



## The meaning of wild: Genetic and adaptive consequences from large-scale releases of domestic mallards

Philip Lavretsky <sup>1✉</sup>, Jonathon E. Mohl <sup>2</sup>, Pär Söderquist<sup>3</sup>, Robert H. S. Kraus<sup>4</sup>, Michael L. Schummer<sup>5</sup> & Joshua I. Brown<sup>1</sup>

The translocation of individuals around the world is leading to rising incidences of anthropogenic hybridization, particularly between domestic and wild congeners. We apply a landscape genomics approach for thousands of mallard (*Anas platyrhynchos*) samples across continental and island populations to determine the result of over a century of supplementation practices. We establish that a single domestic game-farm mallard breed is the source for contemporary release programs in Eurasia and North America, as well as for established feral populations in New Zealand and Hawaii. In particular, we identify central Europe and eastern North America as epicenters of ongoing anthropogenic hybridization, and conclude that the release of game-farm mallards continues to affect the genetic integrity of wild mallards. Conversely, self-sustaining feral populations in New Zealand and Hawaii not only show strong differentiation from their original stock, but also signatures of local adaptation occurring in less than a half-century since game-farm mallard releases have ceased. We conclude that ‘wild’ is not singular, and that even feral populations are capable of responding to natural processes. Although considered paradoxical to biological conservation, understanding the capacity for wildness among feral and feral admixed populations in human landscapes is critical as such interactions increase in the Anthropocene.

<sup>1</sup>Department of Biological Sciences, University of Texas at El Paso, El Paso, TX 79668, USA. <sup>2</sup>Department of Mathematical Sciences, University of Texas at El Paso, El Paso, TX 79668, USA. <sup>3</sup>Faculty of Natural Sciences, Kristianstad University, SE- 291 88 Kristianstad, Sweden. <sup>4</sup>Department of Migration, Max Planck Institute of Animal Behavior, 78315 Radolfzell, Germany. <sup>5</sup>Department of Environmental Biology, State University of New York College of Environmental Science and Forestry, Syracuse, NY 13210, USA. ✉email: [plavretsky@utep.edu](mailto:plavretsky@utep.edu)

An outstanding but evolving question within evolutionary and conservation biology is what are the genetic consequences of hybridization that result from human-mediated disturbances<sup>1,2</sup>. Unfortunately, with increasing anthropogenic landscape changes, contact between wild and domestic forms seems unavoidable, raising concerns regarding the maladaptive consequences of such interactions on overall fitness and the future adaptive potential of wild populations<sup>1–4</sup>. Alongside changing habitats, domestic releases have become common world-wide, resulting in augmented genetic integrity of various wild populations<sup>5–8</sup>. In fact, in cases where releases have been common and intensive, gene flow from domestic individuals now poses a major threat to these various ecologically and economically important wild species (e.g., rice<sup>9</sup>, trout<sup>10</sup>, salmon<sup>11</sup>, pig<sup>12</sup>, quail<sup>13</sup>, geese<sup>14</sup>, ducks<sup>15–20</sup>, cats<sup>21</sup>, wolves and other wild canids<sup>22,23</sup>). However, genomic data has also revised and refined our understanding where contact between domestic and wild congeners do not always result in wide-spread introgressive hybridization<sup>24</sup>. More generally, predicting the outcomes of domestic introductions remains challenging because introgression of maladaptive traits into wild populations is additive, and negative consequences are not always immediately apparent<sup>25</sup>. With rapidly changing climatic conditions, it is becoming increasingly evident that there is a growing need to understand the genetic underpinning of such events as to predict future challenges in conservation<sup>26–28</sup>. In fact, it is possible that while feral populations are generally unsuited for survival in wild habitats, they may have a greater adaptive potential in human-dominated areas<sup>29</sup>. Here, we employ a landscape genomics approach on global wild and feral mallard (*Anas platyrhynchos*) populations to understand how lineages with varying demographic histories and selection regimes (i.e., natural versus artificial selective pressures) respond to natural versus human influences.

Starting in central China, humans have been domesticating mallards since shortly after 500 BC<sup>30,31</sup>. Since then, domestic mallard supplementation has been practiced around the world, with the most intense release programs beginning more recently in the early 20<sup>th</sup> century<sup>8,32</sup>. While naturally distributed across the Holarctic, the intentional or accidental release of mallards has expanded their range to nearly world-wide outside the Poles<sup>33</sup>. As a result of these efforts, widespread introgression from established feral populations now poses a genetic threat to populations of wild mallards and other closely related waterfowl found throughout Eurasia, North America, and the South Pacific<sup>20,34–37</sup>. Among regions, annual supplementation is currently greatest in Europe and eastern North America where nearly five-million<sup>38,39</sup> and at least two-hundred thousand (500,000 from 1920s to 1960s;<sup>40–42</sup>) game-farm mallards, respectively, are annually being released. Note that game-farm mallards are captive bred to be released on shooting preserves for hunting and dog training purposes<sup>43</sup> (Fig. 1). This substantial influx of domesticated individuals has already affected the genetic composition of Eurasian<sup>44</sup>, and North American<sup>16</sup> wild mallards. In fact, DNA sequencing of historical mallards pre-dating large scale game-farm augmentation events (pre-1920) confirmed that the genetic constitution of today's eastern North American mallard has been fundamentally altered through extensive gene flow with game-farm mallards<sup>43</sup>.

The [un]intentional release of domestic mallards across regions of varying habitat and ecology provides powerful experiments that can help us to understand how introduced population(s) respond under differing selective pressures, which is an underlying mechanism of local adaptation. We examine the genetic and adaptive consequences from large-scale supplementation practices of domestic mallards by coupling partial genomic sequencing of thousands of samples representing wild and feral

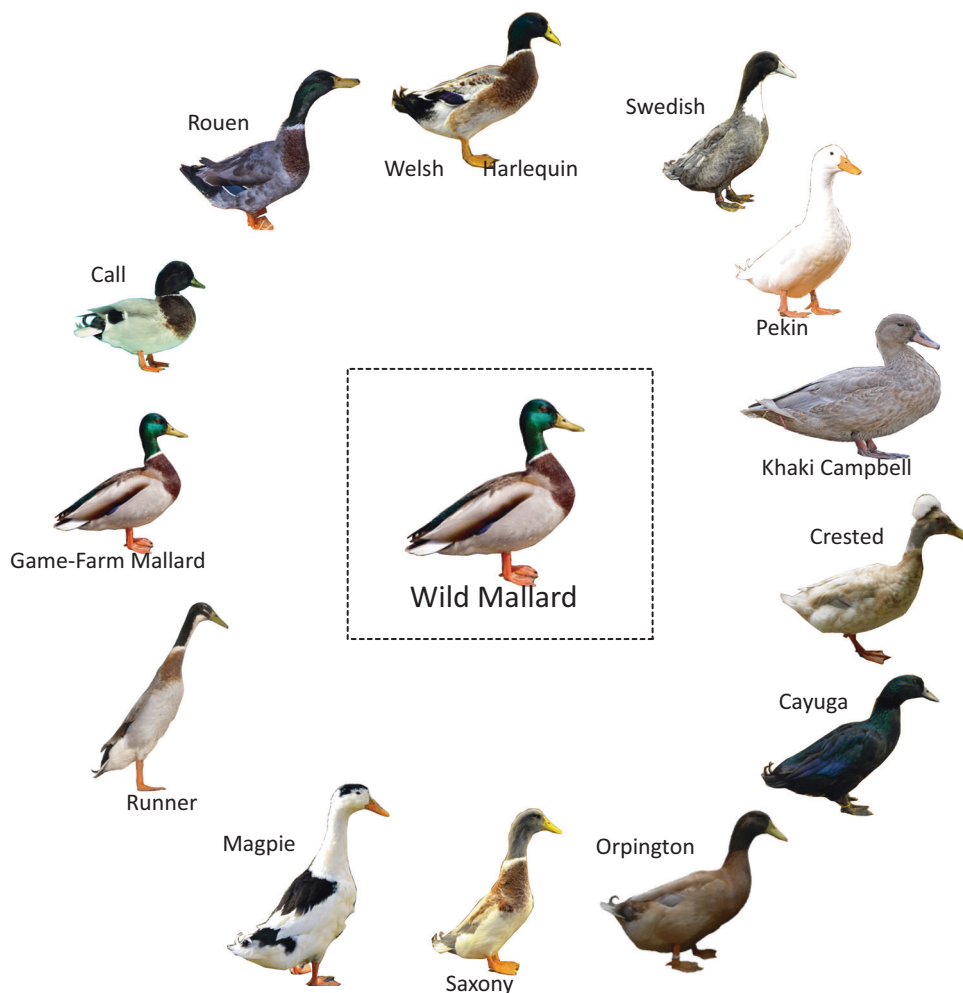
populations throughout Eurasia, North America, and New Zealand. After establishing ancestry across wild and feral populations, we couple demographic and phylogenetic analyses to identify the sources of stock currently being released (i.e., in mainland North America and Eurasia) and of two self-sustaining feral populations (Hawaii and New Zealand). Next, we use the program GRADIENT FOREST to reconstruct the adaptive landscape of wild and feral populations by testing for genotype-environment associations including both natural and anthropogenic indices<sup>45–47</sup>, and thus, reconstructing their genetic niche space<sup>47</sup>. Additionally, we determine whether the association with environmental factors is driven by many or few loci, as well as whether the same loci are consistently under selective pressures from each wild/feral population. In particular, with wild mallards being innately migratory, we predict that associations with environmental cues (e.g., mean diurnal range, temperature, and precipitation) will be consistent across loci for wild populations<sup>47</sup>. Furthermore, we posit that genetic diversity within feral mallard populations that have been shown to take advantage of more human-dominated habitats will be dictated by loci associated with anthropogenic indices (e.g., intensity of urbanization).

## Results

A total of 2951 ddRAD-seq loci (358,999 base-pairs (bp)) were recovered across chromosomes, including 2751 autosomal (333,731 bp), 193 Z-sex chromosome linked (24,331 bp), and 7 W-sex chromosome linked (937 bp) that met our sequencing coverage and missing data criteria for 1916 samples (Supplementary Fig. S1). We obtained an average median depth of 167 sequences/locus, with a median sequencing depth range of 26–235 sequences across samples. Importantly, sex was reliably identified across samples based on sequencing depth ratios of the sex and autosomal chromosomes (Supplementary Data 1 and Supplementary Fig. S2). Finally, a total of 600 overlapping base-pairs of the mtDNA control region were also obtained across the 1862 samples (of 1916); note once again that mtDNA was lacking for the feral mallards of Hawaii (see specifics in Supplementary Data 1).

**Mitochondrial DNA.** We recovered known Old World (OW) A and New World (NW) B haplogroups in our mtDNA haplotype network (Fig. 2a;<sup>48–50</sup>). First, all domestic, feral Khaki Campbell's, and all but one Eurasian wild mallard were recovered within the OW A haplogroup. The one wild mallard from China possessed a NW B haplotype, which may represent a North American migrant; however, being unique, it is also possible that this haplotype represents introgression with endemic spot-billed ducks (*A. poecilorhyncha*) that also possess B haplogroup variants<sup>51</sup>. Additionally, a single wild mallard from the Netherlands was recovered within the Khaki Campbell's haplotype. Although New Zealand's feral mallards primarily carried OW A haplotypes (91% of samples), eight samples carried six NW B haplotypes not shared with any other wild mallard (Fig. 2a; see Supplementary Data 1 for sample specifics).

In contrast to Eurasian mallards, 56% and 44% of North American wild mallards were recovered in the OW A versus NW B haplogroups, respectively. More importantly, 75% (568 of 791) of all OW A haplotypes recovered in North America were within the same 15 OW A haplotypes carried among sampled game-farm mallards. Similarly, while 13 OW A mtDNA haplotypes were found among New Zealand's feral mallards, 65% of these samples carried the same major OW A haplotype (Fig. 2a; see Supplementary Data 1 for sample specifics). After accounting for game-farm derived haplotypes, wild mallards within the OW ( $N_{hap} = 151$ ) and NW ( $N_{hap} = 297$ ) haplogroups showed high



**Fig. 1 The wild mallard and representation of mallard domestics.** Representation of breeds resulting from the domestication of the wild mallard.

