696 | Hatchling |
| Minnesota | 0629-48700 | Hatchling |
| Minnesota | 0629-48705 | Hatchling |
| Minnesota | 0629-48708 | Hatchling |
| Minnesota | 0629-48720 | Hatchling |
| Minnesota | 0629-48726 | Hatchling |
| Minnesota | 0629-48736 | Hatchling |
| Minnesota | 0629-48740 | Hatchling |
| Minnesota | 0629-48778 | Hatchling |
| Minnesota | 0679-03821 | Hatchling |
| New Jersey | 629-39883 | Hatchling |
| New Jersey | 629-39895 | Hatchling |
| New Jersey | 629-05434 | Hatchling |

| | | |
|---|---|---|
| New Jersey | 629-39893 | Hatchling |
| New Jersey | 629-05426 | Hatchling |
| New Jersey | 629-45894 | Hatchling |
| New Jersey | 629-05418 | Hatchling |
| New Jersey | 629-46837 | Hatchling |
| New Jersey | 629-46841 | Hatchling |
| New Jersey | 629-45811 | Hatchling |
| New Jersey | 0679-01751 | Hatchling |
| New Jersey | 629-18087 | Hatchling |
| New Jersey | 629-45855 | Hatchling |
| New Jersey | 0679-01784 | Hatchling |
| New Jersey | 629-18089 | Hatchling |
| New Jersey | 629-45897 | Hatchling |
| New Jersey | 0709-01559 | Hatchling |
| New Jersey | 0679-01735 | Hatchling |
| New Jersey | 0679-01766 | Hatchling |
| New Jersey | 629-46862 | Hatchling |
| Nunavut | 0709-02722 | GPS |
| Nunavut | 0709-04901 | GPS |
| Oklahoma | 0709-04951 | Hatchling |
| Oklahoma | 0709-04952 | Hatchling |
| Oklahoma | 0709-04953 | Hatchling |
| Oklahoma | 0709-04954 | Hatchling |
| Oklahoma | 0709-04975 | Hatchling |
| Oklahoma | 709-04977 | Hatchling |
| Oklahoma | 709-04979 | Hatchling |
| Oklahoma | E27 | Hatchling |
| Oklahoma | E28 | Hatchling |
| Oklahoma | E30 | Hatchling |
| Oklahoma | E31 | Hatchling |
| Oklahoma | E48 | Hatchling |
| Oklahoma | E58 | Hatchling |
| Oklahoma | E59 | Hatchling |
| Oklahoma | E61 | Hatchling |
| Oklahoma | E63 | Hatchling |
| Oklahoma | E7 | Hatchling |
| Oklahoma | E9 | Hatchling |
| Oklahoma | Sutton1 | Hatchling |
| Ontario | 0709-02723 | GSP |
| Ontario | 0709-02727 | GSP |
| Ontario | 0709-02728 | GSP |
| Ontario | 0709-04912 | GSP |
| Saskatchewan | 0709-04905 | GSP |
| Virginia | 0709-04915 | Hatchling |

| | | |
|---|---|---|
| Virginia | 0709-04922 | Hatchling |
| Virginia | 0709-04964 | Hatchling |
| Virginia | 0709-04965 | Hatchling |
| Virginia | 0709-04967 | Hatchling |
| Virginia | 0709-00739 | Hatchling |
| Virginia | 0709-00741 | Hatchling |
| Virginia | 0709-00742 | Hatchling |
| Virginia | 0709-00743 | Hatchling |
| Virginia | 0709-00744 | Hatchling |
| Virginia | 0709-00745 | Hatchling |
| Virginia | 0709-00747 | Hatchling |
| Virginia | 0709-00748 | Hatchling |
| Virginia | 0709-02702 | Hatchling |
| Virginia | 0709-02704 | Hatchling |
| Wisconsin | 0629-48638 | Hatchling |
| Wisconsin | 0629-48657 | Hatchling |
| Wisconsin | 0629-48661 | Hatchling |
| Wisconsin | 0629-48668 | Hatchling |
| Wisconsin | 0629-48670 | Hatchling |
| Wisconsin | 0629-48717 | Hatchling |
| Wisconsin | 0629-48751 | Hatchling |
| Wisconsin | 0629-48811 | Hatchling |
| Wisconsin | 0629-48833 | Hatchling |
| Wisconsin | 0629-49322 | Hatchling |
| Wisconsin | 0629-49370 | Hatchling |
| Wisconsin | 0629-49388 | Hatchling |
| Wisconsin | 0629-49391 | Hatchling |
| Wisconsin | 0629-49393 | Hatchling |
| Wisconsin | 0629-49397 | Hatchling |
| Wisconsin | 0679-03826 | Hatchling |
| Wisconsin | 0679-03911 | Hatchling |
| Wisconsin | 0679-03915 | Hatchling |
| Wisconsin | 0679-03922 | Hatchling |
| Wisconsin | 0709-02710 | GPS |
| Wisconsin | 0709-02721 | GSP |
| Wisconsin | GLKN2014B046 | Hatchling |
| Wyoming | 629-44403 | Hatchling |
| Wyoming | 629-44405 | Hatchling |
| Wyoming | 629-44407 | Hatchling |
| Wyoming | 629-44408 | Hatchling |
| Wyoming | 629-44409 | Hatchling |

Supplementary Table 2 GO Terms associated with the genes annotated to the single nucleotide polymorphism under selection that are in coding regions. Genes are categorize by the GO Molecular Function and GO Biological Process.

| Gene | GO Molecular Function | GO Biological Process |
|---|---|---|
| PDE4D | metal ion binding(GO:0046872)3',5'-cyclic-nucleotide phosphodiesterase activity(GO:0004114);metal ion binding(GO:0046872) | signal transduction(GO:0007165) |
| ADIPOR2 | receptor activity(GO:0004872);identical protein binding(GO:0042802);protein heterodimerization activity(GO:0046982);adiponectin binding(GO:0055100);adipokinetic hormone receptor activity(GO:0097003) | G-protein coupled receptor signaling pathway(GO:0007186);;adiponectin-activated signaling pathway(GO:0033211);glucose homeostasis(GO:0042593);vascular wound healing(GO:0061042) |
| ASB12 | - | intracellular signal transduction(GO:0035556) |
| GNE | hydrolase activity, hydrolyzing O-glycosyl compounds(GO:0004553); | UDP-N-acetylglucosamine metabolic process(GO:0006047) |
| FHIT | ubiquitin protein ligase binding(GO:0031625);identical protein binding(GO:0042802);bis(5'-adenosyl)-triphosphatase activity(GO:0047710) | purine nucleotide metabolic process(GO:0006163);negative regulation of proteasomal ubiquitin-dependent protein catabolic process(GO:0032435);intrinsic apoptotic signaling pathway by p53 class mediator(GO:0072332) |
| GALNT10 | polypeptide N-acetylgalactosaminyltransferase activity(GO:0004653);polypeptide N-acetylgalactosaminyltransferase activity(GO:0004653);carbohydrate binding(GO:0030246) | O-glycan processing(GO:0016266) |
| EIF4EBP1 | eukaryotic initiation factor 4E binding(GO:0008190);translation repressor activity(GO:0030371) | G1/S transition of mitotic cell cycle(GO:0000082);IRES-dependent translational initiation of linear mRNA(GO:0002192);insulin receptor signaling pathway(GO:0008286);TOR signaling(GO:0031929);positive regulation of mitotic cell cycle(GO:0045931);negative regulation of translational initiation(GO:0045947);cellular response to dexamethasone stimulus(GO:0071549) |

