

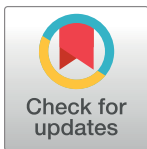
RESEARCH ARTICLE

Population genetics and geographic origins of mallards harvested in northwestern Ohio

Michael L. Schummer^{1*}, John Simpson², Brendan Shirkey², Samuel R. Kucia¹, Philip Lavretsky³, Douglas C. Tozer⁴

1 Department of Environmental Biology, State University of New York College of Environmental Science and Forestry, Syracuse, New York, United States of America, **2** Winous Point Marsh Conservancy, Port Clinton, Ohio, United States of America, **3** Department of Biological Sciences, University of Texas at El Paso, El Paso, Texas, United States of America, **4** Long Point Waterfowl and Wetlands Research Program, Birds Canada, Port Rowan, Ontario, Canada

* mlschumm@esf.edu



Abstract

The genetic composition of mallards in eastern North America has been changed by release of domestically-raised, game-farm mallards to supplement wild populations for hunting. We sampled 296 hatch-year mallards harvested in northwestern Ohio, October–December 2019. The aim was to determine their genetic ancestry and geographic origin to understand the geographic extent of game-farm mallard introgression into wild populations in more westward regions of North America. We used molecular analysis to detect that 35% of samples were pure wild mallard, 12% were early generation hybrids between wild and game-farm mallards (i.e., F1–F3), and the remaining 53% of samples were assigned as part of a hybrid swarm. Percentage of individuals in our study with some form of hybridization with game-farm mallard (65%) was greater than previously detected farther south in the mid-continent (~4%), but less than the Atlantic coast of North America (~92%). Stable isotope analysis using $\delta^2\text{H}_f$ suggested that pure wild mallards originated from areas farther north and west than hybrid mallards. More specifically, 17% of all Ohio samples had $\delta^2\text{H}_f$ consistent with more western origins in the prairies, parkland, or boreal regions of the mid-continent of North America, with 55%, 35%, and 10% of these being genetically wild, hybrid swarm, and F3, respectively. We conclude that continued game-farm introgression into wild mallards is not isolated to the eastern population of mallards in North America, and may be increasing and more widespread than previously detected. Mallards in our study had greater incidence of game-farm hybridization than other locales in the mid-continent but less than eastern North American regions suggesting further need to understand game-farm mallard genetic variation and movement across the continent.

OPEN ACCESS

Citation: Schummer ML, Simpson J, Shirkey B, Kucia SR, Lavretsky P, Tozer DC (2023) Population genetics and geographic origins of mallards harvested in northwestern Ohio. PLoS ONE 18(3): e0282874. <https://doi.org/10.1371/journal.pone.0282874>

Editor: Tzen-Yuh Chiang, National Cheng Kung University, TAIWAN

Received: August 25, 2022

Accepted: February 25, 2023

Published: March 15, 2023

Copyright: © 2023 Schummer et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All raw Illumina sequences are available from the National Center for Biotechnology Information Sequence Read Archive (accession no. 32175707–32176002, BioProject accession no. PRJNA911314), with sample-specific information found in Supporting Information [S1 Table](#). All relevant isotopic data can be found in Supporting Information [S1 Table](#).

Funding: PL and MLS National Science Foundation (DEB Grant ID 2010704), nsf.gov, MLS, JS, and BS funding by Winous Point Marsh Conservancy, MLS

Introduction

Humans and mallards (*Anas platyrhynchos*) have a closely linked history since their domestication in China around 500 B. C. [1, 2]. Nearly all domestic ducks (e.g., Pekin duck) are

and DZ funding by Birds Canada. Sponsors and funders did not play a role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

decendents of wild, Eurasian mallards [1]. The intentional supplementation of wild populations is common practice for many game species [3–6] including waterfowl [7, 8]. Although wild mallard populations naturally have a Holarctic distribution, the intentional introduction of wild and domesticated forms has artificially increased their distribution to nearly everywhere but the Poles [9]. Among domestic mallards, the intentional release of the game-farm mallard breed is responsible for the majority of their range expansion. Started in the early 20th century, the release of game-farm mallards continues intensively today in Europe (6-million annually) [10, 11] and eastern North America ($\geq 200,000$ annually) [12–14] largely for the supplementation of wild mallard populations for hunting purposes [14, 15]. Much like other domesticated mallard forms, all game-farm mallards have a Eurasian origin, and those being released in Europe and North America are from the same genetic source [10–12]. Introduction of game-farm mallards were thought to have no impact on wild populations, but only recently have molecular studies begun to estimate levels of introgression of game-farm genes into wild mallard populations. Wild mallards are genetically panmictic [16], but release of game-farm mallards has changed the genetic ancestry of wild populations in Europe [17] and North America [18, 19]. Increasing prevalence of game-farm genes within wild populations is predicted to decrease mallard fitness because of introduction of artificially selected, maladaptive traits [17, 19], and this has been hypothesized to at least partially explain recent mallard population declines in eastern North America [19] and central Europe [17].

In North America, studies continue to find a close association among game-farm mallard ancestry prevalence and regions of intense game-farm mallard release programs [14, 19]. Among regions, > 90% of releases of game-farm mallards have occurred along the Atlantic coast of North America: ~500,000 game-farm mallards were released annually from 1920–1950 [13], and $\geq 200,000$ have been released annually since, although these numbers are not well-documented [14, 15]. Elsewhere in North America release of game-farm mallards were at a smaller scale (e.g., 16,983 released as part of population supplementation programs in Ohio from 1959–1963) [20]. By 2010, only ~8% of mallards on the Atlantic coast were pure wild [18] and the prevalence of game-farm $\times$ wild mallard hybrids generally decreases westward [18], but finer scale studies of mallard genetics are needed to understand the movement and spatial distribution of game-farm $\times$ wild mallard hybrids. Specifically, Lavretsky et al. [18] detected that ~40% of mallards sampled throughout the Mississippi flyway of North America had substantial game-farm mallard introgression, whereas southern parts of the same flyway only reported ~4% as hybrids [21]. Thus, the distribution of and mechanisms for game-farm mallard ancestry throughout North America remains a critical information gap.

Mallards in North America are managed as western, mid-continent, and eastern breeding populations (Fig 1). Among these, the eastern mallard population steadily increased until 1999, with populations declining thereafter by 40% [22, 23]. However, observed declines were unequal across their eastern range [23]. The Atlantic Flyway Breeding Waterfowl Plot Survey (AFBWPS), which occurs from Virginia to New Hampshire, identified mallard population declines of ~50% (880,000 in 1999 to 448,500 in 2017), whereas the eastern survey area of the Waterfowl Breeding Population and Habitat Survey (WBPHS) that includes eastern Canada and Maine suggested stable to slightly increasing abundance from 171,000 to 204,200 during the same time period [23]. The combined population estimates from the AFBWPS and WBPHS are used to monitor the eastern mallard population and suggest an ~38% decline. The mid-continent population of mallards is described as those in the traditional WBPHS area (strata 13–18, 20–50, and 75–77) and the Great Lakes (Michigan, Minnesota, and Wisconsin) (Fig 1). The mallard population in the traditional WBPHS area increased by 75% from 2005 to 2016 and although it has since declined from this historic high (11.8 million), it remains 19% above the long-term average of 7.9 million [24]. In contrast, the mallard population in the

