Taylor & Francis
Taylor & Francis Group

Ostrich

Journal of African Ornithology

ISSN: 0030-6525 (Print) 1727-947X (Online) Journal homepage: www.tandfonline.com/journals/tost20

Evolution, population structure and morphology of the African Black Duck *Anas sparsa* and Yellow-billed Duck *A. undulata*

Philip Lavretsky, Ramsey Russell, Sara Gonzalez, Vergie M Musni, Alexis Díaz & Joshua I Brown

To cite this article: Philip Lavretsky, Ramsey Russell, Sara Gonzalez, Vergie M Musni, Alexis Díaz & Joshua I Brown (19 Jun 2025): Evolution, population structure and morphology of the African Black Duck *Anas sparsa* and Yellow-billed Duck *A. undulata*, Ostrich, DOI: 10.2989/00306525.2025.2509215

To link to this article: https://doi.org/10.2989/00306525.2025.2509215

View supplementary material

Published online: 19 Jun 2025.

Submit your article to this journal

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tost20

Ostrich 2025, 96(2): xxx–xxx
Printed in South Africa — All rights reserved

This is the final version of the article that is published ahead of the print and online issue

Copyright © 2025 NISC (Pty) Ltd
OSTRICH
ISSN 0030–6525 EISSN 1727-947X
https://doi.org/10.2989/00306525.2025.2509215

Evolution, population structure and morphology of the African Black Duck *Anas sparsa* and Yellow-billed Duck *A. undulata*

Philip Lavretsky[1]*, Ramsey Russell[2], Sara Gonzalez[1], Vergie M Musni[1], Alexis Díaz[3] and Joshua I Brown[1,4]

[1] *Department of Biological Sciences, University of Texas at El Paso, El Paso, Texas, United States*
[2] *GetDucks.com, Brandon, Mississippi, United States*
[3] *Centro de Ornitología y Biodiversidad (CORBIDI), Lima, Peru*
[4] *Department of Life, Earth, and Environmental Sciences, West Texas A&M University, Canyon, Texas, United States*
* *Correspondence: plavretsky@utep.edu*

Biological conservation requires a fundamental understanding of evolutionary history and established contemporary population genetics. Here, we sequenced mitochondrial DNA (mtDNA) and thousands of nuclear loci across individuals of the African Black Duck *Anas sparsa* and Yellow-billed Duck *A. undulata* to understand the evolutionary histories and to establish their current population structure. In addition to testing for possible hybridisation between these two African species, we compared the genetic ancestries with known wild and domestic Mallards *A. platyrhynchos* to understand whether the presence of that species in Africa is resulting in elevated interspecific hybridisation. Finally, we assessed morphological variation within the two African species; although the sample sizes limited inferences for the African Black Duck, we were able to demarcate trait cut-offs for field identification of the sexes of the Yellow-billed Duck. We recovered strong population structure between the two species and the Mallard for both mtDNA and nuclear loci. Whereas we recovered high levels of co-ancestry and low levels of nucleotide diversity among the African Black Duck samples, our demographic analysis estimated a contemporary effective population size of ~2.5 million, which was equivalent to the estimate for the Yellow-billed Duck. We found that the increase in effective population size of both African species coincides with the onset of the last glacial cycle, with numbers peaking during the last glacial maximum. We recovered two genetic hybrids from the samples: a single hybrid between Yellow-billed Duck and wild Mallard, and the first genetically vetted Yellow-billed Duck × African Black Duck hybrid. Our study not only sheds light on the current population structures and standing genetic variation, but outlines methodologies to build morphological cut-offs for species and sex identification, which could be applied in future conservation efforts for Yellow-billed Ducks and African Black Ducks.

Évolution, structure des populations et morphologie des Canards noirâtres *Anas sparsa* et Canards à bec jaune *A. undulata*

La conservation des espèces nécessite une compréhension fondamentale de l'histoire de l'évolution et de la génétique des populations contemporaines. Ici, nous avons séquencé l'ADN mitochondrial et des milliers de loci nucléaires afin de comprendre l'histoire évolutive et d'établir la structure actuelle des populations du Canard noirâtre *Anas sparsa* et du Canard à bec jaune *A. undulata*. En plus de tester l'hybridation possible entre ces deux espèces africaines, nous avons comparé les ascendances génétiques avec des Canards colverts *A. platyrhynchos* sauvages et domestiques pour comprendre si leur présence entraîne une hybridation interspécifique élevée. Enfin, nous avons évalué la variation morphologique au sein des deux canards africains et, bien que la taille des échantillons ait limité les conclusions pour le Canard noirâtre, nous avons été en mesure de définir des seuils de traits pour l'identification sur le terrain des sexes du Canard à bec jaune. Nous avons constaté une forte structure populationnelle entre les trois espèces, au niveau de tous les marqueurs génétiques. Alors que nous avons trouvé des faibles niveaux de diversité nucléotidique et un haut niveau de parenté parmi les échantillons de Canards noirâtres, les analyses démographiques ont estimé une taille de population effective contemporaine de ~2.5 millions, équivalente à l'estimation pour les Canards à bec jaune. Nous avons constaté que l'augmentation de la taille effective de la population des deux espèces africaines coïncide avec le début du dernier cycle glaciaire, le nombre d'individus atteignant son maximum lors du dernier maximum glaciaire. Nous avons trouvé deux hybrides génétiques: un hybride entre un Canard à bec jaune et un Canard colvert sauvage, et le premier hybride documenté entre un Canard à bec jaune et un Canard noirâtre. Notre étude sur la structure actuelle et la variation génétique des populations a aussi fourni des méthodologies pour établir des seuils morphologiques pour l'identification des espèces et du sexe, qui peuvent être appliqués dans des futurs efforts de conservation des Canards à bec jaune et des Canards noirâtres.

Keywords: co-ancestry, hybridisation, introgression, median-joining network, morphometrics, nucleotide diversity, population genetics, South Africa

Supplementary material: available at https://doi.org/10.2989/00306525.2025.2509215

Ostrich is co-published by NISC (Pty) Ltd and Informa UK Limited (trading as Taylor & Francis Group)

Published online 19 Jun 2025

Introduction

Understanding evolutionary histories and establishing contemporary genetic uniqueness of taxa is fundamental to biodiversity conservation (Gaitán-Espitia and Hobday 2021; Hohenlohe et al. 2021; Nielsen et al. 2023). Insufficient understanding of these aspects of species often results in biased and poorly executed conservation actions (Hoelzel et al. 2019; Holderegger et al. 2019). Fortunately, advances in molecular methods over the last decades have made it possible to access genomes of non-model organisms to conduct population and landscape genetics analyses, obtaining fine-scale understandings of any organism (Leipold et al. 2020). In particular, as extinction rates continue to climb in the Anthropocene (Corlett 2015), such data is increasingly required not only to catalogue standing genetic variation, but also to understand how species are responding to ever-changing landscapes (Caro et al. 2014). In addition to direct consequences from human-induced ecological changes that can result in complete or local extinction, as well as range shifts, such events are also resulting in secondary contact between many once allopatric species, raising the frequency and incidence of anthropogenic hybridisation (Lavretsky et al. 2023a, 2024). Here, we coupled genomic and morphological data to understand the evolutionary histories and establish the current population structure between two sister species of African dabbling ducks (subfamily Anatinae) in the genus *Anas*; furthermore, we also assessed morphological variations in an attempt to develop characters for field identification.

The African Black Duck *Anas sparsa* and Yellow-billed Duck *A. undulata* are two of 14 species of waterfowl comprising the Mallard *A. platyrhynchos* species complex that radiated around the world over the last 2 million years (Lavretsky et al. 2014a). In fact, the complexes inception originates in Africa from one of these two species or an unknown ancestor (Palmer 1976; Johnson and Sorenson 1999; Lavretsky et al. 2014a). Whereas, both species are phenotypically distinguishable, each is considered sexually monomorphic. Although the sexes differ in size, other diagnostic morphological features that can be used in the field have not been studied (Day 1977; Frost et al. 1979). Generally, both species are considered nonmigratory but opportunistically nomadic (Johnsgard 1978). Whereas the Yellow-billed Duck prefers most types of freshwater aquatic habitats (e.g. flooded grasslands, permanent and seasonal lakes, streams, marshes, brackish coastal lagoons and artificial reservoirs: Dean and Skead 1989), the African Black Duck is the only habitat specialist of the species complex preferring rivers in wooded and mountainous regions (McKinney et al. 1978). Both species primarily occur in South Africa, but can also be found in Angola, Botswana, Burundi, The Democratic Republic of the Congo, Ethiopia, Kenya, Malawi, Mozambique, Namibia, Rwanda, Tanzania, Uganda, Zambia and Zimbabwe (Johnsgard 1978; Gebrekidan et al. 2023). Population sizes are roughly estimated at ~10 000 and ~200 000 for the African Black Duck and Yellow-billed Duck, respectively (Dean and Skead 1989); however, studies into their basic ecology and detailed census estimates remain limited. Similarly, few studies looking into population structure, standing genetic diversity, and evolutionary histories exist for these species, with the most recent studies largely focusing on the Yellow-billed Duck (Brown et al. 2019; de Souza et al. 2019; Stephens et al. 2020).

