GBE

Temporal Genomics Reveal a Century of Genomic Diversity Shifts Across a Biodiversity Hotspot Avian Assemblage

Joseph D. Manthey [1,*], Amie E. Settlecowski [2,3], Yonas Meheretu [4,5], Garrett J. Behrends[1,6], Yann Bourgeois [7], Luke C. Campillo [8], Stéphane Boissinot [9,10], Ben D. Marks[2]

[1]Department of Biological Sciences, Texas Tech University, Lubbock, TX, USA
[2]Bird Collection Gantz Family Collections Center, The Field Museum, Chicago, IL, USA
[3]Office of Shared Research Facilities, University of Chicago, Chicago, IL, USA
[4]Department of Wildlife, Fish and Environmental Studies, Swedish University of Agricultural Sciences, Umeå, Sweden
[5]Department of Biology & Institute of Mountain Research & Development, Mekelle University, Mekelle, Tigray, Ethiopia
[6]Department of Biology, University of Missouri – St. Louis, St. Louis, MO, USA
[7]DIADE, University of Montpellier, CIRAD, IRD, Montpellier, France
[8]Hawaii Cooperative Studies Unit, University of Hawai'i at Hilo, Hilo, HI, USA
[9]New York University Abu Dhabi, Abu Dhabi, United Arab Emirates
[10]Center for Genomics and Systems Biology, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates

*Corresponding author: E-mail: jdmanthey@gmail.com.

Accepted: August 14, 2025

Abstract

Biodiversity has experienced tremendous shifts in community, species, and genetic diversity during the Anthropocene. Understanding temporal diversity shifts is especially critical in biodiversity hotspots, i.e., regions that are exceptionally biodiverse and threatened. Here, we use museomics and temporal genomics approaches to quantify temporal shifts in genomic diversity in an assemblage of eight generalist highland bird species from the Ethiopian Highlands (part of the Eastern Afromontane Biodiversity Hotspot). With genomic data from contemporary and historical samples, we demonstrate an assemblage-wide trend of increased genomic diversity through time, potentially due to improved habitat connectivity within highland regions. Genomic diversity shifts in these generalist species contrast with general trends of genomic diversity declines in specialist or imperiled species. In addition to genetic diversity shifts, we found an assemblage-wide trend of decreased realized mutational load, indicative of overall trends for potentially deleterious variation to be masked or selectively purged. Across this avian assemblage, we also show that shifts in population genomic structure are idiosyncratic, with species-specific trends. These results are in contrast with other charismatic and imperiled African taxa that have largely shown strong increases in population genetic structure over the recent past. This study highlights that not all taxa respond the same to environmental change, and generalists, in some cases, may even respond positively. Future comparative conservation genomics assessments on species groups or assemblages with varied natural history characteristics would help us better understand how diverse taxa respond to anthropogenic landscape changes.

Key words: museomics, temporal genomics, mutational load, conservation genetics.

Significance

Understanding how biodiversity is changing is essential given a continually changing planet, particularly in regions with high diversity. Here, we used genomic data from modern and historical samples of eight bird species from the Ethiopian Highlands to document shifts in genomic variation over the past hundred years. We identify that shifts in genomic variation may be consistent across species or idiosyncratic depending on the type of variation studied and provide a baseline for genomic diversity shifts in Eastern Afromontane birds.

Introduction

Humans have influenced aspects of the three hierarchical levels of biodiversity throughout the Anthropocene; some communities have experienced homogenization (Capinha et al. 2015; Li et al. 2020; Nogué et al. 2021), species diversity has decreased due to extinctions (Turvey and Crees 2019), and species have exhibited shifts in intraspecific genetic diversity globally (van der Valk et al. 2019; Benham et al. 2024; Blanchet et al. 2024; Shaw et al. 2025). Using genomic data from contemporary and historical samples (i.e., museomics) provides opportunities to assess how genetic diversity and structure have changed through time (Bieker and Martin 2018; Schmitt et al. 2019). Museomics studies assessing shifts in genetic diversity and structure have primarily focused on specialist or endangered vertebrates, often identifying reductions in genetic diversity, increases in mutational load, and increased population genetic structure through time (Feng et al. 2019; van der Valk et al. 2019; Curry et al. 2021; Dussex et al. 2021; Mathur and DeWoody 2021; Jackson et al. 2022; Dussex et al. 2023a; Sánchez-Barreiro et al. 2023; Blanchet et al. 2024). What remains lacking is an understanding of how genomic trajectories are shifting through time in generalist species and assemblages. A better understanding of genomic variation and its changes through time in taxa with varied natural history characteristics is necessary for effective comparative conservation genomics (Teixeira and Huber 2021).

Assessing assemblage-wide shifts in genomic trajectories is particularly important in biodiversity hotspots (Myers et al. 2000; Zachos and Habel 2011), regions of exceptionally high biodiversity and threats to that biodiversity. Because biodiversity hotspots are threatened with previous or ongoing loss of geographic area and have high endemicity, quantifying the genomic trajectories of representative taxa in these regions may provide snapshots into overall population trends useful for conservation biology and conservation genomics. The Eastern Afromontane Biodiversity Hotspot has high endemicity, with hundreds of endemic birds and mammals and more than 2,000 endemic plants, among others (Gordon et al. 2012). The Ethiopian Highlands make up a large portion of the geographic area of the Eastern Afromontane Biodiversity Hotspot and is a species-rich (Yalden and Largen 1992; Friis et al. 2001; Largen and Spawls 2010), largely contiguous region of tropical highland habitat. The Ethiopian Highlands are composed of two large massifs separated by the Great Rift Valley (GRV): the Harar Massif to the southeast and the Abyssinian Massif to the northwest (Fig. 1). The Abyssinian Massif is further divided by the Blue Nile Valley (BNV) separating the Choke Mountains from the Central Highlands (Fig. 1). These lowland biogeographic barriers have shaped intraspecific phylogeographic structure in many Ethiopian Highlands taxa, including birds (Manthey et al. 2022; Behrends et al. 2024), mammals (Gottelli et al. 2004; Belay and Mori 2006; Bryja et al. 2018; Razgour et al. 2019; Kostin et al. 2020; Mizerovská et al. 2020; Komarova et al. 2021), frogs (Evans et al. 2011; Freilich et al. 2016; Manthey et al. 2017; Reyes-Velasco et al. 2018; Reyes-Velasco et al. 2018), and plants (Kebede et al. 2007; Silvestrini et al. 2007). As such, species in this community provide an opportunity to assess not only genomic diversity shifts, but also changes in genomic structure across distinct highland regions through time.

Here, we use whole-genome sequencing of historical and modern samples of eight Ethiopian Highlands bird species to quantify population genomic changes over the past century. The focal species are all high elevation generalists (i.e., they can thrive in a variety of habitats in the higher elevations of the Ethiopian Highlands); they are found in forests, woodlands, shrub, edge habitats, and in some cases grasslands (Kittelberger et al. 2021) in both undisturbed and disturbed areas (Asefa et al. 2017). Additionally, they can be found in highland farmlands, seminatural habitats, and settlements and cities with trees and shrubs (Gove et al. 2013; Buechley et al. 2015; Marcacci et al. 2020; Shiferaw and Yazezew 2021). None of the focal species are listed as threatened or endangered on the International Union for Conservation of Nature Red List (IUCN 2024). We chose the eight focal species because they are highland generalists that can often be found co-occurring in the same habitats and most of them are Horn of Africa endemics. The Ethiopian Highlands have a history of thousands of years of agricultural land conversion (Nyssen et al. 2004; Hurni et al. 2010) and recent varied land management and conservation practices in different areas (Munro et al. 2008; Hurni et al. 2010; Nyssen et al. 2015). It is therefore difficult to a priori predict whether

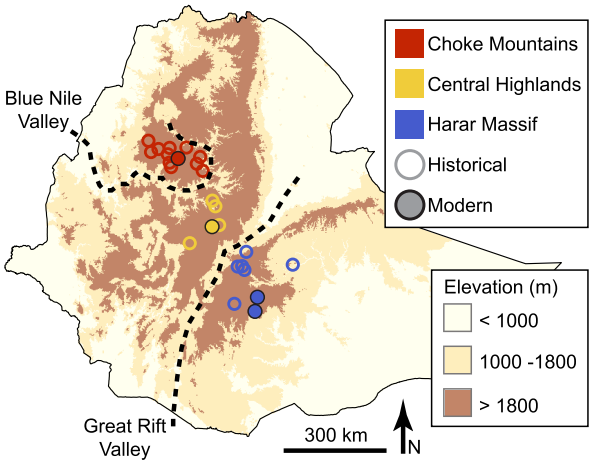

Fig. 1. Historical (1925 to 1931; $n = 51$) and modern (2016 to 2017; $n = 55$) sampling map of Ethiopian Highlands birds. Dotted lines represent approximate regions of biogeographic barriers. The GRV separates the Ethiopian Highlands into the Abyssinian and Harar Massifs. The BNV is an additional biogeographic barrier separating the Choke Mountains from the Central Highlands on the Abyssinian Massif.

we expect generalists' population genomic trajectories to have improved or worsened over the past century. However, we may predict that the assemblage overall has consistent shifts or lack thereof in genomic diversity and structure over the past century of change.

