Article

Climate-driven flyway changes and memory-based long-distance migration

https://doi.org/10.1038/s41586-021-03265-0

Received: 22 January 2020

Accepted: 20 January 2021

Published online: 3 March 2021

Check for updates

Zhongru Gu[1,2,3,18], Shengkai Pan[1,2,18], Zhenzhen Lin[1,2,18], Li Hu[1,2,3], Xiaoyang Dai[4], Jiang Chang[5], Yuanchao Xue[6], Han Su[1,2,3], Juan Long[1,2,3], Mengru Sun[3,6], Sergey Ganusevich[7], Vasiliy Sokolov[8], Aleksandr Sokolov[9], Ivan Pokrovsky[9,10,11], Fen Ji[12], Michael W. Bruford[2,13], Andrew Dixon[2,14,15,16] & Xiangjiang Zhan[1,2,3,17]✉

Millions of migratory birds occupy seasonally favourable breeding grounds in the Arctic[1], but we know little about the formation, maintenance and future of the migration routes of Arctic birds and the genetic determinants of migratory distance. Here we established a continental-scale migration system that used satellite tracking to follow 56 peregrine falcons (*Falco peregrinus*) from 6 populations that breed in the Eurasian Arctic, and resequenced 35 genomes from 4 of these populations. The breeding populations used five migration routes across Eurasia, which were probably formed by longitudinal and latitudinal shifts in their breeding grounds during the transition from the Last Glacial Maximum to the Holocene epoch. Contemporary environmental divergence between the routes appears to maintain their distinctiveness. We found that the gene *ADCY8* is associated with population-level differences in migratory distance. We investigated the regulatory mechanism of this gene, and found that long-term memory was the most likely selective agent for divergence in *ADCY8* among the peregrine populations. Global warming is predicted to influence migration strategies and diminish the breeding ranges of peregrine populations of the Eurasian Arctic. Harnessing ecological interactions and evolutionary processes to study climate-driven changes in migration can facilitate the conservation of migratory birds.

Global climate change and anthropogenic development are expected to affect the annual adaptive movements of migratory Arctic birds[1–3], leading to potential fitness effects imposed by inhospitable routes and temporally mismatched breeding[2,4]. Next-generation genome sequencing has facilitated studies of the interaction between genomic variation and environment in migratory birds[5]. However, to our knowledge, there is to date no published research on the role of climate-driven genomic responses in shaping differences in migratory strategy among bird populations. Here we combined the use of satellite tracking of 56 peregrine falcons from migratory Arctic populations[6] (Fig. 1a, Extended Data Fig. 1, Supplementary Table 1) with genome data to explore the demographic history of these populations, and the spatiotemporal dynamics of their migrations.

Migration patterns of Arctic peregrines

From 41 individual birds, we identified 150 completed migration paths (Supplementary Table 2). Peregrines initiated their autumn migration mainly in September, travelled 2,280–11,002 km in about 27 days (95% confidence interval, 14–46 days) (covering 213 km per day (95% confidence interval, 49–420 km per day)) and arrived at their wintering areas mainly in October. Peregrines migrate solitarily; birds that depart from different breeding grounds use different routes (except for those departing from Kola and Kolguev, which use the same route), and winter at widely distributed sites across four distinct regions (Fig. 1a, Extended Data Fig. 2). Individual birds ($n = 26$) that were tracked for more than one year exhibited strong path repeatability during migration ($R = 0.45$, $P < 0.001$), complete fidelity to wintering locations and limited breeding dispersal (5.37 km on average) (Fig. 1b, Supplementary Table 3). All populations demonstrated a high degree of migratory connectivity ($R^2 = 0.86$, $P < 0.001$) (Fig. 1c), which suggests strong selection for long-term memory.

We performed a principal component analysis and identified two main groups with migratory distance as their most significant differentiation, identified by our random forest modelling and *t*-tests (Fig. 1d, e, Extended Data Fig. 3, Supplementary Table 4). Birds that took eastern

[1]Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing, China. [2]Cardiff University–Institute of Zoology Joint Laboratory for Biocomplexity Research, Chinese Academy of Sciences, Beijing, China. [3]University of the Chinese Academy of Sciences, Beijing, China. [4]School of Biological Sciences, University of Bristol, Bristol, UK. [5]State Key Laboratory of Environmental Criteria and Risk Assessment, Chinese Research Academy of Environmental Sciences, Beijing, China. [6]Key Laboratory of RNA Biology, CAS Center for Excellence in Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China. [7]Wild Animal Rescue Centre, Moscow, Russia. [8]Institute of Plant and Animal Ecology, Ural Division Russian Academy of Sciences, Ekaterinburg, Russia. [9]Arctic Research Station of the Institute of Plant and Animal Ecology, Ural Division Russian Academy of Sciences, Labytnangi, Russia. [10]Department of Migration, Max Planck Institute of Animal Behavior, Radolfzell, Germany. [11]Laboratory of Ornithology, Institute of Biological Problems of the North FEB RAS, Magadan, Russia. [12]State Key Laboratory of Stem Cell and Reproductive Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing, China. [13]School of Biosciences and Sustainable Places Institute, Cardiff University, Cardiff, UK. [14]Emirates Falconers' Club, Abu Dhabi, United Arab Emirates. [15]Reneco International Wildlife Consultants, Abu Dhabi, United Arab Emirates. [16]International Wildlife Consultants, Carmarthen, UK. [17]Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming, China. [18]These authors contributed equally: Zhongru Gu, Shengkai Pan, Zhenzhen Lin. ✉e-mail: zhanxj@ioz.ac.cn

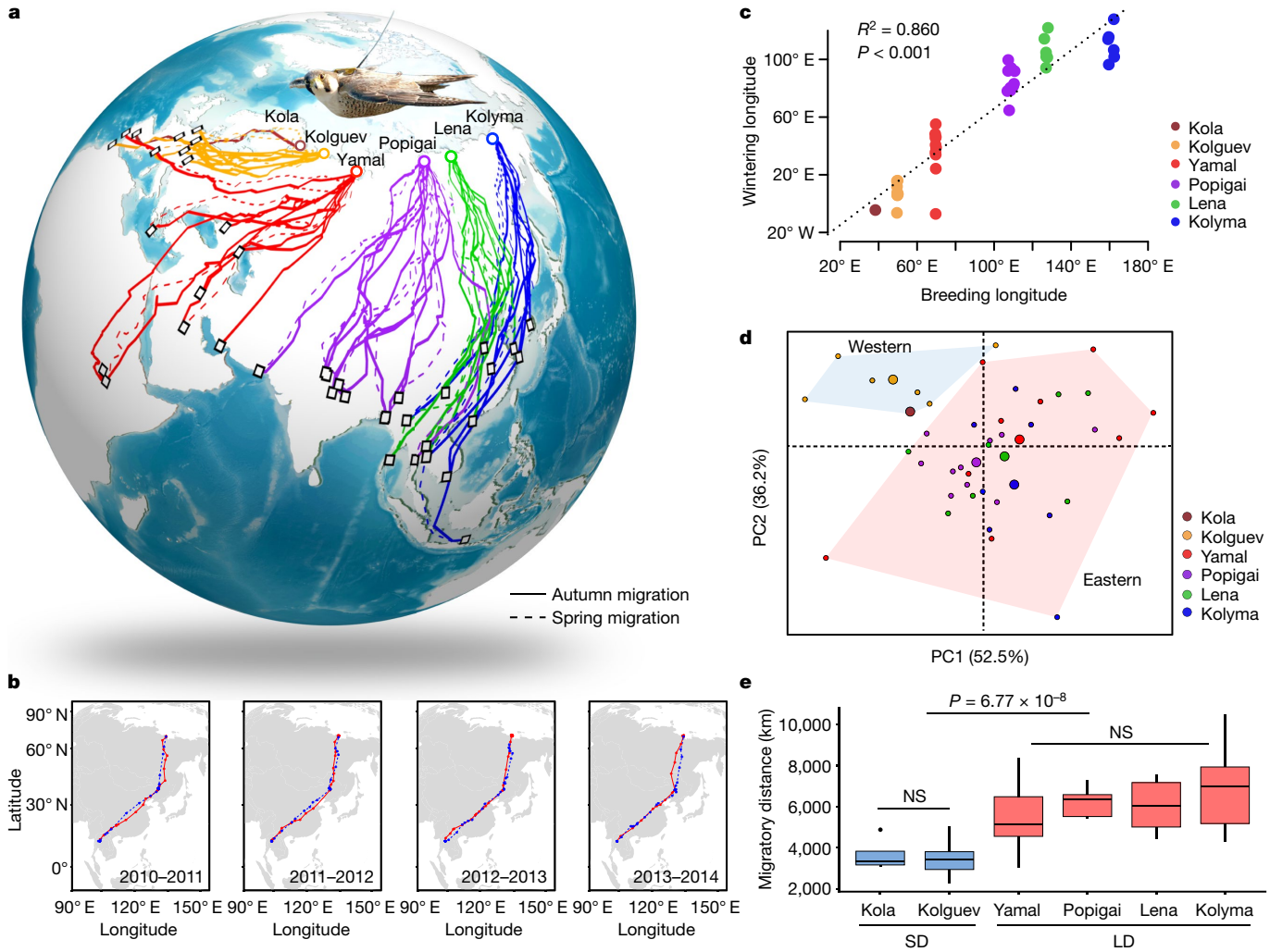


Fig. 1 | Migration system. a, Five migration routes for 56 peregrine falcons tracked by satellite. Only complete migration paths are shown (n = 41). Public ETOPO1 data (downloaded from https://ngdc.noaa.gov/mgg/global/relief/ETOPO1/image/) were used to plot the ocean bathymetry map in R. **b**, Individual migration-path fidelity of one representative individual bird (from the Lena Delta) that was tracked for four years (2010–2014). **c**, Migration connectivity at the population level, as shown in a linear regression analysis. Significance was calculated using F test ($P = 2.2 \times 10^{-16}$). **d**, Principal component analysis (PCA) of migratory strategy. Small and large dots represent individuals (n = 41) and centroids of the minimum convex polygon for each population (n = 6), respectively. Principal component (PC)1 is the mix of autumn departure and arrival dates, and PC2 is the migration distance. **e**, Comparison of migratory distance ($P < 1 \times 10^{-6}$, two-sided t-tests, effect size = 1.43) between short-distance (SD) (n = 12 individual birds; 3,680 km: 95% confidence interval, 2,443–5,018 km) and long-distance (LD) groups (n = 32 individual birds; 6,134 km: 95% confidence interval, 3,282–8,828 km). In the box plots, the centre line represents the median, whiskers represent maximum and minimum values, and box boundaries represent 75th and 25th percentiles. The P values for the comparisons for any two populations within short-distance and long-distance were not significant (NS) ($P > 0.05$, two-sided t-tests). The tracking data of three additional Kola peregrines from a previous study (Methods) were also used for this analysis.

routes flew significantly farther than did birds that travelled along western routes (6,134 km versus 3,680 km, respectively) ($P < 1 \times 10^{-6}$) (Fig. 1e). We therefore classified these two groups as long-distance migrants (birds from Kolyma, Lena, Popigai and Yamal) and short-distance migrants (birds from Kolguev and Kola).

Historical formation of migration routes

We sequenced the genomes of 35 peregrines and obtained 6,328,655 high-quality single-nucleotide polymorphisms (SNPs) (Supplementary Table 5). We used several analytical approaches that found consistent support for four distinct genetic clusters, which correspond to the sequenced populations; the Yamal and Kolyma populations are inferred to have diverged after the separation of their ancestors from the ancestor of the Kola and Kolguev populations (Fig. 2a). Our sequential Markovian coalescent analysis revealed that the effective population size (N_e) of the ancestral lineage increased over time from about 100 thousand years ago (ka) to a peak at 20–30 ka (Fig. 2b) around the Last Glacial Maximum (LGM)[7]. To resolve uncertainties in the recent demographic history (Fig. 2a), we developed an approximate Bayesian computation (ABC) approach. Our choice of ABC–random forest model (Fig. 2c, Supplementary Figs. 1–4, Supplementary Tables 6, 7) confirmed the divergence pattern of the four studied populations, and ABC simulations further found that long-distance and short-distance populations started to separate during the LGM (23.03 ka; 95% confidence interval, 17.67–32.94 ka), which was followed by a split between the Yamal and Kolyma populations at 11.30 ka (95% confidence interval, 9.14–14.29 ka) and between the Kola and Kolguev populations at 10.53 ka (95% confidence interval, 9.18–12.90 ka) (Fig. 2d, Extended Data Fig. 4, Supplementary Table 8).