Here, we assessed ancestry across multiple individuals using the mitochondrial DNA control region and thousands of nuclear loci to attempt to fill in critical knowledge gaps in the evolutionary histories and population genetics of the African Black Duck and Yellow-billed Duck. In addition to establishing the standing genetic diversity, we also took the opportunity to determine whether hybrids between these two sister species exist. Additionally, we test whether the burgeoning feral Mallard populations of South Africa are a cause for conservation concern via introgressive hybridisation. Although we expect that limited instances of hybridisation occur between the African species due to ecological isolation, the presence of Mallards across this landscape may result in elevated hybridisation (Brown et al. 2019; de Souza et al. 2019; Stephens et al. 2020). Finally, we compare morphometrics among samples with genetically vetted ancestry and sex to identify trait measurements that can be used for field identification. Together, these data shed insight into these genetically poorly understood species, including potential conservation concerns, while developing tools for better conservation management in the future.

Materials and methods

Sampling and DNA extraction

A total of 94 and 5 muscle tissue samples were salvaged from hunter-harvested Yellow-billed Ducks (YBDU) and African Black Ducks (AFBD), respectively, from three provinces in South Africa: Mpumalanga (N_{YBDU} = 46; N_{AFBD} = 4), Free State (N_{YBDU} = 49; N_{AFBD} = 1) and KwaZulu-Natal (N_{YBDU} = 32; N_{AFBD} = 0) (Supplementary Table S1). Genomic DNA from the 99 samples was extracted using a DNeasy Blood and Tissue Kit, following the manufacturer's protocol (Qiagen, Valencia, California). DNA quality was based on the presence of a high-molecular-weight band visualised using gel electrophoresis with a 1% agarose gel (Graham et al. 2015), and quantified using a Qubit 3 Fluorometer (Invitrogen, Carlsbad, California) to ensure a minimum concentration of 20 ng μl^{-1}.

Mitochondrial DNA

The mitochondrial DNA (mtDNA) control region was polymerase chain reaction (PCR) amplified using primers L78 and H774 (Sorenson and Fleischer 1996; Sorenson et al. 1999), and sequenced using an optimised Sanger sequencing protocol described in Lavretsky et al. (2014a). PCR products were visualised via agarose electrophoresis, purified using ExoSAP-IT treatment (ThermoFisher), and then sequenced using primer L78 with the 3130xl Genetic Analyzer (Applied Biosystems, Waltham, Massachusetts), performed by the Genomic Analysis Core Facility at the Border Biomedical Research Center of the University of Texas at El Paso. Sequences were aligned and edited

using Sequencher 4.8 (Gene Codes Corporation, Ann Arbor, Michigan). All sequences were submitted to GenBank (accession numbers PV543677–PV543775 (Supplementary Table S1).

Prior to analyses, overlapping mtDNA control region sequences for the reference wild and game-farm Mallard samples were included (Lavretsky et al. 2014a, 2014b, 2019a, 2020). A mtDNA haplotype network was inferred using a median-joining algorithm in the program POPART 1.7 (Leigh and Bryant 2015). Haplotype diversity was obtained by simple counting. Finally, composite averaged pairwise population estimates of relative differentiation (F_{ST}) were calculated between wild Mallard, Yellow-billed Duck and African Black Duck, in DnaSP 6 (Rozas et al. 2017).

ddRAD-seq library preparation and sequencing

Double digest restriction-site associated DNA (ddRAD-seq) libraries were prepared following protocols in DaCosta and Sorenson (2014; also see Lavretsky et al. 2015) using the restriction enzymes SbfI and EcoRI, but with a modified bead-based size-selection protocol as described in Hernandez et al. (2021). Libraries were pooled in equimolar concentrations and sent for 150-base pair (bp) single-end sequencing on the Illumina HiSeq X system (Novogenetics Ltd). Once again, previously published raw ddRAD sequences for the same samples representing wild and game-farm Mallards used in the mtDNA analyses were used as nuclear references, and were included in subsequent bioinformatics steps (Lavretsky et al. 2019a, 2020). All new Illumina raw reads are deposited in the NCBI Sequence Read Archive (http://www.ncbi.nlm.nih.gov/sra; BioProject PRJNA1249530, accession numbers SAMN47904803–SAMN47904901) (Supplementary Table S1).

Bioinformatics included de-multiplexing Illumina raw reads using the *ddRADparser.py* script of the BU-ddRAD-seq pipeline (DaCosta and Sorenson 2014) based on perfect barcode/index matches. Next, custom in-house Python scripts (available at https://github.com/jonmohl/PopGen; see Lavretsky et al. 2020) were used to automate sequence filtering, alignment and genotyping, using a combination of the tools Trimmomatic (Bolger et al. 2014), Burrows-Wheeler Aligner 07.15 (Li and Durbin 2011) and Samtools 1.7 (Bolger et al. 2014) (detailed methods can be found in the Supplementary Material). All sequences were aligned to a recently published and chromosomally assembled wild North American Mallard genome (Lavretsky et al. 2023b). Variant call format (VCF) files were further filtered for any base pair missing >10% of samples, which also included a minimum base-pair depth of 5X (i.e. 10X per genotype) and per base PHRED quality scores of ≥30, using VCFtools 0.1.15 (Danecek et al. 2011).

Nuclear population structure

The population structure of nuclear DNA variation was assessed using independent bi-allelic autosomal ddRAD-seq single nucleotide polymorphisms (SNPs). We completed all analyses without *a priori* information on population or species identity. We used VCFtools 0.1.15 (Danecek et al. 2011) to extract bi-allelic SNPs, and then PLINK 1.9 (Purcell et al. 2007) to filter for singletons (i.e. minimum allele frequency [MAF] ≥ 0.0015), any SNP missing ≥10% of data across samples, and linkage disequilibrium (LD). A significant LD correlation factor (r^2) of ≥0.5 resulted in randomly excluding one of the two ddRAD-seq SNPs.

Population structure was first visualised implementing a principal components analysis (PCA) in PLINK 1.9 (Purcell et al. 2007). We then calculated a matrix of co-ancestry coefficients with the package 'fineRADstructure' (Malinsky et al. 2018) and individual maximum-likelihood estimation of admixture probabilities as implemented in ADMIXTURE 1.3 (Alexander et al. 2009; Alexander and Lange 2011; Shringarpure et al. 2016). A co-ancestry matrix is based on the distribution of identical or nearest-neighbour haplotypes among samples with recent co-ancestry emphasised by rare SNPs (Kimura and Ohta 1973), and thus an increase in these SNPs is expected to correspond with relatedness. The fineRADstructure program was run with a burn-in of 100 000 iterations, followed by 100 000 Markov chain Monte Carlo iterations, and tree-building using default parameters. The co-ancestry matrix was visualised using R scripts *fineradstructureplot.r* and *finestructurelibrary.r* (https://www.milan-malinsky.org/fineradstructure). Finally, individual assignment probabilities (*Q*-values) were estimated with ADMIXTURE 1.3 (Alexander et al. 2009; Alexander and Lange 2011; Shringarpure et al. 2016), including standard errors based on 100 bootstrap replicates for each evaluated *K*-selection population model (of 1–5). Each ADMIXTURE analysis was run with a 10-fold cross-validation and with a quasi-Newton algorithm employed to accelerate convergence (Zhou et al. 2011). Each analysis used a block relaxation algorithm for point estimation and was terminated once the change (i.e. delta) in the log-likelihood of point estimations increased by less than 0.0001.

Finally, locus-by-locus and a composite of average pairwise population estimates of genetic differentiation (represented by F_{ST}) between wild Mallard, Yellow-billed Duck and African Black Duck were calculated in VCFtools 0.1.15 (Danecek et al. 2011).