| | | |
|---|---|---|
| CEP120 | protein C-terminus binding(GO:0008022) | centrosome cycle(GO:0007098);positive regulation of centrosome duplication(GO:0010825);cerebral cortex development(GO:0021987);neurogenesis(GO:0022008);interkinetic nuclear migration(GO:0022027);astral microtubule organization(GO:0030953);regulation of protein localization(GO:0032880);positive regulation of cilium assembly(GO:0045724) |
| NTRK2 | GPI-linked ephrin receptor activity(GO:0005004);voltage-gated ion channel activity(GO:0005244);ATP binding(GO:0005524);protein homodimerization activity(GO:0042803);neurotrophin binding(GO:0043121);brain-derived neurotrophic factor binding(GO:0048403);brain-derived neurotrophic factor-activated receptor activity(GO:0060175):voltage-gated ion channel activity(GO:0005244) | vasculogenesis(GO:0001570);neuron migration(GO:0001764);neuromuscular junction development(GO:0007528);learning(GO:0007612);circadian rhythm(GO:0007623);feeding behavior(GO:0007631);positive regulation of cell proliferation(GO:0008284);positive regulation of neuron projection development(GO:0010976);glutamate secretion(GO:0014047);positive regulation of phosphatidylinositol 3-kinase signaling(GO:0014068);peptidyl-tyrosine phosphorylation(GO:0018108);central nervous system neuron development(GO:0021954);cerebral cortex development(GO:0021987);neuron differentiation(GO:0030182);brain-derived neurotrophic factor receptor signaling pathway(GO:0031547);positive regulation of peptidyl-serine phosphorylation(GO:0033138);regulation of ion transmembrane transport(GO:0034765);neurotrophin signaling pathway(GO:0038179);mechanoreceptor differentiation(GO:0042490);regulation of GTPase activity(GO:0043087);positive regulation of MAPK cascade(GO:0043410);negative regulation of neuron apoptotic process(GO:0043524);retinal rod cell development(GO:0046548);protein autophosphorylation(GO:0046777);ephrin receptor signaling pathway(GO:0048013);collateral sprouting(GO:0048668);positive regulation of collateral sprouting(GO:0048672);oligodendrocyte differentiation(GO:0048709);peripheral nervous system neuron development(GO:0048935);positive regulation of axonogenesis(GO:0050772);regulation of protein kinase B signaling(GO:0051896);positive regulation of synapse assembly(GO:0051965);retina development in camera-type eye(GO:0060041);long-term synaptic potentiation(GO:0060291);cellular response to amino acid stimulus(GO:0071230);trans-synaptic signaling by neuropeptide, modulating synaptic transmission(GO:0099551);regulation of dendrite extension(GO:1903859);negative regulation of anoikis(GO:2000811);regulation of ion transmembrane transport(GO:0034765);regulation of ion transmembrane transport(GO:0034765);regulation of dendrite extension(GO:1903859) |
| FBN2 | | |

| | | |
|---|---|---|
| DSCAML1 | protein homodimerization activity(GO:0042803) | cell adhesion(GO:0007155);central nervous system development(GO:0007417);homophilic cell adhesion via plasma membrane adhesion molecules(GO:0007156);synapse assembly(GO:0007416);brain development(GO:0007420);retina layer formation(GO:0010842);embryonic skeletal system morphogenesis(GO:0048704);camera-type eye photoreceptor cell differentiation(GO:0060219) |
| DEPTOR | | negative regulation of protein kinase activity(GO:0006469);negative regulation of TOR signaling(GO:0032007);intracellular signal transduction(GO:0035556) |
| ST8SIA4 | alpha-N-acetylneuraminate alpha-2,8-sialyltransferase activity(GO:0003828) | ganglioside biosynthetic process(GO:0001574);protein glycosylation(GO:0006486);N-glycan processing(GO:0006491);oligosaccharide metabolic process(GO:0009311);sialylation(GO:0097503);sialylation(GO:0097503) |
| BAMBI | frizzled binding(GO:0005109) | positive regulation of cell proliferation(GO:0008284);regulation of cell shape(GO:0008360);positive regulation of epithelial to mesenchymal transition(GO:0010718);cell migration(GO:0016477);negative regulation of transforming growth factor beta receptor signaling pathway(GO:0030512);positive regulation of protein binding(GO:0032092);positive regulation of catenin import into nucleus(GO:0035413);positive regulation of transcription, DNA-templated(GO:0045893);positive regulation of canonical Wnt signaling pathway(GO:0090263) |
| TMEM38B | calcium-activated potassium channel activity(GO:0015269);cation channel activity(GO:0005261) | potassium ion transmembrane transport(GO:0071805);monovalent inorganic cation transport(GO:0015672);cation transmembrane transport(GO:0098655) |
| PCGF3 | - | - |
| BTF3 | - | - |
| SLCO4C1 | sodium-independent organic anion transmembrane transporter activity(GO:0015347) | sodium-independent organic anion transport(GO:0043252) |

| | | |
|---|---|---|
| AMOT | receptor activity(GO:0004872);angiostatin binding(GO:0043532) | vasculogenesis(GO:0001570);gastrulation with mouth forming second(GO:0001702);establishment of cell polarity involved in ameboidal cell migration(GO:0003365);chemotaxis(GO:0006935);negative regulation of angiogenesis(GO:0016525);actin cytoskeleton organization(GO:0030036);regulation of cell migration(GO:0030334);negative regulation of GTPase activity(GO:0034260);cellular protein localization(GO:0034613);hippo signaling(GO:0035329);positive regulation of embryonic development(GO:0040019);cell migration involved in gastrulation(GO:0042074);blood vessel endothelial cell migration(GO:0043534);regulation of small GTPase mediated signal transduction(GO:0051056) |
| ATP7A | nucleotide binding(GO:0000166);cation-transporting ATPase activity(GO:0019829);copper-dependent protein binding(GO:0032767);cuprous ion binding(GO:1903136) | copper ion transport(GO:0006825);cellular copper ion homeostasis(GO:0006878);positive regulation of oxidoreductase activity(GO:0051353);ATP hydrolysis coupled cation transmembrane transport(GO:0099132) |
| FKTN | | negative regulation of cell proliferation(GO:0008285);protein O-linked mannosylation(GO:0035269);negative regulation of JNK cascade(GO:0046329) |
| ELAVL2 | RNA binding(GO:0003723) | |
| PARD3B | | |
| TRABD2A | metalloendopeptidase activity(GO:0004222);Wnt-protein binding(GO:0017147) | proteolysis(GO:0006508);negative regulation of Wnt signaling pathway(GO:0030178) |
| PATZ1 | RNA polymerase II core promoter proximal region sequence-specific DNA binding(GO:0000978);transcriptional activator activity, RNA polymerase II core promoter proximal region sequence-specific binding(GO:0001077);chromatin binding(GO:0003682);nucleic acid binding(GO:0003676);nucleic acid binding(GO:0003676) | transcription from RNA polymerase II promoter(GO:0006366);spermatogenesis(GO:0007283);male gonad development(GO:0008584);T cell differentiation(GO:0030217);positive regulation of transcription from RNA polymerase II promoter(GO:0045944) |
| RNF170 | metal ion binding(GO:0046872) | |
| SLC44A1 | choline transmembrane transporter activity(GO:0015220) | choline transport(GO:0015871) |

| | | |
|---|---|---|
| HOOK3 | microtubule binding(GO:0008017);identical protein binding(GO:0042802);dynein light intermediate chain binding(GO:0051959) | endosome organization(GO:0007032);lysosome organization(GO:0007040);endosome to lysosome transport(GO:0008333);interkinetic nuclear migration(GO:0022027);cytoskeleton-dependent intracellular transport(GO:0030705);cytoplasmic microtubule organization(GO:0031122);microtubule anchoring at centrosome(GO:0034454);early endosome to late endosome transport(GO:0045022);negative regulation of neurogenesis(GO:0050768);Golgi localization(GO:0051645);protein localization to centrosome(GO:0071539);neuronal stem cell population maintenance(GO:0097150) |
| SHB | SH3/SH2 adaptor activity(GO:0005070) | signal transduction(GO:0007165);positive regulation of signal transduction(GO:0009967) |
| NDUFAF2 | NADH dehydrogenase (ubiquinone) activity(GO:0008137);electron carrier activity(GO:0009055) | respiratory electron transport chain(GO:0022904);cellular respiration(GO:0045333);negative regulation of insulin secretion involved in cellular response to glucose stimulus(GO:0061179);reactive oxygen species metabolic process(GO:0072593) |
| PSD3 | ARF guanyl-nucleotide exchange factor activity(GO:0005086) | regulation of ARF protein signal transduction(GO:0032012);positive regulation of GTPase activity(GO:0043547) |
| IFT172 | | neural tube closure(GO:0001843);heart looping(GO:0001947);Notch signaling pathway(GO:0007219);smoothened signaling pathway(GO:0007224);brain development(GO:0007420);epidermis development(GO:0008544);dorsal/ventral pattern formation(GO:0009953);protein processing(GO:0016485);spinal cord motor neuron differentiation(GO:0021522);cytoplasmic microtubule organization(GO:0031122);positive regulation of smoothened signaling pathway(GO:0045880);negative regulation of epithelial cell proliferation(GO:0050680);palate development(GO:0060021);limb development(GO:0060173);cilium assembly(GO:0060271);bone development(GO:0060348);hindgut development(GO:0061525);left/right axis specification(GO:0070986);non-motile cilium assembly(GO:1905515) |

