MOLECULAR ECOLOGY WILEY

Assessing changes in genomic divergence following a century of human-mediated secondary contact among wild and captive-bred ducks

Philip Lavretsky[1] | Nancy Rotzel McInerney[2] | Jonathon E. Mohl[3] | Joshua I. Brown[1] | Helen F. James[4] | Kevin G. McCracken[5,6,7,8] | Robert C. Fleischer[2]

[1]Department of Biological Sciences, The University of Texas at El Paso, El Paso, TX, USA

[2]Center for Conservation Genomics, Smithsonian Conservation Biology Institute, Smithsonian Institution, Washington, DC, USA

[3]Border Biomedical Research Center, The University of Texas at El Paso, El Paso, TX, USA

[4]Department of Vertebrate Zoology, National Museum of Natural History, Smithsonian Institution, Washington, DC, USA

[5]Department of Biology, University of Miami, Coral Gables, FL, USA

[6]Department of Marine Biology and Ecology, Rosenstiel School of Marine and Atmospheric Sciences, University of Miami, Miami, FL, USA

[7]Human Genetics and Genomics, Hussman Institute for Human Genomics, University of Miami Miller School of Medicine, Miami, FL, USA

[8]Institute of Arctic Biology, University of Alaska Museum, University of Alaska Fairbanks, Fairbanks, AK, USA

Correspondence
Philip Lavretsky, Department of Biological Sciences, The University of Texas at El Paso, El Paso, TX 79668, USA.
Email: plavretsky@utep.edu

Funding information
NIH, Grant/Award Number: 5G12RR007592; James A. Kushlan endowment in Waterbird Biology and Conservation at the University of Miami; Black Duck Joint Venture

Abstract

Along with manipulating habitat, the direct release of domesticated individuals into the wild is a practice used worldwide to augment wildlife populations. We test between possible outcomes of human-mediated secondary contact using genomic techniques at both historical and contemporary timescales for two iconic duck species. First, we sequence several thousand ddRAD-seq loci for contemporary mallards (*Anas platyrhynchos*) throughout North America and two domestic mallard types (i.e., known game-farm mallards and feral Khaki Campbell's). We show that North American mallards may well be becoming a hybrid swarm due to interbreeding with domesticated game-farm mallards released for hunting. Next, to attain a historical perspective, we applied a bait-capture array targeting thousands of loci in century-old (1842–1915) and contemporary (2009–2010) mallard and American black duck (*Anas rubripes*) specimens. We conclude that American black ducks and mallards have always been closely related, with a divergence time of ~600,000 years before present, and likely evolved through prolonged isolation followed by limited bouts of gene flow (i.e., secondary contact). They continue to maintain genetic separation, a finding that overturns decades of prior research and speculation suggesting the genetic extinction of the American black duck due to contemporary interbreeding with mallards. Thus, despite having high rates of hybridization, actual gene flow is limited between mallards and American black ducks. Conversely, our historical and contemporary data confirm that the intensive stocking of game-farm mallards during the last ~100 years has fundamentally changed the genetic integrity of North America's wild mallard population, especially in the east. It thus becomes of great interest to ask whether the iconic North American mallard is declining in the wild due to introgression of maladaptive traits from domesticated forms. Moreover, we hypothesize that differential gene flow from domestic game-farm mallards into the wild mallard population may explain the overall temporal increase in differentiation between wild black ducks and mallards, as well as the uncoupling of genetic diversity and effective population size estimates across time in our results. Finally, our findings highlight how genomic methods can recover complex population histories by capturing DNA preserved in traditional museum specimens.

1 | INTRODUCTION

Natural or anthropogenically induced changes in species' ranges (expansion, contraction, or shifts) have recently brought many closely related species into secondary contact, raising concern over the possibility of extinction by introgressive hybridization in various taxa (Abbott, Barton, & Good, 2016; Payseur & Rieseberg, 2016; Todesco et al., 2016; Wayne & Shaffer, 2016). Such interactions are further complicated when artificially selected-on domesticated forms come into contact with their wild congeners (Domyan et al., 2014; Kidd, Bowman, Lesbarreres, & Schulte-Hostedde, 2009; Makino et al., 2018; Skoglund, Ersmark, Palkopoulou, & Dalén, 2015; Tufto, 2017; Wu et al., 2018). Where the release of domesticated stocks is common and intensive, interbreeding with released or feral conspecifics can cause a decrease in genetic variation, leading to a loss of adaptive potential and overall fitness within the wild population (Frankham, 2005; Laikre, Schwartz, Waples, & Ryman, 2010). Here, we use DNA from archived specimens to evaluate whether the human-mediated introduction of domestic and movement of wild mallard (*Anas platyrhynchos*) populations has altered the genetic integrity of mallard populations.

The mallard is a Holarctic species whose range has expanded globally due to changing habitat and direct augmentation of populations with domesticated variants (Baldassarre, 2014). In North America, range expansions since the early 1900s have extended the mallard's range east of the Mississippi to include much of eastern North America (Osborne, Swift, & Baldassarre, 2010; USFWS, 2013). This expansion was likely due to (a) direct augmentation by game managers, sportsmen and others, who have released ~500,000 captive-bred mallards per year along the eastern seaboard since the 1920s (Hepp, Novak, Scribner, & Stangel, 1988; Heusmann, 1974; Soutiere, 1986) and have continued to release ~210,000 birds annually in recent years (Osborne et al., 2010; USFWS, 2013), and (b) conversion of boreal forests into open, prairie-like habitat, allowing mallards to naturally expand eastward in the 1950s (e.g., southern Ontario; Hanson, Rogers, & Rogers, 1949). Using genomewide markers, Lavretsky, Janzen, and McCracken (2019) recently identified two mallard genetic clusters in North America, a "Western" mallard genetic cluster to which all mallards west of the Mississippi River were assigned and a "Non-Western" mallard cluster that was largely recovered in mallards from eastern North America. Their results suggest that historical releases and habitat changes have fostered gene flow between game-farm and wild populations of mallards, and even with their eastern North American sister species, the American black duck (*Anas rubripes*; "black duck").

Prior to expansion of the mallard's range in the 1950s, mallards and American black ducks were largely allopatric, being found west and east of the Mississippi River, respectively (Johnsgard, 1967; Merendino & Ankney, 1994; Snell, 1986). Reports of hybridization between the species in secondary contact raised concerns that the iconic eastern black duck could face genetic extinction (Mank, Carlson, & Brittingham, 2004; Rhymer, 2006; Rhymer & Simberloff, 1996). Employing three microsatellite markers and ancient DNA (aDNA) methods, Mank et al. (2004) reported an 18-fold reduction in differentiation (Gst) between mallards and black ducks sampled in 1998 (0.008) versus 1940 (0.146) and concluded that a century of hybridization must have led to a loss of genetic distinctiveness. More recently, Lavretsky, Janzen, et al. (2019) applied double-digest restriction site-associated DNA sequencing (ddRAD-seq) to a large and geographically broad sample set and achieved the ability to genetically distinguish between contemporary mallards, black ducks and their hybrids despite their overall low levels of genomic differentiation (composite Φ_{ST} ~ 0.01). The authors demarcated putatively non-neutrally evolving (i.e., outlier) regions across several chromosomes. These regions of high differentiation appear to persist despite potentially a century of extensive introgressive hybridization between mallards (wild and/or domestic) and black ducks. This suggests that if secondary contact has amalgamated much of the genomes of black ducks and mallards, it has perhaps also allowed the divergence and speciation process to progress (Edwards et al., 2005; Gourbière & Mallet, 2010; Nosil, Harmon, & Seehausen, 2009; Nosil & Schluter, 2011; Servedio & Noor, 2003).

In this article, we first aim to determine the source of the recently described non-Western mallard nuclear genetic signature now widespread in eastern North American (Lavretsky, Janzen, et al., 2019; also see Figure 2a). We note that 75% of North American mallards with substantial genetic assignment (≥10%) to the non-Western mallard nuclear cluster also possessed Old World (OW) A as opposed to New World (NW) B mitochondrial haplotypes (Lavretsky, Janzen, et al., 2019). The Eurasian origins of the OW A mitochondrial haplogroup (Avise, Ankney, & Nelson, 1990; Johnson & Sorenson, 1999; Kulikova et al., 2005; Lavretsky, McCracken, & Peters, 2014) suggest that non-Western nuclear variation is likely also of Eurasian descent. We explicitly test whether the two genetic mallard clusters in North America are due to natural (i.e., divergence within wild mallards) or artificial (i.e., anthropogenic release of domesticated game-farm mallards) processes. The former seems unlikely as wild mallards show no discernible structure across their Holarctic distributions (Champagnon et al., 2013; Kraus et al., 2011; Söderquist et al., 2017), with the only genetic structure within Eurasian mallards due to hybridization with captive-bred mallards (Söderquist et al., 2017). Furthermore, tracking data show that when mallards migrate between Eurasia

and North America, the path used is eastward through Alaska and not westward across the Atlantic Ocean (Hupp et al., 2011; Lam et al., 2012; Shin et al., 2015; Yamaguchi et al., 2008). The number of mallards moving between continents is fairly limited (Pearce et al., 2011), and thus, gene flow between mallards moving across the Atlantic Ocean and extensively interbreeding is unlikely. Together, these movement patterns would not explain either the geographical location or high representation of the secondary mallard cluster in North America.

To test whether game-farm mallards released by hunting preserves have bred with wild populations and evaluate the origin of the secondary "Non-Western" mallard cluster in North America, we sequenced thousands of ddRAD-seq nuclear loci and the mitochondrial control region across known domestic mallard breeds (i.e., game-farm and feral "park" mallards) and compared them with previously published wild North American mallard sequences (Lavretsky, Janzen, et al., 2019). If gene flow between game-farm and wild mallards underlies the non-Western cluster in eastern North America, we expect (A) all game-farm mallards will possess OW A mitochondrial haplotypes and population assignment to the non-Western nuclear mallard cluster, (B) recovered OW A mitochondrial haplotypes within contemporary mallards will be shared with those found in game-farm mallards, and (C) both OW A mitochondrial haplotypes and non-Western nuclear mallard variation will be absent in historical mallards from before large-scale game-farm releases (i.e., pre-1920).

Next, we revisit the hypothesis that the black duck has become, or is in danger of becoming, a hybrid swarm (Mank et al., 2004). Using a bait-capture array based on ddRAD sequences, we targeted thousands of loci in century-old museum specimens, plus contemporary mallards and American black ducks. We evaluate four possible scenarios in which secondary contact over the last century results in (A) a temporal decrease in genomic differences except at regions under strong selection, (B) a decrease in differentiation across the genome, (C) a history of divergence-with-gene flow (Feder, Egan, & Nosil, 2012) or (D) recurring cases of secondary contact. If contemporary gene flow (i.e., in the last 100–200 years) is the proximate cause for the shared genetic variation between mallards and black ducks, then either scenario A or B should be supported. Under scenario A, we also expect to see similar or higher estimates of differentiation at non-neutral, outlier loci in historical and contemporary data sets, consistent with species reinforcement. Conversely, an overall decrease in genetic differentiation across the genomes of black ducks and mallards between the historical and modern time periods would support genomic swamping (scenario B). Gene flow between mallards and black ducks that is continuous (i.e., divergence-with-gene flow) or intermittent (i.e., recurring secondary contact) through evolutionary time would support scenario C or D, respectively. In addition to comparing the genomic landscapes of historical versus contemporary mallards and black ducks, we more formally tested evolutionary hypotheses by independently analysing historical and
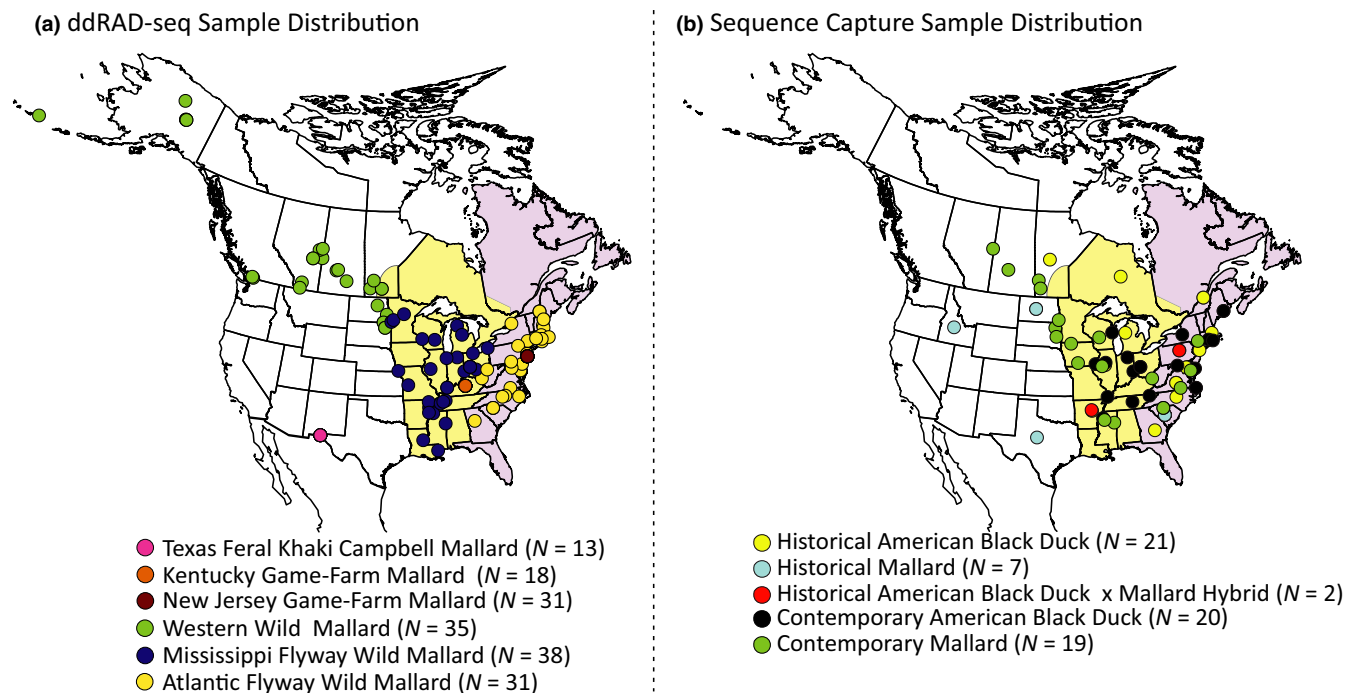

