Discordance between taxonomy and population genomic data: An avian example relevant to the United States Endangered Species Act

Andrew N. Black [ID][a,b], Andrew J. Mularo [ID][c], Jong Yoon Jeon [ID][a], David Haukos [ID][d], Kristin J. Bondo[e], Kent A. Fricke[f], Andy Gregory [ID][g], Blake Grisham[e], Zachary E. Lowe [ID][b] and J. Andrew DeWoody [ID][a,c,*]

[a]Department of Forestry and Natural Resources, Purdue University, West Lafayette, IN 47907, USA
[b]Western Association of Fish and Wildlife Agencies, Boise, ID 83719, USA
[c]Department of Biological Sciences, Purdue University, West Lafayette, IN 66506, USA
[d]U.S. Geological Survey, Kansas Cooperative Fish and Wildlife Research Unit, Kansas State University, Manhattan, KS 66506, USA
[e]Department of Natural Resources Management, Texas Tech University, Lubbock, TX 79410, USA
[f]Kansas Department of Wildlife and Parks, Emporia, KS 66801, USA
[g]Department of Biological Sciences, University of North Texas, Denton, TX 76201, USA
*To whom correspondence should be addressed: Email: dewoody@purdue.edu
Edited By Li-Jun Ma

Abstract

Population genomics can reveal cryptic biological diversity that may impact fitness while simultaneously serving to delineate relevant conservation units. Here, we leverage the power of whole-genome resequencing for conservation by studying 433 individual lesser prairie-chicken (*Tympanuchus pallidicinctus*; LEPC, a federally endangered species of conservation concern in the United States) and greater prairie-chicken (*Tympanuchus cupido*; GRPC, a legally huntable species throughout much of its range). The genomic diversity of two formally recognized distinct population segments (DPSs) of LEPCs is similar, but they are genetically distinct. Neither DPS is depleted of its genomic diversity, neither is especially inbred, and temporal diversity is relatively stable in both conservation units. Interspecific differentiation between the two species was only slightly higher than that observed between LEPC DPSs, due largely to bidirectional introgression. The high resolution provided by our dataset identified a genomic continuum between the two species such that individuals sampled from the hybrid zone were imperfectly assigned to their presumptive species when considering only their physical characteristics. The admixture between the two species is reflected in the spectrum of individual ancestry coefficients, which has legal implications for the "take" of individuals under the Endangered Species Act. Overall, our data highlight the recurring dissonance between static policies and dynamic species boundaries that are increasingly obvious in the population genomic era.

Keywords: Endangered Species Act, genetic diversity, introgression, prairie-chicken, whole-genome resequencing

Significance Statement

By providing legal protection to species at risk of extinction, the 1973 United States Endangered Species Act (ESA) represents a cornerstone for conservation. This research provides (i) a genomic assessment of population structure and diversity among two ESA-listed populations and (ii) an evaluation of introgression between two hybridizing species to inform policy and management. The data demonstrate recent differentiation between species and highlight the difficulties of implementing a binary framework (i.e. Latin binomials) for categorizing individuals that vary dramatically in their hybrid ancestry. As population genomics becomes more mainstream, we expect that individual introgression will become more evident. That poses significant legal and policy challenges for agencies that must walk the line with respect to the "take" of threatened and nonthreatened species.

Introduction

The Endangered Species Act (ESA) is designed to help protect species of conservation concern, and it turned 50 years old in 2023. Many authors have pointed out the practical problems with language in the ESA, particularly as it pertains to hybrids (e.g. (1–4)). These problems are exacerbated by the rapidly increasing resolution of population genomic data that allows biologists to more clearly identify and quantify signatures of effective hybridization (i.e. introgression). In sexual organisms, genomic data indicate that many related species are obvious evolutionary amalgamations produced largely by horizontal gene flow (5–7).
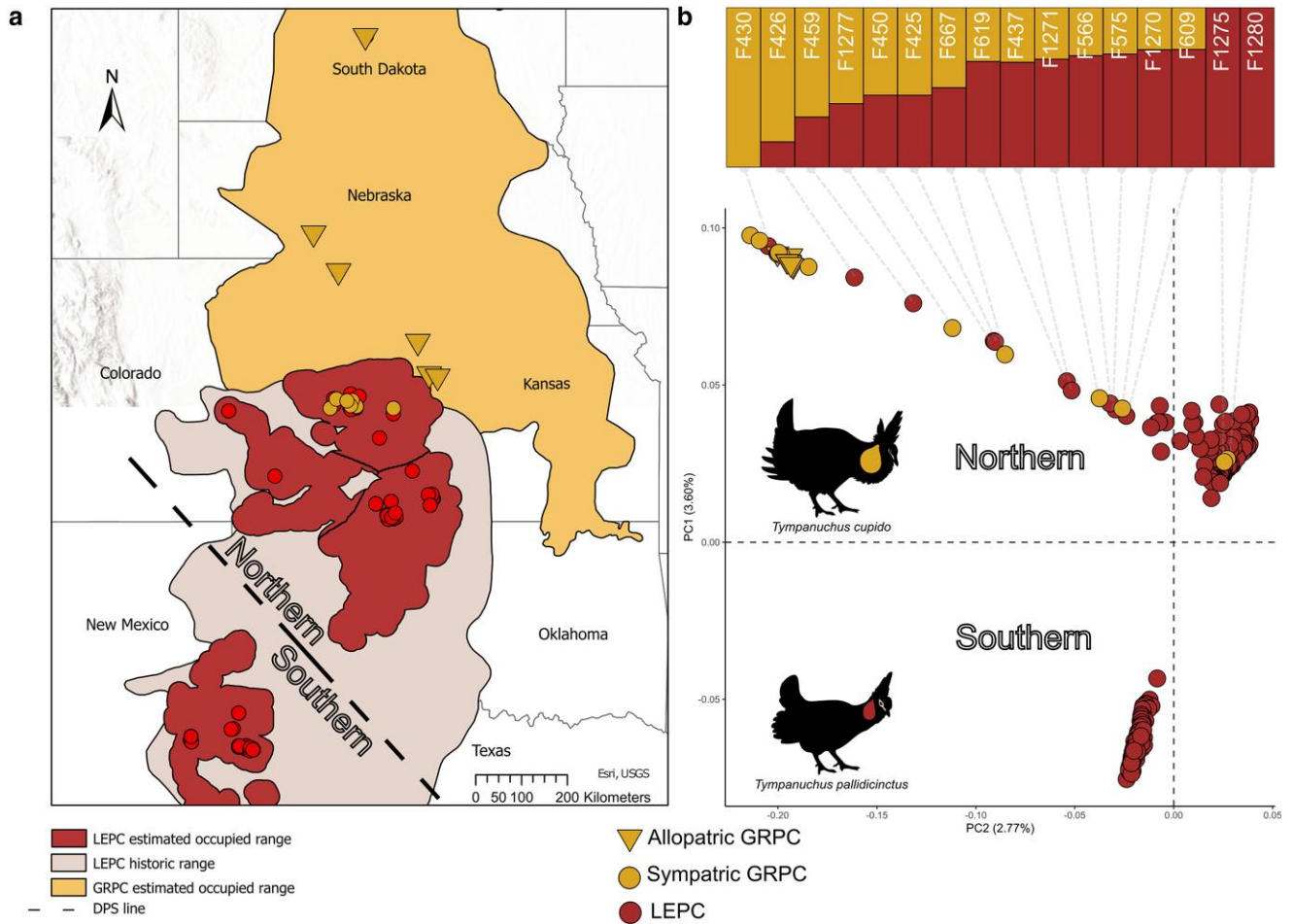
Fig. 1. Geographic sampling regime and Principal Component Analysis (PCA) of whole-genome sequences. a) Sampling locations of 411 LEPC (*T. pallidicinctus*) collected from the Northern and Southern DPSs in the LEPC current range. Additionally, 22 GRPC (*T. cupido*) were collected from sympatric (within LEPC range) and allopatric (outside of LEPC range) locations. Sampling points largely correspond to sampling lek, with 1–20 birds collected from each, but some individuals were obtained outside of leks. The black borders in the Northern DPS and Southern DPS correspond to the four different ecoregions. b) PCA of allele frequencies derived from ~ 18 M polymorphic (maf > 0.01) sites across the nuclear genome. Principal component 1 is labeled on the *Y*-axis, and principal component 2 is labeled on the *X*-axis to better orient the points geographically. Less than 10% of the overall variation is explained by taxonomy or geography. Above the PCA, individual admixture proportions ($K = 3$) for mixed ancestry samples (> 0.20) identify three birds (at ends of bar plot; see Figure 2 for additional information) where genomic identities did not correspond to morphological species assignments (sample numbers included in white font). For symbols in the map and PCA, colors in the diagrammed silhouettes correspond to the phenotypic species identification of LEPC (brown) and GRPC (goldenrod). The dashed diagonal (a) and horizontal line (b) depicts the separation between the Northern and Southern DPSs.