| | | |
|---|---|---|
| CDKAL1 | N6-threonylcarbomyladenosine methylthiotransferase activity(GO:0035598);4 iron, 4 sulfur cluster binding(GO:0051539);catalytic activity(GO:0003824);catalytic activity(GO:0003824);iron-sulfur cluster binding(GO:0051536);4 iron, 4 sulfur cluster binding(GO:0051539) | tRNA methylthiolation(GO:0035600);maintenance of translational fidelity(GO:1990145) |
| SAP30L | transcription cofactor activity(GO:0003712);histone deacetylase activity(GO:0004407);zinc ion binding(GO:0008270);phosphatidylinositol-5-phosphate binding(GO:0010314);nucleosome binding(GO:0031491);histone binding(GO:0042393);non-sequence-specific DNA binding, bending(GO:0044378) | regulation of transcription, DNA-templated(GO:0006355);histone deacetylation(GO:0016575) |
| GPR98 | - | - |
| FSD1L | - | - |
| ACSS1 | acetate-CoA ligase activity(GO:0003987);ATP binding(GO:0005524);AMP binding(GO:0016208) | acetyl-CoA biosynthetic process from acetate(GO:0019427) |
| TBC1D24 | - | - |
| TRIM14 | zinc ion binding(GO:0008270) | negative regulation of viral transcription(GO:0032897);innate immune response(GO:0045087);positive regulation of NF-kappaB transcription factor activity(GO:0051092) |
| GOLPH3 | phosphatidylinositol-4-phosphate binding(GO:0070273) | retrograde vesicle-mediated transport, Golgi to ER(GO:0006890);Golgi organization(GO:0007030);Golgi to plasma membrane protein transport(GO:0043001);Golgi vesicle budding(GO:0048194) |
| PTPRE | protein tyrosine phosphatase activity(GO:0004725) | peptidyl-tyrosine dephosphorylation(GO:0035335) |

| | | |
|---|---|---|
| MTMR8 | phosphatidylinositol-3-phosphatase activity(GO:0004438);protein tyrosine phosphatase activity(GO:0004725);phosphatidylinositol-3,5-bisphosphate 3-phosphatase activity(GO:0052629);phosphatidylinositol-3-phosphatase activity(GO:0004438);protein tyrosine phosphatase activity(GO:0004725);phosphatidylinositol-3,5-bisphosphate 3-phosphatase activity(GO:0052629) | phosphatidylinositol dephosphorylation(GO:0046856);regulation of autophagy(GO:0010506);peptidyl-tyrosine dephosphorylation(GO:0035335) |
| NANS | N-acylneuraminate-9-phosphate synthase activity(GO:0047444) | carbohydrate biosynthetic process(GO:0016051) |
| ZMPSTE24 | metalloendopeptidase activity(GO:0004222) | nuclear envelope organization(GO:0006998);prenylated protein catabolic process(GO:0030327);CAAX-box protein processing(GO:0071586) |
| ATG10 | Atg12 transferase activity(GO:0019777) | protein lipidation(GO:0006497);autophagy(GO:0006914);ER overload response(GO:0006983);protein modification by small protein conjugation(GO:0032446) |
| PPP1R9A | - | - |
| ANKRD31 | - | - |
| ITPRIP | - | - |
| LUZP1 | | ventricular septum development(GO:0003281);neural fold bending(GO:0021503);artery development(GO:0060840) |
| NODAL | cytokine activity(GO:0005125);transforming growth factor beta receptor binding(GO:0005160);growth factor activity(GO:0008083) | positive regulation of pathway-restricted SMAD protein phosphorylation(GO:0010862);BMP signaling pathway(GO:0030509);regulation of apoptotic process(GO:0042981);regulation of MAPK cascade(GO:0043408);cell development(GO:0048468);SMAD protein signal transduction(GO:0060395) |
| FUT10 | alpha-(1->3)-fucosyltransferase activity(GO:0046920) | cerebral cortex radially oriented cell migration(GO:0021799);fucosylation(GO:0036065);neuronal stem cell population maintenance(GO:0097150);protein glycosylation(GO:0006486);fucosylation(GO:0036065);protein glycosylation(GO:0006486) |
| RUFY3 | metal ion binding(GO:0046872) | positive regulation of cell migration(GO:0030335) |
| TPPP2 | tubulin binding(GO:0015631) | - |

| | | |
|---|---|---|
| MAN2A1 | alpha-mannosidase activity(GO:0004559);hydrolase activity, hydrolyzing N-glycosyl compounds(GO:0016799);carbohydrate binding(GO:0030246);metal ion binding(GO:0046872) | liver development(GO:0001889);mannose metabolic process(GO:0006013);N-glycan processing(GO:0006491);protein deglycosylation(GO:0006517);mitochondrion organization(GO:0007005);vacuole organization(GO:0007033);respiratory gaseous exchange(GO:0007585);lung alveolus development(GO:0048286);positive regulation of neurogenesis(GO:0050769);retina morphogenesis in camera-type eye(GO:0060042) |
| KCNB2 | voltage-gated potassium channel activity(GO:0005249) | regulation of ion transmembrane transport(GO:0034765);protein homooligomerization(GO:0051260);potassium ion transmembrane transport(GO:0071805) |
| PHAX | RNA binding(GO:0003723) | snRNA export from nucleus(GO:0006408);protein transport(GO:0015031) |
| TBCD | beta-tubulin binding(GO:0048487);GTPase activator activity(GO:0005096);GTPase activator activity(GO:0005096);beta-tubulin binding(GO:0048487);GTPase activator activity(GO:0005096);beta-tubulin binding(GO:0048487) | mitotic cell cycle(GO:0000278);tubulin complex assembly(GO:0007021);post-chaperonin tubulin folding pathway(GO:0007023);negative regulation of microtubule polymerization(GO:0031115);positive regulation of GTPase activity(GO:0043547);cell morphogenesis involved in neuron differentiation(GO:0048667);microtubule cytoskeleton organization(GO:0000226);protein folding(GO:0006457);tubulin complex assembly(GO:0007021);post-chaperonin tubulin folding pathway(GO:0007023);negative regulation of cell-substrate adhesion(GO:0010812);negative regulation of microtubule polymerization(GO:0031115);adherens junction assembly(GO:0034333);positive regulation of GTPase activity(GO:0043547);bicellular tight junction assembly(GO:0070830);tubulin complex assembly(GO:0007021);post-chaperonin tubulin folding pathway(GO:0007023);positive regulation of GTPase activity(GO:0043547) |
| CHD1 | DNA binding(GO:0003677);helicase activity(GO:0004386);ATP binding(GO:0005524);ATP binding(GO:0005524);ATP binding(GO:0005524);ATP binding(GO:0005524);ATP binding(GO:0005524) | transcription, DNA-templated(GO:0006351);regulation of transcription, DNA-templated(GO:0006355) |

| | | |
|---|---|---|
| GCNT4 | acetylglucosaminyltransferase activity(GO:0008375) | inter-male aggressive behavior(GO:0002121);thyroid hormone metabolic process(GO:0042403);tissue morphogenesis(GO:0048729);homeostasis of number of cells(GO:0048872);kidney morphogenesis(GO:0060993) |
| KLF8 | nucleic acid binding(GO:0003676) | |
| TXNRD 3 | thioredoxin-disulfide reductase activity(GO:0004791);electron carrier activity(GO:0009055);protein disulfide oxidoreductase activity(GO:0015035);flavin adenine dinucleotide binding(GO:0050660);thioredoxin -disulfide reductase activity(GO:0004791);electron carrier activity(GO:0009055);protein disulfide oxidoreductase activity(GO:0015035);flavin adenine dinucleotide binding(GO:0050660) | oxidation-reduction process(GO:0055114);cellular oxidant detoxification(GO:0098869);cell redox homeostasis(GO:0045454);oxidation-reduction process(GO:0055114) |