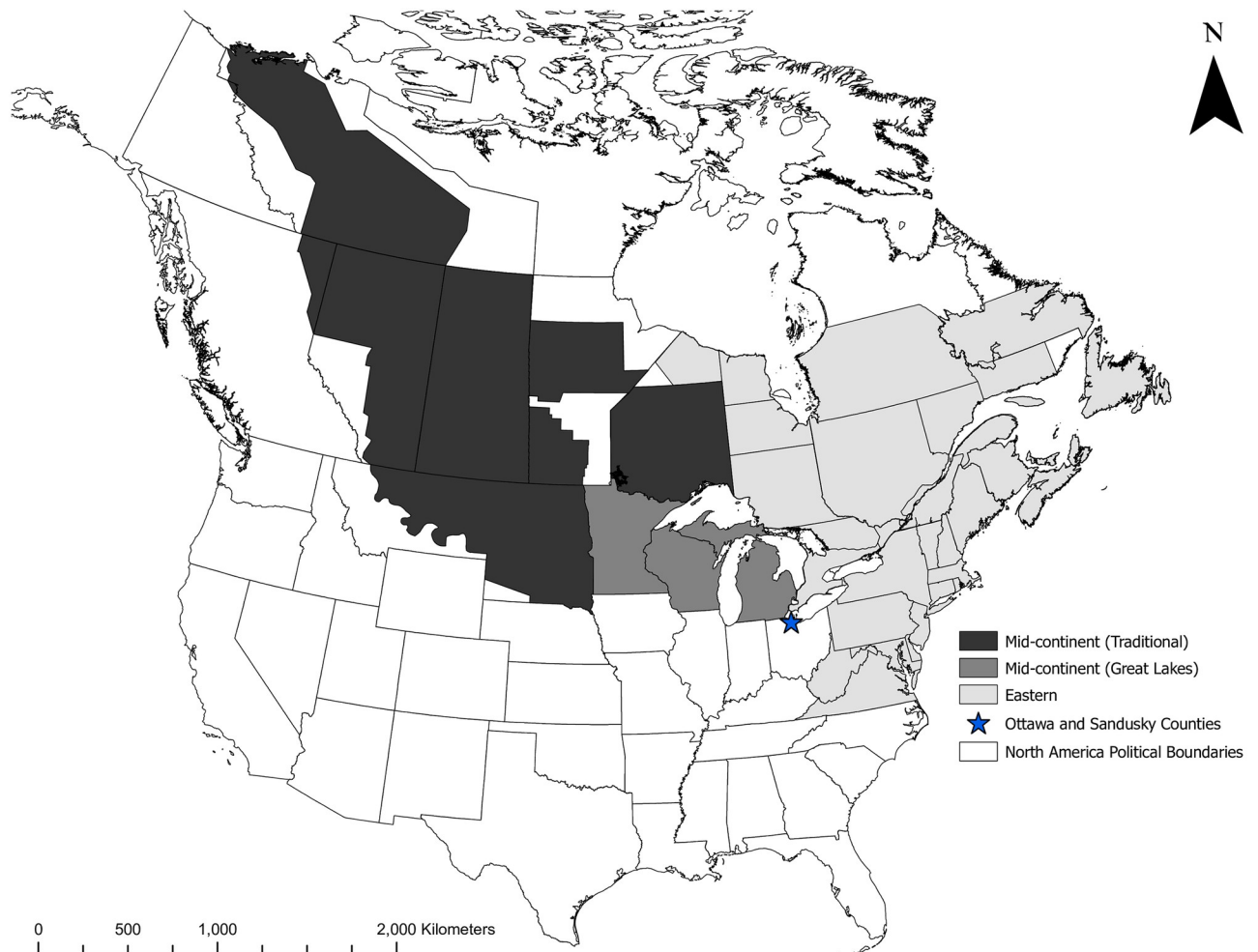


Fig 1. Waterfowl survey areas currently assigned to the eastern and mid-continent stocks of mallards as identified by the United States Fish and Wildlife Service (USFWS) for management purposes. Spatial layers of populations were provided by the USFWS and base outline political boundary layers are free, open-source from ESRI. WBPBS, Eastern Survey, and Great Lakes shapefiles were provided by USFWS. U.S. political boundaries were obtained through the U.S. Census Bureau (<https://www.census.gov/geographies/mapping-files/time-series/geo/cartto-boundary-file.html>). Canadian province political boundaries were obtained through Open Canada census Bureau (<https://open.canada.ca/data/en/dataset/a883eb14-0c0e-45c4-b8c4-b54c4a819edb>). Except for the shapefiles listed above, no external data was used for the creation of the figure. ESRI ArcMap 10.8 was used for the creation of the figure.

<https://doi.org/10.1371/journal.pone.0282874.g001>

Great Lakes portion of the mid-continent area was about 28% lower from 2006 to 2020 (mean = 0.72 million) than from 1995 to 2005 (mean = 1.00 million) [24]. Prevalence of game-farm genes in mallards does decrease moving westward, but game-farm and hybrid mallards still made up ~40% of samples of the mid-continent population [18]. More recently low rates of game-farm ancestry were detected at southern portions of the mid-continent during winter (~4%) [21], suggesting substantial regional variation in game-farm introgression. The Great Lakes region is geographically at the intersection where mallards from the mid-continent and eastern populations mix [25–27]. Consequently, we thought the Great Lakes region would have a greater proportion of mallards carrying game-farm mallard ancestry as compared to the overall mid-continent population. Declining mallard breeding abundance in the Great Lakes region is consistent with hypothesis that game-farm traits are maladaptive when they enter the wild mallard population and at least partially explain population declines in eastern North

America. Here, we not only focus on describing the genetic composition of mallards in the Great Lakes region, but link genetic ancestry with geographical breeding origins through isotopic analyses.

We focused on determining the genetic composition and natal origins of mallards migrating through the Great Lakes region by analyzing mallards harvested at private-marsh complexes in northwestern Ohio during the 2019 autumn hunting season. We assayed thousands of nuclear loci and assigned genetic ancestry across samples, enabling us to determine the proportions comprised of wild, game-farm, and filial classes of game-farm $\times$ wild hybrids (e.g., F1, F2, etc.) of our sample. We also assigned these mallards to geographic natal origins using the ratio of stable isotopes of hydrogen ($^2\text{H}/^1\text{H}$) from feathers (hereafter $\delta^2\text{H}_f$) grown on natal areas because these isotopes produce a strong, predictable latitudinal gradient across broad geographic regions [28–31]. Unlike band-recovery data that can be spatially limited by accessibility and not representative of origins if juvenile mallards are not banded before they could fly [25, 32], a stable isotope approach decreases spatial biases associated with specific banding locations because sampling of metabolically inert flight feathers grown during development can be used to link ducks to remote areas of origin [33]. We sampled hatch-year (HY) mallards because they grow feathers on breeding grounds prior to fledging. As such, we were able to estimate breeding locations where game-farm derived genetic variation may be more likely to originate, and their proportional contributions to the mallard population migrating through an important stop-over location in the Great Lakes region. Mallards using the marshes of northwestern Ohio during their autumn-winter migration originate from mid-continent and eastern breeding populations [25–27]. Consequently, mallards harvested in this region provided a random sampling opportunity of individuals from much of the eastern half of North America during their autumn-winter migration. Similar to Palumbo et al. [34], we also determined percentages of mallards that could be classified as being produced locally (i.e., ≤ 1 standard deviation in $\delta^2\text{H}_f$ of our sampling locations) and those with an isotopic signature of an immigrant to determine the extent that we sampled mallards from a broad geographic area. We predicted that pure wild mallards would originate from farther north and west (i.e., mid-continent population) than game-farm admixed individuals that primarily exist farther east in North America. We also predicted more northern mallards with increasingly wild genetic ancestry and respective stable hydrogen isotopic values to be present in later months of our sampling period as they migrated through and became increasingly available for harvest.

