Hybridization between the American Wigeon and the Eurasian Wigeon in North America

Andrew Correll

Faculty Advisor: Dr. Jeff Peters

Honors Research Thesis

Submitted August 17th, 2019

Introduction

Properly diagnosing species forms the foundation for our understanding of much of ecology and evolution. Almost all biologists agree on what a species fundamentally is (Giraud *et al.* 2008), which the Evolutionary Species Concept sums up—"a single lineage of ancestral descendant populations of organisms which maintains its identity from other such lineages" (Wiley 1978). However, numerous theories and criteria have been posited to more precisely define the concept. The Morphological Species Concept focuses on differing traits, but a species categorized using this definition often comprises multiple cryptic species when different definitions are applied, such as the Biological Species Concept or the Phylogenetic Species Concept (Giraud *et al.* 2008). The most influential species concept was defined by Ernst Mayr in the Biological Species Concept (BSC), which defines a species as "groups of interbreeding natural populations that are reproductively isolated from other such groups" (Mayr 1963). However, gene flow sometimes occurs between species, as Ernst Mayr himself admitted was the case between coyotes and wolves, oaks and birches, and many other species. Therefore, he revised his BSC definition to "biological properties of individuals which prevent the interbreeding (fusion) of populations" (Mayr 1996). This resembles the commonly-agreed-upon Evolutionary Species Concept but similarly falls short in not providing much information on

what constitutes those biological properties. Utilizing genetic information provides more clarity, and numerous genetic species concepts have been proposed, such as the Phylogenetic Species Concept (PSC), the Genealogical Concordance Phylogenetic Species Recognition criteria (GCPSR), and the Genic Species Concept (GSC) (Giraud *et al.* 2008; Wu 2001). The PSC examines a single gene's genealogy to determine evolutionary relationships, while the GCPSR emphasizes evolutionary concordance among a number of such examined genes.

Next-generation sequencing has taken the GCPSR to the next level and has allowed for widespread examination of thousands of loci, providing much more resolution than previous methods (Rašic *et al.* 2014). This allows for the detection of specific segments of DNA that separate species and cause decreased hybrid fitness—a tenet of the relatively new genic species model (Wu 2001). The genic species model—an alternative to the BSC and the other genetic speciation models—emphasizes genes that contribute to nonrandom mating and maintain species identity, thus incorporating the maintenance of genetic differentiation idea that is important in the widely held evolutionary species concept (Wiley 1978; Wu 2001). Multiple genes with a small detrimental effect in hybrids may add up, as in multifarious speciation, or one or a few genes can have strong selective effects in this model (Nosil 2012). This more modern view of speciation allows for introgression by fertile hybrids by moving away from complete reproductive isolation—the BSC—and towards differentiation and partial isolation at the genic level, defining species on their ability to not lose differentiation and to continue to diverge (Wu 2001).

Wu's (2001) definition is consistent with the possibility of introgression, which is important because gene flow might be natural and historically stable (Moore & Buchanan 1985). Furthermore, hybridization could be an important part of speciation itself (Allendorf *et al.* 2001),

allowing for the creation of species (DeMarais *et al.* 1992) and the exchange of beneficial genes, such as mimicry traits in butterflies, through genetic leakage (Dasmahapatra *et al.* 2012). While strong selection on a small number of traits is more likely to lead to two groups of organisms diverging when there is a lot of gene flow, multifarious selection on many traits may be more likely to cause reproductive isolation, as an important gene might diverge and cause weaker hybrid fitness (Nosil 2012).

Genes that contribute to reproductive isolation or cause hybrids to have reduced fitness, (but function normally in "pure" individuals) are considered speciation genes (Nei *et al.* 1983). For example, the two mice subspecies *Mus musculus musculus* and *Mus m. domesticus* have such different alleles for Prdm9 (encoding a histone methyltransferase) that hybrids are infertile. This one gene separates the two subspecies, as placing the original Prdm9 gene back into the mouse line (transgenetically) restores fertility in the F1 hybrids (Mihola *et al.* 2009).

A greater understanding of species integrity is needed in studies of speciation, as the "decision about where to place the limit of the species is subjective" (Taylor *et al.* 2000). Pairwise kinship coefficients and ln Pr(X| K) calculations can be used to try to determine the number of independently evolving populations (Harvey & Brumfield 2015; Peters *et al.* 2016), but the current amorphous speciation criterion in the scientific community has led to the idea of "evolutionarily significant units" for legal purposes and to avoid the controversial issue of defining subspecies or species (Pennock & Dimmick 1997). While genetic species concepts are gaining prominence in biological use, geographic data and tenets of the MSC and BSC are still used in conglomeration with genetic data when defining species boundaries (Harvey & Brumfield 2015). The use of multiple lines of evidence leads to less subjectivity in the definition of a species, but understanding about the consequences of hybridization and gene flow, and how

it contributes to or impedes the speciation process, is lacking. For example, in hybridizing species, it is often unknown whether hybrids backcross into parental species, which is necessary for gene flow.