Ecological niche modelling on the basis of present and palaeoclimate datasets showed that the distribution of potential breeding range

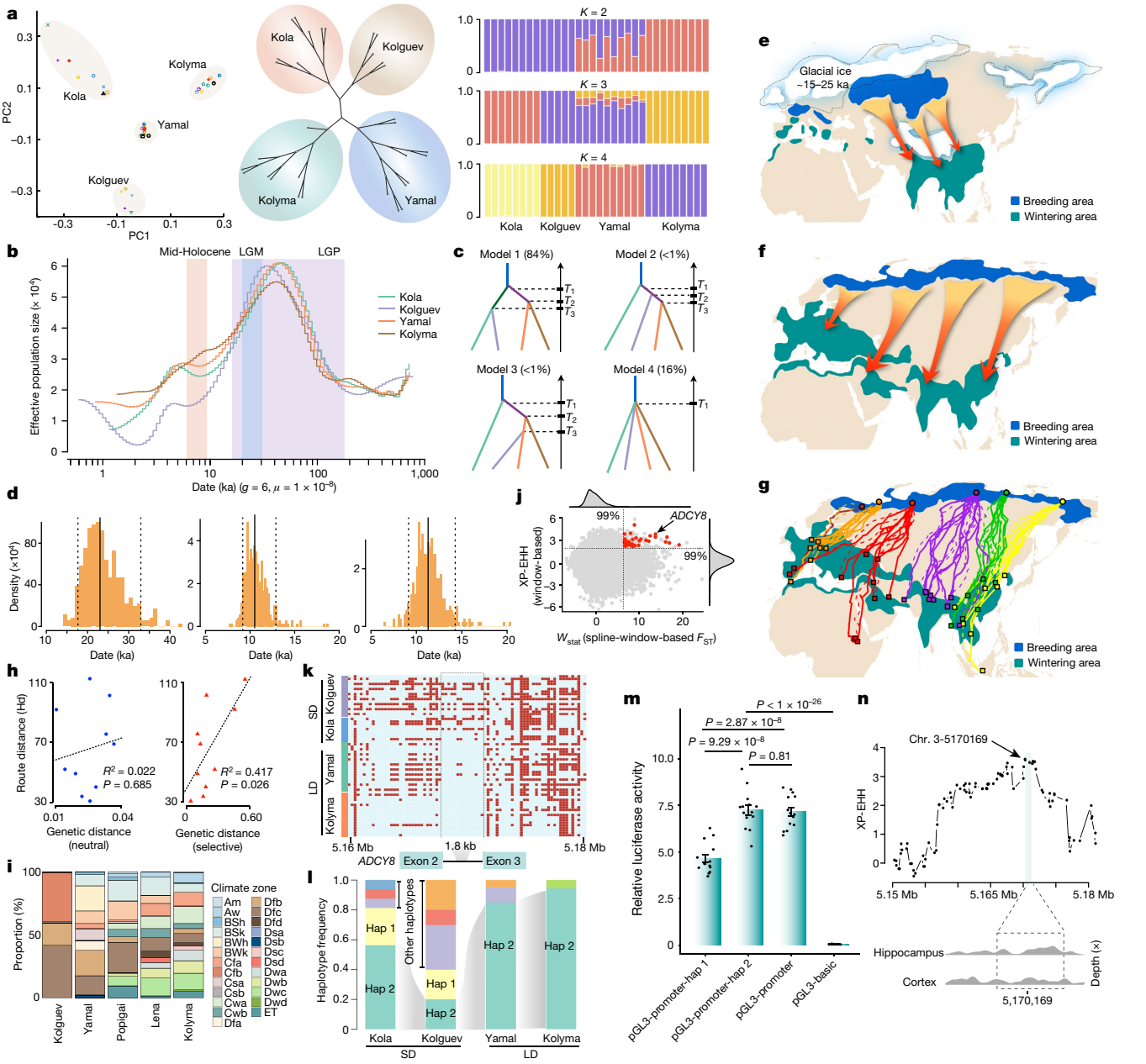


Fig. 2 | Past formation and present maintenance of migration routes, and genetic basis for differences in migration distance. a, PCA, neighbour-joining tree and frappe analysis, showing the evolutionary relationship of the four genome-enabled populations. **b**, Reconstruction of demographic history for each population using sequential Markovian coalescent analysis. LGP, Last Glacial Period. **c**, Four candidate models for model choice in ABC. Eighty-four per cent of the total of 313 chunks (Methods) support model 1. T_1, T_2 and T_3 are divergence times. **d**, Posterior distribution of divergence time estimates for model 1 in ABC. Left, T_1; middle, T_2; right, T_3. Solid line, median; dashed lines, 95% confidence interval. **e**–**g**, Species distributions predicted during the LGM (**e**), middle of the Holocene (**f**) and present (**g**). **h**, The relationship between the route distance (Hd) and neutral genetic distance ($F_{ST}/(1-F_{ST})$) (left) and genetic distance on the basis of selected loci (right). The dashed line is the linear regression line. Significance levels were calculated using F test. **i**, Proportion of grids (0.083° × 0.083°) with Köppen–Geiger climate zones within each migration route. Full names for the climate zone abbreviations are provided in Supplementary Information. **j**, Spline-window-based F_{ST} and XP-EHH to detect selective sweeps. Red points indicate windows that contain selected genes. **k**, *ADCY8*-haplotype (hap) heat map. The dashed rectangle marks the focal 1.8-kb fragment in *ADCY8*. Red and light blue squares symbol different alleles in each column (SNP). **l**, Haplotype frequency in the identified segment. **m**, Results of dual-luciferase reporter assay in chicken hippocampus primary cells. Data are mean ± s.e.m. Significance levels were calculated using two-sided t-test ($n = 14$ replicates for each of the first 3 groups, and $n = 6$ for the pGL3-basic group). **n**, XP-EHH results for every selected SNP within 1.8-kb flanking regions (top) and ATAC-seq results confirming the existence of CRE-motif (bottom).

positively correlated with fluctuations in N_e (Supplementary Figs. 5, 6). A much larger area of Siberia was suitable for breeding during the LGM than in the last interglacial period (120–140 ka)[8] or middle of the Holocene (5–7 ka), coinciding with the largest estimate of N_e (Fig. 2e, f, Supplementary Fig. 5). Arctic-dwelling peregrines occupy mainly tundra habitat[9], and we found a close coincidence between reconstructed tundra habitat and the distribution of peregrine breeding range in the LGM (Supplementary Fig. 7), which suggests that the enlargement of tundra habitat underpinned the expansion of the peregrine population during the LGM. Conversely, population declines and gradual divergences after the LGM mirror the large-scale loss and northward contraction of tundra. Recent population declines after the middle of

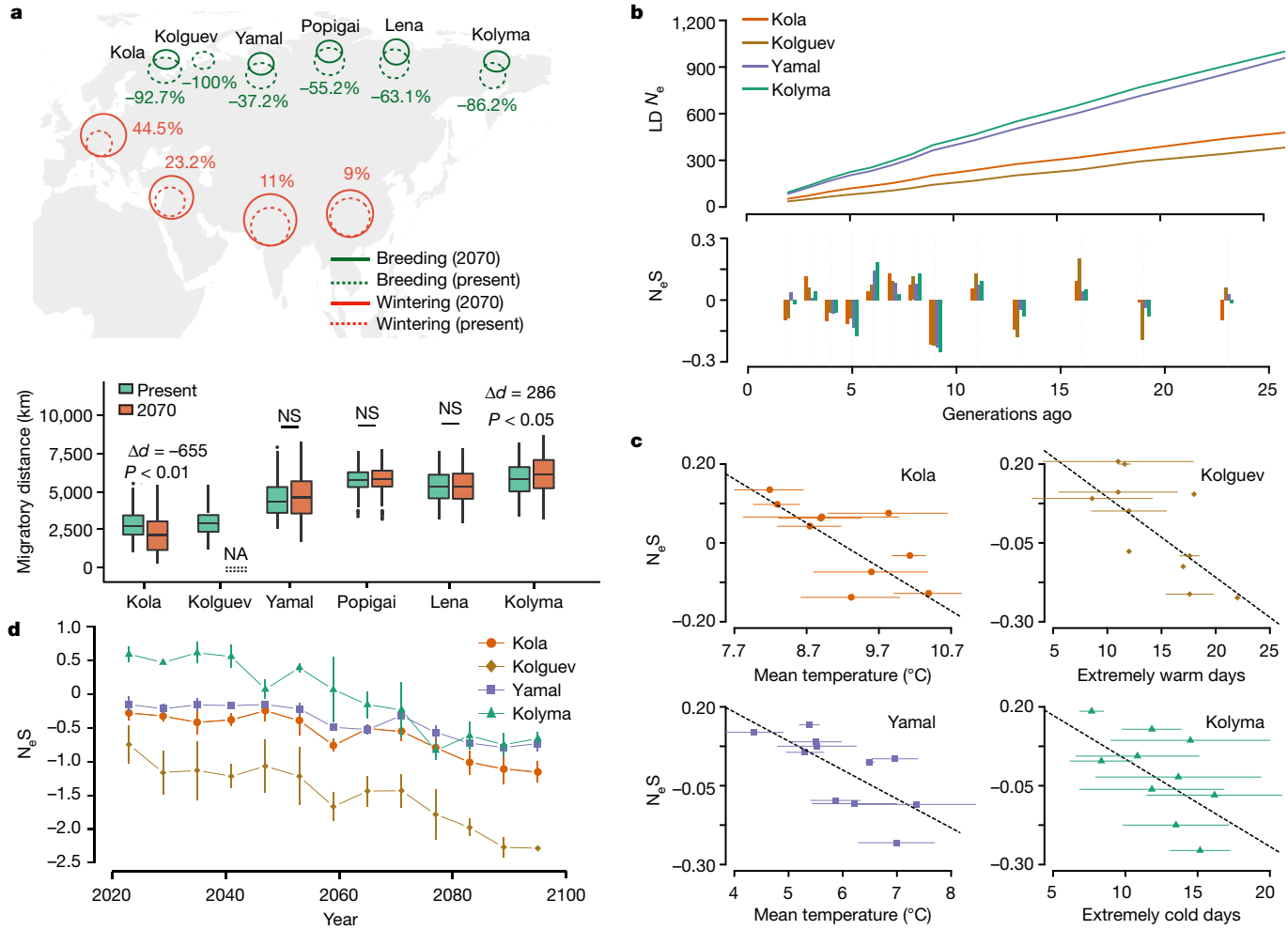


Fig. 3 | Shortened migration route and population decline in European populations owing to global warming. a, Area changes in breeding and wintering areas (top) between present and future (2070, RCP 8.5) and migratory distance comparisons (bottom) in the six peregrine populations. NA, no predicted future breeding areas. *Δd*, mean change of migration distances. In the box plots, the centre line represents the median, whiskers represent maximum and minimum values, and box boundaries represent 75th and 25th percentiles. $n = 200$ for each comparison. Significance and effect size were calculated using two-sided *t*-tests and Cohen's *d*, respectively (Kola, $P = 3.4 \times 10^{-9}$, effect size = 0.580; Yamal, $P = 0.090$, effect size = 0.170; Popigai, $P = 0.676$, effect size = 0.042; Lena, $P = 0.868$, effect size = 0.017; Kolyma, $P = 0.015$, effect size = 0.243). **b**, Estimates of N_e and N_eS for long-distance peregrines for the most recent 25 generations. **c**, The linear regression between N_eS and the most significant environmental variable identified in the generalized linear modelling analysis. Data are median ± 95% confidence interval. Dashed lines represent linear regression lines. **d**, N_eS changes of each population predicted in the future. Data are median ± 95% confidence interval.

the Holocene (Fig. 2b) may have also resulted from anthropogenic factors, as habitat distributions have remained relatively stable (Fig. 2f, g).

Our ecological niche modelling simulations suggest that peregrines had a much smaller potential western wintering area during the LGM, whereas the eastern wintering areas remained stable (Fig. 2e, Supplementary Fig. 5). During the LGM, peregrines therefore probably migrated to a wintering area across India and Southeast Asia—a south-eastward migration (Fig. 2e) that is distinct from the current south-westerly migration route that formed during the middle of the Holocene (Fig. 2f, g). Furthermore, because breeding areas are inferred to have shifted northward in the middle of the Holocene compared with the LGM (resulting in a longer migratory route (Fig. 2e, f)), we conclude that glacial cycles can regulate both migratory orientation and distance.

Separation of present migration routes

We used the Hausdorff distance (Hd)[10] to quantify the distance between individual migration paths (Methods, Supplementary Information). Mean Hd within populations (17.05 ± 7.20) was significantly lower than that between neighbouring populations (35.83 ± 14.24) ($P < 0.01$). Our cluster analysis largely supported five migratory routes, with Kola and Kolguev using the same route and very few individuals interchanging between populations (Extended Data Fig. 5a).

