



GENETICS

Hybridization of Atlantic puffins in the Arctic coincides with 20th-century climate change

Oliver Kersten^{1*}, Bastiaan Star¹, Anders K. Krabberød², Lane M. Atmore¹, Ole K. Tørresen¹, Tycho Anker-Nilssen³, Sébastien Descamps⁴, Hallvard Strøm⁴, Ulf S. Johansson⁵, Paul R. Sweet⁶, Kjetill S. Jakobsen¹, Sanne Boessenkool^{1*}

The Arctic is experiencing the fastest rates of global warming, leading to shifts in the distribution of its biota and increasing the potential for hybridization. However, genomic evidence of recent hybridization events in the Arctic remains unexpectedly rare. Here, we use whole-genome sequencing of contemporary and 122-year-old historical specimens to investigate the origin of an Arctic hybrid population of Atlantic puffins (*Fratercula arctica*) on Bjørnøya, Norway. We show that the hybridization between the High Arctic, large-bodied subspecies *F. a. naumanni* and the temperate, smaller-sized subspecies *F. a. arctica* began as recently as six generations ago due to an unexpected southward range expansion of *F. a. naumanni*. Moreover, we find a significant temporal loss of genetic diversity across Arctic and temperate puffin populations. Our observations provide compelling genomic evidence of the impacts of recent distributional shifts and loss of diversity in Arctic communities during the 20th century.

INTRODUCTION

Global warming is driving flora and fauna toward the poles in search of more favorable environmental conditions (1–3). Such poleward shifts in species' distributions increase the potential for previously geographically isolated taxa or populations to hybridize. The evolutionary consequences of hybridization are complex, for instance, leading to decreased hybrid fitness, the extinction of parental and/or hybrid taxa, or to increased levels of genomic diversity promoting the resilience of a species (4–6). As the Arctic is experiencing the fastest rates of ecological and environmental change in the world (1, 3, 7), the potential for inter- and intraspecific hybridization is expected to be particularly high in this region (8, 9). Hybridization can be detected through temporal, whole-genome sequencing of historical specimens (10–13). Such temporal genomic data are crucial for determining if any observations of hybridization can be directly associated with recent environmental change. Nonetheless, temporal genomic approaches have not yet widely been used in the Arctic region (5, 8, 9, 12, 13) and, despite a few high-profile cases [e.g., polar bear (14) and snow hare (15)], unambiguous examples of recent hybridization in the Arctic remain unexpectedly rare.

Seabirds are a conspicuous component of Arctic biodiversity and are heavily affected by climate change leading to broad population declines and northward distributional shifts of temperate species (16–18). Here, we investigate the origin of a recently recognized Arctic hybrid population of Atlantic puffins [*Fratercula arctica*, Linnaeus, 1789; (19)], an iconic seabird that breeds throughout the North Atlantic and the Arctic (Fig. 1A) and that spends the

nonbreeding season at sea (20, 21). The Atlantic puffin is currently designated as "vulnerable" to extinction globally and "endangered" in Europe (22). Notably, the once world's largest colony at Røst (Norway) has lost more than 80% of its breeding pairs during the past 40 years (20, 23). Similarly, Icelandic and Faroese puffins have experienced low productivity and negative population growth since 2003 (24). These population declines have at least in part been ascribed to changes in prey availability as a result of climate change and overfishing (20, 22, 25, 26).

Partially corroborated by whole-genome analyses (19), Atlantic puffins are broadly classified into three subspecies based on size, i.e., the largest subspecies *F. a. naumanni* (hereafter "naumanni," found in the High Arctic, such as Spitsbergen and northwestern Greenland), the intermediate-sized subspecies *F. a. arctica* (hereafter "arctica," present in Norway, Iceland, and Faroe Islands), and the smallest subspecies *F. a. grabae* [hereafter "grabae," found in France, British Isles, and southern Norway (20); see Fig. 1A]. Recent genomic analyses also identified a distinct Western Atlantic cluster (19). This genetic structure has been attributed to barriers to gene flow on different spatial scales arising from a complex set of ecological drivers, including the interplay between overwintering grounds, natal dispersal, and ocean regimes (19). In addition, in one population in Greenland (Thule), three genetically distinct clusters representing both *naumanni* and *arctica* co-occur during the breeding season, so far without any evidence for hybridization (27). In contrast, a hybrid population of *naumanni* and *arctica* was found on the Arctic Island of Bjørnøya [Norway; (19)]. The evolutionary history of this hybrid population, including the timing and direction of gene flow, remains unknown.

Here, we analyze whole-genome data of modern and historic Atlantic puffin specimens to provide detailed insights into the origin of the hybrid population on Bjørnøya. First, to optimize the analyses of historic and contemporary data, we assembled and annotated a chromosome-level reference genome and performed admixture and demographic analyses using whole-genome data of 18 modern individuals across three breeding colonies that represent the parental

¹Centre for Ecological and Evolutionary Synthesis (CEES), Department of Biosciences, University of Oslo, Oslo, Norway. ²Section for Genetics and Evolutionary Biology (Evogene), Department of Biosciences, University of Oslo, Oslo, Norway. ³Norwegian Institute for Nature Research (NINA), Trondheim, Norway. ⁴Norwegian Polar Institute, Fram Centre, Langnes, Tromsø, Norway. ⁵Swedish Museum of Natural History, Stockholm, Sweden. ⁶American Museum of Natural History, New York, NY, USA.

*Corresponding author. Email: oliver.kersten@ibv.uio.no (O.K.); sanne.boessenkool@ibv.uio.no (S.B.)

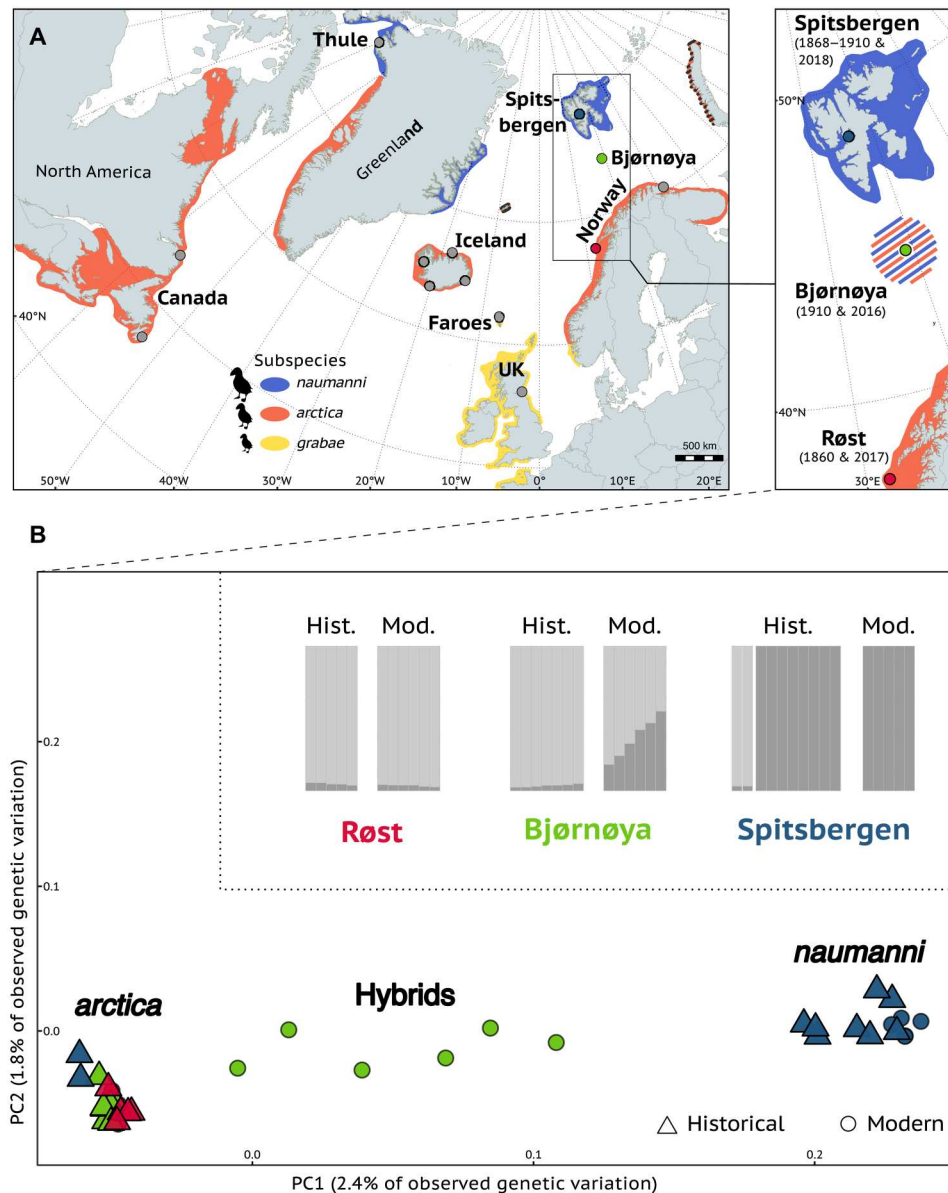


Fig. 1. Genomic population structure of modern and historical Atlantic puffins across three breeding colonies. (A) Distribution of breeding colonies analyzed in this study projected over the range of the three currently recognized subspecies and Atlantic puffin colonies previously analyzed for population genomics (19). The zoomed-in window presents the main study area with the hybrid population on Bjørnøya and the two colonies representing the parental subspecies *F. a. naumanni* and *F. a. arctica* from which historical and modern individuals were analyzed in this study. Sample dates are indicated in parentheses. (B) Genomic population structure of modern and historical Atlantic puffins of Bjørnøya, Røst, and Spitsbergen reconstructed by principal components and admixture analyses, performed with *PCAngsd* and *ngsAdmix*, respectively. The other 10 colonies used for the PCA analysis were removed for visualization purposes (the PCA with all colonies is presented in fig. S5). One related individual was removed from the modern dataset ($N = 17$, Spitsbergen = 5, Bjørnøya = 6, Røst = 6) and the historical dataset consisted of 22 individuals (Røst = 5, Bjørnøya = 7, Spitsbergen = 10). The same color coding for colonies and subspecies is used throughout the figure. Hist., historical; Mod., modern.

and hybrid populations. Second, we used genomic data from modern and historical specimens (dated from 1868 to 1910) to reconstruct the demographic history of the different subspecies of Atlantic puffins and obtain an accurate estimate of the timing of hybridization (10, 11, 13). Third, we investigated evidence for recent genetic impacts, which are of value for the management of species of conservation concern (13, 28–30). Demographic reconstructions show that the different subspecies started diverging approximately 100,000 years ago, most likely due to climatic

oscillations in the Pleistocene (31, 32). Our results show that the origin of the hybrid population on Bjørnøya lies within the past 100 years, coinciding with the rapid 20th-century climate change in the Arctic. Furthermore, we find evidence for losses of genetic variation over the past few decades, which is important for the conservation of the species due to the potential, associated decline in population viability (13, 30). Given that ecosystems worldwide are changing at unprecedented rates, our work highlights the power and importance of temporal genomics to detect and link population-

scale genetic changes with environmental and anthropogenic stressors of the past few decades.

RESULTS

A chromosome-level reference genome of the Atlantic puffin

Building on and enriching a previous, publicly available reference genome assembled with 10x Genomics data (table S1) (19), a high-quality, chromosome-level reference genome of the Atlantic puffin was assembled into 24 autosomes, 2 sex chromosomes, 1 mitochondrial genome, and 1 "unplaced scaffolds" sequence (229 unplaced scaffolds merged into a single sequence with a length of 6.1 Mb), using PacBio (CLR, Continuous Long Read), 10x Genomics, and Hi-C sequencing data (see Materials and Methods and Supplementary Text). The assembly is 1.215 Gb long, with an N50 contig/scaffold length of 16 Mb/82 Mb, holds 97.4% of protein-coding sequences from the avian set of the OrthoDB v10 database (BUSCO score), and displays substantial synteny with the razorbill (*Alca torda*; fig. S1 and table S2). The assembly has a quality value score of 37.63, an error rate of 0.017%, and a k-mer completeness rate of over 90% (fig. S2 and table S2). We observe a bimodal distribution of 10x Genomics k-mer read data (fig. S2), which may be driven by diploidy where some k-mers occur in both haplotypes and some occur in just haplotype one. As the 10x Genomics data were used for polishing, the impact on the completeness and correctness of the assembly is expected to be negligible. The annotation based on whole-genome alignments with other avian species includes 15,523 protein-coding genes, with an average coding DNA sequence length of 2316 base pairs (bp) and an average intron length of 4088 bp (table S2). The mitochondrial genome is 19,084 bp long, which is 2 kb longer than the currently published *F. a. arctica* mitogenome, now containing a large pseudo-duplicated region that was previously not assembled (fig. S3). The inclusion of this region emphasizes the impact of long reads in improving the assembly of mitogenomes that were previously obtained with short reads only, especially given that repeats and duplications are inherent to mitogenomes (33).