Nucleotide diversity and demographic analyses

Per locus and average nucleotide diversity was calculated across autosomal and Z-sex chromosome linked ddRAD-seq loci for wild Mallard, Yellow-billed Duck and African Black Duck in VCFtools 0.1.15 (Danecek et al. 2011). Additionally, we compared long-term demographic histories across autosomal nuclear ddRAD-seq loci for those same groups, following the approach of Hernandez et al. (2021). In short, changes in effective population size through time are modelled by first creating single-species ∂a∂i input files of the one-dimensional site-frequency spectrum (SFS), calculated from Nexus-formatted concatenated sequencing data using custom python scripts (available at https://github.com/jibrown17/Dove_dadi.demographics; Hernández et al. 2021). Note that we re-filtered datasets for missingness (≥10%) per species, set to maximise the total base pairs prior to analyses in VCFtools 0.1.15 (Danecek et al. 2011). Each species' SFS was then folded and masked before being projected down to account for missing data between

groups ($N_{\text{Wild MALL}}$ = 46 alleles, N_{YBDU} = 120 alleles, N_{AFBD} = 10 alleles). Next, based on a custom demographic model that uses 100 integration steps (Hernández et al. 2021; https://github.com/jibrown17/Dove_dadi.demographics), ∂a∂i creates a model SFS that is used to estimate the optimum parameters of an effective population ($N_n = v_n \times N_{Anc}$; where N_n = effective population size at the n^{th} time interval) and time intervals ($t_n = T_n \times 2 \times N_{Anc} \times G$; where t_n = total years before present at the n^{th} time interval, and G = generation time) for each integration step.

Optimised parameters are then scaled to the empirical data using θ ($\theta = 4N_{ANC} \times \mu$; where N_{ANC} = ancestral effective population size), and subsequently used to calculate biologically informative values of effective population size through time (Hernández et al. 2021). We used the geometric mean calculated across 50 replicates of parameter optimisation for each model and estimated the goodness-of-fit for each model by calculating the log-likelihood of the model given the empirical data. Finally, we estimated confidence intervals (CI) using parameter uncertainty metrics included in ∂a∂i (Gutenkunst et al. 2009; Coffman et al. 2016). Briefly, ∂a∂i calculates uncertainty values using a Fisher information matrix (FIM) that provides a calculation of variance by measuring how much information can be derived from the data with respect to an unknown parameter. The FIM requires a step size (ε) to be chosen for the calculation of the numerical derivatives. We note that ∂a∂i is unable to calculate an uncertainty value for a parameter if the numerical derivative of the parameter is negative; therefore, we maximised the number of parameters for which ∂a∂i is able to return a true estimate of uncertainty by calculating uncertainty across a range of step sizes ($\varepsilon = 10^{-2} – 10^{-9}$: Coffman et al. 2016; Blischak et al. 2020; also see detailed methods on uncertainty metrics in Hernández et al. 2021).

Sex determination

Although sex was assigned across the samples when collected in the field, we wanted to verify sex with molecular methods (Supplementary Table S1). First, we assigned sex based on differences in sequencing depth across autosomal and sex chromosome-linked ddRAD-seq loci (Lavretsky et al. 2019b). In short, the homogametic sex (i.e. males = ZZ) is expected to have near-zero levels of sequencing depth across W-linked loci but near equal depth for Z-linked loci when compared with autosomal loci. Conversely, we expected about half the sequencing depth at both the W- and Z-linked loci as compared with the autosomal loci for the heterogametic sex (i.e. females = ZW).

In addition to assigning sex based on ddRAD-seq loci, sex was determined by amplifying homologous CHD genes found on both sex chromosomes of birds (Çakmak et al. 2017). In short, the amplified CHD gene found on the Z versus W chromosome differed by several bases, resulting in the PCR amplification and gel electrophoresis of one band in the homogametic sex versus two bands in the heterogametic sex. Primers for PCR amplification were based on Çakmak et al. (2017), except the PCR mixture and thermocycler conditions were re-optimised as described in Lavretsky et al. (2023c). Amplification was verified using gel electrophoresis with 4% agarose gel.

Morphological trait measurements and analyses

Six morphological traits were measured in the field, for both the Yellow-billed Ducks and African Black Ducks, although for statistical purposes we had sufficient sampling of both males and females for only the Yellow-billed Duck. Mass (g) was measured with a Pesola Medio-Line spring scale (2 500 g × 20 g); wing chord was measured using an Avinet Universal Wide Wing & Tail Ruler (30 mm); and a digital calliper was used to measure bill length, bill width, culmen length and tarsus length (mm) (see Supplementary Table S1 for sample-specific morphological measurements). For each of the six traits, statistical differences ($p < 0.001$) between males and females of the Yellow-billed Duck were tested in SPSS (IBM Corp 2020). Note that the total number of males and females differed between each trait because of missingness (Supplementary Table S1). We then performed an independent *t*-test for importance across the six traits (Table 1), whereby doing this assured that we used the correct morphological traits that would yield salient results. All significant traits were then analysed by PCA by first log-transforming them for standardisation, and this provided further information regarding which of the morphological traits were most responsible for data variability.

Next, we used two machine learning methods to determine the most important variables and accuracy probability to distinguish between sexes. A random-forest analysis was first implemented in the R package 'randomForest' (Liaw and Weiner 2002) based on a total of 1 000 trees without replacement. The analysis optimises the number of measurements and assignment probability by sequentially removing variables from the random-forest analysis until the most parsimonious number of traits is reached. In addition, we built a regression decision tree using recursive partitioning analysis as implemented in the R package 'rpart' (Therneau et al. 2015). This machine learning-based classification method produces binary trees that best explain and partitions between classes (e.g. sexes). Model fit and optimisation were based on complexity values set to increase by a minimum of 0.0001, and with cross-validation errors used to determine the best complexity value eventually used to produce the final pruned tree.

Results

Mitochondrial haplogroups

A total of 594 overlapping base-pairs of the mtDNA control region were aligned across samples. A total of four major haplogroups were recovered, with Yellow-billed Duck and African Black Duck making up their own respective monophyletic and highly divergent groups (Figure 1). Although we recovered the two Old World (OW; Eurasian origin) A and New World (NW; North American origin) B divergent mtDNA haplogroups among Mallards (Ankney et al. 1986; Avise et al. 1990; Lavretsky et al. 2014), none of the African samples were found carrying either of those haplogroups. A total of three haplotypes were recovered across the five samples of African Black Ducks, whereas 16 haplotypes were recovered across the 94 samples of Yellow-billed Ducks. Whereas most haplotypes differed by a maximum of two mutations, the Yellow-billed

Table 1: Measurements by sex of six morphological traits in the Yellow-billed Duck *Anas undulata* and African Black Duck *A. sparsa*. Average, range and sample size are presented

Trait	Yellow-billed Duck		African Black Duck	
	Males	Females	Males	Females
Mass (g)	1 021.64	867.12	1 160	925
	875–1 250	750–1 000	1 100–1 200	n/a
	31	66	3	1
Bill length (mm)	59.03	53.94	57.35	53.00
	53.80–62.00	52.00–57.50	55.10–59.50	n/a
	32	65	4	1
Bill width (mm)	20.17	18.85	20.50	19.10
	15.80–21.60	17.10–21.20	19.90–21.10	n/a
	30	61	4	1
Culmen–Nares (mm)	51.18	46.66	48.45	44.10
	43.40–50.80	45.60–56.70	47.70–49.20	n/a
	30	61	4	1
Tarsus length (mm)	53.85	50.32	56.53	51
	49.10–56.70	38.60–54.80	54.60–57.50	n/a
	30	61	4	1
Wing length (mm)	27.08	25.45	27.45	25.50
	25.50–28.50	24.20–26.70	26.80–27.90	n/a
	32	65	4	1