The proposition that migration routes are genetically determined is mostly based on migratory restlessness[11], displacement experiments[12] and correlations between genetic background and migration route[13]. We addressed this fundamental question by randomly selecting 93 putatively neutrally evolved SNP loci and 75 loci under positive selection (Methods). Using these markers, we genotyped nine and six individuals, respectively, for the Popigai and Lena populations, from which we had obtained insufficient DNA from shed feathers for genome resequencing. Combining these genotypes of 15 additional individuals with those from the population genomic data, we measured genetic differentiation (F_{ST}) among 5 peregrine populations and tested their relationship with the mean Hd for each migration route. We found a nonsignificant correlation between route Hd and F_{ST} ($R^2 = 0.02$, $P = 0.69$) on the basis of neutral loci, which suggests that demographic isolation

is not the major factor that isolates migration routes. By contrast, we found a significant relationship for loci under selection ($R^2 = 0.42$, $P = 0.03$), which suggests a role for adaptive genomic regions in the maintenance of separated migration routes (Fig. 2h).

Environmental divergence was evident among the migration routes, as Köppen–Geiger climate zones were significantly different between adjacent routes ($P < 0.01$, χ^2 test) (Fig. 2i, Extended Data Fig. 5b) and during most breeding and wintering periods (Supplementary Fig. 8). We found a coincidence between abrupt changes in climate zone and the boundaries of the routes used by different populations (Extended Data Fig. 5c, d). Migration cost (on the basis of least-cost paths) was significantly lower for birds that followed their population-specific route ($P = 0.01$, effective size = 0.45) (Extended Data Fig. 5e, f). We conclude that current migration routes are maintained mainly by environmental constraints, with a synergistic contribution from local adaptation.

Genes for differences in migratory distance

We used window-based F_{ST} and cross-population extended haplotype homozygosity (XP-EHH) to detect selection signals across the genomes of short-distance and long-distance peregrines. Combined, these methods identified 149 selection sweeps (37 genes) between the 2 groups (Supplementary Table 9), and found that the most significant outlier occurred at the *ADCY8* locus (Fig. 2j). We narrowed the signal down to a 1.8-kb region that contains 14 linked SNPs, in the second intron of the gene (Fig. 2k). Haplotype frequency analysis demonstrated a positive-selection signature of *ADCY8* in the long-distance group (Fig. 2l). The dominant haplotype (which we designate haplotype 2) was at a high frequency (34 out of 38 haplotypes) in the long-distance group, whereas the short-distance group had six haplotypes at a range of frequencies (haplotype 2, 11 out of 26 haplotypes; haplotype 1, 6 out of 26 haplotypes; the other four haplotypes occurred 4, 2, 2 and 1 times) (Fig. 2l). We investigated the potential functional importance of haplotype 2 in the long-distance peregrines and a randomly selected haplotype-1 sequence from the short-distance peregrines by designing a dual luciferase reporter assay for functional analysis in cultured chicken hippocampus primary cells (Methods). Haplotype-1 insert cells showed a significantly lower luciferase activity than that of haplotype-2 insert cells ($P < 0.001$) (Fig. 2m, Supplementary Fig. 9), which suggests that the peregrine haplotype-1 sequence has a suppressing effect.

Of the 14 loci that we identified within the *ADCY8* locus, the SNP (C/T) in position 5,170,169 of peregrine chromosome 3 produced the largest XP-EHH value (Fig. 2n) and the allele T was 100% fixed in the long-distance group, which suggests that this standing variation is under strongest selection and may have a major role in functional differentiation. A search of the focal fragment against the motif database found that the ancestral short-distance group had a 5′-CGTCA-3′ motif (a canonical half-site cAMP-responsive element (CRE) that is a binding site for the transcription factor CREB[14,15]), whereas the fixed chromosome 3 5170169*T changes the first nucleotide of this motif. We used assay for transposase-accessible chromatin using sequencing (ATAC-seq)[16] to sequence both hippocampus and cerebral cortex tissues from a short-distance peregrine. Our ATAC-seq analysis detected a peak that coincided with the position of this motif (Fig. 2n), which provided experimental support for its existence as a functional element.

Previous studies have suggested that CREB can regulate gene expression by binding to the CRE element through the CREB basic region/leucine zipper domain (bZIP), which can be regulated by its DNA methylation level[17,18]. In peregrines, the identified substitution from C to T in *ADCY8* creates a novel transcriptional binding site (5′-TGTCA-3′) that potentially disrupts the DNA methylated site (a CpG island) on the canonical motif. Moreover, we found that *CREB1* (the transcription activator) expressed the bZIP domain in the peregrine brain, but that the conditional repressor *CREB2*[19] did not (Supplementary Fig. 10). Our results suggest that the new CRE motif may be free from DNA methylation and facilitate the binding of CREB1 to *ADCY8*, and could consequently maintain a higher activity of ADCY8 in long-distance peregrines. Our comparison of expression levels of *ADCY8* in the peregrine brain with two genotypes of the focal SNP provided supporting evidence (reads per kilobase of transcript per million mapped reads of 81.11 in CT, versus 73.96 in CC; $P = 3.73 \times 10^{-5}$, hypergeometric test).

Previous empirical evidence indicates that *ADCY8* is involved in long-term memory[20,21]. *ADCY8* encodes adenylyl cyclase type 8 that catalyses the conversion of ATP to cAMP, which acts as a secondary messenger and regulates downstream memory-related genes[22,23]. We found that the mean migration-path fidelity was significantly higher in long-distance peregrines (which require strong long-term memory) than in short-distance peregrines ($P < 0.001$) (Supplementary Fig. 11). Together with previous numerical evidence that CREB genes determine the development of long-term memory[24], our work suggests that the *ADCY8* and CREB genes have a key role in influencing the distance of migratory flights via coregulation of the capacity for long-term memory. For the T allele (which is positively selected in the long-distance populations), we estimated a date for selection of 18.87 ka (Supplementary Information)—this is after the separation of two migratory groups, which strengthens our conclusion that the regulation of migratory distances is the result of natural selection.

Predicted effect of global warming

We used ecological niche modelling simulations to project future (2070) breeding and wintering distributions for each population of Arctic peregrines under Representative Concentration Pathway (RCP) 8.5. The breeding and wintering distributions of all populations would shift poleward by 2.08° (95% confidence interval, 0.31–3.44°) and 1.47° (0.11–11.06°) latitude, respectively (Extended Data Fig. 6, Supplementary Fig. 12), which is consistent with observations for most Arctic shorebirds[25] and congruent with the climatic envelope that corresponds to tundra habitats. The greatest reduction is predicted to occur in the Kolguev and Kola populations, which would lose 100% and 93% of their suitable breeding habitats, respectively (Fig. 3a). We also found short-distance peregrines may have a much shorter migration route (655 km; 95% confidence interval, 442–868 km), whereas long-distance peregrines may have a longer route (286 km; 95% confidence interval, 56–515 km) (Fig. 3a, Supplementary Fig. 13). If the climate warms at the same rate as it has over recent decades, peregrines in western Eurasia may stop migrating altogether and Eastern peregrines may face greater risks, as mortality is positively associated with migratory distance[26].

Recent population declines in migratory Arctic birds have been attributed to the amplification of global warming in the high Arctic[1,4,25]. Climate change may have already affected peregrine populations, so we compared N_e changes of each population with local temperature during breeding periods (May–July) since 1840. Our analysis using SNeP showed that each population has undergone declines during the past 25 generations (about 150 years). However, analysis of the N_e slope (N_eS) revealed variation with recovery detected during relatively cooler summers (Fig. 3b), and that four populations showed the largest negative N_eS about eight or nine generations ago (corresponding to the 1960s, and coinciding widespread use of organochlorine pesticides[27]). Our generalized linear modelling analysis found that change in N_e was negatively correlated with mean breeding season temperature in Kola ($R^2 = 0.46$, $P = 0.03$) and Yamal ($R^2 = 0.39$, $P = 0.02$), with mean breeding season temperature and number of extremely warm days in Kolguev ($R^2 = 0.61$, $P = 0.01$), and with number of extremely cold and warm days in Kolyma ($R^2 = 0.68$, $P = 0.02$) (Fig. 3c, Supplementary Table 10). Our predicted N_eS in the future (that is, from 2020 to 2100) showed a continuing decreasing trend, but short-distance migrants in Kola and Kolguev will suffer the highest probability of population decline (Fig. 3d, Supplementary Fig. 14).

Discussion

Spatiotemporal changes in the migration behaviours of animals are thought to be related to climatic changes, anthropogenic effects[28] and the evolutionary responses of migrants. As these dynamic processes can leave a footprint on the genome, we were able to combine animal movement and population genomic data to identify a major role of climate in the formation and maintenance of the migration patterns of peregrines (Fig. 2, Extended Data Fig. 5).

Previous studies have identified several candidate genes that may regulate migration[29,30]. The higher activity of *ADCY8* that we identified in long-distance peregrine migrants (Fig. 2) may increase their long-term memory. Our analysis reveals a unique mutation that facilitates the binding of the transcription factor CREB1 to *ADCY8*, and fixation of this variation happened after the divergence of long-distance and short-distance populations. Our work thus not only reveals a causative gene that may explain migratory differences, but also provides a mechanistic basis for these differences.

In a changing global climate, peregrines may move to new wintering areas and adjust their migration routes. However, our prediction of marked shrinkage in Arctic breeding areas—together with a predicted collapse of peregrine populations in the European Arctic—represents a clear threat to peregrines, and possibly many other migratory Arctic species. Our study demonstrates the value of an integrated approach that combines satellite telemetry, population and functional genomics, and modelling to untangle questions related to migration, and lays a cornerstone for conserving migratory species on the basis of understanding ecological interactions and evolutionary processes in conjunction.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-021-03265-0.

1. McRae, L. et al. *Arctic Species Trend Index 2010. Tracking Trends in Arctic Wildlife* (CAFF International Secretariat, 2010).
2. Lameris, T. K. et al. Potential for an Arctic-breeding migratory bird to adjust spring migration phenology to Arctic amplification. *Glob. Change Biol.* **23**, 4058–4067 (2017).
3. Trautmann, S. in *Bird Species* (ed. Tietze, D. T.) 217–234 (Springer, 2018).
4. Zurell, D., Graham, C. H., Gallien, L., Thuiller, W. & Zimmermann, N. E. Long-distance migratory birds threatened by multiple independent risks from global change. *Nat. Clim. Change* **8**, 992–996 (2018).
5. Bay, R. A. et al. Genomic signals of selection predict climate-driven population declines in a migratory bird. *Science* **359**, 83–86 (2018).
6. White, C. M., Cade, T. J. & Enderson, J. H. *Peregrine Falcons of the World* (Lynx, 2013).
7. Clark, P. U. et al. The last glacial maximum. *Science* **325**, 710–714 (2009).
8. Otto-Bliesner, B. L., Marshall, S. J., Overpeck, J. T., Miller, G. H. & Hu, A. Simulating Arctic climate warmth and icefield retreat in the last interglaciation. *Science* **311**, 1751–1753 (2006).
9. Brambilla, M., Rubolini, D. & Guidali, F. Factors affecting breeding habitat selection in a cliff-nesting peregrine *Falco peregrinus* population. *J. Ornithol.* **147**, 428–435 (2006).
10. Hausdorff, F. Bemerkung über den Inhalt von Punktmengen. *Math. Ann.* **75**, 428–433 (1914).
11. Pulido, F. The genetics and evolution of avian migration. *Bioscience* **57**, 165–174 (2007).
12. Perdeck, A. C. An experiment on the ending of autumn migration in starlings. *Ardea* **52**, 133–139 (1964).
13. Delmore, K. E., Toews, D. P., Germain, R. R., Owens, G. L. & Irwin, D. E. The genetics of seasonal migration and plumage color. *Curr. Biol.* **26**, 2167–2173 (2016).
14. Impey, S. et al. Stimulation of cAMP response element (CRE)-mediated transcription during contextual learning. *Nat. Neurosci.* **1**, 595–601 (1998).
15. Bourtchuladze, R. et al. Deficient long-term memory in mice with a targeted mutation of the cAMP-responsive element-binding protein. *Cell* **79**, 59–68 (1994).
16. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–1218 (2013).
17. Mayr, B. & Montminy, M. Transcriptional regulation by the phosphorylation-dependent factor CREB. *Nat. Rev. Mol. Cell Biol.* **2**, 599–609 (2001).
18. Iguchi-Ariga, S. M. & Schaffner, W. CpG methylation of the cAMP-responsive enhancer/promoter sequence TGACGTCA abolishes specific factor binding as well as transcriptional activation. *Genes Dev.* **3**, 612–619 (1989).
19. Bartsch, D. et al. Aplysia CREB2 represses long-term facilitation: relief of repression converts transient facilitation into long-term functional and structural change. *Cell* **83**, 979–992 (1995).
20. Wieczorek, L. et al. Absence of Ca^{2+}-stimulated adenylyl cyclases leads to reduced synaptic plasticity and impaired experience-dependent fear memory. *Transl. Psychiatry* **2**, e126 (2012).
21. Rosenegger, D., Wright, C. & Lukowiak, K. A quantitative proteomic analysis of long-term memory. *Mol. Brain* **3**, 9 (2010).
22. Ferguson, G. D. & Storm, D. R. Why calcium-stimulated adenylyl cyclases? *Physiology (Bethesda)* **19**, 271–276 (2004).
23. Zhang, M. et al. Ca-stimulated type 8 adenylyl cyclase is required for rapid acquisition of novel spatial information and for working/episodic-like memory. *J. Neurosci.* **28**, 4736–4744 (2008).
24. Yin, J. C. & Tully, T. CREB and the formation of long-term memory. *Curr. Opin. Neurobiol.* **6**, 264–268 (1996).
25. Wauchope, H. S. et al. Rapid climate-driven loss of breeding habitat for Arctic migratory birds. *Glob. Change Biol.* **23**, 1085–1094 (2017).
26. Lok, T., Overdijk, O. & Piersma, T. The cost of migration: spoonbills suffer higher mortality during trans-Saharan spring migrations only. *Biol. Lett.* **11**, 20140944 (2015).
27. Brown, J. W. et al. Appraisal of the consequences of the DDT-induced bottleneck on the level and geographic distribution of neutral genetic variation in Canadian peregrine falcons, *Falco peregrinus*. *Mol. Ecol.* **16**, 327–343 (2007).
28. Wilcove, D. S. & Wikelski, M. Going, going, gone: is animal migration disappearing. *PLoS Biol.* **6**, e188 (2008).
29. Mueller, J. C., Pulido, F. & Kempenaers, B. Identification of a gene associated with avian migratory behaviour. *Proc. R. Soc. Lond. B* **278**, 2848–2856 (2011).
30. Peterson, M. P. et al. Variation in candidate genes *CLOCK* and *ADCYAP1* does not consistently predict differences in migratory behavior in the songbird genus *Junco*. *F1000Res.* **2**, 115 (2013).