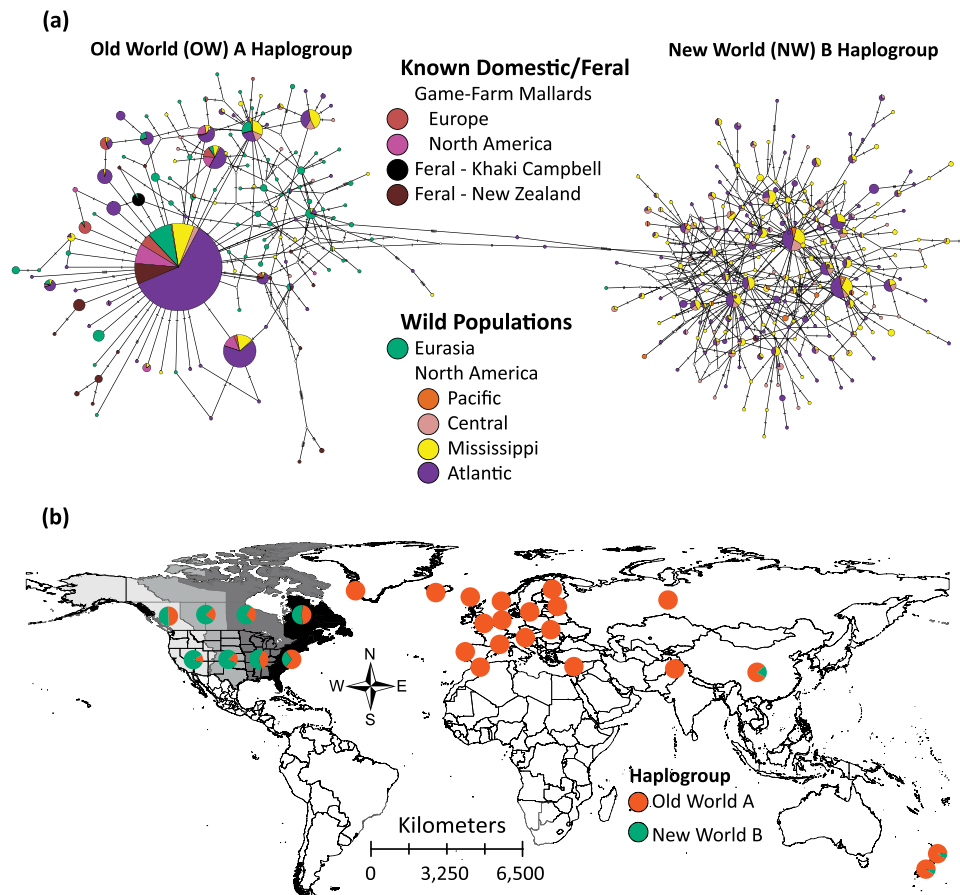
uniqueness and variability, whereas game-farm and feral populations showed the complete opposite trend with only 32 haplotypes found across 249 samples (Fig. 2a).

Although plotting the proportion of samples carrying OW A versus NW B haplotypes identified all European countries except one mallard from China as OW A, the presence of OW A mtDNA haplotypes increased eastward in North America<sup>16</sup> (Fig. 2b).

**Nuclear population structure and phylogenetics.** Population structure analyses were based on 36,637 (of 36,641) independent bi-allelic SNPs. Plotting the first two components of the PCA provided clear groupings of wild and domestic ancestry mallards, with individuals in intermediate space between them as hybrids (Fig. 3a, Supplementary Fig. S3A). Additional separation in PCA space was recovered for wild mallards from Greenland, and for all three of the sampled feral populations (i.e., New Zealand, Hawaii, and Khaki Campbell; Fig. 3a). Next, although an optimum  $K$  population model of 4 (Supplementary Fig. S3B) was recovered for our ADMIXTURE analysis across all samples, additional resolution was obtained up to a  $K$  population value of 7 (Fig. 3b). First, we found three wild mallard populations that included Greenland, remaining Eurasia, and North America. Whereas feral populations of Khaki Campbell's and mallards of New Zealand comprised unique genetic clusters, we confirmed that the game-farm mallard stocks currently being released in Eurasia and North America, as well as the ancestry of Hawaii's feral population

assign to the same genetic cluster (i.e., game-farm mallard genetic cluster; Fig. 3b). Finally, we found that reference game-farm mallards sourced in Eurasia<sup>52</sup> and from one breeder in North America<sup>53</sup> have been more recently admixed with wild mallards from their respective regions and consistent with their histories of directed efforts to admix them with wild birds.

Next, because the recovery of all unique genetic clusters in ADMIXTURE required us to evaluate increasing values of  $K$  populations up to seven, we tested whether ancestry assignments changed among admixed individuals by excluding feral populations. We analyzed known game-farm and wild mallards of Eurasia and mainland North America using a 36,637 independent bi-allelic SNP dataset and following the same protocols in obtaining assignment probabilities from ADMIXTURE as outlined in the methods for  $K$  population value of 2–4 (Fig. 3c). We again recovered unique genetic structuring between wild mallards of Eurasia and North America, but also found stability in assignment probabilities across samples once interspecific assignments were >10%. In fact, bootstrapped standard-deviations (SD) surrounding individual point estimates to the game-farm mallard genetic cluster was most concerted, averaging  $\pm 3\%$ , with a maximum SD of  $\sim 6\%$  regardless of which ADMIXTURE dataset was analyzed (Supplementary Fig. S4); and making any sample with  $\leq 6\%$  assignment indistinguishable from pure wild mallard. Conversely, SD values surrounding individual point estimates to wild mallard clusters, particularly as we forced additional population  $K$  values to be evaluated were much more variable,



**Fig. 2 Mitochondrial DNA population structure.** **a** A haplotype network based on 600 base-pairs of the mitochondrial control region and sequenced across wild mallards, known game-farm mallards, and feral mallard populations. **b** Plotting the proportion of samples recovered with either the Old World (OW) A or New World (NW) B haplotype across sampled locations. Note that North America was grouped by the four known migratory flyways that waterfowl use during annual migrations.

often deviating by >10% (Supplementary Fig. S4). Thus, we used our ADMIXTURE analysis under a  $K$  of 3 population model, and assigned individuals as hybrid if they possessed >10% interspecific assignment and all others as pure wild (i.e., assignment probability of <10% to game-farm mallard ancestry) or feral (i.e., assignment probability of  $\geq 90\%$  to game-farm mallard ancestry). We acknowledge that while some samples with assignment probability of 90–95% wild potentially includes late generational backcrosses, we were more concerned as to not create noise in our hybrid sample set by including individuals likely already biologically wild; the same ancestry cutoffs were used in previous studies<sup>16,43,54,55</sup>.

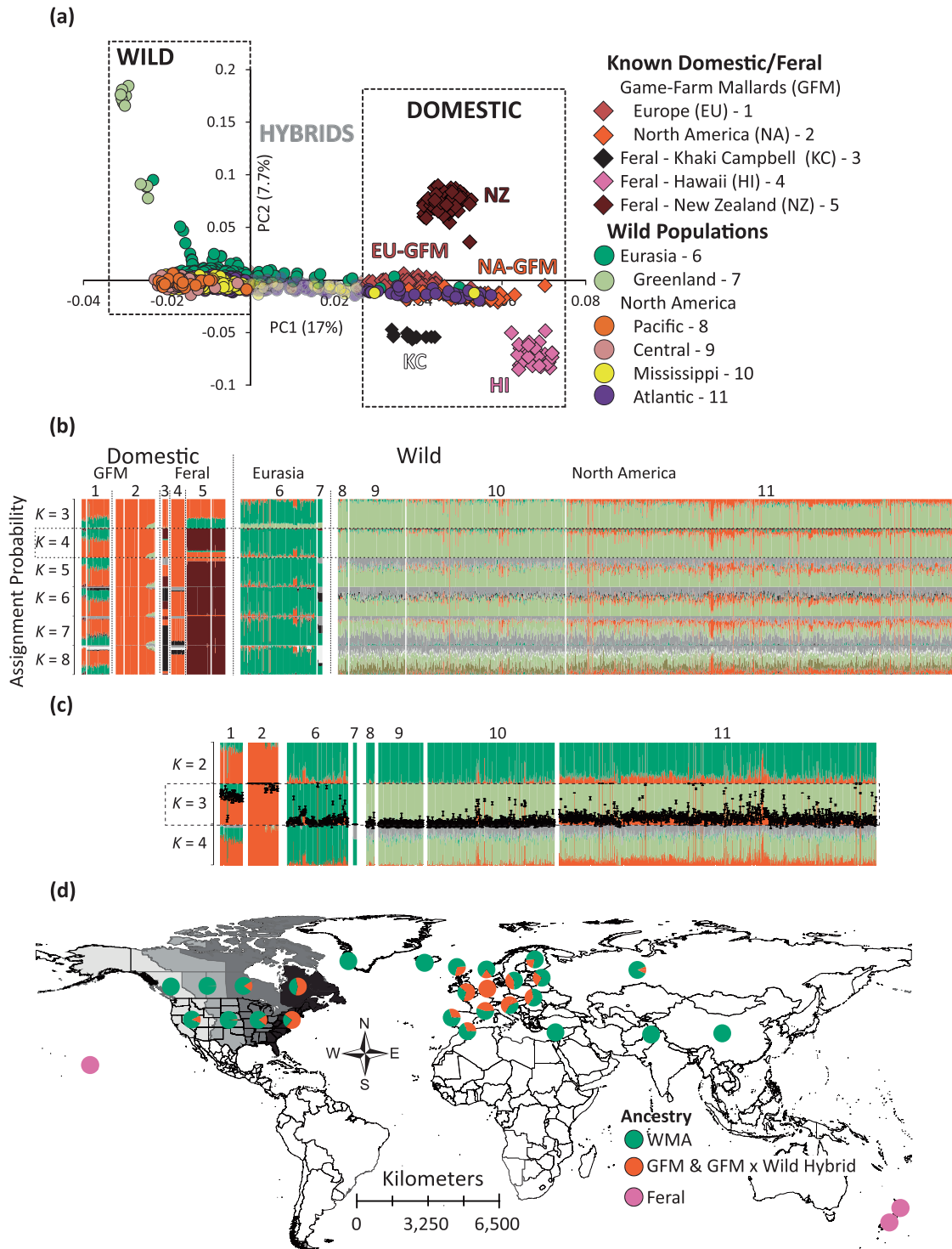
Based on our set ancestry cutoff, only 16 and three North American and Eurasian wild caught samples, respectively, were identified as feral game-farm mallards (Fig. 3d; Supplementary Data 1), representing ~1% of the total sample sets from their respective regions. Plotting the proportion of feral, wild, and hybrid individuals on the landscape established the greatest rates of hybrid prevalence in central Europe and eastern North America, with the proportion of wild ancestry increasing with distance from these areas (Fig. 3d). The westward decline in game-farm  $\times$  wild mallard hybrid prevalence in North America was concordant with mtDNA patterns (Fig. 2b).

Finally, nine groups were analyzed in TREEMIX: (1) three wild mallard genetic clusters including mainland Eurasia, Greenland, and North America, (2) a single domestic game-farm mallard cluster from Eurasia, and two from North America, and (3) feral populations of Khaki Campbell's, as well as those on Hawaii and

New Zealand. An optimum TREEMIX tree and >99% of the variance explained was recovered when including up-to two migration edges (Fig. 4; Supplementary Fig. S3C). Two major clades representing wild or domestically-derived groups were recovered. Among these, the Eurasian wild mallard was recovered as basal to all others and consistent with this being the ancestor<sup>30,31</sup>. Next, Khaki Campbell's were recovered basal to the remaining groups within the domestic clade, which included the Eurasian game-farm mallard basal to North American game-farm mallards and feral populations of Hawaii and New Zealand (Fig. 4). Interestingly, whereas the drift parameter generally increased among groups within the game-farm mallard clade, Khaki Campbell's and Greenland mallards had comparable and highest drift parameters. Finally, the two statistically significant gene flow events included, gene flow from the wild mallard ancestor to the second North American game-farm mallard group (i.e. GFM 2; Fig. 4), and from the basal domestic lineage into New Zealand mallards; both consistent with their respective mixed supplementation histories<sup>53,56,57</sup>.