The U.S. Fish and Wildlife Service (USFWS) recently listed the lesser prairie-chicken (LEPC) under the ESA, identifying two distinct population segments (DPSs). The more isolated Southern DPS is now federally endangered and the Northern DPS is now threatened (Figure 1a; (8, 9)). The LEPC is one of three lekking prairie grouse species in the genus *Tympanuchus*, which also includes the greater prairie-chicken (GRPC) (10–12). The LEPC and GRPC diverged from one another 0.6–0.9 Mya (13, 14), and they can hybridize in regions of geographic sympatry (Figure 1a). Hybrids are apparently fertile as indicated by the direct evidence of controlled crosses that have produced F1 and F2 interspecific hybrids (15) as well as by indirect genetic evidence of introgression in wild birds (16, 17).

Previous genetic studies of LEPC using a few genetic markers revealed moderate genetic differentiation and similar diversity levels in samples from different ecoregions (11, 16, 18, 19), but the recent USFWS (8) Final Rule states that "There are concerns about the implications of genetic introgression (dilution) of LEPC genes, particularly given that potential effects are poorly understood." Here, we performed a high-resolution assessment of LEPC population structure and potential interspecific introgression that are relevant to new conservation policies. Our whole-genome sequence data from 433 individual LEPCs and GRPCs demonstrate that: (i) the LEPC has modest levels of genomic diversity compared to other avian species; (ii) genomic diversity has not substantively changed over time in either species; (iii) there is distinct but moderate differentiation between LEPC sampled from the Northern and Southern DPSs; (iv) interspecific hybridization between LEPC and GRPC has resulted in bidirectional introgression that blurs assignment of a few individuals to a species; and (v) interspecific divergence between LEPC and GRPC is very low. The boundary between these two species is still porous (11), and the current scientific nomenclature (i.e. discrete Latin binomials for each species), which is critical for effective conservation policy, poorly reflects their ongoing evolutionary dynamics.