Methods

No statistical methods were used to predetermine sample size. The experiments were randomized, and investigators were blinded to allocation during experiments and outcome assessment.

Tracking peregrine migration

We used satellite-received Argos platform transmitter terminals and GSM-received GPS transmitters to track 56 peregrines from 6 breeding regions in Arctic Eurasia (Supplementary Table 1). From this, we obtained data of at least 1 full migration for 41 peregrines: Kola Peninsula (n = 1 autumn migration and 0 spring migration), Kolguev Island (n = 8 autumn migration and 4 spring migration), Yamal Peninsula (n = 9 autumn migration and 5 spring migration), Popigai River (n = 10 autumn migration and 9 spring migration), Lena Delta (n = 6 autumn migration and 6 spring migration) and Lower Kolyma River (n = 7 autumn migration and 5 spring migration) (Supplementary Information). Permits to trap, collect blood samples and deploy satellite transmitters on peregrines were provided by the relevant authorities in Russia. All laboratory experiment procedures were under the guidance of the Ethics Committee of the Institute of Zoology, Chinese Academy of Sciences. Studies in this Article that involved peregrine brain-tissue collection and analysis were in full compliance with the Institutional Animal Care and Use Committee of the Institute of Zoology, Chinese Academy of Sciences.

For tracking data processed by the Argos system, we removed duplicate timestamps and used the Douglas Argos filter algorithm designed to retain points (Supplementary Fig. 15) that correspond to a realistic rate of movement (≤100 km h^{-1}) and that do not form tight angles between successive locations (≤15°)[31]. Migration strategy[32,33] was quantified using departure and arrival date, duration and migratory distance (migration path distance) for all individuals, where possible (Supplementary Fig. 16). We defined the start and end of migration as the day that birds moved more than 40 km from their breeding (natal) range and the day they arrived at the wintering range, respectively.

Population migration connectivity and wintering distribution pattern

To quantify the degree of migratory connectivity[34], we extracted the longitudes of breeding and wintering sites for each tracked individual and used a linear regression to explore the correlation between breeding sites and wintering sites, as used in a previous study[33]. The coefficient of determination (R^2) was used to proxy the migratory connectivity. For the spatial distribution pattern of wintering sites, we obtained the minimum convex polygon (MCP) of wintering sites and used a G function from the R package spatstat[35] to conduct a point pattern analysis. A greater empirical $\hat{G}(x)$ than the theoretical function suggests that the sites tend to be closer than expected, in contrast to a dispersed pattern. More details are provided in Supplementary Information.

Individual migration-path fidelity

We estimated the path fidelity of each bird by assessing the individual repeatability[36] of migration paths across several years ($n \geq 2$). To quantify the consistency of these migratory paths, we first calculated the track deviation from the great circle distance for each path in each year and then evaluated the repeatability of deviation after standardized measurement on the basis of latitude. Specifically, we measured the total migratory distance (sum of all distances between successive positions along a migration path). Then, the straightness was calculated as the distance between the start and end locations of the path divided by total migratory distance[37]. The repeatability of path straightness was calculated using a linear mixed model implemented in the R package rptR[38], followed by a significance testing through a permutation test.

To compare the path fidelity between short-distance and long-distance migratory groups, we considered the latitudes and longitudes of all sites along each migration path as response and independent variables, respectively, and conducted a linear regression analysis to estimate the regression coefficient α of an individual in different migration periods. We calculated the standard variation of α (α_{sd}), as a proxy of individual path fidelity (that is, a lower α_{sd} indicates that an individual uses a more similar path across migration periods). We compared differences in the mean α_{sd} estimates between short-distance and long-distance groups using a t-test.

Migratory strategy comparison

PCA was used to cluster individuals on the basis of their migratory strategy (that is, migratory distance, duration, and departure and arrival date during autumn migration). We did not use the spring departure and arrival date, because we did not obtain these data for the Kola population. In our study, only individuals that completed at least one migration route were included, and for individuals that were tracked for several years, we used mean migratory values to control pseudoreplication. We removed one of the variables if they were highly related ($|r| > 0.75$). For the comparison of migratory strategy between short-distance and long-distance groups, we used a random forest model in R and a t-test to detect the most significantly different strategy parameter. We calculated the effect size (Cohen's d[39]) for the t-test. In the comparisons, as we obtained the data of only one complete autumn migration in the Kola population (Supplementary Table 2), we also used the tracking data of three Kola peregrines from a previous study[40].

Sample collection, and genomic DNA extraction and sequencing

Blood samples were collected for the genome resequencing of 35 peregrines across the Eurasian Arctic (10 from Kola, 5 from Kolguev, 11 from Yamal and 9 from Kolyma). We also obtained nine and six feather samples from Popigai and Lena, respectively. Genomic DNA was extracted from blood and feather samples using the Blood & Cell Culture DNA Midi Kit and Blood & Tissue Extraction Kit (Qiagen), respectively. Paired-end libraries with insert size of 170 bp for blood DNA were constructed and subjected to sequencing on an Illumina HiSeq 2000 platform at BGI and Novogene. The feather DNA was used for the following PCR experiments.

Sequencing data filtering and SNP calling

An average of 68.75 Gb clean data (55.79×) were generated for 35 individuals. We used FASTQC and trimmomatic[41] to remove reads with low quality, as previously described[42]. We used the chromosome-level peregrine genome assembly[43] as the reference genome, in which the original assembly[44] was upgraded to chromosomal fragments. We then used the Burrows–Wheeler alignment[45] to map the filtered reads from each individual onto the autosomal reference genome with Z-chromosome fragments excluded (Supplementary Information). Finally, we used the pipeline in Genome Analysis Toolkit[46] (version 3.5) to call SNPs.

Population genomic analysis

Because close relatedness can bias population assignment[47], pairwise identity-by-state scores among all individuals were estimated using PLINK[48] (version 1.9) and a threshold of 0.1 was applied, resulting in the removal of three closely related individuals for the following analysis (Supplementary Fig. 17). PCA was conducted on the autosomal biallelic SNPs for the remaining 32 individuals using in-house scripts. To reconstruct the phylogenetic tree, we used PLINK to calculate genetic distances among the studied peregrines on the basis of the identified SNPs with default settings. A neighbour joining unrooted tree was then obtained using the R package phangorn with upgma function[49]. Analysis of genetic structure was implemented in frappe[50] which uses an expectation maximization algorithm. The number of genetic clusters (K) was set to range from 2 to 5.

Demographic history reconstruction

To perform sequential Markovian coalescent analysis, we used SMC++ (version 1.10.0)[51] to model historical N_e for each peregrine population, with the mutation rate and generation time derived from previous estimates[44]. To date the divergence among the peregrine populations, we developed an ABC approach (Supplementary Information). In brief, we established four candidate historical demographic models for model choice, according to our phylogenetic results. The prior N_e distributions and divergence time parameters were set from a range of 1,000 to 100,000 and 1,000 to 10,000 generations ago, respectively, according to the SMC++ results. The peregrine genome was divided into chunks with a size of 2 Mb, and genes in each chunk were counted and ranked. We conducted 100,000 simulations for each candidate model using the rapid coalescent scrm simulator[52] and summarized them using 95 different summary statistics (for example, the total number of segregating sites and summarized site frequency spectrum). A machine learning tool, ABC random forest[53], was used to conduct model choice. For the selected model with the highest approximated posterior probability (model 1), we further simulated 1,000,000 datasets for parameter inferences. Neural network methods were applied for the inference in the R package abc[54]. To evaluate the ABC performance, we applied the cross validation on model choice and parameter inference.

Ecological niche modelling

We used MaxEnt (version 3.3.3k), in the R dismo[55] package, to predict breeding and wintering distributions under present environment conditions. On the basis of satellite-tracking data, we randomly selected presence data within the MCPs of the summer (June to August) and wintering areas (December to January) of individuals. The 90% MCP was calculated for each individual using the MCP function of the adehabitatHR[56] package in R. For climate data, we downloaded 19 present and palaeo-bioclimatic variables (Supplementary Table 11) from WorldClim[57] and a previous study[58]. To reconstruct breeding and wintering distributions in the past, we projected the ecological niche models built under present climate to palaeoclimates during the middle of the Holocene (about 6 ka), LGM (22 ka) and last interglacial period (120–140 ka). The correlation between N_e estimates and predicted areas were fitted using a linear regression analysis at five periods (100, 40, 20, 15, 6 ka).

Palaeo-vegetation data analysis

To examine whether the predicted LGM breeding areas mostly consisted of tundra biome, we obtained Eurasian palaeo-pollen data from a previous study[59]. Palaeo-pollen data classified as tundra biome by the biome_2000 model were extracted for the LGM, and then mapped to our predicted breeding areas to estimate the overlapping extent.

Quantification of inter-route distances

Hausdorff distances were used to quantify the dissimilarity between migratory paths of individual pairs of peregrines. The approach measures how far apart two subsets of a space metric are from each other[10]. In our study, a migratory path was treated as positional distribution of bird movement points in time and space, and the distance between the migration path *A* and *B* was obtained using hausdorff_dist function in R package pracma[60] (Supplementary Information). The mean Hd estimates were compared within and between neighbouring populations using a *t*-test. The generated Hd matrix was then used for the route clustering analysis using a hclust method in the R package heatmap. The inter-route distance was finally calculated as the mean Hd between pairs of individual paths from different migration routes.

Maintenance mechanisms of present migration routes

To investigate maintenance mechanisms of present migration routes, we first checked the influence of neutral genetic and selective genetic distance on the route distances estimated as in 'Quantification of inter-route distances'. For the estimation of genetic distance, we randomly selected 93 putatively neutrally evolving and 75 selected SNP loci on the basis of selection analysis of 32 resequenced peregrine genomes described in 'Identification of selective sweeps and detection of selected SNPs between short-distance and long-distance peregrines'. Then, we genotyped nine and six individuals (shed feathers) sampled in Popigai and Lena, respectively. The detailed PCR amplifications and sequencing of these feather DNA extracts are described in Supplementary Information. With the combined genotypes from blood and feather samples, we calculated the F_{ST} for each locus among five populations (Kolguev, Yamal, Popigai, Lena and Kolyma) using vcftools[61]. The genetic distance was then calculated as $F_{ST}/(1-F_{ST})$[5]. We fitted the relationship between migration route separation (Hd) and neutral or selective genetic distance, using a linear regression model.