(a) ddRAD-seq Sample Distribution

- 🔴 Texas Feral Khaki Campbell Mallard (*N* = 13)
- 🟠 Kentucky Game-Farm Mallard (*N* = 18)
- 🟤 New Jersey Game-Farm Mallard (*N* = 31)
- 🟢 Western Wild Mallard (*N* = 35)
- 🔵 Mississippi Flyway Wild Mallard (*N* = 38)
- 🟡 Atlantic Flyway Wild Mallard (*N* = 31)

(b) Sequence Capture Sample Distribution

- 🟡 Historical American Black Duck (*N* = 21)
- 🔵 Historical Mallard (*N* = 7)
- 🔴 Historical American Black Duck x Mallard Hybrid (*N* = 2)
- ⚫ Contemporary American Black Duck (*N* = 20)
- 🟢 Contemporary Mallard (*N* = 19)

FIGURE 1 Sample size and distributions of (a) ddRAD-seq and (b) sequence capture data sets. For reference, North America is portioned into three geographical regions for waterfowl: areas west of the Mississippi River (a.k.a "Western"; white), the Mississippi Flyway (light orange) and the Atlantic Flyway (light purple). Note that only the ddRAD-seq mallard samples are partitioned geographically, whereas the sequence capture samples are partitioned by species (American black duck vs. mallard) and time (Historical vs. Contemporary). See Table S1 for additional sample information
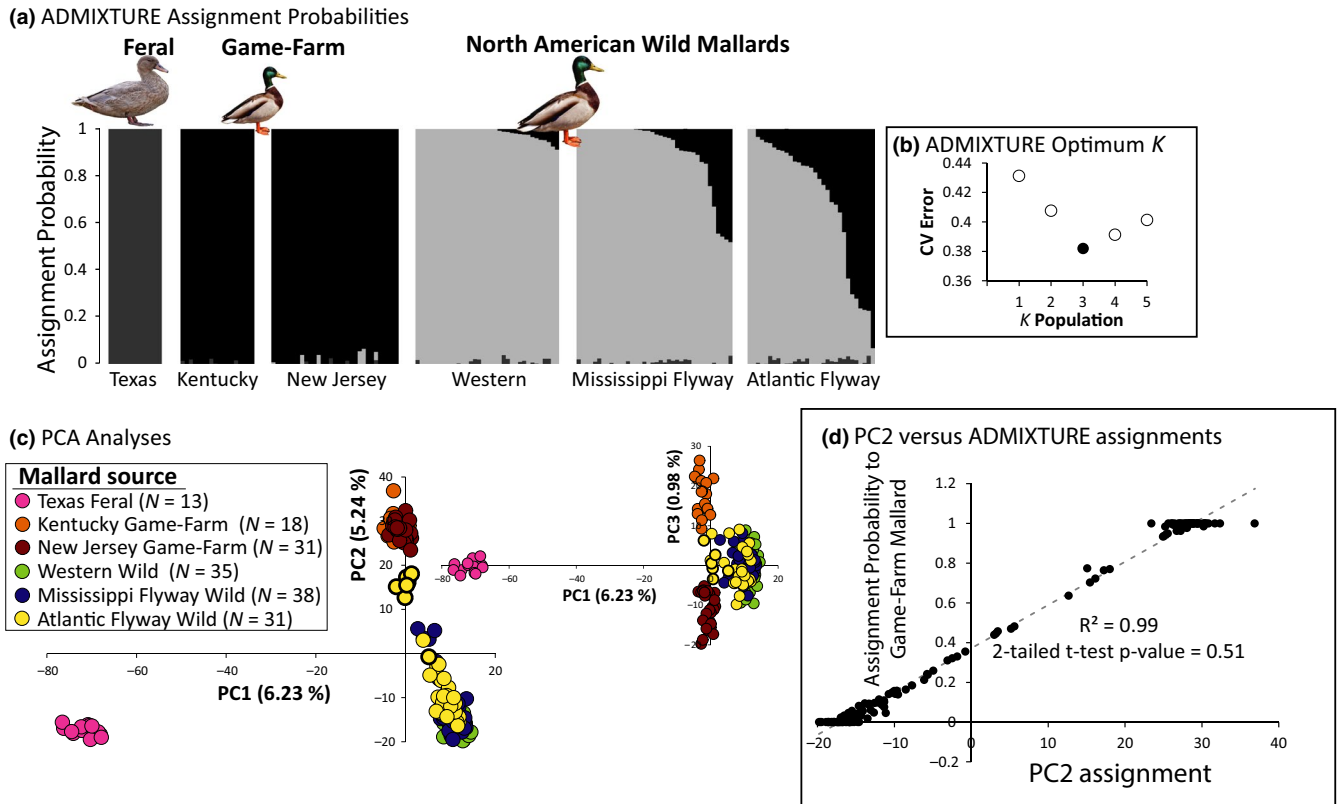
FIGURE 2 (a) ADMIXTURE assignment probabilities for known domestic lineages and wild samples of the mallard estimated (b) at the optimum K population of 3. Wild mallard samples are geographically categorized as "Western," or from the Mississippi Flyway (MISS), or from the Atlantic Flyway (ATL). Above ADMIXTURE results are photographs representative of the male plumage for the domestic and wild mallard lineages. (c) PCA results for the same set of samples, with plots comparing PC1 versus PC2 and PC1 versus PC3 scores. Darkened circles in the PCA plot denote the six wild Atlantic flyway mallards for which sequence capture data were also obtained and analysed (Figure 5b). (d) R^2 correlation between ADMIXTURE assignment probability and PC2 scores across samples. Note that the two-tailed t test finds that PCA and ADMIXTURE assignment per sample are statistically similar

contemporary data sets against various evolutionary models (e.g., split-migration, isolation-with-migration) using the program $\partial a \partial i$ (Gutenkunst, Hernandez, Williamson, & Bustamante, 2009, 2010). If contemporary (i.e., from secondary contact in last 100–200 years) gene flow between black ducks and mallards has not been the amalgamating force it was once thought to be, or gene flow has occurred over evolutionary time (scenarios C or D), then we expect the optimum model and respective estimated parameters to be the same when analysing historical or contemporary data sets.

2 | MATERIALS AND METHODS

2.1 | Domestic mallard sampling, DNA extraction, ddRAD-seq library preparation and mitochondrial DNA sequencing

Tissue or blood for a total of 62 mallards of known domestic origins was obtained from three states and included (a) game-farm mallards from New Jersey that are actively being released for hunting (N = 31), (b) game-farm mallards from Kentucky raised for dog training

and actively released (N = 18) and (c) other feral mallards from Texas (N = 13) that represent "other mallards" that can potentially interbreed with wild populations (Figure 1; Table S1). We phenotypically identified the feral mallards from Texas to be Khaki Campbell mallards (Figure 2a), which are domesticated mallards that are often associated with and identified as park ducks (i.e., farm ducks that people release into "natural" environments where they are able to interact with wild populations, henceforward referred to as park mallards). These birds were found living alongside wild mallards and Mexican ducks (*Anas platyrhynchos diazi*) during the breeding season and thus are a source of possible hybridization in a park setting. Blood was obtained during banding operations from live birds for both Texas feral Khaki Campbell's and game-farm mallards from Kentucky, while tissue was opportunistically sampled from hunters who shot game-farm mallards released at a New Jersey hunting preserve.

DNA extraction was done using a DNeasy Blood & Tissue kit, following the manufacturer's protocols (Qiagen). ddRAD-seq library preparation for multiplexing followed steps outlined in Lavretsky et al. (2015) (also see DaCosta & Sorenson, 2014). The samples were pooled in equimolar concentrations, and 150 base pair, single-end sequencing was completed on an Illumina HiSeq 4000 at

the University of Oregon's Genomics & Cell Characterization Core Facility. Raw Illumina reads have been deposited in NCBI's Sequence Read Archive (SRA; http://www.ncbi.nlm.nih.gov/sra; SRA data PRJNA591912). We included sequenced mallard samples from Lavretsky, Janzen, et al. (2019) that were genetically characterized as either "Pure Western Mallard," "Pure Non-Western Mallard" or hybrids of the two (N = 104; NCBI's SRA data PRJNA516035: see Table S1 for BioSample Accession Numbers). These latter samples were from birds harvested by hunters in the wild. Samples spanned Western (west of the Mississippi River), Mississippi Flyway and Atlantic Flyway mallards (Figure 1; Table S1).

Processing of raw Illumina reads followed the computational pipeline described by DaCosta and Sorenson (2014; Python scripts available at http://github.com/BU-RAD-seq/ddRAD-seq-Pipe-line) and following the steps outlined in Lavretsky, Janzen, et al. (2019). Final output files (e.g., FASTA, ADMIXTURE, FINERADSTRUC-TURE) were generated with custom python scripts (Lavretsky et al., 2016), in which alleles were scored as missing if they had less than 5× coverage. Loci with <10% missing genotypes were retained for downstream analyses. Chromosomal positions across loci were attained by aligning a reference sequence across ddRAD markers to a wild mallard genome (Hou et al., unpublished; SRA BioProject PRJNA554956 & BioSample SAMN12287667). Only ddRAD loci that had a single hit to an autosomal chromosome Z sex chromosome were retained for population structure analyses (thus, we excluded Z sex chromosome-linked markers).

Finally, primers L78 and H774 were used to PCR amplify and sequence 653 bp of the mtDNA control region (Sorenson, Ast, Dimcheff, Yuri, & Mindell, 1999; Sorenson & Fleischer, 1996) following dideoxy sequencing methods described in Lavretsky, McCracken, et al. (2014). The PCR products were sequenced on a 3130XL Genetic Analyzer at the University of Texas at El Paso, Border Biomedical Research Center's Genomic Analysis Core Facility. Sequences were aligned and edited using Sequencher v. 4.8 (Gene Codes, Inc). All sequences have been submitted to GenBank (*Accession numbers* MN734444–MN734576).

2.2 | Historical DNA sampling, DNA extraction and library preparation

A total of 40 contemporary (N = 20 black ducks; N = 20 mallards; sample years range from 2009 to 2010) and 43 historical (N = 25 black ducks from 1842 to 1910; N = 15 mallards from 1856 to 1915; and N = 3 putative mallard × black duck hybrids from 1840 to 1910) individuals were sampled (Figure 1; Table S1). Our sampling scheme included regions in which mallards and black ducks are currently sympatric (the Mississippi and Atlantic Flyways), as well as mallards from west of the Mississippi River that are beyond the range of black ducks (Figure S1 and Table S1). All contemporary and historical mallards sampled for sequence capture were from the wild.

For contemporary samples (Table S1), either breast or wing muscle was used for DNA extraction using a DNeasy Blood & Tissue

kit and following the manufacturer's protocols (Qiagen). Extractions were quantified using a NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific Inc.) to ensure a minimum concentration of 0.02 μg/μl. Following extraction, modern DNA samples were sheared with the QSonica R800 Sonicator for 1 hr consisting of 15 s on/off at 25% amplitude to sufficiently fragment the DNA to a size range of 300–600 bp.

For historical samples, toe pad tissue was carefully sampled at the National Museum of Natural History (USNM; Table S1). Considering the limited amount of endogenous nuclear DNA and its potentially degraded condition in historical samples, aDNA extractions were performed in a dedicated ancient laboratory at the National Zoo's Center for Conservation Genomics. We further reduced the chance of cross-contamination by conducting extractions and library preparation on separate days for historical mallards, black ducks and putative hybrids. All tools, buffers and surfaces were first sanitized using standard ancient DNA procedures (e.g., the use of UV irradiation and bleach to destroy contaminating DNA and kill bacteria and fungi) (Keyser-Tracqui & Ludes, 2005; Knapp, Clarke, Horsburgh, & Matisoo-Smith, 2012). DNA isolation followed McDonough, Parker, Rotzel McInerney, Campana, and Maldonado (2018) using silica columns with vacuum extraction, yet with the initial sample incubation in a Tris/EDTA/NaCl buffer and rotated at 55°C overnight. For library preparation of historical and contemporary samples, we used a Kapa Biosystems Library Preparation kit (Illumina, Inc) with custom dual indexed adapters (Faircloth & Glenn, 2012), following protocols by McDonough et al. (2018). We visualized the samples on a 2% agarose gel to determine the success of each library.

2.3 | Sequence capture methods and data processing

A commercially synthesized capture array (MYbaits, Arbor Biosciences) comprised of high fidelity RNA baits was designed to target 3,446 ddRAD-seq autosomal and Z sex chromosome loci previously isolated within a variety of mallard-like ducks, including black ducks and mallards (Lavretsky et al., 2015; Lavretsky, Janzen, et al., 2019; Peters et al., 2016). Given that ddRAD-seq methods provide a pseudorandom sample of loci across genomes, using consensus sequences from several *Anas* species for bait development provided a pseudorandom sampling of loci in our sequence capture data set. We had no a priori knowledge regarding polymorphism of the loci captured in black ducks or mallards, but expected that a subset of markers would be polymorphic. The chromosome position of each sequence capture was known because the ddRAD-seq loci used to build the capture array were previously assigned chromosomal positions by blasting to a chromosome-aligned mallard genome (Hou et al., unpublished; SRA BioProject PRJNA554956 & BioSample SAMN12287667). This made it possible to align and quantify differentiation across chromosomes, and compare results with previously aligned mallard

and black duck genomes (Lavretsky, Janzen, et al., 2019). We also tested for the presence of mitochondrial DNA (mtDNA) by-catch in our data sets. mtDNA is often well represented in historical samples (Picardi & Pesole, 2012) and is captured even though not targeted with nuclear baits (Gasc, Peyretaillade, & Peyret, 2016; Griffin et al., 2014). To further limit contamination, sequence capture was performed independently on modern and historical samples and separated by species.