Results

Population structure and introgression

LEPC samples were collected over more than a decade from the Southern DPS ($n = 150$) and the Northern DPS ($n = 261$); GRPC
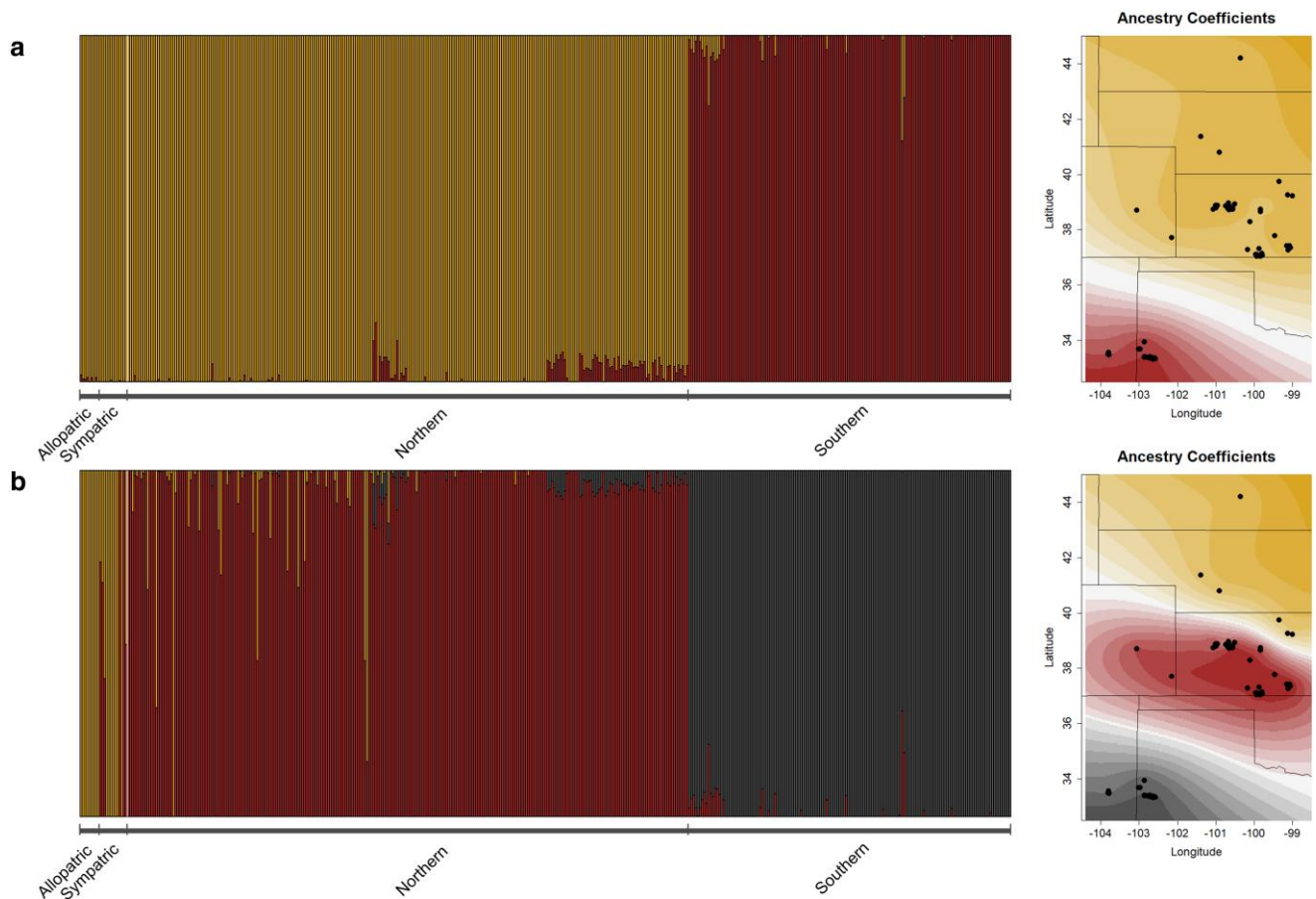
Fig. 2. Admixture plots of whole-genome sequence data. Individual admixture proportions for each bird, illustrating the assignment to the optimal number of clusters ($K = 2$) and next best model ($K = 3$; Figure S2), according to the method of Evanno et al. (20). Colors of individual bars correspond to genetic clusters, not the phenotypic species assignment. Admixture proportions illustrated in bar plots for $K = 2$ (a) and $K = 3$ (b) are sorted from north to south latitude (left to right). For each bar plot, individual ancestry coefficients were also continuously plotted over geographic space (maps). The vertical white solid lines in the bar plots correspond to the separation between field assigned species, with 411 LEPC (*T. pallidicinctus*) collected from the Northern and Southern DPSs and 22 GRPC (*T. cupido*) that based upon collection local were sympatric (within LEPC range) and allopatric (outside of LEPC range). For $K = 2$, the admixture analysis groups the Southern DPS separately from the Northern DPS + GRPC. The second most likely number of clusters ($K = 3$) is mostly consistent with DPS and taxonomic integrity while identifying biparental introgression between LEPC and GRPC.

samples were collected from areas of sympatry with LEPC ($n = 13$) and areas of allopatry ($n = 9$; Figure 1a). A principal component analysis (PCA) and admixture analysis were performed on $\sim$ 18 M variable sites among 433 individual samples. A distance-based approach (i.e. model-free) using allele frequencies across the entire genome explained only a small proportion of the overall variability in the dataset (principal component 1 = 3.60%, *Y*-axis). The PCA clearly separated the Southern DPS and revealed a transect of genetic connectivity between LEPC from the Northern DPS to the allopatric GRPC (Figure 1b). Samples along principal component 2 (2.77% explained variation, *X*-axis) poorly discriminated LEPC and GRPC sampled north of the DPS line (Figure 1b). A virtually identical PCA was generated when using the chicken (*Gallus gallus*) reference genome (GCA_016700215.1), demonstrating that there was minimal ascertainment bias due to reference genome selection (Figure S1). Admixture analysis using a model-based (Hardy–Weinberg) approach identified two genomic groups ($K = 2$; Figure S2) that clustered individuals by collection location above or below the official DPS boundary line but did not discriminate between phenotypic species assignment in the Northern part of the range (Figure 2a). However, the next most likely number of genomic groups ($K = 3$) categorized individuals based upon DPS and species, revealing bidirectional interspecific

introgression (63 individuals with co-ancestry proportions > 0.001; Figure 2a). The $K = 3$ model revealed taxonomic discrepancies between field (phenotypic) species assignments and genomic affiliations in two GRPC individuals and one LEPC. One nominal LEPC from the northern DPS (F430) was apparently misidentified or misrecorded as a GRPC. One sample (F1270) identified in the field as a putative hybrid was of GRPC (0.73) and Northern DPS LEPC ancestry (0.27).

Of the remaining 410 LEPC samples, 63 (15.4%) individuals had some degree of GRPC hybrid ancestry. These hybrid ancestry individuals were collected from all three ecoregions in the Northern DPS, but primarily from the Short-Grass Prairie/Conservation Reserve Program (CRP) Mosaic Ecoregion. The 63 birds with hybrid ancestry comprised 24.1% of the putative LEPCs that we sampled in the Northern DPS. The mean proportion of Northern DPS LEPC genomes attributable to GRPC (i.e. the degree of introgression) was 4.5% among these 63 individuals. Out of 13 sympatric GRPCs, six had LEPC hybrid ancestry but two were likely misidentifications in the field. The remaining four admixed GRPCs have a mean of 42% of their genome from Northern DPS LEPC. The ABBA–BABA tests and associated D-statistics found significant but modest levels of gene flow among all examined groups, including between species (Table S1). Global weighted estimates of pairwise F_{ST}

Table 1. Global weighted pairwise F_{ST} values between the four examined groups for *T. cupido* (GRPC) and the Northern and Southern DPSs of *T. pallidicinctus* (LEPC).

Species		GRPC		LEPC	
		Allopatric	Sympatric	Northern DPS	Southern DPS
GRPC	Allopatric	—	0.0405	0.1087	0.1447
	Sympatric		—	0.0558	0.0946
LEPC	Northern DPS			—	0.044
	Southern DPS				—

Allele frequency differences between the Northern and Southern DPSs of LEPC account for 4.4% of the overall variance in the dataset.

between LEPC collected from the Northern and Southern DPSs were low to moderate ($F_{ST} = 0.044$), demonstrating that 4% of the total variation across the genome was partitioned between the two DPSs. The greatest differentiation occurred between LEPC from the Southern DPS vs allopatric ($F_{ST} = 0.145$) and sympatric ($F_{ST} = 0.095$) GRPC. As expected with interspecific gene flow, levels of differentiation were reduced when comparing LEPC from the Northern DPS to allopatric ($F_{ST} = 0.109$) and sympatric ($F_{ST} = 0.06$) GRPC (Table 1).