#### Genomic differentiation and selectively non-neutral loci.

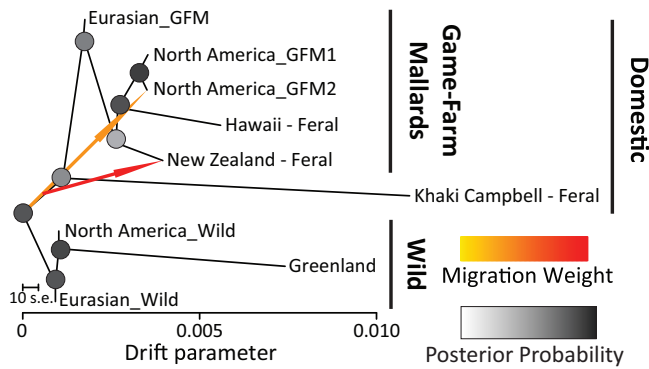
Genomic comparisons were done on the same nine mallard groups as in phylogenetic analyses (see above). Given that we were interested in general differences across groups, we obtained an average and standard deviation ( $\Phi_{ST} \sim 0.071 \pm 0.12$ ) across pair-wise and locus-by-locus comparisons to set an arbitrary two-standard deviation  $\Phi_{ST}$  threshold of 0.31 for outlier detection (Supplementary Fig. S5). First, composite and locus-by-locus  $\Phi_{ST}$  estimates



**Fig. 3 Nuclear population structure analyses.** Nuclear population structure analyses based on 36,637 independent bi-allelic nuclear SNPs assayed across 1916 samples comprising wild mallards, known game-farm mallards, and feral mallard populations, and visualized based on a **(a)** Principal Component Analysis, and with ADMIXTURE assignment probabilities obtained for **(b)** all samples or **(c)** just known game-farm mallards and wild mallards from Eurasia and North America. **d** We then plotted the proportion of samples recovered as wild (WMA), GFM-feral /GFM-feral × WMA, or feral across sampled locations. Note that North America was grouped by the four known migratory flyways that waterfowl use during annual migrations.

recovered little overall genomic differentiation (composite  $\Phi_{ST} \sim 0.004$ ; Supplementary Fig. S5) and no putative outliers when comparing the genomes of wild Eurasian and North American mallards (Fig. 5; Supplementary Fig. S6; Supplementary Data S2). Conversely, Greenland’s wild mallards were most differentiated (composite  $\Phi_{ST} \sim 0.15\text{--}0.20$ ; Supplementary Fig. S5A), including

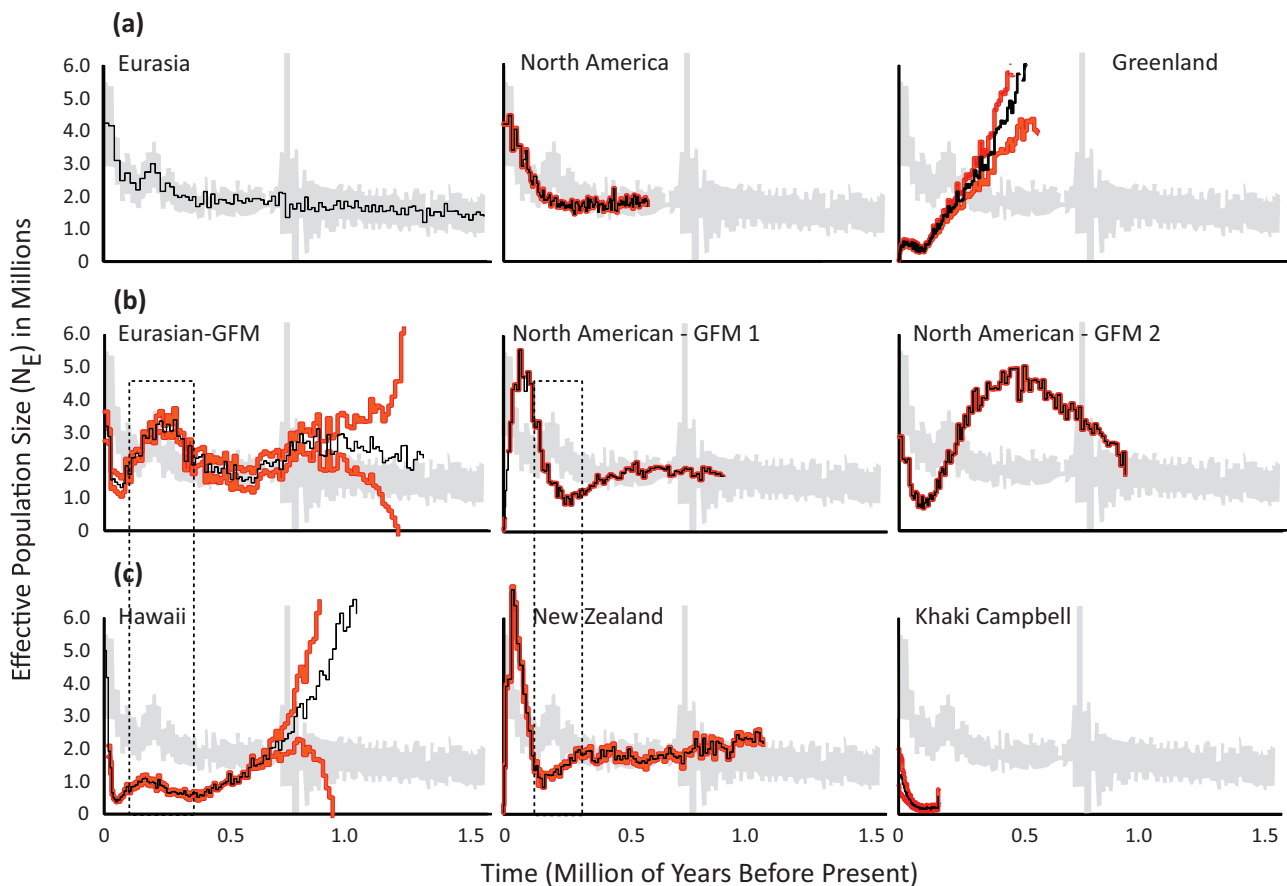
highly elevated composite  $\Phi_{ST}$  estimates with 123 – 670 putative outlier loci across pair-wise comparisons (Supplementary Fig. S6; Supplementary Data S2). While putative outliers were found across the entire genomic landscape of Greenland mallards, 66 loci were recovered across pair-wise comparisons as potential ‘islands of differentiation’<sup>58</sup> (Supplementary Fig. S6; Supplementary Data S2).



**Fig. 4 TREEMIX relationships and gene flow events.** Relationships among sampled wild, game-farm, and feral mallard groups, as well as gene flow events recovered in tree reconstructions as implemented in the program TREEMIX and based on 36,637 independent bi-allelic nuclear SNPs. Grouping of samples were based on respective ADMIXTURE assignment probabilities (Fig. 3; also see Supplementary Data 1). Note that contemporary game-farm × wild hybrids identified in North America and Eurasia were excluded from analyses.

Next, composite  $\Phi_{ST}$  estimates among game-farm mallards recovered little genetic differentiation between the two North American groups (composite  $\Phi_{ST} \sim 0.013$ ), but slightly greater values from Eurasian groups (composite  $\Phi_{ST} \sim 0.03$ ; Supplementary Fig. S5A), including 0–2 putative outliers among pair-wise comparisons (Supplementary Fig. S6; Supplementary Data S2). Finally, when comparing between and within feral lineage, we found fairly consistent composite estimates of relative differentiation respective to each group, with most ranging between 0.05 and 0.20, except for feral Khaki Campbell’s that had consistently high estimates (composite  $\Phi_{ST} > 0.20$ ) across pair-wise comparisons (Supplementary Fig. S5A). The relatively high composite  $\Phi_{ST}$  estimates of Khaki Campbell’s translated across the genome with >10% of markers identified as putative outliers, including 89 loci spanning the Z-sex and 22 autosomal chromosomes (Supplementary Fig. S6; Supplementary Data S2).

**Genetic diversity & demographics.** Similar levels of nucleotide diversity were recovered across groups (average  $\pi$  range = 0.0044–0.0055), except for feral Khaki Campbell’s that had nearly a third of the nucleotide diversity ( $\sim 0.0016$ ; Supplementary Fig. S5B). Despite having similar levels of genetic diversity, we recovered substantially different demographic histories among groups (Fig. 5). First, near identical demographic



**Fig. 5 Per species time-series demographic analyses.** Per species time-series demographic analyses estimated using the single species *∂a*di** model based on 36,637 independent bi-allelic nuclear SNPs across groups of samples that were partitioned based on their respective ADMIXTURE assignment probabilities (Fig. 3; also see Supplementary Data 1). Contemporary game-farm × wild hybrids identified in North America and Eurasia were excluded from analyses. Eight groups were analyzed: (a) wild mallards from Eurasia, North America, and Greenland, (b) known game-farm mallards (GFM) Eurasia and North America, as well as (c) feral Khaki Campbell and mallard populations on Hawaii and New Zealand. For each of the analyzed groups, we overlay the average (black line) and respective 95% confidence intervals (vermillion lines) over Eurasian wild mallard 95% confidence intervals (in gray) for comparative purposes of estimated effective population size. Boxes denote examples of how differing population histories depress (Hawaii), accentuate positively (Eurasian-GFM), or accentuate negatively (North American -GFM 1 & New Zealand) the ancestral signal.

**Table 1 List and description of the 27 environmental and 2 indices of human impact variables used for genotype-environment association testing in GRADIENT FOREST analyses.**

Variable	Description	Unit	Eur.	NA	HI	NZ
BIO1	Mean annual temperature	°C				
<i>BIO2</i>	Mean diurnal range	°C			2	
BIO3	Isothermality	°C				
<i>BIO4</i>	Temperature seasonality	°C	2		3	
BIO5	Max temperature of warmest month	°C				
BIO6	Min temperature of coldest month	°C				
BIO7	Temperature annual range	°C				
<i>BIO8</i>	Mean temperature of wettest quarter	°C		1		3
<i>BIO9</i>	Mean temperature of driest quarter	°C	5			4
BIO10	Mean temperature of warmest quarter	°C				
<i>BIO11</i>	Mean temperature of coldest quarter	°C				5
BIO12	Annual precipitation	mm				
BIO13	Precipitation of wettest month	mm				
BIO14	Precipitation of driest month	mm				
<i>BIO15</i>	Precipitation seasonality	mm	4		1	
BIO16	Precipitation of wettest quarter	mm				
<i>BIO17</i>	Precipitation of driest quarter	mm			4	
<i>BIO18</i>	Precipitation of warmest quarter	mm	1	3		
BIO19	Precipitation of coldest quarter	mm				
EVI_ANNUAL	Enhanced Vegetation Index Annual; MOD13A3					
<i>EVI_SUMMER</i>	Enhanced Vegetation Index June; MOD13A3				5	
<i>EVI_WINTER</i>	Enhanced Vegetation Index December; MOD13A3			5		
NDVI_ANNUAL	Normalized Difference Vegetation Index Annual; MOD13A3					
NDVI_SUMMER	Normalized Difference Vegetation Index June; MOD13A3					
NDVI_WINTER	Normalized Difference Vegetation Index December; MOD13A3					
NPP_ANNUAL	Net Primary Productivity Annual; MOD17A2H					
<i>SRTM</i>	Shuttle radar topography mission; Elevation	m	3			
<i>Human.Footprint</i>	Anthropogenic footprint; NASA SEDAC data			2		1
<i>Human.Index</i>	Anthropogenic index; NASA SEDAC data			4		2

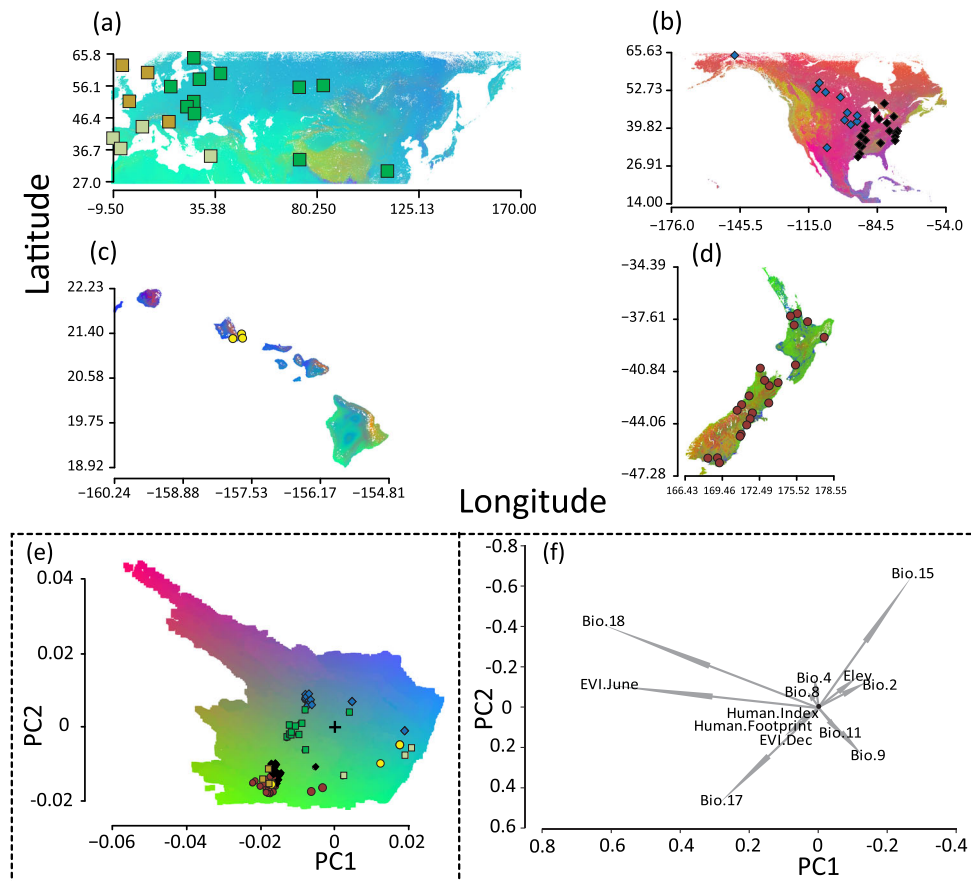
The top five predictive variables based on the cumulative  $R^2$  weighted importance ranking (Supplementary Fig. S7) used in the respective analysis of wild mallards from Eurasia (Eur.) and North America (NA), as well as feral populations on Hawaii (HI) and New Zealand (NZ) are numerically denoted. The 13 italicized variables were used in the combined analysis (Fig. 6e).