| | | |
|---|---|---|
| SYK | protein tyrosine kinase activity(GO:0004713);non-membrane spanning protein tyrosine kinase activity(GO:0004715);receptor binding(GO:0005102);ATP binding(GO:0005524);phosphotyrosine residue binding(GO:0001784);protein serine/threonine kinase activity(GO:0004674);non-membrane spanning protein tyrosine kinase activity(GO:0004715);signal transducer, downstream of receptor, with protein tyrosine kinase activity(GO:0004716);integrin binding(GO:0005178);ATP binding(GO:0005524);protein kinase binding(GO:0019901);phosphatase binding(GO:0019902);Toll-like receptor binding(GO:0035325);SH2 domain binding(GO:0042169) | lymph vessel development(GO:0001945);adaptive immune response(GO:0002250);macrophage activation involved in immune response(GO:0002281);neutrophil activation involved in immune response(GO:0002283);leukocyte activation involved in immune response(GO:0002366);serotonin secretion by platelet(GO:0002554);protein phosphorylation(GO:0006468);inflammatory response(GO:0006954);leukocyte cell-cell adhesion(GO:0007159);transmembrane receptor protein tyrosine kinase signaling pathway(GO:0007169);integrin-mediated signaling pathway(GO:0007229);regulation of platelet activation(GO:0010543);cell migration(GO:0016477);peptidyl-tyrosine phosphorylation(GO:0018108);cell differentiation(GO:0030154);neutrophil chemotaxis(GO:0030593);regulation of superoxide anion generation(GO:0032928);positive regulation of cell adhesion mediated by integrin(GO:0033630);positive regulation of cell adhesion mediated by integrin(GO:0033630);intracellular signal transduction(GO:0035556);peptidyl-tyrosine autophosphorylation(GO:0038083);defense response to bacterium(GO:0042742);positive regulation of mast cell degranulation(GO:0043306);regulation of neutrophil degranulation(GO:0043313);innate immune response(GO:0045087);innate immune response(GO:0045087);positive regulation of B cell differentiation(GO:0045579);positive regulation of bone resorption(GO:0045780);positive regulation of alpha-beta T cell proliferation(GO:0046641);blood vessel morphogenesis(GO:0048514);regulation of phagocytosis(GO:0050764);B cell receptor signaling pathway(GO:0050853);cellular response to hydrogen peroxide(GO:0070301);regulation of ERK1 and ERK2 cascade(GO:0070372);cellular response to molecule of fungal origin(GO:0071226);regulation of arachidonic acid secretion(GO:0090237);regulation of platelet aggregation(GO:0090330);regulation of platelet aggregation(GO:0090330);lymph vessel development(GO:0001945);positive regulation of receptor internalization(GO:0002092);macrophage activation involved in immune response(GO:0002281);neutrophil activation involved in immune response(GO:0002283);serotonin secretion by platelet(GO:0002554);leukocyte cell-cell adhesion(GO:0007159);integrin-mediated signaling pathway(GO:0007229);activation of JUN kinase activity(GO:0007257);regulation of tumor necrosis factor-mediated signaling pathway(GO:0010803);peptidyl-serine phosphorylation(GO:0018105);peptidyl-tyrosine phosphorylation(GO:0018108);leukotriene biosynthetic process(GO:0019370);neutrophil |

| | | |
|---|---|---|
| | | chemotaxis(GO:0030593);receptor internalization(GO:0031623);positive regulation of type I interferon production(GO:0032481);regulation of superoxide anion generation(GO:0032928);positive regulation of cell adhesion mediated by integrin(GO:0033630);collagen-activated tyrosine kinase receptor signaling pathway(GO:0038063);defense response to bacterium(GO:0042742);transcription factor import into nucleus(GO:0042991);positive regulation of mast cell degranulation(GO:0043306);regulation of neutrophil degranulation(GO:0043313);beta selection(GO:0043366);innate immune response(GO:0045087);positive regulation of interleukin-3 biosynthetic process(GO:0045401);positive regulation of granulocyte macrophage colony-stimulating factor biosynthetic process(GO:0045425);positive regulation of B cell differentiation(GO:0045579);positive regulation of B cell differentiation(GO:0045579);positive regulation of gamma-delta T cell differentiation(GO:0045588);positive regulation of bone resorption(GO:0045780);positive regulation of alpha-beta T cell differentiation(GO:0046638);positive regulation of alpha-beta T cell proliferation(GO:0046641);protein autophosphorylation(GO:0046777);blood vessel morphogenesis(GO:0048514);positive regulation of cytokine secretion(GO:0050715);regulation of phagocytosis(GO:0050764);positive regulation of calcium-mediated signaling(GO:0050850);B cell receptor signaling pathway(GO:0050853);regulation of sequence-specific DNA binding transcription factor activity(GO:0051090);regulation of ERK1 and ERK2 cascade(GO:0070372);cellular response to molecule of fungal origin(GO:0071226);cellular response to low-density lipoprotein particle stimulus(GO:0071404);regulation of arachidonic acid secretion(GO:0090237);regulation of platelet aggregation(GO:0090330);positive regulation of peptidyl-tyrosine autophosphorylation(GO:1900086) |
| XRCC4 | DNA binding(GO:0003677);protein C-terminus binding(GO:0008022);identical protein binding(GO:0042802);DNA binding(GO:0003677) | double-strand break repair via nonhomologous end joining(GO:0006303);DNA recombination(GO:0006310);response to X-ray(GO:0010165);DNA ligation involved in DNA repair(GO:0051103);positive regulation of ligase activity(GO:0051351);double-strand break repair(GO:0006302);DNA recombination(GO:0006310) |
| IDUA | L-iduronidase activity(GO:0003940) | carbohydrate metabolic process(GO:0005975);dermatan sulfate catabolic process(GO:0030209) |

| | | |
|---|---|---|
| PTPRA | protein tyrosine phosphatase activity(GO:0004725);protein binding(GO:0005515) | peptidyl-tyrosine dephosphorylation(GO:0035335) |
| ATG3 | Atg8 ligase activity(GO:0019776);Atg12 transferase activity(GO:0019777);enzyme binding(GO:0019899) | autophagosome assembly(GO:0000045);autophagy of mitochondrion(GO:0000422);protein transport(GO:0015031);protein ubiquitination(GO:0016567);mitochondrial fragmentation involved in apoptotic process(GO:0043653);autophagy of nucleus(GO:0044804);negative regulation of phagocytosis(GO:0050765);regulation of cilium assembly(GO:1902017);autophagy(GO:0006914) |
| RNF38 | ubiquitin protein ligase activity(GO:0061630) | protein ubiquitination(GO:0016567);proteasome-mediated ubiquitin-dependent protein catabolic process(GO:0043161) |
| FREM1 | | cell communication(GO:0007154);cell-matrix adhesion(GO:0007160);craniofacial suture morphogenesis(GO:0097094) |
| PAM | peptidylglycine monooxygenase activity(GO:0004504);peptidylamidoglycolate lyase activity(GO:0004598);copper ion binding(GO:0005507);zinc ion binding(GO:0008270) | peptide amidation(GO:0001519);oxidation-reduction process(GO:0055114) |
| GMDS | - | |
| CLTA | structural molecule activity(GO:0005198);clathrin heavy chain binding(GO:0032050) | intracellular protein transport(GO:0006886);clathrin coat assembly(GO:0048268);clathrin-dependent endocytosis(GO:0072583) |
| MTMR12 | phosphatase regulator activity(GO:0019208) | regulation of catalytic activity(GO:0050790);toxin transport(GO:1901998) |
| DEPDC1B | GTPase activator activity(GO:0005096) | cell migration(GO:001+B2:C746477); positive regulation of Wnt signaling pathway(GO:0030177);intracellular signal transduction(GO:0035556);signal transduction(GO:0007165);positive regulation of GTPase activity(GO:0043547) |

# CHAPTER IV

## RELATEDNESS AND TURNOVER RATE OF A REINTRODUCED POPULATION OF BALD EAGLES (*HALIAEETUS LEUCOCEPHALUS) IN NEW JERSEY*


### Abstract

Avian species utilize both monogamous and polygamous mating systems, with advances in molecular technologies revealing extra pair copulations in species that were once thought to be sexually monogamous. Raptors are thought to exhibit some of the highest rates of sexual monogamy among bird species, but few raptor species have been studied in light of genetics to determine relatedness and nest turnover rates. Bald eagles (*Haliaeetus lecocephalus*) are long-lived raptors that are believed to be sexually monogamous and have high rates of site fidelity. Bald eagles suffered a severe population bottleneck in the middle of the 19$^{th}$ century resulting in many reintroduction efforts to help increase the numbers of individuals across their range. This study evaluates the nest turnover rate and relatedness in offspring bald eagles across the state of New Jersey, an area where high numbers of reintroduced individuals were released. The results reveal situations where siblings raised together in the nest exhibited half sibling and unrelated

relatedness values. The study also documented a 36.3% nest turnover rate as well as high levels of inbreeding across the population.