Genome-wide patterns of diversity

Genome-wide mean heterozygosity (H) was 0.0031 ± 0.0005 (Figure 3a) across all 411 nominal LEPC, with reduced levels of diversity in the Southern DPS ($H = 0.00295 \pm 3.27e-4$) compared to the Northern DPS ($H = 0.00340 \pm 5.30e-4$). There were no obvious trends in temporal H among samples collected from 2008 to 2023, either between DPSs or species (Figure 3b). Overall LEPC inbreeding was low (fROH$_{total}$ = 0.067 ± 0.062) and similar to Greater Sage-Grouse (*Centrocercus urophasianus*; fROH$_{total}$ = 0.072 ± 0.06), Gunnison Sage-Grouse (*Centrocercus minimus*; fROH$_{total}$ = 0.03 ± 0.025), and other galliform species (21). Inbreeding was statistically greater in the Southern DPS than in the Northern DPS (*P*-values < $2e-16$). One LEPC individual (out of 411 in total) sampled in the Northern DPS was inbred (fROH$_{total}$ = 0.61). No significant differences in inbreeding were detected when comparing the two species for any fROH category (*P*-values = 0.056–0.83; Figure 3c).

A distance-based nuclear tree largely resolved each species (i.e. it clustered GRPC together within a clade) and DPS, but it revealed some GRPC interspersed among LEPC collected from the Northern DPS (Figure 4a). Mixing of putative species in the nuclear tree can be partially explained based upon the 16 samples that were misidentified in the field or showed hybrid ancestry (Figure 1b). In contrast, the maximum likelihood mtDNA tree provided no resolution as it detected no clear structure between LEPC from either DPS or between species and neither did the haplotype network (Figures 4b and S12). There was no definitive evidence of reciprocal monophyly, one common yardstick of "good" species (22), in our analyses of mtDNA genomes (Figure 4). Haplotype diversity statistics associated with the mtDNA genomes reveal reduced haplotype and nucleotide diversity for LEPC sampled from the Southern DPS, consistent with the nuclear genomic results (Table 1). Furthermore, despite the larger sample size, LEPC contained less than half the level of haplotype diversity observed in GRPC (Table 1).

Species divergence and effective population sizes

Grouse have traditionally been classified using phenotypic and behavioral characteristics, but species boundaries are ill-defined due to hybridization among species (17, 23). To focus on the greatest divergence between LEPC and GRPC, the two extremes of our sampling distribution were utilized (i.e. Southern DPS of LEPC vs allopatric GRPC). A nonoverlapping 50 kb sliding window analysis of genome-wide pairwise nucleotide diversity (ϕ_π) showed relatively stable and highly correlated (Pearson's $r = 0.968$) ϕ_π for both LEPC (mean = 0.00361, range = 0.00003–0.03952) and GRPC (0.00386, 0.00036–0.03683), illustrating similar genomic diversity and architecture (Figure S3). Mean ϕ_π among both species was 0.00366 (0.00012–0.03609), signifying there were not any more notable differences occurring between species relative to within species. Under neutrality, $\phi_\pi = 4N_e(\mu)$ and equivalently, $N_e = \phi_\pi / 4\mu$, where N_e is the effective population size and μ is the mutation rate. Using a grouse-specific mean mutation rate of 3.28×10^{-9} (24), we estimate a long-term $N_e = 276,152$ in LEPC and $N_e = 294,207$ in GRPC.

For context, we quantified divergence and differentiation using data from the Gunnison and Greater Sage-Grouse (25). Those results (Figure S4) reveal substantially greater diversity observed among the Greater Sage-Grouse population ($\phi_\pi = 0.00211$) relative to the Gunnison Sage-Grouse populations ($\phi_\pi = 0.00081$). Thus, Sage-Grouse have roughly half as much nucleotide diversity as prairie-chickens. Low levels of interspecific differentiation ($F_{ST} = 0.073$ across ~18 M sites) existed between all LEPC and all GRPC samples, whereas the two Sage-Grouse species showed ~5× more interspecific genomic differentiation ($F_{ST} = 0.384$).

Discussion

Conservation policies ostensibly rely on the best available science. Whole-genome resequencing studies like ours have the capacity to capture relevant patterns of genomic diversity and to clarify the underlying evolutionary processes. Genomic diversity is positively associated with evolutionary fitness, with a lack of diversity often seen in populations of conservation concern (11, 26). Our study of whole-genome sequences has revealed no obvious red flags in observed levels of diversity in the LEPC. We explicitly note that this situation could change quickly in the face of rapid and severe demographic bottlenecks due to external factors such as infectious disease, additional habitat loss, and climate change. We certainly advocate the continuation of best management practices to help sustain LEPC populations for the foreseeable future.