In this study, we examine the genomes of putative hybrids between two species of duck, the American Wigeon (*Anas americana*) and Eurasian Wigeon (*A. penelope*), to determine whether backcrossing is occurring. The Eurasian Wigeon is native to Europe and Asia, but the overwintering population has been increasing in North America since at least the 1970's (Michael 1984). Phenotypic hybrids have been observed in the wild, but these observations do not necessarily imply gene flow, because hybrids could be infertile, as in the case of mules (the offspring of a female horse and male donkey). Indirect evidence suggests, however, that American x Eurasian F1 hybrids are indeed fertile, because the two species share a low frequency of similar mitochondrial DNA (mtDNA) haplotypes (sets of DNA variations maternally inherited as a single unit in birds) (Kulikova & Zhuravlev 2010; Peters *et al.* 2005). Furthermore, an analysis of nuclear DNA suggests that models including gene flow are a better explanation of genetic variation than models that do not include gene flow (Peters *et al.* 2014b). If gene flow was occurring then this would be consistent with Anseriformes' tendency for high introgression (Fowler *et al.* 2009; Kulikova *et al.* 2004; Peters *et al.* 2014a; Rhymer *et al.* 1994; Williams *et al.* 2005). However, the studies showing gene flow—and any genetic analysis program for that matter—make many assumptions (Hey & Nielsen 2004), and only provide indirect evidence of gene flow. In this study, the aim is to obtain *direct* evidence of backcrossing by using genomic data to test for the existence of F2-backcrosses (offspring of a hybrid and "pure" individual) in a sample of putative hybrids. We will use programs specifically designed to detect hybridization—ADMIXTURE (Alexander *et al.* 2009) and STRUCTURE (Pritchard *et al.*

2000)—that assign samples percentages of the genome to each of the parental populations or species. For example, an F1 hybrid would have ~50% assignment to each of the parental species, whereas an F2-backcross would have 25% assignment to one species and 75% assignment to the other (an F3 would have a ratio of 12.5%:87.5%, and so on).

This information would be useful in the taxonomy of these ducks, because morphologically the Eurasian Wigeon and the American Wigeon are distinct but more similar to each other than other species of duck (Livezey 1991). Based on mitochondrial DNA, they are not sister species, with the American Wigeon being more closely related to the South American Chiloe Wigeon (*A. sibilatrix*) (Peters *et al.* 2005; Johnson & Sorenson 1999; Gonzalez *et al.* 2009). Therefore, with varying lines of evidence, the objective of this study is to determine whether *Anas americana* x *Anas penelope* F1 hybrids are backcrossing into the parental populations.

Methods

Muscle samples from 89 individuals were used in this study. Thirty-four putatively "pure" American Wigeons and 19 "pure" Eurasian Wigeons (2 from Russia and 17 from North America) were arbitrarily subsampled from Peters et al. (Peters *et al.* 2005; Peters *et al.* 2014b). In addition, 36 putative hybrids (from plumage characteristics) were obtained from the Burke Museum of Ornithology, University of Washington. Sampling locations included numerous locations around the United States, though mostly from Western states including Washington, California, Oregon, and Alaska. A DNeasy Blood and Tissue Kit was used to extract the DNA following the manufacturer's protocol.

Once the purified DNA was obtained, a double-digest restriction-associated DNA sequencing (ddRAD-seq) protocol was used to sample thousands of loci scattered across the genome (DaCosta & Sorenson 2014). In brief, this method included, 1) digesting the genomic DNA with restriction enzymes, 2) tagging each sample with a unique DNA sequence, 3) selecting fragments between 300 and 450 base pairs, 4) using PCR to amplify selected fragments, and 5) pooling all samples in equimolar concentrations.

The double digest is a procedure that improves upon the original RAD-seq method (Miller *et al.* 2007) by reducing the time it takes to create a fragment library (DaCosta & Sorenson 2014). It generates a smaller subset of genes (for practicality) than RAD-seq by only selecting those loci which have a relatively long restriction site (the longer the required sequence for the site then the fewer loci that will be cut overall). Also, using two enzymes further reduces the number of genes obtained by only selecting those loci that have the two enzyme's restriction site in close proximity to each other (DaCosta & Sorenson 2014).

To place a DNA tag on each sample, the restriction enzymes SbfI (leaves a TGCA overhang) and EcoRI (leaves an AATT overhang) were used to cut the DNA, leaving portions of DNA with two overhangs. Two adaptors were ligated, a P1 and a P2; the P1 had an overhang with reverse complementarity to SbfI's cut site overhang while the P2 had an overhang with complementarity to EcoRI's cut site overhang (DaCosta & Sorenson 2014). The adaptors contained a "barcode" sequence and enabled the selection of only fragments with EcoRI's cut site on one end and SbfI's cut site on the other end by using a "divergent Y" in the P2 adaptor (DaCosta & Sorenson 2014). This adaptor also contained Illumina forward (in P1) and reverse (in P2) amplification sequences to be used in a later PCR step. Then, to make sure that all the DNA pieces were of relatively similar lengths, a fine-sieve 2% gel was run to separate all the

DNA strands by length. This was done by using a size standard of known length and cutting out DNA fragments (visualized using SYBR green and UV light) that traveled the same distance as the 300-450 base pair (bp) standard. The gel segments were purified using a MinElute Gel extraction Kit (Qiagen, Valencia, CA).

With each sample barcoded and of similar lengths, a PCR amplification procedure was completed (to obtain enough DNA for the later sequencing procedure to function properly) using Phusion high-fidelity DNA polymerase (Thermo Scientific, Pittsburgh, PA), and those products were purified using magnetic AMPure XP beads (Beckman Coulter, Inc., Indianapolis, IN). A Quantstudio 7 real-time PCR machine quantified the products. These data were used to dilute some samples so that there were relatively equal amounts of DNA per individual (as too little of a sample might provide an incomplete sequence but too much might waste sequencing resources on one sample). The final barcoded samples with equimolar concentrations were pooled and sequenced at the University of Oregon.