Using genomic data from eight generalist highland species, we show an assemblage-wide trend of increase in genomic diversity and decrease in realized mutational load through time. In contrast, we found that shifts in genomic structure through time were idiosyncratic, with different trends across species.

Results

Temporal Genomics of Eight Ethiopian Highlands Passerines

We obtained whole-genome sequencing data at ~5 to 30× coverage for three populations of eight montane passerine (Aves: Passeriformes) bird species from the Ethiopian Highlands (Fig. 1; Tables S1 and S2). Our focal taxa are Rüppell's Robin-Chat (*Cossypha semirufa*), Streaky Seedeater (*Crithagra striolata*), Brown-Rumped Seedeater (*Crithagra tristriata*), Abyssinian Slaty-Flycatcher (*Melaenornis chocolatinus*), Tacazze Sunbird (*Nectarinia tacazze*), Abyssinian Catbird (*Sylvia galinieri*), Abyssinian Thrush (*Turdus abyssinicus*), and Ethiopian White-Eye (*Zosterops poliogastrus*). Our temporal sampling included 55 modern (sampled 2016 to 2017) and 51 historical samples (sampled 1925 to 1931). We used specimens from the 1920s for this effort because this is among the earliest series of specimens of the focal taxa that included multiple samples per species per locality.

Assemblage-Wide Shifts in Genomic Diversity and Mutational Load

We measured genetic diversity as observed heterozygosity per individual (Fig. 2; i.e., number of heterozygous sites/total sites genotyped) and used a linear mixed effects model (LMEM) to test for changes through time while accounting for differences between species and localities. Here, we identified an assemblage-wide increase of genomic diversity through time ($\chi^2 = 30.2$, $P < 0.001$; Table 1). We quantified runs of homozygosity (ROH) in 100 kbp windows as an indicator for the presence of inbreeding; we found no large ROH across the study species, suggesting little to no inbreeding (results not shown).

As a measure of shifts in genomic health through time, we estimated potential mutational load and realized mutational load (Mathur et al. 2023). Here, potential load is the proportion of functional nucleotide substitutions in the genome that are potentially deleterious (or weakly deleterious), whereas realized load is the proportion of potentially deleterious variants that are found in the homozygous state. We identified an assemblage-wide decrease in realized load through time (Fig. 2; Table 1; $\chi^2 = 5.2$, $P = 0.022$) and no significant shift in potential load through time (Fig. 2; Table 1; $\chi^2 = 0.2$, $P = 0.655$).

Idiosyncratic Genomic Structure Shifts Through Time

We used genome-wide single nucleotide polymorphisms (SNPs) from each species to estimate genetic structure using phylogenetic networks and principal components analyses (PCA). Concordant with two previous phylogeographic studies on some of our focal species (Manthey et al. 2022; Behrends et al. 2024), most species exhibited genomic structure consistent with differentiation between the three highland regions (Fig. 3), with weak genomic structure in the Tacazze Sunbird, Abyssinian Thrush, and Ethiopian White-Eye. We used variance partitioning to quantify contributions of geographic distance, biogeographic barriers, and time on genetic structure, followed by assessing statistical significance of explanatory variables on genetic structure using multiple regression of distance matrices (Fig. 3). Across species, geography, biogeographic barriers, time, and the interactions of these variables explained between 31% and 94% of the variance in genetic structure (Fig. 3). However, the relative contributions of the explanatory factors varied considerably by species. Most species had a significant impact of geography (either geographic distance or biogeographic barriers or both) on genetic structure (excepting the Tacazze Sunbird). Similarly, most species also had an impact of time on genetic structure (excepting the Rüppell's Robin-Chat). Taken together,

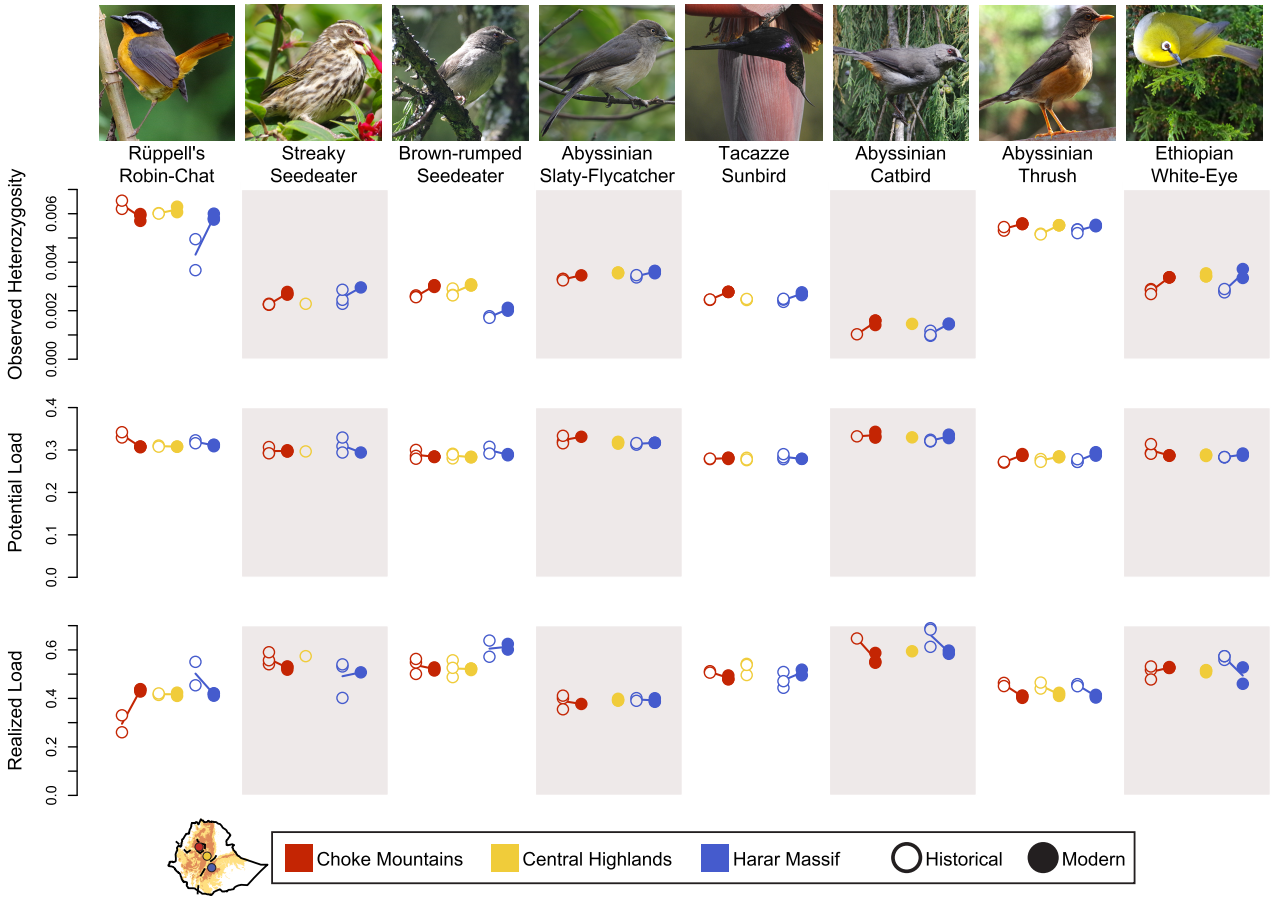

Fig. 2. Genomic diversity shifts over the past century in eight Ethiopian Highlands forest bird species. Points represent estimates per individual and lines represent averages per population per time period. Relative to historical populations, contemporary populations have higher genetic diversity (results from LMEM; $\chi^2 = 30.2$, $P < 0.001$), no consistent changes in potential load ($\chi^2 = 0.2$, $P = 0.655$), and lower realized load ($\chi^2 = 5.2$, $P = 0.022$). Key for geographic localities at bottom of figure. Tacazze Sunbird photo by Asrat Ayalew. All other photos by Joseph Manthey.

Table 1 Results of LMEM testing whether response variables have changed through time while using different alignment datasets (full, using MapDamage, and downsampled), and including different types of polymorphisms

Response variable	Alignment dataset	Sites used	LMEM results	Difference through time
Observed heterozygosity	Full	All sites	$\chi^2 = 30.2$, $P < 0.001$	+0.00035
	MapDamage	All sites	$\chi^2 = 39.3$, $P < 0.001$	+0.00040
	Downsampled	All sites	$\chi^2 = 65.6$, $P < 0.001$	+0.00084
	Full	Transversions only	$\chi^2 = 60.3$, $P < 0.001$	+0.027
	Full	Biallelic sites (Structure Dataset)	$\chi^2 = 29.7$, $P < 0.001$	+0.023
Potential load	Full	Polymorphic sites in genes	$\chi^2 = 0.2$, $P = 0.655$	…
	MapDamage	Polymorphic sites in genes	$\chi^2 = 0.2$, $P = 0.628$	…
	Downsampled	Polymorphic sites in genes	$\chi^2 = 88.2$, $P < 0.001$	+0.015
	Full	Transversions only in genes	$\chi^2 = 12.0$, $P < 0.001$	+0.004
Realized load	Full	Polymorphic sites in genes	$\chi^2 = 5.2$, $P = 0.022$	−0.018
	MapDamage	Polymorphic sites in genes	$\chi^2 = 11.7$, $P < 0.001$	−0.024
	Downsampled	Polymorphic sites in genes	$\chi^2 = 51.7$, $P < 0.001$	−0.044
	Full	Transversions only in genes	$\chi^2 = 34.8$, $P < 0.001$	−0.041

P values below 0.05 indicate a significant change through time. Using the different datasets, contemporary populations have higher genetic diversity and lower realized load, while there is mixed support for either no change or increased potential load in contemporary populations. Note that absolute values of differences through time should not be directly compared among all datasets because of the varying types of sites included in each dataset.