Genomic population structure and diversity

Atlantic puffins on Spitsbergen and Røst belong to distinct genomic clusters and subspecies, as previously identified from low-coverage whole-genome data (Fig. 1 and figs. S4 and S5) (19). Bjørnøya, by contrast, represents an admixed population where *arctica* and *naumanni* interbreed (Fig. 1 and figs. S4 and S5). Principal components analysis (PCA) and ADMIXTURE ($K = 2$) based on high-coverage resequencing of Spitsbergen, Røst, and Bjørnøya individuals support two main clusters with Bjørnøya forming an intermediate group (Fig. 1 and fig. S4). The presence of a hybridization zone on Bjørnøya and gene flow between *naumanni* and *arctica* (table S3) is confirmed by significantly negative f_3 -statistics ($Z < -3$) for the (Spitsbergen, Røst; Bjørnøya) topology, significantly positive D statistics caused by an excess of ABBA sites in the {(Røst, Bjørnøya), Spitsbergen}, Razorbill topology, and an f_4 -ratio of 0.297 in this topology (table S3). Spitsbergen is most distinct from the two other populations with significantly higher between-colony F_{ST} values (Dunn test with Holm correction, $P < 0.05$; fig. S6 and table S4), significantly lower nucleotide diversity and heterozygosity (Dunn test with Holm correction, $P < 0.05$; figs. S6 and S7 and table

S4) and positive Tajima's D values. Bjørnøya and Røst have the lowest between-colony F_{ST} values, similar levels of nucleotide diversity and heterozygosity, and negative Tajima's D values (figs. S6 and S7 and table S4).

Historical Røst and Spitsbergen museum specimens cluster with their respective modern counterparts (Fig. 1B), apart from two individuals from Spitsbergen that cluster with Røst (Fig. 1B and figs. S5, S8, and S9). These individuals were substantially smaller in size compared to other Atlantic puffins in Spitsbergen (fig. S10). At $K = 2$, which was the optimal model fit for the ngsAdmix analyses (fig. S11), both Røst and Spitsbergen show very similar ancestry proportions throughout time (Fig. 1B and fig. S8). In contrast, none of the historical Bjørnøya puffins cluster with their modern counterparts (Fig. 1B and fig. S5). All seven individuals cluster with the Røst samples in the PCA and admixture analysis and do not display ancestry proportions indicative of admixture (Fig. 1B and figs. S8 and S9).

Throughout the past 108 to 157 years, Spitsbergen and Røst puffins have experienced a significant loss of genetic diversity ($W_{\text{Spitsbergen}} = 135$, $P = 0.0008$; $W_{\text{Røst}} = 0$, $P = 1.5 \times 10^{-9}$; Fig. 2A and tables S5 and S7) and increase in inbreeding ($W_{\text{Spitsbergen}} = 41$, $P = 0.001$; $W_{\text{Røst}} = 30$, $P = 0.004$; Fig. 2B and fig. S12). Historically, Røst had the highest genetic diversity characterized by a large number of rare alleles indicating a large population size (figs. S6 and S13 and table S6). Contemporary Spitsbergen and Røst individuals showed no significant differences in levels of heterozygosity or runs of homozygosity (RoH) lengths (figs. S14 and S15), yet significantly elevated Tajima's D values ($W_{\text{Spitsbergen}} = 576$, $P = 1.5 \times 10^{-9}$; $W_{\text{Røst}} = 576$, $P = 1.5 \times 10^{-9}$; figs. S14 and S16 and table S5), compared to their historical conspecifics. Tajima's D was positive in modern Spitsbergen individuals while negative in all other colonies independent of the sampling period (table S5). Concurrent with the changes in Spitsbergen and Røst, Bjørnøya shows a significant decrease in nucleotide diversity ($W_{\text{Bjørnøya}} = 9$, $P = 4.7 \times 10^{-9}$; Fig. 2A and table S5) and an increase in inbreeding ($W_{\text{Bjørnøya}} = 41$, $P = 0.003$; Fig. 2B and fig. S12) and Tajima's D ($W_{\text{Bjørnøya}} = 576$, $P = 1.5 \times 10^{-9}$; figs. S14 and S16 and table S5) throughout 106 years. Moreover, the introgression of Spitsbergen ancestry into Bjørnøya has resulted in modern individuals on Bjørnøya having significantly longer RoHs than their historical conspecifics ($W_{\text{Bjørnøya}} = 42$, $P = 0.0017$; fig. S15).

Reconstructing demographic history

We reconstructed demographic trajectories of effective population size (N_e) for the Spitsbergen (*naumanni*) and Røst (*arctica*) colonies using a combination of methods that cover a range of temporal scales (Fig. 3). The common ancestral effective population size of *naumanni* and *arctica* was relatively stable between 100 thousand years (ka) and 500 ka ago at ~60,000 to 70,000 individuals (Fig. 3 and figs. S17 to S19) before both subspecies start to display independent trajectories between 40 ka and 100 ka ago (Fig. 3 and figs. S17 to S19). From around 15 ka to 40 ka ago, the demographic trajectories for *naumanni* and *arctica* diverged substantially, coinciding with the Last Glacial Maximum (LGM) and High Arctic, pre-LGM, Weichselian glacial cycles (Fig. 3 and figs. S17 to S19). The Røst population remained relatively constant in effective size with 50,000 to 60,000 individuals until 10 ka ago, whereas Spitsbergen decreased to about 8000 to 13,000 individuals (Fig. 3 and figs. S17 to S19). The timing of the split of *naumanni* from the ancestral

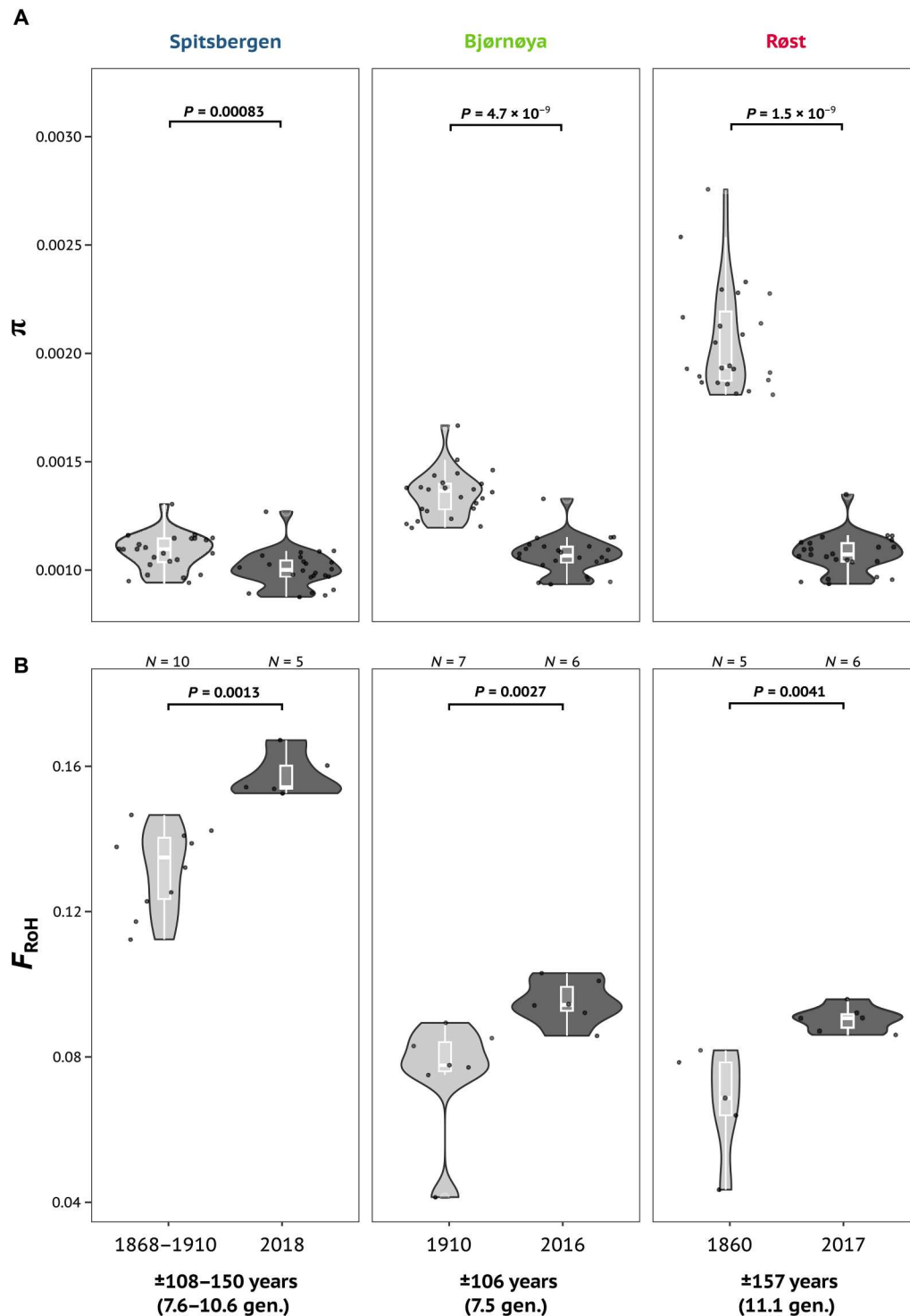


Fig. 2. Temporal changes in genomic diversity across three Atlantic puffin populations. (A) Nucleotide diversity and (B) inbreeding (F_{RoH}) were estimated for the historical and modern populations of Spitsbergen, Røst, and Bjørnøya representing the subspecies *F. a. naumanni*, *F. a. arctica*, and their hybrids, respectively. Nucleotide diversity for each population was calculated per chromosome ($N = 24$, black dots) with the program ANGSD using a site frequency spectrum. Individual inbreeding coefficients, F_{RoH} , were defined as the fraction of the individual genomes falling into runs of homozygosity (RoHs) of a minimum length of 150 kbp (19). RoHs were declared as all regions with at least two subsequent 100-kbp windows harboring heterozygosity below 0.482×10^{-3} . The x axes present sampling years and their differences in years and generations are given. The significance of differences is indicated by P values.

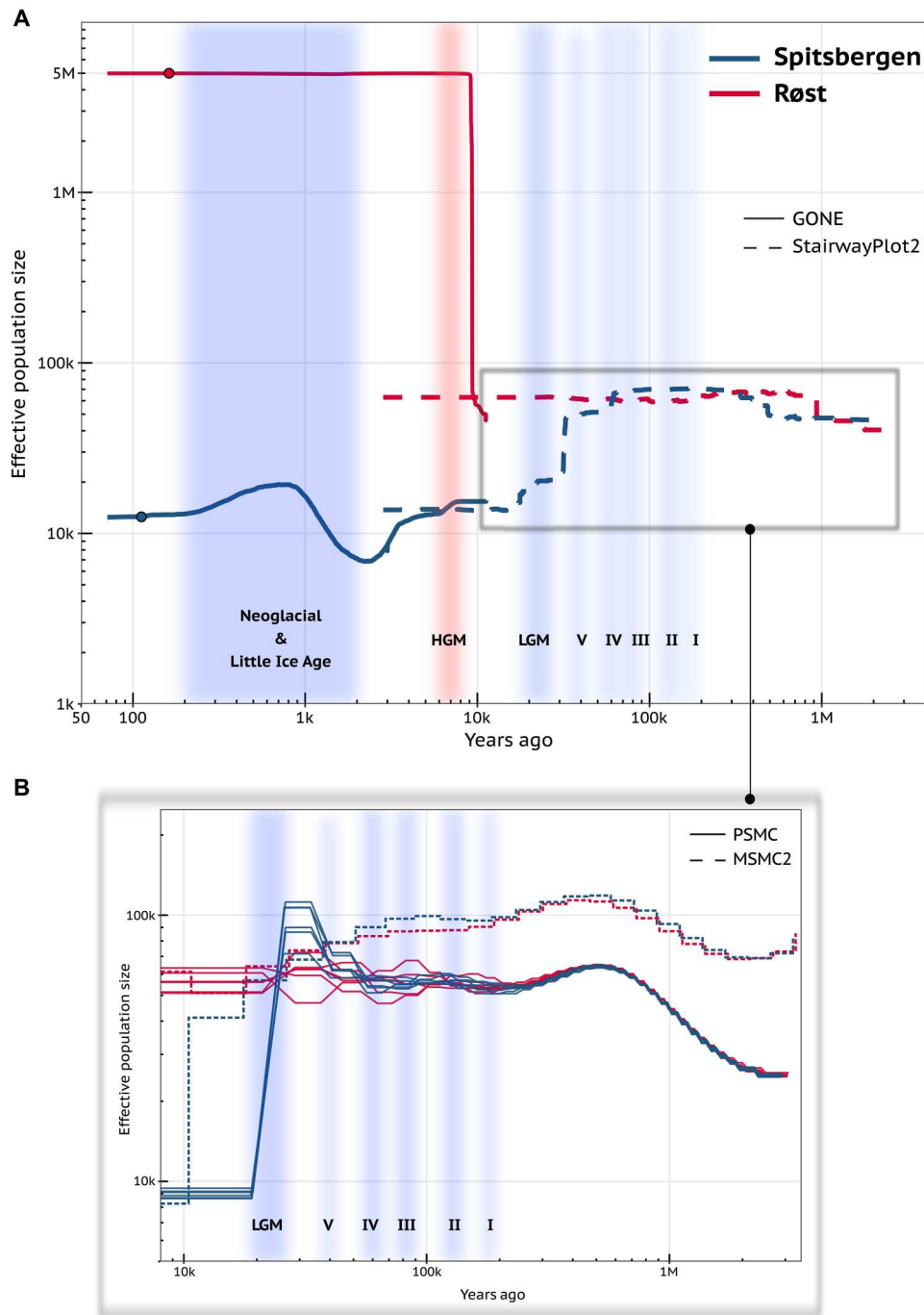


Fig. 3. Temporal reconstruction of effective population size of two Atlantic puffin populations across the past 2 million years. (A) The demographic histories of the Spitsbergen (*F. a. naumanni*) and Røst (*F. a. arctica*) colonies throughout the past 2 million years were estimated with the program GONE (straight) and StairwayPlot2 (dashed). (B) Changes in the effective population size in the period spanning 10,000 to 2 million years ago were also determined with the programs PSMC and MSMC2. Red and blue periods represent climatic warm and cold periods, respectively. HGM, Arctic Holocene Glacial Minimum; LGM, Last Glacial Maximum. Periods I to V = regional and local cold periods on Spitsbergen reconstructed by multiple proxies (31).

arctica effective population is corroborated by the relative cross-coalescence rate, which falls below 1.0 100 ka ago and below 0.5 (populations separated) 15 ka ago (fig. S20). While trends in effective population size before 10 ka ago were consistently reconstructed across all methods, multiple sequentially Markovian coalescent (MSMC2) displayed substantially higher N_e of the ancestral (*naumanniarctica*) population compared to the other methods (Fig. 3B and fig. S19), while pairwise sequentially Markovian coalescent (PSMC) showed a spike in effective population size of Spitsbergen 30 ka ago followed by a drop in the subsequent 10 ka, supporting the independence of the colony from the *arctica* population (Fig. 3B and fig. S18).