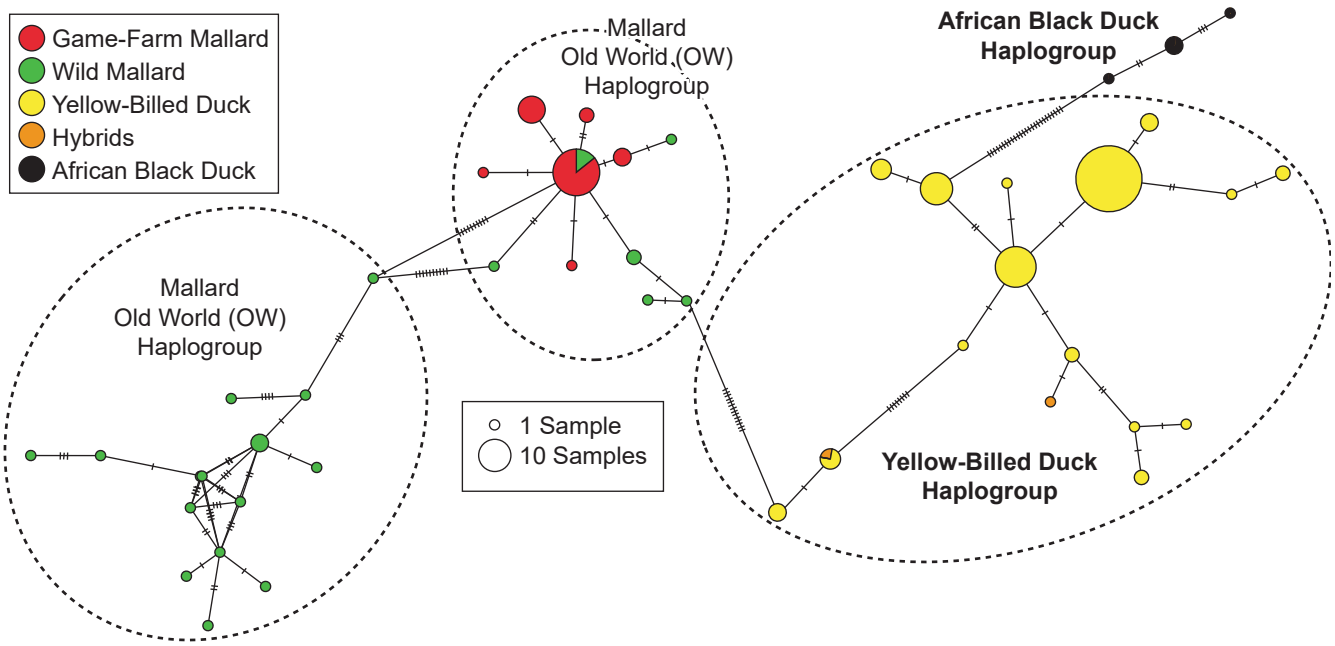

Figure 1: Mitochondrial DNA median-joining network for the Yellow-billed Duck *Anas undulata*, African Black Duck *A. sparsa*, and haplogroups of Old World and New World Mallards *A. platyrhynchos*. Network reconstructed in POPART 1.7 (Leigh and Bryant 2015). Circle size corresponds with sample size represented within each haplotype, with the number of hash marks on branches representing the number of mutations separating each haplotype

Duck haplogroup also included two divergent lineages separated by eight mutations (Figure 1). Finally, pairwise estimates of relative differentiation recovered similar levels between wild Mallard and the Yellow-billed Duck (F_{ST} = 0.59) and African Black Duck (F_{ST} = 0.69), which was slightly lower than recovered between the two African species (F_{ST} = 0.86).

Nuclear sequencing and population structure

A total of 106 672 base pairs (bp) were recovered across autosomal (100 250 bp) and Z-linked (6 422 bp) ddRAD-seq loci that met our sequencing coverage and missing data criteria across 157 genotyped samples. An average sequencing depth of 126 sequences and range of 15–2 227 sequences per locus were recovered across samples.

Population structure analyses of all 157 samples were based on 14 498 (of 15 674) independent bi-allelic autosomal ddRAD-seq SNPs. Plotting the first three principal components in the PCA explained 53.3% of the variation (Supplementary Figure S2A), providing four clear genetic groups distinguishing between wild Mallard, game-farm Mallard, Yellow-billed Duck and African Black Duck (Figure 2). The same genetic clusters were also recovered in our co-ancestry matrix (Figure 3a), in which we recovered the highest co-ancestry among African Black Ducks as compared with the other species. The associated dendrogram had strong bootstrap nodal support, and identified the Yellow-billed Duck as the most divergent lineage, with the African Black Duck being an outgroup to the clade of wild and game-farm Mallards. Our ADMIXTURE analyses recovered an optimum population model of three groups (Supplementary Figure S2b) that separated Mallards (wild and game-farm), Yellow-billed Duck and African Black Duck; separation of wild and game-farm Mallards was achieved under a five-population model (K = 5) (Figure 3b).

We found a single sample that was in an intermediate PCA space between Yellow-billed Duck and wild Mallard (Figure 2), and which was recovered has having mixed co-ancestry and ADMIXTURE individual assignment

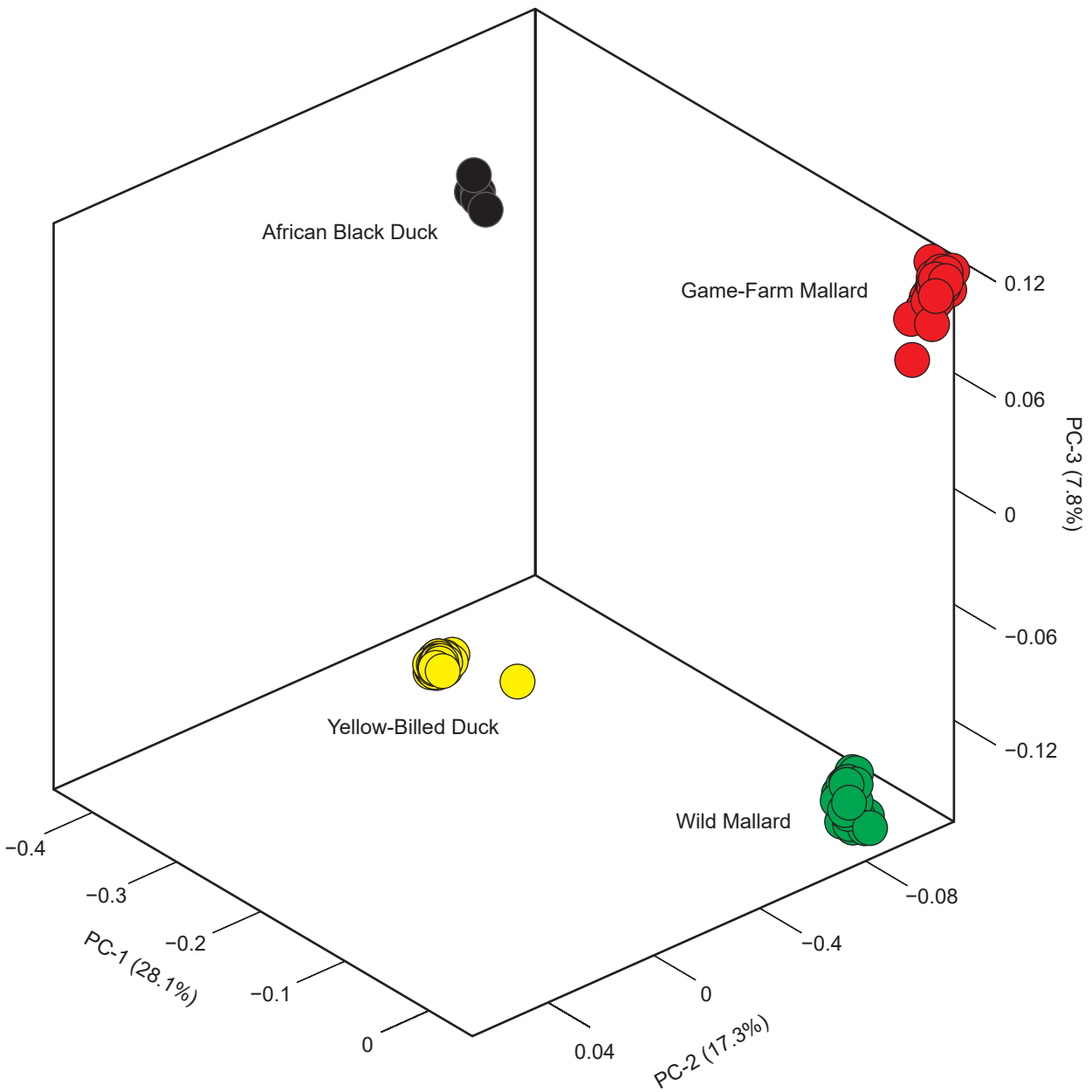

Figure 2: Plot of the first three principal components derived from a principal components analysis (PCA) of 157 samples and 14 498 independent bi-allelic autosomal ddRAD-seq SNPs for the African Black Duck *Anas sparsa* (*n* = 5) and Yellow-billed Duck *A. undulata* (*n* = 94), as well as samples of wild (*n* = 26) and game-farm (*n* = 32) Mallards *A. platyrhynchos*. PCA computed in PLINK 1.9 (Purcell et al. 2007)

probability from the two species (Figure 3a). Additionally, this sample was found to be outside the other Yellow-billed Duck samples in the *fineRADstructure* dendrogram. The average genomic assignment probabilities of ~75% and ~25% to the Yellow-billed Duck and wild Mallard genetic clusters, respectively, were consistent with this single sample being an F2 or later Yellow-billed Duck × wild Mallard hybrid. A second sample with similar genetic