Once sequencing data were obtained, a bioinformatics pipeline (DaCosta & Sorenson 2014) was used to process the data. This processing involved condensing identical reads while retaining the highest quality score for each individual, filtering low-quality reads (average PHRED score < 20), and clustering similar reads among individuals into putative loci (Peters *et al.* 2016). For each cluster, the read with the highest quality score was localized to a specific chromosome and position in the mallard reference genome (accession numbers SS263068950-SS263191362) using BLASTN v. 2 (Altschul *et al.* 1990). Reads were then aligned using MUSCLE v. 3 (Edgar 2004) and genotyped using the RADGenotypes.py script (DaCosta & Sorenson 2014). Sequences were manually edited (using Geneious), which involved truncating sequences that had overhanging end gaps from indels and sequences that had two or more

polymorphisms within the last five nucleotides of a sequence. Edited genotypes were imported into the original data files, and genotyping was repeated to generate input files for downstream analyses.

Individuals were sexed by comparing read depth of putative Z chromosome genes (those that aligned to the mallard Z chromosome) with autosomal genes. Because females only have a single copy of a Z-chromosome (ZW) and males have two copies (ZZ), females are expected to have half the number of reads for Z-loci compared to autosomal loci, whereas males will have similar number of reads between these marker types. Heterozygosity was also examined to determine sex, as females should theoretically have zero heterozygous Z-loci.

Using the program Adegenet in R (Jombart 2008), a principal component analysis (PCA) was performed and the first two principal components plotted. The program ADMIXTURE 1.3 (Alexander *et al.* 2009) and STRUCTURE 2.2.3 (Pritchard *et al.* 2000) were used to assign individuals to populations (Peters *et al.* 2016). We included all parsimony informative loci (at least two alleles were represented by a minimum of two copies) while excluding loci with greater than 10% missing data. No a priori information about species was used. Bi-allelic single nucleotide polymorphisms (SNP) were used in ADMIXTURE (Lavretsky *et al.* 2015), whereas the full alleles were used for STRUCTURE (Pritchard *et al.* 2000). We ran 250,000 burn-in and 1,000,000 sampling generations in STRUCTURE and replicated the analysis three times.

Results

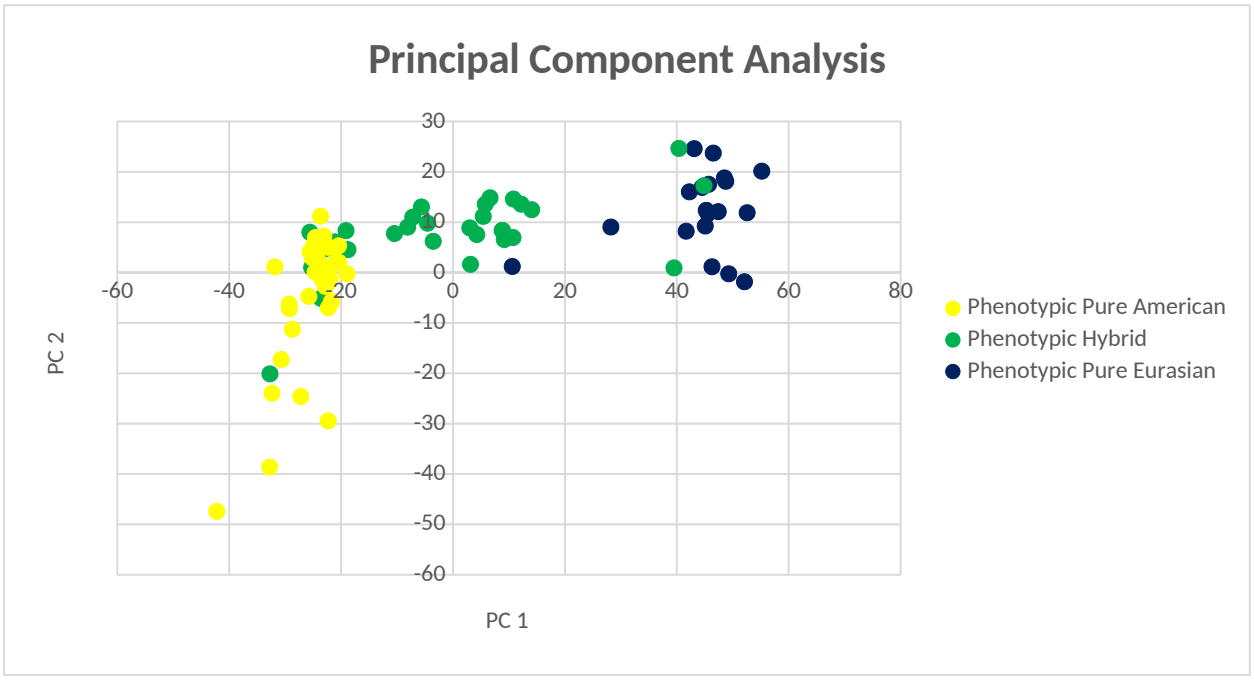
A total of 3095 loci were used that were of sufficient quality—present in >90% of individuals and containing ≤5% flagged genotypes. In the PCA plot (fig. 1a), the first principal component (PC1) clearly separated the two species: the “pure” American Wigeons clustered on

the left of the graph and the “pure” Eurasian Wigeons clustered on the right. Putative hybrids showed mixed results; 19 clustered at an intermediate distance between the two parentals, 12 clustered with American Wigeons, and 3 clustered with Eurasian Wigeons. The large number of putative hybrids grouping with the two parentals suggests that they were misidentified. Two phenotypically “pure” Eurasian wigeons appear clustered with the intermediate-distance hybrids.

Subjectively examining the grouping of the hybrids suggests a mix of F1 hybrids and F2 backcrosses (fig. 1b). Assuming that the central-most cluster of 12 individuals represents F1 hybrids, the six individuals that cluster between this group and “pure” American Wigeons are likely F2-backcrosses into American Wigeon (F2-American). Likewise, there is evidence of one F2-backcross into Eurasian Wigeon (F2-Eurasian) that had an intermediate PC1 value between the F1 hybrid and Eurasian Wigeon clusters.