Within the past 10 ka, the Spitsbergen population remained relatively constant in effective size (10,000 to 11,000 individuals) for approximately 7 ka before dropping to 7000 individuals during the onset of the Neoglacial period followed by a threefold increase and subsequent decrease to ca. 12,000 individuals (Fig. 3A and fig. S21). Simultaneously, the Røst population increased in effective size from 50,000 to 5 million individuals at the onset of the Arctic Holocene Glacial Minimum (HGM; 6 ka to 8 ka ago) and has remained at this level ever since (Fig. 3A and fig. S21).

Admixture timing of *arctica* and *naumannii* at Bjørnøya

The onset of admixture between *arctica* and *naumannii* on Bjørnøya was timed using two different approaches that use local ancestry probabilities along the hybrid genomes (Fig. 4). *Naumannii* tracts that introgressed into *arctica* genomes on Bjørnøya, as determined by RFMix2, account for 7.2 to 48.9% of the hybrid genomes, which is similar to levels identified by the ADMIXTURE analysis (Fig. 4A and fig. S22). The length distribution of introgressed *naumannii* tracts is dominated by tracts with a length of 2 to 3 Mb and contains segments up to 60 Mb (fig. S23A). Combining all Bjørnøya individuals, and using the length and proportion of introgressed *naumannii* tracts per chromosome (fig. S23) and a chromosome-specific recombination rate (table S7), the most probable time of admixture onset on Bjørnøya was estimated to be 18.1 generations (272 years) ago ($CI_{95\%} = [17.3 \text{ to } 20.8]$; Fig. 4B). Similarly, individual-based admixture timings applying the same method range between 16.4 and 23.2 generations ago (fig. S23B). FastGlobetrotter uses local ancestry paintings of hybrid individuals generated by ChromainterV2 and suggests a one-pulse over a two-pulse admixture scenario due to the better model fit (fig. S24 and table S8). The bootstrapped estimates of admixture timings from fastGlobetrotter predicted an onset of admixture 8.9 generations (126 years) ago ($CI_{95\%} = [5.8 \text{ to } 12.8]$; Fig. 4B and table S8). The 95% confidence intervals of the two approaches (RFMix2 versus fastGlobetrotter) span a combined period between 82 and 295 years ago overlapping with the first 30 years following the collection of our historical specimens (112 years ago; Fig. 4B).

DISCUSSION

We provide genomic insights into the demographic history of the Atlantic puffin throughout the past 500,000 years, which is a period of marked phylogeographical separation between avian species (34). We show that the currently recognized subspecies *F. a. arctica* and *F. a. naumannii* diverged from a common ancestor at least 40,000 years ago. Despite this evolutionary divergence, a hybrid population on the Arctic island of Bjørnøya originated

only very recently. The absence of historical Bjørnøya hybrids combined with the estimated recent onset of admixture (Fig. 4) implies that hybridization between *arctica* and *naumannii* must have been initiated at least after 1910. Although hybridized individuals have been observed in several Arctic species [e.g., polar bear (14) and beluga whale (35)], we provide the first genomic evidence for the recent establishment of an entire hybrid population driven by a southward distributional shift. Our findings present a rare observation of a population-scale response to the rapid environmental changes that the Arctic ecosystem has started experiencing within the past century (1, 7, 36, 37).

Although the predominant distributional shift of flora and fauna in response to global warming is toward the poles (1–3), we here instead observe a southwards expansion of an Arctic subspecies. There may be two reasons for such an unexpected direction. First, *naumannii* colonies on Spitsbergen could have experienced favorable conditions in the past 100 years and the increase in population sizes could have led to the emigration of Atlantic puffins as a response to density-dependent factors (20). However, while there is no data on population trends of Atlantic puffin colonies on Spitsbergen within the past 100 years (38), the significant reduction in genetic diversity and increase in inbreeding over time (Fig. 2) do not support such an increase in population size. Alternatively, migration away from Spitsbergen could be attributed to detrimental environmental conditions at the natal colony, leading to Atlantic puffins establishing alternative colonies further south that may have been characterized by more favorable conditions (20, 39). While fishery-induced food shortages or pollution are harmful factors that cannot be ruled out, the estimated onset of hybridization notably coincides with the beginning of the 20th-century Arctic warming, where sea-surface temperature increased by up to 1.5°C (36, 37). This period is associated with large changes in plankton and fish communities and/or abundance (40) that are thought to have had detrimental effects on seabirds in the North Atlantic and the Arctic (18, 25, 41). Also, in part, driven by the multidecadal, cyclic pattern of ocean temperatures (37), the warming of the Arctic in the early 20th century likely played a substantial role in affecting the breeding success of puffins, as changes in sea temperatures can alter the food web and with that abundance and distribution of key prey for Atlantic puffins (25, 42). For example, the distribution of age-1 herring, an important prey for puffins during chick rearing, shifted eastward from the western Barents Sea to the Murman Coast in 1920 (26). Regardless of the ecological driver, our observations show that the responses and distributional shifts of Arctic fauna to global warming are not unidirectional, and attempts to understand or predict faunal responses to the rapidly changing climate need to incorporate such ecological complexity.

One common mechanism explaining phylogeographic patterns in Arctic seabirds, such as the European shag [*Gulosus aristotelis*; (43)] or common murre [*Uria aalge*; (44)], is allopatric divergence through isolation into different refugia [reviewed in (42)] driven by multiple regional and local glacial periods with adaptation to different habitat and climatic conditions. Nevertheless, evidence for such allopatric divergence has been difficult to obtain, given that many seabirds have colonized the Arctic within the past 100,000 years, leading to shallow evolutionary genomic divergence (19, 45, 46). High-resolution genomic data are therefore essential to detect such recent divergence, yet such data remain scarce for seabirds. This lack of data greatly limits our ability to infer basic population

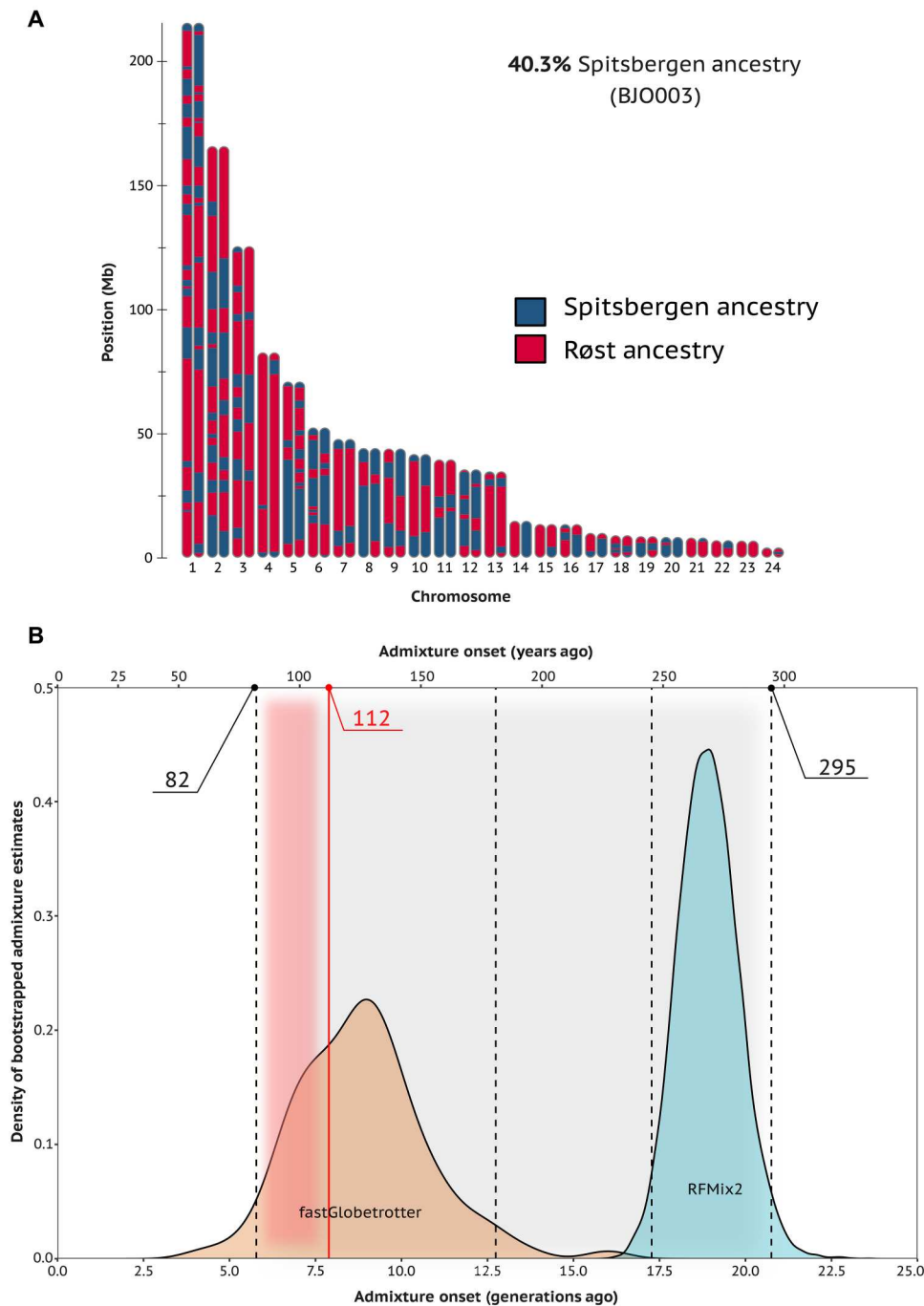


Fig. 4. Estimation of the onset of admixture between two Atlantic puffin subspecies using genomic data. Length and distribution of genomic regions assigned with Spitsbergen (*F. a. naumanni*) ancestry across all 24 autosomes of six hybrid genomes from Bjørnøya, as depicted in (A), were used to determine the 95% confidence intervals of admixture onset estimates, as shown in (B). (A) The ideogram shows 48 haplotypes of one example hybrid individual reconstructed by the program RFMix2 (the ideograms of all hybrids are presented in fig. S22). (B) The onset of admixture in years ago (top axis) and the number of generations ago (bottom axis) were estimated using fastGlobetrotter and RFMix2. FastGlobetrotter uses ancestry probabilities determined by Chromopainter2. Numbers in black at the top (82 and 295) represent the lower and upper limit of the 95% confidence intervals of the two approaches. The number in red (112) presents the sampling year of the historical specimens on Bjørnøya. The shift of the distribution of admixture timing estimates by RFMix2 further to the past is likely due to an underestimation of the length of the introgressed tracts, which is caused by switch errors during phasing chunking up the tracts.

structure, gene flow or demographic history, and the mechanisms that drive these patterns. Here, we observe that *naumanni* and *arctica* start diversifying from a common ancestor ~100 ka ago, representing the minimum age of divergence. This gradual separation of these subspecies throughout several Pleistocene glacial cycles (31) culminates in a complete differentiation during the LGM when *naumanni* may have inhabited a cryptic glacial refugium in the Arctic (47, 48). At the same time, *arctica* was likely confined to lower latitudes, such as Iceland or the Faroes (49). The high genetic diversity observed among Atlantic puffins in this area suggests that these populations remained larger than the more confined *naumanni* (19, 47). *Arctica* subsequently colonized the North Atlantic following the melting of the ice sheets during the Arctic HGM (32). The post-LGM radiation within the North Atlantic allowed this subspecies with colonies throughout Norway, Iceland, and the Faroe Islands—here represented by the colony on Røst—to reach its recent population size, while the Spitsbergen population remained multiple orders of magnitude smaller (Fig. 3). Such a process of divergence has been proposed for other Arctic seabirds (43, 45, 50, 51), which, for puffins, is now supported by a detailed genomic reconstruction of such recent divergence based on high-quality genomic data. Herewith, we provide evidence for the defining influence environmental changes have had in shaping the demographic histories of an Arctic seabird.