To investigate the environmental divergence among migration routes, we randomly sampled 200 grids (0.083° × 0.083°) from the 90% MCP of each route and extracted the variable of climate zones (referring to the Köppen–Geiger classification system[62]) from each grid. χ^2 tests were applied for testing differences in climate zones between adjacent migration routes (Extended Data Fig. 5b). To check the environmental boundaries, we further divided the Eurasian continent into geographical bands (width of 2° longitude, on the basis of the estimated migration distance per day in the studied peregrines) with direction parallel to the mean migration angle of all individuals and calculated the median value of climate zones of the grids (0.083° × 0.083°) at regular intervals (1° in latitude) from neighbouring bands for paired comparisons (Extended Data Fig. 5c). Paired *t*-tests were used to check the abrupt change of climate zone (boundary) (Extended Data Fig. 5d) between adjacent bands. During comparisons, we used the same latitude range between the pairwise routes.

To test whether there is more benefit from being a conventional migrator (migrating within a population route) or an unconventional migrator (migrating across routes), we first proved that the tracked peregrines migrated in a least-cost manner (Supplementary Fig. 18). We then simulated scenarios in which peregrines depart from their actual breeding sites and fly along least-cost paths, but winter in the actual wintering sites of neighbouring routes (Extended Data Fig. 5e). Taking account of migration path length, we estimated the relative least cost of cross-route migration (Supplementary Information). A *t*-test was used to compare the difference in the relative least cost between the actual within-route and simulated cross-route migration. For the *t*-test, the effect size *d* was calculated.

Identification of selective sweeps and detection of selected SNPs between short-distance and long-distance peregrines

We used two methods (a window-based F_{ST} and XP-EHH[63]) to identify selective sweeps between the short-distance and long-distance groups. The F_{ST} of each locus was calculated using vcftools. A smoothed spline technique in R package GenWin[64] (version 1.0.1) was implemented to determine the window boundary, and *w*-statistic was used as a proxy of windowed F_{ST}. The XP-EHH value of each locus was calculated using selscan[65] (version 1.1) with BEAGLE-phased[66] SNPs (*n* iterations = 100). Outlier regions (top 1%) detected by both methods were considered as selective sweeps. For the sweeps identified on the focal gene, we calculated the nucleotide diversity (θ_π) of each locus using vcftools and narrowed the sweep down to a specific region. In addition, to verify the phased haplotype of the *ADCY8* gene, we chose three individuals from each population for 10x genomics sequencing, followed by linked-reads phasing using the Long Ranger[67] (version 2.2.2) (Supplementary Table 12).

We then used an integrative method[68] to detect the selected SNPs between the pairwise groups (Kolguev versus long-distance and Yamal versus Kolyma), which integrated the results obtained from three selection tests: the FLK[69] based on comparing different patterns of allele

frequencies among populations to the values expected under a scenario of neutral evolution[70], latent factor mixed models for which the environment is used as a fixed effect and latent factors used to infer environmental associations[71], and pcadapt (version 3.03) analysis based on a Bayesian factor model. Detailed settings have previously been described[72]. The adjusted *z*-scores were calculated for each of three above tests, and the calibrated *P* values were obtained as previously reported[71]. The candidate SNP loci were ultimately determined using Benjamini–Hochberg false discovery rate (FDR) control. The level of FDR was set to 0.05.

Luciferase reporter assay for the focal *ADCY8* haplotype

To investigate the potential functional importance of different *ADCY8* haplotypes, a randomly selected haplotype 1 in short-distance peregrines and the dominant haplotype (haplotype 2) in long-distance peregrines (Fig. 2l) were fully synthesized by SinoGenoMax and inserted into the pGL3-promoter backbone using KpnI/XhoI restriction sites. Positive (pGL3-promoter) and negative (pGL3-basic) controls were also constructed. The activity of haplotype 1 or haplotype 2 was examined in primary cells cultured from chicken embryonic hippocampus tissues. We isolated and cultured neurons from fertilized chicken eggs (Boehringer Ingelheim) (Supplementary Information) for lipofection and luciferase reporter assays, by referring to previous studies[73]. pGL3-haplotype 1 or pGL3-haplotype 2 were cotransfected with pRL-TK into the neurons using Lipofectamine 2000 (Invitrogen). After a 48-h incubation, the dual-luciferase activity was measured using Dual-Glo Luciferase Assay kit (Promega). A *t*-test was used to compare the fluorescence intensity among experimental groups. At least three independent experiments of each assay were performed, with a minimum of six replicates. To predict the motif in this focal region, we extracted the sequences (10 bp) around each SNP and searched the sequences against the TFBS motif database[74].

ATAC-seq and RNA-sequencing analysis

Hippocampus and cortex tissues were collected from a peregrine that died of natural causes in the Chongqing Zoo. For the ATAC experiment (Supplementary Information), the samples were prepared according to the manual[75] in Shanghai Jiayin Biotechnology, followed by library construction, and subjected to sequencing on an Illumina NovaSeq 6000 (Novogene). The raw reads generated were further quality-controlled. The clean reads were mapped to a reference peregrine genome, and sequencing depth around the target region was evaluated.

For the RNA-sequencing (RNA-seq), brain tissues were collected from two humanely euthanized peregrines from the Beijing Raptor Rescue Center. Total RNA was extracted from the samples of two peregrines (chromosome 3-5170169 SNP genotype CC and CT) using the TRIzol reagent according to the user guide (Invitrogen). For each brain RNA sample, library with insert size of 350 bp was constructed and then sequenced on a HiSeq 2500 (Novogene). Reads per kilobases per million reads of *ADCY8* with different genotypes were calculated by mapping RNA-seq reads to the peregrine gene set[44] using SOAP[76] (version 2.22). Expression difference between these genotypes was compared using a hypergeometric test.

Detailed descriptions of the ATAC-seq are provided in Supplementary Information.

Global warming effects and prediction of migratory distance in the future

On the basis of the inferred present peregrine distribution, we predicted future (2070) potential breeding and wintering distributions under RCP 8.5, a scenario in which emissions continue to rise through the twenty-first century. Occurrence probabilities were transformed into binary maps using true skill statistic-maximizing values as thresholds. Differences between present and future distributions were investigated using two parameters: area change and latitude shift. We counted the grids (0.083° × 0.083°) within the nonoverlapping regions between present and future, and the shifted latitude was represented as the ratio of area and longitude (per degree). To compare the migratory distance between the present and future, we randomly selected 200 sites per population within breeding or wintering ranges and calculated great circle distances between corresponding sites. The confidential interval of these distance estimates were calculated as the minimum and maximum of great circle distances among all the sites in each of breeding and wintering ranges. For comparison, we conducted the same analyses under RCP 4.5, a mitigation scenario in which emissions peak around 2040 (Supplementary Information).

Prediction of changes in N_e in the future

To study how N_e will change under future climate change, we initially investigated the association between the most recent changes of N_e and climate variables since the industrial revolution (1840). We used a linkage-disequilibrium-based method, SNeP[77] (version 1.11), to reconstruct recent N_e changes in generations for each population. For this analysis, we randomly selected 5,000 loci per chromosome and used a N_eS to investigate the rate of N_e changes[78]. SNeP could reliably examine the changes or trends of N_e, rather than the actual N_e(ref. [77]).

For the climate data, the monthly temperature of the Community Climate System Model (version 4) output was downloaded from Coupled Model Intercomparison Project 5. The yearly mean temperature was calculated as the average of monthly temperature in the breeding season (May to July), a period that is vital to the breeding success of peregrines[79,80]. To quantify historical weather extremes, we downloaded the gridded daily minimum and maximum temperature data from Berkeley Earth Surface Temperature. Days that exceeded the 95% upper threshold of average temperature were considered as extreme hot days, and days below the 5% lower threshold were considered as extreme cold days. The simulated future daily temperature (2020 to 2100) was downloaded from the NASA Earth Exchange Global Daily Downscaled Projections. The inferring method for future weather extremes was the same as historical extremes.

We constructed a generalized linear model to model the relationship between N_eS and changes of climate variables (Supplementary Information), and finally predicted the future N_eS under future climate using the fitted generalized linear model.

Statistical analysis

All reported *P* values were from Student's *t*-tests (two-sided) unless otherwise specified. All assays were performed in at least three independent experiments, with a minimum of six replicates. In the analysis of individual path repeatability, *P* values were calculated using a permutation test. The comparisons of climate zones between migration routes were calculated using a χ^2 test. The *P* values of band comparisons of climate zones were calculated using a paired *t*-test with latitude differences controlled. For the *t*-test, Cohen's *d* is determined by calculating the mean difference between two groups, and then dividing the result by the pooled s.d. The expression difference between different alleles of the focal *ADCY8* SNP was compared using a hypergeometric test.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

All of the sequenced genome data have been deposited in the GenBank under accession number PRJNA686418. The tracking data are included in the Arctic Animal Movement Archive and in Movebank under the identifiers 103426553 and 934079034. Source data are provided with this paper.