**Introduction**

Although it was once believed that most bird species exhibited sexual monogamy (Lack 1969), molecular technologies have revealed that extra pair copulations occur in a variety of bird species at rates higher than originally expected (Griffith *et al.* 2002). By utilizing extra pair copulations as a mating strategy, several benefits can be incurred by both the male and the female including genetic benefits, a reduction in cost if a mate is lost, and the ability to gain extra resources (Petrie and Kempenaers 1998). The benefits of extra pair copulations does not come without cost, as time spent away from a mate in pursuit of extra pair copulations allows the unsupervised mate to pursue its own extra pair copulations. Moreover, life history has an impact on the frequency of extra pair copulations as extra pair copulations are more prevalent in species that nest at high densities and reproduce in synchrony (Stutchbury 1998; Arlt *et al.* 2004).

Raptors are long-lived species that have high rates of parental investments by the males (Gavin *et al.* 1998). It is thought that this high parental investment also contributes to the fact that raptors have some of the highest intra-pair copulation rates among bird species (Birkhead and Moller 1992). Although raptors are thought to be highly monogamous, as more studies are conducted a range of extra pair paternity rates for raptors are being documented. Current studies of raptors failed to detect extra pair paternity events in merlins (*Falco columbarius*, Warkentin *et al.* 1994) and flammulated owls (*Otus flammeolus*, Arsenault *et al.* 2002), low rates of extra pair paternities in tawney owls (*Strix aluco*, Saladin *et al.* 2007), northern goshawks (*Accipiter gentilis*

*atricapillus*, Gavin *et al.* 2017), and lesser kestrels (*F. naumanni*, Negro *et al.* 1996), and a single species with a high rate as shown in Cooper's hawks (*A. cooperii*, Rosenfield *et al.* 2015). The vastness of rates revealed in these studies demonstrate that more genetic based studies are need as specific species may exhibit higher rates than previously thought. Eagles, class Accipitridae, are assumed to exhibit high rates of monogamy, but this has not been thoroughly studied from a genetics standpoint. The only published study examining extra pair paternity in an eagle species evaluated nest turnover rates and pairwise relatedness of the eastern imperial eagle (*Aquila heliacal*) (Rudnick et al. 2009). Based on the results of their study, they determined that this species is highly monogamous.

Raptors are known to utilize the same nesting sites in successive years, but the identity and relationship of the nesting adults is typically unknown (Newton 1979). Although nesting success and quality of nesting site can have an impact on nest turnover rates, the impact of these qualities as well as the average nest turnover rate vary between species and even populations of the same species. Current studies have shown various turnover rates when comparing peregrine falcons (*F. peregrinus,* Court *et al.* 1989, Ponnikas *et al.* 2017), prairie falcons (*F. mexicanus,* Lehman *et al.* 2000), golden eagles (*Aquila chrysaetos,* Kochert & Steenhof 2012), and Eastern imperial eagles (Rudnick *et al.* 2005).

The bald eagle (*Haliaeetus leucocephalus*) is found throughout North America and despite the species' unique history, has been minimally studied from a genetic standpoint. Bald eagles suffered a severe population bottleneck in the middle of the 19$^{th}$ century due to the use of dichlorodiphenyltrichloroethane (DDT) and as a result many

populations were involved in reintroduction programs with the genetic consequences of the bottleneck and reintroductions unknown. The bald eagle is thought to be sexually monogamous, but there are several occasions documented instances in which three bald eagles were caring for the young in a single nest, with the overall rate and understanding of this behavior unknown (Frazer 1983). In the state of New Jersey, the bottleneck reduced the bald eagle population to one nesting pair located in Cumberland County. After this pair failed to produce offspring for several years, a hacking project was conducted in which 60 eaglets from Manitoba and Nova Scotia were hacked into the New Jersey population over an eight-year period (Smith and Clark 2016). Now, there are 172 known eagle nests in New Jersey, but the degree of relatedness, inbreeding, and nest turnover rates are undocumented (Smith and Clark 2016).

This study utilizes long-term sampling (1997 – 2016) of 18 nest in New Jersey with the purpose of estimating the incidence of extra pair paternity events as well as the type and frequency of nest turnover rates within this population. The results from this study will provide the first genetic parentage and nest turnover analysis of bald eagles, thereby providing a clearer picture of the social structure of this species.

## Materials and Methods

### *Samples*

Whole blood sample aliquots (~0.5 ml) were collected from previously frozen samples and from hatchling bald eagles, stored in lysis buffer, and shipped to our lab at Oklahoma State University for standard DNA extraction (Longmire *et al.* 1997). Whole blood samples represented 162 hatchling individuals, sampled at 18 nest locations from

1997 to 2016 (Figure 1, Table 1). Due to sample availability and DNA quality, not all nests were represented in each year. DNA quality was assessed by running an aliquot on a 1% agarose gel and quantified using a NanoDrop 3300 spectrophotometer (Thermo Scientific). DNA was subsequently shipped to Eurofins (River Falls, WI) where it was run on our custom Axiom myDesign array for bald and golden eagles (Van Den Bussche *et al.* 2017, Judkins 2017 Chapter 3). The Axiom Analysis Suite v2.0.0.35 was used to score the raw data provided by Eurofins using the 45,952 SNPs developed by Judkins (2017 Chapter 3) for bald eagles.

*Genetic Statistics*

Genepop 4.4.3 (Raymond & Rousset 1995) was used to test for Hardy-Weinberg equilibrium (HWE) for each SNP within each nest site. A sequential Bonferroni correction was used with a p-value of $< 0.05$ to indicate significance. Genepop 4.4.3 was also used to determine the inbreeding coefficient for each nest site.

*Relationship Analyses*

To determine the relationship and kinship values between the 164 samples, KING (Manichaikul *et al.* 2010), Colony2 (Wang 2004), and ML-relate (Kalinowski *et al.* 2006) were used. In KING, the 45,952 SNP dataset was used and the kinship option was utilized to calculate the kinship coefficient between each individual within the dataset. In Colony2, the 45,952 SNP dataset was run using a full-likelihood comparison with a high accuracy run. A polygamous system was allowed since the dataset spanned many years and nest turnover events were expected. The allelic dropout and error rates for this

analysis were calculated for each locus using Plink (Purcell *et al.* 2007). Five replicate analyses were used to ensure consistency across runs. A subset of SNPs was created for ML-relate as the 45,952 SNP dataset was too computationally intensive for the program. To create the reduced dataset, Plink was utilized to filter the 45,952 SNP dataset to only include SNPs with a minor allele frequency > 0.4. Furthermore, a 50 base pair sliding window was used and no more than two SNPs were allowed within this window. For the final dataset, 300 SNPs were selected at random from the remaining 2,463 filtered SNPs. The estimate relatedness option was used in ML-relate to calculate pairwise relationship values.

To determine the relationship among individuals and nest turnover rates, the results from all three analyses were compared. When there was a discrepancy in a pairwise relationship value between programs, the relationship (full sibling, half sibling, unrelated) with the majority of program's support was used. To determine when a nest turnover event occurred, a single year's samples was compared to the previous year's samples and a relationship type was determined. If the relationship between the samples were half-siblings it was assumed that a single individual of the pair was replaced. If the relationship between the samples was unrelated, it was assumed that both individuals of the pair were replaced.

**Results**

DNA was successfully extracted and genotyped for 162 individuals sent to Eurofins (Table 1). Plink determined that the average allelic dropout rate ranged from 0.0 - 0.079 across all SNPs and these calculations were used in Colony2. When using a

sequential Bonferroni correction, all loci were in HWE. The hatchlings at the nest sites were found to be highly inbred with the inbreeding coefficients ranging from 0.29-0.35 (Table 2). The high levels of inbreeding coupled with the lack of parental samples is attributed to some of discrepancies in relationships between programs.