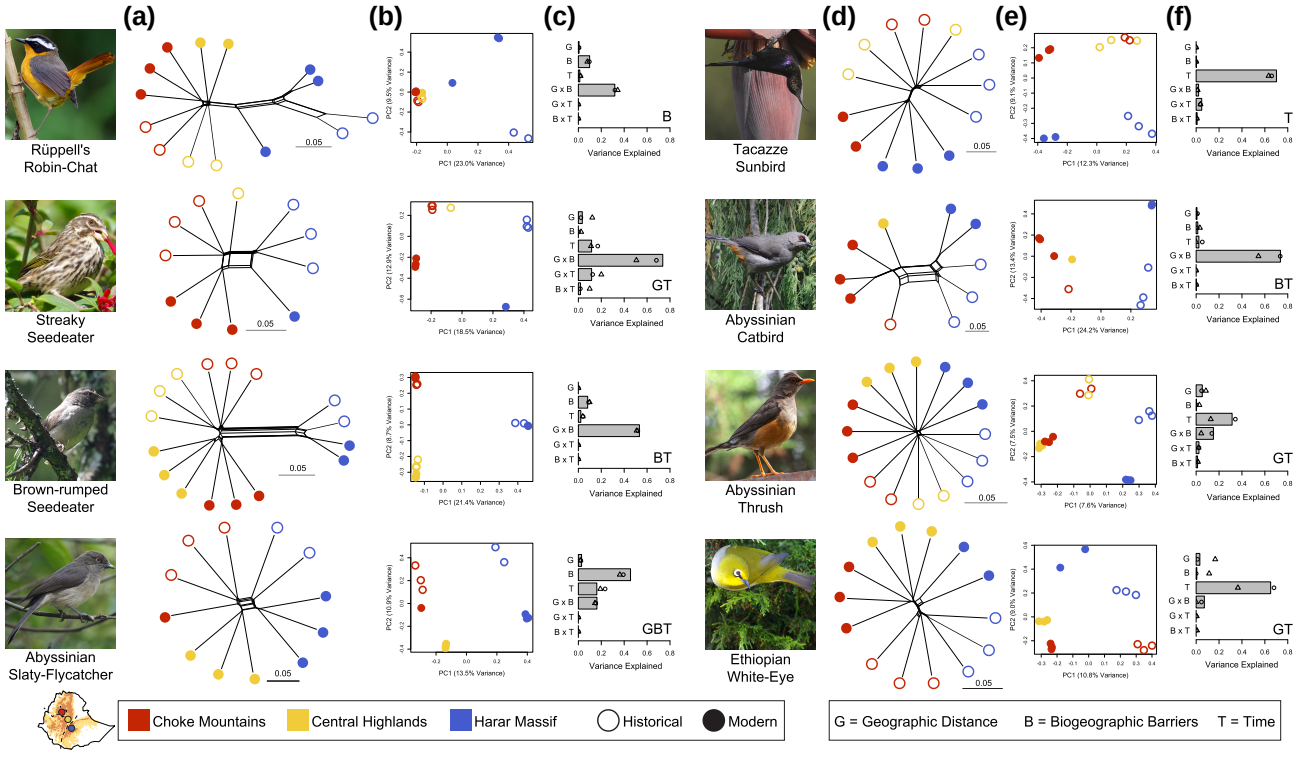

Fig. 3. Population genomic structure of historical and modern samples for the eight focal Ethiopian Highlands bird species. a, d) SplitsTree phylogenetic networks. b, e) PCA. c, f) Variance partitioning of genetic distance (gray bars). Explanatory variables are geographic distance, biogeographic barriers, time between sampling, and their interactions (denoted by x). In addition to the full dataset, variance partitioning results are shown for the MapDamage rescaled alignment dataset (circles) and for the full dataset including only transversions (triangles). In bottom right corner of each plot are the factors that significantly contributed ($P < 0.05$) to genetic distances based on multiple regression of distance matrices. Key for geographic localities and explanatory factors at bottom of figure. Tacazze Sunbird photo by Asrat Ayalew. All other photos by Joseph Manthey.

geographic and temporal separation have both structured population genomic variation in this assemblage, but trends are not consistent in magnitude across species.

Consistency of Results Using Different Datasets and Filtering Strategies

When comparing modern and historical sequencing datasets, it is important to thoroughly assess if any results are artefacts of systematic dataset differences or due to data filtering strategies. As such, we assessed consistency of results across several datasets, checking for the effects of missing data, sequencing coverage, rescaling base qualities to account for possible DNA damage, and types of polymorphisms included. Altogether, this resulted in 10 datasets per species to assess how the above-mentioned factors did or did not impact results (Table S3).

Missing Data

First, we assessed whether missing data affected results, either by being more prevalent in historical samples or more prevalent in polymorphic versus invariant genotyped sites. Using a LMEM to account for geography and species differences, there was not a significant effect of time on missing proportion of all sites (Fig. S1; $\chi^2 = 1.5$, $P = 0.219$), but somewhat of an effect (marginally nonsignificant) of time on missing proportions of polymorphic sites (Fig. S2; $\chi^2 = 3.4$, $P = 0.064$). The relationship between missing proportion of all sites and of polymorphic sites approximated 1:1 (Fig. S2). Because we found a marginal but nonsignificant effect of time on missing proportions of polymorphic sites, we measured heterozygosity with a biallelic sites dataset including no missing data (same as used for PCA and genetic distances). Here, we found a strong positive association between H_O of the full dataset and of the genetic structure biallelic sites dataset for all species (Fig. S3; all $P < 0.01$) and a LMEM of this H_O dataset identified the same result as the full dataset: an assemblage-wide increase of genomic diversity through time ($\chi^2 = 29.7$, $P < 0.001$; Table 1).

Rescaling Bases With MapDamage

Rescaling base call qualities using MapDamage (Jónsson et al. 2013) has been a common practice in historical and ancient DNA studies. However, recent work has shown

that rescaling base call qualities also has the potential to create reference biases in rescaled sequencing data (Koptekin et al. 2025). Here, we ran our full computational pipelines with and without MapDamage base rescaling. We found that estimates of genetic diversity, mutational load, and genetic structure all showed strongly concordant results between the full dataset and the MapDamage adjusted dataset (Figs. S4 to S6, Table 1). In addition, with variance partitioning we identified similar trends in how each explanatory variable and their interactions shaped genetic distances (Fig. 3).

Sequencing Coverage

Variable genomic sequencing coverage may lead to differential ability to genotype different samples and has the potential to lead to biases in results if differences in coverage are systematic to a single type of sample. Using a LMEM to account for geography and species differences, our modern samples had ∼4× greater coverage than historical samples ($\chi^2 = 13.65$, $P < 0.001$), but this varied per species, with some having more coverage in the historical datasets (e.g., Abyssinian Catbird and Abyssinian Thrush; Fig. S7). Further, genomic coverage was not significantly associated with genetic diversity for five of eight species (Fig. S4). To further explore this association, we used downsampled alignment files (BAM files) for downstream analyses, where each individual was randomly downsampled to ∼6 to 7× genomic coverage. With the downsampled datasets, we found different absolute values, but highly correlated relative values of genetic diversity relative to the full dataset (Fig. S4). In the downsampled datasets, the significant association between diversity and coverage disappeared for the three species for which it was present in the full dataset. A LMEM of this downsampled dataset identified the same result as the full dataset: an assemblage-wide increase of genomic diversity through time ($\chi^2 = 60.3$, $P < 0.001$; Table 1). Similarly, the results and interpretation of realized mutational load values were similar to the full dataset with a decrease in realized mutational load through time ($\chi^2 = 51.7$, $P < 0.001$; Fig. S5; Table 1). In contrast, potential load estimates with the downsampled dataset were not consistent with the full dataset (Fig. S5; Table 1), and the LMEM showed a significant increase in potential load through time, whereas the full dataset did not (Table 1).