Materials and methods

Sample collection

We obtained HY mallards that were legally harvested at private-hunting clubs in Ottawa and Sandusky counties in northwestern Ohio during waterfowl hunting season, 4 October–19 December 2019. Use of animals in this study followed all federal and state guidelines regarding the ethics of animal welfare. All ducks were aged and sexed using plumage characteristics [35]. We sampled wing muscle and P1 (i.e., the primary feather adjacent to the secondaries) from each mallard for molecular and isotopic analyses, respectively.

DNA extraction, ddRAD-seq libraries, sequencing, and bioinformatics

Genomic DNA was extracted from 296 mallards using and following manufacturer's protocols for the DNeasy Blood and Tissue kit (Qiagen, Valencia, CA, USA). DNA quality was ensured based on the presence of high molecular weight bands visualized using gel electrophoresis with a 1% agarose gel, and quantified DNA concentration using a Qubit 3 Fluorometer (Invitrogen, Carlsbad, CA) to ensure a minimum concentration of 20 ng/ μL . Following, double-digestion

we constructed restriction-site associated DNA sequencing (ddRAD-seq) fragment libraries following Lavretsky et al. [36], but with fragment size selection for average fragment lengths of 350 base-pairs followed Hernandez et al. [37]. In short, genomic DNA was enzymatically fragmented using SbfI and EcoRI restriction enzymes (New England Biolabs, Ipswich, MA), and Illumina TruSeq (Illumina, San Diego, CA) compatible barcodes ligated, permitting future de-multiplexing. Samples were pooled in equimolar concentrations and sent for 150 base pair, single-end sequencing on an Illumina HiSeq X with Novogenetics (Sacramento, CA, USA). Raw Illumina reads were deposited in the National Center for Biotechnology Information's Sequence Read Archive (SRA; <http://www.ncbi.nlm.nih.gov/sra>; BioProject TBD).

We de-multiplexed raw Illumina reads using the *ddRADparser.py* script of the BU ddRAD-seq pipeline [38] based on perfect barcode/index matches. We note that previously published ddRAD raw sequence data generated using the same protocols as described above for other wild mallards [39] and game-farm mallards [19] were also included in analyses, serving as references in molecular analyses. Bioinformatics followed Lavretsky et al. [19] with all steps through genotyping automated using a custom in-house Python script (Python scripts available at <https://github.com/jonmohl/PopGen>). We first trimmed or discarded poor quality sequences using trimmomatic [40] with the remaining quality reads aligned to a reference mallard genome [2] using the Burrows Wheeler Aligner v. 07.15 (bwa; [41]). Next, samples were sorted and indexed in Samtools v. 1.6 [42] 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 only high-quality sequences are retained. Finally, we further filtered for any base-pair missing $>5\%$ of samples, which also required a minimum base-pair sequencing depth coverage of 5X (i.e., 10X per genotype) and quality per base PHRED scores of ≥ 30 using VCFtools v. 0.115 [43].

Though only autosomal loci were used in subsequent population structure analyses, the alignment of ddRAD-seq loci to our reference genome also provided positions across sex chromosomes, allowing us to use differences in expected sequencing depth to demarcate sex across samples. The heterogametic sex (i.e., female; ZW) is expected to have half the sequencing depth across both sex-chromosomes as compared to autosome-linked loci, whereas the homogametic sex (i.e., males; ZZ) showing the same sequencing depth at Z-sex chromosome and autosome linked loci and no W-linked loci (also see [36]). Raw Illumina reads were deposited in the National Center for Biotechnology Information's Sequence Read Archive (SRA; <http://www.ncbi.nlm.nih.gov/sra>; BioProject PRJNA911314; See S1 Table for sample specific accession numbers).

Population structure

Population structure was evaluated using autosomal ddRAD-seq bi-allelic single nucleotide polymorphisms (SNPs) only. Prior to analyses, datasets were filtered for singletons (i.e., minimum allele frequency [maf] = 0.0023) and any SNP missing $\geq 5\%$ of data across samples using PLINK v. 0.70 [44]. Independence between SNPs was based on pair-wise analysis of linkage disequilibrium (LD) (—indep-pairwise 2 1 0.5) in which 1 of 2 linked SNPs are randomly excluded if an LD correlation factor (r^2) > 0.5 was obtained. All analyses were done without *a priori* information on population or species identity.

Population structure was visualized using the PCA function in PLINK to perform a principal component analysis (PCA). Next, we used ADMIXTURE version 1.3 [45, 46] to attain maximum likelihood estimates of population assignments for each individual, following dataset preparation steps as outlined in Alexander et al. [47]. ADMIXTURE was run with a 10-fold cross validation, incorporating a quasi-Newton algorithm to accelerate convergence [48]. Each

analysis used a block relaxation algorithm for point estimation and terminated once the change in the log-likelihood of the point estimations increased by <0.0001 . Each analysis was run for K populations of 1 through 5, and with 100 iterations per each value of K . The optimum K in each analysis was based on the average of cross-validation errors across the iterations per K value; however, we examined additional values of K to test for further structural resolution across analyses. We used the R package PopHelper [49] to convert ADMIXTURE outputs into CLUMPP input files at each K value, and to determine the robustness of assignments of individuals to populations at each K value with the program CLUMPP version 1.1 [50]. In CLUMPP, we employed the Large Greedy algorithm and 1,000 random permutations. Final admixture proportions for each K value and per sample assignment probabilities (Q estimates; the log likelihood of group assignment) were based on CLUMPP analyses of all 100 replicates per K value.

Establishing hybrid indices and hybrid assignment

To assess the effectiveness of our molecular dataset in distinguishing between classes of generational hybrids, and to more directly assign putative hybrids to those classes, we simulated expected assignment probabilities with our empirical data for first generation hybrids (F1) and 9 generations of backcrosses (F2–F10) based on methods outlined in Lavretsky et al. [51]. We used the same ddRAD-seq bi-allelic nuclear SNP set analyzed for population structure in simulations, but only including reference wild mallard and game-farm mallards. We first generated 10 F1 hybrids by randomly sampling an allele from the wild mallard and game-farm mallard gene pools across bi-allelic SNP positions; we randomly sampled each position based on a probability proportional to the allelic frequency in each respective gene pool. We then backcrossed 5 hybrids to the parental gene pool for 9 generations. Following, we ran 10 independent simulations. Simulations were then combined with the original reference set and inputted into ADMIXTURE to estimate assignment probabilities for a K of 2; which was the optimum K population value (see results). We ran 25 iterations for each of the 10 simulated datasets. Subsequently, we combined all 250 ADMIXTURE outputs and converted in PopHelper into CLUMPP input files. We employed the Large Greedy algorithm and 1,000 random permutations with final admixture proportions for each K and per sample assignment probabilities based on CLUMPP analyses of all 250 replicates per K . Per generation expected assignment probabilities were based on the average of all 10 (F1) or each of the 5 (F2–F10) backcrosses, along with each lower and upper limit.

Hydrogen stable isotope analysis

We sent feather samples for $\delta^2\text{H}$ analysis at the Cornell Stable Isotope Laboratory in Ithaca, NY. After washing in a 2:1 chloroform:methanol solution, feather samples were air dried under a fume hood and a subsample (0.35 mg) of vane material was loaded into silver capsules, crushed, and placed with internal lab standards into a desiccator for a minimum of three days prior to analysis. Samples were then loaded into a Zero Blank carousel under helium flow. Pyrolysis combustion on glassy carbon was at 1350°C in a Thermo Scientific Temperature Conversion Elemental Analyzer (TC/EA; Bremen, Germany) coupled via a Conflo IV (Thermo Scientific) to a Thermo Scientific Delta V Advantage isotope ratio mass spectrometer. Analysis of $\delta^2\text{H}$ was conducted using the comparative equilibration method of Wassenaar and Hobson [52] with 2 calibrated keratin reference materials (CBS, $\delta^2\text{H} = -197\text{‰}$; KHS, $\delta^2\text{H} = -54.1\text{‰}$; SPK, $\delta^2\text{H} = -121.6\text{‰}$) corrected for linear instrumental drift. Based on within-run analyses of a third keratin standard, measurement error was approximately $\pm 3\text{‰}$ for

hydrogen isotopes in feather ($\delta^2\text{H}_f$). All $\delta^2\text{H}$ values are reported relative to the Vienna Standard Mean Ocean Water–Standard Light Antarctic Precipitation scale.