*Relatedness*

All 18 nests sampled had years in which multiple hatchlings hatched at the nest (n=55). Of these occurrences, 46 (84%) of the relationships were full sibling relationships and 3 (5%) were half sibling relationships (Table 3). The half siblings occurred at Merrill Creek (1) and Wheaton Island (2). Upon further investigation of the half sibling relationships, it was determined that in all three comparisons one individual was not related to the previous year's individuals at that nest, thereby creating an outlier. When comparing the outliers to each other at Wheaton Island in 2010 and 2012, it was determined that these individuals were half siblings themselves.

In four comparisons of samples from the same nest and the same year (7%), all three programs determined the individuals sampled together were not related (Table 3). These comparisons were found at Manasaquan River (1) Mannington 2 (2), and Newton (1). When comparing the outlier individual of the pair for these individuals, it was determined that the outliers for 2005 and 2006 at Mannington 2 were half siblings.

In two cases (4%), the programs were not able to come to a consensus for the relationship between the individuals (Table 3). In both of these cases, three individuals were raised together at the Alloways nest (1999, 2016). In these situations, two individuals were determined to be full siblings with the outlier individual being a half

sibling with one of the full sibling individuals and a full sibling with the other. These pairwise relationships types were the same in all three programs, therefore the true genetic relationship of the three individuals could not be determined.

*Nest turnover rates*

When comparing the relatedness values of individuals in a single nest on a year to year basis, it was determined that there were 20 nest turnover events. Three of the events (15%) were a turnover in which both of the nesting individuals the year before were replaced, 15 (75%) were a turnover in which one individual was replaced, and two (10%) were a turnover in which the data was unclear as to if one or both individuals were replaced (Table 4). The highest number of replacements at a nest was three over a nine year period and each of these were turnovers in which only one individual was replaced. Five nests did not experience any turnover events in the three to seven year time period when the nest was monitored.

**Discussion**

This study aimed to evaluate the relatedness of hatchlings and nest turnover rates for a population of bald eagles in New Jersey. Despite bald eagle being thought as sexually monogamous, three pairs of half siblings were documented as well as four occurrences of siblings that were not related. The study also documented the nest turnover rates which appear to be higher than expected when taking into account the lifespan of bald eagles. Finally, all nests in the population exhibit high rates of inbreeding.

*Sibling Relatedness and Inbreeding*

Of the 55 occurrences in which more than one individual was raised in a nest in a given year, three occurrences resulted in half sibling relationships while four occurances resulted in siblings that were not related to one another. Based on further evaluation of the outlier hatchlings at Wheaton Island and Mannington 2, it was determined that the outliers were half siblings, indicating that the extra pair copulations were with the same individual over a number of years.

As the paternity of bald eagle hatchlings has not been highly studied we can only speculate as to how and why these occurrences are reflected in the genetics of New Jersey nests. Scenarios that would allow for the half sibling and unrelated individuals observed in these nests would be: female replacement at a nest in which additional eggs were laid at the nest; extra pair copulations by the female; a trio situation in which two males were contributing genes to the offspring; or a trio situation in which two females were contributing genes to the offspring. As the full nesting population of bald eagles in New Jersey is currently estimated at 172 nests and New Jersey is 8,729 $mi^2$, the average nest density of bald eagles in New Jersey is at most one nest per 50.75 $mi^2$. This close proximity of nesting sites would allow for many interactions of individuals between nesting sites allowing for extra pair copulations. Furthermore, a trio of eagles caring for young in single nest has been documented on several occasions and could be an explanation for the genetic patterns observed at these nest sites. For example, in Illinois a webcam has documented three eagles, a single female and two males, which have cared for young between 2013 and 2017 (Stewards of the Upper Mississippi River Refuge 2017). These eagles shared in all aspects of nesting. Furthermore, Frenzel (1983)

documented a nest in which three individuals were at a nest site in Minnesota from at least April until June with the contributions of the third individual unknown. Finally, multiple nesting events in Alaska have been documented with three individuals at a nest site through the entirety of the nesting season (Sherrod *et al.* 1976, Heglund and Reiswig 1980). One of these nests contained four eggs and Bent (1937) proposed that the eggs were laid by more than one female.

Future studies should utilize DNA samples from the suspected parents to provide clarity into these half sibling and unrelated instances. While genotyping error and mishandled samples could be a consideration in these cases of unrelated and half sibling individuals, our data available from other localities also shows a pair of siblings in the same nest during the same year in Oklahoma that are not full or half siblings (Judkins unpublished). Therefore, these events warrant more evaluation from both a field and genetic stance.

When considering the $F_{IS}$ values calculated for the nest sites (Table 2) we find that this population exhibits high degrees of inbreeding. This can be attributed to bottleneck and subsequent reintroduction of 60 eaglets. Other studies of raptors have calculated $F_{IS}$ values of -0.182 - 0.260 for populations of peregrine falcons (Talbot *et al.* 2011), 0.077 for the Mauritius kestrel (*Falco punctatus*, Ewing *et al.* 2008) and 0.014 - 0.098 for populations of Cooper's hawks (Rosenfield *et al.* 2015). When comparing these values to values observed in this study, 0.285-0.347, we find that our study has substantially higher degrees of inbreeding even when compared to the reintroduced population of the Mauritius kestrel and the island populations of the peregrine falcon.

*Nest Turnover*

The data for this study suggested that there were 20 (36.3%) occurrences of nest turnover events with most of the turnovers replacing a single individual of the mating pair (Table 4). This study further revealed that over the 20 year period there were at least 38-40 mating individuals replaced out of the 110 mating individuals at the nest sites. These numbers could be lower than the true replacement values since samples from every year were not available or did not pass quality control. At Union Lake, the type of turnover could not be evaluated as the three programs used had conflicting results. Five nests, Belleplain, Manasquan Reservoir, Primceton, Cohansey-Greenwhich and Duke Farms, did not have any turnover events. These nest were sampled for 3-7 years. The nest that had the most turn over events, Fort Dix, was only represented by nine years of data. Finally, the samples that represented the longest timespan, Galloway (15 years) and Merrill Creek (17 years), only had one or two turnovers respectively. As bald eagles are suspected to have high degrees of nest site fidelity and sexual monogamy, these turnover rates seem high when compared to a bald eagle's 25 year life span.

In other species of raptors, there was substantial differences between species when comparing nest turnover rates. For example, in peregrine falcons (*F. peregrinus),* turnover rates were estimated at 21.7% (Ponnikas *et al.* 2017) and 22.0% (Court *et al.* 1989) in two different populations. Ponnikas *et al.* (2017) noted that the previous year's nesting success did not have an impact on the subsequent year's nest turnover events. Prairie falcons (*F. mexicanus*) in Idaho were documented having a turnover rate of 57% with the high turnover rate attributed to the high density of falcons in the study site (Lehman *et al.* 2000). In a population of golden eagles (Aquila chrysaetos), 40 cases of

turnover events occurred between 1970-2011 (Kochert & Steenhof 2012), while in a population of Eastern imperial eagles 20 turnover events occurred between 1999-2002 (Rudnick *et al.* 2005). Because of the range of documented turnover rates across raptors, species specific studies, as well as studies specific to a specific population studies of a species, are necessary and assumptions about nest turnover rate without these studies are cautioned.

## Conclusions

This study represents the first genetic analyses of the New Jersey population of bald eagles. The results of this study reveal a highly inbred population with the first genetically based occurrences of half siblings and unrelated bald eagles being raised together in a nest. Furthermore, based on our limited knowledge of genetically tested nest turnover rates in raptors, this population appears to have a moderate degree of nest turnover rate. Future studies of relatedness and nest turnover rates in bald eagles should utilize samples not only from the hatchlings, but also adults to aid in clarity. Furthermore, using a combination of field observations, marking of individuals, and genetics would help provide clarity into these occurrences in other populations.

## Acknowledgements

We would like to thank Erica Miller for providing the idea, samples, and her knowledge about this population of bald eagles. We would also like to thank Kathy Clark for her information and experience about New Jersey bald eagles. Critical funding for this project was provided by the Iowa Tribe of Oklahoma as well as the Ford Foundation

which funded Megan Judkins' project while this project was being completed. Finally, we thank the members of Megan's dissertation committee, Drs. Meredith Hamilton, Jim Lish, and Andrew Doust for their expertise, guidance, and comments on a previous version of this manuscript.