STRUCTURE assigned nine individuals as F2-backcrosses into Eurasian Wigeons (between 62.5% and 87.5% assignment to the Eurasian population; fig. 2a & 2c), whereas ADMIXTURE assigned two F2-backcrosses (fig. 2b & 2c). In addition, ADMIXTURE assigned 14 F1-hybrids (between 37.5% and 62.5% assignment probability), whereas STRUCTURE assigned six individuals to this group. Finally, ADMIXTURE and STRUCTURE both found five F2-backcrosses into American Wigeons (between 12.5% and 37.5%).

a)



b)

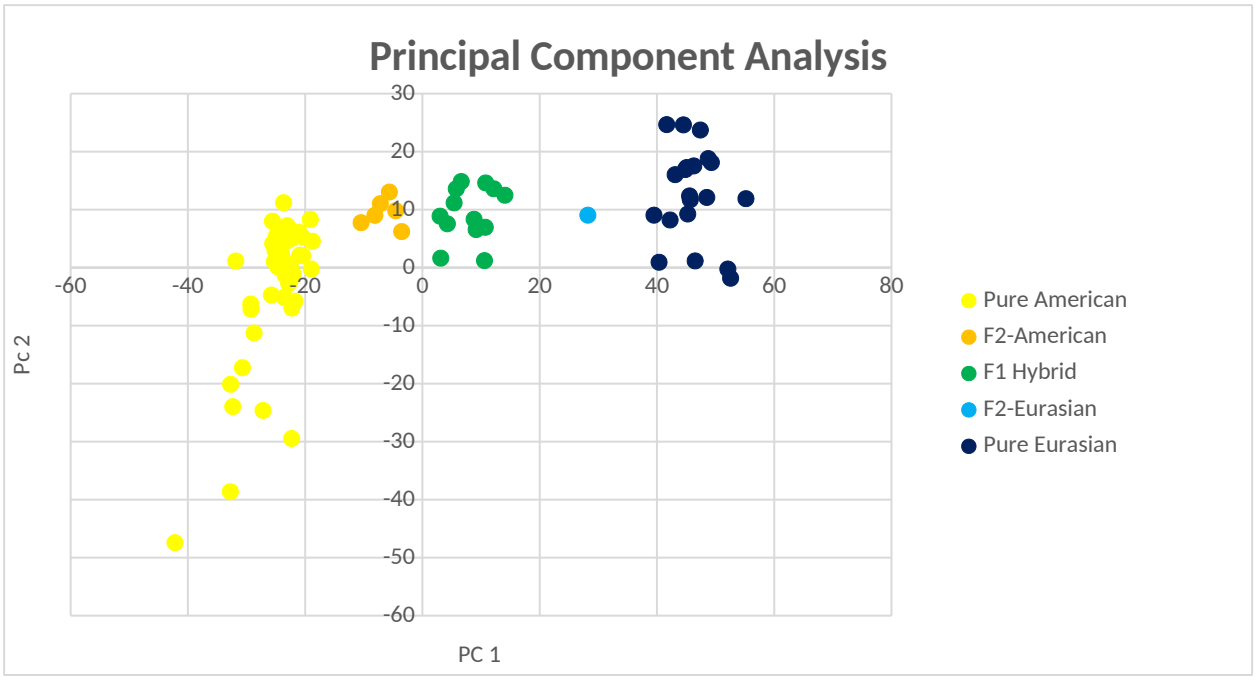
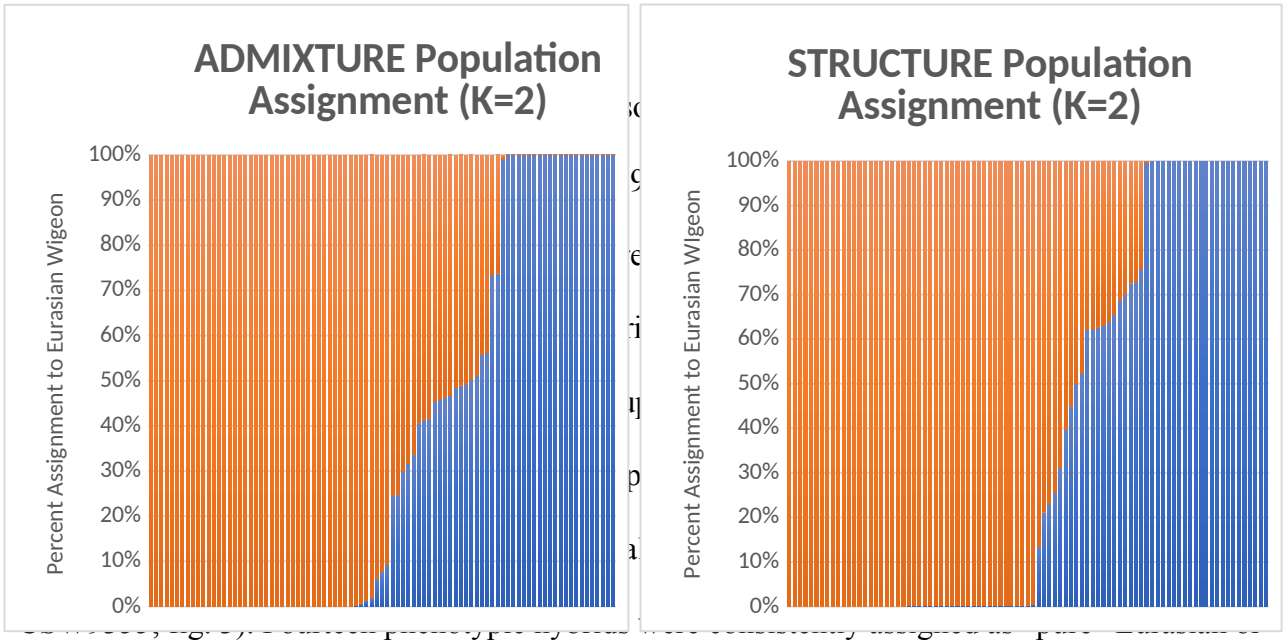
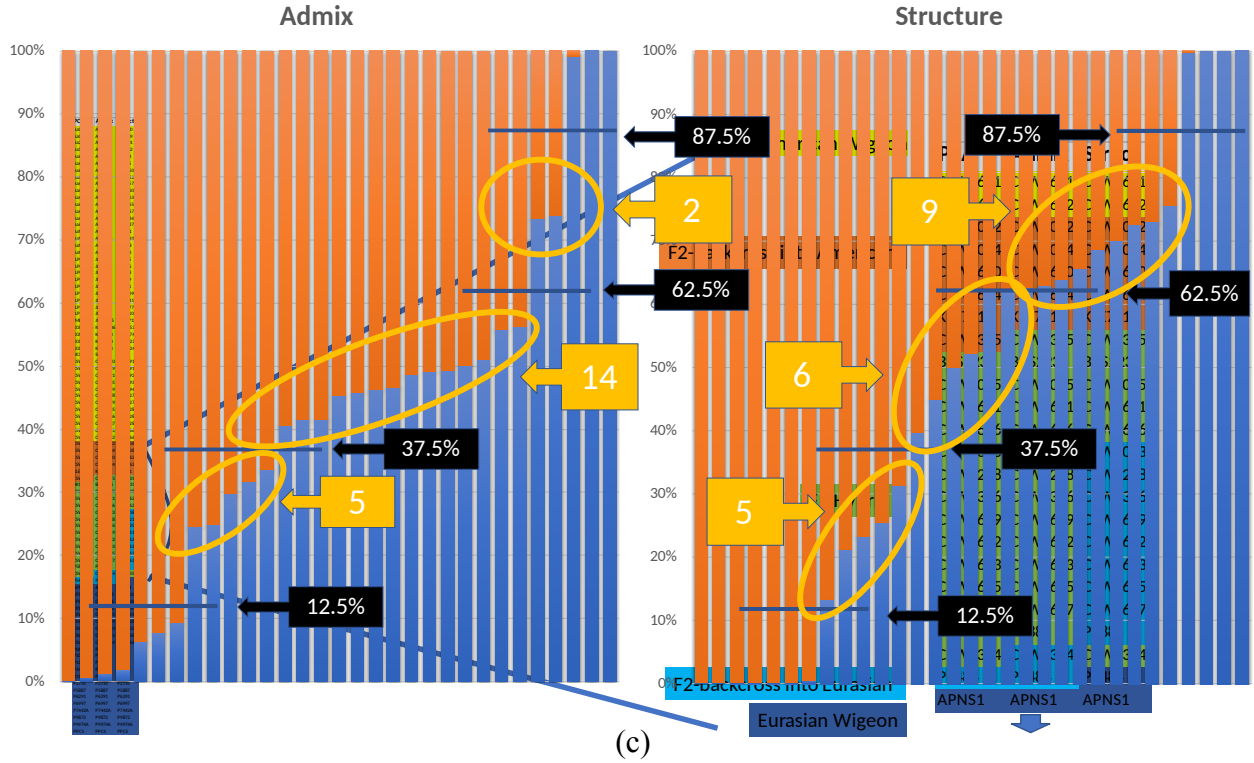