Analyses of historical and ancient DNA enable a direct assessment of changes in genetic diversity over time and allow the quantification of genetic impacts that are highly relevant for conservation efforts (11, 13, 28–30). While small populations are generally thought to be at higher risk of extinction (30, 52), it has been shown that the correlation between modern genome-wide diversity and population sizes of endangered species, for example, the tiger (53) or various apes (54), can be weak (30). Moreover, the relationship between genetic estimates of effective population size (N_e) and census size (N) is complex and influenced by factors, such as gene flow and genomic population structure, which renders relative comparisons of N_e within a species more useful. We find that modern genomic data identified a large effective population size for the Atlantic puffin colony of Røst (effective population size, 5 million), likely reflecting high gene flow between several colonies across Norway, Iceland, and the Faroe Islands, which, together with Røst, comprise a single genomic cluster (19). Similarly, the census size at Røst (226,000 breeding pairs in 2022) is orders of magnitude larger than the total Spitsbergen population [$<10,000$ breeding pairs (20)]. Yet, comparing historical to modern specimens, we nonetheless detected a significant loss of genetic diversity and an increase in inbreeding across the Atlantic puffin colonies of Spitsbergen and Røst within the past 100 years. Such genomic erosion can lead to a loss of adaptive potential and reduced population viability and is concerning from a conservation standpoint (13, 30, 55). Our observations therefore underline the importance of temporal genomic sampling together with a high-quality reference genome (56) to evaluate the status of species that are subject to rapid ecological changes and are of conservation concern.

To date, concrete genetic evidence that inter- or intraspecies hybridization is occurring on a population scale as a result of recent climatic changes in the Arctic is rare. By using modern and historical genomic data, we show that the onset of admixture between two Atlantic puffin subspecies has occurred within the past 100 years providing an unprecedented scale, resolution, and accuracy for

the timing of population-scale hybridization. Moreover, the emergence of this hybrid population coincides precisely with the anthropogenic warming of the Arctic. The unexpected southward range shift of a High Arctic subspecies thereby highlights how anthropogenic stressors can potentially force distributional changes resulting in hybridizing species or subspecies, essentially displacing genomic variation of a species whose populations have co-evolved over thousands of years (4, 57). The hybridization of Atlantic puffins in the East Atlantic Arctic may be a forecast of future scenarios throughout other parts of the Arctic illustrated by the sympatry of genetically distinct, but non-admixing, Atlantic puffin subspecies within a single High Arctic colony on the west coast of Greenland (27). Both the timing of admixture, as well as genomic erosion across multiple Atlantic puffin colonies and subspecies, present another line of compelling evidence that the past 100 years have been tremendously impactful on Arctic communities and highlight the importance and power of analyzing and linking contemporary and historical genomic data.

MATERIALS AND METHODS

Reference genome assembly and annotation

For the construction of an improved and highly contiguous and complete de novo Atlantic puffin genome assembly, we used a combination of PacBio, 10x Genomics Chromium, and Hi-C sequencing data. In total, six SMRT cells, sequenced on the PacBio Sequel II platform in CLR mode, generated 3.98 million reads with an average length of 13 kbp (N50: 24 kbp) and a coverage of 43.7X. A single 10x Genomics Chromium library was sequenced on three Illumina HiSeqX lanes generating approximately 2 billion reads (500.4X coverage). A cross-linked chromatin interaction (Hi-C) library was built with the Arima Genomic Hi-C kit and sequenced on an Illumina NovaSeq sequencer generating ca. 120 million unique Hi-C reads (157.1X coverage) of which ca. 50 million fragments were longer than 20 kb. The initial assembly was performed following the recommendations from the Vertebrate Genomes Project (58) with some adjustments. We used Flye (59), Falcon-Unzip (60), and Canu (61) for the initial assemblies to find the strategy with the optimal performance for the Atlantic puffin genome (table S9). Subsequent refinement and manual curation of the initial assemblies followed published recommendations and pipelines (see, e.g., <https://github.com/VGP/vgp-assembly/tree/master/pipeline>, accessed December 2022), as performed for previous genome assemblies (58). We chose the best (highest % of complete BUSCO score, longest N50, fewest N 's per 100 kb) assembly for further refinement (table S9). First, using mitoVGP (33) in combination with the PacBio and 10x Genomics reads, the mitogenome was identified among all scaffolds and subsequently separated from the assembly. Subsequently, the nuclear scaffolding was improved with HiGlass (62) and PretextView (<https://github.com/wtsi-hpag/PretextView>, v0.2.5, accessed December 2022). The resulting 26 largest super-scaffolds were visually distinct in their length distribution (fig. S2) and designated as "chromosomes." The remaining sequences were labeled as "unplaced scaffolds" and merged into a single sequence with 200 N 's in between. The final assembly consisted of 24 autosomes, 2 sex chromosomes, 1 mitogenome, and 1 merged-unplaced scaffold sequence. The mitochondrial genome was annotated with the MITOS2 (63) web server. The nuclear genome was annotated as

in (64). Further details on the reference genome assembly and annotation steps can be found in Supplementary Text.

Sequencing and single-nucleotide polymorphism identification of modern individuals

Samples from a total of 18 adult Atlantic puffin individuals collected across three breeding colonies (Spitsbergen, Bjørnøya, and Røst; Fig. 1) between 2012 and 2018 were sequenced for the present study to a genome-wide average depth of coverage of 21.6 (high-coverage dataset, >10 \times ; data S1). These 18 individuals were sequenced and analyzed at a lower coverage for a previous whole-genome study of Atlantic puffins (19). These samples were collected and made available by SEAPOPOP (www.seapop.no/en), SEATRACK (www.seapop.no/en/seatrack/), and ARCTOX (www.arctox.cnrs.fr/en/home) using the appropriate permits for collection and approved by the appropriate ethics committees. Spitsbergen and Røst were selected as representative colonies of the subspecies *naumanni* and *arctica*, respectively, and Bjørnøya was previously identified as a hybrid population between the two (19). We processed sequencing reads and mapped them to the Atlantic puffin assembly in PALEOMIX v1.2.14 (65).

Genotypes at autosomal single-nucleotide polymorphisms (SNPs) were jointly called with GATK v4.2.0 (66) and subsequently filtered (see Supplementary Text). Removal of one (of two) related individuals from the SNP dataset and pruning of SNPs to account for linkage disequilibrium resulted in a total of three different SNP datasets (table S10). Two additional SNP panels containing nonvariant sites were produced by applying the flag “--include-non-variant-sites” in GATK *GenotypeGVCFs* (table S10). Last, three SNP datasets with an outgroup were built by mapping sequencing data of the razorbill (*A. torda*, GCA_008658365.1) to the Atlantic puffin reference and either calling genotypes with GATK followed by combining SNPs with the Atlantic puffin genotypes or sampling the consensus base for each individual (Atlantic puffins and razorbill) at each site with ANGSD v0.935 (table S10) (67). A detailed description of laboratory methods and SNP calling can be found in Supplementary Text.

Sequencing and genotype likelihood estimation of historical individuals

Samples from a total of 22 adult Atlantic puffins from the three breeding areas [Spitsbergen ($N = 10$), Bjørnøya ($N = 7$), Røst/Lofoten archipelago ($N = 5$)] dating from 1860 to 1910 were sequenced for this study to a genome-wide average depth of coverage of 5.7 (data S1). These areas were chosen to represent the historical counterparts to the analyzed modern individuals of the subspecies *naumanni*, *arctica*, and of the hybrid population. Two of the Spitsbergen and one of the Røst individuals were substantially smaller in size and had therefore been assigned to be “mutants” or “migrants” by the original authors and collectors of the specimens (fig. S10) (68). All historical samples were made available by natural history museums using the appropriate permits for collection and import and export, and approved by the appropriate ethics committees. Before sequencing, DNA was extracted following protocols customized for historical avian specimens, and libraries were built with the single-stranded Santa Cruz Reaction library protocol (69). Sequencing reads from the 22 historical specimens together with all 77 modern Atlantic puffins sequenced to date (data S1) (19, 27) were processed in PALEOMIX.

After removing transitions and one related individual, an ANGSD pre-run determined filter settings (fig. S25), which were subsequently applied to calculate genotype likelihoods for SNPs covered in at least 95 of 98 individuals. The dataset was pruned to account for linkage disequilibrium (table S10). A detailed description of laboratory methods applied to historical specimens and genotype likelihood estimation can be found in Supplementary Text.

Genomic population structure

The population structure identified in (19) was validated with the 18 high-coverage individuals. A PCA was conducted with SmartPCA (70). Individual ancestry proportions were estimated with ADMIXTURE v1.3.0 (71).

Investigation of the genomic population structure was expanded with the genotype likelihood panel including 76 medium-coverage (5 \times to 10 \times) modern and 22 historical samples (table S10). A PCA was performed with PCAnsd v0.982 (72) and individual ancestry proportions were estimated in ngsAdmix v32 (73). The fit of the number of ancestral populations to the dataset was evaluated with evalAdmix v0.95 (74). A thorough description of the genomic population structure methodology can be found in Supplementary Text.

Genomic differentiation, diversity, and inbreeding

Genomic differentiation and diversity between and within the three Atlantic puffin breeding colonies at Spitsbergen, Bjørnøya, and Røst were assessed from the high-coverage dataset. Heterozygosity was quantified with VCFtools v0.1.16 (75); Tajima's D , π , and F_{ST} were calculated as the average of all 50-kb sliding window values calculated with the *tajima* function of the utility program VCF-kit (<https://vcf-kit.readthedocs.io/en/latest/>, accessed December 2022) and the script *popgenWindows.py* (https://github.com/simonhmartin/genomics_general, accessed December 2022). Levels of heterozygosity in modern Bjørnøya individuals were validated using the theoretical relationship between heterozygosity, F_{ST} , ancestry fraction, and heterozygosity of the parental populations (fig. S26) (76).

Genomic differentiation, diversity, and inbreeding across the entire Atlantic puffin distribution and between modern and historical specimens were analyzed from the medium-coverage dataset. Using an unpruned set of sites without transitions, covered in at least 95 individuals and passing several additional quality filters without removing rare alleles, we generated folded one-dimensional (1D) site frequency spectra (SFS) for each population and genomic cluster with ANGSD and winsfs v0.6.0 (77). These were used to compute Tajima's D and π per chromosome for each entity. We estimated per-individual, folded, 1D-SFS with ANGSD and winsfs with the same set of sites. Heterozygosity was calculated by dividing the number of polymorphic sites by the number of total sites present in the SFS. The proportion of runs of homozygosity (RoH) within each individual Atlantic puffin genome was computed as in (19). The cutoff for a “low heterozygosity region” was set to 0.482×10^{-3} (fig. S27). Individual inbreeding coefficients, F_{RoH} , were also calculated as in (19). More details on the assessment of genomic differentiation, diversity, and inbreeding can be found in Supplementary Text.

Demographic history

Before reconstructing the demographic history of the two Atlantic puffin subspecies, we generated estimates of the mutation rate of the

Atlantic puffin genome. The mutation rate was calculated by (i) counting the number of polymorphic sites versus total sites after pseudo-haploidization (78) and (ii) transforming pairwise genetic distances obtained via identity-by-state (IBS) sampling [see, e.g., (79)]. IBS sampling identified the consensus base at each of the 343,369,191 filtered sites (table S10) to convert each sample including the outgroup to a pseudo-haploid genome. For both approaches, the razorbill genome was used as an outgroup combined with an approximated divergence time between the razorbill and Atlantic puffin of 34 million years ago (80) and a generation time of the Atlantic puffin of 14.2 years (81). We used a final generational mutation rate of 1.7125×10^{-8} for all subsequent analyses, which was the mean of the mutation rates estimated by both approaches.

The demographic histories of the Spitsbergen and Røst populations were reconstructed using four programs with different temporal resolutions and different utilization of genomic signals. For the PSMC (82), a consensus fastq file was generated for each of the 12 Atlantic puffins, filtering out sites with a depth of less than 10 or more than double the mean coverage of that Atlantic puffin and with a root mean square mapping quality of reads covering the site under 25 (82, 83). After converting the fastq file to a PSMC input file and confirming the suitability of parameters and settings for the data (83), PSMC was run with “-N30 -t5 -r5 -p 4+30*2+4+6+10” for all individuals, including 100 subsequent bootstrap runs for each individual (82).

Before performing the StairwayPlot2 (84) analyses, a folded population-based SFS was generated by splitting the “NonVariant” dataset into populations and running the *easySFS.py* script (<https://github.com/isaacovercast/easySFS>, v0.0.1, accessed December 2022). StairwayPlot2 was run on each SFS with default settings, a mutation rate of 1.7125×10^{-8} and a generation time of 14.2 years (81).

MSMC2 (85) analyses were performed with phased SNP data. To phase the genomes, phase sets were identified in each individual Atlantic puffin genome by WhatsHap v0.18 (86) and genomes were subsequently statistically phased with Shapeit4 v4.1 (table S11) (87). After producing mappability files for each chromosome, individual per-chromosome calling mask files and individual MSMC2-specific vcf files were created. All samples were combined into specific MSMC2 per-chromosome input files. MSMC2 was run with “-s -p 1*2+22*1+1*2+1*6” for all unique combinations of haplotypes within each population as well as for unique combinations of two (out of five) different individuals from Spitsbergen versus two (out of six) different individuals from Røst, estimating within and across population coalescence rates (85).