Neither LEPC DPS is depauperate of genomic diversity, neither is particularly inbred, temporal diversity is stable in each DPS (Figure 3), and diversity is similar across ecoregions (Figure S5), almost certainly because of the large (relative to contemporary size estimates (27)) long-term N_e. There is modest but distinct genome-wide differentiation (Table 1) that reflects both the anthropogenic habitat fragmentation that has gradually isolated each DPS since the 19th century (28) and any natural interspecific isolation that might have arisen prior to anthropogenic habitat
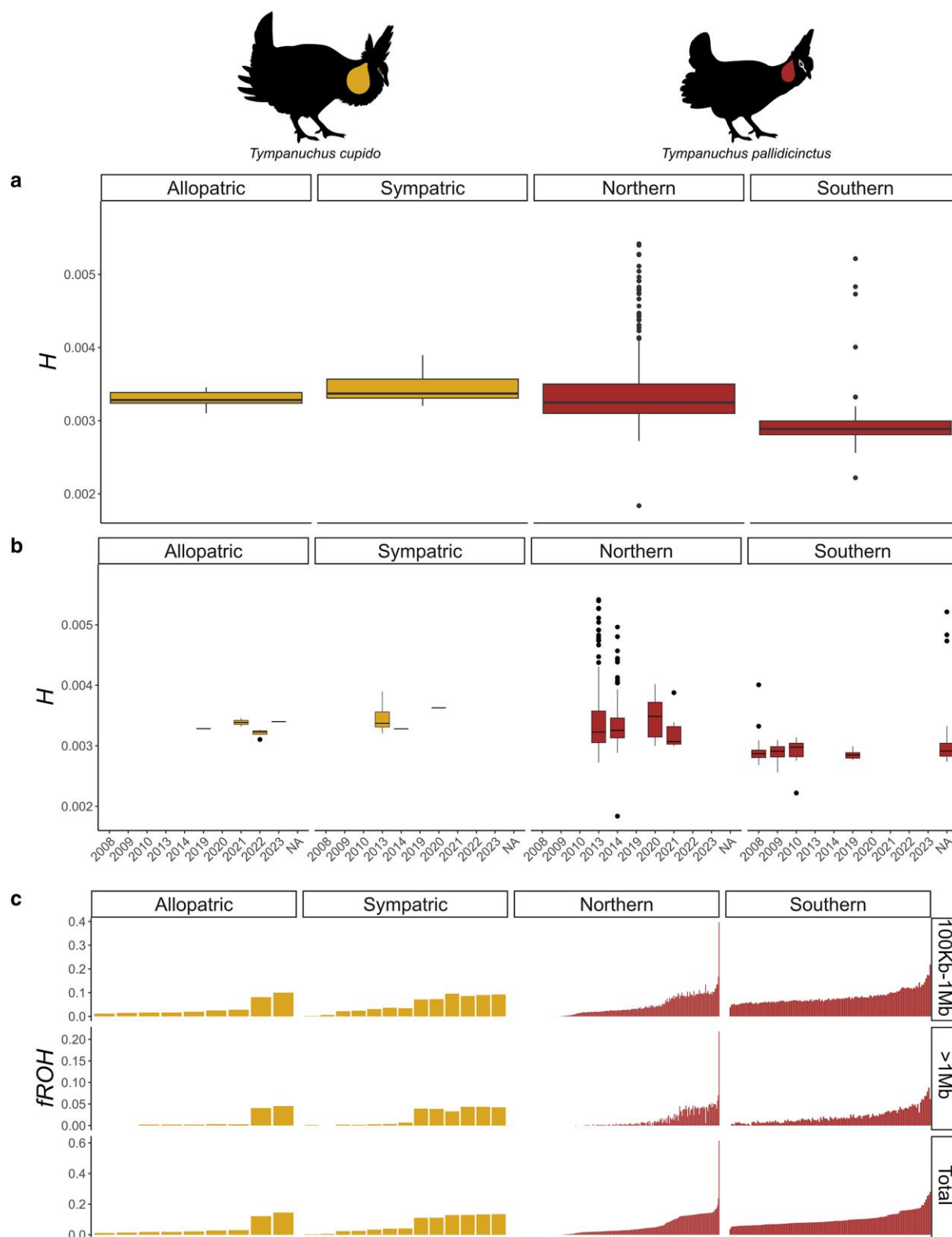
Fig. 3. Genomic diversity and autozygosity. 411 nominal LEPC (*T. pallidicinctus*) were collected from the Northern and Southern DPSs and 22 GRPC (*T. cupidio*) that based upon collection local were sympatric (within LEPC range) and allopatric (outside of LEPC range). a) Box plot of 433 individual heterozygosity (*H*) estimates derived from whole-genome resequencing, grouped by sampling region (allopatric, sympatric, Northern DPS, and Southern DPS), and colored by species. b) Box plot of individual heterozygosity estimates for each sampled region, factored by the year the sample was collected. Note that several birds were missing the sampling year in the attribute data, so were labeled as NA (not applicable). There is very little temporal variation in heterozygosity regardless of geographic or taxonomic affiliation. c) The proportion of each individual genome with runs of homozygosity (*fROH*), binned by length (100 kb–1 Mb; > 1 Mb; total) and grouped by sampling region. For all Panels (a–c), colors in the diagramed silhouettes correspond to LEPC (brown) and GRPC (goldenrod) samples.
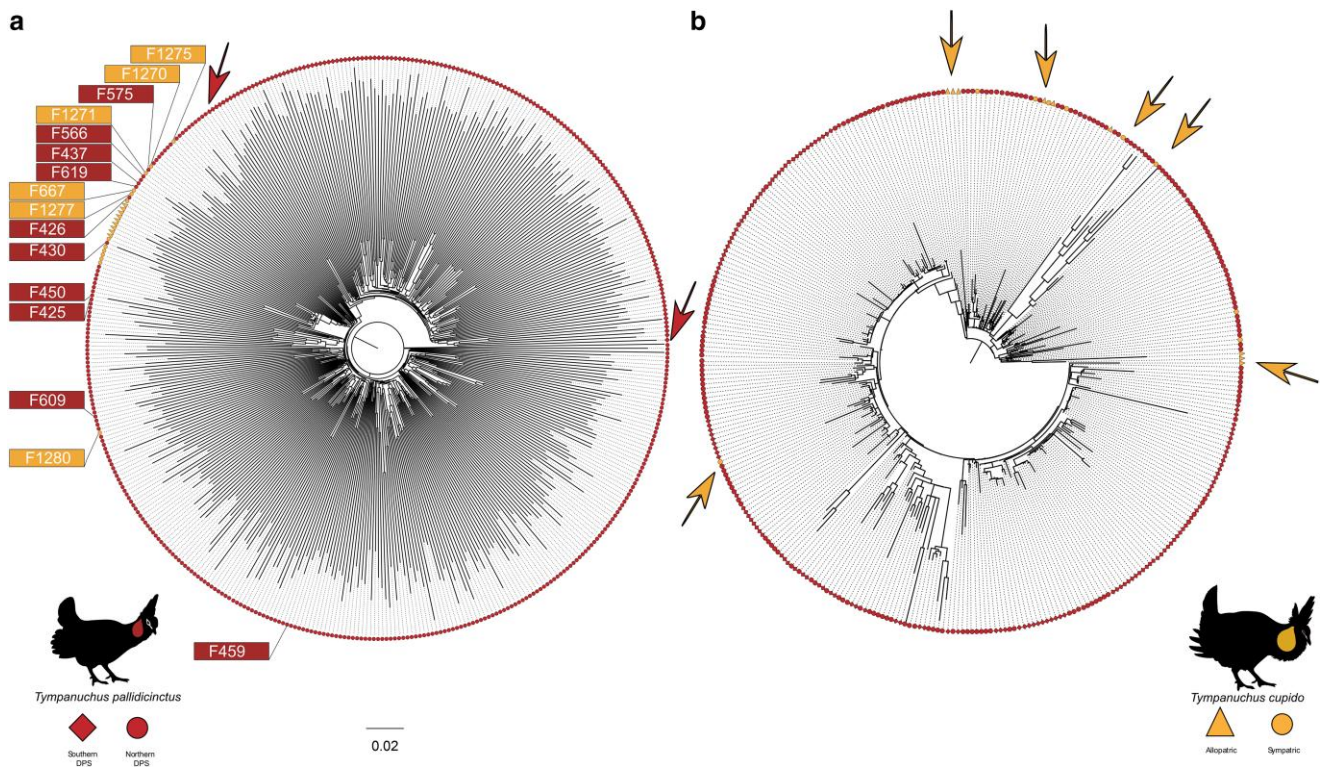
Fig. 4. Nuclear and mitochondrial phylogeographic trees. Midrooted distance tree derived from filtered nuclear genotype likelihoods (a) and a maximum likelihood tree from aligned reconstructed mitochondrial assemblies (b), with colors representing the two species and shapes corresponding to the source location (Northern DPS, Southern DPS, allopatric, and sympatric). For both trees, colors in the diagrammed silhouettes correspond to the LEPC (*T. pallidicinctus*) and the GRPC (*T. cupido*). The red arrows in a) identify the division between LEPC collected from the Northern and Southern DPSs, and the labeled sample numbers represent the potentially misidentified or introgressed birds illustrated in the bar plot above Figure 1b. The goldenrod arrows in b) are illustrated to orient the location of several GRPC samples throughout the mtDNA tree. The nuclear tree (a) shows increased species discrimination compared to the mitochondrial tree (b), but neither exhibits a clear pattern of reciprocal monophyly that is often associated with distinct species or evolutionarily significant units (22) driven in part by bidirectional introgression between species.