Fig. 1. (a) Principal Component Analysis (PCA) of the 3095 ddRAD-seq loci recovered from 89 American and Eurasian Wigeons and their hybrids. Each circle represents an individual, and the color of circles indicate phenotypic assignments from prepared wings and skins. Two prominent outliers (less than -150 on the PC2 axis) were removed. (b) Colors indicate our classification of "pure" parentals, F1 hybrids, and F2-backcrosses based on clusters within the PC1 plot.



(a) (b)

American in all three tests.



(c)

Fig. 2. (a) ADMIXTURE population assignments based off 3095 ddRAD-seq loci recovered from the 89 American, Eurasian, and hybrid individuals. Colors indicate population assignment, with American Wigeon in orange and Eurasian Wigeon in blue. (b) STRUCTURE population assignments. (c) Comparison of population assignments among the different hybrid classifications with exclusion of individuals near 100% assignment probability.

Fig. 3. Comparison of the hybrid grouping assignments among the three tests; the full results in color-coordinated columns on the left and an enlargement zooming in on assignment disagreement on the right. The PCA colors were assigned using cluster location as in fig. 1b. Yellow indicates an assignment closest to "pure" American Wigeon (100-87.5% American assignment); orange indicates F2-American (87.5-62.5%); green indicates an F1 hybrid (62.5-37.5%); light blue indicates F2-Eurasian (37.5-12.5%); and dark blue indicates "pure" Eurasian Wigeon (12.5-0%). Each row is an individual sample.

Discussion

The objective of this study was accomplished as PCA, ADMIXTURE, and STRUCTURE all agreed on five F2-American Wigeon backcrosses (PCA was the only analysis that differed by assigning a sixth individual; fig. 3). This confirmed the hypothesis that American x Eurasian F1 hybrids backcross into parental populations. In addition, one F2-Eurasian Wigeon was detected by both PCA and ADMIXTURE, although it was categorized as a pure Eurasian using STRUCTURE (Fig. 3).