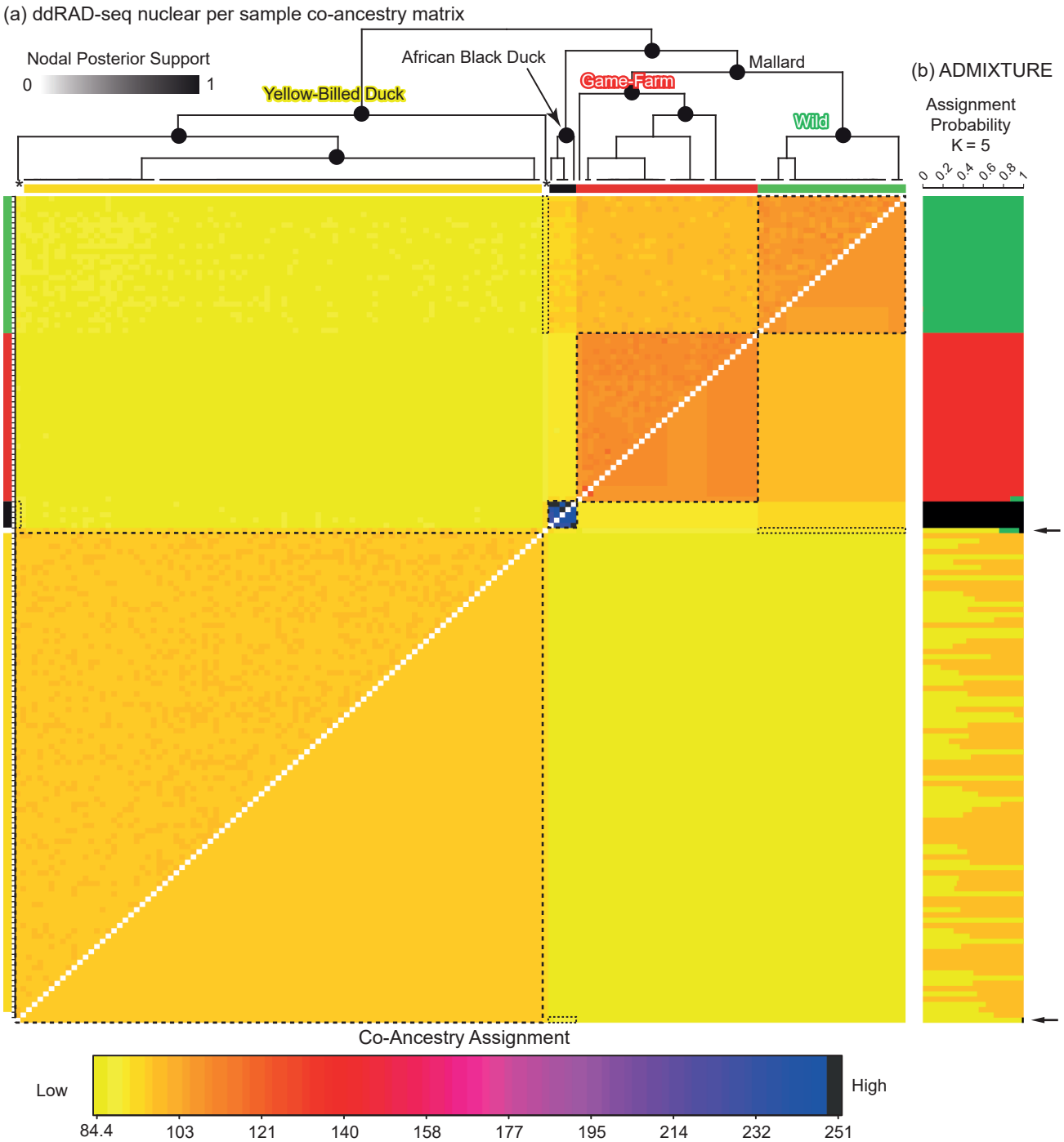

Figure 3: (a) Individual (above diagonal) and average (below diagonal) co-ancestry coefficient matrix estimated across 157 samples and 14 498 independent bi-allelic autosomal ddRAD-seq SNPs in the African Black Duck *Anas sparsa* (*n* = 5) and Yellow-billed Duck *A. undulata* (*n* = 94), as well as samples of wild (*n* = 26) and game-farm (*n* = 32) Mallards *A. platyrhynchos*, computed in R using the package 'fineRADstructure' (Malinsky et al. 2018). The level of recent co-ancestry is colour-coded from low (yellow) to high (blue). (b) Additionally, we provide individual assignment probabilities across samples calculated across the same SNP set in ADMIXTURE 1.3 (Alexander et al. 2009; Alexander and Lange 2011; Shringarpure et al. 2016) under a *K*-selected population of five. The two putative hybrids are highlighted in both analyses

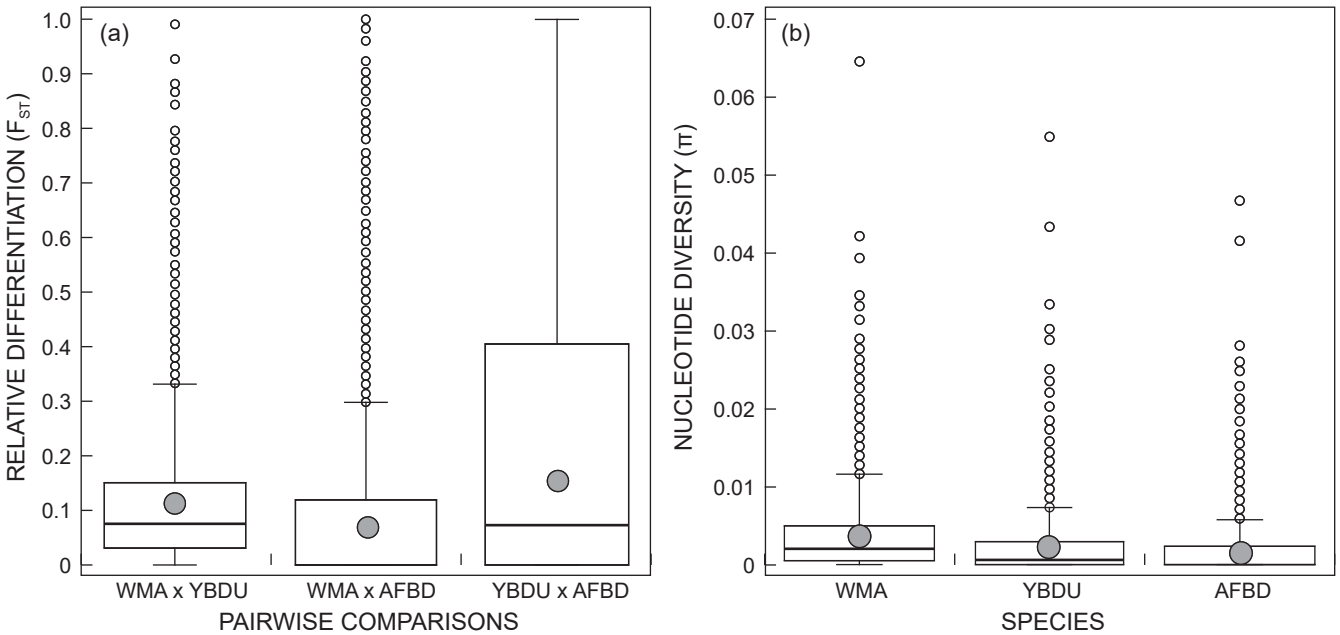

Figure 4: Box-plot distribution of (a) relative differentiation (F_{ST}) across pairwise comparisons across, as well as (b) nucleotide diversity across ddRAD-seq loci in wild Mallards *Anas platyrhynchos* (WMA), Yellow-billed Ducks *A. undulata* (YBDU), and African Black Ducks *A. sparsa* (AFBD). Grey circles denote the mean value for each calculation

intermediacy signatures in the co-ancestry plot was consistent with being a highly back-crossed Yellow-billed Duck × African Black Duck hybrid (Figure 3).

The two putative hybrids were excluded when estimating the locus-by-locus and composite average pairwise population estimates of relative differentiation (F_{ST}) between wild Mallard, Yellow-billed Duck and African Black Duck. Generally, Yellow-billed Duck was most genomically differentiated from African Black Duck (avg. F_{ST} = 0.23) and wild Mallard (avg. F_{ST} = 0.11) as compared with African Black Duck versus wild Mallard (avg. F_{ST} = 0.093) (Figure 4a).

Genetic diversity and demographic histories

A total of 96 326 base pairs of autosomal loci only were used in ∂A∂I demographic analyses. Once again, both hybrids were excluded from the datasets. First, we recovered the lowest range of nucleotide diversity for African Black Ducks (avg. π = 0.0015), followed by Yellow-billed Ducks (avg. π = 0.0021) and wild Mallards (avg. π = 0.0037) (Figure 4b). Such differences in genomic diversity translated to variance in the demographic results. Specifically, both the Yellow-billed Duck and African Black Duck exhibited substantially lower overall effective population sizes as compared with wild Mallard across time (Figure 5). Despite these differences in effective population sizes, the wild Mallard and Yellow-billed Duck had similar trends, with their first population peaks of ~13.2 million and ~3 million individuals, respectively, occurring between 400 000 and 500 000 years before present (ybp), followed by a continuous population decline until c. 100 000 ybp. After this period, the Yellow-billed Duck and Mallard experienced a second increase in effective population sizes, peaking at ~3 million and ~10 million individuals today, respectively. Next, unlike the other two species whose estimates went into millions of years, the effective population size of the African Black Duck could only be assessed through ~230 000 ybp (Figure 5). Inferred trends for the African Black Duck followed those recovered for the Yellow-billed Duck after 100 000 ybp, including reaching a peak effective population size of ~2.5 million individuals today (Figure 5).