31. Douglas, D. C. et al. Moderating Argos location errors in animal tracking data. *Methods Ecol. Evol.* **3**, 999–1007 (2012).
32. Mueller, T., O'Hara, R. B., Converse, S. J., Urbanek, R. P. & Fagan, W. F. Social learning of migratory performance. *Science* **341**, 999–1002 (2013).
33. Trierweiler, C. et al. Migratory connectivity and population-specific migration routes in a long-distance migratory bird. *Proc. R. Soc. Lond. B* **281**, 20132897 (2014).
34. Ambrosini, R., Møller, A. P. & Saino, N. A quantitative measure of migratory connectivity. *J. Theor. Biol.* **257**, 203–211 (2009).
35. Baddeley, A., Rubak, E. & Turner, R. *Spatial Point Patterns: Methodology and Applications with R* (Chapman and Hall/CRC, 2015).
36. López-López, D. P., García-Ripollés, C. & Urios, V. Individual repeatability in timing and spatial flexibility of migration routes of trans-Saharan migratory raptors. *Curr. Zool.* **60**, 642–652 (2014).
37. Benhamou, S. How to reliably estimate the tortuosity of an animal's path: straightness, sinuosity, or fractal dimension? *J. Theor. Biol.* **229**, 209–220 (2004).
38. Stoffel, M. A., Nakagawa, S. & Schielzeth, H. rptR: repeatability estimation and variance decomposition by generalized linear mixed-effects models. *Methods Ecol. Evol.* **8**, 1639–1644 (2017).
39. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences* (Routledge Academic, 1988).
40. Ganusevich, S. A. et al. Autumn migration and wintering areas of peregrine falcons *Falco peregrinus* nesting on the Kola Peninsula, northern Russia. *Ibis* **146**, 291–297 (2004).
41. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
42. Zhao, S. et al. Whole-genome sequencing of giant pandas provides insights into demographic history and local adaptation. *Nat. Genet.* **45**, 67–71 (2013).
43. Damas, J. et al. Upgrading short-read animal genome assemblies to chromosome level using comparative genomics and a universal probe set. *Genome Res.* **27**, 875–884 (2017).
44. Zhan, X. et al. Peregrine and saker falcon genome sequences provide insights into evolution of a predatory lifestyle. *Nat. Genet.* **45**, 563–566 (2013).
45. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
46. DePristo, M. A. et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498 (2011).
47. Rodríguez-Ramilo, S. T. & Wang, J. The effect of close relatives on unsupervised Bayesian clustering algorithms in population genetic structure analysis. *Mol. Ecol. Resour.* **12**, 873–884 (2012).
48. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
49. Schliep, K. P. phangorn: phylogenetic analysis in R. *Bioinformatics* **27**, 592–593 (2011).
50. Tang, H. et al. Genetic structure, self-identified race/ethnicity, and confounding in case–control association studies. *Am. J. Hum. Genet.* **76**, 268–275 (2005).
51. Terhorst, J., Kamm, J. A. & Song, Y. S. Robust and scalable inference of population history from hundreds of unphased whole genomes. *Nat. Genet.* **49**, 303–309 (2017).
52. Staab, P. R., Zhu, S., Metzler, D. & Lunter, G. scrm: efficiently simulating long sequences using the approximated coalescent with recombination. *Bioinformatics* **31**, 1680–1682 (2015).
53. Pudlo, P. et al. Reliable ABC model choice via random forests. *Bioinformatics* **32**, 859–866 (2016).
54. Csilléry, K., François, O. & Blum, M. G. abc: an R package for approximate Bayesian computation (ABC). *Methods Ecol. Evol.* **3**, 475–479 (2012).
55. Hijmans, R. J., Phillips, S., Leathwick, J. & Elith, J. dismo: species distribution modeling. R package version 1.3-3 https://cran.r-project.org/package=dismo (2020).
56. Calenge, C. adhabitatHR: home range estimation. R package version 0.4.19 https://cran.r-project.org/package=adehabitatHR (2021).
57. Fick, S. E. & Hijmans, R. J. WorldClim2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **37**, 4302–4315 (2017).
58. Beyer, R. M., Krapp, M. & Manica, A. High-resolution terrestrial climate, bioclimate and vegetation for the last 120,000 years. *Sci. Data* **7**, 236 (2020).
59. Cao, X., Tian, F., Dallmeyer, A. & Herzschuh, U. Northern Hemisphere biome changes (> 30° N) since 40 cal ka BP and their driving factors inferred from model-data comparisons. *Quat. Sci. Rev.* **220**, 291–309 (2019).
60. Borchers, H. W. pracma: practical numerical math functions. R package version 2.3.3 https://cran.r-project.org/package=pracma (2021).
61. Danecek, P. et al. The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
62. Beck, H. E. et al. Present and future Köppen–Geiger climate classification maps at 1-km resolution. *Sci. Data* **5**, 180214 (2018).
63. Sabeti, P. C. et al. Genome-wide detection and characterization of positive selection in human populations. *Nature* **449**, 913–918 (2007).
64. Beissinger, T. M., Rosa, G. J., Kaeppler, S. M., Gianola, D. & de Leon, N. Defining window-boundaries for genomic analyses using smoothing spline techniques. *Genet. Sel. Evol.* **47**, 30 (2015).
65. Szpiech, Z. A. & Hernandez, R. D. selscan: an efficient multithreaded program to perform EHH-based scans for positive selection. *Mol. Biol. Evol.* **31**, 2824–2827 (2014).
66. Browning, S. R. & Browning, B. L. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am. J. Hum. Genet.* **81**, 1084–1097 (2007).
67. Zheng, G. X. et al. Haplotyping germline and cancer genomes with high-throughput linked-read sequencing. *Nat. Biotechnol.* **34**, 303–311 (2016).
68. François, O., Martins, H., Caye, K. & Schoville, S. D. Controlling false discoveries in genome scans for selection. *Mol. Ecol.* **25**, 454–469 (2016).
69. Fariello, M. I., Boitard, S., Naya, H., SanCristobal, M. & Servin, B. Detecting signatures of selection through haplotype differentiation among hierarchically structured populations. *Genetics* **193**, 929–941 (2013).
70. Bonhomme, M. et al. Detecting selection in population trees: the Lewontin and Krakauer test extended. *Genetics* **186**, 241–262 (2010).
71. Frichot, E. & François, O. LEA: an R package for landscape and ecological association studies. *Methods Ecol. Evol.* **6**, 925–929 (2015).
72. Pan, S. et al. Population transcriptomes reveal synergistic responses of DNA polymorphism and RNA expression to extreme environments on the Qinghai–Tibetan Plateau in a predatory bird. *Mol. Ecol.* **26**, 2993–3010 (2017).
73. Zhang, Y. et al. An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *J. Neurosci.* **34**, 11929–11947 (2014).
74. Yang, L. et al. TFBSshape: a motif database for DNA shape features of transcription factor binding sites. *Nucleic Acids Res.* **42**, D148–D155 (2014).
75. Buenrostro, J. D., Wu, B., Chang, H. Y. & Greenleaf, W. J. ATAC-seq: a method for assaying chromatin accessibility genome-wide. *Curr. Protoc. Mol. Biol.* **109**, 21–29 (2015).
76. Li, R. et al. De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res.* **20**, 265–272 (2010).
77. Barbato, M., Orozco-terWengel, P., Tapio, M. & Bruford, M. W. SNeP: a tool to estimate trends in recent effective population size trajectories using genome-wide SNP data. *Front. Genet.* **6**, 109 (2015).
78. Pitt, D. et al. Demography and rapid local adaptation shape Creole cattle genome diversity in the tropics. *Evol. Appl.* **12**, 105–122 (2019).
79. Carlzon, L., Karlsson, A., Falk, K., Liess, A. & Møller, S. Extreme weather affects peregrine falcon (*Falco peregrinus tundrius*) breeding success in South Greenland. *Ornis Hungarica* **26**, 38–50 (2018).
80. Franke, A. et al. Status and trends of circumpolar peregrine falcon and gyrfalcon populations. *Ambio* **49**, 762–783 (2020).

Acknowledgements This study was supported by National Natural Science Foundation of China (31821001, 31930013, 31911530186 and 91740201), Strategic Priority Program of Chinese Academy of Sciences (XDB31000000), the National Key Program of Research and Development, Ministry of Science and Technology (2016YFC0503200), Youth Innovation Promotion Association of Chinese Academy of Sciences (2020086) to S.P., Second Tibetan Plateau Scientific Expedition and Research Program (STEP) (2019QZKK0501), Water Ecological Security Assessment, the Major Research Strategy for Middle and Lower Yangtze River (ZDRW-ZS-2017-3), Biodiversity Survey, Monitoring and Assessment Project (2019–2023) of Ministry of Ecology and Environment, China, the Royal Society to X.Z. and M.W.B., and the CAS President's International Fellowship Initiative for Visiting Scientists to M.W.B. Funding for satellite tracking, sampling and partial genome resequencing was provided by the Environment Agency-Abu Dhabi. Additional tracking and sampling from Kolguev and Kola was funded by the Max Planck Institute of Animal Behavior to I.P. and Earthspan (www.earthspan.foundation). We thank J. M. Graves for her suggestions; M. Al Bowardi and M. Al Mansouri for their support; X. Li, Q. Dai, X. Liu, Y. Chen, Y. Lin, J. Jiao, G. Wang, X. Guang, W. He and M. Barbato for their help with data analysis; W. Wu, X. Hou, Y. Wang and Y. Zhan for their advice on figure drawing; and O. Kulikova, V. Pozdnyakov, S. Troyev, L. Zhou, Y. Shang, C. Dai, Y. Yin, C. Li and N. C. Fox for assistance during fieldwork.

Author contributions X.Z. led the project. X.Z. and A.D. conceived and designed the study. A.D., S.G., V.S., A.S., I.P., J.L. and Z.L. conducted the fieldwork and sample collection. X.Z. and A.D. examined migration paths, migration connectivity and genetic structure of peregrines across Eurasia. X.Z. and M.W.B. supervised the population genomic research. Z.G., S.P., L.H., J.C. and X.D. performed the data analyses. Z.L., Y.X., M.S., H.S. and F.J. conducted the molecular experiments. X.Z. and Z.G. wrote the manuscript, with contributions from M.W.B., S.K. and A.D.

Competing interests The authors declare no competing interests.

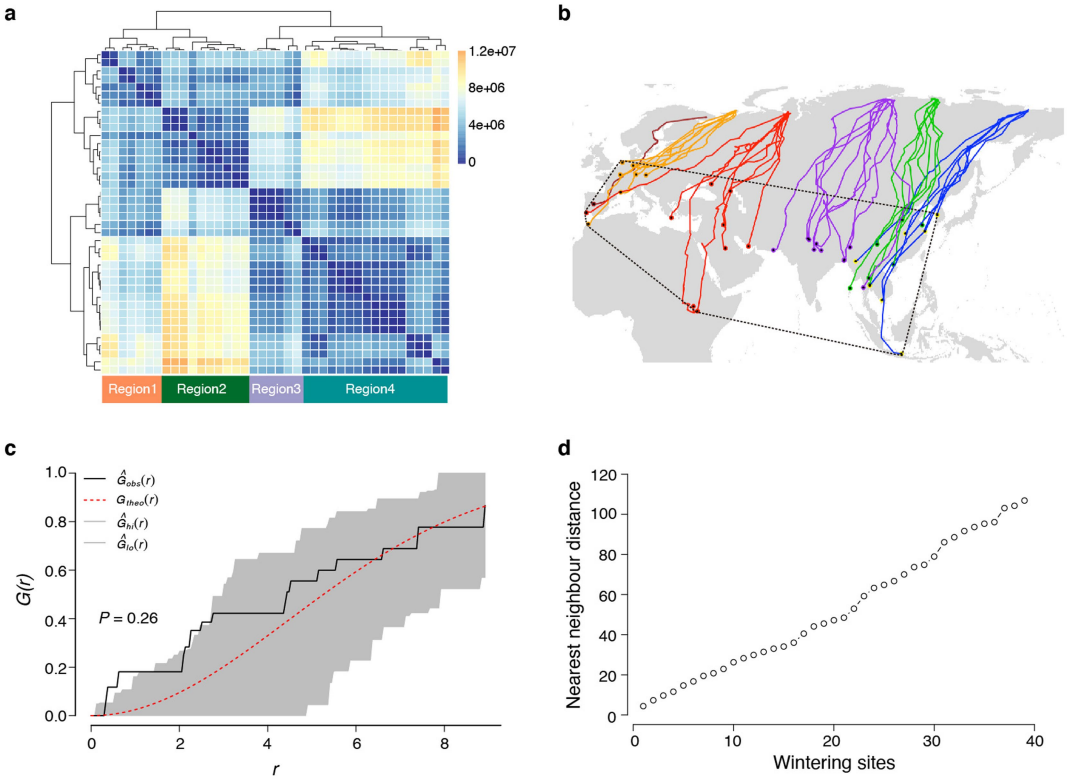
Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-021-03265-0.

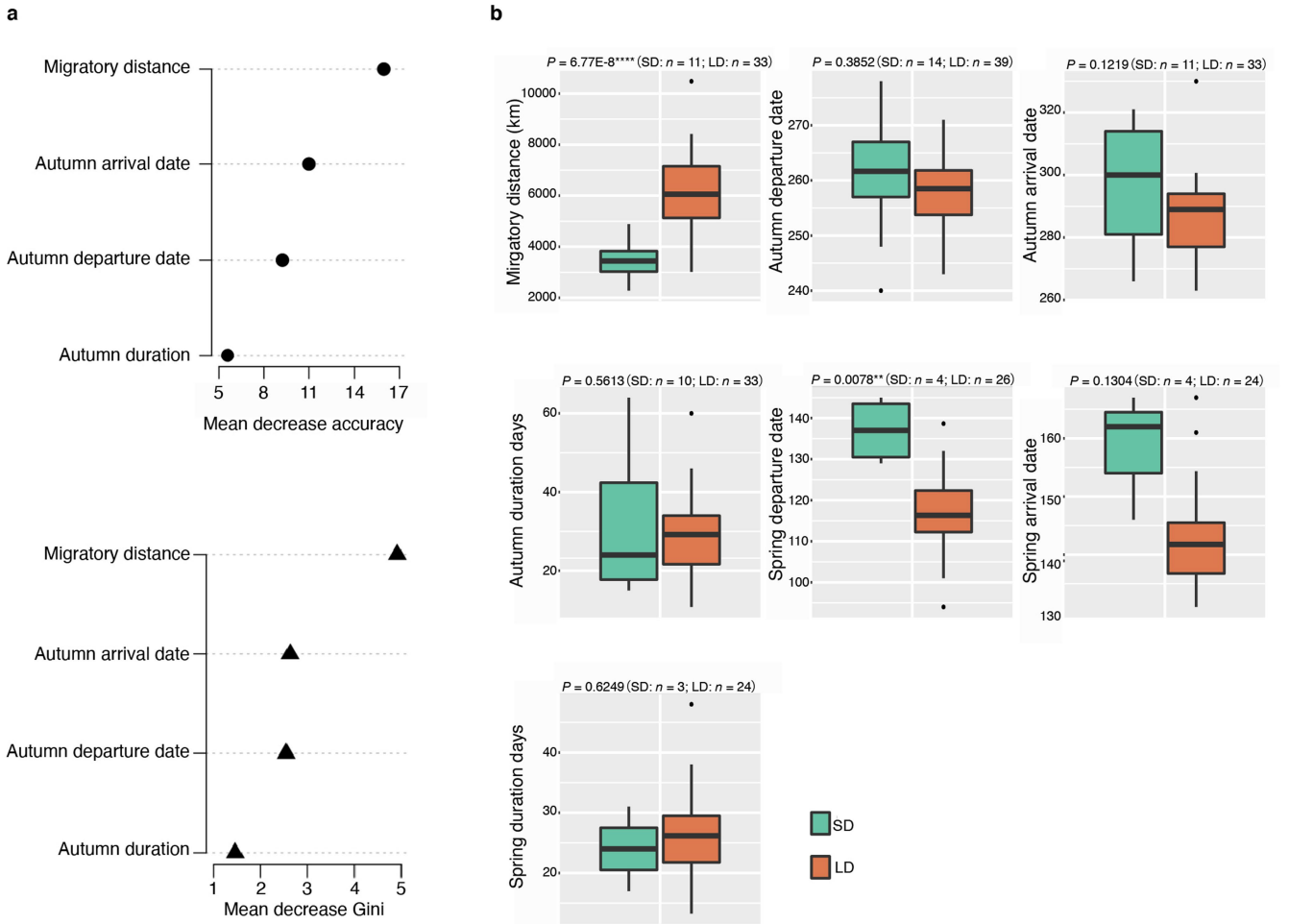
Correspondence and requests for materials should be addressed to X.Z.