Table 2. Number of individuals, segregating sites, haplotypes and Haplotype (H_d) and nucleotide (π) diversity statistics for the Northern and Southern DPSs of *T. pallidicinctus* (LEPC) and the GRPC (*T. cupido*).

	No. of individuals	No. of segregating sites	No. of haplotypes	H_d	Mean no. of differences between haplotypes	π
LEPC	339	55	32	0.27205	0.49210	0.00039
Northern DPS	199	44	23	0.34993	0.66398	0.00052
Southern DPS	140	14	10	0.15128	0.24203	0.00019
GRPC	18	10	8	0.69935	1.30719	0.00103
Sympatric	9	5	5	0.72222	1.11111	0.00088
Allopatric	9	6	5	0.72222	1.50000	0.00118

modification. Our data reveal lower levels of heterozygosity (Figure 3a) and haplotype diversity (Table 2) in the Southern DPS compared to all other examined groups. Mean levels of genomic diversity ($H = 0.003$) observed in our prairie-chicken samples are similar to Greater Sage-Grouse ($H = 0.002$) as well as many other common avian species (29). We found limited evidence for inbreeding among LEPC samples (mean fROH$_{total}$ = 6%), which were similar to those estimated in Gunnison (fROH$_{total}$ = 3%) and Greater Sage-Grouse (fROH$_{total}$ = 7%; (25)) but substantially lower than in a critically imperiled avian species like the Stewart Island Kākāpō (fROH$_{total}$ ~80%; (30)). The lack of appreciable inbreeding and the modest level of temporally stable diversity in our dataset (Figure 3) indicates that genomic imprints of demographic reductions and geographic isolation (12, 31) have not yet manifested themselves in prairie-chickens (though see (32)).

Hybridization and subsequent backcrossing that leads to introgression is well-documented in many avian sub(species) and can be resolved using genomic approaches (32, 33). According to the standard method for identifying the optimal *K*-value (20), the most likely model of admixture identified two main groups ($K = 2$) of prairie-chickens with relatively little differentiation among birds sampled north of the DPS line (Figure 2a). This lack of differentiation is likely driven by recent or ongoing gene flow among leks, ecoregions (Figure S6), and between putative species (Figures 1b and 2a, b). A transect of genomic variability between allopatric GRPC and LEPC collected from the Northern DPS (Figure 1a) and our phylogeographic trees (Figure 4a and especially b) reveal that individual GRPC grouped with LEPC from the Northern DPS. We cannot disregard the influence of incomplete lineage sorting in the observed topography, but the lack of clear

monophyletic groups is consistent with the ABBA–BABA tests and associated D-statistics (Table S1).

Our second-best model of admixture identified three main groups ($K = 3$) of prairie-chickens (Figure 2b). To us, the $K = 3$ model makes the most biological sense because it best demarcates each species and each DPS while also quantifying bidirectional, interspecific introgression. Much of the interspecific gene flow is likely due to ongoing hybridization and introgression between species in the Short-Grass Prairie/CRP Mosaic Ecoregion of the Northern DPS (Table S1; Figure S7). Lineages are expected to show extensive genetic similarity under a model of speciation with ongoing gene flow (34), as observed between LEPC and GRPC when comparing levels of overall diversity (Figure S3), sequence divergence (0.003%; see Supporting Information), and differentiation (Table 1). Collectively, our results indicate there have been hybridization and introgression between the two species, consistent with behavioral observations (35, 36).

Our genomic analyses indicate that LEPC genes effectively flow among leks, ecoregions, and species as reflected by a lack of differentiation in the PCA (Figures 1b and S4), the admixture analyses (Figures 2 and S7), the analysis of variance ($F_{ST} = 0.044$), and the phylogeographic trees (Figure 4). We note that our mtDNA tree (Figure 4b) exhibits little genetic structure between LEPC sampled from the Northern and Southern DPSs or between species. This lack of structure reflected by the maternally inherited mtDNA tree may in part be the result of effective female-biased dispersal. Introgression between LEPC and GRPC has long been suspected (16, 17), and our genomic data confirm and quantify it (Figures 1b, 2b, S4, S7, and S9; (32)). GRPC and LEPC are estimated to have diverged from one another 0.6 – 0.9 Mya (13, 14). This relatively recent divergence has led to prairie grouse phylogenies that are fraught with uncertainty because of incomplete lineage sorting and further uncertainty exists because of secondary contact and introgression among lineages after the Pleistocene glaciers receded from the Great Plains (17, 37).

In the future, we expect that bidirectional introgression could continue or even expand in geographic scope and/or in the extent of effective hybridization. Environmental niche models indicate that over the last 130,000 years, there has been substantial overlap in potential habitat for each species (11). Furthermore, the area of sympatry harbors the largest contemporary population of LEPC (27, 38). If LEPCs expand habitat occupancy northward through management efforts like the Conservation Reserve Program (39), there may be increasing opportunities for future hybridization and introgression. Effective gene flow within and between species should continue to help maintain existing genomic variation and temper demographic bottlenecks that can otherwise linger and lead to genomic erosion (40).