As opposed to the other three tools, the software package GONE (88) is able to estimate fluctuations in N_e in the very recent past. Before running GONE, all chromosomes were randomly downsampled to a maximum of 500,000 SNPs per chromosome for computational purposes. GONE was run per population with default settings (88) and a mean genome-wide recombination rate of 1.63 centimorgan (cM)/Mb (table S7). The results of 100 bootstrap runs provided a median estimate of N_e and 95% confidence intervals.

Fluctuations in N_e reconstructed by either method were plotted using the mutation rate of 1.7125×10^{-8} and a generation time of 14.2 years (81). Each population trajectory was offset by the year the sampling took place (2016–2018; data S1). Last, major past warming and cooling periods on Svalbard and in the North Atlantic [see, e.g., (31, 32)] were added to the resulting plots. A more detailed

description of the analyses used for the reconstruction of demographic history can be found in Supplementary Text.

Admixture validation and timing estimate

Corroborating the detection of admixture on Bjørnøya by previous research (19), f_3 -statistics, D statistics, and the f_4 -ratio were calculated. f_3 -statistics were determined with *threepop* in Treemix v1.13 (89) investigating the (P1 and P2; Admixed Pop) topology. The D statistics and the f_4 -ratio were calculated for the most significant (the highest proportion ABBA over BABA sites) combination of Spitsbergen, Bjørnøya, and Røst in the topology {(P1, P2), P3}, Outgroup} with the program Dsuite (90).

Two different approaches requiring phased genomes (table S10) were applied to estimate the time or onset of admixture between the two subspecies on Bjørnøya. The first approach used RFmix v2 (91) and the theoretical relationship between the length of introgressed tracts, time of admixture, recombination rate, and admixture fractions (92). The second approach combined Chromopainter2 (93) and fastGlobetrotter (94), which exploits the fact that the probability of inheriting two DNA segments from the same ancestral source along the genome of an admixed individual decays exponentially with a rate proportional to the date of admixture and genetic distance between the segments (94).

RFmix v2 was run on all Bjørnøya individuals per chromosome accounting for the fact that reference haplotypes (Spitsbergen and Røst) may not be of “pure” ancestry. Following previous research [see, e.g., (95)], the length of introgressed Spitsbergen tracts (L) and global diploid Spitsbergen ancestry estimates (f) in Bjørnøya individuals generated by RFmix, together with the chromosome-specific recombination rate (r in morgans per base pair; table S7) were supplied to the equation $L = [(1 - f)^*r*(T_{ADMIX} - 1)]^{-1}$ (92) to estimate the individual-based and chromosome-specific age of admixture. The distribution of T_{ADMIX} across all 24 chromosomes and six individuals was bootstrapped 10,000 times to identify the average time of admixture and the 95% confidence intervals [see, e.g., (95)]. For the second approach, the genomes of the Bjørnøya individuals were “painted” (i.e., local ancestry probabilities for each SNP per haplotype were determined) with Chromopainter2. FastGlobetrotter was run with the painted genomes, default settings, and chromosome-specific recombination rates (table S7). One hundred bootstrap replicates of the proposed admixture scenario (here: one-pulse) were combined to determine 95% confidence intervals for the estimated admixture timings. More details on the validation and timing of admixture can be found in Supplementary Text.

Ethical statement

Feather, blood, and toepad samples of puffins included in this study were collected and handled under the following permits.

1. Gåsøyane, Røst, Hornøya, Bjørnøya (Norway): FOTS ID #15602 and #15603 from the Norwegian Food Safety Authority for SEATRACK and SEAPOP; Permit 2018/607 from Miljødirektoratet (Norwegian Environment Agency), dated 4 May 2018.

2. Gannet and Gull Island (Canada): Canadian Wildlife Service Migratory Bird Banding Permit 10559 G, approved Animal Use Protocol (AUP) by Eastern Wildlife Animal Care Committee (17GR01 and 18GR01), Newfoundland and Labrador Wilderness and Ecological Reserves Permit - Scientific Research (DOC/2017/

02003), Canadian Wildlife Service Scientific Permit ST2785 (to M.L.M.), Canadian Wildlife Service Banding Permit 10694, and Acadia University Animal Care Committee Permits ACC 02-15 and 06-15 (to M.L.M.).

3. Isle of May (Scotland): Scottish Natural Heritage licence 2014/MON/RP/156 and Ringing Permit A400 (to M.P.H.).

4. Vestmannaeyjar, Papey, Breiðafjörður, Grimsey (Iceland): Icelandic puffins were legally hunted during the hunting period of 1 July to 15 August.

5. Faroe: Feathers came from predated birds collected in the field after the predator was finished with them.

6. Thule: All research in Greenland was conducted after ethical approval and issuance of permits by the Government of Greenland, Department of Fisheries, Hunting and Agriculture (High Arctic Institute permit numbers: Sags nr. 2012–065141, Dok. nr. 888887, Sags nr. 2013–083369, Dok. nr. 1204884, Sags nr. 2014–099682, Dok nr. 1594176, and Sags nr. 2015–115204, Dok. nr. 1975643).

7. Historical specimens: All specimens were sampled by the respective museums following their ethical guidelines and loan policies for tissue samples.

Supplementary Materials

This PDF file includes:

Supplementary Text

Figs. S1 to S27

Tables S1 to S11

References

Other Supplementary Material for this manuscript includes the following:

Data S1

REFERENCES AND NOTES

- C. Parmesan, Ecological and evolutionary responses to recent climate change. *Annu. Rev. Ecol. Evol. Syst.* **37**, 637–669 (2006).
- E. S. Poloczanska, C. J. Brown, W. J. Sydeman, W. Kiessling, D. S. Schoeman, P. J. Moore, K. Brander, J. F. Bruno, L. B. Buckley, M. T. Burrows, C. M. Duarte, B. S. Halpern, J. Holding, C. V. Kappel, M. I. O'Connor, J. M. Pandolfi, C. Parmesan, F. Schwing, S. A. Thompson, A. J. Richardson, Global imprint of climate change on marine life. *Nat. Clim. Chang.* **3**, 919–925 (2013).
- G. T. Pecl, M. B. Araújo, J. D. Bell, J. Blanchard, T. C. Bonebrake, I.-C. Chen, T. D. Clark, R. K. Colwell, F. Danielsen, B. Evengård, L. Falconi, S. Ferrier, S. Frusher, R. A. Garcia, R. B. Griffis, A. J. Hobday, C. Janion-Scheepers, M. A. Jarzyna, S. Jennings, J. Lenoir, H. I. Linnetved, V. Y. Martin, P. C. McCormack, J. McDonald, N. J. Mitchell, T. Mustonen, J. M. Pandolfi, N. Pettorelli, E. Popova, S. A. Robinson, B. R. Scheffers, J. D. Shaw, C. J. B. Sorte, J. M. Strugnell, J. M. Sunday, M.-N. Tuanmu, A. Vergés, C. Villanueva, T. Wernberg, E. Wapstra, S. E. Williams, Biodiversity redistribution under climate change: Impacts on ecosystems and human well-being. *Science* **355**, eaai9214 (2017).
- M. Todesco, M. A. Pascual, G. L. Owens, K. L. Ostevik, B. T. Moyers, S. Hübner, S. M. Heredia, M. A. Hahn, C. Casesy, D. G. Bock, L. H. Rieseberg, Hybridization and extinction. *Evol. Appl.* **9**, 892–908 (2016).
- J. Ottenburghs, The genic view of hybridization in the Anthropocene. *Evol. Appl.* **14**, 2342–2360 (2021).
- C. J. Brauer, J. Sandoval-Castillo, K. Gates, M. P. Hammer, P. J. Unmack, L. Bernatchez, L. B. Beheregaray, Natural hybridization reduces vulnerability to climate change. *Nat. Clim. Chang.*, 282–289 (2023).
- M. C. Serreze, R. G. Barry, Processes and impacts of Arctic amplification: A research synthesis. *Glob. Planet Change* **77**, 85–96 (2011).
- B. P. Kelly, A. Whiteley, D. Tallmon, The Arctic melting pot. *Nature* **468**, 891 (2010).
- J. P. Colella, S. L. Talbot, C. Brochmann, E. B. Taylor, E. P. Hoberg, J. A. Cook, Conservation genomics in a changing arctic. *Trends Ecol. Evol.* **35**, 149–162 (2020).
- S. A. Taylor, E. L. Larson, R. G. Harrison, Hybrid zones: Windows on climate change. *Trends Ecol. Evol.* **30**, 398–406 (2015).
- K. Bi, T. Linderoth, D. Vanderpool, J. M. Good, R. Nielsen, C. Moritz, Unlocking the vault: Next-generation museum population genomics. *Mol. Ecol.* **22**, 6018–6032 (2013).
- D. Nogués-Bravo, F. Rodríguez-Sánchez, L. Orsini, E. de Boer, R. Jansson, H. Morlon, D. A. Fordham, S. T. Jackson, Cracking the code of biodiversity responses to past climate change. *Trends Ecol. Evol.* **33**, 765–776 (2018).
- E. L. Jensen, D. M. Leigh, Using temporal genomics to understand contemporary climate change responses in wildlife. *Ecol. Evol.* **12**, e9340 (2022).
- J. D. Pongracz, D. Paetkau, M. Branigan, E. Richardson, Recent hybridization between a polar bear and grizzly bears in the Canadian Arctic. *Arctic* **70**, 151–160 (2017).
- M. R. Jones, L. S. Mills, P. C. Alves, C. M. Callahan, J. M. Alves, D. J. R. Lafferty, F. M. Jiggins, J. D. Jensen, J. Melo-Ferreira, J. M. Good, Adaptive introgression underlies polymorphic seasonal camouflage in snowshoe hares. *Science* **360**, 1355–1358 (2018).
- M. Paleczny, E. Hammill, V. Karpouzi, D. Pauly, Population trend of the world's monitored seabirds, 1950–2010. *PLoS ONE* **10**, e0129342 (2015).
- Conservation of Arctic Flora and Fauna, "State of the Arctic Marine Biodiversity Report Update: Seabirds" (Conservation of Arctic Flora and Fauna International Secretariat, 2021).
- S. Descamps, H. Strøm, As the Arctic becomes boreal: Ongoing shifts in a high-Arctic seabird community. *Ecology* **102**, e03485 (2021).
- O. Kersten, B. Star, D. M. Leigh, T. Anker-Nilssen, H. Strøm, J. Danielsen, S. Descamps, K. E. Erikstad, M. G. Fitzsimmons, J. Fort, E. S. Hansen, M. P. Harris, M. Irested, O. Kleven, M. L. Mallory, K. S. Jakobsen, S. Boessenkool, Complex population structure of the Atlantic puffin revealed by whole genome analyses. *Commun. Biol.* **4**, 922 (2021).
- M. P. Harris, S. Wanless, *The Puffin* (T & AD Poyser, Bloomsbury Publishing, 2011).
- A. L. Fayet, R. Freeman, T. Anker-Nilssen, A. Diamond, K. E. Erikstad, D. Fifield, M. G. Fitzsimmons, E. S. Hansen, M. P. Harris, M. Jessopp, A.-L. Kouwenberg, S. Kress, S. Mowat, C. M. Perrins, A. Petersen, I. K. Petersen, T. K. Reiertsen, G. J. Robertson, P. Shannon, I. A. Sigurðsson, A. Shoji, S. Wanless, T. Guilford, Ocean-wide drivers of migration strategies and their influence on population breeding performance in a declining seabird. *Curr. Biol.* **27**, 3871–3878.e3 (2017).
- BirdLife International, *Fratercula arctica*. *The IUCN Red List of Threatened Species 2017* (BirdLife International, 2017).
- T. Anker-Nilssen, S. A. Hanssen, B. Moe, G. H. R. Systad, R. Barrett, J. O. Bustnes, S. Christensen-Dalsgaard, N. Dehnhard, S. Descamps, K. E. Erikstad, A. Follstad, M. Langset, K. Layton-Matthews, S. H. Lorentsen, E. Lorentzen, T. K. Reiertsen, H. Strøm, "Key-site monitoring in Norway 2021, including Svalbard and Jan Mayen," (Short Report no. 1–2022, SEAPOP, 2022).
- K. Lilliendahl, E. S. Hansen, V. Bogason, M. Sigursteinsson, M. L. Magnúsdóttir, P. M. Jónsson, H. H. Helgason, G. J. Óskarsson, P. F. Óskarsson, Ó. J. Sigurðsson, Recruitment failure of Atlantic puffins *Fratercula arctica* and sandeels *Ammodytes marinus* in Vestmannaeyjar Islands. *Náttúrufræðingurinn* **83**, 65–79 (2013).
- E. S. Hansen, H. Sandvik, K. E. Erikstad, N. G. Yoccoz, T. Anker-Nilssen, J. Bader, S. Descamps, K. Hodges, M. D. S. Mesquita, T. K. Reiertsen, Ø. Varpe, Centennial relationships between ocean temperature and Atlantic puffin production reveal shifting decennial trends. *Glob. Chang. Biol.* **27**, 3753–3764 (2021).
- K. F. Drinkwater, The regime shift of the 1920s and 1930s in the North Atlantic. *Prog. Oceanogr.* **68**, 134–151 (2006).
- D. M. Leigh, O. Kersten, B. Star, T. Anker-Nilssen, K. Burnham, J. Johnson, J. Provencher, S. Boessenkool, Sympatry of genetically distinct Atlantic Puffins (*Fratercula arctica*) in the High Arctic. *IBIS* **165**, 1022–1030 (2023).
- N. Nakahama, Museum specimens: An overlooked and valuable material for conservation genetics. *Ecol. Res.* **36**, 13–23 (2021).
- P. Wandeler, P. E. A. Hoeck, L. F. Keller, Back to the future: Museum specimens in population genetics. *Trends Ecol. Evol.* **22**, 634–642 (2007).
- D. Díez-del-Molino, F. Sánchez-Barreiro, I. Barnes, M. T. P. Gilbert, L. Dalén, Quantifying temporal genomic erosion in endangered species. *Trends Ecol. Evol.* **33**, 176–185 (2018).
- H. Alexanderson, M. Henriksen, H. T. Ryen, J. Y. Landvik, G. Peterson, 200 ka of glacial events in NW Svalbard: An emergence cycle facies model and regional correlations. *arktos* **4**, 1–25 (2018).
- W. R. Farnsworth, L. Allaart, Ó. Ingólfsson, H. Alexanderson, M. Forwick, R. Noormets, M. Retelle, A. Schomacker, Holocene glacial history of Svalbard: Status, perspectives and challenges. *Earth Sci. Rev.* **208**, 103249 (2020).
- G. Formenti, A. Rhie, J. Balacco, B. Haase, J. Mountcastle, O. Fedrigo, S. Brown, M. R. Capodiferro, F. O. Al-Ajli, R. Ambrosini, P. Houde, S. Koren, K. Oliver, M. Smith, J. Skelton, E. Betteridge, J. Doluacan, C. Corton, I. Bista, J. Torrance, A. Tracey, J. Wood, M. Uliano-Silva, K. Howe, S. McCarthy, S. Winkler, W. Kwak, J. Korlach, A. Functammasan, D. Fordham, V. Costa, S. Mayes, M. Chiara, D. S. Horner, E. Myers, R. Durbin, A. Achilli, E. L. Braun, A. M. Phillippy, E. D. Jarvis; Vertebrate Genomes Project Consortium, Complete vertebrate mitogenomes reveal widespread repeats and gene duplications. *Genome Biol.* **22**, 120 (2021).