Assignment to geographic origin using stable isotopes

To estimate natal origins of harvested birds, we used likelihood-based algorithms to produce spatially explicit assignments of source areas based on analysis of feather $\delta^2\text{H}$ data [53, 54]. Similar to prior studies [34, 55–57], we created a map (hereafter isoscape) by applying a rescaling function derived by regressing $\delta^2\text{H}_f$ of known-origin mallards against amount-weighted growing-season average precipitation ($\delta^2\text{H}_p$) [58] ($\delta^2\text{H}_f = -27.4 + 1.28 * \delta^2\text{H}_p$). We acquired the breeding range of mallards from BirdLife International [58], and the spatially explicit assignments were restricted (i.e., masked) to only those areas of the continent occurring in this range. Finally, mallards originating from outside the mid-continent and eastern populations made up < 1% of direct band recoveries (i.e., recoveries in the hunting season following banding) in Ottawa and Sandusky counties of Ohio, so we clipped the breeding range to remove the western breeding range (Fig 1). Additionally, to eliminate erroneous assignments, we clipped the breeding range along the 100th meridian because precipitation regimes in the western U.S. interfere with longitudinal $\delta^2\text{H}_p$ gradients. We estimated the likelihood that individual cells (i.e., pixels) within the isoscape represented a potential source area for a given sample by comparing the observed $\delta^2\text{H}_f$ against the isoscape predicted $\delta^2\text{H}_f$ using a normal probability density function [34, 53]. We then applied Bayes' Theorem to assess the posterior probability that an individual pixel within the isoscape was the putative source area of a given sample [34, 53]. After normalizing the probabilities to sum to one, we assigned individuals to source areas within the isoscape by selecting the raster cells that were consistent with the upper 67% of estimated probabilities of origin for each individual and coded those as one and all others as zero, consistent with 2:1 odds. We subsequently summed the results of the assignments over all individuals by addition of the surfaces to attain a final surface of the distribution in the number of individuals assigned to each pixel [34, 53].

In addition to assigning birds to their natal origins, we classified birds as locals versus immigrants using likelihood-based assignments [34]. We defined local birds as any individual whose $\delta^2\text{H}_f$ was consistent with the expected value for birds ($\delta^2\text{H}_{fd}$) originating in Ottawa and Sandusky counties, Ohio. Using the rescaled $\delta^2\text{H}_p$, we predicted the mean $\delta^2\text{H}_{fd}$ to be -82.5‰ to -84.5‰. We used a normal probability density function to compare the observed $\delta^2\text{H}_f$ for unknown-origin birds against the predicted mean $\delta^2\text{H}_{fd}$ to assess the likelihood that a given bird had grown its feathers in Ottawa or Sandusky counties of Ohio, given expected variation in $\delta^2\text{H}_f$ [34]. If $\delta^2\text{H}_f$ was within 12.8‰ (1 standard deviation) of the expected $\delta^2\text{H}_f$ for Ottawa and Sandusky counties of Ohio ($\delta^2\text{H}_{fd}$), we classified individual ducks as locals; others were classified as immigrants.

Weighting isotope results with band recoveries

Stable hydrogen isotopes (i.e., $\delta^2\text{H}_p$) produce a predictable latitudinal gradient but provide little information longitudinally [53]. Thus, we weighted isotope results by applying a band recovery distribution from the eastern and mid-continent populations. To do so, we used available band recovery data to inform weighting of source population origination to the mid-continent or eastern populations. We summarized available (through 14 July 2021) band recovery data of HY mallards from the U.S. Geological Survey Gamebirds database. The database was queried for all reported direct recoveries of HY mallards in Ottawa and Sandusky counties of Ohio, as well as for all banding stations and number of bands placed on HY mallards in the mid-continent and eastern population areas. The recovery location and their

corresponding banding location were displayed in a geographic information system (hereafter GIS) (ArcGIS 10.8, Environmental Systems Research Institute Inc., Redlands, California, USA). We selected all reported recoveries through time and used their associated banding location to assign origin to each sample to the mid-continent or eastern population. A polygon encompassing banding coordinates associated with direct recoveries served as the geographic area where harvested mallards potentially originated. To determine banding effort, we extracted all banding stations and summed the number of bands placed on HY mallards in the polygon where potential direct recoveries could have originated. We then calculated the proportion of band recoveries derived from the mid-continent and eastern populations and adjusted for banding effort by dividing the total number of recovered bands by the total number of bands placed on mallards that would be available for recovery. New rasters and an assignment to origin map were then created where the pixel values within the mid-continent and eastern populations represented the banding effort informed proportion of harvest. We created assignment to origin maps for each genetic assignment (i.e., pure wild, hybrid swarm, and filials) with sufficient sample sizes.

Statistical analyses for associating genetic assignment with geographical origin

We used linear regression models to determine if δ^2H_f was influenced by genetic assignment (pure wild, hybrid swarm $\times$ wild, and F3), sex, and harvest date (4 October = day 0). When genetic assignment was included in the final model, we used a Tukey Honest Difference test to determine differences among means between categories at $\alpha = 0.05$. We provided descriptive statistics for genetic group(s) with insufficient sample sizes to obtain statistical significance (see results).

Results

Sequencing and molecular data output

A total of 2,951 ddRAD-seq autosomal loci (692,460 base-pairs (bp)) were recovered across samples that included 2,751 autosomal (658,008 bp), 193 Z-sex chromosome linked (33,509 bp), and 7 W-sex chromosome linked (943 bp) ddRAD-seq loci that met our criteria for sequencing coverage and missing data. We attained an average depth of 118 sequences/locus/sample (depth range 18–163 sequences/locus/sample). Moreover, we confirmed sex identity across samples by plotting sequencing depth ratios between autosomal and recovered W- and Z-sex chromosome linked ddRAD-seq loci (S1 Fig). Finally, we attained 593 base-pairs of the mtDNA control region across all samples.

Nuclear simulations and population structure

A total of 22,136 (of 22,139) independent bi-allelic ddRAD-seq SNPs were retained for nuclear population structure analyses. Plotting the first two principal components explained a total of 22.66% of the variance (S2A Fig), and which separated reference wild and game-farm mallards with Ohio samples scattered from the wild reference mallard set towards game-farm mallards (Fig 2A). Plotting ADMIXTURE assignment probabilities under the optimum K population of two (S2B Fig) distinguished between wild and game-farm mallards, while Ohio samples possessed assignment probabilities to either the wild mallard genetic cluster only or had assignments to both clusters (Fig 2B).