**Literature Cited**

Arlt D, Hansson B, Bensch S, Von Schantz T, Hasselquist D (2004) Breeding synchrony does not affect extra-pair paternity in great reed warblers, *Behaviour*, 141: 863-880.

Arsenault DP, Stacey PB, Hoelzer GA (2002) No extra-pair fertilization in flammulated owls despite aggregated nesting, *The Condor*, **104**, 197–201.

Bent AC. (1937) Life histories of North American birds of prey. Part I. U.S. National Museum Bulletin 167. Dover Publisher, New York.

Birkhead TA, Moller AP (1992) Sperm competition in birds: evolutionary causes and consequences. Academic Press, San Diego, CA.

Court, GS, Bradley, DM, Gates, CC, Boag, D.A. 1989. Turnover and recruitment in a tundra population of Peregrine Falcons, *Ibis*, 131: 487–496.

Ewing SR, Nager RG, Nicoll MAC, Aumjaud A, Jones CG, Keller LF (2008) Inbreeding and loss of genetic variation in a reintroduced population of Mauritius kestrel. *Conservation Biology*, 22, 395-404.

Frazer JD (1983) Three adult bald eagles at an active nest, *Journal of Raptor Research*, 17: 29 -30. Unpublished report, U.S.D.I. Fish and Wildlife Service., Aleutian Islands National Wildlife Refuge

Gavin TA, Reynolds RT, Joy SM. Leslie D, May B (2017) Genetic evidence for low frequency of extra-pair fertilizations in northern goshawks, *American Ornithological*

*Society*, **100**, 556–560.

Griffith SC, Owens IPF, Thuman KA, Griffith, Simon C. Owens, IPF, Thuman KA (2002) Extra pair paternity in birds : a review of interspecific. *Molecular ecology*, **11**, 2195–2212.

Heglund, PJ, Reiswig B (1980) Raptor survey, the breeding Bald Eagle population, Amchitka Island, Alaska.

Lack D (1968) Ecological adaptations for breeding in birds. Methuen Ltd, London

Kalinowski ST, Wagner AP, Taper ML (2006) ML-RELATE: A computer program for maximum likelihood estimation of relatedness and relationship, *Molecular Ecology Notes*, **6**, 576–579.

Kochert MN, Steenhof K (2012) Frequency of nest use by golden eagles in southwestern Idaho, *Journal of Raptor Research*, **46**, 239–247.

Lehman RN, Steenhof K, Carpenter LB, Kochert MN (2000) Turnover and dispersal of Prairie Falcons in southwestern Idaho, *Journal of Raptor Research*, **34**, 262–269.

Longmire JL, Maltbie M, Baker RJ (1997) Use of "lysis buffer" in DNA isolation and its implication for museum collections, *Occasional Papers of the Museum of Texas Tech University*, **163**, 1–3.

Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM (2010) Robust relationship inference in genome-wide association studies, *Bioinformatics*, **26**, 2867–2873.

Negro JJ, Villarroel M, Tella JL, Kuhnlein U, Hiraldo F, Donazar JA, Bird DM (1996) DNA fingerprinting reveals a low incidence of extra-pair fertilizations in the lesser kestrel, *Animal Behaviour*, **51**, 935–943.

Petrie M, Kempenaers B (1998) Extra-pair paternity in birds: Explaining variation between species and populations. Trends in Ecology and Evolution, 13, 52–57.

Ponnikas S, Ollila T, Kvist L (2017) Turnover and post-bottleneck genetic structure in a recovering population of Peregrine Falcons Falco peregrinus, *Ibis*, **159**, 311–323.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreir MAR, Bender D, Maller J, Sklar P, Bakker PIW, Daly MJ, Sham PC (2007) REPORT PLINK : A tool set for whole-genome association and population-based linkage analyses, *The American Journal of Human Genetics*, **81**, 559–575.

Raymond M, Rousset F (1995) Genepop population-genetics software for exact tests and ecumenicism, *Journal of Heredity*, **86**, 248–249.

Rosenfield RN, Sonsthagen SA, Stout WE, Talbot SL (2015) High frequency of extra-pair paternity in an urban population of Cooper's Hawks, *Journal of Field Ornithology*, **86**, 144–152.

Rudnick JA, Katzner TE, Bragin EA, Rhodes OE, Dewoody JA (2005) Using naturally shed feathers for individual identification, genetic parentage analyses, and population monitoring in an endangered Eastern imperial eagle population from Kazakhstan, *Molecular Ecology*, **14**, 2959–2967.

Rudnick JA, Katzner TE, Dewoody JA (2009) Genetic analyses of noninvasively

collected feathers can provide new insights into avian demography and behavior, *Handbook of Nature Conservation*, 181–197.

Saladin V, Ritschard M, Roulin A, Bize P, Richner H (2007) Analysis of genetic parentage in the tawny owl (Strix aluco) reveals extra-pair paternity is low, *Journal of Ornithology*, **148**, 113–116.

Sherrod SK, White CM, Williamson FSI (1976) Biology of the bald eagle on Amchitka Island, Alaska. Living Bird.

Smith L, Clark KE (2016) New Jersey bald eagle project 2016. Retrieved from http://www.state.nj.us/dep/fgw/ensp/pdf/eglrpt16.pdf.

Stewards of the Upper Mississippi River Refuge (2017) Trio bald eagle nest cam. Retrieved from http://stewardsumrr.org/webcams/bald-eagle-nest-cam-live/.

Stutchbury BJM (1998) Breeding synchrony best explains variation in extra-pair mating system among avian species, *Behavioral Ecology and Sociobiology*, 43: 221-222.

Talbot SL, Palmer AG, Sage GK, Sonsthagen SA, Swem T, Brimm DJ, White CM (2011) Lack of genetic polymorphism among peregrine falcons Falco peregrinus of Fiji. Journal of Avian Biology, 42, 415–428.

Van Den Bussche RA, Judkins ME, Montague MJ, Warren WC (2017) A resource of genome-wide single nucleotide polymorphisms ( Snps ) for the conservation and management of golden eagles, *Journal of Raptor Research*, **51**, 368–377.

Wang J (2004) Sibship reconstruction from genetic data with typing errors, *Genetics*,

**166**, 1963–1979.

Warkentin IG, Curzon AD, Carter RE, Wetton JH, James PC, Oliphant LW, Parkin DT (1994) No evidence for extrapair fertilizations in the merlin revealed by DNA fingerprinting, *Molecular Ecology*, **3**, 229–234.

**Figure Legends**

Figure 1 Map of New Jersey with diamonds representing the location of nesting sites used in this study.