34. J. C. Avise, *Phylogeography: The History and Formation of Species* (Harvard Univ. Press, 2000).
35. M. Skovrind, J. A. S. Castruita, J. Haile, E. C. Treadaway, S. Gopalakrishnan, M. V. Westbury, M. P. Heide-Jørgensen, P. Szpak, E. D. Lorenzen, Hybridization between two high Arctic cetaceans confirmed by genomic analysis. *Sci. Rep.* **9**, 7729 (2019).
36. T. Yamanouchi, Early 20th century warming in the Arctic: A review. *Polar Sci.* **5**, 53–71 (2011).
37. H. Tokinaga, S.-P. Xie, H. Mukougawa, Early 20th-century Arctic warming intensified by Pacific and Atlantic multidecadal variability. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 6227–6232 (2017).
38. H. Løvenskiold, "Avifauna Svalbardensis: With a discussion on the geographical distribution of the birds in Spitsbergen and adjacent islands" (Norsk Polarinstittut Skrifter nr. 129, Norsk Polarinstittut, 1964).
39. J. C. Coulson, A review of philopatry in seabirds and comparisons with other waterbird species. *Waterbirds* **39**, 229–240 (2016).
40. G. Beaugrand, P. C. Reid, Relationships between North Atlantic salmon, plankton, and hydroclimatic change in the Northeast Atlantic. *ICES J. Mar. Sci.* **69**, 1549–1562 (2012).
41. S. Descamps, T. Anker-Nilssen, R. T. Barrett, D. B. Irons, F. Merkel, G. J. Robertson, N. G. Yoccoz, M. L. Mallory, W. A. Montevecchi, D. Boertmann, Y. Artukhin, S. Christensen-Dalsgaard, K.-E. Erikstad, H. G. Gilchrist, A. L. Labansen, S.-H. Lorentsen, A. Mosbech, B. Olsen, A. Petersen, J.-F. Rail, H. M. Renner, H. Strøm, G. H. Systad, S. I. Wilhelm, L. Zelenskaya, Circumpolar dynamics of a marine top-predator track ocean warming rates. *Glob. Chang. Biol.* **23**, 3770–3780 (2017).
42. M. P. Harris, T. Anker-Nilssen, R. H. McCleery, K. E. Erikstad, D. N. Shaw, V. Grosbois, Effect of wintering area and climate on the survival of adult Atlantic puffins *Fratercula arctica* in the eastern Atlantic. *Mar. Ecol. Prog. Ser.* **297**, 283–296 (2005).
43. E. Thanou, S. Sponza, E. J. Nelson, A. Perry, S. Wanless, F. Daunt, S. Cavers, Genetic structure in the European endemic seabird, *Phalacrocorax aristotelis*, shaped by a complex interaction of historical and contemporary, physical and nonphysical drivers. *Mol. Ecol.* **26**, 2796–2811 (2017).
44. J. A. Morris-Pocock, S. A. Taylor, T. P. Birt, M. Dams, J. F. Piatt, K. I. Warheit, V. L. Friesen, Population genetic structure in Atlantic and Pacific Ocean common murre (*Uria aalge*): Natural replicate tests of post-Pleistocene evolution. *Mol. Ecol.* **17**, 4859–4873 (2008).
45. K. Wojczulanis-Jakubas, A. Kilikowska, A. M. A. Harding, D. Jakubas, N. J. Karnovsky, H. Steen, H. Strøm, J. Welcker, M. Gavrilov, J. T. Lifjeld, A. Johnsen, Weak population genetic differentiation in the most numerous Arctic seabird, the little auk. *Polar Biol.* **37**, 621–630 (2014).
46. J. E. Thomas, G. R. Carvalho, J. Haile, N. J. Rawlence, M. D. Martin, S. Y. Ho, A. P. Sigfússon, V. A. Jósefsson, M. Frederiksen, J. F. Linnebjerg, J. A. S. Castruita, J. Niemann, M.-H. S. Sinding, M. Sandoval-Velasco, A. E. Soares, R. Lacy, C. Barilaro, J. Best, D. Brandis, C. Cavallo, M. Elorza, K. L. Garrett, M. Groot, F. Johansson, J. T. Lifjeld, G. Nilson, D. Serjeanston, P. Sweet, E. Fuller, A. K. Hufthammer, M. Meldgaard, J. Fjeldså, B. Shapiro, M. Hofreiter, J. R. Stewart, M. T. P. Gilbert, M. Knapp, Demographic reconstruction from ancient DNA supports rapid extinction of the great auk. *eLife* **8**, e47509 (2019).
47. J. R. Stewart, A. M. Lister, I. Barnes, L. Dalén, Refugia revisited: Individualistic responses of species in space and time. *Proc. Biol. Sci.* **277**, 661–671 (2010).
48. J. Y. Landvik, E. J. Brook, L. Gualtieri, G. Raisbeck, O. Salvigsen, F. Yiou, Northwest Svalbard during the last glaciation: Ice-free areas existed. *Geology* **31**, 905–908 (2003).
49. I. Newton, *Speciation and Biogeography of Birds* (Academic Press, 2003).
50. A. Tigano, M. Dams, T. P. Birt, J. A. Morris-Pocock, Y. B. Artukhin, V. L. Friesen, The Arctic: Glacial refugium or area of secondary contact? Inference from the population genetic structure of the thick-billed murre (*Uria lomvia*), with implications for management. *J. Hered.* **106**, 238–246 (2015).
51. T. Moum, E. Arnason, Genetic diversity and population history of two related seabird species based on mitochondrial DNA control region sequences. *Mol. Ecol.* **10**, 2463–2478 (2001).
52. G. Caughley, Directions in conservation biology. *J. Anim. Ecol.* **63**, 215–244 (1994).
53. Y. S. Cho, L. Hu, H. Hou, H. Lee, J. Xu, S. Kwon, S. Oh, H.-M. Kim, S. Jho, S. Kim, Y.-A. Shin, B. C. Kim, H. Kim, C.-U. Kim, S.-J. Luo, W. E. Johnson, K.-P. Koepfli, A. Schmidt-Küntzel, J. A. Turner, L. Marker, C. Harper, S. M. Miller, W. Jacobs, L. D. Bertola, T. H. Kim, S. Lee, Q. Zhou, H.-J. Jung, X. Xu, P. Gadhvi, P. Xu, Y. Xiong, Y. Luo, S. Pan, C. Gou, X. Chu, J. Zhang, S. Liu, J. He, Y. Chen, L. Yang, Y. Yang, J. He, S. Liu, J. Wang, C. H. Kim, H. Kwak, J.-S. Kim, S. Hwang, J. Ko, C.-B. Kim, S. Kim, D. Bayarikhagva, W. K. Paek, S.-J. Kim, S. J. O'Brien, J. Wang, J. Bhak, The tiger genome and comparative analysis with lion and snow leopard genomes. *Nat. Commun.* **4**, 2433 (2013).
54. J. Prado-Martinez, P. H. Sudmant, J. M. Kidd, H. Li, J. L. Kelley, B. Lorente-Galdos, K. R. Veeramah, A. E. Woerner, T. D. O'Connor, G. Santpere, A. Cagan, C. Theunert, F. Casals, H. Laayouni, K. Munch, A. Hobolth, A. E. Halager, M. Malig, J. Hernandez-Rodriguez, I. Hernando-Herraez, K. Prüfer, M. Pybus, L. Johnstone, M. Lachmann, C. Alkan, D. Twigg, N. Petit, C. Baker, F. Hormozdiari, M. Fernandez-Callejo, M. Dabad, M. L. Wilson, L. Stevison, C. Camrubi, T. Carvalho, A. Ruiz-Herrera, L. Vives, M. Mele, T. Abello, I. Kondova, R. E. Bontrop, A. Pusey, J. Lankester, J. A. Kiyang, R. A. Bergl, E. D. Lonsdorf, S. Myers, M. Ventura, P. Gagneux, D. Comas, H. Siegmund, J. Blanc, L. Agueda-Calpena, M. Gut, L. Fulton, S. A. Tishkoff, J. C. Mullikin, R. K. Wilson, I. G. Gut, M. K. Gonder, O. A. Ryder, B. H. Hahn, A. Navarro, J. M. Akey, J. Bertranpetit, D. Reich, T. Mailund, M. H. Schierup, C. Hvilson, A. M. Andrés, J. D. Wall, C. D. Bustamante, M. F. Hammer, E. E. Eichler, T. Marques-Bonet, Great ape genetic diversity and population history. *Nature* **499**, 471–475 (2013).
55. M. H. Kohn, W. J. Murphy, E. A. Ostrander, R. K. Wayne, Genomics and conservation genetics. *Trends Ecol. Evol.* **21**, 629–637 (2006).
56. K. Theissinger, C. Fernandes, G. Formenti, I. Bista, P. R. Berg, C. Bleidorn, A. Bombarely, A. Crottini, G. R. Gallo, J. A. Godoy, S. Jentoft, J. Malukiewicz, A. Mouton, R. A. Oomen, S. Paez, P. J. Palsbøll, C. Pampoulie, M. J. Ruiz-López, S. Secomandi, H. Svardal, C. Theofanopoulou, J. de Vries, A.-M. Waldvogel, G. Zhang, E. D. Jarvis, M. Balcint, C. Ciofi, R. M. Waterhouse, C. J. Mazzoni, J. Höglund; European Reference Genome Atlas Consortium, How genomics can help biodiversity conservation. *Trends Genet.* **39**, 545–559 (2023).
57. A. J. Chuncu, Hybridization in a warmer world. *Ecol. Evol.* **4**, 2019–2031 (2014).
58. A. Rhie, S. A. McCarthy, O. Fedrigo, J. Damas, G. Formenti, S. Koren, M. Uliano-Silva, W. Chow, A. Functamman, J. Kim, C. Lee, B. J. Ko, M. Chaisson, G. L. Gedman, L. J. Cantin, F. Thibaud-Nissen, L. Haggerty, I. Bista, M. Smith, B. Haase, J. Mountcastle, S. Winkler, S. Paez, J. Howard, S. C. Vernes, T. M. Lama, F. Grutzner, W. C. Warren, C. N. Balakrishnan, D. Burt, J. M. George, M. T. Biegler, D. Iorns, A. Digby, D. Eason, B. Robertson, T. Edwards, M. Wilkinson, G. Turner, A. Meyer, A. F. Kautt, P. Franchini, H. W. Detrich 3rd, H. Svardal, M. Wagner, G. J. P. Naylor, M. Pippel, M. Malinsky, M. Mooney, M. Simbirsky, B. T. Hannigan, T. Pesout, M. Houck, A. Misuraca, S. B. Kingan, R. Hall, Z. Kronenberg, I. Sović, C. Dunn, Z. Ning, A. Hastie, J. Lee, S. Selvaraj, R. E. Green, N. H. Putnam, I. Gut, J. Ghurye, E. Garrison, Y. Sims, J. Collins, S. Pelan, J. Torrance, A. Tracey, J. Wood, R. E. Dagnew, D. Guan, S. E. London, D. F. Clayton, C. V. Mello, S. R. Friedrich, P. V. Lovell, E. Osipova, F. O. Al-Ajli, S. Secomandi, H. Kim, C. Theofanopoulou, M. Hiller, Y. Zhou, R. S. Harris, K. D. Makova, P. Medvedev, J. Hoffman, P. Masterson, K. Clark, F. Martin, K. Howe, P. Flicek, B. P. Walenz, W. Kwak, H. Clawson, M. Diekhans, L. Nassar, B. Paten, R. H. S. Kraus, A. J. Crawford, M. T. P. Gilbert, G. Zhang, B. Venkatesh, R. W. Murphy, K.-P. Koepfli, B. Shapiro, W. E. Johnson, F. Di Palma, T. Marques-Bonet, E. C. Teeling, T. Warnow, J. M. Graves, O. A. Ryder, D. Haussler, S. J. O'Brien, K. Lorch, H. A. Lewin, K. Howe, E. W. Myers, R. Durbin, A. M. Phillippy, E. D. Jarvis, Towards complete and error-free genome assemblies of all vertebrate species. *Nature* **592**, 737–746 (2021).
59. M. Kolmogorov, J. Yuan, Y. Lin, P. A. Pevzner, Assembly of long, error-prone reads using repeat graphs. *Nat. Biotechnol.* **37**, 540–546 (2019).
60. C.-S. Chin, P. Peluso, F. J. Sedlazeck, M. Nattestad, G. T. Concepcion, A. Clum, C. Dunn, R. O'Malley, R. Figueroa-Balderas, A. Morales-Cruz, G. R. Cramer, M. Delledonne, C. Luo, J. R. Ecker, D. Cantu, D. R. Rank, M. C. Schatz, Phased diploid genome assembly with single-molecule real-time sequencing. *Nat. Methods* **13**, 1050–1054 (2016).
61. S. Koren, B. P. Walenz, K. Berlin, J. R. Miller, N. H. Bergman, A. M. Phillippy, Canu: Scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res.* **27**, 722–736 (2017).
62. P. Kerpedjiev, N. Abdennur, F. Leckschas, C. McCallum, K. Dinkla, H. Strobelt, J. M. Luber, S. B. Ouellette, A. Azhir, N. Kumar, J. Hwang, S. Lee, B. H. Alver, H. Pfister, L. A. Mirny, P. J. Park, N. Gehlenborg, HiGlass: Web-based visual exploration and analysis of genome interaction maps. *Genome Biol.* **19**, 125 (2018).
63. A. Donath, F. Jühling, M. Al-Arab, S. H. Bernhart, F. Reinhardt, P. F. Stadler, M. Middendorf, M. Bernt, Improved annotation of protein-coding genes boundaries in metazoan mitochondrial genomes. *Nucleic Acids Res.* **47**, 10543–10552 (2019).
64. C. L. C. Sætre, F. Eroukhanoff, K. Rönkä, E. Klun, R. Thorogood, J. Torrance, A. Tracey, W. Chow, S. Pelan, K. S. Jakobsen, O. K. Tøresen, A chromosome-level genome assembly of the reed warbler (*Acrocephalus scirpaceus*). *Genome Biol. Evol.* **13**, evab212 (2021).
65. M. Schubert, L. Ermini, C. Der Sarkissian, H. Jónsson, A. Ginolhac, R. Schaefer, M. D. Martin, R. Fernández, M. Kircher, M. McCue, E. Willerslev, L. Orlando, Characterization of ancient and modern genomes by SNP detection and phylogenomic and phylogenomic analysis using PALEOMIX. *Nat. Protoc.* **9**, 1056–1082 (2014).
66. A. McKenna, M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytzky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, M. A. DePristo, The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
67. T. S. Korneliusson, A. Albrechtsen, R. Nielsen, ANGSD: Analysis of next generation sequencing data. *BMC Bioinformatics.* **15**, 356 (2014).