histories were recovered for genetically pure wild mallards from North America and Eurasia, which retained an ancestral effective population size of ~1.5 million until ~200,000 years before present when they began to exponentially increase to a contemporary effective population size of ~4.2 million (Fig. 5a). Considering the demographic history of wild mallards as a reference ancestral state, we found the most discordant demographic histories were recovered for Greenland's wild mallards and feral Khaki Campbell's (Fig. 5c). Among game-farm mallards, the Eurasian group and the second North American group (i.e., GFM-2) not only possessed exaggerated cyclical demographic patterns in deeper time, but also had a sudden increase in effective population size at time zero (Fig. 5b). Conversely, the North American game-farm mallard that did not show admixed ancestry (Fig. 2; i.e., GFM-1) had exaggerated patterns in more recent time, including a population crash at time zero. Next, New Zealand's feral mallards had a nearly identical demographic history as North American game-farm mallards, whereas the feral population in Hawaii more closely resembled Greenland's wild mallards until recently where they experienced substantial growth. Interestingly, despite mallard domestication being initiated around 500 BC<sup>30,31</sup>, demographic deviations from the ancestral state suggests deep time divergence across domestic and feral groups.

**Genetic niche modeling.** The mean  $R^2$  value for GF models of wild mainland North American ( $R^2 = 0.10$ ,  $t\text{-test}(62) = 11.19$ ,  $p < 0.001$ ) and Eurasian ( $R^2 = 0.15$ ,  $t\text{-test}(39) = 2.94$ ,  $p < 0.01$ ), as well as feral populations of mallards in New Zealand ( $R^2 = 0.13$ ,  $t\text{-test}(99) = -2.6528$ ,  $p < 0.01$ ) and Hawaii ( $R^2 = 0.47$ ;

$t\text{-test}(261) = 35.44$ ,  $p < 0.0001$ ) all performed better than their respective randomized datasets (Supplementary Fig. S7). For each group, GF outputs were based on the top five environmental factors with highest adjusted  $R^2$  values (Table 1; Supplementary Fig. S7). No environmental variable was recovered as significant across all groups, with all but five variables being specific to a single group (Table 1). However, we note that GF recovered a limited accuracy importance (i.e., the out of bag error for GF; Supplementary Fig. S7) for the top environmental variables, suggesting it was unable to parse minute differences between the effects of individual predictors. While this creates some uncertainty with regards to the influence of individual variables, consistency among the top 10 variables identified in individual models suggests that the overall relationship between the environment and genotypic turnover is being reflected. Therefore, in order to account for a lack of more precise differentiation, we included all 13 unique environmental variables (Table 1) when creating the combined GF model (Fig. 6e). Note that GF models were built without human indices for Hawaii's feral mallards as there was not enough variability among sites for proper analysis (Fig. 6c); however, the combined function of cumulative importance calculated among the other groups were extended to the unknown population sites (i.e., Hawaii; Fig. 6e).

The top five environmental variables for wild mallards in Eurasia included general seasonality in temperature and precipitation, and more specifically, variables associated with average temperature and precipitation during the warmest quarter of the year, and elevation. Although graphing PCs of GF outputs recovered what appeared to be fairly constrained genetic space



**Fig. 6 Mapped genotype-environment associations from GRADIENT FOREST analyses.** Mapped genotype-environment associations from GRADIENT FOREST (GF) based on the top five most predictive environmental and/or human index variables (Table 1; Supplementary Fig. S7) across (a) Eurasia (wild), (b) mainland North America (wild), (c) Hawaii (feral), and (d) New Zealand (feral). Note that GF models are unitless, and changes in color represent expected changes in allele frequency. Finally, a (e) combined GF model overlapping all four groups based on the (f) 11 environmental variables and 2 human indices found to be significant in at least one of the four analyses is provided (Table 1). The + in the combined plot denotes the center placement for the PCA of the top predictor variables for reference. Moreover, note that sampling sites for each of the independent analyses are color coded based on their relationship recovered in the combined plot.

(Supplementary Fig. S7A), three groups were recovered when plotting genotypic turnover across Eurasia (Fig. 6a), including: (1) Faroe Islands, Britain, Norway, and Slovenia, (2) Russia, Ukraine, Finland, Sweden, Estonia, Pakistan, and China, and (3) regions from France and Spain to Cyprus and Portugal. Interestingly, group one were most distinct from the other two Eurasian groups due to their response to precipitation and anthropogenic influences (Fig. 6e, f).

Next, the top five variables for wild mallards collected during autumn and winter months in North America included weather (i.e., mean temperature during the wettest quarter and precipitation levels during the warmest quarter) and the vegetation index during winter, as well as the human-impact indices (Table 1); graphing PCs of GF outputs showed that wild mallards occupied a fairly diverse adaptive space (Supplementary Fig. S7B). Genotypic turnover was recovered in the northern most arctic regions of Canada, western Rockies, southern regions from Mexico to Florida, and many parts east of the Mississippi River, including northeastern coastal habitat (Fig. 6b). More generally, the primary adaptive range reconstructed from wintering wild mallards recapitulated regions stretching southward from the prairie pothole region of central Canada into the USA, and that are known to be critical migratory and wintering regions for the majority of North America's wild mallards<sup>59</sup>. In fact, recovery of temperature, precipitation, and vegetative indices as important variables is consistent with other models identifying

these as key signals used in migratory initiation in North America<sup>60–62</sup>. Conversely, in the combined GF plot, genotypic turnover became clear between eastern and western North America, explained by the positive or negative association with human-impact indices, respectively (Fig. 6).

Among the two feral populations, GF outputs for New Zealand were based on the average temperature ranges tied to wet-dry cycles, as well as both human indices (Table 1; Supplementary Fig. S7D). Projecting genotypic turnover across New Zealand recovered western parts of the North and South Islands at the edge of the adaptive range for these mallards (Fig. 6d). New Zealand's mallards clustered with wild mallards of western Europe and eastern North America in the combined GF model; suggesting these groups to be similarly responding to human indices (Fig. 6e, f). Finally, GF analysis of Hawaii's feral mallards revealed much more reduced genetic space of all groups (Supplementary Fig. S7C), and did not make any predictable clustering in the combined model (Fig. 6e).

## Discussion

**Mallard begets mallard and the proliferation of the game-farm mallard breed.** Applying a landscape genomics approach, we establish that genetically pure North American and Eurasian wild mallards are not only structured at mitochondrial DNA as



expected (Fig. 2;<sup>50,63,64</sup>), but also identify fine nuclear sub-structuring (Fig. 3) previously undetected<sup>65</sup>. Although Greenland mallards have previously been found as genetically unique<sup>64,65</sup>, we report previously undetected genetic structure between wild mallards from the rest of Eurasia and North America that is attributable to small frequency differences found across their genomes (Supplementary Fig. S6)<sup>65</sup>. In general, differential patterns of mito-nuclear structure are explained by known life-history traits of mallards that include strong female breeding phylopatriy and male-biased dispersal<sup>33</sup> that often results in strong mtDNA and weak or no nuclear structuring, respectively<sup>48,66</sup>. Despite the population structure, wild Eurasian and North American mallards were generally similar in genetic diversity, with no significant outlier loci detected, and identical demographic histories (Figs. 2–3 and 5; Supplementary Fig. S6). Nevertheless, the strong structuring of Greenland's mallards suggests that population subdivision has occurred across some parts of the mallard's Holarctic distribution; but this will require additional sampling to confirm whether Greenland's mallards diverged through local selective pressures and/or gone through severe genetic drift via founder events.

Next, we provide further evidence that the wild mallard are the primary ancestor of domesticated breeds (Fig. 4;<sup>67–69</sup>), and that game-farm mallards from Eurasia were the ancestral eve of stocks around the world. Whereas wild mallards show a high degree of mtDNA haplotype diversity, samples collected from feral populations and regions where releases are common contain a much smaller set of unique haplotypes (Fig. 2). For example, in eastern North America, 75% of samples with an OW A haplotype were contained within only 15 unique haplotypes. This creates the kind of star-like pattern we recovered in the network (Fig. 2a), which is consistent with a history of game-farm mallard stocks coming from a single or few common lineage(s)<sup>48</sup>. Moreover, we contend that these results clarify the debate about how OW A haplotypes arrived and spread in North America. Whereas past hypotheses suggested more natural pathways for OW A introduction to North America<sup>50,51,70–72</sup>, the geographic distribution and association with game-farm nuclear ancestry strongly support that it was rather a direct consequence of gene flow with introduced game-farm mallards<sup>16</sup>, and thereby anthropogenically induced.

Among attempts at domestication, mallards have been one of the most successful<sup>73</sup> with over 20 breeds being artificially selected for various uses and functions (Fig. 1). Interestingly, hybridization with domestic breeds, specifically our reference park mallard (i.e., Khaki Campbell's), is extremely limited, as evidenced by the fact that almost no wild individuals shared any nuclear ancestry with them (Fig. 3). It has become clear, rather, that game-farm mallard releases are the primary pathway of introgression in Eurasia and North America and have already led to self-sustaining and potentially locally adapted feral populations in Hawaii and New Zealand (Fig. 6). In New Zealand, despite releases being discontinued ~70 years ago<sup>34–36</sup> the self-sustained feral population is now comparable to wild North American mallards in breeding and survival<sup>74</sup>. Moreover, New Zealand mallards seem to be uniquely adapted to their current environment and landscape (Fig. 6e), making them a conservation paradox as their presence remains a major conservation concern for endemic New Zealand grey ducks (*A. superciliosa superciliosa*), but are potentially better suited to the increasingly agricultural and urban landscapes of this region (Supplementary Fig. S6;<sup>75–77</sup>). Importantly, our results contrast the general consensus among wildlife managers that game-farm mallards rarely survived long enough to leave the area they are released in, because it is evident from our findings that they survived at great enough rates to have their

genetic variation readily found everywhere they were and are currently being released (Figs. 2 and 3;<sup>78</sup>).

Despite opportunities to interact, causes for why domestic × wild mallard hybrids are the result of game-farm mallards and not alternative domestic breeds remains unknown. We posit that it is possible that phenotypic similarities between wild and game-farm mallards may not result in negative assortative mating that may exist between wild mallards and other domestic breeds like Khaki Campbell's (Fig. 1). Moreover, it is possible that the domestication process results in highly variable genomic changes in extent and type (e.g., chromosomal inversions) that may confer reproductive isolation; but which is evidently absent between wild and game-farm mallards. In fact, Khaki Campbell's are highly divergent across ddRAD-seq loci as compared to game-farm mallards when compared to wild mallards (Supplementary Fig. S6). Behavioral work and full genome data will be required to test between these respective hypotheses, and in order to better understand potential mechanism(s) that result in variable levels of interactive potential and/or offspring viability that may exist between wild mallards and their various domesticated breeds.