Sampling putative hybrids within North America may have caused the detection of relatively fewer F2-Eurasians than F2-Americans (one compared to five, respectively; fig. 3). This is because the Eurasian Wigeon population in North America is significantly smaller than the American Wigeon population, so an F1 hybrid would have many more mating opportunities with "pure" American Wigeons. The low number of F2-Eurasians detected could have also resulted from decreased fitness when F1 hybrids mate with parental Eurasians, a potential form of hybrid breakdown (Oka *et al.* 2004). This breakdown often involves a complex interaction of different genes (Oka *et al.* 2004) and can be the result of an incompatibility between cytoplasmic genes and nuclear genes (Breeuwer & Werren 1995). However, STRUCTURE assigned nine F2-Eurasians, whereas PCA and ADMIXTURE assigned one and two, respectively. Although the reason for the disagreeing results requires additional work, it is possible that STRUCTURE is less suited to identifying hybridization when there are fewer "examples" of a particular hybrid class.

There were two genetically dissimilar outliers (less than -150 on the PC2 axis) in the original data set that were labeled as American x Eurasian Wigeon hybrids by our collaborators but were likely hybrids of an American or Eurasian Wigeon with a third species, like a mallard

(*A. platyrhynchos*). In addition, there is some significant spread along the PC2 axis observed among the "pure" American group; this higher variation in PC2 for the American Wigeon population indicates greater genetic variance among North American *Anas americana* than among North American *Anas penelope* (potentially because of a larger effective population size for American Wigeons).

Many putative hybrids were assigned as genetically "pure" in fig. 1a. (12 into American and three into Eurasian). Also, ADMIXTURE and STRUCTURE assigned these 15 individuals to parental populations with high posterior probability, which suggests some individuals were misidentified as hybrids and that phenotypic characteristics can be an unreliable indicator of

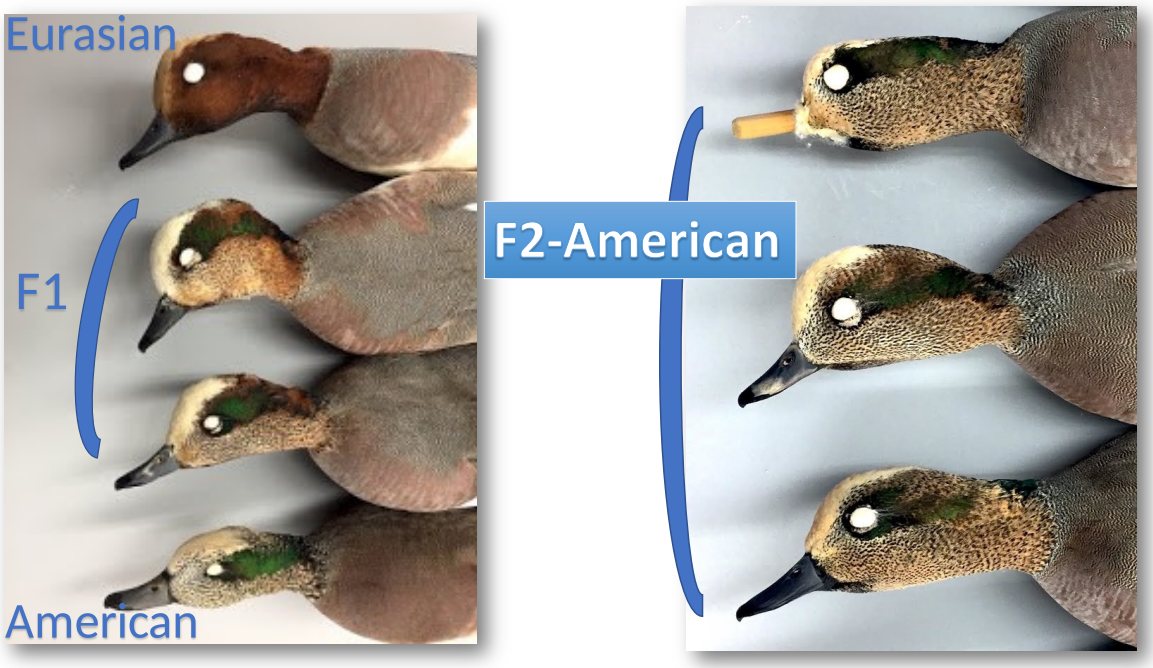


Fig. 4. Images of a subsample of specimens used in this study (starting at top left and going down) of a "pure" Eurasian male, two F1 hybrids, and a "pure" American male. The three samples on the right are images of male F2-American samples. Notice the similarity between the F2-American and F1 hybrids with the green coloration, and also notice the strange occurrence (which may be due to varying lighting) of the F2-American hybrids having *less* green on the head than the F1 hybrids. Adapted from images found at the Burke Museum's website: http://hybridduck.blogspot.com/p/hybrid-gallery.html

hybridization. This underlines the importance of this genetic study because it is likely not accurate (only 58% of the time) to rely on phenotypic identification of these hybrids that look very similar to each other (see fig. 4).