DNA–RNA hybridization via RNA baits was used to enrich libraries consisting of ≥200 ng of starting genomic DNA for markers of interest following the manufacturer's instructions. The concentration of purified library products was estimated with quantitative PCR using an Illumina library quantification kit (KAPA Biosystems). Two equimolar concentration pooled solutions of contemporary or historical samples were prepared for Illumina sequencing. Multiplexed libraries were 150 base pair paired-end sequenced on an Illumina HiSeq 4000 at University of Oregon's Genomics & Cell Characterization Core Facility. Raw Illumina reads have been deposited in NCBI's Sequence Read Archive (SRA; http://www.ncbi.nlm.nih.gov/sra; SRA data PRJNA591912).

All libraries were dual-barcoded with unique barcodes ligated to both ends of each sequence and demultiplexed based on perfect barcode matches only, limiting any potential for barcode sequence jumping. Across samples, Geneious version 10 (http://www.geneious.com; Kearse et al., 2012) was used to merge pair-end files, filter for quality (PHRED score ≥30) per base and total sequence (average PHRED score ≥30), and remove PCR duplicates. Samples were then checked for per base sequence quality, per sequence GC content, sequence duplication levels, over-represented sequences and adapter content in FastQC version 0.11.5 (Andrews, 2010). Trimmomatic (Bolger, Lohse, & Usadel, 2014) was used to trim or discard poor quality sequences. The Burrows Wheeler Aligner v. 07.15 (bwa; Li & Durbin, 2009) program was used to index and align per sample sequences to our reference—comprised of initial ddRAD sequences used to build the capture array and a consensus mtDNA sequence of mallards and black ducks (Lavretsky, Hernández Baños, & Peters, 2014; Lavretsky, McCracken, et al., 2014)—using optimized parameters as described in Schubert et al (2012; -T 25 -I 1024). Due to known postmortem degradation effects on DNA, we used MapDamage v 2.0 (Jónsson, Ginolhac, Schubert, Johnson, & Orlando, 2013) to rescale the base scores of historical samples based on patterns of DNA fragmentation and cytosine deamination. Rescaled sequences were used for samples that were flagged by MapDamage as having high levels of damage. Samples were then sorted and indexed in Samtools v. 1.6 (Li et al., 2009) and combined using the "mpileup" function with the following parameters "-c –A -Q 30 -q 30," which set a base pair and overall sequence PHRED score of ≥30 to ensure that high-quality sequences only are retained. All steps through mpileup were automated using a custom in-house script (Python scripts available at https://github.com/jonmohl/PopGen). Next, VCF files for each marker, as well as concatenated autosomal and Z sex chromosome markers, were converted to FASTA file formats using the program PGDSPIDER v. 2.1.1.2 (Lischer & Excoffier, 2011) with base pair retention based on sequence coverage of a minimum allele depth of 5× (10× per genotype) and quality PHRED scores of ≥30. FASTA files were further filtered for samples having <50% sequence coverage and base positions having <80% of sample representation (Python scripts available at https://github.com/jonmohl/PopGen).

2.4 | Population structure and molecular diversity of domestic versus wild mallards

We visualized population structure using biallelic SNPs without a priori assignment of individuals to populations or species. In-house scripts were used to extract biallelic SNPs from a FASTA file with all autosomal loci concatenated using a minimum allele frequency (MAF) of 0.02, which retained all biallelic SNPs except for singletons (i.e., SNPs observed in only one individual). Maximum-likelihood estimates of population assignments for each individual were obtained with ADMIXTURE v.1.3 (Alexander & Lange, 2011; Alexander, Novembre, & Lange, 2009). Autosomal biallelic SNPs were formatted for analyses using plink v. 0.67 (Purcell et al., 2007), following steps outlined by Alexander, Novembre, and Lange (2012). We ran 100 iterations at each population of K (from K of 1–5). Each analysis used a block relaxation algorithm for point estimation and terminated once the change (i.e., delta) in the log likelihood of the point estimations increased by <0.0001. The optimum K was based on the lowest average of CV errors across the 100 analyses per K. We used the r program PopHelper (Francis, 2016) to convert ADMIXTURE outputs into CLUMPP input files at each K. We determined the robustness of individual assignments to populations at each K employing the Greedy algorithm and 1,000 random permutations in the program clumpp v.1.1 (Jakobsson & Rosenberg, 2007). In addition to ADMIXTURE, we tested for relationships and admixture among samples with the program fineRADstructure (Malinsky, Trucchi, Lawson, & Falush, 2018) that includes RADpainter v 0.1 and finestructure (Lawson, Hellenthal, Myers, & Falush, 2012). fineRADstructure derives a coancestry matrix by calculating the number of SNPs between samples, comparing these to all individuals, or potential "donors," and deriving the most recent coalescent events among samples. We employed a burn-in period of 100,000, followed by 100,000 Markov chain Monte Carlo iterations; tree building used default parameters. We visualized the results using the r scripts fineradstructureplot.r and finestructurelibrary.r (available at http://cichlid.gurdon.cam.ac.uk/fineRADstructure.html). We also visualized population structure using principal components analysis (PCA) in the r package Adegenet (Jombart, 2008), using the "dudi.pca" function on the same set of biallelic SNPs.

Next, composite pairwise estimates of relative divergence (Φ_{ST}) and nucleotide diversity (π) for ddRAD-seq loci were calculated across the domestic types and geographical locations for wild mallards in the r package PopGenome (Pfeifer, Wittelsbürger,

Ramos-Onsins, & Lercher, 2014) using concatenated data sets, and with indel positions treated as missing.

2.5 | Sequence capture-based genetic diversity and population structure of contemporary versus historical samples

For historical and contemporary data sets, we calculated pairwise locus-by-locus and composite (concatenated data sets for autosomal and Z sex chromosome-linked markers) estimates of relative divergence (Φ_{ST}), as well as nucleotide diversity (π) across sequence-captured markers in the R package PopGenome (Pfeifer et al., 2014). Inter- and intraspecific Manhattan plots comprising Φ_{ST} estimates across chromosomes permitted the visualization and comparison of previously described putative outlier regions (Lavretsky, Janzen, et al., 2019) between historical and contemporary populations.

For all analyses of population structure, we used all biallelic SNPs with singletons excluded by filtering out SNPs with minor-allele frequencies ≤0.02. First, we tested for and visualized population structure using PCA with the "dudi.pca" function in the R package Adegenet (Jombart, 2008), without a priori information on population or species identity. We further visualized population structure with fineRADstructure (Malinsky et al., 2018), following the steps described above. Because Z-linked loci have potential issues of heterogamety (i.e., males are ZZ and females are ZW) and sexes are unknown for 11 samples (Table S1), both PCA and fineRADstructure analyses were restricted to autosomal markers.

To gauge the success of mtDNA by-catch and ensure no nuclear copy was captured, we compared all contemporary individuals to their respective published sequences based on Sanger sequencing (Lavretsky, Hernández Baños, et al., 2014; Lavretsky, McCracken, et al., 2014). We then calculated pairwise Φ_{ST} and diversity estimates in the R package PopGenome (Pfeifer et al., 2014) and created a haplotype network using the median-joining algorithm as inferred in the program Network v. 4.5.1.0 (Bandelt, Forster, & Röhl, 1999). We included mtDNA sequences for the 63 domesticated mallard samples in the haplotype network. The New World mallard clade is characterized by two divergent mtDNA haplogroups: Old World (OW) A and New World (NW) B (Ankney, Dennis, Wishard, & Seeb, 1986; Avise et al., 1990; Lavretsky, Hernández Baños, et al., 2014). Given that domesticated lineages of the mallard are thought to be derived from Eurasian mallards and that the OW mtDNA A haplogroup originated in Eurasia (Avise et al., 1990; Johnson & Sorenson, 1999; Kulikova et al., 2005; Lavretsky, McCracken, et al., 2014), we expected all domesticated mallard samples to carry OW mtDNA A haplotypes.

2.6 | Testing among evolutionary scenarios between mallards and black ducks

Having the same sequence capture data for historical and contemporary mallards and black ducks allowed us to formally test each data set against evolutionary scenarios using the program ∂a∂i (Gutenkunst, Hernandez, Williamson, & Bustamante, 2009, 2010). ∂a∂i implements an efficient diffusion-based approach to test the empirical data against specified evolutionary models (e.g., isolation-with-migration) with the best-fit model determined using a log likelihood-based multinomial approach. Using ∂a∂i, a frequency spectrum was derived across all sequence-captured autosomal sites, with only minor alleles considered, and variants observed in all samples ignored ("masked"). Data sets were projected down to a total of 14 and 20 alleles per species for the historical and contemporary data sets, respectively. We tested the empirical data against three evolutionary models: isolation-with-migration, split-migration and split-no migration (Gutenkunst et al., 2009, 2010). The best-fit model was based on the log likelihood using a multinomial approach, where model parameters were optimized prior to calculating the likelihood, which is the product of Poisson likelihoods for each parameter given an expected value from the model frequency spectrum. The optimum model was reanalysed 20 times to obtain an average and ranges across runs. Demographic parameters, including population sizes ($n_i = [(N_i/N_{ref}) * N_{Anc}]$), migration rates ($M_{i \leftarrow j} = 2N_{Anc} m_{i \leftarrow j}$) and divergence times ($t = T/2N_{ref}$; N_{ref} = reference effective population size; N_{Anc} = Ancestral effective populations size; t = time in generations), are estimated by the different evolutionary models.

To convert the parameter estimates from ∂a∂i to biologically informative values, we estimated generation time (G) and mutation rates (μ, per locus). Generation time (G) was calculated as $G = \alpha + (s/(1 - s))$, where α is the age of maturity and s is the expected adult survival rate (Sæther et al., 2005). The age of maturity for mallard-like ducks generally is one year (i.e., $\alpha = 1$), and the average adult survival rate is 0.54 (range: 0.34–0.74) and 0.54 (range: 0.4–0.70) for mallards and black ducks, respectively (Nichols, Obrecht, & Hines, 1987). Using an overall survival rate average of 0.54 for the two species, we estimated the generation time to be 4.0 years. Next, to attain a mutation rate for nuclear genes, we multiplied a 1.2×10^{-9} substitutions/site/year mutation rate—previously calculated for nuclear genes in other ducks (Peters, Zhuravlev, Fefelov, Humphries, & Omland, 2008)—by generation time to attain a 4.83×10^{-9} substitutions/site/generation (s/s/g) mutation rate. A final mutation rate was calculated as the product of the above mutation rate and the total number of base pairs.

3 | RESULTS

3.1 | Data filtering

After quality filtering, we recovered 2,274 ddRAD-seq autosomal loci (290,622 base pairs; 6,964 biallelic single nucleotide polymorphisms [SNPs]) that met our coverage and missing data criteria.

For our sequence capture data set, four historical black ducks, eight historical mallards, one historical hybrid and one contemporary mallard did not meet our criteria. Thus, a total of 69 samples (21 historical black ducks from 1842 to 1910; seven historical

Mallards from 1872 to 1915; and two historical putative mallard × black duck hybrids from 1840 to 1910; 20 contemporary black ducks; and 19 contemporary mallards) were retained for all downstream analyses (Figure 1; Table S1). For the 69 samples, a total of 2,202 markers (140,477 base pairs (bp) across the Z sex (99 markers; 6,122 bp) and 28 autosomal (2,103 markers; 134,355 bp) chromosomes (Figure S1) passed quality filtering. Given an original bait-capture array of 3,446 ddRAD-seq loci, we were able to achieve a recovery rate of 64%, which is similar to reported recovery rates using RAD-based bait arrays (Souza et al., 2017). Additionally, we succeeded with off-target recovery of 641 base pairs of the mtDNA control region for 69 samples, with >90% of sites represented by >10× coverage.

3.2 | ddRAD-seq results between domestic game-farm and wild mallards

Pairwise comparisons of composite Φ_{ST} estimates for ddRAD-seq loci across the three domestic mallard types found feral mallards from Texas to be most different (composite Φ_{ST} estimate ~ 0.27), and game-farm mallards from Kentucky and New Jersey to be most similar (composite Φ_{ST} estimates = 0.022; Figure S2a). This suggests that the Kentucky and New Jersey samples are likely of the same domestic lineage. Comparing these domestic lineages to wild mallards, we found that game-farm mallards have lower levels of relative differentiation (range of composite Φ_{ST} estimates = 0.053–0.11) than Texas feral mallards (range of composite Φ_{ST} estimates = 0.22–0.24). With wild mallards grouped by geographical region, we recovered twofold higher estimates of relative differentiation between game-farm mallards and both Western and Mississippi Flyway mallards (Φ_{ST} estimates ranging from 0.083–0.11) than game-farm and Atlantic Flyway mallards (Φ_{ST} estimates ranging from 0.053–0.065) (Figure S2a). We found overall lower levels of nucleotide diversity in domestic versus wild mallards (Figure S2a) although wild mallards had only slightly higher nucleotide diversity than game-farm mallards from Kentucky or New Jersey.