### Sample identity is critical in demographic reconstructions.

Given that coalescent events dictate demographic histories<sup>79,80</sup>, we demonstrate how severe loss and the influx of novel genetic diversity via genetic drift (e.g., inbreeding, domestication) and gene flow, respectively, can cause deviations in allelic histories from their true ancestral states, resulting in distorted demographic histories<sup>47,81</sup>. First, we find extreme demographic changes as compared to the wild ancestral state to be among highly bottlenecked groups (i.e., Greenland, Khaki Campbell & Hawaii's feral mallards; Fig. 5; also see drift values in Fig. 4). Next, the presence and exaggeration of cyclical demographic patterns in deeper time generally corresponds with extent and recency of gene flow (Fig. 5). For example, whereas all the three analyzed game-farm mallard groups showed variable but exaggerated cyclical patterning in deep time, each group either showed a sudden increase (i.e., Eurasian GFM & GFM-2) or crash (i.e., GFM-1) at time zero (Fig. 5b) that followed whether the group experienced recent wild introgression or not, respectively (Fig. 3). In fact, similar patterns of distortion were found in PSMC analyses of full genomes among domestic and wild mallards<sup>68,69</sup>. For example, Guo et al.<sup>67</sup> inferred domestic lineages to have diverged over 40,000 years ago based on demographic reconstructions, and thus, the domestication of mallards must have been even earlier than expected. However, given that animal husbandry and domestication among human civilizations generally started somewhere between 15,000 and 36,000 years ago with domestication of fowl occurring over the last 5000 years<sup>73,82–84</sup>, we argue that these inferences of demographic history were biased due to distortions created by the severe bottlenecks that occurred among domestic mallard lineages. The same patterning can be observed for game-farm mallards. Although, the true date of mallard propagation remains unknown, raising mallards for sport hunting, and thus the likely rise of the game-farm mallard, was first reported in 1631 England for King Charles II<sup>85</sup>, and the use of game-farm mallard first reported in breeding and ringing operations in 1890s England<sup>86</sup>. Consequently, our demographic results suggesting deep time divergence between game-farm and wild mallards (Fig. 4) is clearly inaccurate based on records indicating that this breed, as it occurs today, ranges from 150 to 400 years old. Overall, we warn that future studies interested in reconstructing demographic histories, which are often used in conservation assessments and decision making, need to be cautious about the individuals they select to reduce bias, especially in

lineages that are naturally inbred, domestic, or contain highly admixed individuals<sup>47,81</sup>.

**The conservation paradox of feral populations and the meaning of wild.** Manipulating biodiversity through land use change or various human interactions plays an important role across many aspects of society including agriculture, urban development, wildlife management, and nearly all forms of conservation<sup>87,88</sup>; animal domestication in particular, has been essential for the rise of many societies<sup>82–84</sup>. However, the traits favored by artificial selection during domestication can often be deleterious in wild settings, thereby reducing the adaptive potential, fitness, and overall survival of a wild population<sup>8,89–94</sup>. Unfortunately, wild × feral interactions have become more common, and while generally these lead to negative outcomes for the native wild taxa, in some cases, feral populations are able to take advantage and thrive in increasingly urbanized landscapes<sup>29</sup>. Such scenarios could call into question long held conservation priorities; as in, if we are working to preserve biodiversity, should we begin to embrace the potential for introduced feral taxa to thrive in more urban and agricultural habitats, and if so, would this come at the cost or addition to locally adaptive wild diversity? In particular, feral taxa that are already in the focus of conservation efforts that represent ancient lineages with relevant genetic or cultural importance may be perfect cases of naturalization, with the potential of being awarded recognition and protection<sup>95</sup>.

Feral populations are innately different from native wild ones due to their history of artificial selection, and thus, the probability of a group of domestic individuals becoming feral is dependent on their response to the selective pressures of their new environments<sup>27</sup>. Towards this end, the history of domestic mallard releases and established feral populations provides natural experiments that can be used to advance our understanding of ‘feralization’ as a process, as well as species establishment and expansion. First, we determined that the genetic niche space of New Zealand’s feral mallards, as well as wild mallards in Europe and those in eastern North America are explained by precipitation levels during drought years, winter vegetation, and are responding to human disturbances unlike other groups (Supplementary Fig. S7). In fact, maps of genotypic turnover for each of these three groups recapitulate regions in the world with a strong human footprint (Fig. 6a, b, d;<sup>26</sup>). Conversely, the genetic adaptive niche space of wild mallards in eastern Eurasia and western North America, as well as those from France, Spain, and Portugal are largely responding to temperature and precipitation regimes of their regions (Fig. 6). Moreover, we find that within feral populations, New Zealand mallards occupy a unique adaptive space, which could be the result of extensive introgression from locally adapted grey ducks<sup>36</sup>. Overall, these results demonstrate, that in the Anthropocene, the definition of wild may not be singular, and that feral populations are capable of responding relatively quickly to human-induced and natural ecological changes.

While we recovered 1141 and 1360 outlier loci from pair-wise  $\Phi_{ST}$  and environment association, respectively (Supplementary Figs. S6 and S7), there was a general lack of consistency across analyses; though we found concordance among 23% (750 loci) of recovered putative outlier loci among analyses (Supplementary Data 2). Nevertheless, the lack in  $\Phi_{ST}$  outlier consistency across respective pair-wise comparisons for wild Eurasian, wild North American, and New Zealand’s feral mallards suggests that experienced selective pressures from their respective niche spaces are differentially impacting parts of their respective genomes. In fact, ~57% of loci found to be significantly associated in GF models were group-specific, suggesting that loci are being independently affected by distinct selective pressures among groups. A lack of overlapping candidate loci is unsurprising given

the nature of our ddRAD-seq dataset, which is limited to non-coding regions, meaning that we are only able to detect the effects of selection through hitchhiking<sup>96–98</sup>. Thus, while we find many loci of putatively small effect, identifying a common genetic underpinning of adaptation among these groups will require full genome data and more information regarding protein functions and their response to these various selective pressures<sup>99,100</sup>.

## Conclusions

Going forward, conservation science will need new approaches and practices when contending with challenges presented by climate change and our growing human footprint<sup>100,101</sup>. Mallard introductions in New Zealand and Hawaii provide evidence that feral populations can stabilize and become self-sustaining, as the supplementation and introduction of domestic variants allows adaptive alleles to become prevalent once natural selection becomes a more dominant force. However, it is possible that these feral populations are the product of favorable island conditions where predation and seasonal changes are limited. Thus, we posit that the increasing frequency of game-farm mallard ancestry among wild continental mallards is concerning given the more challenging aspects of migration, predation, and severe seasonal changes. Together, these results beg the question of whether currently declining mallard populations of Eurasia and North America are due to the influx of maladaptive traits from released game-farm mallards, or simply the inability of feral and feral admixed individuals to properly adapt to ecosystems far more complex than captive settings. Determining how domestication changed the morphology, behavior, and/or ecology of game-farm mallards will be required to test whether introgression of such traits is maladaptive for wild populations in ways that result in declining population parameters (e.g., survival and fecundity) and sizes. Regardless, the rising prevalence of feral and non-native populations is becoming a major conservation concern, including being the primary causes of biotic homogenization<sup>102,103</sup>, especially as more endemic species become imperiled in the Anthropocene<sup>29</sup>. Therefore, interaction of wild and domestic animals, as well as the potential establishment of feral or feral admixed populations, require continued monitoring due to the dynamic nature of environmental selective pressures, urbanization, and climate change.

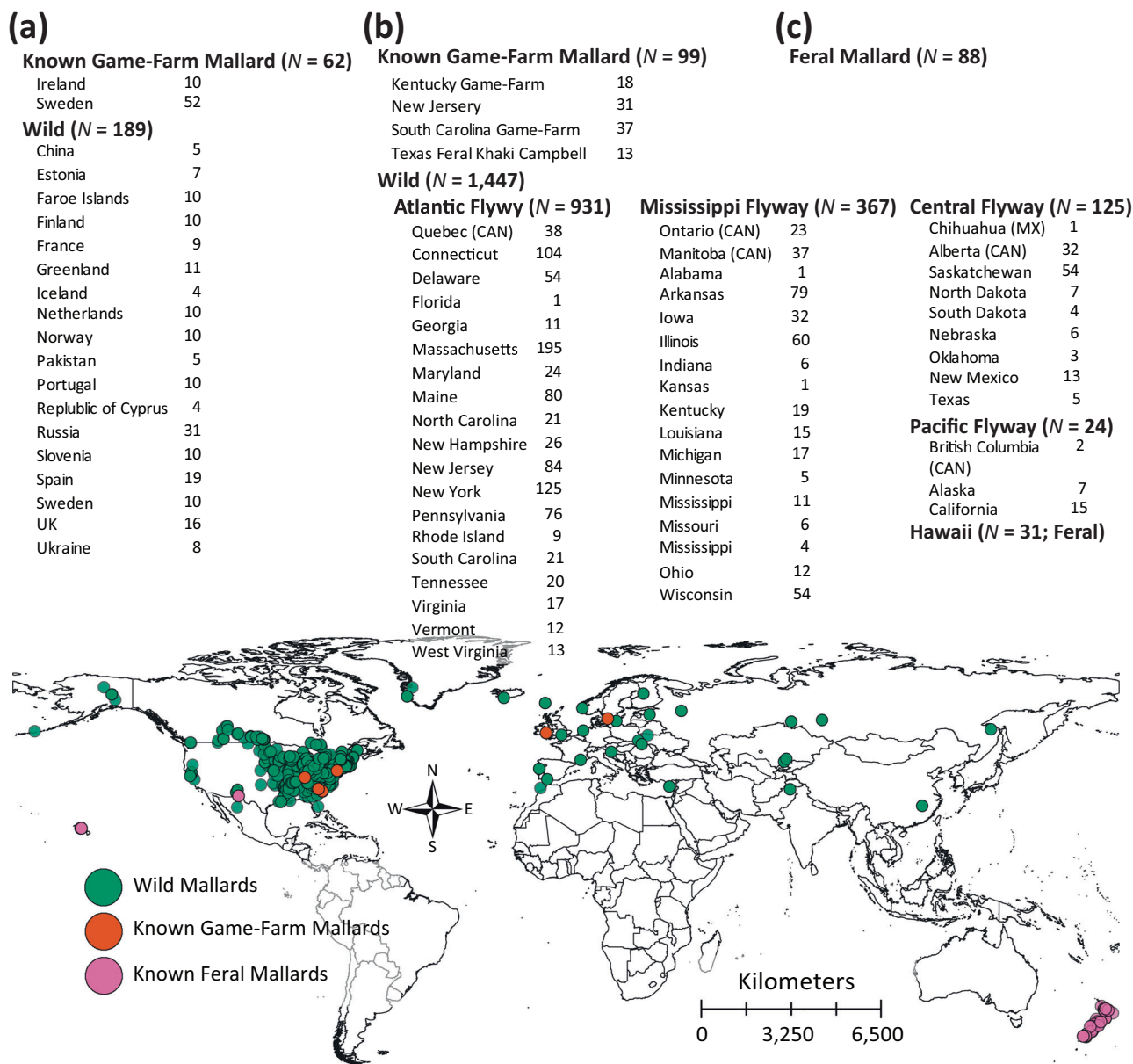
## Methods

**Sampling & DNA extraction.** Tissue, blood, DNA, or comparable published sequences were obtained for a total of 1916 samples of wild and feral mallard populations representing their ranges in mainland North America, Eurasia, Hawaii, and New Zealand (Fig. 7; Supplementary Data 1). In addition, known domestic stocks of game-farm mallards were sampled from two and three preserves in Eurasia and North America, respectively. Finally, several feral Khaki Campbell ( $N = 13$ ; also see Fig. 1) mallards were also included and served as a proxy of alternative domestic park mallards.

We extracted genomic DNA from blood or tissue using a DNeasy Blood & Tissue kit following the manufacturer’s protocols (Qiagen, Valencia, CA, USA). DNA integrity was based on the presence of a high-molecular weight band as determined with gel electrophoresis using a 1% agarose gel<sup>104</sup>.

**Mitochondrial DNA sequencing & analysis.** We PCR amplified and Sanger sequenced the mitochondrial control region (mtDNA) across samples<sup>105,106</sup> and following protocols outlined in Lavretsky et al.<sup>107</sup> (also see Supplementary Methods for details). We aligned and edited previously published data for 739 samples and new sequences using SEQUENCHER v. 4.8 (Gene Codes Corporation, Ann Arbor, MI, USA). All new sequences are deposited in GenBank (Accession number: OR089157-OR090779; See Supplementary Data 1 for sample specific information).

Mallards are characterized by the old world (OW) A and new world (NW) B mtDNA haplogroups, which distinguish individuals of Eurasian or North American descent, respectively<sup>48,49,108</sup>. Importantly, being of Eurasian descent, all domestically-derived mallards carry OW A haplotypes, and thus, are a distinguishing marker when assessing whether game-farm mallard introgression likely occurred within a wild mallard lineage in North America<sup>43,109</sup>. Thus, samples



**Fig. 7 Total and geographical locations of samples.** Size and geographical locations of samples obtained across (a) Eurasian countries, (b) North American states, and (c) both New Zealand Islands. Wild mallards, known game-farm mallards, and three feral populations are denoted by color on the map.

were visualized and sorted as possessing OW A versus NW B haplogroups using a median-joining network carried out with the program POPART<sup>110</sup>. Note that although no comparable sequences were available for the same mallards from Hawaii, these were previously shown to possess OW A haplotypes<sup>20,111</sup>. Finally, we plotted proportion of samples with OW A versus NW B haplotypes by European country or North American waterfowl flyway region that generally defines migratory pathways in ARCMAP v. 10.7.1 (Esri).