Many closely related taxa (like the LEPC and GRPC) are formally recognized as two distinct species, each subject to different conservation strategies and formal legal frameworks. Unfortunately, the static Latin names of biological entities often do not accurately reflect nature. Incipient species can diverge despite genetic exchange via hybridization and introgression (5), and horizontal gene transfer between any two sexual species may have important policy ramifications. Specifically, the inability of whole-genome sequence data to consistently and clearly delineate individuals into species could raise questions about whether an individual is legally protected, or whether a population segment is in fact "discrete" as defined under the ESA (26). However, this depends entirely on how complex and ongoing evolutionary dynamics (such as speciation, hybridization, and introgression) are interpreted in the context of static conservation policies that vary at the state, federal, or international levels.

Rapid advancements in bioinformatics algorithms and next-generation sequencing technology (41) and the concomitant cost reduction are enabling unprecedented insight into evolutionary processes such as speciation and introgressive hybridization. While this is an exciting time for evolutionary biology, it also poses significant challenges to conservation, policy, and law (42). For example, there is the growing realization that individual eukaryotic organisms are often genomic composites of different species (including humans; (43)), so how does one legally enforce a policy when one cannot consistently define a species? The US ESA, which defines species as "any subspecies of fish or wildlife or plants, and any DPS of any species of vertebrate fish or wildlife which interbreeds when mature," is now over 50 years old and was crafted decades before whole-genome sequencing. If current management practices are to represent the best available science and technology, a consensus is needed on how to best classify conservation units (e.g. which species concept to use). This is especially important when considering the management of populations that exhibit signals of recent and/or ongoing introgressive hybridization. Until legislation better reflects the emerging view of biological reality, species will remain the most important taxonomic unit for the conservation of biodiversity (44)—even when their boundaries are porous and ill-defined. The US ESA is a prime example whereby the lack of a consistent "hybrid policy" means that listing decisions are often handled on a case-by-case basis subject to little formal guidance (4). Increasingly, datasets such as ours indicate that hybridization and introgression are common in sexual organisms, and we think the dense information content contained in whole-genome sequences has the potential to be leveraged into more realistic, consistent, and objective policies where there is the political will to do so.

Materials and methods

Comprehensive methodological details and supplemental results (Tables and Figures S1-S12) are provided in the Supporting Information. Briefly, whole genomes from 433 prairie-chickens were sequenced and aligned to the LEPC reference genome (29). Genotype likelihoods were then leveraged to determine individual heterozygosity, genome-wide pairwise nucleotide diversity, ABBA/BABA tests of introgression, and genomic differentiation using ANGSD (45). Runs of homozygosity were identified using BCFTOOLS (46), and population stratification was assessed using PCANGSD (47) and NGSADMIX (48). A midpoint rooted nuclear genomic tree was created using NGSDIST (49), and a mtDNA tree was created using IQTREE (50).

Acknowledgments

Supplementary Material

Supplementary material is available at PNAS Nexus online.

Author Contributions

A.N.B., A.J.M., J.Y.J., and J.A.D. conceived the ideas and designed methodology. D.H., K.J.B., K.A.F., A.G., and B.G. collected and/or processed the samples. A.N.B., A.J.M., and J.Y.J. analyzed the data. A.N.B. and J.A.D. led the writing of the manuscript. Z.E.L. and J.A.D. acquired funding and provided mentorship. All authors contributed critically to the drafts and gave final approval for publication.

Data Availability

Paired-end Illumina shotgun sequence data from 481 prairie-chicken samples has been deposited in the NCBI Sequence Read Archive under Bioproject PRJNA986511. All code, attribute data, and scripts have been uploaded to a publicly accessible repository (https://github.com/Andrew-N-Black/LEPC-popgen).

References

1 O'Brien SJ, Mayr E. 1991. Bureaucratic mischief: recognizing endangered species and subspecies. *Science*. 251:1187–1188.

2 National Research Council. 1995. *Science and the Endangered Species Act*. Chicago (USA): The National Academies Press.

3 Allendorf FW, Leary RF, Spruell P, Wenburg JK. 2001. The problem with hybrids: setting conservation guidelines. *Trends Ecol Evol*. 16:613–622.

4 Lind-Riehl JF, Mayer AL, Wellstead AM, Gailing O. 2016. Hybridization, agency discretion, and implementation of the U.S. Endangered Species Act. *Conserv Biol*. 30:1288–1296.

5 Arnold ML. 2015. *Divergence with genetic exchange*. Oxford (UK): Oxford University Press.

6 Etherington GJ, *et al*. 2022. Extensive genome introgression between domestic ferret and European polecat during population recovery in Great Britain. *J Hered*. 113:500–515.

7 Frei D, De-Kayne R, Selz OM, Seehausen O, Feulner PGD. 2022. Genomic variation from an extinct species is retained in the extant radiation following speciation reversal. *Nat Ecol Evol*. 6: 461–468.

8 U.S. Fish and Wildlife Service. 2023. Endangered and threatened species: lesser prairie-chicken; threatened status with section 4(d) rule for the northern distinct population segment and endangered Status for the southern distinct population segment. *Fed Regist*. 87:72674–72755.

9 U.S. Fish and Wildlife Service. 2022. Species status assessment report for the lesser prairie chicken. Tympanuchus pallidicinctus. *Fed Regist*. 2.3:1–109.

10 Taylor MA, Guthery FS. 1980. *Status, ecology, and management of the lesser prairie chicken*. Fort Collins (USA): U.S. Department of Agriculture. p. 1–15.

11 DeYoung RW, Williford DL. 2016. Genetic Variation and Population Structure in the Prairie Grouse. In: Haukos DA, Boal CW, editors. *Ecology and conservation of lesser prairie-chickens*. 1st ed. Boca Raton (FL): Studies in Avian Biology. p. 77–97.

12 Boal CW, Haukos DA. 2016. The Lesser Prairie-Chicken. In: Haukos DA, Boal CW, editors. *Ecology and conservation of lesser prairie-chickens*. 1st ed. Boca Raton (FL): Studies in Avian Biology. p. 1–14.

13 Stein RW, Brown JW, Mooers AO. 2015. A molecular genetic time scale demonstrates Cretaceous origins and multiple diversification rate shifts within the order Galliformes (Aves). *Mol Phylogenet Evol*. 92:155–164.

14 Persons NW, Hosner PA, Meiklejohn KA, Braun EL, Kimball RT. 2016. Sorting out relationships among the grouse and ptarmigan using intron, mitochondrial, and ultra-conserved element sequences. *Mol Phylogenet Evol*. 98:123–132.