Polymorphism Type

In historical samples, postmortem DNA damage often manifests as transitions at the ends of sequencing reads (e.g., C to T or G to A transitions). Here, we visualized relative nucleotide substitution profiles in modern and historical sequencing reads (Fig. S8). The nucleotide substitution profiles suggested enzymatic repair during laboratory work largely reduced any substitution pattern differences between modern and historical samples (Fig. S1). In addition, we repeated measures of genetic distances and mutational load using only transversions to eliminate potential effects of erroneous transition genotyping calls. Here, we found generally strong correlations between genetic distances calculated with all polymorphism types versus just transversions (all $r \geq 0.6$; Fig. S6) and with variance partitioning we identified similar trends in how each explanatory variable and their interactions shaped genetic distances (Fig. 3). Similarly, mutational load estimates were consistent between the transversions only datasets and the full datasets in all but one species (Streaky Seedeater; Fig. S5). However, the LMEM showed a significant increase in potential load through time, where the full dataset did not (Table 1).

Overall Patterns

In sum, across all datasets, genetic diversity increased through time and realized mutational load decreased through time (Fig. 2; Table 1). In contrast, potential realized load patterns varied between datasets, where calculations either showed no significant change through time or a slight increase through time (Fig. 2; Table 1). Genetic structure results, as measured with genetic distances, were similar across datasets (Fig. S6). Additionally, the results of variance partitioning showed similar forces shaping genetic structure across geography and through time (Fig. 3).

Note on Harar Massif Sampling

In the Choke Mountains and Central Highlands, the modern and historical sampling were from the same mountain ranges (Fig. 1). In the Harar Massif, it is important to note that our modern samples were only from the Bale Mountains, while most of the historical samples were from the Arsi Mountains (points closest to GRV in Fig. 1; ∼100 to 120 km from our Bale Mountains sampling sites). Some small mammals exhibit genetic differentiation between these two mountain ranges in the Harar Massif (Kostin et al. 2019), but it has not previously been explored in birds. Based on our results for genetic diversity and structure discussed above, we believe these different sampling locations through time in the Harar Massif only impacted our results for the Rüppell's Robin-Chat in this region. This species exhibited strong differences in genetic diversity in this region through time (Fig. 2), large differences in genetic structure through time (based on long branches in network [Fig. 3] and large shifts in F_{ST} through time [Fig. S9]). In contrast, all other species did not show extreme differences in Harrar Massif genetic diversity or structure through time relative to the Choke Mountains or Central Highlands populations (Figs. 2 and 3).

Discussion

Shifts in genomic diversity through time have generally been studied in charismatic, threatened, or specialist

species (Feng et al. 2019; van der Valk et al. 2019; Curry et al. 2021; Dussex et al. 2021; Mathur and DeWoody 2021; Jackson et al. 2022; Dussex et al. 2023a; Sánchez-Barreiro et al. 2023; Blanchet et al. 2024). Globally, species with genetic diversity measurements spanning several decades have generally shown a trend for genetic erosion (i.e., genetic diversity loss) (Shaw et al. 2025). For example, African megafauna such as cheetahs, gorillas, and lions have experienced population declines resulting in reduced genomic diversity through time (Terrell et al. 2016; van der Valk et al. 2019; Curry et al. 2021). In bird museomics studies across the world, there has been documentation of mixed patterns in single species studies; as examples, mutational load is increasing through time in an island endemic pigeon (Jackson et al. 2022), varied trends of genomic diversity loss or maintenance were observed in different populations of a continental sparrow (Benham et al. 2024), and an island crow has experienced genomic diversity loss through time (Blanchet et al. 2024). In contrast, we find increases in population genomic diversity over the past century (Fig. 2). This trend was consistent across datasets testing for impacts of confounding factors shaping the trends seen here (Figs. S3 and S4; Table 1). Additionally, we found a lack of strong shifts in phylogeographic structure through time (Fig. 3) in our focal taxa. We hypothesize that the contrasting results for our focal species is likely due to their generalist nature.

Two mechanisms that could increase genomic diversity through time are (i) an increase in census population sizes or (ii) increased connectivity across partially connected and fragmented subpopulations. Both these mechanisms could be due to new or improved favorable habitats for these generalists. An increase in census population size could also be due to reduced competition with specialists not included in our study. Increased local or regional connectivity could also be due to reduction in small-scale habitat fragmentation. As the causes for these two mechanisms overlap, they are not mutually exclusive.

Increased genomic diversity through time due to increases in census population sizes would likely be a slow process, as there is generally a lag time between changes in census population sizes to realized changes in genomic diversity (Gargiulo et al. 2025). Generally, these lag times would be associated with mutation rates and life history traits, such as generation time and reproductive output. All the focal species are passerine birds and likely have moderately short generation times (e.g., 2 to 3 yr; Reid et al. 2019), potentially providing up to 50 generations since initial sampling for genomic diversity to increase. In contrast, increased connectivity across partially connected and fragmented subpopulations could lead to genomic mixing in few generations. At large scales, we did not find evidence for increased connectivity through time across biogeographic barriers (Fig. S9) and we did not find large consistent increases in LD decay between time periods (e.g., LD decay shifts upward would indicate very recent admixture among distinct lineages; Fig. S10). At small spatial scales, we do not have sufficient spatial and temporal sampling to directly measure increased connectivity. Because of the timescales of these two mechanisms, we hypothesize that increased connectivity between slightly isolated subpopulations more plausibly drove shifts in genomic diversity.

Regardless of the mechanism, an increased amount of habitat and connectivity would likely be required for these species to exhibit increased genomic diversity through time. Because of the generalist nature of the focal species, an increase in any forest, woodland, or shrubby habitat in natural or partially natural settings (e.g., complex farmlands or vegetated settlements) would likely exert a positive effect for population sizes in these taxa. However, identifying a habitat quality baseline from a century ago to compare with today is difficult; the Ethiopian Highlands have a complex history of increases and decreases in forest, secondary woodland, and scrub habitats over the past couple millennia (Darbyshire et al. 2003; Nyssen et al. 2004; Hurni et al. 2010). Accounts from the mid-20th century indicate there was little pristine forest cover in the Ethiopian Highlands at that time (Logan 1946; Ritler 1997), and tree cover has likely been consistent or increasing in settlements throughout the 1900s (McCann 1997). Additionally, land management and conservation practices are varied across regions, but some photographic resurvey work has demonstrated improved vegetative cover since the 1970s in some locales, while other areas have experienced continued degradation (Munro et al. 2008; Nyssen et al. 2015). These combined accounts suggest no simple relationship in landscape change across the heterogeneous Ethiopian Highlands landscape over the past century. Regardless, the changes have increased the genomic diversity in this assemblage of generalist highland passerines.

In concert with our observations of genomic diversity increases through time, we generally found a decrease in realized mutational load through time (Fig. 2), but mixed results for potential load among datasets where we found either no significant change or an increase through time (Fig. 2; Table 1). These patterns are consistent with both mechanisms for increased genomic diversity mentioned above, where increased genomic diversity is associated with more potential load, while there is also increased sheltering of potentially deleterious homozygous variants. Genome-wide signatures and temporal shifts in mutational load are important to understand from a conservation perspective (Blomqvist et al. 2010; Mathur and DeWoody 2021; Bertorelle et al. 2022; Dussex et al. 2023b; Bourgeois et al. 2024); these measures represent populations' trends and tendencies for potentially deleterious variation to be masked or selectively purged in larger or increasing populations but largely succumb to randomness of drift in smaller populations, potentially leading to

population-wide decreases in fitness (Blomqvist et al. 2010; Bertorelle et al. 2022; Dussex et al. 2023b). Overall, patterns of increasing genomic diversity and decreasing realized mutational load through time suggest that Ethiopian Highlands generalist birds are on positive population genomic trajectories. Though many conservation studies in Africa focus on declining species that ultimately show negative biodiversity trends, this study offers a more hopeful perspective; some species may respond positively to environmental change, provided that habitat connectivity is maintained. Similar studies on other generalist species and species with varied natural history characteristics are needed to better understand how diverse taxa respond to anthropogenic landscape changes.

Materials and Methods

Study System

Our focal taxa are eight montane passerine birds (Aves: Passeriformes), including the Rüppell's Robin-Chat (*C. semirufa*), Streaky Seedeater (*C. striolata*), Brown-Rumped Seedeater (*C. tristriata*), Abyssinian Slaty-Flycatcher (*M. chocolatinus*), Tacazze Sunbird (*N. tacazze*), Abyssinian Catbird (*S. galinieri*), Abyssinian Thrush (*T. abyssinicus*), and Ethiopian White-Eye (*Z. poliogastrus*). These species' geographic distributions are generally limited to high elevations in the Horn of Africa; the most widespread species is the Abyssinian Thrush that is found in montane regions from Malawi to Eritrea, and the most geographically restricted species is the Abyssinian Catbird, which is endemic to the Ethiopian Highlands. Although the realized niches of all these species are not identical, all these species are forest generalists, in that they may be found in interior forest, forest edge, stunted or regrowing forest, and even urban or inhabited areas with sparse trees. Some of the species can also be found frequently in nonforested urban areas; for example, the seedeaters or thrush can be found in the Addis Ababa Airport parking lot, which is sparsely vegetated. These species often co-occur, and indeed all the species were captured in the same locality for modern sampling in the Choke Mountains.