**ddRAD-seq library preparation & bioinformatics.** We followed ddRAD-seq library protocols using *SbfI* and *EcoRI* restriction enzymes, and as outlined in DaCosta and Sorenson<sup>112</sup> (also see ref. <sup>96</sup>). Size selection followed gel electrophoresis protocols as outlined in DaCosta and Sorenson<sup>112</sup> for samples collected before 2016, with samples collected thereafter size selected following optimized double-sided magnetic bead selection as described in Hernandez et al.<sup>113</sup>. Because samples were collected over the last decade, multiplexed libraries have been sent to various genomic facilities for 150 base-pair, single-end chemistry sequencing across Illumina platforms including, HiSeq 2000, HiSeq 2500, HiSeq 4000, HiSeq X, and Novoseq. Additional details on DNA extraction and ddRAD-seq library preparation protocols can be found in Supplementary Methods. All raw Illumina sequences are available from the National Center for Biotechnology Information

Sequence Read Archive (BioProject accession numbers. PRJNA980669, PRJNA516035<sup>16</sup>, PRJNA911832<sup>114</sup>, PRJNA800412<sup>115</sup>, PRJNA591912<sup>43</sup>, PRJNA847792<sup>47</sup>, PRJNA745366<sup>53</sup>, PRJNA577581<sup>20</sup>, PRJNA718623<sup>116</sup>; Sample specific accession numbers can be found in Supplementary Data 1). In all cases, raw Illumina reads were de-multiplexed using the *ddRADparser.py* script of the BU ddRAD-seq pipeline<sup>112</sup> based on perfect barcode/index matches.

Across 1916 samples, we used custom in-house Python scripts (Python scripts available at <https://github.com/jonmohl/PopGen>; see<sup>43</sup>) to automate sequence filtering, alignment, and genotyping using a combination of TRIMMOMATIC<sup>117</sup>, BURROWS WHEELER ALIGNER v. 07.15 bwa<sup>118</sup>, and SAMTOOLS v. 1.7<sup>117</sup> (detailed methods can be found in Supplementary Methods). Note that reads were aligned to a chromosomal-level reference wild mallard genome<sup>119</sup>. We further filtered VCF files for any base-pair missing >5% of samples that also included a minimum base-pair depth of 5X (i.e., 10X per genotype), and quality per base PHRED scores of ≥30 using VCFTOOLS v. 0.1.15<sup>120</sup>.

Finally, sex was assigned to each sample based on differences in sequencing depth across autosomal and sex chromosome-linked loci<sup>121</sup>. Specifically, for the homogametic sex (i.e., males = ZZ), we expect to find near-zero levels of sequencing depth across W-sex chromosome linked loci but near equal depth for Z-sex chromosome linked loci when compared to autosomal loci. For the heterogametic sex (i.e., females = ZW), we expect to recover about half the

sequencing depth at both W- and Z-sex chromosome linked loci as compared to autosomal loci.

**Population structure & phylogenetics.** All nuclear population structure was based on independent bi-allelic ddRAD-seq autosomal single nucleotide polymorphisms (SNPs), and without using a priori assignment of individuals to populations or species. The final dataset was obtained by using VCFTOOLS v. 0.1.15<sup>120</sup> to first extract bi-allelic SNPs, and then PLINK v. 1.9<sup>122</sup> to filter for singletons, any SNP missing  $\geq 5\%$  of data across samples or in linkage disequilibrium (LD) (detailed methods can be found in Supplementary Methods).

Population structure was first visualized with a Principal Components Analysis (PCA) as implemented in PLINK v. 1.9<sup>122</sup>. Next, assignment probabilities (Q values) were estimated with the program ADMIXTURE v. 1.3<sup>123–125</sup>. In addition to obtaining average Q values across 100 iterations, standard errors for each analysis were based on 100 bootstrap replicates for each evaluated K population model. We evaluated whether the averaged Q score and respective standard error overlapped  $\geq 98\%$  population assignment that we considered to represent a genetically pure parental, whereas those individuals assigned to multiple genetic clusters determined to be as hybrids<sup>16</sup>. Detailed methods can be found in Supplementary Methods. We plotted the proportion of samples determined as hybrid or wild by European country and North American waterfowl flyway in ARCMAP v. 10.7.1 (Esri).

Next, to reduce the effects of contemporary gene flow events on inferences, we excluded any hybrid identified in population structure analyses (i.e., contemporary hybrids) when estimating relative differentiation ( $\Phi_{ST}$ ), nucleotide diversity, and reconstructing phylogenetic relationships of major wild and domestic mallard lineages. To do so, we used recovered genetic clusters from the above population genetics analyses to categorize samples, demarcating those samples that represent parental versus contemporary hybrids. Once done, we estimated pair-wise population relative differentiation ( $\Phi_{ST}$ ) and per population nucleotide diversity across ddRAD-seq loci using the POPGENOME package in the program R<sup>126</sup>. In addition to composite pair-wise population estimates using concatenated autosomal and Z-chromosome datasets, per-locus estimates of relative differentiation were also estimated across ddRAD-seq loci. A two-standard deviation threshold from the average of  $\Phi_{ST}$  estimates obtained across pair-wise and locus-by-locus comparisons was used for outlier detection. Although we used a generally arbitrary cutoff, doing so still allowed us to investigate our primary interest of whether the same or different loci showed elevated estimates of relative differentiation across pair-wise comparisons<sup>127,128</sup>.

Finally, the program TREEMIX version 1.12<sup>129</sup> was used to reconstruct and compare evolutionary histories among major wild and domestic mallard lineages. In addition to reconstructing relationships, historical gene flow was also inferred in TREEMIX. In short, TREEMIX simultaneously estimates a maximum likelihood (ML) species tree, along with the direction and weight ( $w$ ) of gene flow among taxa that best explains analyzed allele frequencies among groups. The optimum number of migration edges was determined by sequentially adding migration events up to 36 ( $-m$  0–36), and then evaluating the proportion of the variance explained by each migration model. Standard errors ( $-se$ ) were calculated to assess significance among recovered migration edges. In order to limit overconfidence in the tree model, migration edges were added until  $>99\%$  of the variance in the tree model was explained. Finally, likelihood ratios were calculated using likelihood estimates to assess significance between possible tree models.

**Demographic analyses.** Long-term demographic histories of each mallard population was estimated following the approach of Hernandez et al.<sup>113</sup>, which uses  $\partial\text{AdI}$  to model changes in effective population size through time based on multi-individual partial genome data. Each species' SFS was folded and masked before being projected down to account for missing data between groups<sup>113,130,131</sup>. We then estimated confidence intervals (CI) using parameter uncertainty metrics included in  $\partial\text{AdI}$ <sup>131,132</sup>. Briefly,  $\partial\text{AdI}$  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. We maximized the number of parameters for which  $\partial\text{AdI}$  is able to return a true estimate of uncertainty by calculating uncertainty across a range of step sizes ( $\epsilon = 10^{-2}$ – $10^{-9}$ <sup>132,133</sup>; also see detailed methods on uncertainty metrics<sup>113</sup>).  $\partial\text{AdI}$  parameters were then transformed into biologically informative values (see Supplementary Methods for details).

**Genotype-environment association modeling with gradient forest.** We obtained high resolution (i.e.,  $\sim 1$  km<sup>2</sup>) global environmental data from several public databases, with a focus on 27 annual and seasonal environmental variables shown to have impacts on bird physiology and ecology (Table 1;<sup>47,134</sup>). Specifically, we included 19 climate variables from the worldclim version 1.4 database (<https://www.worldclim.org/version1>;<sup>134</sup>); Landsat Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI) and Net Primary Productivity (NPP) data from the USGS AppEEARS database (<https://lpdaacsvs.cr.usgs.gov/appeears>); and elevation data from the Global Land Cover Facility (<http://www.landcover.org/>). To test the importance of human disturbance on genetic diversity within wild versus feral populations, we additionally downloaded data from the

Human Impact Wildlife Conservation Society<sup>135</sup> and the Human Footprint Wildlife Conservation Society<sup>136</sup> indices.

Game-farm admixed individuals were excluded from these analyses based on ancestry estimates from ADMIXTURE; additionally, only winter collected samples were used for Eurasia and North America (Hawaii and New Zealand mallards are non-migratory). Following the approach of Bay et al. (also see refs. 47,137), we used a genotype-environment association analysis as implemented in the R package GRADIENTFOREST (GF)<sup>46</sup>. GF analysis was originally created to detect effects of environmental predictor variables on species turnover across a landscape<sup>46</sup>, but has since been adapted for identifying and modeling changes in allele frequency<sup>47,137,138</sup>. In short, we visualized genetic-environmental associations for each of the populations<sup>138</sup>, as well as used the combined GRADIENTFOREST function to determine differences in genetic niche space among the groups (see ref. 47; also see Supplementary Methods for details).

**Statistics and reproducibility.** Statistical analyses were performed in either software packages or in R, with run-ready input files deposited in figshare (accession <https://doi.org/10.6084/m9.figshare.23396225>). This study included a total of 1916 samples, but various subsets of this total were selected to answer more specific questions (also see Supplementary Methods for details).

**Reporting summary.** Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

## Data availability

Mitochondrial DNA sequences are deposited in GenBank (Sample specific accession numbers can be found in Supplementary Data 1). All raw Illumina sequences are available from the National Center for Biotechnology Information Sequence Read Archive (BioProject accession numbers. PRJNA980669, PRJNA516035, PRJNA911832, PRJNA800412, PRJNA591912, PRJNA847792, PRJNA745366, PRJNA577581, PRJNA718623; Sample specific accession numbers can be found in Supplementary Data 1). All other source data (e.g. FASTA (Fig. 2), ADMIXTURE input (Fig. 3), TREEMIX input (Fig. 4),  $\partial\text{AdI}$  input (Fig. 5), and GRADIENT FOREST input (Fig. 6) files) used for all figures and tables are also available on Figshare (accession <https://doi.org/10.6084/m9.figshare.23396225>). All other data are available from the corresponding author (or other sources, as applicable) on reasonable request.

Received: 25 October 2022; Accepted: 24 July 2023;

Published online: 05 August 2023

## References

- Grabenstein, K. C. & Taylor, S. A. Breaking barriers: causes, consequences, and experimental utility of human-mediated hybridization. *Trends Ecol. Evol.* **33**, 198–212 (2018).
- Crispo, E., Moore, J. S., Lee-Yaw, J. A., Gray, S. M. & Haller, B. C. Broken barriers: human-induced changes to gene flow and introgression in animals. *BioEssays* **33**, 508–518 (2011).
- Frankham, R. Genetics and extinction. *Biol. Conserv.* **126**, 131–140 (2005).
- Araki, H., Cooper, B. & Blouin, M. S. Carry-over effect of captive breeding reduces reproductive fitness of wild-born descendants in the wild. *Biol. Lett.* **5**, 621–624 (2009).
- Laikre, L. & Ryman, N. Effects on intraspecific biodiversity from harvesting and enhancing natural populations. *Ambio* **25**, 504–509 (1996).
- Lichatowich, J. & Lichatowich, J. A. *Salmon without rivers: a history of the Pacific salmon crisis* (Island Press, 2001).
- Waples, R. S. & Drake, J. Risk/benefit considerations for marine stock enhancement: a Pacific salmon perspective. In: *Stock Enhancement and Sea Ranching: Developments, Pitfalls and Opportunities*, 2nd ed., 260–306 (Wiley, 2004).
- Champagnon, J., Elmer, J., Guillemain, M., Gauthier-Clerc, M. & Lebreton, J.-D. Conspecifics can be aliens too: a review of effects of restocking practices in vertebrates. *J. Nat. Conserv.* **20**, 231–241 (2012).
- Wang, H., Vieira, F. G., Crawford, J. E., Chu, C. & Nielsen, R. Asian wild rice is a hybrid swarm with extensive gene flow and feralization from domesticated rice. *Genome Res.* **27**, 1029–1038 (2017).
- Leitwein, M., Gagnaire, P. A., Desmarais, E., Berrebi, P. & Guinand, B. Genomic consequences of a recent three-way admixture in supplemented wild brown trout populations revealed by local ancestry tracts. *Mol. Ecol.* **27**, 3466–3483 (2018).
- Bolstad, G. H. et al. Gene flow from domesticated escapes alters the life history of wild Atlantic salmon. *Nat. Ecol. Evol.* **1**, 0124 (2017).
- Canu, A., Iacolina, L., Apollonio, M. & Scandura, M. Lack of polymorphism at the MC1R wild-type allele and evidence of domestic allele introgression