The observations that no single sample had three different assignments among the three tests and that the assignment of 77 of the 89 total samples (86.5%) agreed among all tests suggests our methods have substantial power for detecting hybrids. The "pure" populations also showed strong agreement with only one "pure" result being disputed among the tests (P5484 in fig. 3.). Despite generally strong agreement among programs, the tests showed some points of contention. STRUCTURE had 10 unique assignments that did not agree with any other test, whereas ADMIXTURE and PCA had zero and one, respectively. Overall, STRUCTURE had the most disagreement out of the tests and detected more Eurasian-backcrosses than did the other two programs. The ducks that STRUCTURE assigned as F2-Eurasian were assigned as F1 hybrids in the other two tests (fig. 2c). Furthermore, the one F2-Eurasian sample PCA and ADMIXTURE detected was assigned as a "pure" Eurasian individual by STRUCTURE with high posterior probability (99.7%). The difference in assignments likely cannot be explained by differences in data input; PCA and STRUCTURE examined whole loci whereas ADMIXTURE examined individual SNPs (single nucleotide polymorphisms), yet PCA and ADMIXTURE had the highest agreement. Thus, the STRUCTURE results seem to be biased toward over-assignment to the Eurasian population. Additional research is needed to determine which assumptions are causing the differences observed.

A weakness of this study was that the ddRAD-seq protocol resulted in less than 0.1% of the genome being examined. A more comprehensive (and more resource-heavy) analysis may be needed if the genetic basis upon which these species are separated (speciation genes in the genic

species model) is to be investigated. Another potential weakness was that individuals that may have been F3 (and above) backcrosses might not have been detected. Specifically, ADMIXTURE identified 3 individuals that had 6.2%, 7.7%, and 9.3% assignment to the Eurasian Wigeon population, so these samples may have been F3 (expected genomic ratio of 12.5%:87.5%) or F4 (6.25%:93.75%) backcrosses into American Wigeon. Methodology exists to identify these individuals in future analyses (Lavretsky *et al.* 2016).

With the agreement among the three tests on five F2-Americans, introgression is supported, which can lead to some different scenarios in the future. First, reinforcement may occur if these hybrids have lower survival rates (lower fitness) than their "pure" counterparts. This kind of occurrence could come from one gene that, when combined between species in a hybrid, malfunctions (a speciation gene), thus promoting reproductive isolation. Another possibility is the merging of the two species if the hybrids are equally fit. Such speciation reversal has been proposed to have happened between the California and Holarctic Ravens (*Corvus corax*) (Kearns *et al.* 2018). Even if the hybrids are equally fit, though, Wigeon mating preferences may discourage interspecific mating, so the two species may still not experience much gene flow (even if no physical or genetic barriers exist). A third potential outcome is competitive exclusion, where one species outcompetes the other and excludes it from North America. Given the prevalence of Eurasian Wigeon eBird sightings in North America, complete competitive exclusion does not seem likely. However, North American competition with the potentially better-adapted American Wigeon may still be a major limiting factor for the expansion of Eurasian Wigeon. Further research should focus on detecting unequally segregated speciation genes between the bird species, as this will provide more information on the nature by which these birds differ evolutionarily.

The finding that gene flow occurs between two separate species is consistent with the genic species concept (which allows for varying levels of introgression) and is not consistent with the functional definition of Mayr's Biological Species Concept (which relies on sterile F1 hybrids). According to the popular BSC, American and Eurasian Wigeons should be the same species since they produce fertile F1 hybrids. When using the principle that geographic data and tenets of the MSC and BSC should be used in addition to genetic data (Harvey & Brumfield 2015), a complicated picture is created. One mtDNA analysis suggests they are not even sister species (Johnson & Sorenson 1999), while another finds their nuclear DNA to be indistinguishable from each other (Peters *et al.* 2014b). This study found direct evidence of gene flow and F2-backcrossing, but it is quite apparent the two ducks are phenotypically distinct. With increasing geographic overlap resulting in more opportunities for hybridization, future examination of possible speciation genes will provide further insight into the process of biological speciation for these ducks and for biodiversity in general.

Bibliography

Alexander DH, Novembre J, Lange K (2009) Fast model-based estimation of ancestry in unrelated individuals. *Genome Research,* **19**, 1655-1664.

Allendorf F, Leary R, Spruell P, Wenburg J (2001) The problems with hybrids: setting conservation guidelines. *Trends in Ecology & Evolution,* **16**, 613-622.

Altschul S, Gish W, Miller W, Myers E, Lipman D (1990) Basic local alignment search tool. *Journal of Molecular Biology,* **215**, 403-410.

Breeuwer J, Werren J (1995) Hybrid breakdown between 2 haplodiploid species - the role of nuclear and cytoplasmic genes. *Evolution,* **49**, 705-717.

DaCosta JM, Sorenson MD (2014) Amplification biases and consistent recovery of loci in a double-digest rad-seq protocol. *Plos one,* **9**, 1-14.

Dasmahapatra KK, Walters JR, Briscoe AD, et al (2012) Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature,* **487**, 94-98.

DeMarais BD, Dowling TE, Douglas ME, Minckley WL, Marsh PC (1992) Origin of Gila seminuda (teleostei: cyprinidae) through introgressive hybridization: implications for evolution and conservation. *Proceedings of the National Academy of Sciences of the United States of America,* **89**, 2747.

Edgar R (2004) Muscle: a multiple sequence alignment method with reduced time and space complexity. *Bmc Bioinformatics,* **5**, 1-19.

Fowler AC, Eadie JM, Engilis A, Jr. (2009) Identification of endangered Hawaiian ducks (*Anas wyvilliana*), introduced North American Mallards (*A. platyrhynchos*) and their hybrids using multilocus genotypes. *Conservation Genetics,* **10**, 1747-1758.

Giraud T, Refrégier G, Le Gac M, de Vienne DM, Hood ME (2008) Speciation in fungi. *Fungal Genetics & Biology,* **45**, 791-802.

Gonzalez J, Duettmann H, Wink M (2009) Phylogenetic relationships based on two mitochondrial genes and hybridization patterns in Anatidae. *Journal of Zoology,* **279**, 310-318.