Whole-Genome Sequencing of Modern Samples

We used a QIAGEN DNeasy blood and tissue extraction kit to obtain genomic DNA from 9 blood samples of *C. striolata* and *N. tacazze*. We sent these DNA extracts to the Texas Tech University Center for Biotechnology and Genomics, where they used the Illumina DNA Prep kit to create sequencing libraries. After quality checking with the Agilent TapeStation 4200, the samples were sequenced on part of a single lane of an Illumina NovaSeq6000 S4 flow cell (150 × 150 bp) with other samples from unrelated projects. For an additional 46 individuals, we obtained sequencing data that were generated and used for previous phylogeographic studies in the focal taxa (Manthey et al. 2022; Behrends et al. 2024). Total numbers per species are presented in Table S1.

Whole-Genome Sequencing of Historical Samples

We sampled toepads from 51 historical specimens from eight species (Table S1) following extensive precautions to limit contamination of the toepads by modern DNA. We did not prepare or handle fresh specimens or enter a modern molecular laboratory prior to cutting toepads on sampling days. We cut toepads in a separate room from the specimen preparation laboratory at a collections bench that was thoroughly cleaned by dusting and wiping down with freshly prepared 10% bleach followed by 70% ethanol. While cutting toepads we wore disposable sleeves, a surgical mask, and two pairs of gloves. For each toepad we replaced our top pair of gloves, used a fresh razor blade, and cut the largest possible wedge (mean, $M = 1.76$ mg; standard deviation, $SD = 0.91$) from the toepad of the more exposed hallux as long as it was accessible. We deposited the samples in sterile microcentrifuge tubes that were not opened again prior to sample processing.

We completed all pre-PCR molecular laboratory work following ancient DNA protocols (Fulton and Shapiro 2019) in a positively pressurized clean laboratory to minimize contamination by modern DNA. To monitor for contamination, we introduced a negative control for every batch of 11 toepad samples at each stage of processing that we then carried through sequencing, resulting in 18 total negative control libraries. We attempted to minimize potential contaminating DNA on the exterior of the toepad samples by performing a brief enzymatic predigestion following the methods of Settlecowski et al. (2023). Briefly, we digested the outer layer of each toepad in 180 µL of digestion buffer (30 mM Tris–HCl, 10 mM EDTA, 1% SDS) and 20 µL of proteinase K for 3 min at 37 °C and 1,000 RPM. After the predigestion, we discarded the digestion solution and successively washed the toepad for 5 min at room temperature and 1,000 RPM, first with 500 µL of 70% ethanol and then 500 µL STE. We then purified DNA from each toepad via phenol chloroform DNA extraction followed by ethanol precipitation following the methods of Tsai et al. (2020) with a few modifications. Our modifications were to use the digestion buffer described above in place of Qiagen Buffer ATL, begin each digestion with 40 µL rather than 20 µL of proteinase K, exclude dithiothreitol following overnight digestion, and to intermittently vortex samples rather than mash with forceps during the digestion period. We resuspended the precipitated DNA in 45 µL of 10 mM Tris–HCl and used 2 µL to measure the DNA concentration via Qubit High Sensitivity dsDNA assay. We treated each toepad DNA sample with NEB PreCR Repair Mix to repair

deaminated cytosines, DNA nicks, among other types of DNA damage expected in historical samples. We followed the sequential reaction protocol provided with the kit and performed up to two treatment reactions per sample, inputting no more than the maximum 500 ng of DNA per reaction. We cleaned up the repaired DNA via Qiagen MinElute columns, eluted the DNA in 17 µL of 10 mM Tris–HCl, and used 2 µL to measure the DNA concentration via Qubit High Sensitivity dsDNA assay.

We prepared a genomic sequencing library for each toepad sample using the IDT xGen ssDNA & Low Input DNA Library Prep Kit, because a prior study suggested that this kit returned a higher proportion of target historical DNA from bird toepad of the same age of samples herein when compared to the commonly used KAPA Hyper Prep Kit and another single-stranded DNA library preparation kit (Settlecowski et al. 2023). We prepared libraries following manufacturer protocol with several modifications. We performed larger ratio SPRI bead cleanups to avoid removing smaller DNA fragments expected from historical samples, using a homebrew SPRI bead solution (Rohland and Reich 2012). We performed a 1.8× SPRI cleanup following extension, 1.6× SPRI cleanup following ligation, and 1.4× SPRI cleanup following library amplification based on expected input DNA and library fragment sizes (Settlecowski et al. 2023). Following ligation, we amplified each library in triplicate with six cycles per PCR, using KAPA HiFi HotStart Uracil + ReadyMix rather than the IDT kit-provided PCR reagents to facilitate amplification of any library molecules with remaining uracils. Lastly, we used generic iTru5 and iTru7 indexed primers (Glenn et al. 2019) rather than IDT xGen indexed primers to index each library with a unique i5 and i7 sequence. Following library amplification, pooling by sample, and cleanup we measured the mean library molecule size via Agilent Bioanalyzer High Sensitivity DNA Kit assay and calculated the concentration of adapter-ligated library molecules by qPCR with KAPA Library Quantification Kit. We combined all sample libraries in a 10 nM pool that was sent to the Texas Tech University Center for Biotechnology and Genomics for sequencing on three NovaSeq6000 S4 flow cells (100 × 100 bp).

Quality Control and Alignment of Historical Sequence Data

First, we used seqtk v1.3 (Shen et al. 2016) to trim adaptase tails attached during sequencing library preparation. We then used hts_SuperDeduper v1.3.2 (Petersen et al. 2015) to remove PCR duplicates from the raw sequencing reads. Next, we merged any reads with small insert sizes using SeqPrep (Robbins et al. 2011) and removed low complexity reads using the remove_low_complex.py script from the nf-polish pipeline (available at: github.com/MozesBlom/nf-polish). We then aligned both merged and unmerged sets of reads to the *Ficedula albicolis* reference genome (ENSEMBL release FicAlb1.5 v105, GCA_000247815.2) (Ellegren et al. 2012) using the mem algorithm of the BWA v2.2.1 program (Li and Durbin 2009). We used this genome because it is a songbird (as are all our focal taxa), it is a chromosome-scale assembly, there is generally high synteny among birds (Derjusheva et al. 2004; Griffin et al. 2008), and this genome was already annotated for mutational load calculations in a SnpEff (Cingolani et al. 2012) database. We used samtools v1.6 (Li et al. 2009) to convert SAM files to BAM format and to merge the alignments for the merged and unmerged read sets. We cleaned and sorted the BAM files with Picard Tools (available at: broadinstitute.github.io/picard).

Quality Control and Alignment of Modern Sequence Data

We used bbduk (Bushnell 2014) to quality filter the raw sequencing data and then used the mem algorithm of the BWA v2.2.1 program (Li and Durbin 2009) to align the filtered reads to the *F. albicolis* reference genome. We used samtools v1.6 (Li et al. 2009) to convert the SAM file to BAM format, followed by cleaning and sorting the BAM files with Picard Tools. We used samtools to measure the genome-wide depth of sequencing coverage for each individual.

Different Alignment Datasets

To test for impacts of rescaling base qualities or sequencing coverage differences, we created three alignment datasets per species: (i) the full dataset, (ii) an alignment dataset with rescaled bases, and (iii) a downsampled alignment dataset. Here, we used samtools to measure the genome-wide depth of sequencing coverage for each individual. With this information, we used the "DownsampleSAM" function of Picard Tools to create a downsampled dataset. Lastly, we used mapDamage v2.3.0 (Jónsson et al. 2013) to quantify DNA damage patterns in the BAM files and rescale base qualities for any reads with substitutions that were likely the result of DNA damage (MapDamage dataset). All three alignment datasets were used in downstream genotyping and filtering schemes described below and summarized in Table S3. We also used output from mapDamage to visualize any excess of C to T transitions at the ends of sequencing reads in historical relative to modern samples. Overall, we found no excess of C to T transitions at the ends of reads, and historical and modern samples had similar substitution patterns in each species (Fig. S8). These results strongly suggest the UDG enzymatic treatment of historical samples, and the bioinformatics filtering schemes largely removed any signatures of post-mortem damage in the sequences. Additionally, we observed a decrease in substitution rates overall toward

read ends, which we hypothesize is due to a combination of quality filtering and proximity to soft clipped bases because of alignment to a nonconspecific reference. The Rüppell's Robin-chat and the Abyssinian Slaty Flycatcher are both in the same family as the reference genome (Muscicapidae) and show the least reduction in substitutions near the beginning of the reads, supporting this hypothesis.

Genotyping and Genotype Filtering

Using the final BAM files, we genotyped each individual using the bcftools v1.17 (Li 2011) mpileup command and retained sites that had a minimum sequencing depth of eight reads. Variant call format (VCF) files from all individuals were merged using the bcftools merge command. On the combined VCF files, we used vcftools v0.1.16 (Danecek et al. 2011) to further filter to 10 datasets per species (Table S3), with three main types of filtering. For genetic structure analyses, we filtered for no missing data allowed, only including biallelic sites, a minimum minor allele count of two (i.e., removing singletons), and a minor allele frequency less than 0.5. For genetic diversity estimates, we allowed up to two individuals with missing data per site, included variant and invariant sites, no minimum minor allele count, and a minor allele frequency less than 0.5. For mutational load analyses, we allowed up to two individuals with missing data per site, included only biallelic sites found in genes, a minimum minor allele count of two, and a minor allele frequency less than 0.5. Further, to identify if polymorphism types included shaped any trends in our datasets, some of these datasets were filtered again to only include transversions. For all datasets, we removed the sex chromosomes to remove any impacts of different ploidy between sexes.