- across European wild boar populations. *Mammalian Biology* **81**, 477–479 (2016).
13. Puigcerver, M. et al. Decreased fitness of restocked hybrid quails prevents fast admixture with wild European quails. *Biol. Conserv.* **171**, 74–81 (2014).
  14. Heikkinen, M. E. et al. Long-term reciprocal gene flow in wild and domestic geese reveals complex domestication history. *G3: Genes, Genomes, Genetics* **10**, 3061–3070 (2020).
  15. Ottenburghs, J. Multispecies hybridization in birds. *Avian Res.* **10**, 20 (2019).
  16. Lavretsky, P., Janzen, T. & McCracken, K. G. Identifying hybrids & the genomics of hybridization: Mallards & American black ducks of eastern North America. *Ecol. Evol.* **9**, 3470–3490 (2019).
  17. Söderquist, P. et al. Admixture between released and wild game birds: a changing genetic landscape in European mallards (*Anas platyrhynchos*). *Eur. J. Wildl. Res.* **63**, 98 (2017).
  18. Rueness, E. K. 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, 23. (2017).
  19. Söderquist, P., Norrström, J., Elmberg, J., Guillemain, M. & Gunnarsson, G. Wild mallards have more “goose-like” bills than their ancestors: a case of anthropogenic influence? *PLoS One* **9**, e115143 (2014).
  20. Wells, C. P. et al. Persistence of an endangered island endemic, an introduced congener, and multiple hybrid swarms across the main Hawaiian Islands. *Mol. Ecol.* **28**, 5203–5216 (2019).
  21. Mattucci, F. et al. Genomic approaches to identify hybrids and estimate admixture times in European wildcat populations. *Sci. Rep.* **9**, 1–15 (2019).
  22. Pilot, M. et al. Widespread, long-term admixture between grey wolves and domestic dogs across Eurasia and its implications for the conservation status of hybrids. *Evolut. Appl.* **11**, 662–680 (2018).
  23. Leonard, J. A., Echegaray, J., Randi, E., Vilà, C. & Gompper, M. Impact of hybridization with domestic dogs on the conservation of wild canids. In *Free-ranging Dogs and Wildlife Conservation*, 170–184 (2014).
  24. Forcina, G. et al. Genome-wide markers redeem the lost identity of a heavily managed gamebird. *Proc. R. Soc. B* **288**, 20210285 (2021).
  25. Tufto, J. Domestication and fitness in the wild: a multivariate view. *Evolution* **71**, 2262–2270 (2017).
  26. Venter, O. et al. Sixteen years of change in the global terrestrial human footprint and implications for biodiversity conservation. *Nat. Commun.* **7**, 1–11 (2016).
  27. Gering, E., Incorvaia, D., Henriksen, R., Wright, D. & Getty, T. Maladaptation in feral and domesticated animals. *Evolut. Appl.* **12**, 1274–1286 (2019).
  28. Makino, T. et al. Elevated proportions of deleterious genetic variation in domestic animals and plants. *Genome Biol. Evol.* **10**, 276–290 (2018).
  29. Hirashiki, C., Kareiva, P. & Marvier, M. Concern over hybridization risks should not preclude conservation interventions. *Conserv. Sci. Pract.* **3**, e424 (2021).
  30. Kiple, K. F. *The Cambridge World History of Food*. 2. Vol. 2 (Cambridge University Press, 2001).
  31. Huang, Y. et al. The duck genome and transcriptome provide insight into an avian influenza virus reservoir species. *Nat. Genet.* **45**, 776–783 (2013).
  32. Heusmann, H. W. The history and status of the Mallard in the Atlantic Flyway. *Wildl. Soc. Bull.* **19**, 14–22 (1991).
  33. Baldassarre, G. *Ducks, Geese, and Swans of North America* (Johns Hopkins University Press, 2014).
  34. Hitchmough, R. A., Williams, M. & Daugherty, C. H. A genetic analysis of Mallards, Grey Ducks, and their hybrids in New Zealand. *N. Z. J. Zool.* **17**, 467–472 (1990).
  35. Williams, M. The changing relative abundance of grey duck (*Anas superciliosa*) and mallard (*A. platyrhynchos*) in New Zealand. *NOTORNIS* **64**, 211–228 (2017).
  36. Brown, J. I. et al. A multi-locus approach to discern conservation units and adaptive potential of Pacific black ducks across Australia and surrounding islands. *Emu-Austral Ornithol.* **121**, 124–135 (2021).
  37. Guay, P. J. & Tracey, J. P. Feral Mallards: a risk for hybridisation with wild Pacific black ducks in Australia? *Vic. Naturalist* **126**, 87–91 (2009).
  38. Madden, J. R. How many gamebirds are released in the UK each year? *Eur. J. Wildl. Res.* **67**, 1–14 (2021).
  39. Champagnon, J., Gauthier-Clerc, M., Lebreton, J., Mouronval, J. & Guillemain, M. Les canards colverts lâchés pour la chasse interagissent-ils avec les populations sauvages. *Faune Sauvage* **298**, 4–9 (2013).
  40. Lincoln, F. C. Restocking of marshes with hand-reared mallards not proved practicable. *Yearb Agric.* 310–313 (1934).
  41. U.S. Fish and Wildlife Service (USFWS). Review of captive-reared Mallard regulations on shooting preserves. Final Report, U.S. Fish and Wildlife Service, Office of Migratory Bird Management, Washington, D.C., USA. (2013)
  42. Heusmann, H. W. Mallard-black duck relationships in the Northeast. *Wildl. Soc. Bull.* **2**, 171–177 (1974).
  43. Lavretsky, P. et al. Assessing changes in genomic divergence following a century of human mediated secondary contact among wild and captive-bred ducks. *Mol. Ecol.* **29**, 578–595 (2020).
  44. Champagnon, J. et al. Assessing the genetic impact of massive restocking on wild mallard. *Anim. Conserv.* **16**, 295–305 (2013).
  45. Fitzpatrick, M. C., Chhatre, V. E., Soolanayakanahally, R. Y. & Keller, S. R. Experimental support for genomic prediction of climate maladaptation using the machine learning approach Gradient Forests. *Mol. Ecol. Resour.* **21**, 2749–2765 (2021).
  46. Ellis, N., Smith, S. J. & Pitcher, C. R. Gradient forests: calculating importance gradients on physical predictors. *Ecology* **93**, 156–168 (2012).
  47. Brown, J. I., Harrigan, R. J. & Lavretsky, P. Evolutionary and ecological drivers of local adaptation and speciation in a North American Avian Species Complex. *Mol. Ecol.* **31**, 2578–2593 (2022).
  48. Lavretsky, P., Hernández Baños, B. E. & Peters, J. L. Rapid radiation and hybridization contribute to weak differentiation and hinder phylogenetic inferences in the new world mallard complex (*Anas* spp.). *Auk* **131**, 524–538 (2014).
  49. Avise, J. C., Ankney, D. C. & Nelson, W. S. Mitochondrial gene trees and the evolutionary relationship of mallard and black ducks. *Evolution* **44**, 1109–1119 (1990).
  50. Kulikova, I. V. et al. Phylogeography of the Mallard (*Anas platyrhynchos*): hybridization, dispersal, and lineage sorting contribute to complex geographic structure. *Auk* **122**, 949–965 (2005).
  51. Kulikova, I. V., Zhuravlev, Y. N. & McCracken, K. G. Asymmetric hybridization and sex-biased gene flow between eastern spot-billed ducks (*Anas zonorhynchos*) and mallards (*A. platyrhynchos*) in the Russian Far East. *Auk* **121**, 930–949 (2004).
  52. Čížková, D., Javůrková, V., Champagnon, J. & Kreisinger, J. Duck’s not dead: does restocking with captive bred individuals affect the genetic integrity of wild mallard (*Anas platyrhynchos*) population? *Biol. Conserv.* **152**, 231–240 (2012).
  53. Lavretsky, P., Duenez, E., Kneee, M. & Kaminski, R. M. Population genetics of a translocated population of mottled ducks and allies. *J. Wildl. Manag.* **85**, 1616–1627 (2021).
  54. Davis, J. B., Outlaw, D. C., Ringelman, K. M., Kaminski, R. M. & Lavretsky, P. Low levels of hybridization between domestic and wild Mallards wintering in the lower Mississippi Flyway. *Ornithology* **139**, ukac034 (2022).
  55. Lawson, D. M., Williams, C. K., Lavretsky, P., Howell, D. L. & Fuller, J. C. Mallard-black duck hybridization and population genetic structure in North Carolina. *J. Wildl. Manag.* **85**, 1243–1255 (2021).
  56. Dyer, J. & Williams, M. An introduction most determined: Mallard (*Anas platyrhynchos*) to New Zealand. *Notornis* **57**, 178–195 (2010).
  57. Williams, M. *The Duckshooter’s Bag: An Understanding of New Zealand’s Wetland Gamebirds* (Wetland Press, 1981).
  58. Nosil, P. & Schluter, D. The genes underlying the process of speciation. *Trends Ecol. Evol.* **26**, 160–167 (2011).
  59. Lincoln, F. C. The waterfowl flyways of North America. *USDA Circ.* **342**, 12 (1935).
  60. Notaro, M., Lorenz, D., Hoving, C. & Schummer, M. Twenty-first-century projections of snowfall and winter severity across central-eastern North America. *J. Clim.* **27**, 6526–6550 (2014).
  61. Schummer, M. L., Kaminski, R. M., Raedeke, A. H. & Graber, D. A. Weather-related indices of autumn–winter dabbling duck abundance in middle North America. *J. Wildl. Manag.* **74**, 94–101 (2010).
  62. Schummer, M. L., Coluccy, J. M., Mitchell, M. & Van Den Elsen, L. Long-term trends in weather severity indices for dabbling ducks in eastern North America. *Wildl. Soc. Bull.* **41**, 615–623 (2017).
  63. Peters, J. L. et al. Mito-nuclear discord in six congeneric lineages of Holarctic ducks (genus *Anas*). *Mol. Ecol.* **23**, 2961–2974 (2014).
  64. Kraus, R. H. et al. Evolution and connectivity in the world-wide migration system of the mallard: inferences from mitochondrial DNA. *BMC Genet.* **12**, 1–15 (2011).
  65. Kraus, R. H. et al. Global lack of flyway structure in a cosmopolitan bird revealed by a genome wide survey of single nucleotide polymorphisms. *Mol. Ecol.* **22**, 41–55 (2013).
  66. Sonsthagen, S. A., Wilson, R. E., Lavretsky, P. & Talbot, S. L. Coast to coast: high genomic connectivity in North American scoters. *Ecol. Evol.* **9**, 7246–7261 (2019).
  67. Guo, X. et al. Revisiting the evolutionary history of domestic and wild ducks based on genomic analyses. *Zool. Res.* **42**, 43 (2021).
  68. Feng, P. et al. Whole-genome resequencing provides insights into the population structure and domestication signatures of ducks in eastern China. *BMC Genom.* **22**, 1–13 (2021).
  69. Wang, L. et al. Understanding the genetic domestication history of the Jianchang duck by genotyping and sequencing of genomic genes under selection. *G3 Genes Genomes Genet.* **10**, 1469–1476 (2020).