Molecular sex determination

Sex across the samples of Yellow-billed Ducks and African Black Ducks was successfully determined based on sequencing depth among sex and autosomal ddRAD-seq loci (Supplementary Figure S1) and with optimised PCR protocols (Supplementary Table S1). Both molecular methods yielded identical sex IDs across the samples, indicating that ~13% (13 of 99) of the samples were incorrectly identified in the field (Supplementary Table S1).

Morphological trait measurements and a dichotomous key

Note that both putative hybrids were again excluded from the morphological analysis, resulting in a dataset that included 5 African Black Ducks and 92 Yellow-billed Ducks. First, we found that males on average were heavier, had longer and wider bills, as well as longer tarsi and wings than females regardless of species (Table 1). When comparing between species, we found that male Yellow-billed Ducks are heavier (by ~139 g) and have longer bills (by 1.68 mm and 2.73 mm based on length and culmen nares, respectively) than male African Black Ducks (Table 1). Conversely, we found that male African Black Ducks have wider bills (by 0.33 mm), longer tarsi (by 2.68 mm) and longer wings (by 0.37 mm) than male Yellow-billed Ducks (Table 1). The reduced number of African

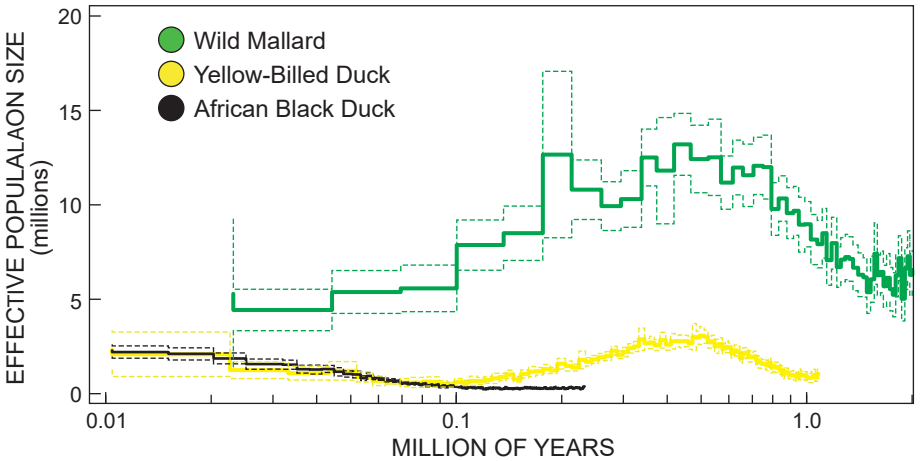

Figure 5: Time-series demographic results for wild Mallards *Anas platyrhynchos*, African Black Ducks *A. sparsa* and Yellow-billed Ducks *A. undulata*, with the associated 95% confidence intervals, estimated using single-species ∂a∂i models (Hernández et al. 2021) based on bi-allelic ddRAD-seq autosomal SNPs

Table 2: Mean difference and standard error (SE) difference between sexes of the Yellow-billed Duck *Anas undulata* for six morphometric traits, together with the *p*-value from the *t*-test for equality of means and SE differences. Principal component analysis loadings for the first two PCs (see Figure 6) for the sexes. Bold font denotes traits that contributed the most to each PC

	Mean difference	SE difference	*p*-value	PC1	PC2
Mass (g)	149.18	17.40	<0.001	**0.43**	0.095
Bill length (mm)	4.52	0.43	<0.001	**0.45**	−0.26
Bill width (mm)	1.14	0.19	<0.001	0.35	**0.84**
Culmen–Nares (mm)	3.79	0.57	<0.001	0.40	−0.022
Tarsus length (mm)	3.28	0.64	<0.001	0.38	**−0.47**
Wing length (mm)	1.15	0.15	<0.001	**0.43**	−0.083

Black Duck samples (Table 1) did not allow us to test for significance of sex differences for this species.

When comparing between sexes of Yellow-billed Ducks, we found that all six morphological traits were statistically different ($p < 0.001$) (Table 2). Analysing the log-transformed data in our PCA recovered significant loadings for all traits except culmen length, with the first and second principal components accounting for 67% and 9.67% of the variation within the data, respectively. Plotting out the PCA recovered 92% (61 of 66) of the males and 86% (24 of 28) of the females within their respective 95% ellipsis (Figure 6); 1 male and 2 females were recovered within the 95% ellipsis of the alternative sex. Generally, the right-skewed distribution of Yellow-billed Duck females denotes overall smaller morphological traits as compared with males.

Next, employing machine learning to optimise among morphological traits in distinguishing between male and female Yellow-billed Ducks, we recovered bill length as the most-informative single variable in our recursive partitioning tree, with 94% accuracy (Figure 7a), whereas our random-forest analysis recovered a minimum of two terminal nodes yielding 95.45% accuracy (Figure 7b). In fact, adding additional morphological variables to our random-forest analysis never increased accuracy past 95.45%. Analyses using the R package 'reprtree' (Banerjee et al. 2012) recovered mass and bill length as the two measurements that maximised accuracy based on Mallows L2 distance. In short, birds that are >937.5 g were immediately

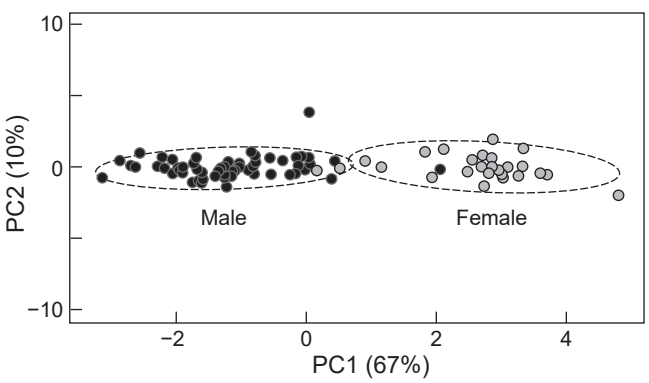

Figure 6: Plot of the first two principal components of the principal component analysis (PCA) of morphometric data of the sexes of Yellow billed Ducks *Anas undulata*. Bill length and bill width accounted for ~67% of the variation in PC1 (*x*-axis), whereas mass and bill length accounted for ~10% of the variation of PC2 (*y*-axis) (see Table 2)

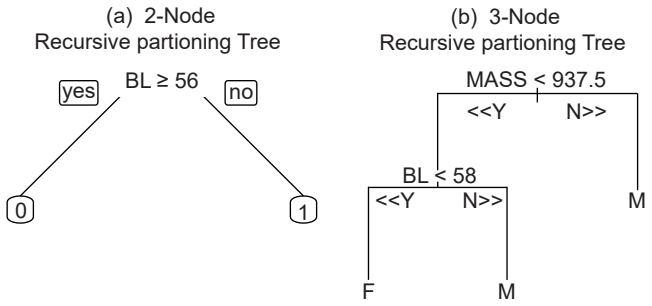

Figure 7: Sexing of the Yellow-billed Duck *Anas undulata* based on morphometric traits: (a) the 2-node recursive partitioning tree yielded an accuracy of 94% based on bill length (BL) only; (b) however, in the 3-node recursive partitioning tree an accuracy of 95.45% was achieved when combining BL and mass. If the mass was >937.5 g, the Yellow-billed Duck is likely a male; if the mass was <937.5 g and the bill length <58 mm, the Yellow-billed Duck is likely a female

characterised as male, whereas lighter birds required the bill length to be >58 mm to be identified as male (Figure 7b). Interestingly, we found that the single bill length cut-off of 56 mm identified in our recursive partitioning tree went to 58 mm in our random-forest analysis (Figure 7). Sources of inaccuracy were based on differences in weight, including two heavy females (1 adult and 1 hatch-year) and two smaller hatch-year males (Supplementary Table S1). Finally, we got 100% correct assignment when applying the same morphological cut-offs to the four African Black Duck samples with complete weight and bill length data (Supplementary Table S1).