Genetic Diversity and Mutational Load

We measured observed heterozygosity (H_O) for each individual as an estimate of genetic diversity. For each individual, H_O is measured as the number of heterozygous sites divided by the total number of sites genotyped. Because this is a genetic diversity measure of individual diploid genomes, estimates should be representative of population-level nucleotide diversity. Additionally, H_O should be less biased than population-level estimates of genetic diversity in cases of unequal sample sizes or slight population structure across regions. We quantified ROH using ROHan (Renaud et al. 2019), which uses BAM files to estimate ROH directly from alignment data.

To quantify mutational load, we first used SnpEff (Cingolani et al. 2012) to annotate estimated functional effects of SNPs in coding regions based on the precomputed database for the reference genome. We extracted all variant sites that were annotated by SnpEff as having low, moderate, or high impacts for input into calculations of mutational load. Here, we estimated two measures of load, potential load and realized load, based on modifications to equations presented in Mathur et al. (2023):

$$PL_i = \frac{\sum V_{Hi} + \sum V_{Mi}}{\sum V_{Hi} + \sum V_{Mi} + \sum V_{Li}}$$

$$RL_i = \frac{\sum HDV_{Hi} + \sum HDV_{Mi}}{\sum V_{Hi} + \sum V_{Mi}}$$

Here, potential load (PL) is the sum of all high (H) and moderate (M) impact variants (V) found in individual i, divided by the sum of all high (H), moderate (M), and low (L) impact variants (V) (i.e., all nonsynonomous variants) found in individual i. PL is a proportion representative of functional mutations in the genome that are potentially deleterious (or weakly deleterious). Realized load (RL) is the sum of all high (H) and moderate (M) impact homozygous derived variants (HDV) found in individual i, divided by the sum of all high (H) and moderate (M) impact variants (V) found in individual i. RL is representative of the proportion of potentially deleterious variants that are found in the homozygous state.

To test whether diversity or load shifted through time in this avian assemblage, we used LMEM with the R package *lme4* (Bates et al. 2015). Here, we used a likelihood ratio test of null and alternative models to assess if time significantly impacted any of these measurements. In the null model, we treated species and geographic region as random effects. In the alternative model, we added time as a fixed effect. We had three different response variables: H_O, PL, and RL. We did not test for shifts in ROH because all individuals exhibited less than 1% of the genome in ROH.

Genetic Structure

We estimated genetic structure with two methods. First, we estimated genetic structure using PCA in PLINK v1.9 (Chang et al. 2015). Second, we estimated a phylogenetic network in SplitsTree v4.14.6 (Huson and Bryant 2006). As input for SplitsTree, we used Nei's D genetic distances (Nei 1972) calculated with the R package *StAMPP* (Pembleton et al. 2013). As part of this conversion, we also used the R packages *vcfR* (Knaus and Grünwald 2017) and *adegenet* (Jombart and Ahmed 2011) for manipulation of the genotype files. To identify any large shifts in genetic structure through time, we estimated genetic differentiation (F_{ST}) (Reich et al. 2009) for all population pairwise comparisons that included three samples using an F_{ST} estimator that works well with small sample sizes (Willing et al. 2012). For all populations with three samples, we estimated linkage disequilibrium (LD) decay using PopLDDecay (Zhang et al. 2019), because we would expect shifts in LD decay if there

were large decreases in population size through time or changes in connectivity between distinct genetic lineages. Our goal here was not to identify absolute values of LD decay (because of small sample sizes), but rather shifts through time as visualized by LD decay curves.

We also aimed to identify which factors impacted genetic structure for each species. Here, we used multiple regression of distance matrices (MRM) to assess how genetic distance (Nei's D) between samples could be explained by (i) time period sampled, (ii) geographic distance between samples, or (iii) number of biogeographic barriers separating samples (as in Fig. 1). Here, we estimated distances between points using the R package *fossil* (Vavrek 2011) and used the MRM function of the *ecodist* R package (Goslee and Urban 2007) with 100,000 permutations to assess statistical significance. Because multiple explanatory variables explained genetic distances (see Results), we used variance partitioning implemented in the R package *vegan* (Oksanen et al. 2007) to assess how each explanatory variable and their interactions shaped genetic distances.

Supplementary Material

Supplementary material is available at *Genome Biology and Evolution* online.

Acknowledgments

We thank scientific collectors that have contributed to scientific collections; without past scientific collection, this work would not have been possible. We thank the collections' staff and curators at the Field Museum of Natural History and the Museum of Comparative Zoology at Harvard University for allowing us to sample toepads. This work was supported by NSF Grants 1953688 to J.D.M. and 1953796 to B.D.M. We would like to thank Dr. Maanasa Raghavan for providing access to the ancient DNA facility at the University of Chicago and the Raghavan Genoscape Lab, especially Dr. Constanza de la Fuente, for facilitating our pre-PCR laboratory work. Mohamed Fokar at the TTU Center for Biotechnology & Genomics provided sequencing support. The TTU Center for Biotechnology & Genomics acquisition of the NovaSeq6000 was supported by NIH grant 1S10OD025115-01. The High-Performance Computing Center (HPCC) at Texas Tech University provided computational resources. The Pritzker DNA Lab at the Field Museum of Natural History provided resources for molecular work. We thank Asrat Ayalew for allowing us to use the Tacazze Sunbird photograph in this manuscript. Collecting modern samples was supported by assistance and permissions by the Ethiopian Wildlife Conservation Authority and the Oromia Forest and Wildlife Enterprise. We thank Rock Hewn Tours and personnel, especially Megersa Kelbessa for transportation, translation, and field assistance.

Author Contributions

Conceptualization: J.D.M., B.D.M., Y.M. Lab Work: A.E.S., J.D.M. Field Work: J.D.M, Y.B., L.C.C. Bioinformatics: J.D.M., A.E.S., G.J.B. Intellectual Contributions: All authors. Contributed Funding: J.D.M., B.D.M., S.B. First draft writing: J.D.M., A.E.S. Edits and contributions to final manuscript: All authors.

Data Availability

All raw sequence data is available from NCBI SRA. All newly generated sequence data is available on the NCBI BioProject database, under BioProject # PRJNA1300819. Previously generated sequence data from Manthey et al. (2022) and Behrends et al. (2024) can be found under BioProject #s PRJNA605410 and PRJNA948542. All code for analyses in this paper can be found on GitHub: https://github.com/jdmanthey/ETH_historical.

Literature Cited

Asefa A, Davies AB, McKechnie AE, Kinahan AA, van Rensburg BJ. Effects of anthropogenic disturbance on bird diversity in Ethiopian montane forests. Condor. 2017:119:416–430. https://doi.org/10.1650/CONDOR-16-81.1.

Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J Stat Softw. 2015:67:1–48. https://doi.org/10.18637/jss.v067.i01.

Behrends GJ, Meheretu Y, Manthey JD. The Great Rift Valley is a greater biogeographic barrier than the Blue Nile Valley for six Ethiopian Highland passerines in the eastern Afromontane biodiversity hotspot. Ornithology. 2024:141:ukae030. https://doi.org/10.1093/ornithology/ukae030.

Belay G, Mori A. Intraspecific phylogeographic mitochondrial DNA (D-loop) variation of gelada baboon, Theropithecus gelada, in Ethiopia. Biochem Syst Ecol. 2006:34:554–561. https://doi.org/10.1016/j.bse.2006.01.004.

Benham PM, Walsh J, Bowie RC. Spatial variation in population genomic responses to over a century of anthropogenic change within a tidal marsh songbird. Glob Chang Biol. 2024:30:e17126. https://doi.org/10.1111/gcb.17126.

Bertorelle G et al. Genetic load: genomic estimates and applications in non-model animals. Nat Rev Genet. 2022:23:492–503. https://doi.org/10.1038/s41576-022-00448-x.

Bieker VC, Martin MD. Implications and future prospects for evolutionary analyses of DNA in historical herbarium collections. Bot Lett. 2018:165:409–418. https://doi.org/10.1080/23818107.2018.1458651.

Blanchet G et al. Reduction of genetic diversity in 'Alalā (Hawaiian crow; *Corvus hawaiiensis*) between the late 1800s and the late 1900s. J Hered. 2024:115:32–44. https://doi.org/10.1093/jhered/esad063.

Blomqvist D, Pauliny A, Larsson M, Flodin L-Å. Trapped in the extinction vortex? Strong genetic effects in a declining vertebrate population. BMC Evol Biol. 2010:10:33–39. https://doi.org/10.1186/1471-2148-10-33.

Bourgeois Y, Warren BH, Augiron S. The burden of anthropogenic changes and mutation load in a critically endangered harrier from the reunion biodiversity hotspot, *Circus maillardi*. Mol Ecol. 2024:33:e17300. https://doi.org/10.1111/mec.17300.