70. Johnson, K. P. & Sorenson, M. D. Phylogeny and biogeography of dabbling ducks (genus: *Anas*): a comparison of molecular and morphological evidence. *Auk* **116**, 792–805 (1999).
71. McCracken, K. G., Johnson, W. P. & Sheldon, F. H. Molecular population genetics, phylogeography, and conservation biology of the mottled duck (*Anas fulvigula*). *Conserv. Genet.* **2**, 87–102 (2001).
72. Omland, K. E. Examining two standard assumptions of ancestral reconstructions: repeated loss of dichromatism in dabbling ducks (*Anatini*). *Evolution*, **51**, 1636–1646 (1997).
73. Sossinka, R. Domestication in birds. *Avian Biol.* **6**, 373–403 (1982).
74. Sheppard, J. L., Amundson, C. L., Arnold, T. W. & Klee, D. Nesting ecology of a naturalized population of Mallards *Anas platyrhynchos* in New Zealand. *IBIS* **161**, 504–520 (2019).
75. Gillespie, G. D. Hybridization, introgression, and morphometric differentiation between Mallard (*Anas platyrhynchos*) and Grey Duck (*Anas superciliosa*) in Otago, New Zealand. *Auk* **102**, 459–469 (1985).
76. Rhymer, J. M., Williams, M. J. & Kingsford, R. T. Implications of phylogeography and population genetics for subspecies taxonomy of grey (Pacific black) duck *Anas superciliosa* and its conservation in New Zealand. *Pac. Conserv. Biol.* **10**, 57–66 (2004).
77. Rhymer, J. M., Williams, M. J. & Braun, M. J. Mitochondrial analysis of gene flow between New Zealand Mallards (*Anas platyrhynchos*) and Grey Ducks (*A. superciliosa*). *Auk* **111**, 970–978 (1994).
78. Söderquist, P., Gunnarsson, G., Elmberg, J. & Dessborn, L. Survival of wild and farmed-released mallards: the Swedish example. *Eur. J. Wildl. Res.* **67**, 1–12 (2021).
79. Griffiths, R. C. & Tavaré, S. Sampling theory for neutral alleles in a varying environment. *Philos. Trans. R. Soc. Lond. Ser. B* **344**, 403–410 (1994).
80. Donnelly, P. & Tavaré, S. Coalescents and genealogical structure under neutrality. *Annu. Rev. Genet.* **29**, 401–421 (1995).
81. Lu, C. W., Yao, C. T. & Hung, C. M. Domestication obscures genomic estimates of population history. *Mol. Ecol.* **31**, 752–766 (2022).
82. Grayson, D. K. The archaeological record of human impacts on animal populations. *J. World Prehistory* **15**, 1–68 (2001).
83. Vigne, J.-D. The origins of animal domestication and husbandry: a major change in the history of humanity and the biosphere. *Comptes Rendus Biologies* **334**, 171–181 (2011).
84. DeMello, M. In *Animals and Society* 102–116 (Columbia University Press, 2021).
85. Leopold, A. *Game Management* (Univ of Wisconsin Press, 1987).
86. Sellers, R. M. & Greenwood, J. J. Sir Richard Graham and the marking of ducks at Netherby, Cumberland, 1908–1933—an early bird-ringing project. *Ring. Migr.* **33**, 86–93 (2018).
87. Armstrong, D. P. & Seddon, P. J. Directions in reintroduction biology. *Trends Ecol. Evol.* **23**, 20–25 (2008).
88. Ewen, J. G., Armstrong, D. P., Parker, K. A. & Seddon, P. J. *Reintroduction Biology: Integrating Science and Management* (John Wiley & Sons, 2012).
89. Perez, R. M., Wilson, D. E. & Gruen, K. D. Survival and flight characteristics of captive-reared and wild northern bobwhite in Southern Texas. *National Quail Symposium Proceedings* **5**, 81–85 (2002).
90. Evans, C. S. & Evans, L. Representational signalling in birds. *Biol. Lett.* **3**, 8–11 (2007).
91. Corbi, J., Baack, E. J., Dechaine, J. M., Seiler, G. & Burke, J. M. Genome-wide analysis of allele frequency change in sunflower crop–wild hybrid populations evolving under natural conditions. *Mol. Ecol.* **27**, 233–247 (2018).
92. Haccou, P. et al. Introgression of crop alleles into wild and weedy populations. *Annu. Rev. Ecol. Evol. Syst.* **44**, 325–345 (2013).
93. Araki, H., Ardren, W. R., Olsen, E., Cooper, B. & Blouin, M. S. Reproductive success of captive-bred steelhead trout in the wild: evaluation of three hatchery programs in the Hood River. *Conserv. Biol.* **21**, 181–190 (2007).
94. Champagnon, J., Guillemain, M., Elmberg, J., Folkesson, K. & Gauthier-Clerc, M. Changes in Mallard *Anas platyrhynchos* bill morphology after 30 years of supplemental stocking. *Bird. Study* **57**, 344–351 (2010).
95. Rodríguez-Rodríguez, E. J., Gil-Morión, J. & Negro, J. J. Feral animal populations: separating threats from opportunities. *Land* **11**, 1370 (2022).
96. Lavretsky, P. et al. Speciation genomics and a role for the Z chromosome in the early stages of divergence between Mexican ducks and mallards. *Mol. Ecol.* **24**, 5364–5378 (2015).
97. McKinney, G. J., Larson, W. A., Seeb, L. W. & Seeb, J. E. RADseq provides unprecedented insights into molecular ecology and evolutionary genetics: comment on Breaking RAD by Lowry et al. (2016). *Mol. Ecol. Resour.* **17**, 356–361 (2017).
98. Lowry, D. B. et al. Breaking RAD: an evaluation of the utility of restriction site-associated DNA sequencing for genome scans of adaptation. *Mol. Ecol. Resour.* **17**, 142–152 (2017).
99. Rhie, A. et al. Towards complete and error-free genome assemblies of all vertebrate species. *Nature* **592**, 737–746 (2021).
100. Paez, S. et al. Reference genomes for conservation. *Science* **377**, 364–366 (2022).
101. Ottenburghs, J. et al. Avian introgression in the genomic era. *Avian Res.* **8**, 30 (2017).
102. Olden, J. D. & Rooney, T. P. On defining and quantifying biotic homogenization. *Glob. Ecol. Biogeogr.* **15**, 113–120 (2006).
103. McFarlane, S. E. & Pemberton, J. M. Detecting the true extent of introgression during anthropogenic hybridization. *Trends Ecol. Evol.* **34**, 315–326 (2019).
104. Graham, C. F. et al. Impacts of degraded DNA on restriction enzyme associated DNA sequencing (RADSeq). *Mol. Ecol. Resour.* **15**, 1304–1315 (2015).
105. Sorenson, M. D., Ast, J. C., Dimcheff, D. E., Yuri, T. & Mindell, D. P. Primers for a PCR-based approach to mitochondrial genome sequencing in birds and other vertebrates. *Mol. Phylogenetics Evol.* **12**, 105–114 (1999).
106. Sorenson, M. D. & Fleischer, R. C. Multiple independent transpositions of mitochondrial DNA control region sequences to the nucleus. *Proc. Natl Acad. Sci. USA* **93**, 15239–15243 (1996).
107. Lavretsky, P., McCracken, K. G. & Peters, J. L. Phylogenetics of a recent radiation in the mallards and allies (Aves: *Anas*): Inferences from a genomic transect and the multispecies coalescent. *Mol. Phylogenetics Evol.* **70**, 402–411 (2014).
108. Ankney, C. D., Dennis, D. G., Wishard, L. N. & Seeb, J. E. Low genic variation between black ducks and mallards. *Auk* **103**, 701–709 (1986).
109. Lavretsky, P. In *Population Genomics: Wildlife, Population Genomics* (eds P. Hohenlohe & O. P. Rajora) (Springer, Cham, 2020).
110. Leigh, J. W. & Bryant, D. POPART: full-feature software for haplotype network construction. *Methods Ecol. Evol.* **6**, 1110–1116 (2015).
111. Fowler, A., Eadie, J. & Englis, A. Identification of endangered Hawaiian ducks (*Anas wyvilliana*), introduced North American mallards (*A. platyrhynchos*) and their hybrids using multilocus genotypes. *Conserv. Genet.* **10**, 1747–1758 (2009).
112. DaCosta, J. M. & Sorenson, M. D. Amplification biases and consistent recovery of loci in a double-digest RAD-seq protocol. *PLoS One* **9**, e106713 (2014).
113. Hernández, F., Brown, J. I., Kaminski, M., Harvey, M. G. & Lavretsky, P. Genomic evidence for rare hybridization and large demographic changes in the evolutionary histories of four North American dove species. *Animals* **11**, 2677 (2021).
114. Schummer, M. L. et al. Population genetics and geographic origins of mallards harvested in northwestern Ohio. *PLoS One* **18**, e0282874 (2023).
115. Brown, J. I. et al. Genomic and morphological data shed light on the complexities of shared ancestry between closely related duck species. *Sci. Rep.* **12**, 1–14 (2022).
116. Lavretsky, P., Wilson, R. E., Talbot, S. L. & Sonsthagen, S. A. Phylogenomics reveals ancient and contemporary gene flow contributing to the evolutionary history of sea ducks (Tribe Mergini). *Mol. Phylogenetics Evol.* **161**, 107164 (2021).
117. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
118. Li, H. & Durbin, R. Inference of human population history from individual whole-genome sequences. *Nature* **475**, 493 (2011).
119. Lavretsky, P., Hernández, F., Swale, T. & Mohl, J. E. Chromosomal-level reference genome of a wild North American mallard (*Anas platyrhynchos*). *G3 Genes|Genomes|Genetics* jkad171 (2023).
120. Danecek, P. et al. The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
121. Lavretsky, P., DaCosta, J. M., Sorenson, M. D., McCracken, K. G. & Peters, J. L. 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.* **28**, 2594–2609 (2019).
122. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
123. Shringarpure, S. S., Bustamante, C. D., Lange, K. & Alexander, D. H. Efficient analysis of large datasets and sex bias with ADMIXTURE. *BMC Bioinforma.* **17**, 218 (2016).
124. Alexander, D. H. & Lange, K. Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinforma.* **12**, 246 (2011).
125. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).
126. Pfeifer, B., Wittelsbürger, U., Ramos-Onsins, S. E. & Lercher, M. J. PopGenome: an efficient swiss army knife for population genomic analyses in R. *Mol. Biol. Evol.* **31**, 1929–1936 (2014).
127. Wolf, J. B. & Ellegren, H. Making sense of genomic islands of differentiation in light of speciation. *Nat. Rev. Genet.* **18**, 87 (2017).
128. Campagna, L. et al. Repeated divergent selection on pigmentation genes in a rapid finch radiation. *Sci. Adv.* **3**, e1602404 (2017).
129. Pickrell, J. K. & Pritchard, J. K. Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet.* **8**, e1002967 (2012).

130. Gutenkunst, R. N., Hernandez, R. D., Williamson, S. H. & Bustamante, C. D. *Diffusion Approximations for Demographic Inference: DaDi* (2010).
131. Gutenkunst, R. N., Hernandez, R. D., Williamson, S. H. & Bustamante, C. D. Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genet.* **5**, e1000695 (2009).
132. Coffman, A. J., Hsieh, P. H., Gravel, S. & Gutenkunst, R. N. Computationally efficient composite likelihood statistics for demographic inference. *Mol. Biol. Evol.* **33**, 591–593 (2016).
133. Blischak, P. D., Barker, M. S. & Gutenkunst, R. N. Inferring the demographic history of inbred species from genome-wide SNP frequency data. *Mol. Biol. Evol.* **37**, 2124–2136 (2020).
134. Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. & Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **25**, 1965–1978 (2005).
135. *Human Impact Wildlife Conservation Society - WCS & Center for International Earth Science Information Network - CIESIN - Columbia University* (NASA Socioeconomic Data and Applications Center (SEDAC), Palisades, New York, 2005).
136. *Human Footprint Wildlife Conservation Society - WCS & Center for International Earth Science Information Network - CIESIN - Columbia University* (NASA Socioeconomic Data and Applications Center (SEDAC), Palisades, New York, 2005).
137. Bay, R. A. et al. Genomic signals of selection predict climate-driven population declines in a migratory bird. *Science* **359**, 83–86 (2018).
138. Fitzpatrick, M. C. & Keller, S. R. Ecological genomics meets community-level modelling of biodiversity: mapping the genomic landscape of current and future environmental adaptation. *Ecol. Lett.* **18**, 1–16 (2015).

## Acknowledgements

We want to especially thank all of the United States' state and federal agencies, Canadian Wildlife Services, New Zealand Fish and Game, and all of the private individuals across the world that provided samples and made this project possible. We want to thank E. Fricke, M. Kaminski, F. Hernandez, V. Musni, M. Cone, E. Duenez, and P. Cavarrubias for their help in the lab. Funding was provided by Ducks Unlimited Canada Institute for Wetlands and Waterfowl Research, the Waterfowl Research Foundation, the National Institutes of Health (NIH- NIMHD grant 5U54MD007592), and the National Science Foundation (DEB Grant ID 2010704). Additionally, PS was funded by the Swedish Environmental Protection Agency (Grant ID V-205-09), and RHSK was funded by Martin Wikelski at the Max Planck Institute of Animal Behavior. Finally, we want to thank the AE and two anonymous reviewers for their comments.

## Author contributions

P.L. conceived the study; P.L. & M.L.S. funded the project; P.L. & J.I.B. designed experiments; Sample acquisition was done by P.L., P.S., R.H.S.K., M.L.S. & J.I.B.; P.L., J.M. & J.I.B. obtained molecular data; P.L., J.M. & J.I.B. performed bioinformatics and molecular analyses; P.L., J.M., P.S., R.H.S.K., M.L.S. & J.I.B. equally contributed to the writing of the manuscript.

## Competing interests

All authors have no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s42003-023-05170-w>.

**Correspondence** and requests for materials should be addressed to Philip Lavretsky.

**Peer review information** *Communications Biology* thanks Filippo Barbanera, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Luke Grinham and David Favero.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023