15 Crawford JA. 1978. Morphology and behavior of greater × lesser prairie chicken hybrids. *Southwest Nat*. 23:591–596.

16 Oyler-McCance SJ, *et al*. 2016. Rangewide genetic analysis of lesser prairie-chicken reveals population structure, range expansion, and possible introgression. *Conserv Genet*. 17:643–660.

17 Galla SJ, Johnson JA. 2015. Differential introgression and effective size of marker type influence phylogenetic inference of a recently divergent avian group (Phasianidae: Tympanuchus). *Mol Phylogenet Evol*. 84:1–13.

18 Van Den Bussche RA, Hoofer SR, Wiedenfeld DA, Wolfe DH, Sherrod SK. 2003. Genetic variation within and among fragmented populations of lesser prairie-chickens (*Tympanuchus pallidicinctus*). *Mol Ecol*. 12:675–683.

19 Bouzat JL, Johnson K. 2004. Genetic structure among closely spaced leks in a peripheral population of lesser prairie-chickens. *Mol Ecol*. 13:499–505.

20 Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol*. 14:2611–2620.

21 Mathur S, Tomeček JM, Tarango-Arámbula LA, Perez RM, DeWoody JA. 2023. An evolutionary perspective on genetic load in small, isolated populations as informed by whole genome resequencing and forward-time simulations. *Evolution*. 77:690–704.

22 Moritz C. 1994. Defining 'evolutionarily significant units' for conservation. *Trends Ecol Evol (Amst)*. 9:373–375.

23 Oyler-McCance SJ, John J, Quinn TW. 2010. Rapid evolution in lekking grouse: implications for taxonomic definitions. *Ornithol Monographs*. 67:114–122.

24 Kozma R, Melsted P, Magnússon KP, Höglund J. 2016. Looking into the past–the reaction of three grouse species to climate change over the last million years using whole genome sequences. *Mol Ecol*. 25:570–580.

25 Oh KP, Aldridge CL, Forbey JS, Dadabay CY, Oyler-McCance SJ. 2019. Conservation genomics in the Sagebrush Sea: population divergence, demographic history, and local adaptation in sage-grouse (*Centrocercus* spp.). *Genome Biol Evol*. 11:2023–2034.

26 Funk WC, Forester BR, Converse SJ, Darst C, Morey S. 2019. Improving conservation policy with genomics: a guide to integrating adaptive potential into U.S. Endangered Species Act decisions for conservation practitioners and geneticists. *Conserv Genet*. 20:115–134.

27 Nasman K, Rintz T, DiDonato G, Kulzer F. 2022. Range-wide population size of the lesser prairie-chicken: 2012 to 2022. *Western Association of Fish and Wildlife Agencies*. 1–43.

28 Rodgers RD. 2016. A history of Lesser Prairie-Chickens. In: Haukos DA, Boal CW, editors. *Ecology and conservation of lesser prairie-chickens*. 1st ed. Studies in Avian Biology. p. 15–38.

29 Black AN, et al. 2023. A highly contiguous and annotated genome assembly of the lesser prairie-chicken (Tympanuchus pallidicinctus). *Genome Biol Evol.* 15:evad043.

30 Dussex N, et al. 2021. Population genomics of the critically endangered kākāpō. *Cell Genomics.* 1:100002.

31 Hagen CA, Grisham BA, Boal CW, Haukos DA. 2013. A meta-analysis of lesser prairie-chicken nesting and brood-rearing habitats: implications for habitat management. *Wildl Soc Bull.* 37: 750–758.

32 Johnson JA, et al. 2023. Phylogenomics of the extinct heath hen provides support for sex-biased introgression among extant prairie grouse. *Mol Phylogenet Evol.* 189:107927.

33 Ottenburghs J, et al. 2017. Avian introgression in the genomic era. *Avian Res.* 8:1–11.

34 Via S. 2001. Sympatric speciation in animals: the ugly duckling grows up. *Trends Ecol Evol.* 16:381–390.

35 Bain MR, Farley GH. 2002. Display by apparent hybrid prairie-chickens in a zone of geographic overlap. *Condor.* 104:683–687.

36 Gould GM, Augustine JK. 2020. Multiple signals predict male mating success in the lek-mating lesser prairie-chicken (Tympanuchus pallidicinctus). *Behav Ecol Sociobiol (Print).* 74:1–17.

37 Ellsworth DL, Honeycutt RL, Silvy NJ, Rittenhouse KD, Smith MH. 1994. Mitochondrial-DNA and nuclear-gene differentiation in North American prairie grouse (Genus *Tympanuchus*). *Auk.* 111: 661–671.

38 McDonald L, et al. 2014. Range-wide population size of the lesser prairie-chicken: 2012 and 2013. *Wildl Soc Bull.* 38:536–546.

39 Fields TL, White GC, Gilgert WC, Rodgers RD. 2006. Nest and brood survival of lesser prairie-chickens in west central Kansas. *J Wildl Manage.* 70:931–938.

40 Leroy G, et al. 2018. Next-generation metrics for monitoring genetic erosion within populations of conservation concern. *Evol Appl.* 11:1066–1083.

41 Satam H, et al. 2023. Next-generation sequencing technology: current trends and advancements. *Biology (Basel).* 12:997.

42 Cook CN, Redford KH, Schwartz MW. 2023. Species conservation in the era of genomic science. *Bioscience.* 73:885–890.

43 Ackermann RR, et al. 2019. Hybridization in human evolution: insights from other organisms. *Evol Anthropol.* 28:189–209.

44 Stanton DWG, et al. 2019. More grist for the mill? Species delimitation in the genomic era and its implications for conservation. *Conserv Genet.* 20:101–113.

45 Korneliussen TS, Albrechtsen A, Nielsen R. 2014. ANGSD: analysis of next generation sequencing data. *BMC Bioinformatics.* 15: 356.

46 Narasimhan V, et al. 2016. BCFtools/RoH: a hidden Markov model approach for detecting autozygosity from next-generation sequencing data. *Bioinformatics.* 32:1749–1751.

47 Meisner J, Albrechtsen A. 2018. Inferring population structure and admixture proportions in low depth NGS data. *Genetics.* 210:719–731.

48 Skotte L, Korneliussen TS, Albrechtsen A. 2013. Estimating individual admixture proportions from next generation sequencing data. *Genetics.* 195:693–702.

49 Vieira FG, Lassalle F, Korneliussen TS, Fumagalli M. 2015. Improving the estimation of genetic distances from next-generation sequencing data. *Biol J Linn Soc.* 117:139–149.

50 Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol.* 32:268–274.