Formal assignment of samples of admixed ancestry were based on hybrid filial class cut-offs established through our simulations that were run under a K population model of two in

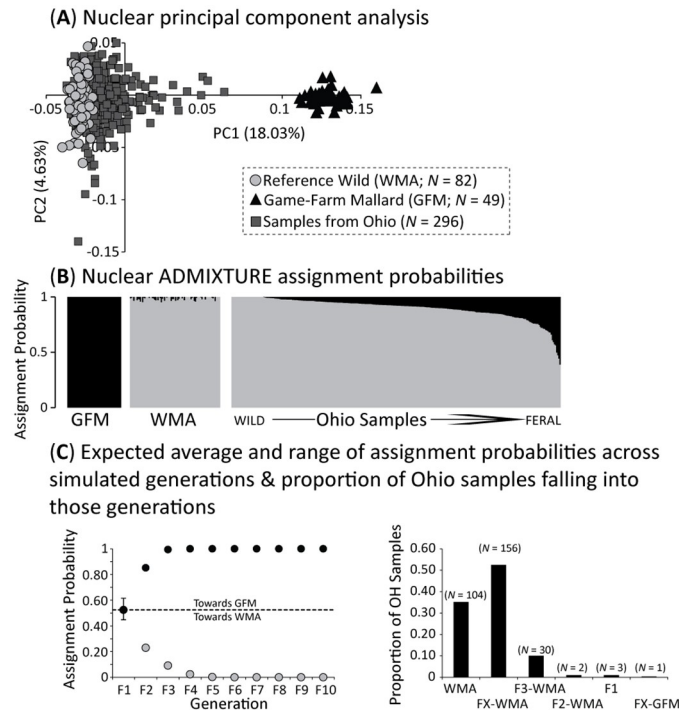


Fig 2. Population structure analyses based on 22,136 independent bi-allelic ddRAD-seq SNPs assayed across Ohio (OH) samples, as well as reference wild (WMA) and game-farm (GFM) mallards, including (A) a plot of the first two principal components of the principal components analysis (PCA) and (B) ADMIXTURE assignment probabilities attained under an optimum K populations of 2 (S2B Fig). Finally, we provide the (C) expected average and range of assignment probabilities across simulated generations (left), as well as the proportion of Ohio samples falling into those generations (right).

<https://doi.org/10.1371/journal.pone.0282874.g002>

ADMIXTURE (Fig 2C). First, we designated any sample within the assignment probabilities range recovered among reference wild mallard set as pure wild mallards, which included any sample with $\leq 5\%$ alternative genetic cluster assignment. For the remaining admixed samples, we were able to reliably identify between F1 hybrids, as well as F2 and F3 backcrossed generations (S2 Table). Note that any sample with mixed ancestry assignments intermediate to those recovered in simulated F1–F3 generations represented backcrosses of unknown age (e.g., FX hybrid $\times$ hybrid crosses, parental hybridization switching between generations; [18]; and were designated as either backcrossed with more wild (i.e., FX-WMA) or game-farm (i.e., FX-GFM) mallard, respectively (Fig 2C). Categorization of samples based on simulations recovered 35% ($n = 104$) of Ohio samples as pure wild mallard, while the majority (53%; $n = 156$) being within the FX-WMA category (Fig 2C). A total of three (1%), two (0.7%), and 30 (10%) of Ohio samples were designated as F1, F2-WMA, and F3-WMA, respectively (S1 Table). No Ohio sample was found to be game-farm mallard (i.e., $\geq 95\%$ to the game-farm mallard genetic cluster), but one sample was designated as a hybrid backcross towards game-farm mallard (i.e., FX-GFM; Fig 2C; S1 Table).

Stable isotope results for local vs. migrant

We analyzed 296 HY mallards of unknown origin and classified 24% ($n = 71$) of individuals as local whose $\delta^2\text{H}_f$ was consistent with feather growth in Ottawa or Sandusky counties of Ohio, and 76% ($n = 225$) of individuals as immigrants whose $\delta^2\text{H}_f$ was consistent with feather growth outside these counties.

Relationship of stable isotopes and genetics

We detected that $\delta^2\text{H}_f$ values were explained by genetic assignment ($F_{2, 286} = 8.80$, $P < 0.01$) and harvest date ($F_{1, 286} = 8.33$, $P < 0.01$) (Table 1), but not sex ($P = 0.25$). Model predicted $\delta^2\text{H}_f$ changed by -16.6‰ from 4 October to 19 December; approximately the northern distance from our northwestern Ohio study area to Port Austin, MI, ~ 288 km or 2.5° latitude. Among genetic assignments, we detected that pure wild mallards ($n = 104$) had more negative (i.e., more northern) $\delta^2\text{H}_f$ values ($P < 0.01$; -121.05 ± 2.43 [SE]) than hybrid swarm mallards (-108.23 ± 1.99 ; $n = 156$), but F3 (-113.33 ± 4.53 ; $n = 30$) did not differ from pure wild or hybrid swarm mallards ($P \geq 0.29$; Table 1, Fig 3). Values for $\delta^2\text{H}_f$ from genetic groups with insufficient samples sizes for statistical testing included, F1 ($n = 3$, mean = -105.86 ± 2.80 , min = -110.76 , max = -101.06), F2 ($n = 2$, min = -111.64 , max = -106.91), and hybrid $\times$ game-farm ($n = 1$, -99.46).

Discussion

Studies of mallard genetics have become increasingly common because of concern about introgression of potentially maladaptive genes into the wild population from breeding with released game-farm mallards [17, 19]. Most releases of game-farm mallards in North America have occurred on the United States portion of the Atlantic coast (USFWS 2003). Across broad regions, estimates are that $\sim 92\%$ of mallards in Atlantic coast states and $\sim 40\%$ in the mid-continent had substantial ($> 10\%$ assignment probabilities) game-farm mallard derived ancestry in 2010 [18]. In more western locales of North America only 3% of mallards possessed substantial game-farm mallard ancestry [18]. Similarly, at southern latitudes in the mid-continent only $\sim 4\%$ of samples were game-farm hybrids in 2021 [21]. We detected that 65% of sampled mallards had substantive amounts of game-farm genetic ancestry at our Ohio study area. We posit that the greater hybrid prevalence in Ohio was from movement of eastern mallards into the Great Lakes region on their way to more southern locales. Differences between our results and those of Davis et al. [21] suggest differences in migration and wintering areas for mid-continent and eastern population mallards. Moreover, hybrids reported by Davis et al. [21] in southern parts of the mid-continent were late backcrosses as compared to the wide variance of hybrid ancestry detected in the current study (Fig 2B). We suspect this regional variation in percentage of mallards with game-farm genes exists because those with greater percentage of game-farm ancestry are less likely to migrate to traditional wintering areas at southern latitudes of North America because they lack this innate behavior. Combined, results provide evidence of meta-population dynamics in the mid-continent that will require greater efforts coupling genetics and movement ecology studies to understand variation in migration patterns among mallard genotypes.

Table 1. Parameter estimates $\pm$ standard error (SE) and 95% confidence limits (CL) of factors affecting $\delta^2\text{H}_f$ of hatch-year mallards (*Anas platyrhynchos*) harvested in northwestern Ohio, 4 October–19 December 2019.