Peer review information *Nature* thanks Simon G. Sprecher and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

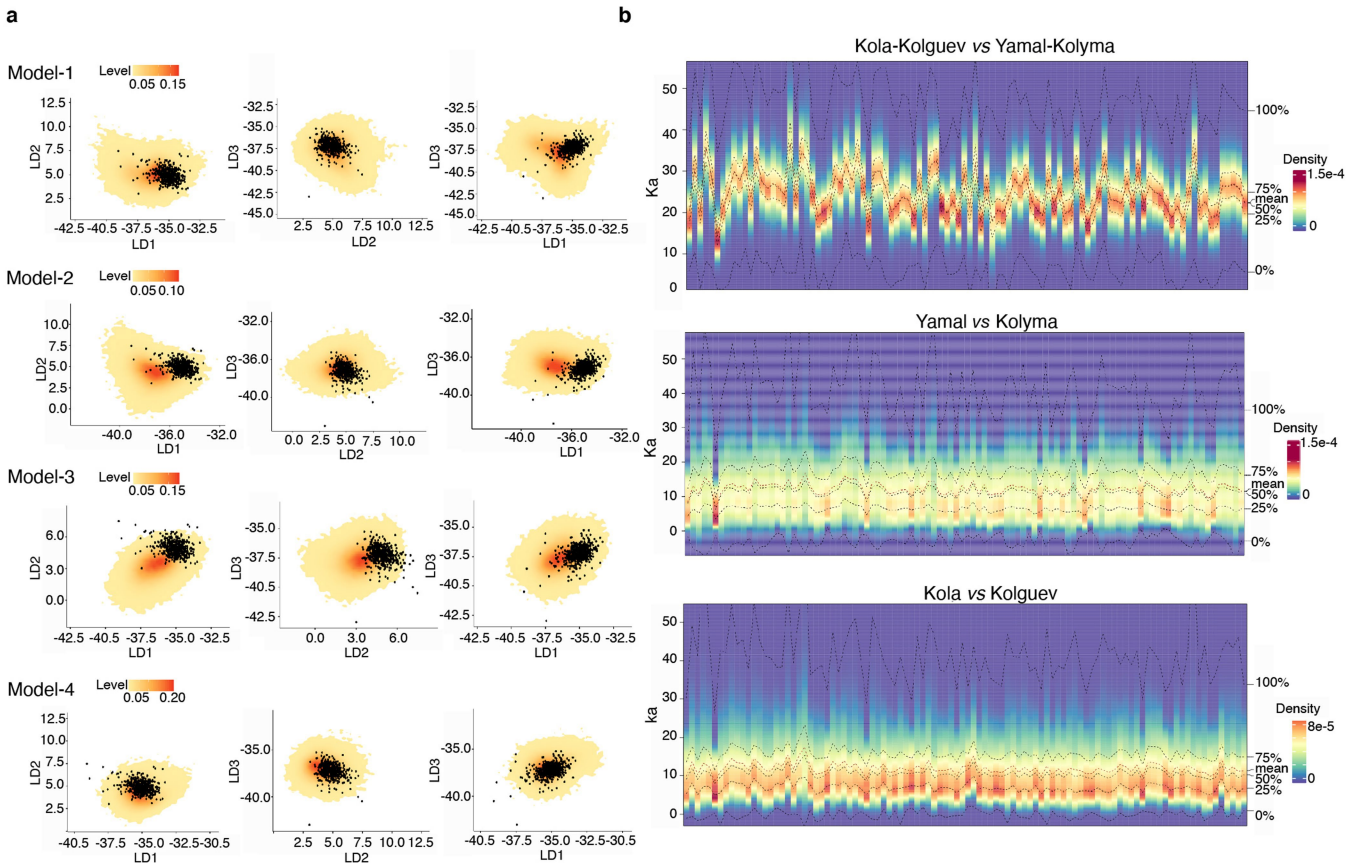
Extended Data Fig. 1 | Sampling sites for tracking peregrines in the Arctic. The sample size, visit years for each place and the peregrines equipped with Argos satellite transmitters are shown.



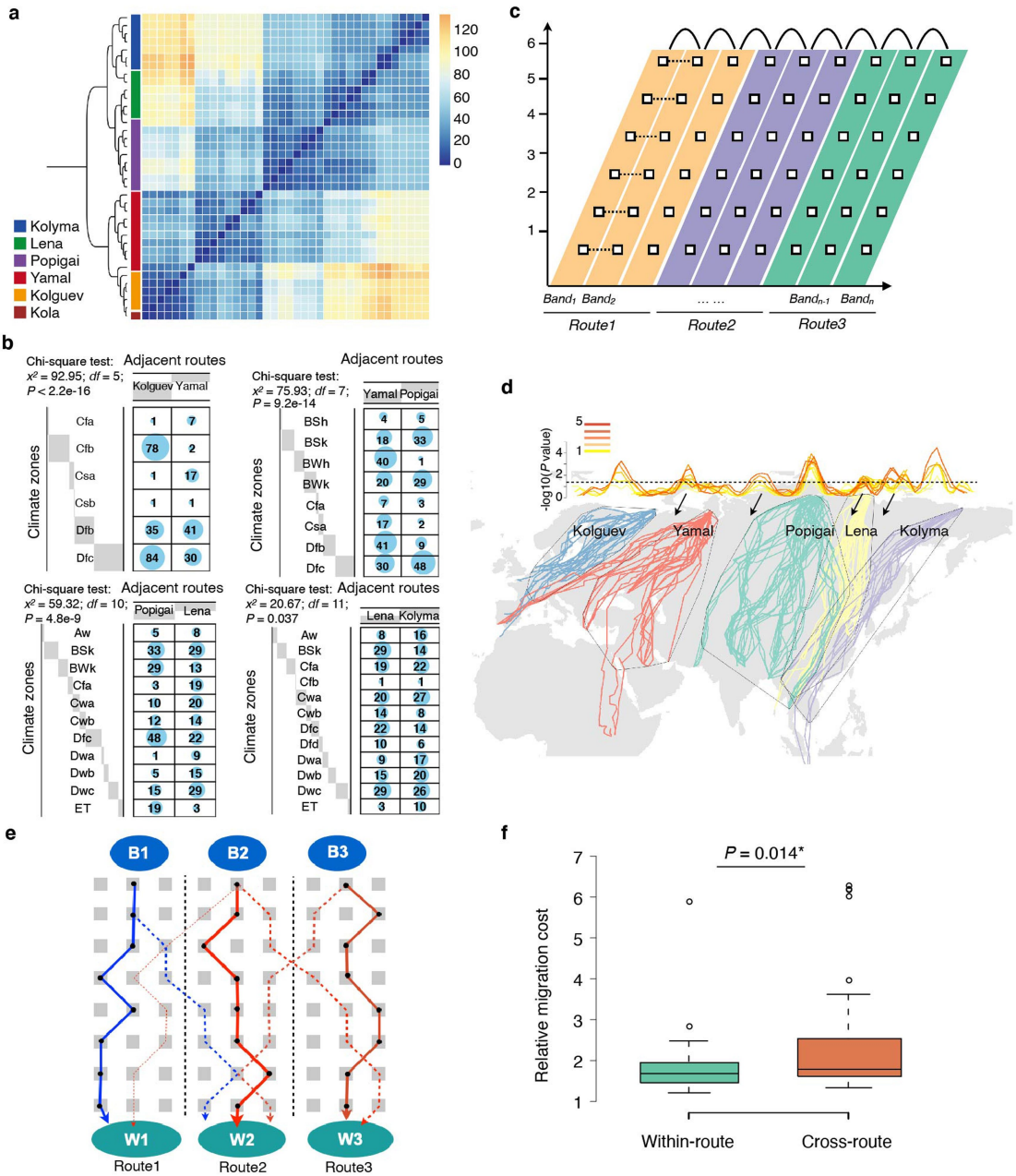
Extended Data Fig. 2 | The broad-front migration pattern of peregrines. **a**, Four main wintering regions identified in the cluster analysis. **b**, Migration paths with the centroids of breeding and wintering MCP for each bird, and the MCP of wintering ranges for all birds (dashed line), are shown. **c**, *G* function results in the point pattern analysis, showing a broad-front wintering distribution. The solid and dashed line denote the observed and theoretical value of *G*, respectively. The 95% confidence interval of theoretical *G* value is shadowed. The *P* value was calculated for the statistic of maximum absolute deviation using Monte Carlo simulations ($n = 100$). **d**, The distance from each winter centroid to its nearest neighbour centroid (nearest neighbour distance) is shown ($n = 40$).



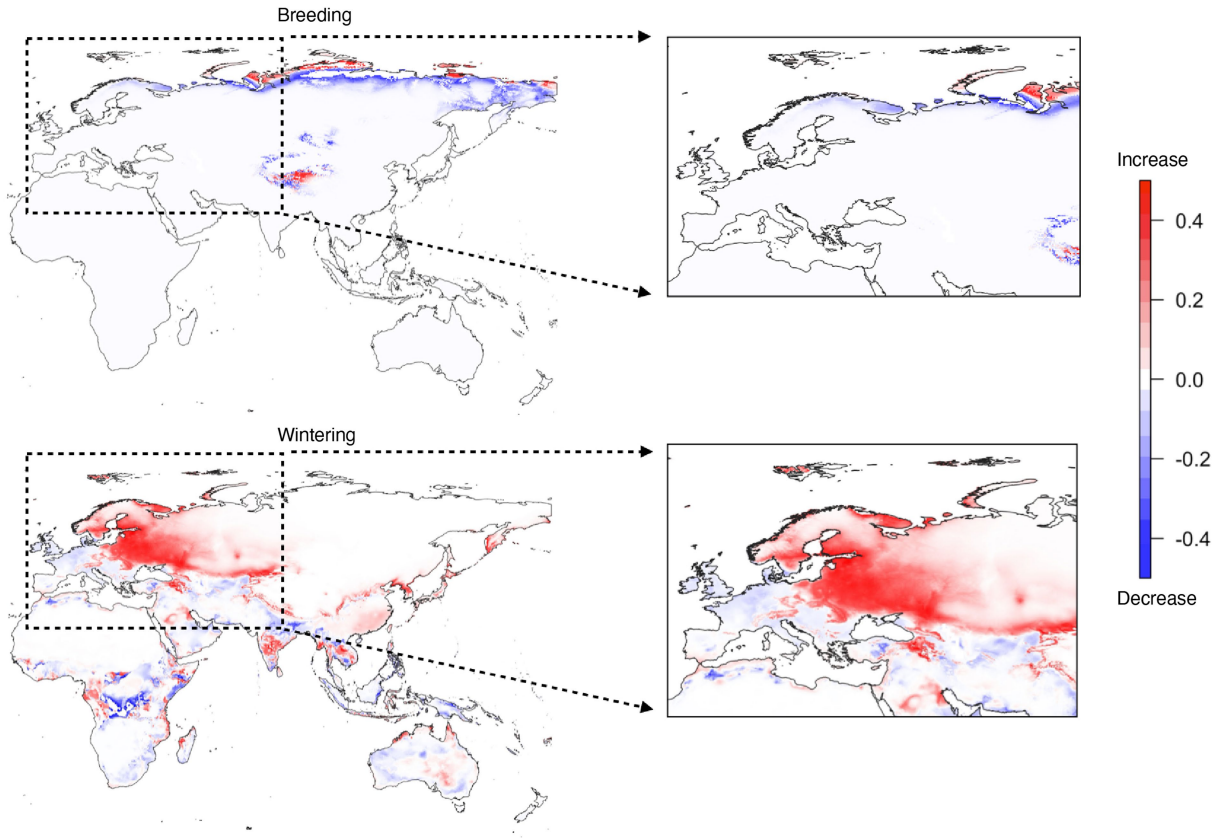
Extended Data Fig. 3 | Comparison of migration strategy between short-distance and long-distance groups. a, Variable importance estimated by random forest modelling. **b**, Comparisons of migratory strategy between the short-distance (SD) and long-distance (LD) groups. Significance was determined by a two-sided *t*-test. Sample size (*n*) for each comparison is shown. In the box plots, the centre line represents the median, whiskers represent maximum and minimum values, and box boundaries represent 75th and 25th percentiles.