68. F. Salomonsen, *The Atlantic Alcidae: The Seasonal and Geographical Variation of the Auks Inhabiting the Atlantic Ocean and the Adjacent Waters* (Elanders boktryckeri aktiebolag, 1944), vol. 6 of Göteborgs kungl. vetenskaps- och vitterhets-samhälles Handlingar.
69. J. D. Kapp, R. E. Green, B. Shapiro, A fast and efficient single-stranded genomic library preparation method optimized for ancient DNA. *J. Hered.* **112**, 241–249 (2021).
70. N. Patterson, A. L. Price, D. Reich, Population structure and eigenanalysis. *PLOS Genet.* **2**, e190 (2006).
71. D. H. Alexander, J. Novembre, K. Lange, Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).
72. J. Meisner, A. Albrechtsen, Inferring population structure and admixture proportions in low-depth NGS data. *Genetics* **210**, 719–731 (2018).
73. L. Skotte, T. S. Korneliussen, A. Albrechtsen, Estimating individual admixture proportions from next generation sequencing data. *Genetics* **195**, 693–702 (2013).
74. G. Garcia-Erill, A. Albrechtsen, Evaluation of model fit of inferred admixture proportions. *Mol. Ecol. Resour.* **20**, 936–949 (2020).
75. P. Danecek, A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, R. Durbin; 1000 Genomes Project Analysis Group, The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
76. S. M. Boca, L. Huang, N. A. Rosenberg, On the heterozygosity of an admixed population. *J. Math. Biol.* **81**, 1217–1250 (2020).
77. M. S. Rasmussen, G. Garcia-Erill, T. S. Korneliussen, C. Wiuf, A. Albrechtsen, Estimation of site frequency spectra from low-coverage sequencing data using stochastic EM reduces overfitting, runtime, and memory usage. *Genetics* **222**, iyac148 (2022).
78. D. L. Duc, A. Velluva, M. Cassatt-Johnstone, R.-A. Olsen, S. Baleka, C.-C. Lin, J. R. Lemke, J. R. Southon, A. Burdin, M.-S. Wang, S. Grunewald, W. Rosendahl, U. Joger, S. Rutschmann, T. B. Hildebrandt, G. Fritsch, J. A. Estes, J. Kelso, L. Dalén, M. Hofreiter, B. Shapiro, T. Schöneberg, Genomic basis for skin phenotype and cold adaptation in the extinct Steller's sea cow. *Sci. Adv.* **8**, eabl6496 (2022).
79. M. V. Westbury, S. Hartmann, A. Barlow, I. Wiesel, V. Leo, R. Welch, D. M. Parker, F. Sicks, A. Ludwig, L. Dalén, M. Hofreiter, Extended and continuous decline in effective population size results in low genomic diversity in the World's Rarest Hyena species, the Brown Hyena. *Mol. Biol. Evol.* **35**, 1225–1237 (2018).
80. S. Kumar, G. Stecher, M. Suleski, S. B. Hedges, TimeTree: A resource for timelines, time-trees, and divergence times. *Mol. Biol. Evol.* **34**, 1812–1819 (2017).
81. J. P. Bird, R. Martin, H. R. Akçakaya, J. Gilroy, I. J. Burfield, S. T. Garnett, A. Symes, J. Taylor, Ç. H. Şekercioğlu, S. H. M. Butchart, Generation lengths of the world's birds and their implications for extinction risk. *Conserv. Biol.* **34**, 1252–1261 (2020).
82. H. Li, R. Durbin, Inference of human population history from individual whole-genome sequences. *Nature* **475**, 493–496 (2011).
83. K. Nadachowska-Brzyska, C. Li, L. Smeds, G. Zhang, H. Ellegren, Temporal dynamics of avian populations during Pleistocene revealed by whole-genome sequences. *Curr. Biol.* **25**, 1375–1380 (2015).
84. X. Liu, Y.-X. Fu, Stairway Plot 2: Demographic history inference with folded SNP frequency spectra. *Genome Biol.* **21**, 280 (2020).
85. S. Schiffels, K. Wang, MSMC and MSMC2: The Multiple Sequentially Markovian Coalescent, in *Statistical Population Genomics*, J. Y. Duthiel, Ed. (Springer, 2020), pp. 147–166.
86. M. Patterson, T. Marschall, N. Pisanti, L. van Iersel, L. Stougie, G. W. Klau, A. Schönhuth, WhatsHap: Weighted haplotype assembly for future-generation sequencing reads. *J. Comput. Biol.* **22**, 498–509 (2015).
87. O. Delaneau, J.-F. Zagury, M. Robinson, J. Marchini, E. Dermizakis, Integrative haplotype estimation with sub-linear complexity. bioRxiv 493403 (2018). <https://doi.org/10.1101/493403>.
88. E. Santiago, I. Novo, A. F. Pardiñas, M. Saura, J. Wang, A. Caballero, Recent demographic history inferred by high-resolution analysis of linkage disequilibrium. *Mol. Biol. Evol.* **37**, 3642–3653 (2020).
89. J. K. Pickrell, J. K. Pritchard, Inference of population splits and mixtures from genome-wide allele frequency data. *PLOS Genet.* **8**, e1002967 (2012).
90. M. Malinsky, M. Matschiner, H. Svardal, Dsuite - Fast π -statistics and related admixture evidence from VCF files. *Mol. Ecol. Resour.* **21**, 584–595 (2021).
91. B. K. Maples, S. Gravel, E. E. Kenny, C. D. Bustamante, RFMix: A discriminative modeling approach for rapid and robust local-ancestry inference. *Am. J. Hum. Genet.* **93**, 278–288 (2013).
92. F. Racimo, S. Sankararaman, R. Nielsen, E. Huerta-Sánchez, Evidence for archaic adaptive introgression in humans. *Nat. Rev. Genet.* **16**, 359–371 (2015).
93. D. J. Lawson, G. Hellenthal, S. Myers, D. Falush, Inference of population structure using dense haplotype data. *PLOS Genet.* **8**, e1002453 (2012).
94. P. Wangkumhang, M. Greenfield, G. Hellenthal, An efficient method to identify, date, and describe admixture events using haplotype information. *Genome Res.* **32**, 1553–1564 (2022).
95. M. Duranton, F. Bonhomme, P.-A. Gagnaire, The spatial scale of dispersal revealed by admixture tracts. *Evol. Appl.* **12**, 1743–1756 (2019).
96. H. Li, Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997 (2013).
97. X. Zhang, S. Zhang, Q. Zhao, R. Ming, H. Tang, Assembly of allele-aware, chromosomal-scale autopolyploid genomes based on Hi-C data. *Nat. Plants* **5**, 833–845 (2019).
98. J. Ghurye, M. Pop, Modern technologies and algorithms for scaffolding assembled genomes. *PLoS Comput. Biol.* **15**, e1006994 (2019).
99. D. Guan, S. A. McCarthy, J. Wood, K. Howe, Y. Wang, R. Durbin, Identifying and removing haplotypic duplication in primary genome assemblies. *Bioinformatics* **36**, 2896–2898 (2020).
100. N. C. Durand, J. T. Robinson, M. S. Shamim, I. Machol, J. P. Mesirov, E. S. Lander, E. L. Aiden, Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. *Cell Syst.* **3**, 99–101 (2016).
101. A. Bishara, Y. Liu, Z. Weng, D. Kashef-Haghighi, D. E. Newburger, R. West, A. Sidow, S. Batzoglu, Read clouds uncover variation in complex regions of the human genome. *Genome Res.* **25**, 1570–1580 (2015).
102. E. Garrison, G. Marth, Haplotype-based variant detection from short-read sequencing. arXiv:1207.3907 [q-bio.GN] (2012).
103. K. Howe, W. Chow, J. Collins, S. Pelan, D.-L. Poinon, Y. Sims, J. Torrance, A. Tracey, J. Wood, Significantly improving the quality of genome assemblies through curation. *Gigascience* **10**, gjaai153 (2021).
104. D. E. Wood, J. Lu, B. Langmead, Improved metagenomic analysis with Kraken 2. *Genome Biol.* **20**, 257 (2019).
105. R. Challis, E. Richards, J. Rajan, G. Cochrane, M. Blaxter, BlobToolKit - Interactive quality assessment of genome assemblies. *G3 (Bethesda)* **10**, 1361–1374 (2020).
106. D. Guan, S. A. McCarthy, J. M. D. Wood, Y. Sims, W. Chow, Z. Ning, K. Howe, G. Wang, Y. Wang, R. Durbin, Genome sequence assembly evaluation using long-range sequencing data. bioRxiv 2022.05.10.491304 (2022). <https://doi.org/10.1101/2022.05.10.491304>.
107. M. Manni, M. R. Berkeley, M. Seppey, F. A. Simão, E. M. Zdobnov, BUSCO update: Novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol. Biol. Evol.* **38**, 4647–4654 (2021).
108. A. Mikheenko, A. Pribelski, V. Saveliev, D. Antipov, A. Gurevich, Versatile genome assembly evaluation with QUAST-LG. *Bioinformatics* **34**, i142–i150 (2018).
109. M. Goel, H. Sun, W.-B. Jiao, K. Schneeberger, SyRi: Finding genomic rearrangements and local sequence differences from whole-genome assemblies. *Genome Biol.* **20**, 277 (2019).
110. F. Cabanettes, C. Klopp, D-GENIES: Dot plot large genomes in an interactive, efficient and simple way. *PeerJ* **6**, e4958 (2018).
111. A. Rhie, B. P. Walenz, S. Koren, A. M. Phillippy, Merqury: Reference-free quality, completeness, and phasing assessment for genome assemblies. *Genome Biol.* **21**, 245 (2020).
112. G. Formenti, L. Abueg, A. Brajuka, N. Brajuka, C. Gallardo-Alba, A. Giani, O. Fedrigo, E. D. Jarvis, Gfastats: Conversion, evaluation and manipulation of genome sequences using assembly graphs. *Bioinformatics* **38**, 4214–4216 (2022).
113. M. Al Arab, C. H. Z. Siederdisen, K. Tout, A. H. Sahyoun, P. F. Stadler, M. Bernt, Accurate annotation of protein-coding genes in mitochondrial genomes. *Mol. Phylogenet. Evol.* **106**, 209–216 (2017).
114. V. Peona, O. M. Palacios-Gimenez, J. Blommaert, J. Liu, T. Haryoko, K. A. Jönsson, M. Irested, Q. Zhou, P. Jern, A. Suh, The avian W chromosome is a refugium for endogenous retroviruses with likely effects on female-biased mutational load and genetic incompatibilities. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **376**, 20200186 (2021).
115. B. D. Ondov, T. J. Treangen, P. Melsted, A. B. Mallonee, N. H. Bergman, S. Koren, A. M. Phillippy, Mash: Fast genome and metagenome distance estimation using MinHash. *Genome Biol.* **17**, 132 (2016).
116. M. Simonsen, T. Mailund, C. N. S. Pedersen, Rapid neighbour-joining, in *Algorithms in Bioinformatics* (Springer Berlin Heidelberg, 2008), pp. 113–122.
117. J. Armstrong, G. Hickey, M. Diekhans, I. T. Fiddes, A. M. Novak, A. Deran, Q. Fang, D. Xie, S. Feng, J. Stiller, D. Genereux, J. Johnson, V. D. Marinescu, J. Alfoldi, R. S. Harris, K. Lindblad-Toh, D. Haussler, E. Karlsson, E. D. Jarvis, G. Zhang, B. Paten, Progressive Cactus is a multiple-genome aligner for the thousand-genome era. *Nature* **587**, 246–251 (2020).
118. I. T. Fiddes, J. Armstrong, M. Diekhans, S. Nachtweide, Z. N. Kronenberg, J. G. Underwood, D. Gordon, D. Earl, T. Keane, E. E. Eichler, D. Haussler, M. Stanke, B. Paten, Comparative annotation Toolkit (CAT)-simultaneous clade and personal genome annotation. *Genome Res.* **28**, 1029–1038 (2018).
119. M. Stanke, M. Diekhans, R. Baertsch, D. Haussler, Using native and syntentically mapped cDNA alignments to improve de novo gene finding. *Bioinformatics* **24**, 637–644 (2008).
120. P. Jones, D. Binns, H.-Y. Chang, M. Fraser, W. Li, C. McAnulla, H. McWilliam, J. Maslen, A. Mitchell, G. Nuka, S. Pesseat, A. F. Quinn, A. Sangrador-Vegas, M. Scheremetjew, S.-