Parameter _{df}	$\beta \pm \text{SE}$	95% CL
Intercept	-113.26 ± 3.61	-120.35 to -106.16
Date _{1, 287}	-0.22 ± 0.07	-0.36 to -0.07
Genomics _{2, 287}		
F3	7.72 ± 5.15	-2.41 to 17.85
Hybrid swarm mallard	12.82 ± 3.14	6.64 to 19.01
Pure wild mallard	0.00	0.00

<https://doi.org/10.1371/journal.pone.0282874.t001>

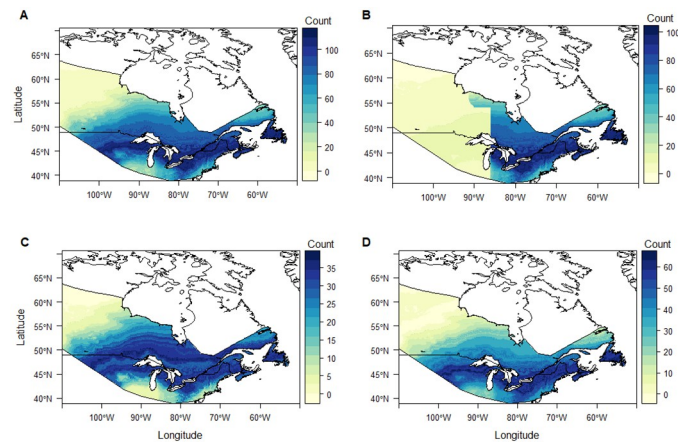


Fig 3. Geographic depiction of probable origins of hatch-year mallards harvested in Ottawa and Sandusky counties, Ohio, USA, from October to December 2019. Legend numbers correspond to the number of individuals assigned to each pixel based on a 2:1 odds ratio. A) all samples in our study without application of a banding data informed weighted spatial mask, B) origins constrained by a banding data informed weighted spatial mask, C) Origins of pure wild mallards without application of a banding data informed weighted spatial mask, and D) Origins of hybrid-swarm mallards without application of a banding data informed weighted spatial mask. U.S. and Canadian political boundaries were obtained through the R package ‘rnatrualearth’. Atlantic and Mississippi flyway shapefiles were obtained from the metadata of Weltzin et al. 2018 (<https://www.sciencebase.gov/catalog/item/5b75d10ee4b0f5d5787fea6f>). Atlantic and Mississippi flyway boundaries and data were used from Weltzin et al. 2018 (<https://www.sciencebase.gov/catalog/item/5b75d10ee4b0f5d5787fea6f>). The calibration equation used to produce the assignment to origin was obtained from van Dijk et al. 2014. Except for the shapefiles listed above, no external data was used for the creation of the figure. Figures were created using the R statistical framework (version 4.0.2, R Core Team 2018).

<https://doi.org/10.1371/journal.pone.0282874.g003>

Nearly all possible genetic combinations of mallards were obtained at our study area. Most of our samples were assigned as hybrid swarm (53%) or pure wild (35%), but we also obtained three F1 hybrids (pure game-farm × pure wild cross), which can only result from annual interbreeding between game-farm and wild mallards. We conclude that continued game-farm introgression into the wild mallard population is not isolated to the eastern population of mallards in North America, and may be increasing and more widespread than previously reported. More intensive and geographically broad sampling of the mid-continent region will be required to better understand whether the apparent ongoing interbreeding between game-farm and wild mallards is due to local release efforts or immigrants from the eastern population of mallards. Additionally, a thorough survey of the number and locations of game-farm releases throughout North America would help refine our understanding of the spatial potential of game-farm mallard introgression.

Pure wild mallards in our sample had $\delta^2\text{H}_f$ values that indicated they geographically originated from farther north than hybrid swarm mallards. Stable hydrogen isotopes (i.e., $\delta^2\text{H}_p$) produce a predictable latitudinal gradient, but provide little information longitudinally [59]. We used weighting of the isotope results by applying a band recovery distribution from the eastern and mid-continent populations, which weighted a greater portion of our sample longitudinally towards the eastern population. However, 17% of our samples had $\delta^2\text{H}_f < -140.0$, which could only have originated from the mid-continent population (i.e., prairie, parkland, or boreal) or from areas north of the applied mallard range in Quebec. As such our weighted origin maps should be interpreted with caution because they can introduce error from the uneven spatial distribution or lack of banding effort in regions where mallards originated. Further, although not significant, mean $\delta^2\text{H}_f$ from F1 through F3 hybrids in our study followed a predictable pattern consistent with individuals possessing greater levels of game-farm genetic

assignment originating at more southern latitudes (i.e., F1) and future hybrid filial generations occurring in more northern locales (i.e., F2 through F3). For mallards with $\delta^2\text{H}_f < -140.0$, most were wild (55%), 35% were hybrid swarm, and 10% were F3 which is consistent with introgression of game-farm genetic variation moving north and west in a generational pattern consistent with breeding ecology of dabbling ducks whereby males follow females to breeding areas after pairing during the non-breeding period [9]. We speculate that we did not detect differences in $\delta^2\text{H}_f$ between hybrid filials and other genetic categories because of small sample sizes. Our results suggest game-farm mallard introgression into the mid-continent population of mallards in the prairie, parkland, and boreal regions. This could be cause for concern because other areas with widespread introgression of game-farm genes have seen declining mallard breeding abundance [22, 23].

Mallards are considered facultative migrants [9], whereby they typically migrate when weather or deteriorating environmental conditions make them move during autumn migration [60, 61]. Consistent with this pattern, we detected that mallards originating from more northern latitudes were harvested later in the duck hunting season in northwestern Ohio. Further, our result that 76% of harvested mallards in northwestern Ohio were classified as immigrants is similar to findings by Palumbo et al. [34] from Lake St. Clair, Ontario, approximately 100km north of our study area. These results contrast with a sample of 1,254 HY mallards from throughout the Atlantic coast states that indicated that origin, as determined by $\delta^2\text{H}_f$, was unrelated to timing of harvest [62, 63]. Differences between our northwestern Ohio study site and those from farther east could be related to differences in genetic composition of mallards between these regions, whereby seasonal movement patterns differ because eastern population mallards have greater genetic input from game-farm mallards. Wild mallards may be adapted to migrate when weather conditions deteriorate and move facultatively towards ancestral wintering grounds [9, 60]. In contrast, mallards bred through artificial selection on game-farms may not have a similar phenotypic response to migrate, which may produce less predictable seasonal movements [64]. Early studies during a time of large releases of game-farm mallards suggest failure of these ducks to migrate similar to wild congeners [65], but the hypothesis that percentage of game-farm genes influences migratory behavior needs further testing. Knowing the genetics of a banded sample of mallards or tracked them by telemetry would aid in our understanding of migratory movements of different genotypes of mallards.

We found that a surprisingly high proportion of HY mallards harvested in northwestern Ohio had some form of hybridization with game-farm mallards and that pure wild mallards originated from areas farther north (and also likely west) compared to hybrid, admixed individuals. Our findings add to growing concern that continued and possibly increasing introgression of game-farm mallard genes into the wild mallard breeding population may be causing population declines in some regions, particularly in the U.S. portion of eastern North America. We advocate for expanded use of stable feather isotopes in combination with genomic methods, as we have done here, to aid in disentangling the extent of the apparently negative influence of game-farm genes on the health of the North American mallard population. Specifically, focusing these efforts in the mid-continent region should be priority because of the apparent western movement of game-farm genetic variation into this mallard population that we detected. Together, coordinated efforts to collect stable isotope data, genetics, and movements of mallards would provide strong inference capacity about the distribution and effect on migratory movement of game-farm mallard genetic variation across North America.

Supporting information

S1 Table. Raw genetic, isotope, and additional metadata from hatch-year mallards (*Anas platyrhynchos*) harvested in northwestern Ohio, 4 October–19 December 2019.
(XLSX)

S2 Table. Simulated expected average, minimum, and maximum assignment probability values to wild mallard (WMA) genetic cluster based on ddRAD-seq autosomal bi-allelic SNPs for 10 hybrid classes determined between wild and game-farm (GFM) mallards. The parental backcross for each generation is provided, and those generations that are genetically distinguishable denoted.
(PNG)

S1 Fig. Plot of Z- or W-Sex chromosome versus autosomal (A) sequencing depth to identify sex across samples.
(TIF)

S2 Fig. (A) Proportion of variation explained by each principal components of the principal components analysis (PCA; Fig 1A). (B) CV-Error values averaged across the 100 ADMIXTURE analysis replicates for each of the evaluated *K* population values of 1–5; the optimum *K* population of two is denoted.
(TIF)

Acknowledgments

We thank the private duck hunting clubs in Sandusky and Ottawa counties, Ohio, that contributed harvested mallards for sampling and J. Schmit for processing collected mallards. Lab assistance was provided by L. Wallace, T. Hodges, R. Stedman, E. Breault, and M. Sturtz. The Scientific Advisory Committee of the Long Point Waterfowl and Wetlands Research Program of Birds Canada and two anonymous reviewers provided comments to improve the paper.