Similarity between wild mallards and game-farm mallards from Kentucky and New Jersey was further supported across ADMIXTURE (Figure 2a), PCA (Figure 2c) and FINERADSTRUCTURE analyses (Figure S3). First, at the optimum population K of 3 (Figure 2b), ADMIXTURE grouped all feral mallards from Texas into a single population that differed from game-farm mallards from Kentucky and New Jersey (the game-farm mallard cluster), which comprised a single population. We recovered the previously described two-population structure in wild mallards (i.e., Western vs. non-Western; Lavretsky, Janzen, et al., 2019) and determined that the non-Western mallard nuclear signature in wild mallards is due to the genetic contribution from game-farm mallards (Figure 2a). Geographically separating wild mallards, we report that none of the Western, 24% (9 of 38) of Mississippi Flyway, and 71% (22 of 31) of Atlantic Flyway wild mallards sampled here

had substantial assignment probabilities (>10%) to the game-farm mallard cluster. PCA and ADMIXTURE results corresponded well with PC1 (explaining 6.23% of the variation), clearly differentiating Texas's feral mallards from all other samples, while PC2 (explaining 5.24% of the variation) differentiated wild mallards from game-farm mallards (Figure 2c). In general, we find a strong correlation between PC2 scores and ADMIXTURE-based individual assignment probabilities to game-farm mallards (Figure 2d). Those samples with substantial assignment to game-farm mallards in ADMIXTURE analyses made up the distribution of samples connecting the primary cluster of wild mallards and game-farm mallards in our PCA (Figure 2c), supporting the interpretation of these as intermediates (i.e., hybrids) between the two. PC3 (explaining 0.98% of variation) showed some support for separating Kentucky and New Jersey game-farm mallards.

Finally, the FINERADSTRUCTURE dendrogram recovered two major lineages that corresponded well with domesticated and wild versions of the mallard (Figure S3), with the exception of seven wild mallards from the Atlantic Flyway that clustered in the domestic lineage. Investigating coancestry assignments recovered the same three primary clusters of samples as in ADMIXTURE results (Figure 2a), which include feral mallards from Texas, game-farm mallards from Kentucky and New Jersey, and wild mallards. Feral mallards from Texas showed the highest estimates of coancestry, which likely reflects a combination of significant genetic drift due to the domestication process, and sibling relationships—all of the feral mallards were sampled from the same wetland and were often seen grouped together. For wild mallards, we recovered low levels of coancestry with feral mallards from Texas, but found several individuals from the Mississippi Flyway and Atlantic Flyway that had relatively higher coancestry with game-farm mallards. All samples with wild mallard plus game-farm mallard coancestry, and the seven that clustered with game-farm samples, had mixed assignment probabilities in ADMIXTURE analyses and were found as intermediates in our PCA (Figure 2), consistent with these being feral × wild mallard hybrids.

3.3 | Interspecific sequence capture results

Based on our sequence capture data sets, composite Φ_{ST} estimates between contemporary (composite $\Phi_{ST\ AUTOSOMAL}$ = 0.012; composite $\Phi_{ST\ Z\ SEX\ CHROMOSOME}$ = 0.034; $\Phi_{ST\ mtDNA}$ = 0.39) and historical (composite $\Phi_{ST\ AUTOSOMAL}$ = 0.000065; composite $\Phi_{ST\ Z\ SEX\ CHROMOSOME}$ = 0.018; $\Phi_{ST\ mtDNA}$ = 0.0020) mallard and black duck data sets showed a 182-, 1.8- and 199-fold increase in overall differentiation at autosomal, Z-linked, and mtDNA markers, respectively (Figure S2b). Φ_{ST} estimates from sequence capture techniques were identical to those calculated with ddRAD-seq loci for contemporary mallards and black ducks (Lavretsky, Janzen, et al., 2019). Plotting Φ_{ST} estimates per sequence capture locus and across chromosomes between historical and contemporary black ducks and mallards demarcated several regions of high differentiation, all of which were
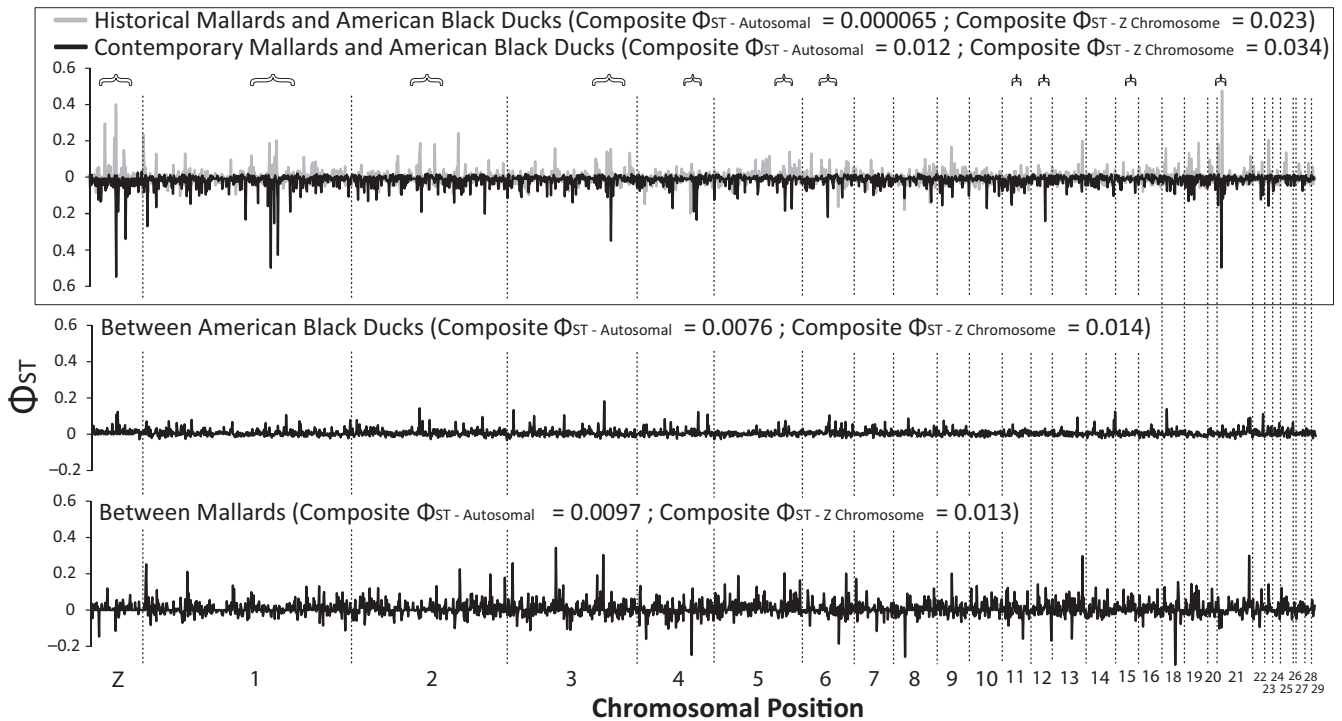
FIGURE 3 Chromosomally aligned Φ_{ST} estimates of sequence-captured 2,103 autosomal and 99 Z-linked loci for pairwise inter- and intraspecific comparisons between historical and contemporary mallards and American black ducks. For easier comparison, interspecific comparisons are broken up by historical (above the zero line) and contemporary (below the zero line) data sets. Note that brackets denote chromosome regions previously found to harbour genes that were under divergent selection within American black ducks and/or mallards by Lavretsky, DaCosta, et al. (2019), and Lavretsky, Janzen, et al. (2019). Composite Φ_{ST} estimates for either autosomal or Z sex chromosome-linked loci are also provided for each respective comparison

within regions previously identified as putatively under divergent selection between the two species (top panel in Figure 3; Lavretsky, DaCosta, Sorenson, McCracken, & Peters, 2019; Lavretsky, Janzen, et al., 2019).

3.4 | Intraspecific sequence capture results

With the exception of mtDNA, estimated intraspecific composite Φ_{ST} estimates were low between historical and contemporary mallards (composite $\Phi_{ST\ AUTOSOMAL}$ = 0.0097; composite $\Phi_{ST\ Z\ SEX\ CHROMOSOME}$ = 0.013; $\Phi_{ST\ mtDNA}$ = 0.24) and black ducks (composite $\Phi_{ST\ AUTOSOMAL}$ = 0.0076; composite $\Phi_{ST\ Z\ SEX\ CHROMOSOME}$ = 0.014; $\Phi_{ST\ mtDNA}$ = 0.10) (Figure S2b). Intraspecific Manhattan Φ_{ST} plots of autosomal and Z-linked markers demarcated no marker with a Φ_{ST} > 0.20 in black ducks (highest Φ_{ST} = 0.18; middle plot in Figure 3), but twelve loci in mallards (highest Φ_{ST} = 0.34; bottom plot in Figure 3). Finally, historical mallards ($\pi_{AUTOSOMAL}$ = 0.018; $\pi_{Z\ SEX\ CHROMOSOME}$ = 0.0085; π_{mtDNA} = 0.0068) and black ducks ($\pi_{AUTOSOMAL}$ = 0.016; $\pi_{Z\ SEX\ CHROMOSOME}$ = 0.011; π_{mtDNA} = 0.0096) had approximately twofold higher calculated nucleotide diversity across markers than contemporary mallards ($\pi_{AUTOSOMAL}$ = 0.0074; $\pi_{Z\ SEX\ CHROMOSOME}$ = 0.0048; π_{mtDNA} = 0.011) and black ducks ($\pi_{AUTOSOMAL}$ = 0.0078; $\pi_{Z\ SEX\ CHROMOSOME}$ = 0.0049) (Figure S2b). Once again, π values calculated from sequence capture techniques were similar to those calculated

with ddRAD-seq loci for contemporary mallards and black ducks (Lavretsky, DaCosta, et al., 2019; Lavretsky, Janzen, et al., 2019; also see Figure 3a). Similar estimates were achieved across estimated parameters whether contemporary sets were analysed together or separated into the three regional bins (Figure S2b). In addition to having lower nucleotide diversity and moderately higher relative differentiation as compared to historical samples, contemporary samples showed substantial intraspecific differentiation in mtDNA, particularly when comparing Atlantic Flyway mallards to all other groups (Figure S2b).

3.5 | Population structure from sequence capture of mitochondrial and nuclear DNA

Mitochondrial haplotype tree reconstruction (Figure 4a) identified the well-known Old World (OW) A and New World (NW) B haplogroups (Avise et al., 1990; Johnson & Sorenson, 1999; Kulikova et al., 2005; Kulikova, Zhuravlev, & McCracken, 2004; Lavretsky, McCracken, et al., 2014; McCracken, Johnson, & Sheldon, 2001). All historical mallards and 91% (19/21) of historical black ducks grouped within the NW B mtDNA haplogroup (Figure 4b). All contemporary black ducks grouped in the NW B mtDNA haplogroup, but nearly 50% of contemporary mallards grouped in the OW A mtDNA haplogroup. More specifically, 57% of western region and 86% of Atlantic Flyway
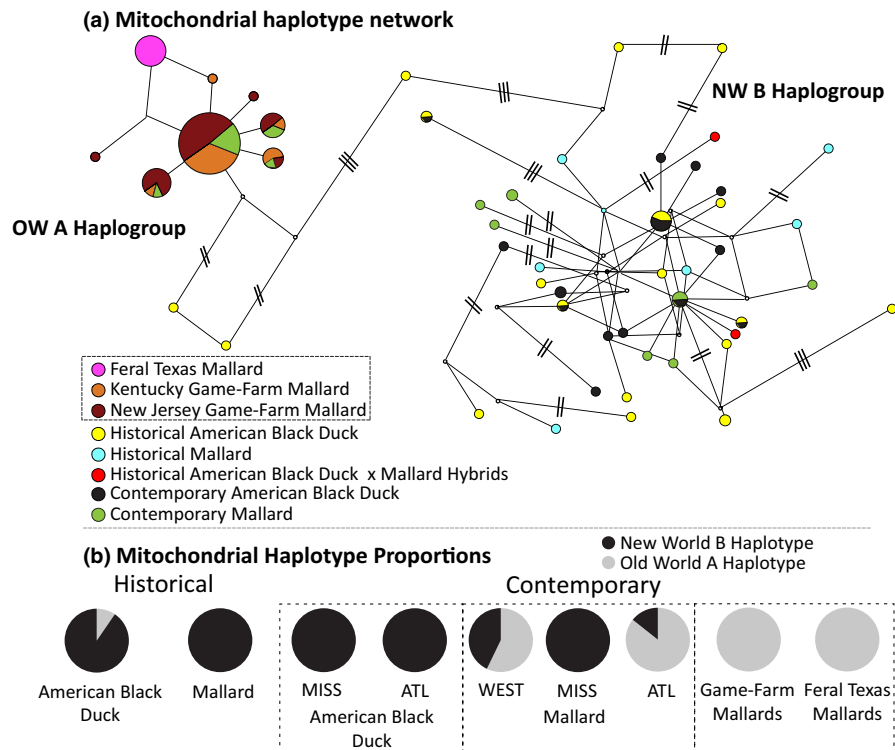
(a) Mitochondrial haplotype network

FIGURE 4 (a) A haplotype network based on sequence capture of 641 base pairs of the mitochondrial control region for historical and contemporary mallards and American black ducks, as well as two historical putative American black duck × mallard hybrids. Aligned domestic mallard (i.e., feral park and game-farm mallards) Sanger sequenced mtDNA are also included. Size of circles is proportional to number of samples carrying that haplotype, and branch lengths indicate the number of mutations separating haplotypes. Note that the number of mutations for those branches representing >1 mutation is denoted with dashes. (b) Proportion of domestic mallards, historical mallards, historical American black ducks, contemporary mallards, contemporary American black ducks—within contemporary samples, we compared mallards and black ducks that were recovered in the Mississippi (MISS) or Atlantic (ATL) Flyways, as well as mallards from west of the Mississippi River (WEST)—with either OW A or NW B Haplotypes. See Figure 1 for specifics on geographical designations

mallards had OW A mtDNA haplotypes, whereas all Mississippi Flyway mallards had NW B mtDNA haplotypes. Lavretsky, Janzen, et al. (2019) reported OW A haplotypes in 38%, 40% and 74% of western region, Mississippi flyway and Atlantic flyway mallards, respectively. Thus, the proportions reported here based on fewer samples should be taken with caution, especially the absence of OW A haplotypes in Mississippi flyway. Finally, all domestic mallards carried OW A haplotypes, including a single haplotype found in all feral mallards from Texas, and six haplotypes found across game-farm mallards from Kentucky and New Jersey (Figure 4b). Importantly, four major OW A haplotypes included game-farm mallards from Kentucky and New Jersey, as well as all of the contemporary mallards that carried OW A haplotypes (Figure 4a).