Discussion

Our study is the first to apply a population genomics approach to understanding fundamental aspects of the population structure and evolution of the Yellow-billed Duck and African Black Duck. We report equally strong population structures within the Mallard, Yellow-billed Duck and African Black Duck. While generally similar in trend, our estimates of relative differentiation across pairwise comparisons were about half in comparison to those reported in previous studies (cf. Lavretsky et al. 2014; Brown et al. 2019). In addition to lower relative differentiation, a closer relationship between the African Black Duck and Mallard than between the African Black Duck and Yellow-billed Duck was supported in the derived dendrogram in the co-ancestry analysis. However, the heatmap identified the five African Black Duck samples as possessing the highest co-ancestry estimates, despite these individuals being collected from different locations and at different times (Supplementary Table S1). Rather, the elevated level of shared ancestry is likely explained by the reduction in genomic diversity in the African Black Duck as compared to in the Mallard and Yellow-billed Duck, by ~60% and ~31%, respectively. In fact, such a reduction likely further affected our ability to attain sufficient genomic signatures to assess effective population sizes back in time (also see Lavretsky et al. 2023a), with demographic analyses only going to ~200 000 ybp. Together, we posit that the close evolutionary relationship between the African Black Duck and Mallard is likely reflecting more stochastic processes (i.e. genetic drift) than true evolutionary history, which was further supported by similar and highly divergent mtDNA haplogroups. Given the evident reduction in genomic diversity among African Black Ducks likely biasing inferences, future work will require additional sampling of individuals and genomes to fully infer their histories.

Despite a 31% reduction in average genomic diversity in the African Black Duck as compared with the Yellow-billed Duck, our demographic results recover near identical effective population size estimates of ~2.5 million for these species today. Although, we acknowledge that reduced genomic diversity can bias such inferences, previous studies show that recent and severe bottlenecking often negatively impact effective population size inferences in more recent time-points where they appear to be near zero (Lavretsky et al. 2023a). Thus, we posit that the evolutionary history of the African Black Duck is likely best explained by retained long-term small population sizes. Although the African Black Duck census size is thought to be 100-fold smaller than that of the Yellow-billed Duck, these numbers remain largely unknown (Dean and Skead 1989). Moreover, we note that genomic representation is unlikely biasing results, as estimates of effective population size obtained here were highly concordant with previous estimates using low-coverage full genome data for these species (Camacho 2024). Whereas the low genetic diversity of the species is cause for conservation concern, it is likely linked to these being river specialists that are often found in small, localised populations. In fact, ignoring any skewness in sex ratios and potential overlapping generations, we estimate a census population (N_C) as four-times the effective population size (N_E) (Wright 1931; Charlesworth 2009) which puts the census size of each species at ~10 million individuals. These, estimates are substantially higher than current census sizes for both species, and thus begs the question of whether the extensive diversity is a product of retained ancestry from a more-diverse time and/or ancestral gene flow (Lavretsky et al. 2014, 2015; Peters et al. 2014). We do note that the increase in effective population sizes of the Yellow-billed Duck and African Black Duck after 100 000 ybp coincide with the onset of the last glacial period (i.e. 120 000–11 500 ybp) (Clark et al. 2009), including them peaking during the ending of the last glacial maximum (i.e. 29 000–19 000 ybp), suggesting that environments during this time favoured both species. Once again, increased genomic coverage will help shed light into these species' evolution.

Conservation considerations

In addition to building on our understanding of the evolutionary and population genetics of the Yellow-billed Duck and African Black Duck, we attempted to understand potential conservation concerns and build management tools. First, we confirm that the two species are as genomically distinct from one another as they are from the Mallard. Such genetic distinctiveness is evident not only in the substantially different mtDNA haplotypes, but also from nuclear-based population structure analyses.

Such distinctiveness permitted us to evaluate for possible admixture. Specifically, the burgeoning feral Mallard populations of Africa continue to be an existential threat for many species, including wild Mallards worldwide (Lavretsky et al. 2023a). Fortunately, we recovered none of the African Black Duck samples and only ~1% (1 of 94) of the Yellow-billed Duck samples as presenting an admixed history with Mallards, with the latter result being consistent with recent studies (Brown et al. 2019; de Souza et al. 2019). Importantly, one sample of Yellow-billed Duck was consistent with being a second-generation (F2) backcrossed individual with wild and not feral Mallard ancestry. Thus, while that single sample indicated that there is a limited level of introgressive hybridisation with wild migrating Mallards, it is more important that we did not find a hybrid of feral descent. Although we recovered limited gene flow with Mallards, we acknowledge that our sampling effort limited to South Africa is not representative of either species' true range (Johnsgard 1978; Gebrekidan et al. 2023). Thus, while we infer low levels of hybridisation in our sampling locations, increased sampling across the species' ranges will be necessary to establish true rates, given that other studies in other parts of their ranges have suggested a more elevated level of hybridisation with feral Mallards (Stephens et al. 2020). Next, we identified a single sample that possessed very small amounts of African Black Duck ancestry in the background of a largely Yellow-billed Duck genome. This highly backcrossed individual does suggest that hybridisation between the two species occurs but is likely even more limited than with the Mallard. Future work will require expanded sampling efforts to fully comprehend the extent of hybridisation.

Finally, we took the opportunity to attempt to develop the first set of morphological criteria for field identification between male and female Yellow-billed Ducks. First, we show that field identification of sex had an error rate of ~13%, but that the combination of measuring mass and bill length attained an accuracy of ~95% in diagnosing the sex of Yellow-billed Ducks. Although we were unable to conduct the same level of analysis for African Black Ducks, owing to the reduced sample size, similar keys could be achieved by following our outlined methods. Nevertheless, we did compare data between the two species and found no statistically significant differences in morphometrics between them, although within each species the sexes retained proportionate differences. Together, we provide a more reliable means for sex determination in the field, which can be used in future research and conservation efforts.

Acknowledgements — Funding was provided by the National Institutes of Health (NIH-NIMHD grant no. 5U54MD007592) and the National Science Foundation (DEB Grant ID 2010704).

Ethics statement — All samples were salvaged from legally harvested waterfowl and were imported for scientific purposes under federal USFWS permit no. MB11579C and APHIS USDA permit no. 639-23-143-93936.

Disclosure statement — No potential conflict of interest was reported by the authors.

ORCIDs

Philip Lavretsky: https://orcid.org/0000-0002-5904-8821
Sara Gonzalez: https://orcid.org/0009-0005-7268-1460
Vergie M Musni: https://orcid.org/0009-0002-8388-8297
Alexis Díaz: https://orcid.org/0000-0003-3291-562X
Joshua I Brown: https://orcid.org/0000-0001-6299-9815

References

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

Alexander DH, 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

Ankney CD, Dennis DG, Wishard LN, Seeb JE. 1986. Low genic variation between Black Ducks and Mallards. *The Auk* 103: 701–709. https://doi.org/10.1093/auk/103.4.701

Avise JC, Ankney DC, Nelson WS. 1990. Mitochondrial gene trees and the evolutionary relationship of Mallard and Black Ducks. *Evolution* 44: 1109–1119. https://doi.org/10.2307/2409570

Banerjee M, Ding Y, Noone AM. 2012. Identifying representative trees from ensembles. *Statistics in Medicine* 31: 1601–1616. https://doi.org/10.1002/sim.4492

Blischak PD, Barker MS, Gutenkunst RN. 2020. Inferring the demographic history of inbred species from genome-wide SNP frequency data. *Molecular Biology and Evolution* 37: 2124–2136. https://doi.org/10.1093/molbev/msaa042

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

Brown JI, Lavretsky P, Cumming GS, Peters JL. 2019. Strong population structure and limited gene flow between Yellow-billed Ducks and Mallards in southern Africa. *The Condor* 121: 1–10. https://doi.org/10.1093/condor/duz042

Çakmak E, Akın Pekşen Ç, Bilgin CC. 2017. Comparison of three different primer sets for sexing birds. *Journal of Veterinary Diagnostic Investigation* 29: 59–63. https://doi.org/10.1177/1040638716675197

Camacho FBH. 2024. Signatures of selection, gene flow, and genetic drift on the genomes of a recent avian radiation. PhD thesis, The University of Texas at El Paso, United States.