Bryja J et al. Reticulate Pleistocene evolution of Ethiopian rodent genus along remarkable altitudinal gradient. Mol Phylogenet Evol. 2018:118:75–87. https://doi.org/10.1016/j.ympev.2017.09.020.
Buechley ER, et al. Importance of Ethiopian shade coffee farms for forest bird conservation. Biol Conserv. 2015:188:50–60. https://doi.org/10.1016/j.biocon.2015.01.011.
Bushnell B. BBMap: a fast, accurate, splice-aware aligner. Ernest Orlando Lawrence Berkeley National Laboratory; 2014.
Capinha C, Essl F, Seebens H, Moser D, Pereira HM. The dispersal of alien species redefines biogeography in the anthropocene. Science. 2015:348:1248–1251. https://doi.org/10.1126/science.aaa8913.
Chang CC et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience. 2015:4:7. https://doi.org/10.1186/s13742-015-0047-8.
Cingolani P et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin). 2012:6:80–92. https://doi.org/10.4161/fly.19695.
Curry CJ et al. Spatiotemporal genetic diversity of lions reveals the influence of habitat fragmentation across Africa. Mol Biol Evol. 2021:38:48–57. https://doi.org/10.1093/molbev/msaa174.
Danecek P et al. The variant call format and VCFtools. Bioinformatics. 2011:27:2156–2158. https://doi.org/10.1093/bioinformatics/btr330.
Darbyshire I, Lamb H, Umer M. Forest clearance and regrowth in northern Ethiopia during the last 3000 years. Holocene. 2003:13:537–546. https://doi.org/10.1191/0959683603hl644rp.
Derjusheva S, Kurganova A, Habermann F, Gaginskaya E. High chromosome conservation detected by comparative chromosome painting in chicken, pigeon and passerine birds. Chromosome Res. 2004:12:715–723. https://doi.org/10.1023/B:CHRO.0000045779.50641.00.
Dussex N et al. Population genomics of the critically endangered kākāpō. Cell Genom. 2021:1:100002. https://doi.org/10.1016/j.xgen.2021.100002.
Dussex N et al. Range-wide and temporal genomic analyses reveal the consequences of near-extinction in Swedish moose. Commun Biol. 2023a:6:1035. https://doi.org/10.1038/s42003-023-05385-x.
Dussex N, Morales HE, Grossen C, Dalén L, van Oosterhout C. Purging and accumulation of genetic load in conservation. Trends Ecol Evol. 2023b:38:961–969. https://doi.org/10.1016/j.tree.2023.05.008.
Ellegren H et al. The genomic landscape of species divergence in Ficedula flycatchers. Nature. 2012:491:756–760. https://doi.org/10.1038/nature11584.
Evans BJ, Bliss SM, Mendel SA, Tinsley RC. The Rift Valley is a major barrier to dispersal of African clawed frogs (Xenopus) in Ethiopia. Mol Ecol. 2011:20:4216–4230. https://doi.org/10.1111/j.1365-294X.2011.05262.x.
Feng S et al. The genomic footprints of the fall and recovery of the crested ibis. Curr Biol. 2019:29:340–349.e7. https://doi.org/10.1016/j.cub.2018.12.008.
Freilich X et al. Comparative phylogeography of Ethiopian anurans: impact of the Great Rift Valley and Pleistocene climate change. BMC Evol Biol. 2016:16:206. https://doi.org/10.1186/s12862-016-0774-1.
Friis I, Edwards S, Ensermu K, Sebsebe D. Diversity and endemism in the flora of Ethiopia and Eritrea—what do the published flora volumes tell us. Biol Skr. 2001:54:173–193. https://doi.org/10.1007/s12210-021-01027-8.
Fulton T, Shapiro B. Setting up an ancient DNA laboratory. Methods Mol Biol. 2019:1963:1–13. https://doi.org/10.1007/978-1-4939-9176-1_1.
Gargiulo R, Budde KB, Heuertz M. Mind the lag: understanding genetic extinction debt for conservation. Trends Ecol Evol. 2025:40:228–237. https://doi.org/10.1016/j.tree.2024.1010.1008.
Glenn TC et al. Adapterama I: universal stubs and primers for 384 unique dual-indexed or 147,456 combinatorially-indexed illumina libraries (iTru & iNext). PeerJ. 2019:7:e7755. https://doi.org/10.7717/peerj.7755.
Gordon I et al. Eastern afromontane biodiversity hotspot. In. *Critical ecosystem partnership fund* (CEPF): Birdlife International. 2012.
Goslee SC, Urban DL. The ecodist package for dissimilarity-based analysis of ecological data. J Stat Softw. 2007:22:1–19. https://doi.org/10.18637/jss.v022.i07.
Gottelli D, Marino J, Sillero-Zubiri C, Funk SM. The effect of the last glacial age on speciation and population genetic structure of the endangered Ethiopian wolf (*Canis simensis*). Mol Ecol.2004:13:2275–2286. https://doi.org/10.1111/j.1365-294X.2004.02226.x.
Gove AD, Hylander K, Nemomissa S, Shimelis A, Enkossa W. Structurally complex farms support high avian functional diversity in tropical montane Ethiopia. J Trop Ecol. 2013:29:87–97. https://doi.org/10.1017/S0266467413000023.
Griffin DK et al. Whole genome comparative studies between chicken and Turkey and their implications for avian genome evolution. BMC Genomics. 2008:9:168. https://doi.org/10.1186/1471-2164-9-168.
Hurni H, et al. Land degradation and sustainable land management in the highlands of Ethiopia. In: Hans H, Urs W, editors. Global change and sustainable development: a synthesis of regional experiences from research. Geographica Bernesia; 2010. p. 187–207.
Huson DH, Bryant D. Application of phylogenetic networks in evolutionary studies. Mol Biol Evol. 2006:23:254–267. https://doi.org/10.1093/molbev/msj030.
IUCN Red List of Threatened Species. Version 2024-1. 2024. https://www.iucnredlist.org.
Jackson HA et al. Genomic erosion in a demographically recovered bird species during conservation rescue. Conserv Biol. 2022:36:e13918. https://doi.org/10.1111/cobi.13918.
Jombart T, Ahmed I. *Adegenet 1.3-1*: new tools for the analysis of genome-wide SNP data. Bioinformatics. 2011:27:3070–3071. https://doi.org/10.1093/bioinformatics/btr521.
Jónsson H, Ginolhac A, Schubert M, Johnson PL, Orlando L. mapDamage2. 0: fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics. 2013:29:1682–1684. https://doi.org/10.1093/bioinformatics/btt193.
Kebede M, Ehrich D, Taberlet P, Nemomissa S, Brochmann C. Phylogeography and conservation genetics of a giant lobelia (Lobelia giberroa) in Ethiopian and Tropical East African mountains. Mol Ecol. 2007:16:1233–1243. https://doi.org/10.1111/j.1365-294X.2007.03232.x.
Kittelberger KD, Neate-Clegg MH, Buechley ER, Hakkı Şekercioğlu Ç. Community characteristics of forest understory birds along an elevational gradient in the Horn of Africa: a multi-year baseline. Ornithol Appl. 2021:123:duab009. https://doi.org/10.1093/ornithapp/duab009.
Knaus BJ, Grünwald NJ. vcfr: a package to manipulate and visualize variant call format data in R. Mol Ecol Resour. 2017:17:44–53. https://doi.org/10.1111/1755-0998.12549.
Komarova VA et al. Complex reticulate evolution of speckled brush-furred rats (*Lophuromys*) in the Ethiopian centre of endemism. Mol Ecol. 2021:30:2349–2365. https://doi.org/10.1111/mec.15891.
Koptekin D et al. Pre-processing of paleogenomes: mitigating reference bias and postmortem damage in ancient genome data. Genome Biol. 2025:26:6. https://doi.org/10.1186/s13059-024-03462-w.