A total of 1,920 biallelic autosomal SNPs were used in PCA and FINERADSTRUCTURE (Figure 5) analyses. The first two principal components (PC1 [7.1%] and PC2 [3.9%]) largely separated contemporary Atlantic Flyway mallards from all the other samples, with subtle differences in the clustering of remaining mallard and black duck samples (Figure 5b). FINERADSTRUCTURE analyses confirmed that historical and contemporary black ducks share a relatively high degree of coancestry within the plot. Specifically, the individual coancestry matrix (above diagonal in Figure 5a) recovered equal coancestry across all

but contemporary mallards from the Atlantic Flyway. All Atlantic Flyway mallards showed high individual coancestry among each other. The inability to distinguish between individuals of mallards and black ducks demonstrates the significant proportion of the genome that the two species have always shared. It required sampling many individuals to achieve significant statistical differentiation between them (Lavretsky, Janzen, et al., 2019). Nevertheless, some resolution was attained in the averaged coancestry matrix (below diagonal in Figure 5a), as displayed in the dendrogram, in which three major lineages were recovered: (a) historical and contemporary black ducks, (b) historical and contemporary Western and Mississippi Flyway mallards and (c) Atlantic Flyway mallards. Finally, whereas the PCA showed one historical mallard × black duck hybrid grouping with mallards and the other with black ducks, FINERADSTRUCTURE identified those samples as having highest averaged coancestry with black ducks (Figure 5a). Note that the two historical mallards with substantially higher estimates of individual and averaged coancestry are likely siblings, as they were collected together at the same location (North Dakota) and date in the summer (9 July 1915). Conversely, the high coancestry estimates across Atlantic Flyway mallards are unlikely to be explained by sibling status, as those birds were sampled on different dates in the fall and from six different states (Table S1).
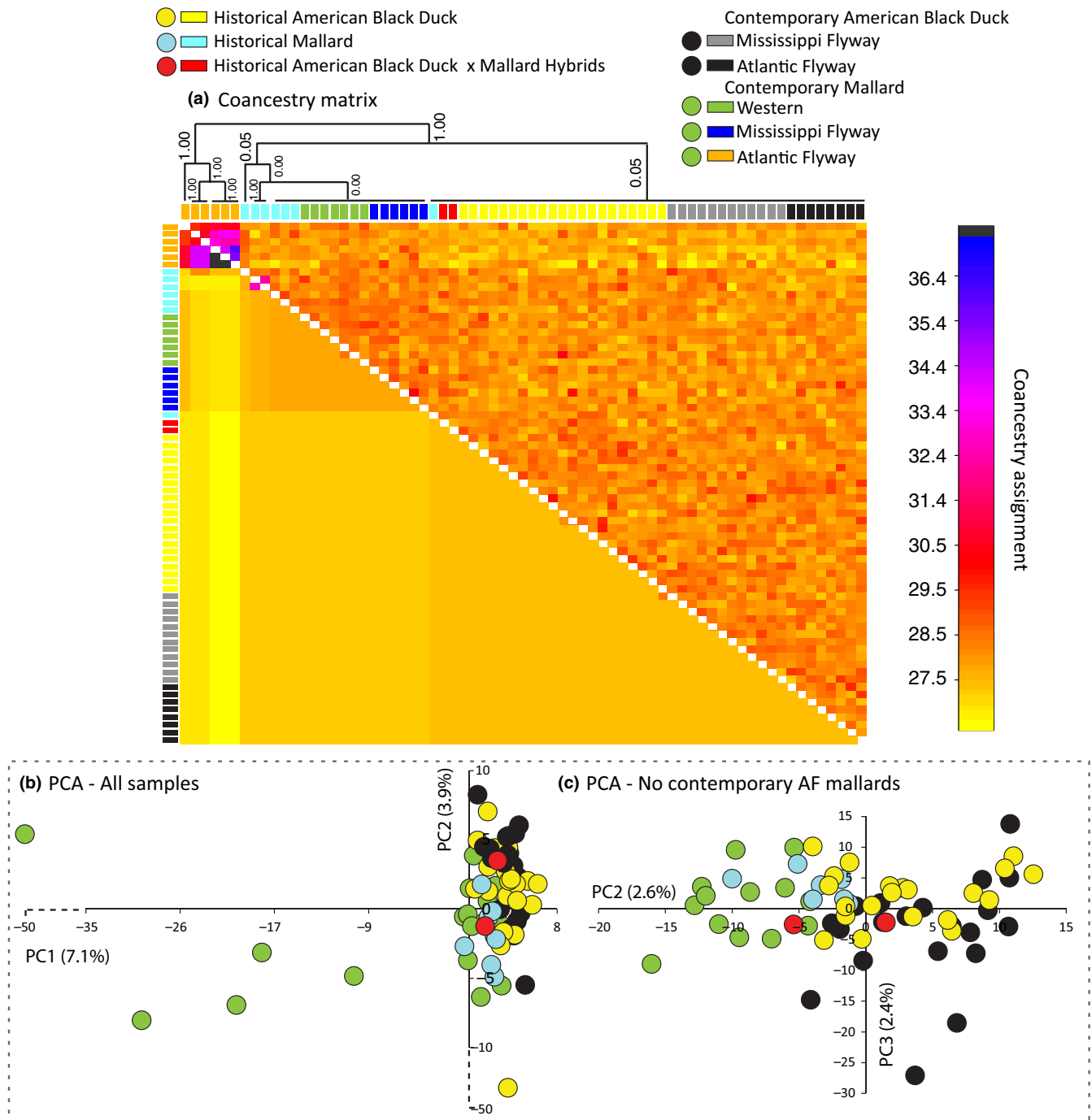
FIGURE 5 (a) fineRADstructure individual (above diagonal) and average (below diagonal) coancestry coefficient matrix estimated using 1,920 sequence capture biallelic autosomal SNPs. The level of recent coancestry is colour-coded from low (yellow) to high (blue), and posterior support across dendrogram branches is provided. Additionally, PCA between contemporary and historical mallards, black ducks and putative hybrids based on the same 1,920 biallelic sequence capture SNP data set, analysed either with (b) all samples or (c) no contemporary mallards from the Atlantic Flyway. Sample assignments are denoted in squares and circles for fineRADstructure and PCA analyses, respectively. Samples were only geographically distinguished in the fineRADstructure analysis; however, note that all five samples at the tail end of the all sample PCA (b) are from the Atlantic flyway (also see Figure 2c)

3.6 | Testing among evolutionary scenarios between mallards and black ducks

Based on log likelihoods, the best-fit model in ∂a∂i was over-whelmingly "Split-Migration" for both historical and contemporary

sequence capture data sets (Figure 6a; Figure S4). Based on the optimum model, we estimated and used an averaged mutation rate of $(4.8 \times 10^{-9} \text{ s s}^{-1} \text{ g}^{-1} \times 134{,}355 \text{ base pairs} =) 6.49 \times 10^{-4} \text{ s s}^{-1} \text{ g}^{-1}$ to transform ∂a∂i results. The average and ranges for estimated effective population size, time since divergence and migration rate were statistically the same between the two data sets (Figure 6b). Although
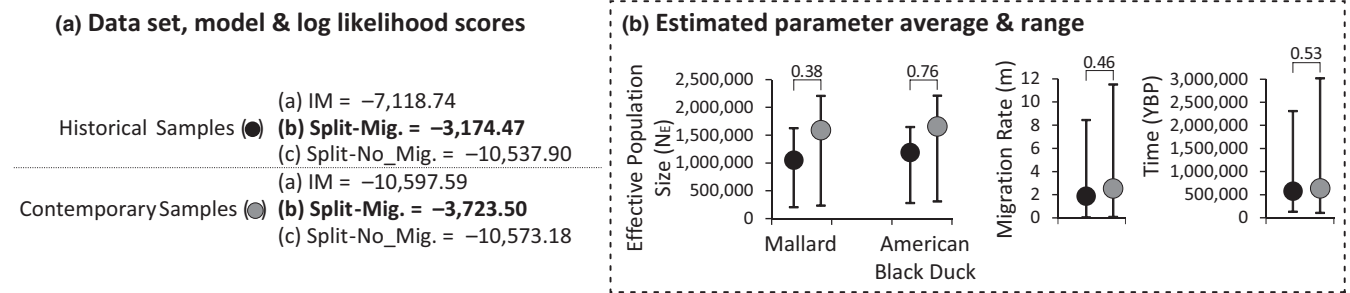
(a) Data set, model & log likelihood scores

Historical Samples (●)
(a) IM = −7,118.74
(b) Split-Mig. = −3,174.47
(c) Split-No_Mig. = −10,537.90

Contemporary Samples (◐)
(a) IM = −10,597.59
(b) Split-Mig. = −3,723.50
(c) Split-No_Mig. = −10,573.18

(b) Estimated parameter average & range

FIGURE 6 Log likelihoods for the comparison of empirical sequence data captured from historical or contemporary mallards and black ducks against the isolation-with-migration (IM), secondary contact (Split-Mig.) and divergence with no gene flow (Split-No_Mig) evolutionary models in the $\partial a \partial i$ program. Based on the optimum model of secondary contact, we provide the average and range across 20 replicates for each estimated parameter, including effective population size (N_E), time since divergence (years before present) and migration rate (m). Note Student's t test p-values calculated between historical and contemporary data sets are nonsignificant across estimated parameters

there was a slight increase in effective population size estimated for black ducks (historical average = 1.051 million, range = 0.21–1.63 million; contemporary average = 1.57 million, range = 0.43–2.00 million) and mallards (historical average = 1.19 million, range = 0.28–1.64 million; contemporary average = 1.37 million, range = 0.27–2.026 million) between data sets, estimates across the respective 20 replicates were not statistically different. An estimated time since divergence between mallards and black ducks was approximately 600,000 years before present (ybp) whether derived from historical (avg. ybp = 574,829, range = 134,078–2,310,814 ybp) or contemporary (avg. ybp = 632,305, range = 166,166–2,076,004 ybp) data sets. Our estimated divergence times between black ducks and wild mallards are on average greater than previously suggested (180,000 ybp (95% CI = 100,000–400,000 ybp); Lavretsky, Hernández Baños, et al., 2014), which may be due to the much smaller set of used loci (17 nuclear introns) and samples (25 per species) analysed previously. Finally, we recovered a relatively low migration rate of approximately two to three migrants per generation whether analysing historical (1.87 migrants per generation, range = 0.04–8.44 migrants per generation) or contemporary (2.82 migrants per generation, range = 0.21–9.72 migrants per generation) data sets.

4 | DISCUSSION

4.1 | The establishment of a feral North American mallard

Historically, the Atlantic Flyway has seen a high rate of game-farm mallard releases (see Introduction), sparking speculation about gene flow with wild populations. Here, we provide strong genetic evidence across population structure analyses (Figure 2; Figure S3) that the non-Western mallard genetic signature found in North American mallards (Lavretsky, Janzen, et al., 2019) is the result of a century of extensive gene flow from game-farm mallards into wild mallard populations. Given that sampled game-farm mallards from Kentucky and New Jersey were assigned to a single genetic cluster in ADMIXTURE analyses (Figure 2a), cluster together in our PCA (Figure 2c), show

high coancestry in fineRADstructure (Figure S3) and share all major mtDNA haplotypes (Figure 4a), we conclude that these likely represent a single domestic lineage of game-farm mallard. Moreover, 100% of samples from Kentucky and New Jersey carry OW A haplotypes (Figure 4b), strongly suggesting that the domestic stock of mallards used by game-farm breeders in North America are originally of Eurasian descent.

There are two well-developed hypotheses to explain the presence of both OW A and NW B haplotypes in New World mallard-like ducks. First, the "mallard invasion hypothesis" posits the speciation of a monochromatic ancestor and the evolution of the NW B haplotype in North America, with subsequent bidirectional hybridization with Eurasian mallards that carry the OW A haplotype only (Avise et al., 1990; Johnson & Sorenson, 1999; Kulikova et al., 2005; McCracken et al., 2001). The "mallard ancestor hypothesis" posits the colonization of North America by Eurasian mallards, followed by vicariant events separating mallard populations, losses of dichromatism in four separate groups and the evolution of the NW B haplotype in the New World (Heusmann, 1974; Omland, 1997). Although we cannot definitively reject either hypothesis, we present compelling evidence that the presence of the predominant OW A mtDNA haplotypes in North America is instead largely the result of a century of game-farm stocking practices. First, whereas all feral mallards from Texas were fixed for a single OW A mtDNA haplotype, the remaining six OW A haplotypes were recovered among game-farm individuals from Kentucky and New Jersey. Most importantly, all OW A haplotypes found in our contemporary mallard samples were not unique, but were shared with game-farm mallards (Figure 4). No historical mallard or black duck sample was found within the set of seven OW A haplotypes recovered in domestic mallard samples. We note that the OW A haplotypes carried by two historical black ducks were still distinct from the primary mtDNA haplogroup currently carried by domestic and wild North American mallards (Figure 4a). Thus, we conclude that game-farm OW A haplotype mitochondrial introgression in contemporary populations can explain much of the data that has long inspired debate about the origins of OW A and NW B mtDNA haplotypes in North America.