Caro T, Darwin J, Forrester T, Ledoux-Bloom C, Wells C. 2014. Conservation in the Anthropocene. In: Wuerthner G, Crist E, Butler T (eds), *Keeping the wild: against the domestication of Earth*. Washington DC: Island Press. pp 109–113. https://doi.org/10.5822/978-1-61091-559-5_9

Charlesworth B. 2009. Effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics* 10: 195–205. https://doi.org/10.1038/nrg2526

Clark PU, Dyke AS, Shakun JD, Carlson AE, Clark J, Wohlfarth B et al. 2009. The last glacial maximum. *Science* 325: 710–714. https://doi.org/10.1126/science.1172873

Coffman AJ, Hsieh PH, Gravel S, Gutenkunst RN. 2016. Computationally efficient composite likelihood statistics for demographic inference. *Molecular Biology and Evolution* 33: 591–593. https://doi.org/10.1093/molbev/msv255

Corlett RT. 2015. The Anthropocene concept in ecology and conservation. *Trends in Ecology and Evolution* 30: 36–41. https://doi.org/10.1016/j.tree.2014.10.007

DaCosta JM, Sorenson MD. 2014. Amplification biases and consistent recovery of loci in a double-digest RAD-seq protocol. *PLoS ONE* 9: e106713. https://doi.org/10.1371/journal.pone.0106713

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. 2011. The variant call format and

VCFtools. *Bioinformatics (Oxford, England)* 27: 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Day D. 1977. A morphological study of Yellow-billed Duck and Red-billed Teal. *Ostrich* 48: 86–96. https://doi.org/10.1080/00306525.1977.9633291

Dean WRJ, Skead DM. 1989. Survival and recovery rates of Yellow-billed Ducks. *The Journal of Wildlife Management* 53: 119–122. https://doi.org/10.2307/3801317

de Souza SG, Symes CT, Smit-Robinson H, Mollett JM. 2019. Minimal evidence of interspecific hybridisation between the Yellow-billed Duck and introduced Mallard in central and northwestern South Africa. *Ostrich* 90: 285–301. https://doi.org/10.2989/00306525.2019.1642252

Frost P, Ball I, Siegfried W, McKinney F. 1979. Sex ratios, morphology and growth of the African Black Duck. *Ostrich* 50: 220–233. https://doi.org/10.1080/00306525.1979.9634117

Gaitán-Espitia JD, Hobday AJ. 2021. Evolutionary principles and genetic considerations for guiding conservation interventions under climate change. *Global Change Biology* 27: 475–488. https://doi.org/10.1111/gcb.15359

Gebrekidan MT, Bekele A, Bogale BA. 2023. Population size and breeding ecology of Yellow-billed Duck (*Anas undulata*, Dubois 1839) in Chelekleka wetland, Ethiopia. *SINET: Ethiopian Journal of Science* 46: 71–80. https://doi.org/10.4314/sinet.v46i1.6

Graham CF, Glenn TC, McArthur AG, Boreham DR, Kieran T, Lance S et al. 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

Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD. 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

Hernández F, Brown JI, Kaminski M, Harvey MG, Lavretsky P. 2021. Genomic evidence for rare hybridization and large demographic changes in the evolutionary histories of four North American dove species. *Animals (Basel)* 11: 2677. https://doi.org/10.3390/ani11092677

Hoelzel AR, Bruford MW, Fleischer RC. 2019. Conservation of adaptive potential and functional diversity. *Conservation Genetics* 20: 1–5. https://doi.org/10.1007/s10592-019-01151-x

Hohenlohe PA, Funk WC, Rajora OP. 2021. Population genomics for wildlife conservation and management. *Molecular Ecology* 30: 62–82. https://doi.org/10.1111/mec.15720

Holderegger R, Balkenhol N, Bolliger J, Engler JO, Gugerli F, Hochkirch A et al. 2019. Conservation genetics: linking science with practice. *Molecular Ecology* 28: 3848–3856. https://doi.org/10.1111/mec.15202

Johnsgard PA. 1978. Ducks, geese and swans of the world. Lincoln, Nebraska: University of Nebraska Press.

Johnson KP, Sorenson MD. 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

Kimura M, Ohta T. 1973. The age of a neutral mutant persisting in a finite population. *Genetics* 75: 199–212. https://doi.org/10.1093/genetics/75.1.199

Lavretsky P, McCracken KG, Peters JL. 2014a. 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, Hernández Baños BE, Peters JL. 2014b. 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. https://doi.org/10.1642/AUK-13-164.1

Lavretsky P, DaCosta JM, Hernández-Baños BE, Engilis A Jr, Sorenson MD, Peters JL. 2015. Speciation genomics and a role for the Z chromosome in the early stages of divergence between Mexican ducks and mallards. *Molecular Ecology* 24: 5364–5378. https://doi.org/10.1111/mec.13402

Lavretsky P, DaCosta JM, Sorenson MD, McCracken KG, Peters JL. 2019b. 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, Janzen T, McCracken KG. 2019a. Identifying hybrids and the genomics of hybridization: mallards and American black ducks of eastern North America. *Ecology and Evolution* 9: 3470–3490. https://doi.org/10.1002/ece3.4981

Lavretsky P, McInerney NR, Mohl J, Brown JI, James H, McCracken KG, Fleischer R. 2020. Assessing changes in genomic divergence following a century of human mediated secondary contact among wild and captive-bred ducks. *Molecular Ecology* 29: 578–595. https://doi.org/10.1111/mec.15343

Lavretsky P, Mohl JE, Söderquist P, Kraus RH, Schummer ML, Brown JI. 2023a. The meaning of wild: genetic and adaptive consequences from large-scale releases of domestic mallards. *Communications Biology* 6: article 819. https://doi.org/10.1038/s42003-023-05170-w

Lavretsky P, Hernández F, Swale T, Mohl JE. 2023b. Chromosomal-level reference genome of a wild North American mallard (*Anas platyrhynchos*). *G3: Genes/Genomes/Genetics* 13: jkad171. https://doi.org/10.1093/g3journal/jkad171

Lavretsky P, Hoyt A, Musni VM, Howell D, Williams CK. 2023c. Frequency and types of alternative breeding strategies employed by nesting American black ducks in North Carolina. *PLoS ONE* 18: e0278905. https://doi.org/10.1371/journal.pone.0278905

Lavretsky P, Kraai KJ, Butler D, Morel J, VonBank JA, Marty JR et al. 2024. Human-induced range expansions result in a recent hybrid zone between sister species of ducks. *Genes* 15: article 651. https://doi.org/10.3390/genes15060651

Leigh JW, Bryant D. 2015. POPART: full-feature software for haplotype network construction. *Methods in Ecology and Evolution* 6: 1110–1116. https://doi.org/10.1111/2041-210X.12410

Leipold M, Tausch S, Hirtreiter M, Poschlod P, Reisch C. 2020. Sampling for conservation genetics: how many loci and individuals are needed to determine the genetic diversity of plant populations using AFLP? *Conservation Genetics Resources* 12: 99–108. https://doi.org/10.1007/s12686-018-1069-1

Li H, Durbin R. 2011. Inference of human population history from individual whole-genome sequences. *Nature* 475: 493–496. https://doi.org/10.1038/nature10231

Liaw A, Weiner M. 2002. Classification and regression by randomForest. *R News* 2/3: 18–22.

Malinsky M, Trucchi E, Lawson D, Falush D. 2018. RADpainter and fineRADstructure: population inference from RADseq data. *Molecular Biology and Evolution* 35: 1284–1290. https://doi.org/10.1093/molbev/msy023

McKinney F, Siegfried WR, Ball IJ, Frost PGH. 1978. Behavioral specializations for river life in the African Black Duck (*Anas sparsa* Eyton). *Zeitschrift für Tierpsychologie* 48: 349–400. https://doi.org/10.1111/j.1439-0310.1978.tb00266.x

Nielsen ES, Hanson JO, Carvalho SB, Beger M, Henriques R, Kershaw F, Von der Heyden S. 2023. Molecular ecology meets systematic conservation planning. *Trends in Ecology and Evolution* 38: 143–155. https://doi.org/10.1016/j.tree.2022.09.006

Palmer RS. 1976. *Handbook of North American birds*, vol. 2. New Haven, Connecticut: Yale University Press.

Peters JL, Sonsthagen SA, Lavretsky P, Rezsutek M, Johnson WP, McCracken KG. 2014. Interspecific hybridization contributes to high genetic diversity and apparent effective

population size in an endemic population of mottled ducks (*Anas fulvigula maculosa*). *Conservation Genetics* 15: 509–520. https://doi.org/10.1007/s10592-013-0557-9

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D et al. 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

Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, Sánchez-Gracia A. 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution* 34: 3299–3302. https://doi.org/10.1093/molbev/msx248

Shringarpure SS, Bustamante CD, Lange K, Alexander DH. 2016. Efficient analysis of large datasets and sex bias with ADMIXTURE. *BMC Bioinformatics* 17: article 218. https://doi.org/10.1186/s12859-016-1082-x

Sorenson MD, Fleischer RC. 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

Sorenson MD, Ast JC, Dimcheff DE, Yuri T, Mindell DP. 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

Stephens K, Measey J, Reynolds C, Le Roux JJ. 2020. Occurrence and extent of hybridisation between the invasive Mallard Duck and native Yellow-billed Duck in South Africa. *Biological Invasions*: 693–707. https://doi.org/10.1007/s10530-019-02122-6

Therneau T, Atkinson B, Ripley B. 2015. rpart: Recursive Partitioning and Regression Trees. R package version 4.1-15. Available at https://cran.r-project.org/package=rpart.

Wright S. 1931. Evolution in Mendelian populations. *Genetics* 16: 97–159. https://doi.org/10.1093/genetics/16.2.97

Zhou H, Alexander D, Lange K. 2011. A quasi-Newton acceleration for high-dimensional optimization algorithms. *Statistics and Computing* 21: 261–273. https://doi.org/10.1007/s11222-009-9166-3

Manuscript received: 11 November 2024, revised: 14 March 2025, accepted: 9 April 2025
Associate Editor: S Miller