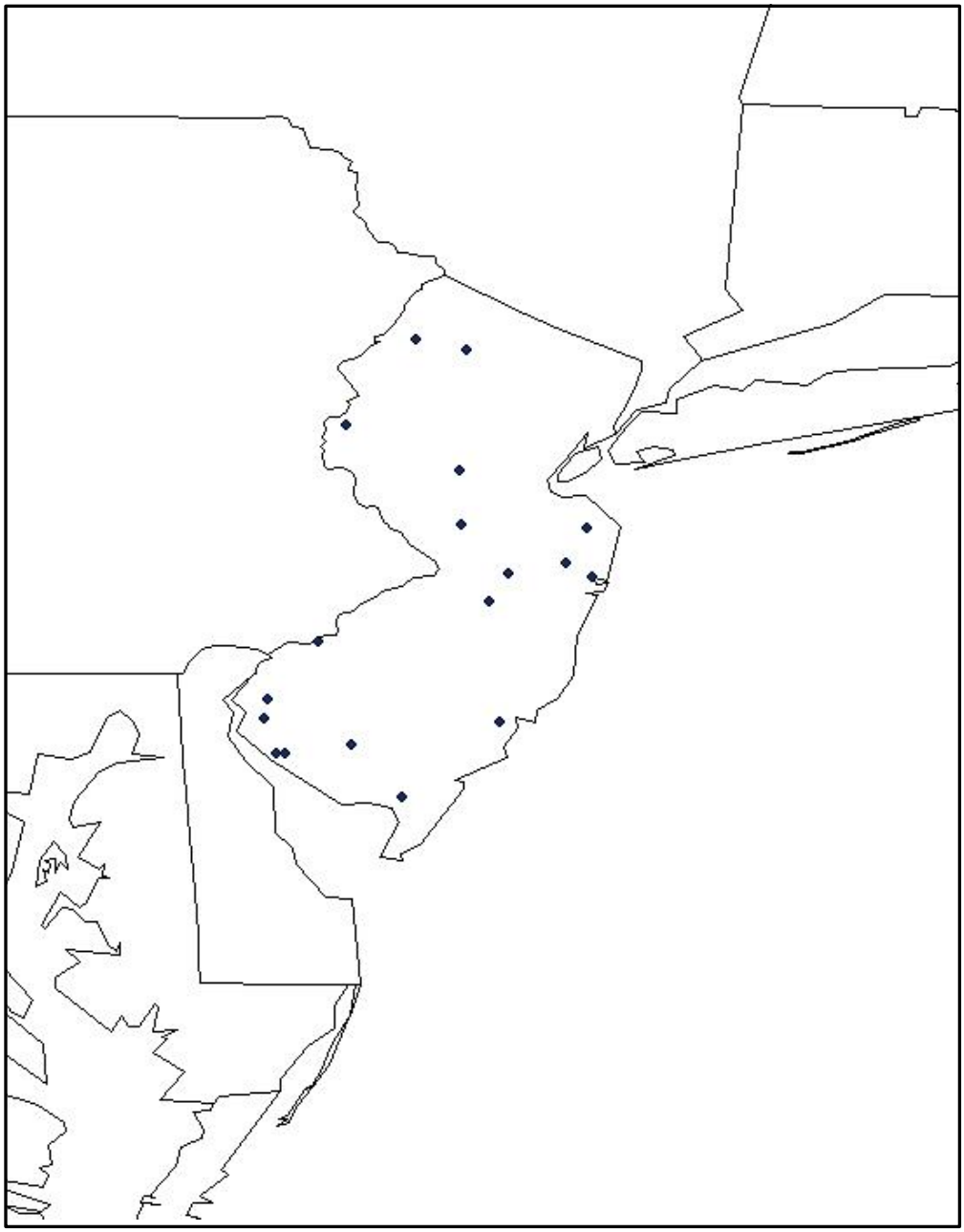


Table 1 List of nesting sites and years in which samples were collected for this study. The numbers associated with each year and nest site represent the number of offspring sampled in the nest.

| | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Alloways | | 3 | 1 | | 1 | | | | | | | | | | | | | | | |
| Belleplain | | 2 | 2 | | 2 | 2 | | 2 | | | | | | | | | | | | |
| Cohansey-Greenwhich | | | | 1 | 2 | | 1 | | 1 | 2 | | | | | | | | | | |
| Duke Farms | | | | | | | | | | | | | 2 | 2 | | | | | | 2 |
| Fort Dix | | | | | | 1 | | | 2 | 1 | 1 | 2 | 1 | | 2 | | | | | |
| Galloway | | | | 3 | 1 | 2 | 2 | 2 | | 1 | | | | | | | 2 | | | 3 |
| Little Swartwood | | | | | | | | | | | 2 | 1 | | 2 | 3 | | | 1 | 2 | 1 |
| Manasquan Reservoir | | | | | | | | | | | 1 | 2 | 1 | 1 | | | | | | |
| Manasquan River | | | | | | | | | | | | | 1 | 1 | 1 | 2 | 1 | | 1 | 1 |
| Mannington 2 | | | | | | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | | | | | | |
| Mantua | | | | | | | 1 | 2 | 2 | | 1 | 1 | | 1 | | 2 | | | | |
| Merrill Creek | | | | 1 | 1 | | | 1 | | 1 | 1 | 1 | 2 | | | 2 | 2 | | 2 | |
| Navesink | | | | | | | 1 | 1 | | | 1 | | | | | | 2 | | | |
| Newton | | | | | | | | | | | | | 1 | | | 1 | | | 2 | 2 |
| Princeton | | | | | | | | | 1 | | 1 | | 2 | | | | | | | |
| Prospertown | | | | | | | | | | | | | | | | 2 | 2 | | | 2 |
| Union Lak | 2 | 1 | | 1 | 1 | | 2 | | | | | 2 | | | | | | | | |
| Wheaton Island | | | | | | | | | 1 | | | 2 | 2 | 2 | | 2 | | | | |

Table 2 Inbreeding coefficients ($F_{IS}$) for the 18 nests used in this study using Genepop 4.4.3.

| Nest | $F_{IS}$ |
|---|---|
| Alloways | 0.3193 |
| Belleplain | 0.3469 |
| Cohansey-Greenwhich | 0.3097 |
| Duke Farms | 0.3043 |
| Fort Dix | 0.3235 |
| Galloway | 0.3139 |
| Little Swartwood | 0.3199 |
| Manasquan Reservoir | 0.3257 |
| Manasquan River | 0.3002 |
| Mannington 2 | 0.2912 |
| Mantua | 0.31 |
| Merril Creek | 0.2851 |
| Navesink | 0.3053 |
| Newton | 0.3182 |
| Princeton | 0.3288 |
| Prospertown | 0.312 |
| Union Lake | 0.3172 |
| Wheaton Island | 0.3222 |

Table 3 The nests and years in which multiple individuals were collected from a single nest site, but the individuals could not be classified as full siblings. The relatedness abbreviations represent half siblings (HS), not related (NR), and unknown (U).

| Nest | Year | N | Relatedness |
|---|---|---|---|
| Alloways | 1998 | 3 | U |
| Galloway | 2016 | 3 | U |
| Manasquan River | 2012 | 2 | NR |
| Mannington 2 | 2002 | 2 | NR |
| | 2005 | 2 | NR |
| Merrill Creek | 2015 | 2 | HS |
| Newton | 2015 | 2 | NR |
| Wheaton Island | 2010 | 2 | HS |
| | 2012 | 2 | HS |

Table 4 Number of nest turnovers at each nesting site broken down by full turnovers which occurred when both of the previous year's nesting pair was replaced and half turnovers which occurred when only one of the previous year's nesting individuals was replaced. The type of turnover was noted as unknown if a turnover event happened, but a consensus for the type of turnover could not be attained.

| Nest | Number of full turn overs | Number of Half Turnovers | Total Turnovers |
|---|---|---|---|
| Alloways | 1 | | 1 |
| Belleplain | | | 0 |
| Cohansey-Greenwhich | | | 0 |
| Duke Farms | | | 0 |
| Fort Dix | | 3 | 3 |
| Galloway | 1 | 1 | 2 |
| Little Swartwood | | 1 | 1 |
| Manasquan Reservoir | | | 0 |
| Manasquan River | 1 | 1 | 2 |
| Mannington 2 | | 1 | 1 |
| Mantua | | 1 | 1 |
| Merril Creek | | 1 | 1 |
| Navesink | | 2 | 2 |
| Newton | | 1 | 1 |
| Princeton | | | 0 |
| Prospertown | | 1 | 1 |
| Union Lake | Unknown | Unknown | 2 |
| Wheaton Island | | 2 | 2 |

VITA

Megan Judkins

Candidate for the Degree of

Doctor of Philosophy

Thesis: CONSERVATION GENOMICS OF NORTH AMERICAN BALD (*HALIAEETUS LEUCOCEPHALUS*) AND GOLDEN (*AQUILA CHRYSAETOS*) EAGLES

Major Field: Zoology

Biographical:

Education:
Completed the requirements for the Doctor of Philosophy in Zoology at Oklahoma State University, Stillwater, Oklahoma in December, 2017.

Completed the requirements for the Master of Business Administration at Marylhurst University, Marylhurst, Oregon in 2012.

Completed the requirements for the Bachelor of Science in Natural Resources Ecology and Management at Oklahoma State University, Stillwater, Oklahoma in 2009.

Experience: Teaching Assistant: Introductory Biology (1 semester), General Genetics (1 semester), Human Heredity (3 semesters)
Graduate Research Assistant: 6 semesters, 3 summers
Grey Snow Eagle House: Assistant Manager (5 years), Interim Director (1 year)

Professional Memberships: Raptor Research Foundation, The Wildlife Society, National Wildlife Rehabilitation Association

Professional Society Involvement: Raptor Research Foundation – Early Career Raptor Researcher's Committee Member, Website Coordinator, Website Committee Member, The Wildlife Society - Native Peoples' Wildlife Management Working Group Board Member