Our results suggest that the continuous input of game-farm mallards has resulted in the genetic swamping of eastern North American mallards and the establishment of a large feral and feral × wild mallard population. Indeed, ADMIXTURE and FIN-ERADSTRUCTURE results from ddRAD-seq identified several contemporary samples from eastern North America that represent game-farm mallards outside of the shooting preserve from which they were released (Figure 2a; Figure S3). Moreover, the six contemporary mallards from the Atlantic Flyway that had high coancestry in FINERADSTRUCTURE results, and clustered away from all other contemporary and historical samples in our PCA (Figure 5), were recovered as feral × wild mallards in the larger ddRAD-seq data set (Figure 2c). Given that all of our wild samples were from birds shot by hunters in natural settings, not on hunting preserves, these findings demonstrate that game-farm mallards are fully capable of dispersing, surviving , and breeding outside the area where they were released. We find little evidence that park ducks represented here by feral Khaki Campbell mallards from Texas are as problematic to the genetic integrity of wild mallard populations as are released game-farm mallards. Given recent findings that ~90% of the 2010 mallards sampled in the Atlantic Flyway (Lavretsky, Janzen, et al., 2019) had substantial (>10% in assignment probabilities) genetic input from game-farm mallards (also see Figure 2), it is likely that the eastern mallard is on the cusp of becoming, or currently already represents, a true hybrid swarm. Future work will benefit from continued sampling, including determining whether the proportions of feral and feral x wild mallard hybrids are increasing, decreasing or are stable. Moreover, molecular assessment of breeding or "local" mallards will be important to establish the breeding distribution of these feral and feral x wild mallard hybrids. We hypothesize that feral and feral x wild mallard hybrids likely breed closer to human-disturbed habitats as compared to wild mallards.

The causes of population decline in black ducks and the eastern North American mallard population (the latter observable only over the past 20 years) are debated (Heusmann, 2017; USFWS, 2013). We predict that the significant input of maladaptive genes from game-farm mallards over the last century is one important cause. In fact, simulating outcomes of breeding between domestic and wild counterparts predicts an additive consequence of gene flow that requires time and continuous genetic input before negative effects are observable (Tufto, 2017). Not only do game-farm mallards differ ecologically (e.g., early nest initiation, lower survival) and physiologically (e.g., smaller body size, bill shape and size, structure of beak lamellae) from wild mallards (Söderquist, Gunnarsson, & Elmberg, 2013; Söderquist, Norrström, Elmberg, Guillemain, & Gunnarsson, 2014), but we show they are also ten-times genetically more distinct from wild mallards than black ducks are (Figure S2a). To what extent current rates of gene flow and the movement of putatively maladaptive traits from domestic populations affect adaptability and future conservation of wild populations remains unknown, and requires careful monitoring.

4.2 | Historical DNA reveals little genetic change between black ducks and wild mallards

We revisit the hypothesis that extensive gene flow between black ducks and mallards has resulted in—or is progressing towards—the genetic extinction of the black duck. Our results conflict with those of Mank et al. (2004), as we report an overall 200-fold increase in relative divergence estimates across mtDNA and autosomal loci, and a twofold increase in Z sex chromosome loci, between historical and contemporary comparisons of the species (Figure S2b). We note that given the evident genomic heterogeneity across our markers (Figure 3), it is unlikely that the few markers used in previous attempts (i.e., 3 microsatellites; Mank et al., 2004) could capture the complete picture. These results demonstrate the value of increased marker coverage, especially when attempting to resolve complex relationships as is often the case for very closely related taxa.

We conclude that mallards and black ducks have always been closely related, as evidenced by the absence of species-specific markers (see top panel in Figure 3), similar diversity estimates (Figure S2b) and high coancestry (Figure 5a). Furthermore, Φ_{ST} estimates across all putative outlier regions previously identified on the Z sex chromosome and autosomal chromosomes 1–5 and 21 (Lavretsky, Janzen, et al., 2019) had nearly identical or slightly higher values in contemporary compared to historical samples (see top panel in Figure 3). Though these findings are consistent with a divergence-with-gene flow evolutionary scenario (Feder et al., 2012), ∂a∂i did not find the isolation-with-migration (IM) model to be the optimum. Instead, support for the same optimum model of secondary contact (i.e., split-migration evolutionary model) was overwhelmingly supported whether analysing historical or contemporary data sets in ∂a∂i, including statistically similar estimates across parameters (Figure 6; Figure S4). Our results support a divergence time of ~600,000 ybp followed by secondary contact with moderate levels of gene flow (2–3 effective migrants per generation). We cannot rule out prolonged period(s) of isolation punctuated by gene flow events between mallards and black ducks across evolutionary time (recurring secondary contact; Figure 1d), but we can reject contemporary gene flow from wild mallards having a significant impact on the black duck's genome (outlined scenarios A & B in the introduction), and a strict evolutionary history of divergence-with-gene flow (scenarios C in the introduction). Thus, despite the high estimates of hybridization between wild mallards and black ducks (i.e., ~25%; Lavretsky, Janzen, et al., 2019) as compared to mallards and mottled ducks (i.e., ~5%–6%; Ford, Selman, & Taylor, 2017; Peters et al., 2016) or Mexican ducks (i.e., ~2%; Lavretsky et al., 2015), together, our results suggest that black ducks and mallards have likely evolved mechanisms to limit actual gene flow (e.g., assortative mating, hybrid breakdown). Finally, given the optimum model of secondary contact (Figure 3; also see in Lavretsky, Janzen, et al., 2019), genetic regions demarcated as under divergent selection are likely the result of postdivergence selection within mallards and/or

black ducks (Cruickshank & Hahn, 2014). These analyses particularly highlight the potential for coupling biological collections with increasingly efficient ancient DNA methods to access the genetics of historical populations and provide unique opportunities to conduct genomic inventories through time (Besnard et al., 2015; Kress, 2014).

4.3 | Impact of differential gene flow on genomic diversity and differentiation estimates

Our study suggests that mallards and black ducks diverged ~600,000 years ago and speciated largely in allopatry punctuated by secondary contact events. This natural process likely proceeded unabated until the early 1920s when the release of game-farm mallards along the eastern seaboard resulted in landscape-level gene flow with wild populations (Figure 7; Lavretsky, Janzen, et al., 2019). A large proportion of black ducks carry OW A haplotypes and game-farm mallard × black duck hybrids have been identified (Lavretsky, Janzen, et al., 2019), yet historical and contemporary black ducks remain genetically similar in nuclear variation (Figure S2b). In black ducks, we observed much less variability in differentiation estimated across the nuclear genome than in historical and contemporary mallard comparisons (middle vs. bottom panel in Figure 3). These results suggest that overall nuclear introgression from game-farm mallards is less extensive and/or that selection is more effective in black ducks than in wild mallards. A higher rate of introgressive hybridization from game-farm mallards, experienced by wild mallards versus black ducks, could explain the nearly

200-fold greater overall differentiation in mallards than in black ducks (Figure 3), as well as the differences in estimated nucleotide diversity (Figure S2b) and effective population sizes (Figure 6b) observed in interspecific historical and contemporary comparisons.

Although effective population size and genetic diversity are typically associated, previous studies have shown that these two diversity estimates can become uncoupled due to differential gene flow and/or demographic processes (e.g., bottlenecking) acting through time (Gasca-Pineda, Cassaigne, Alonso, & Eguiarte, 2013; Lonsinger, Adams, & Waits, 2018; Miller & Waits, 2003). Given that our study deals with gene flow between wild and domestic lineages, we posit that gene flow from an inbred domestic lineage (e.g., game-farm mallards) can move "novel" variation into a population and thus create a signature of increased or similar effective population size as estimated here by $\partial a \partial i$. However, a drop in overall genetic diversity may be expected if those genes being moved have higher rates of homozygosity, as is often the result of artificial selection in domestic lineages. A combination of recent population declines and differential gene flow from an inbred domestic lineage into wild mallards and black ducks may explain why our estimated genetic diversity and effective population size estimates are uncoupled.

Similarly, we hypothesize that differential gene flow of loci across the genome from game-farm mallards into wild mallards may also explain the increased genomic differentiation observed across time between wild mallards and black ducks (Figure 3; Supplementary Materials Figure S2B). Strict assortative mating is unlikely to explain the observed 200-fold change at nuclear and mtDNA considering the recency of extensive contact between wild mallards, black ducks, and
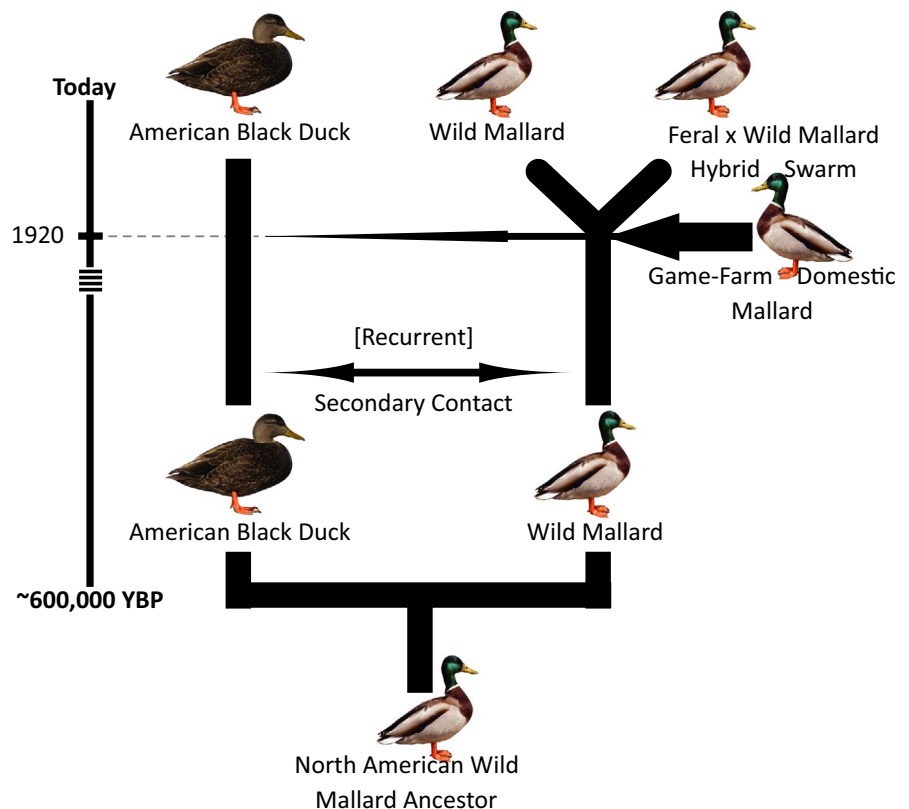

FIGURE 7 Schematic of the putative evolutionary history of American black ducks and mallards, and their interaction with game-farm mallards in North America. Note that Lavretsky, DaCosta, et al. (2019) provided evidence that suggested a wild mallard was likely the ancestor to all monochromatic mallard-like species (e.g., American black ducks) in North America, and we estimated a divergence time between American black ducks and wild mallards of ~600,000 years before present. Widths of arrows represent intensity and directionality of gene flow

game-farm or feral mallards. Several lines of evidence are consistent with the above hypothesis. First, increased mtDNA differentiation is largely explained by the prevalence of OW A haplotypes, which we establish are of game-farm origin (Figure 4). Next, the nuclear landscapes (Figure 3) show higher variation in, and a larger number of highly diverged markers between, historical and contemporary populations of mallards ($N = 12$ with $\Phi_{ST} > 0.20$) than black ducks ($N = 0$ with $\Phi_{ST} > 0.20$). Finally, the same near 200-fold change observed in nuclear and mtDNA is not seen for the Z sex chromosome, or known outlier regions on autosomal chromosomes between wild mallards and black ducks (top panel in Figure 3). Together, these results are consistent with a case of differential introgression, including a scenario of reduced gene flow of domesticated mallard variation at selectively non-neutral genomic locations in wild populations. The latter result not only suggests that postzygotic selection, including Haldane's rule (Haldane, 1922, 1948), may be playing an important role between black ducks and mallards (wild or game-farm) (Kirby, Sargeant, & Shutler, 2004), but also between game-farm and wild mallards.

5 | CONCLUSIONS

We demonstrate that secondary contact has fundamentally changed the genomes of North America's mallard populations. The continued release of captive-bred mallards could not only decrease the natural adaptiveness of North American mallards, and to some degree black ducks, but could replace them through anthropogenic introgressive hybridization; at which time, seeing a pure bird in the wild becomes a pleasant surprise. Given that the release of domestic stock conspecifics into the wild is a practice used worldwide to augment various wildlife populations (Champagnon, Elmberg, Guillemain, Gauthier-Clerc, & Lebreton, 2012; Laikre & Ryman, 1996), monitoring populations to ensure that such interactions do not result in decreased adaptiveness and overall fitness of wild populations is critical (Frankham, 2005; Laikre et al., 2010). The contact between domestic and wild congeners of ducks in eastern North America provides a unique system in which to closely study and monitor the effects of such interactions.