Kostin DS et al. Taxonomic and genetic diversity of rodents from the Arsi Mountains (Ethiopia). Mammalia. 2019:83:237–247. https://doi.org/10.1515/mammalia-2017-0135.
Kostin DS et al. Rodents of choke mountain and surrounding areas (Ethiopia): the Blue Nile gorge as a strong biogeographic barrier. J Vertebr Biol. 2020:69:1–12. https://doi.org/10.25225/jvb.20016.
Largen M, Spawls S. The amphibians and reptiles of Ethiopia and Eritrea. Edition Chimaira; 2010.
Li D et al. Changes in taxonomic and phylogenetic diversity in the anthropocene. Proc R Soc Lond B Biol Sci. 2020:287:20200777. https://doi.org/10.1098/rspb.2020.0777.
Li H et al. The sequence alignment/map format and SAMtools. Bioinformatics. 2009:25:2078–2079. https://doi.org/10.1093/bioinformatics/btp352.
Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics. 2011:27:2987–2993. https://doi.org/10.1093/bioinformatics/btr509.
Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics. 2009:25:1754–1760. https://doi.org/10.1093/bioinformatics/btp324.
Logan WEM. An introduction to the forests of Central and southern Ethiopia. Imperial Forestry Institute, University of Oxford; 1946.
Manthey JD, Bourgeois Y, Meheretu Y, Boissinot S. Varied diversification patterns and distinct demographic trajectories in Ethiopian montane forest bird (Aves: Passeriformes) populations separated by the Great Rift Valley. Mol Ecol. 2022:31:2664–2678. https://doi.org/10.1111/mec.16417.
Manthey JD, Reyes-Velasco J, Freilich X, Boissinot S. Diversification in a biodiversity hotspot: genomic variation in the river frog Amietia nutti across the Ethiopian Highlands. Biol J Linn Soc Lond. 2017:122:801–813. https://doi.org/10.1093/biolinnean/blx106.
Marcacci G et al. Large-scale versus small-scale agriculture: disentangling the relative effects of the farming system and semi-natural habitats on birds' habitat preferences in the Ethiopian highlands. Agric Ecosyst Environ. 2020:289:106737. https://doi.org/10.1016/j.agee.2019.106737.
Mathur S, DeWoody JA. Genetic load has potential in large populations but is realized in small inbred populations. Evol Appl. 2021:14:1540–1557. https://doi.org/10.1111/eva.13216.
Mathur S, Tomeček JM, Tarango-Arámbula LA, Perez RM, DeWoody JA. An evolutionary perspective on genetic load in small, isolated populations as informed by whole genome resequencing and forward-time simulations. Evolution. 2023:77:690–704. https://doi.org/10.1093/evolut/qpac061.
McCann JC. The plow and the forest: narratives of deforestation in Ethiopia, 1840–1992. Environ Hist Durh N C. 1997:2:138–159. https://doi.org/10.2307/3985505.
Mizerovská D et al. Integrative taxonomic revision of the Ethiopian endemic rodent genus Stenocephalemys (Muridae: Murinae: Praomyini) with the description of two new species. J Vertebr Biol. 2020:69:20031. https://doi.org/10.25225/jvb.20031.
Munro RN et al. Soil landscapes, land cover change and erosion features of the Central Plateau region of Tigrai, Ethiopia: photo-monitoring with an interval of 30 years. Catena (Amst). 2008:75:55–64. https://doi.org/10.1016/j.catena.2008.04.009.
Myers N, Mittermeier RA, Mittermeier CG, Da Fonseca GA, Kent J. Biodiversity hotspots for conservation priorities. Nature. 2000:403:853–858. https://doi.org/10.1038/35002501.
Nei M. Genetic distance between populations. Am Nat. 1972:106:283–292. https://doi.org/10.1086/282771.
Nogué S et al. The human dimension of biodiversity changes on islands. Science. 2021:372:488–491. https://doi.org/10.1126/science.abd6706.
Nyssen J et al. Human impact on the environment in the Ethiopian and Eritrean highlands—a state of the art. Earth Sci Rev. 2004:64:273–320. https://doi.org/10.1016/S0012-8252(03)00078-3.
Nyssen J et al. Land degradation in the Ethiopian highlands. In: Billi P, editor. Landscapes and landforms of Ethiopia. Springer; 2015. p. 369–385.
Oksanen J, et al. The vegan package. Community Ecology Package. 2007.
Pembleton LW, Cogan NO, Forster JW. StAMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. Mol Ecol Resour. 2013:13:946–952. https://doi.org/10.1111/1755-0998.12129.
Petersen KR, Streett DA, Gerritsen AT, Hunter SS, Settles ML. Super deduper, fast PCR duplicate detection in fastq files, p 491–492. In: *Proceedings of the 6th ACM conference on bioinformatics, computational biology and health informatics*. Association for Computing Machinery, New York, NY. 2015.
Razgour O et al. Considering adaptive genetic variation in climate change vulnerability assessment reduces species range loss projections. Proc Natl Acad Sci U S A. 2019:116:10418–10423. https://doi.org/10.1073/pnas.1820663116.
Reich D, Thangaraj K, Patterson N, Price AL, Singh L. Reconstructing Indian population history. Nature. 2009:461:489–494. https://doi.org/10.1038/nature08365.
Reid JM, Nietlisbach P, Wolak ME, Keller LF, Arcese P. Individuals' expected genetic contributions to future generations, reproductive value, and short-term metrics of fitness in free-living song sparrows (*Melospiza melodia*). Evol Lett. 2019:3:271–285. https://doi.org/10.1002/evl3.118.
Renaud G, Hanghøj K, Korneliussen TS, Willerslev E, Orlando L. Joint estimates of heterozygosity and runs of homozygosity for modern and ancient samples. Genetics. 2019:212:587–614. https://doi.org/10.1534/genetics.119.302057.
Reyes-Velasco J, Manthey JD, Bourgeois Y, Freilich X, Boissinot S. Revisiting the phylogeography, demography and taxonomy of the frog genus Ptychadena in the Ethiopian highlands with the use of genome-wide SNP data. PLoS One. 2018:13:e0190440. https://doi.org/10.1371/journal.pone.0190440.
Reyes-Velasco J, Manthey JD, Freilich X, Boissinot S. Diversification of African tree frogs (genus *Leptopelis*) in the highlands of Ethiopia. Mol Ecol. 2018:27:2256–2270. https://doi.org/10.1111/mec.14573.
Ritler A. Land use, forests and the landscape of Ethiopia: 1699-1865; an enquiry into the historical geography of central northern Ethiopia. University of Bern; 1997.
Robbins SJ et al. jstjohn/SeqPrep: Tool for stripping adaptors and/or merging paired reads with overlap into single reads. 2011.
Rohland N, Reich D. Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. Genome Res. 2012:22:939–946. https://doi.org/10.1101/gr.128124.111.
Sánchez-Barreiro F et al. Historic sampling of a vanishing beast: population structure and diversity in the black rhinoceros. Mol Biol Evol. 2023:40:msad180. https://doi.org/10.1093/molbev/msad180.
Schmitt CJ, Cook JA, Zamudio KR, Edwards SV. Museum specimens of terrestrial vertebrates are sensitive indicators of environmental change in the anthropocene. Philos Trans R Soc Lond B Biol Sci. 2019:374:20170387. https://doi.org/10.1098/rstb.2017.0387.
Settlecowski AE, Marks BD, Manthey JD. Library preparation method and DNA source influence endogenous DNA recovery from 100-year-old avian museum specimens. Ecol Evol. 2023:13:e10407. https://doi.org/10.1002/ece3.10407.
Shaw RE et al. Global meta-analysis shows action is needed to halt genetic diversity loss. Nature. 2025:638:704–710. https://doi.org/10.1038/s41586-024-08458-x.

Shen W, Le S, Li Y, Hu F. SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. PLoS One. 2016:11:e0163962. https://doi.org/10.1371/journal.pone.0163962.
Shiferaw A, Yazezew D. Diversity, distribution and relative abundance of avifauna at Ansas Dam and surrounding farmland site Debre Berhan Town, Ethiopia. Avian Biol Res. 2021:14:8–17. https://doi.org/10.1177/1758155920963200.
Silvestrini M et al. Genetic diversity and structure of Ethiopian, Yemen and Brazilian Coffea arabica L. accessions using microsatellites markers. Genet Resour Crop Evol. 2007:54:1367–1379. https://doi.org/10.1007/s10722-006-9122-4.
Teixeira JC, Huber CD. The inflated significance of neutral genetic diversity in conservation genetics. Proc Natl Acad Sci U S A. 2021:118:e2015096118. https://doi.org/10.1073/pnas.2015096118.
Terrell KA et al. Continued decline in genetic diversity among wild cheetahs (*Acinonyx jubatus*) without further loss of semen quality. Biol Conserv. 2016:200:192–199. https://doi.org/10.1016/j.biocon.2016.05.034.
Tsai WL, Schedl ME, Maley JM, McCormack JE. More than skin and bones: comparing extraction methods and alternative sources of DNA from avian museum specimens. Mol Ecol Resour. 2020:20:1220–1227. https://doi.org/10.1111/1755-0998.13077.
Turvey ST, Crees JJ. Extinction in the anthropocene. Curr Biol. 2019:29:R982–R986. https://doi.org/10.1016/j.cub.2019.07.040.
van der Valk T, Díez-del-Molino D, Marques-Bonet T, Guschanski K, Dalén L. Historical genomes reveal the genomic consequences of recent population decline in eastern gorillas. Curr Biol. 2019:29:165–170.e6. https://doi.org/10.1016/j.cub.2018.11.055.
Vavrek MJ. Fossil: palaeoecological and palaeogeographical analysis tools. Palaeontol Electronica. 2011:14:1T.
Willing E-M, Dreyer C, Van Oosterhout C. Estimates of genetic differentiation measured by FST do not necessarily require large sample sizes when using many SNP markers. PLoS One. 2012:7:e42649. https://doi.org/10.1371/journal.pone.0042649.
Yalden D, Largen M. The endemic mammals of Ethiopia. Mamm Rev. 1992:22:115–150. https://doi.org/10.1111/j.1365-2907.1992.tb00128.x.
Zachos FE, Habel JC. Biodiversity hotspots: distribution and protection of conservation priority areas. Springer Science & Business Media; 2011.
Zhang C, Dong S-S, Xu J-Y, He W-M, Yang T-L. PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. Bioinformatics. 2019:35:1786–1788. https://doi.org/10.1093/bioinformatics/bty875.

Associate editor: Toni Gossmann