Harvey MG, Brumfield RT (2015) Genomic variation in a widespread neotropical bird (*Xenops minutus*) reveals divergence, population expansion, and gene flow. *Molecular Phylogenetics and Evolution,* **83**, 305-316.

Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics,* **167**, 747-760.

Johnson KP, Sorenson MD (1999) Phylogeny and biogeography of dabbling ducks (genus: Anas): a comparison of molecular and morphological evidence. *Auk.,* **116**, 792-805.

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

Kearns AM, Restani M, Szabo I, et al (2018) Genomic evidence of speciation reversal in ravens. *Nature Communications,* **9**, 906.

Kulikova IV, Zhuravlev YN (2010) Genetic structure of the far eastern population of Eurasian wigeon *Anas penelope* inferred from sequencing of the mitochondrial DNA control region. *Russian Journal of Genetics,* **46**, 976-981.

Kulikova I, Zhuravlev Y, McCracken K (2004) Asymmetric hybridization and sex-biased gene flow between Eastern Spot-billed ducks (*Anas zonorhyncha*) and Mallards (*A-platyrhynchos*) in the Russian far east. *Auk,* **121**, 930-949.

Lavretsky P, Dacosta JM, Hernandez-Banos BE, Engilis A, Jr., Sorenson MD, Peters JL (2015) Speciation genomics and a role for the Z chromosome in the early stages of divergence between Mexican ducks and mallards. *Molecular ecology,* **24**, 5364-5378.

Lavretsky P, Peters JL, Winker K, et al (2016) Becoming pure: identifying generational classes of admixed individuals within lesser and greater scaup populations. *Molecular ecology,* **25**, 661-674.

Livezey B (1991) A phylogenetic analysis and classification of recent dabbling ducks (tribe Anatini) based on comparative morphology. *Auk,* **108**, 471-507.

Mayr E (1996) What is a species, and what is not? *Philosophy of Science,* **63**, 262-277.

Mayr E (1963) *Animal species and evolution* Cambridge, Belknap Press of Harvard University Press, 1963.

Michael CRE (1984) Trans-hemispheric movements of holarctic Anatidae: the Eurasian wigeon (*Anas penelope L.*) in North America. *Journal of Biogeography,* **11**, 27.

Mihola O, Trachtulec Z, Vlcek C, Schimenti JC, Forejt J (2009) A mouse speciation gene encodes a meiotic histone H3 methyltransferase. *Science,* **323**, 373.

Miller MR, Dunham JP, Amores A, Cresko WA, Johnson EA (2007) Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome research,* **17**, 240-248.

Moore WS, Buchanan DB (1985) Stability of the Northern Flicker hybrid zone in historical times: implications for adaptive speciation theory. *Evolution,* **39**, 135.

Nei M, Maruyama T, Wu C (1983) Models of Evolution of Reproductive Isolation. *Genetics,* **103**, 557-579.

Nosil P (2012) The speciation continuum: what factors affect how far speciation proceeds? In: *Ecological speciation* (eds. Harvey P, May R, Godfrary C, Dunne J) pp. 192-213 Oxford University Press, New York.

Oka A, Mita A, Sakurai-Yamatani N, et al (2004) Hybrid breakdown caused by substitution of the X chromosome between two mouse subspecies. *Genetics,* **166**, 913-924.

Pennock DS, Dimmick WW (1997) Critique of the evolutionarily significant unit as a definition for "distinct population segments" under the US Endangered Species Act. *Conservation Biology,* **11**, 611-619.

Peters JL, McCracken KG, Zhuravlev YN, et al (2005) Phylogenetics of wigeons and allies (Anatidae: *Anas*): the importance of sampling multiple loci and multiple individuals. *Molecular Phylogenetics and Evolution,* **35**, 209-224.

Peters JL, Lavretsky P, DaCosta JM, Bielefeld RR, Feddersen JC, Sorenson MD (2016) Population genomic data delineate conservation units in mottled ducks (*Anas fulvigula*). *Biological Conservation,* **203**, 272-281.

Peters JL, Sonsthagen SA, Lavretsky P, Rezsutek M, Johnson WP, McCracken KG (2014a) Interspecific hybridization contributes to high genetic diversity and apparent effective population size in an endemic population of mottled ducks (*Anas fulvigula maculosa*). *Conservation Genetics,* **15**, 509-520.

Peters JL, Winker K, Millam KC, et al (2014b) Mito-nuclear discord in six congeneric lineages of holarctic ducks (genus Anas). *Molecular ecology,* **23**, 2961-2974.

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

Rašic G, Filipovic I, Weeks AR, Hoffmann AA (2014) Genome-wide SNPs lead to strong signals of geographic structure and relatedness patterns in the major arbovirus vector, *Aedes aegypti*. *BMC Genomics,* **15**, 275-275.

Rhymer J, Williams M, Braun M (1994) Mitochondrial analysis of gene flow between New-Zealand Mallards (*Anas-platyrhynchos)* and Grey-Ducks (*A-superciliosa*). *Auk,* **111**, 970-978.

Taylor JW, Jacobson DJ, Kroken S, et al (2000) Phylogenetic species recognition and species concepts in fungi. *Fungal Genetics and Biology,* **31**, 21-32.

Wiley EO (1978) The evolutionary species concept reconsidered. *Systematic Zoology,* **27**, 17.

Williams C, Brust R, Fendley T, Tiller G, Rhodes O (2005) A comparison of hybridization between mottled ducks (*Anas fulvigula*) and mallards (*A-platyrhynchos)* in Florida and South Carolina using microsatellite DNA analysis. *Conservation Genetics,* **6**, 445-453.

Wu C (2001) The genic view of the process of speciation. *Journal of Evolutionary Biology,* **14**, 851-865.