In terms of methodology, we demonstrate the utility of coupling ddRAD markers and sequence capture techniques to obtain genomic markers from century-old duck specimens (Souza et al., 2017), which permitted us to more conclusively examine the potential consequences of secondary contact between closely related species. Furthermore, we report that estimates of overall genetic differentiation, diversity, as well as recovered structure between contemporary mallards and black ducks based on sequence capture techniques were nearly identical to results from previously published ddRAD-seq markers (Lavretsky, Janzen, et al., 2019). ddRAD and related methods (e.g., RAD, GBS, etc.; Andrews, Good, Miller, Luikart, & Hohenlohe, 2016) are inherently biased by the possibility of allelic dropout due to mutations in enzymatic cut-sites (Catchen et al., 2017; Graham et al., 2015; Lowry et al., 2016). These biases are not present in sequence capture data sets as they do not require enzymatic cut-sites to be present or intact to work. Given our highly similar results between the two data sets, and particularly

for species-level comparisons (i.e., no geographical partitioning), we conclude that known biases with restriction enzyme-based techniques (e.g., allelic dropout; Catchen et al., 2017; Graham et al., 2015; Lowry et al., 2016) may have little or no effect on population-level statistics for species with very shallow divergence, such as mallards and black ducks.

AUTHOR CONTRIBUTIONS

P.L., R.C.F. and H.F.J. conceived and designed experiments; P.L. and N.R.M. performed molecular analyses on contemporary samples; P.L. and J.I.B. performed molecular analyses on contemporary samples. P.L. and J.M. performed bioinformatic analyses. P.L., N.R.M., J.M., J.I.B., H.F.J., K.G.M. and R.C.F. wrote the manuscript.

DATA AVAILABILITY STATEMENT

All mitochondrial sequences associated with this study are available on GenBank (Accession nos MN734444–MN734576). All sequence capture sequencing and ddRAD-seq data associated with this study are available from the National Center for Biotechnology Information Sequence Read Archive (Accession nos SAMN13391043–SAMN13391173, BioProject Accession no. PRJNA591912). All other data files (e.g., FASTA files; ADMIXTURE and fineRADstructure input files) are also available on Dryad (Accession https://doi.org/10.5061/dryad.5mkkwh71v).

ORCID

Philip Lavretsky 🆔 https://orcid.org/0000-0002-5904-8821
Helen F. James 🆔 https://orcid.org/0000-0002-2495-6133

REFERENCES

Abbott, R. J., Barton, N. H., & Good, J. M. (2016). Genomics of hybridization and its evolutionary consequences. *Molecular Ecology*, 25, 2325–2332.

Alexander, D. H., & Lange, K. (2011). Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics*, 12, 246. https://doi.org/10.1186/1471-2105-12-246

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, 19, 1655–1664. https://doi.org/10.1101/gr.094052.109

Alexander, D. H., Novembre, J., & Lange, K. (2012). *Admixture 1.22 software manual*.

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and

evolutionary genomics. *Nature Reviews Genetics*, 17, 81–92. https://doi.org/10.1038/nrg.2015.28

Andrews, S. (2010). *FastQC: A quality control tool for high throughput sequence data*.

Ankney, C. D., Dennis, D. G., Wishard, L. N., & Seeb, J. E. (1986). Low genic variation between black ducks and mallards. *The Auk*, 103, 701–709. https://doi.org/10.1093/auk/103.4.701

Avise, J. C., Ankney, D. C., & Nelson, W. S. (1990). Mitochondrial gene trees and the evolutionary relationship of mallard and black ducks. *Evolution*, 44, 1109–1119. https://doi.org/10.1111/j.1558-5646.1990.tb03829.x

Baldassarre, G. (2014). *Ducks, geese, and swans of North America*. Baltimore, MD: Johns Hopkins University Press.

Bandelt, H. J., Forster, P., & Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, 16, 37–48. https://doi.org/10.1093/oxfordjournals.molbev.a026036

Besnard, G., Bertrand, J. A., Delahaie, B., Bourgeois, Y. X. C., Lhuillier, E., & Thébaud, C. (2015). Valuing museum specimens: High-throughput DNA sequencing on historical collections of New Guinea crowned pigeons (Goura). *Biological Journal of the Linnean Society*, 117, 71–82.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Catchen, J. M., Hohenlohe, P. A., Bernatchez, L., Funk, W. C., Andrews, K. R., & Allendorf, F. W. (2017). Unbroken: RADseq remains a powerful tool for understanding the genetics of adaptation in natural populations. *Molecular Ecology Resources*, 17(3), 362–365. https://doi.org/10.1111/1755-0998.12669

Champagnon, J., Crochet, P.-A., Kreisinger, J., Čížková, D., Gauthier-Clerc, M., Massez, G., ... Guillemain, M. (2013). Assessing the genetic impact of massive restocking on wild mallard. *Animal Conservation*, 16, 295–305. https://doi.org/10.1111/j.1469-1795.2012.00600.x

Champagnon, J., Elmberg, J., Guillemain, M., Gauthier-Clerc, M., & Lebreton, J.-D. (2012). Conspecifics can be aliens too: A review of effects of restocking practices in vertebrates. *Journal for Nature Conservation*, 20, 231–241. https://doi.org/10.1016/j.jnc.2012.02.002

Cruickshank, T. E., & Hahn, M. W. (2014). Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology*, 23, 3133–3157. https://doi.org/10.1111/mec.12796

DaCosta, J. M., & Sorenson, M. D. (2014). Amplification biases and consistent recovery of loci in a double-digest RAD-seq protocol. *PLoS ONE*, 9, e106713.

Domyan, E. T., Guernsey, M. W., Kronenberg, Z., Krishnan, S., Boissy, R. E., Vickrey, A. I., ... Shapiro, M. D. (2014). Epistatic and combinatorial effects of pigmentary gene mutations in the domestic pigeon. *Current Biology*, 24(4), 459–464. https://doi.org/10.1016/j.cub.2014.01.020

Edwards, S. V., Kingan, S. B., Calkins, J. D., Balakrishnan, C. N., Jennings, W. B., Swanson, W. J., & Sorenson, M. D. (2005). Speciation in birds: Genes, geography, and sexual selection. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 6550–6557. https://doi.org/10.1073/pnas.0501846102

Faircloth, B. C., & Glenn, T. C. (2012). Not all sequence tags are created equal: designing and validating sequence identification tags robust to indels. *PLoS ONE*, 7, e42543.

Feder, J. L., Egan, S. P., & Nosil, P. (2012). The genomics of speciation-with-gene-flow. *Trends in Genetics*, 28, 342–350. https://doi.org/10.1016/j.tig.2012.03.009

Ford, R. J., Selman, W., & Taylor, S. S. (2017). Hybridization between Mottled Ducks (*Anas fulvigula maculosa*) and Mallards (*A. platyrhynchos*) in the western Gulf Coast region. *The Condor*, 119, 683–696.

Francis, R. M. (2016). Pophelper: An R package and web app to analyse and visualize population structure. *Molecular Ecology Resources*, 17, 27–32.

Frankham, R. (2005). Genetics and extinction. *Biological Conservation*, 126, 131–140. https://doi.org/10.1016/j.biocon.2005.05.002

Gasc, C., Peyretaillade, E., & Peyret, P. (2016). Sequence capture by hybridization to explore modern and ancient genomic diversity in model and nonmodel organisms. *Nucleic Acids Research*, 44, 4504–4518. https://doi.org/10.1093/nar/gkw309

Gasca-Pineda, J., Cassaigne, I., Alonso, R. A., & Eguiarte, L. E. (2013). Effective population size, genetic variation, and their relevance for conservation: The bighorn sheep in Tiburon Island and comparisons with managed artiodactyls. *PLoS ONE*, 8, e78120. https://doi.org/10.1371/journal.pone.0078120

Gourbière, S., & Mallet, J. (2010). Are species real? The shape of the species boundary with exponential failure, reinforcement, and the "missing snowball". *Evolution*, 64, 1–24. https://doi.org/10.1111/j.1558-5646.2009.00844.x

Graham, C. F., Glenn, T. C., McArthur, A. G., Boreham, D. R., Kieran, T., Lance, S., ... Somers, C. M. (2015). Impacts of degraded DNA on restriction enzyme associated DNA sequencing (RADSeq). *Molecular Ecology Resources*, 15, 1304–1315. https://doi.org/10.1111/1755-0998.12404

Griffin, H. R., Pyle, A., Blakely, E. L., Alston, C. L., Duff, J., Hudson, G., ... Chinnery, P. F. (2014). Accurate mitochondrial DNA sequencing using off-target reads provides a single test to identify pathogenic point mutations. *Genetics in Medicine*, 16, 962. https://doi.org/10.1038/gim.2014.66

Gutenkunst, R. N., Hernandez, R. D., Williamson, S. H., & Bustamante, C. D. (2009). Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genetics*, 5, e1000695. https://doi.org/10.1371/journal.pgen.1000695

Gutenkunst, R. N., Hernandez, R. D., Williamson, S. H., & Bustamante, C. D. (2010). *Diffusion approximations for demographic inference: DaDi*.

Haldane, J. B. S. (1922). Sex ratio and unisexual sterility in hybrid animals. *Journal of Genetics*, 12, 101–109. https://doi.org/10.1007/BF02983075

Haldane, J. (1948). The theory of a cline. *Journal of Genetics*, 48, 277–284. https://doi.org/10.1007/BF02986626

Hanson, C. H., Rogers, M., & Rogers, E. S. (1949). Waterfowl of the forested portions of the Canadian pre-Cambrian shield and the Palaeozoic Basin. *The Canadian Field-Naturalist*, 63, 183–204.

Hepp, G. R., Novak, J. M., Scribner, K. T., & Stangel, P. W. (1988). Genetic distance and hybridization of black ducks and mallards: A morph of a different color? *The Auk*, 105, 804–807.

Heusmann, H. W. (1974). Mallard-black duck relationships in the Northeast. *Wildlife Society Bulletin*, 2, 171–177.

Heusmann, H. (2017). Why are numbers of mallards (*Anas platyrhynchos*) declining in the Northeastern United States? The possible role of reduction in winter feeding sites? *The Wilson Journal of Ornithology*, 129, 528–534.

Hou, Z. C., Zhu, F., Yin, Z. T., & Zhang, F. (Unpublished). Three high-quality genomes provides insights into domestication of ducks. BioProject PRJNA554956, BioSample SAMN12287667.

Hupp, J. W., Yamaguchi, N., Flint, P. L., Pearce, J. M., Tokita, K.-I., Shimada, T., ... Higuchi, H. (2011). Variation in spring migration routes and breeding distribution of northern pintails *Anas acuta* that winter in Japan. *Journal of Avian Biology*, 42, 289–300. https://doi.org/10.1111/j.1600-048X.2011.05320.x

Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, 23, 1801–1806. https://doi.org/10.1093/bioinformatics/btm233

Johnsgard, P. A. (1967). Sympatry changes and hybridization incidence in mallards and black ducks. *American Midland Naturalist*, 77, 51–63. https://doi.org/10.2307/2423425

Johnson, K. P., & Sorenson, M. D. (1999). Phylogeny and biogeography of dabbling ducks (genus: *Anas*): A comparison of molecular and morphological evidence. *The Auk*, 116, 792–805. https://doi.org/10.2307/4089339

Jombart, T. (2008). Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24, 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L., & Orlando, L. (2013). mapDamage2.0: Fast approximate Bayesian estimates of

ancient DNA damage parameters. *Bioinformatics*, *29*, 1682–1684. https://doi.org/10.1093/bioinformatics/btt193

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., … Drummond, A. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, *28*, 1647–1649. https://doi.org/10.1093/bioinformatics/bts199

Keyser-Tracqui, C., & Ludes, B. (2005). Methods for the study of ancient DNA. In W. Goodwin (Ed.), *Forensic DNA typing protocols* (pp. 253–264). Springer.

Kidd, A., Bowman, J., Lesbarreres, D., & Schulte-Hostedde, A. (2009). Hybridization between escaped domestic and wild American mink (Neovison vison). *Molecular Ecology*, *18*, 1175–1186.

Kirby, R. E., Sargeant, G. A., & Shutler, D. (2004). Haldane's rule and American black duck × mallard hybridization. *Canadian Journal of Zoology*, *82*, 1827–1831. https://doi.org/10.1139/z04-169

Knapp, M., Clarke, A. C., Horsburgh, K. A., & Matisoo-Smith, E. A. (2012). Setting the stage – Building and working in an ancient DNA laboratory. *Annals of Anatomy-Anatomischer Anzeiger*, *194*, 3–6. https://doi.org/10.1016/j.aanat.2011.03.008

Kraus, R. H. S., Kerstens, H. H. D., Van Hooft, P., Crooijmans, R. P. M. A., Van Der Poel, J. J., Elmberg, J., … Groenen, M. A. M. (2011). Genome wide SNP discovery, analysis and evaluation in mallard (*Anas platyrhynchos*). *BMC Genomics*, *12*, 150. https://doi.org/10.1186/1471-2164-12-150

Kress, W. J. (2014). Valuing collections. *Science*, *346*, 1310.

Kulikova, I. V., Drovetski, S. V., Gibson, D. D., Harrigan, R. J., Rohwer, S., Sorenson, M. D., … McCracken, K. G. (2005). Phylogeography of the mallard (*Anas platyrhynchos*): Hybridization, dispersal, and lineage sorting contribute to complex geographic structure. *The Auk*, *122*, 949–965. https://doi.org/10.1093/auk/122.3.949

Kulikova, I. V., Zhuravlev, Y. N., & McCracken, K. G. (2004). Asymmetric hybridization and sex-biased gene flow between eastern spot-billed ducks (*Anas zonorhyncha*) and mallards (*A. platyrhynchos*) in the Russian Far East. *The Auk*, *121*, 930–949. https://doi.org/10.1093/auk/121.3.930

Laikre, L., & Ryman, N. (1996). Effects on intraspecific biodiversity from harvesting and enhancing natural populations. *Ambio*, *25*, 504–509.