Author Contributions

Conceptualization: Michael L. Schummer, John Simpson, Philip Lavretsky.

Data curation: Michael L. Schummer, John Simpson, Brendan Shirkey.

Formal analysis: Michael L. Schummer, Philip Lavretsky.

Funding acquisition: Michael L. Schummer, John Simpson, Philip Lavretsky, Douglas C. Tozer.

Investigation: Michael L. Schummer, Philip Lavretsky.

Methodology: Michael L. Schummer, John Simpson, Philip Lavretsky.

Project administration: Michael L. Schummer, Brendan Shirkey, Philip Lavretsky.

Resources: Michael L. Schummer, John Simpson, Philip Lavretsky.

Supervision: John Simpson, Philip Lavretsky.

Validation: Michael L. Schummer.

Visualization: Michael L. Schummer, Samuel R. Kucia.

Writing – original draft: Michael L. Schummer, John Simpson, Brendan Shirkey, Philip Lavretsky.

Writing – review & editing: Michael L. Schummer, John Simpson, Brendan Shirkey, Philip Lavretsky, Douglas C. Tozer.

References

1. Kiple KF. The Cambridge world history of food. Cambridge University Press. 2001 Nov 22.
2. Huang Y, Li Y, Burt DW, Chen H, Zhang Y, Qian W, et al. The duck genome and transcriptome provide insight into an avian influenza virus reservoir species. *Nat Genet.* 2013; 45: 776–783. <https://doi.org/10.1038/ng.2657> PMID: 23749191
3. Mock KE, Latch EK, Rhodes OE. Assessing losses of genetic diversity due to translocation: long-term case histories in Merriam's turkey (*Meleagris gallopavo merriami*). *Conserv Genet.* 2004; 5: 631–645.
4. Reynolds M and Klavitter J. Translocation of wild Laysan duck to establish a population at Midway Atoll National Wildlife Refuge, United States and US Pacific Flyway. *Conserv Evid.* 2006; 3: 6–8.
5. Rodríguez-Clark KM, Maldonado JE, Ascanio D, Gamero E, Ovalle Moleiro L, Pérez-Emán J, et al. Using genetics to understand and conserve the Red Siskin (*Carduelis cucullata*). *J of the Nat Finch and Softbill Soc.* 2011; 1:25–30.
6. Gille DA, Buchalski MR, Conrad D, Rubin ES, Munig A, Wakeling BF, et al. Genetic outcomes of translocation of bighorn sheep in Arizona. *J Wildlife Manage.* 2019; 83: 838–854.
7. Hlevinski RA, Malecki RA, and Curtis PD. Can hunting of translocated nuisance Canada geese reduce local conflicts? *Wildlife Soc B.* 2010; 34: 845–849.
8. Weng GJ. Ecology and population genetics of Mottled Ducks within the south Atlantic coastal zone (Doctoral dissertation, University of Georgia). 2006.
9. Baldassarre G. Ducks, geese, and swans of North America. JHU Press; 2014 Nov 15.
10. 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. *J Nat Conserv.* 2012; 20: 231–241.
11. Rueness EK, Cadahia-Lorenzo L, Hoel K, Söderquist P, Velle G, Asmyr MG, et al. Assessment of the risks associated with the import and release of hand-reared mallards for hunting purposes. Opinion of the Panel on Alien Organisms and Trade in Endangered Species (CITES) of the Norwegian Scientific Committee for Food Safety, VKM Report 2017. 2017; 23.
12. Lincoln FC. Restocking of marshes with hand-reared mallards not proved practicable. *Yearb Agric.* 1934; 310–313
13. Heusmann HW. The history and status of the mallard in the Atlantic Flyway. *Wildlife Soc B.* 1991; 19: 14–22.
14. U.S. Fish and Wildlife Service (USFWS). Review of captive-reared mallard regulations on shooting preserves, final draft. U.S. Fish and Wildlife Service, Division of Migratory Bird Management, Washington D.C.; 2003.
15. Osborne CE., Swift BL, Baldassarre GA. Fate of captive-reared and released mallards on eastern Long Island, New York. *Hum-Wildl Interact.* 2010; 4: 266–274.
16. Kraus RH, Zeddeman A, van Hooft P, Sartakov D, Soloviev SA, Ydenberg RC, et al. Evolution and connectivity in the world-wide migration system of the mallard: Inferences from mitochondrial DNA. *BMC genetics.* 2011 Dec; 12: 1–5.
17. Söderquist P, Elmberg J, Gunnarsson G, Thulin CG, Champagnon J, Guillemain M, et al. Admixture between released and wild game birds: a changing genetic landscape in European mallards (*Anas platyrhynchos*). *Eur J Wildl Res.* 2017; 63:1–13.4
18. Lavretsky P, DaCosta JM, Sorenson MD, McCracken KG, Peters JL. ddRAD-seq data reveal significant genome-wide population structure and divergent genomic regions that distinguish the mallard and close relatives in North America. *Mol Ecol.* 2019; 28: 2594–2609. <https://doi.org/10.1111/mec.15091> PMID: 30941840
19. Lavretsky P, McInerney NR, Mohl J, Brown JI, James H, McCracken KG, et al. Assessing changes in genomic divergence following a century of human mediated secondary contact among wild and captive-bred ducks. *Mol Ecol.* 2020; 29: 578–595. <https://doi.org/10.1111/mec.15343> PMID: 31872482
20. Ohio Division of Wildlife. Game Research in Ohio. Vol 3. July, 1965. Ohio Department of Natural Resources, Columbus, OH. 1965. Kenneth W. Laub, Editor.
21. Davis JB, Outlaw DC, Ringelman KM, Kaminski RM, Lavretsky P. Low levels of hybridization between domestic and wild Mallards wintering in the lower Mississippi Flyway. *Ornithology.* 2022 Oct 6; 139: ukac034.