- Y. Yong, R. Lopez, S. Hunter, InterProScan 5: Genome-scale protein function classification. *Bioinformatics* **30**, 1236–1240 (2014).
121. B. Buchfink, K. Reuter, H.-G. Drost, Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat. Methods* **18**, 366–368 (2021).
 122. UniProt Consortium, UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Res.* **49**, D480–D489 (2021).
 123. J. Dainat, D. Hereñú, E. Davis, K. Crouch, LucileSol, N. Agostinho, pascal-git, tayyrov, NBISweden/AGAT: AGAT-v0.9.2 (2022); <https://zenodo.org/record/6621429>.
 124. S. Lindgreen, AdapterRemoval: Easy cleaning of next-generation sequencing reads. *BMC Res. Notes* **5**, 337 (2012).
 125. Broad Institute, Picard Toolkit; <http://broadinstitute.github.io/picard/>.
 126. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin; 1000 Genome Project Data Processing Subgroup, The genome alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
 127. G. A. Van der Auwera, M. O. Carneiro, C. Hartl, R. Poplin, G. Del Angel, A. Levy-Moonshine, T. Jordan, K. Shakir, D. Roazen, J. Thibault, E. Banks, K. V. Garimella, D. Altschuler, S. Gabriel, M. A. DePristo, From FastQ data to high confidence variant calls: The genome analysis toolkit best practices pipeline. *Curr. Protoc. Bioinformatics* **43**, 11.10.1–11.10.33 (2013).
 128. A. Manichaikul, J. C. Mychaleckyj, S. S. Rich, K. Daly, M. Sale, W.-M. Chen, Robust relationship inference in genome-wide association studies. *Bioinformatics* **26**, 2867–2873 (2010).
 129. N. M. Kopelman, J. Mayzel, M. Jakobsson, N. A. Rosenberg, I. Mayrose, Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Mol. Ecol. Resour.* **15**, 1179–1191 (2015).
 130. G. Evanno, S. Regnaut, J. Goudet, Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol.* **14**, 2611–2620 (2005).
 131. A. Prasad, E. D. Lorenzen, M. V. Westbury, Evaluating the role of reference-genome phylogenetic distance on evolutionary inference. *Mol. Ecol. Resour.* **22**, 45–55 (2022).
 132. T. Kawakami, L. Smeds, N. Backström, A. Husby, A. Qvarnström, C. F. Mugal, P. Olason, H. Ellegren, A high-density linkage map enables a second-generation collared flycatcher genome assembly and reveals the patterns of avian recombination rate variation and chromosomal evolution. *Mol. Ecol.* **23**, 4035–4058 (2014).
 133. P. U. Clark, A. S. Dyke, J. D. Shakun, A. E. Carlson, J. Clark, B. Wohlfarth, J. X. Mitrovica, S. W. Hostetler, A. M. McCabe, The last glacial maximum. *Science* **325**, 710–714 (2009).
 134. M. E. Mann, Z. Zhang, S. Rutherford, R. S. Bradley, M. K. Hughes, D. Shindell, C. Ammann, G. Faluvegi, F. Ni, Global signatures and dynamical origins of the little ice age and medieval climate anomaly. *Science* **326**, 1256–1260 (2009).
 135. N. P. McKay, D. S. Kaufman, C. C. Routson, M. P. Erb, P. D. Zander, The onset and rate of Holocene Neoglacial cooling in the arctic. *Geophys. Res. Lett.* **45**, 12,487–12,496 (2018).
 136. A. B. Patil, N. Vijay, Repetitive genomic regions and the inference of demographic history. *Heredity* **127**, 151–166 (2021).
 137. R. S. Taylor, M. Manseau, C. F. C. Klütsch, J. L. Polfus, A. Steedman, D. Hervieux, A. Kelly, N. C. Larter, M. Gamberg, H. Schwantje, P. J. Wilson, Population dynamics of caribou shaped by glacial cycles before the last glacial maximum. *Mol. Ecol.* **30**, 6121–6143 (2021).
 138. C. Pockrandt, M. Alzamel, C. S. Iliopoulos, K. Reinert, GenMap: Ultra-fast computation of genome mappability. *Bioinformatics* **36**, 3687–3692 (2020).
 139. L. Deng, B. Xie, Y. Wang, X. Zhang, S. Xu, A protocol for applying a population-specific reference genome assembly to population genetics and medical studies. *STAR Protoc.* **3**, 101440 (2022).
 140. M. Duranton, F. Allal, S. Valière, O. Bouchez, F. Bonhomme, P.-A. Gagnaire, The contribution of ancient admixture to reproductive isolation between European sea bass lineages. *Evol. Lett.* **4**, 226–242 (2020).
 141. O. G. Zedlitz, Ornithologische Notizen von der "Zeppelin-Studienfahrt" Spitzbergen Sommer 1910 (1911); <https://zenodo.org/record/2361488>.
 142. J. E. McCormack, W. L. E. Tsai, B. C. Faircloth, Sequence capture of ultraconserved elements from bird museum specimens. *Mol. Ecol. Resour.* **16**, 1189–1203 (2016).
 143. S. Feng, Q. Fang, R. Barnett, C. Li, S. Han, M. Kuhlwlilm, L. Zhou, H. Pan, Y. Deng, G. Chen, A. Gamauf, F. Woog, R. Prys-Jones, T. Marques-Bonet, M. T. P. Gilbert, G. Zhang, The genomic footprints of the fall and recovery of the crested Ibis. *Curr. Biol.* **29**, 340–349.e7 (2019).
 144. M. T. P. Gilbert, H.-J. Bandelt, M. Hofreiter, I. Barnes, Assessing ancient DNA studies. *Trends Ecol. Evol.* **20**, 541–544 (2005).
 145. B. Llamas, G. Valverde, L. Fehren-Schmitz, L. S. Weyrich, A. Cooper, W. Haak, From the field to the laboratory: Controlling DNA contamination in human ancient DNA research in the high-throughput sequencing era. *STAR* **3**, 1–14 (2017).
 146. R Core Team, R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing 2020); www.R-project.org/.
 147. E. A. Fox, A. E. Wright, M. Fumagalli, F. G. Vieira, ngsLD: Evaluating linkage disequilibrium using genotype likelihoods. *Bioinformatics* **35**, 3855–3856 (2019).
 148. S. Van Dongen, Graph clustering via a discrete uncoupling process. *SIAM J. Matrix Anal. Appl.* **30**, 121–141 (2008).
 149. L. Li, C. J. Stoeckert Jr., D. S. Roos, OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Res.* **13**, 2178–2189 (2003).
 150. M. V. Matz, Fantastic beasts and how to sequence them: Ecological genomics for obscure model organisms. *Trends Genet.* **34**, 121–132 (2018).
 151. H. B. Mann, D. R. Whitney, On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Stat.* **18**, 50–60 (1947).
- Acknowledgments:** We are grateful to the SEAPOPOP program (www.seapop.no/en, Norwegian Research Council grant number 192141) and SEATRACK (www.seapop.no/en/seatrack) project for collecting and sharing modern Atlantic puffin samples. We also thank the Natural History Museum (NHM) in Stockholm and the American Museum of Natural History (AMNH, New York) for providing museum specimens. AMNH and NHM Stockholm are the holders of the historical Atlantic puffin specimens. In particular, we appreciate the help of T. J. Trombone (AMNH, New York) with sampling and sending historical Atlantic puffin toe pads. Special thanks go to the Norwegian Sequencing Centre (University of Oslo; www.sequencing.uio.no) for the PacBio and Hi-C library preparation and sequencing used for the assembly, as well as for the resequencing of samples analyzed in this study. Computation was performed using the resources and assistance from SIGMA2. **Funding:** This work was financially supported by the Faculty of Mathematical and Natural Sciences, University of Oslo (UiO) to S.B. and from The Research Council of Norway (RCN) to O.K.T. and K.S.J. (grant no. 326819). **Author contributions:** Conceptualization: S.B., B.S., K.S.J., and O.K. Sampling: T.A.-N., S.D., H.S., U.S.J., and P.R.S. Laboratory work: O.K. and L.M.A. Assembly and annotation: A.K.K., O.K.T., and O.K. Genomic analyses: O.K. Visualization: O.K., S.B., B.S., and K.S.J. Ecological context: T.A.-N., S.D., and H.S. Supervision: S.B., K.S.J., and B.S. Writing—original draft: O.K., S.B., B.S., and K.S.J. Writing—review and editing: O.K., S.B., B.S., K.S.J., L.M.A., A.K.K., O.K.T., T.A.-N., S.D., H.S., U.S.J., and P.R.S. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Raw read data analyzed in the current study have been deposited in the European Nucleotide Archive (ENA; www.ebi.ac.uk/ena) under study accession number PRJEB40631 (see data S1 for individual sample accession numbers). Nuclear and mitochondrial sequences of the chromosome-level reference genome (GCA_947846985 and CANTUZ010000001 to CANTUZ010000028) have also been uploaded to ENA (Project PRJEB58718, Sample SAMEA112283621), as well as all unplaced scaffolds (SAMEA113591597). The annotation of the reference genome (.gff file) and all code used for the population genomic analyses are available on the first author's GitHub (<https://github.com/OKersten/BjornoyaHybrids>, accessed December 2022) and on Zenodo under <https://doi.org/10.5281/zenodo.7531157>. This includes versions of any software used, if relevant, and any specific variables or parameters used to generate, test, and process the dataset of this study.

Submitted 17 February 2023

Accepted 6 September 2023

Published 6 October 2023

10.1126/sciadv.adh1407

Hybridization of Atlantic puffins in the Arctic coincides with 20th-century climate change

Oliver Kersten, Bastiaan Star, Anders K. Krabberød, Lane M. Atmore, Ole K. Tørresen, Tycho Anker-Nielsen, Sébastien Descamps, Hallvard Strøm, Ulf S. Johansson, Paul R. Sweet, Kjetill S. Jakobsen, and Sanne Boessenkool

Sci. Adv. **9** (40), eadh1407. DOI: 10.1126/sciadv.adh1407

View the article online

<https://www.science.org/doi/10.1126/sciadv.adh1407>

Permissions

<https://www.science.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of service](#)

Science Advances (ISSN 2375-2548) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.

Copyright © 2023 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution License 4.0 (CC BY).