Laikre, L., Schwartz, M. K., Waples, R. S., Ryman, N., & Group GW (2010). Compromising genetic diversity in the wild: Unmonitored large-scale release of plants and animals. *Trends in Ecology & Evolution*, *25*, 520–529. https://doi.org/10.1016/j.tree.2010.06.013

Lam, T.-Y., Ip, H. S., Ghedin, E., Wentworth, D. E., Halpin, R. A., Stockwell, T. B., … Holmes, E. C. (2012). Migratory flyway and geographical distance are barriers to the gene flow of influenza virus among North American birds. *Ecology Letters*, *15*, 24–33. https://doi.org/10.1111/j.1461-0248.2011.01703.x

Lavretsky, P., Dacosta, J. M., Hernández-Baños, B. E., Engilis, A., Sorenson, M. D., & Peters, J. L. (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. https://doi.org/10.1111/mec.13402

Lavretsky, P., DaCosta, J. M., Sorenson, M. D., McCracken, K. G., & Peters, J. L. (2019). ddRAD-seq data reveal significant genome-wide population structure and divergent genomic regions that distinguish the mallard and close relatives in North America. *Molecular Ecology*, *28*, 2594–2609. https://doi.org/10.1111/mec.15091

Lavretsky, P., Hernández Baños, B. E., & Peters, J. L. (2014). Rapid radiation and hybridization contribute to weak differentiation and hinder phylogenetic inferences in the new world mallard complex (*Anas* spp.). *The Auk*, *131*, 524–538.

Lavretsky, P., Janzen, T., & McCracken, K. G. (2019). Identifying hybrids & the genomics of hybridization: Mallards & American black ducks of eastern North America. *Ecology & Evolution*, *9*, 3470–3490.

Lavretsky, P., McCracken, K. G., & Peters, J. L. (2014). Phylogenetics of a recent radiation in the mallards and allies (Aves: *Anas*): Inferences from a genomic transect and the multispecies coalescent. *Molecular*

Phylogenetics and Evolution, *70*, 402–411. https://doi.org/10.1016/j.ympev.2013.08.008

Lavretsky, P., Peters, J. L., Winker, K., Bahn, V., Kulikova, I., Zhuravlev, Y. N., … McCracken, K. G. (2016). Becoming pure: Identifying generational classes of admixed individuals within lesser and greater scaup populations. *Molecular Ecology*, *25*(3), 661–674. https://doi.org/10.1111/mec.13487

Lawson, D. J., Hellenthal, G., Myers, S., & Falush, D. (2012). Inference of population structure using dense haplotype data. *PLoS Genetics*, *8*, e1002453. https://doi.org/10.1371/journal.pgen.1002453

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

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Lischer, H. E., & Excoffier, L. (2011). PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, *28*, 298–299.

Lonsinger, R. C., Adams, J. R., & Waits, L. P. (2018). Evaluating effective population size and genetic diversity of a declining kit fox population using contemporary and historical specimens. *Ecology and Evolution*, *8*, 12011–12021. https://doi.org/10.1002/ece3.4660

Lowry, D. B., Hoban, S., Kelley, J. L., Lotterhos, K. E., Reed, L. K., Antolin, M. F., & Storfer, A. (2016). Breaking RAD: An evaluation of the utility of restriction site-associated DNA sequencing for genome scans of adaptation. *Molecular Ecology Resources*, *17*(2), 142–152. https://doi.org/10.1111/1755-0998.12635

Makino, T., Rubin, C.-J., Carneiro, M., Axelsson, E., Andersson, L., & Webster, M. T. (2018). Elevated proportions of deleterious genetic variation in domestic animals and plants. *Genome Biology and Evolution*, *10*(1), 276–290. https://doi.org/10.1093/gbe/evy004.

Malinsky, M., Trucchi, E., Lawson, D., & Falush, D. (2018). Radpainter and fineRADstructure: Population inference from RADseq data. *bioRxiv*, *057711*. https://doi.org/10.1093/molbev/msy023

Mank, J. E., Carlson, J. E., & Brittingham, M. C. (2004). A century of hybridization: Decreasing genetic distance between American black ducks and mallards. *Conservation Genetics*, *5*, 395–403. https://doi.org/10.1023/B:COGE.0000031139.55389.b1

McCracken, K. G., Johnson, W. P., & Sheldon, F. H. (2001). Molecular population genetics, phylogeography, and conservation biology of the mottled duck (*Anas fulvigula*). *Conservation Genetics*, *2*, 87–102.

McDonough, M. M., Parker, L. D., Rotzel McInerney, N., Campana, M. G., & Maldonado, J. E. (2018). Performance of commonly requested destructive museum samples for mammalian genomic studies. *Journal of Mammalogy*, *99*, 789–802. https://doi.org/10.1093/jmammal/gyy080

Merendino, M. T., & Ankney, C. D. (1994). Habitat use by mallards and American black ducks breeding in central Ontario. *The Condor*, *96*, 411–421. https://doi.org/10.2307/1369324

Miller, C. R., & Waits, L. P. (2003). The history of effective population size and genetic diversity in the Yellowstone grizzly (*Ursus arctos*): Implications for conservation. *Proceedings of the National Academy of Sciences*, *100*, 4334–4339. https://doi.org/10.1073/pnas.0735531100

Nichols, J. D., Obrecht, H. H. III, & Hines, J. E. (1987). Survival and band recovery rates of sympatric American black ducks and mallards. *The Journal of Wildlife Management*, *51*(4), 700–710. https://doi.org/10.2307/3801730

Nosil, P., Harmon, L. J., & Seehausen, O. (2009). Ecological explanations for (incomplete) speciation. *Trends in Ecology & Evolution*, *24*, 145–156. https://doi.org/10.1016/j.tree.2008.10.011

Nosil, P., & Schluter, D. (2011). The genes underlying the process of speciation. *Trends in Ecology & Evolution*, *26*, 160–167. https://doi.org/10.1016/j.tree.2011.01.001

Omland, K. E. (1997). Examining two standard assumptions of ancestral reconstructions: Repeated loss of dichromatism in

dabbling ducks (*Anatini*). *Evolution*, *51*(5), 1636–1646. https://doi.org/10.1111/j.1558-5646.1997.tb01486.x

Osborne, C. E., Swift, B. L., & Baldassarre, G. A. (2010). Fate of captive-reared and released mallards on eastern Long Island, New York. *Human-Wildlife Interactions*, *4*, 266–274.

Payseur, B. A., & Rieseberg, L. H. (2016). A genomic perspective on hybridization and speciation. *Molecular Ecology*, *25*, 2337–2360.

Pearce, J. M., Reeves, A. B., Ramey, A. M., Hupp, J. W., Ip, H. S., Bertram, M., … Runstadler, J. A. (2011). Interspecific exchange of avian influenza virus genes in Alaska: The influence of trans-hemispheric migratory tendency and breeding ground sympatry. *Molecular Ecology*, *20*, 1015–1025. https://doi.org/10.1111/j.1365-294X.2010.04908.x

Peters, J. L., Lavretsky, P., DaCosta, J. M., Bielefeld, R. R., Feddersen, J. C., & Sorenson, M. D. (2016). Population genomic data delineate conservation units in mottled ducks (*Anas fulvigula*). *Biological Conservation*, *203*, 272–281. https://doi.org/10.1016/j.biocon.2016.10.003

Peters, J. L., Zhuravlev, Y., Fefelov, I., Humphries, E. M., & Omland, K. E. (2008). Multilocus phylogeography of a Holarctic duck: Colonization of North America from Eurasia by gadwalls (*Anas strepera*). *Evolution*, *62*, 1469–1483.

Pfeifer, B., Wittelsbürger, U., Ramos-Onsins, S. E., & Lercher, M. J. (2014). PopGenome: An efficient Swiss army knife for population genomic analyses in R. *Molecular Biology and Evolution*, *31*, 1929–1936. https://doi.org/10.1093/molbev/msu136

Picardi, E., & Pesole, G. (2012). Mitochondrial genomes gleaned from human whole-exome sequencing. *Nature Methods*, *9*, 523. https://doi.org/10.1038/nmeth.2029

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., … Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, *81*, 559–575. https://doi.org/10.1086/519795

Rhymer, J. M. (2006). Extinction by hybridization and introgression in anatine ducks. *Acta Zoologica Sinica*, *52*(Supplement), 583–586.

Rhymer, J. M., & Simberloff, D. (1996). Extinction by hybridization and introgression. *Annual Review of Ecology and Systematics*, *27*, 83–109. https://doi.org/10.1146/annurev.ecolsys.27.1.83

Sæther, B.-E., Lande, R., Engen, S., Weimerskirch, H., Lillegård, M., Altwegg, R., … Visser, M. E. (2005). Generation time and temporal scaling of bird population dynamics. *Nature*, *436*, 99–102. https://doi.org/10.1038/nature03666

Schubert, M., Ginolhac, A., Lindgreen, S., Thompson, J. F., AL-Rasheid, K. A. S., Willerslev, E., … Orlando, L. (2012). Improving ancient DNA read mapping against modern reference genomes. *BMC Genomics*, *13*, 178. https://doi.org/10.1186/1471-2164-13-178

Servedio, M. R., & Noor, M. A. F. (2003). The role of reinforcement in speciation: Theory and data. *Annual Review of Ecology, Evolution, and Systematics*, *34*, 339–364. https://doi.org/10.1146/annurev.ecolsys.34.011802.132412

Shin, J.-H., Lee, K.-S., Kim, S.-H., Hwang, J. K., Woo, C., Kim, J., … Mo, I. P. (2015). Tracking mallards (*Anas platyrhynchos*) with GPS satellite transmitters along their migration route through Northeast Asia. *Avian Diseases*, *60*, 311–315.

Skoglund, P., Ersmark, E., Palkopoulou, E., & Dalén, L. (2015). Ancient wolf genome reveals an early divergence of domestic dog ancestors and admixture into high-latitude breeds. *Current Biology*, *25*, 1515–1519. https://doi.org/10.1016/j.cub.2015.04.019

Snell, E. A. (1986). Wetland loss in southern Ontario. In Directorate L (Ed.), *Wetland distribution and conversion in southern Ontario* (p. 45). Ottawa, Canada: Environment Canada.

Söderquist, P., Elmberg, J., Gunnarsson, G., Thulin, C.-G., Champagnon, J., Guillemain, M., … Kraus, R. H. S. (2017). Admixture between released and wild game birds: A changing genetic landscape in European mallards (*Anas platyrhynchos*). *European Journal of Wildlife Research*, *63*, 98. https://doi.org/10.1007/s10344-017-1156-8

Söderquist, P., Gunnarsson, G., & Elmberg, J. (2013). Longevity and migration distance differ between wild and hand-reared mallards *Anas platyrhynchos* in Northern Europe. *European Journal of Wildlife Research*, *59*, 159–166. https://doi.org/10.1007/s10344-012-0660-0

Söderquist, P., Norrström, J., Elmberg, J., Guillemain, M., & Gunnarsson, G. (2014). Wild mallards have more "goose-like" bills than their ancestors: A case of anthropogenic influence? *PLoS ONE*, *9*, e115143. https://doi.org/10.1371/journal.pone.0115143

Sorenson, M. D., Ast, J. C., Dimcheff, D. E., Yuri, T., & Mindell, D. P. (1999). Primers for a PCR-based approach to mitochondrial genome sequencing in birds and other vertebrates. *Molecular Phylogenetics and Evolution*, *12*, 105–114. https://doi.org/10.1006/mpev.1998.0602

Sorenson, M. D., & Fleischer, R. C. (1996). Multiple independent transpositions of mitochondrial DNA control region sequences to the nucleus. *Proceedings of the National Academy of Sciences of the United States of America*, *93*, 15239–15243. https://doi.org/10.1073/pnas.93.26.15239

Soutiere, E. C. (1986). Hand-reared mallard releases on 3 private farms in Maryland. *Proceedings of the annual conference southeast association*, *40*, 438–445.

Souza, C. A., Murphy, N., Villacorta-Rath, C., Woodings, L. N., Ilyushkina, I., Hernandez, C. E., … Strugnell, J. M. (2017). Efficiency of ddRAD target enriched sequencing across spiny rock lobster species (Palinuridae: Jasus). *Scientific Reports*, *7*, 6781. https://doi.org/10.1038/s41598-017-06582-5

Todesco, M., Pascual, M. A., Owens, G. L., Ostevik, K. L., Moyers, B. T., Hübner, S., … Rieseberg, L. H. (2016). Hybridization and extinction. *Evolutionary Applications*, *9*, 892–908. https://doi.org/10.1111/eva.12367

Tufto, J. (2017). Domestication and fitness in the wild: A multivariate view. *Evolution*, *71*(9), 2262–2270. https://doi.org/10.1111/evo.13307

USFWS (2013). *Review of captive-reared mallard regulations on shooting preserves - FINAL REPORT*. Washington: Division of Migratory Bird Management, US Fish and Wildlife Service.

Wayne, R. K., & Shaffer, H. B. (2016). Hybridization and endangered species protection in the molecular era. *Molecular Ecology*, *25*, 2680–2689.

Wu, D.-D., Ding, X.-D., Wang, S., Wójcik, J. M., Zhang, Y. I., Tokarska, M., … Zhang, Y.-P. (2018). Pervasive introgression facilitated domestication and adaptation in the Bos species complex. *Nature Ecology & Evolution*. https://doi.org/10.1038/s41559-018-0562-y

Yamaguchi, N., Hiraoka, E., Fujita, M., Hijikata, N., Ueta, M., Takagi, K., … Higuchi, H. (2008). Spring migration routes of mallards (*Anas platyrhynchos*) that winter in Japan, determined from satellite telemetry. *Zoological Science*, *25*, 875–881. https://doi.org/10.2108/zsj.25.875

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.