22. Alisauskas RT, Arnold TW, Leafloor JO, Otis DL, Sedinger JS. Lincoln estimates of mallard (*Anas platyrhynchos*) abundance in North America. *Ecol Evol*. 2014; 4: 132–143.
23. U.S. Fish and Wildlife Service (USFWS). Adaptive Harvest Management: 2018 Hunting Season. U.S. Department of Interior, Washington, D.C. 2017.
24. U.S. Fish and Wildlife Service (USFWS). Adaptive Harvest Management: 2021 Hunting Season. U.S. Department of Interior, Washington, D.C. 2020.
25. Munro RE, Kimball CF. Population ecology of the mallard. VII: Distribution and derivation of the harvest. U.S. Fish and Wildlife Service Resource Publication No. 147. 1982.
26. Arnold TW, de Sobrino CN. Derivation of Eastern Population Mallard Harvests, 1995–2009. Final Report to the Canadian Wildlife Service, University of Minnesota, Minneapolis, MN. 2012.
27. Zuverink DA. Changes in the derivation of mallard harvests from the northern US and Canada, 1966–1998 (Doctoral dissertation, The Ohio State University). 2001.
28. Hobson KA. Tracing origins and migration of wildlife using stable isotopes: a review. *Oecologia*. 1999; 120: 314–326. <https://doi.org/10.1007/s004420050865> PMID: 28308009
29. Bowen GJ, Wassenaar LI, Hobson KA. Global application of stable hydrogen and oxygen isotopes to wildlife forensics. *Oecologia*. 2005; 143: 337–348. <https://doi.org/10.1007/s00442-004-1813-y> PMID: 15726429
30. Hobson KA, Van Wilgenburg SL, Wassenaar LI, Larson K. Linking hydrogen isotopes in feathers and precipitation: sources of variance and consequences for assignment to isoscapes. *Plos One*. 2012; 7: e35137.
31. Bowen GJ, West JB. Isoscapes for terrestrial migration. In: editors Hobson KA, Wassenaar LI. Tracking animal migration with stable isotopes. Academic Press, London, United Kingdom. 2019. pp 53–84.
32. Arnold TW, de Sobrino CN. Derivation of Eastern Population Mallard Harvests, 1995–2009. Final Report to the Canadian Wildlife Service, University of Minnesota, Minneapolis, MN. 2012.
33. Hobson KA. Applying isotopic methods to tracking animal movements. *Ter Ecol*. 2008; 2: 45–78.
34. Palumbo M, Tozer D, Hobson K. Origins of harvested Mallards from Lake St. Clair, Ontario: a stable isotope approach. *Avian Conserv Ecol*. 2019; 14: 3.
35. Carney SM. Species, age and sex identification of ducks using wing plumage. U. S. Department of the Interior, U.S. Fish and Wildlife Service, Washington, D.C. Jamestown, ND: Northern Prairie Wildlife Research Center. 1992.
36. Lavretsky P, Dacosta JM, Hernández-Baños BE, Engilis A, Sorenson MD, Peters JL. Speciation genomics and a role for the Z chromosome in the early stages of divergence between Mexican ducks and mallards. *Mol Ecol*. 2015; 24: 5364–5378. <https://doi.org/10.1111/mec.13402> PMID: 26414437
37. Hernández F, Brown JI, Kaminski M, Harvey MG, Lavretsky P. Genomic evidence for rare hybridization and large demographic changes in the evolutionary histories of four North American dove species. *Animals*. 2021; 11: 2677. <https://doi.org/10.3390/ani11092677> PMID: 34573643
38. DaCosta JM, Sorenson MD. Amplification biases and consistent recovery of loci in a double-digest RAD-seq protocol. *PLoS One*. 2014; 9: e106713. <https://doi.org/10.1371/journal.pone.0106713> PMID: 25188270
39. Lavretsky P, Janzen T, McCracken KG. Identifying hybrids & the genomics of hybridization: Mallards & American black ducks of eastern North America. *Ecol Evol*. 2019; 9: 3470–3490.
40. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014; 30: 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170> PMID: 24695404
41. Li H, Durbin R. Inference of human population history from individual whole-genome sequences. *Nature*. 2011; 475: 493. <https://doi.org/10.1038/nature10231> PMID: 21753753
42. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*. 2009; 25: 1754–1760. <https://doi.org/10.1093/bioinformatics/btp324> PMID: 19451168
43. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and VCFtools. *Bioinformatics*. 2011; 27: 2156–2158. <https://doi.org/10.1093/bioinformatics/btr330> PMID: 21653522
44. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007; 81: 559–575. <https://doi.org/10.1086/519795> PMID: 17701901
45. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res*. 2009; 19: 1655–1664. <https://doi.org/10.1101/gr.094052.109> PMID: 19648217
46. Alexander DH, Lange K. Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics*. 2011; 12: 1–6.

47. Alexander, D. H., J. Novembre, K. Lange. Admixture 1.22 software manual. 2012.
48. Zhou H, Alexander D, Lange K. A quasi-Newton acceleration for high dimensional optimization algorithms. *Stat Comput.* 2011; 21: 261–273. <https://doi.org/10.1007/s11222-009-9166-3> PMID: [21359052](#)
49. Francis RM. Pophelper: An R package and web app to analyze and visualize population structure. *Mol Ecol Res.* 2016; 17: 27–32.
50. Jakobsson M, Rosenberg NA. CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics.* 2007; 23: 1801–1806. <https://doi.org/10.1093/bioinformatics/btm233> PMID: [17485429](#)
51. Lavretsky P, Peters JL, Winker K, Bahn V, Kulikova I, Zhuravlev YN, et al. Becoming pure: Identifying generational classes of admixed individuals within lesser and greater scaup populations. *Mol Ecol.* 2016; 25: 661–674. <https://doi.org/10.1111/mec.13487> PMID: [26833858](#)
52. Wassenaar LI, Hobson KA. Comparative equilibration and online technique for determination of non-exchangeable hydrogen of keratins for use in animal migration studies. *Isot Environ Healt S.* 2003; 39: 211–217. <https://doi.org/10.1080/1025601031000096781> PMID: [14521282](#)
53. Hobson KA, Wunder MB, Van Wilgenburg SL, Clark RG, Wassenaar LI. A method for investigating population declines of migratory birds using stable isotopes: origins of harvested lesser scaup in North America. *PLoS One.* 2009; 4: e7915. <https://doi.org/10.1371/journal.pone.0007915> PMID: [19946360](#)
54. Wunder MB. Using isoscapes to model probability surfaces for determining geographic origins. In *Iso-scapes 2010* (pp. 251–270). Springer, Dordrecht.
55. Ashley P, Hobson K, Van Wilgenburg S, North N, Petrie S. Linking Canadian harvested juvenile American Black Ducks to their natal areas using stable isotope (δD , $\delta^{13}C$, and $\delta^{15}N$) methods. *Avian Conserv Ecol.* 2010; 2: 5.
56. Guillemain G, Van Wilgenburg SL, Legagneux P, Hobson KA. Assessing geographic origins of Teal through stable-hydrogen isotope analyses of feathers and ring-recoveries. *J of Ornithol.* 2014; 155: 165–172.
57. Asante CK, Jardine TD, Van Wilgenburg SL, Hobson KA. Tracing origins of waterfowl using the Saskatchewan River Delta: Incorporating stable isotope approaches in continent-wide waterfowl management and conservation. *Condor.* 2017; 119: 261–274.
58. BirdLife International and Handbook of the Birds of the World. Bird species distribution maps of the world. Version 2019.1. 2019. <<http://datazone.birdlife.org/species/requestdis>>. Accessed 27 Sep 2020.
59. Hobson K, Wassenaar L. Tracking animal migration with stable isotopes. Second edition. Academic, London, UK. 2019.
60. Schummer ML, Kaminski RM, Raedeke AH, Graber DA. Weather-related indices of autumn–winter dabbling duck abundance in middle North America. *J Wildlife Manage.* 2010; 74: 94–101.
61. Newton I. The migration ecology of birds. Elsevier; 2010 Aug 4.
62. Kucia SM. Natal origins and breeding habitat associations of the eastern mallard population. MS thesis. State University of New York College of Environmental Science and Forestry, Syracuse, NY. 2021.
63. Kucia SM, Schummer ML, Kusack JW, Hobson KA, Nicolai CA. Natal origins of mallards harvested in the Atlantic flyway of North America: Implications for conservation and management. *Avian Cons Eco.* In press.2023.
64. Errington PL. Banding studies of semidomesticated mallard ducks. *Bird Banding.* 1936; 7: 69–73.
65. Brakhage GK. Migration and mortality of ducks hand-reared and wild-trapped at Delta, Manitoba. *J Wildlife Manage.* 1953; 17: 465–477.