Extended Data Fig. 4 | ABC simulation and parameter inference. a, Linear discriminant summary statistics values of the simulated datasets and the observations given the four ABC candidate models. On the basis of the three statistics (LD1, LD2 and LD3), model 1 is best-supported, because the targets (dark) fit simulated data (shadow) well. **b**, Distribution of divergence times estimated using the chunks supporting model 1. One column represents one chunk; only 100 chunks are shown. The density bar denotes the posterior distribution of inferred divergence time in each chunk.



Extended Data Fig. 5 | Maintenance mechanisms of present migration routes. a, Route cluster analysis on the basis of Hd. **b**, χ^2 testing results of climate zones between adjacent migration routes at the whole-route level. **c**, Schematic of environment comparisons between neighbouring geographical bands. Each route was divided into geographical bands parallel to the main migration direction. Grids at regular intervals were chosen from neighbouring bands for comparison. **d**, Environmental boundaries coinciding with migration route boundaries. The Eurasian continent was divided into geographical bands (at 2° longitude). The *P* values of paired *t*-tests between compared bands are shown, and the dashed line equals 0.05 (top). The bar is scaled to the number of spaces between two targeted bands in a paired comparison. The MCPs (90%) of five migration routes are shaded (bottom). Arrows point to the coincidence between environmental and migration route boundaries. Distinct environment difference within the Popigai route may result from the inclusion of large 'barrier islands' of unsuitable region in the comparison. **e**, Illustration of the model simulating the least-cost migration path. For a typical migration route, we simulated the potential migration path (dashed lines) along which a peregrine departs from its actual breeding site (for example, B1 in route 1) and flies along a least-cost path, but then winters in a wintering site of the neighbouring route (for example, W2 in route 2). B1, B2 and B3 denotes breeding areas; W1, W2 and W3 denote wintering areas. Solid lines are the actual tracked migration path. **f**, Comparison of migration costs between within-route and across-route paths ($P = 0.01$, $t = -2.58$, degrees of freedom = 101.68). Significance was calculated using a two-sided *t*-test ($n = 45$ and 64 for within- and cross-route, respectively). In the box plots, the centre line represents the median, whiskers represent maximum and minimum values, and box boundaries represent 75th and 25th percentiles.



Extended Data Fig. 6 | Differences in breeding and wintering areas between present and future (2070). Predicted changes in breeding (top) and wintering (bottom) area under the RCP 8.5 scenario (left), and zoomed-in Kola and Europe (right).

nature research

Corresponding author(s): Xiangjiang Zhan

Last updated by author(s): Dec 20, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Tracking data of peregrine falcons were generated by our team. Samples for genome sequencing, ATAC-seq and RNA-seq data were prepared in our lab and the sequencing were outsourced to specialized companies. No software was used for data collection.
Data analysis	spatstat package (version 1.58-2), rptR package (version 0.9.21), trimmomatic (version 0.27), Genome Analysis Toolkit (version 3.5), PLINK (version 1.9), phangorn package (2.4.0), frappe (version 1.1), SMC++ (version 1.10.0), BEAGLE (version 4.0), MaxEnt (version 3.3.3k), adehabitatHR package (version 0.3.12), pracma package (version 2.2.2), vcftools (version 0.1.13), GenWin package (version 1.0.1), selscan (version 1.1), SOAP (version 2.22), SNeP (version 1.11), abctools package (version 1.1.3), abc package (version 2.1), scrm simulator (version 1.7.4), Long Ranger (version 2.2.2), hapFLK (version 1.3.0), LFMM (version 1.0), pcadapt (version 3.03), dismo package (version 1.1-4). In ABC analysis, the 95 summary statistics were calculated by using our custom R scripts, which can be found in Github (https://github.com/guzhongru/ABC).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All of the sequenced genome data have been deposited in the GenBank under accession number PRJNA686418 (https://www.ncbi.nlm.nih.gov/bioproject/686418). The tracking data are included in the Arctic Animal Movement Archive and in Movebank under ID numbers 103426553 (https://www.movebank.org/cms/webapp?gwt_fragment=page=studies,path=study103426553) and 934079034 (https://www.movebank.org/cms/webapp?gwt_fragment=page=studies,path=study934079034), which were accessible with written permission. Climate data were downloaded from WorldClim (https://www.worldclim.org), Berkeley Earth Surface Temperature (http://berkeleyearth.org), NASA Earth Exchange Global Daily Downscaled Projections (https://developers.google.com/earth-engine/datasets/catalog/NASA_NEX-GDDP).

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Statistical analyses on the tracking data and genomes of Arctic peregrine falcons to understand the formation, maintenance and future of peregrine migration routes and genetic determinants of migratory distance.
Research sample	56 wild peregrine falcons from the Kola Peninsula (n=7), Kolguev island (n=10), Yamal Peninsula (n=12), Eastern Taimyr (Popigai; n=10), Lena Delta (n=8) and Lower Kolyma River (n=9) were studied. We tracked the falcons using satellite-received Argos Platform Transmitter Terminals (18g Solar PTT100, n = 33; 12g Solar PTT100, n = 10; 22g Solar Argos/GPS PTT n = 3, Microwave Telemetry Inc) and GSM-received transmitters at Kolguev Island (15g UKn Solar/GPS-GSM, n = 10, University of Konstanz). Sequencing and feather samples were also collected from the wild. Sampling was conducted across six regions of the Eurasian Arctic in order to cover continent scale migration pathways and connectivity of the study populations. Furthermore, sampling sites were selected to include the three named subspecies of peregrines that breed in Arctic Eurasia (F. p. peregrinus, F. p. calidus and F. p. hartertii/japonensis). The peregrine falcon Falco peregrinus is an avian Arctic migrant belonging to the family Falconidae. Most individuals fitted with satellite transmitters were adults, chosen over juveniles because of their higher survival rates in order to maximize the amount of tracking data received (41 adults and 15 juveniles). Most adults were females due to the method used to trap birds at the breeding sites during incubation, as most incubation is undertaken by females. Kola Peninsula (n = 7 juveniles: 2 males, 5 females) Kolguev island (n = 4 adult females; n = 6 juveniles: 2 males, 4 females) Yamal Peninsula (n = 10 adults: 1 male, 9 females; n = 2 juvenile males) Eastern Taimyr (Popigai; n = 10 adults: 1 male, 9 females) Lena Delta (n = 8 adult females) Lower Kolyma River (n = 9 adult females) No manipulations were undertaken Blood samples were collected by brachial venipuncture for birds at Kola Peninsula, Kolguev Island, Yamal Peninsula and Lower Kolyma River. Breast feather samples were plucked from adult birds trapped at Eastern Taimyr (Popigai) and Lena Delta.
Sampling strategy	Tracking coordinates and times were obtained for each studied peregrine. Blood samples were collected for the genome resequencing of 35 peregrines across the Eurasian Arctic range (10 from Kola, 5 from Kolguev, 11 from Yamal and 9 from Kolyma). We also obtained nine and six feather samples from Popigai and Lena, respectively. Sample size was limited by the number of satellite transmitters available for the study (n = 56). We attempted to get an even distribution of tracking data across the six sampling locations (range 7-12 transmitters used at each). Exact number deployed depended on logistics of trapping and nest location in each sampling area. For genomics, sample size of ≥ 5 individuals in each study location was considered sufficient due to the large number of SNPs that can be obtained for downstream analysis of the sequence data.
Data collection	Our British and Russian colleagues are mainly responsible for bird trapping, blood samples collection and satellite transmitters deployment. Noose-carpet traps were used to capture adults at nesting sites during the incubation period. For juveniles, the chicks were tagged prior to fledging in nests. Genomic and functional genomic data were generated in our lab.
Timing and spatial scale	Timing: tracking data were obtained between 2009 and 2015. Spatial scale: Kola Peninsula, Kolguev Island, Yamal Peninsula, Eastern Taimyr, Lena Delta and Lower Kolyma River in Russia.
Data exclusions	Argos location data was excluded based on the well-established Douglas Argos Filter algorithm (DAF) (Douglas et al. 2012, Methods in

Data exclusions	Ecology and Evolution) that allows filtering of Argos data by flagging locations that exceed thresholds for distance between consecutive locations, and velocity and bearing between consecutive movement vectors. For tracking data processed by the Argos system, we removed duplicate timestamps and used the DAF designed to retains points, which correspond to a realistic rate of movement (≤ 100 km/h) and which do not form tight angles between successive locations (≤ 15°).
Reproducibility	We performed the luciferase reporter experiments with multiple repeats (n ≥ 6) in each batch and three independent batches in total to confirm reproducibility. All attempts at replication were successful. The number of replications is noted in Methods.
Randomization	As described in Methods, chicken embryos laid by different hens were incubated together till 9 days and then 10-15 embryos were selected randomly to be dissected for primary cell culture. Cells were mixed well and washed by medium three times for further blending before plating into 48 well cell culture plate. Four experimental groups were randomly allocated for the following lipofection and luciferase reporter assay.
Blinding	Data included in this study were generated by different teams, and analysed by different coauthors.

Did the study involve field work? ☒ Yes ☐ No

Field work, collection and transport

Field conditions	Fieldwork was undertaken in the Eurasian Arctic in six different localities in six different years: Yamal Peninsula: 09-18 June 2009 Lena Delta: 25 June -13 July 2010 Popigai River, Eastern Taimyr: 21 June -05 July 2011 Kola Peninsula: 17 July -04 August 2012 Lower Kolyma River: 11 June -02 July 2013 Kolguev Island: 13 July-31 July 2015 In all areas we travelled by boat and camped. Temperatures ranged from -5°C to +30°C, and weather conditions were highly variable with sunshine, fog, and precipitation of rain and snow.
Location	All areas were close to sea level and central locations for each locailite are: Yamal Peninsula: 68°12'38.43"N 68°59'19.23"E Lena Delta: 72°32'3.47"N 126°37'33.85"E Popigai River, Eastern Taimyr: 72°47'2.56"N 107°45'14.04"E Kola Peninsula: 67° 6'5.18"N 39°28'41.99"E Kolguev Island: 69°20'7.14"N 49°18'32.47"E Lower Kolyma River: 68°42'30.99"N 161°17'37.72"E.
Access & import/export	We obtained permits from the Russian authorities: To trap, take blood samples and deploy harness-mounted satellite-received transmitters on peregrine falcons. To use satellite-received transmitters in Russia. To work in the border zone of Russian Federation To work in the Protected Areas of Russian Federation (Lena Delta) To transport blood samples within the Russian Federation To export blood samples from Russia and to import them into UK and China (CITES & State Veterinary permits) Import and export permissions were authorized by Federal Service for Supervision of Natural Resource Management of Russian Federation (No. 12RU00395, 13RU000601), Department for the Environment, Food and Rural Affairs Wildlife Licensing and Registration Service (No. 507316/01, 494040/01), Department for the Environment, Food and Rural Affairs (TARP/2012/306), Animal Health and Veterinary Laboratories Agency (AHVLA) Wildlife Licensing and Registration Service (No. 517558/01), and Agriculture, Fisheries and Conservation Department of Hong Kong Special Administrative Region, China (APO/IL 247/14).
Disturbance	Adult Peregrines were caught at their nests during incubation. In order to avoid damage to eggs, the clutch was substituted with hard-boiled chicken eggs during the trapping process. They were kept warm and subsequently replaced after trapping. If birds were not caught in the snare after two attempts the trapping session was abandoned.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	We used chicken primary cell culture from embryonic hippocampus tissues. We dissected the whole brain from 9 days post-fertilized incubated chicken eggs (Boehringer Ingelheim Vital Biotecnology Co,. Ltd) and took off the hippocampus from both sides of the cerebrum under the dissecting microscope.
Wild animals	Two peregrine brain samples were obtained from the Beijing Raptor Rescue Center. One naturally dead peregrine was collected from the Chongqing Zoo. Peregrine brain tissue collection and analyses were in full compliance with the Institutional Animal Care and Use Committee at the IoZ.
Field-collected samples	Blood samples were collected from 35 peregrines across the Eurasian Arctic range (10 from Kola, 5 from Kolguev, 11 from Yamal and 9 from Kolyma). Nine and six feather samples were collected from Popigai and Lena, respectively.
Ethics oversight	Permits to trap, collect blood samples and deploy satellite transmitters on peregrines were provided by the relevant authorities in Russia. All lab experiment procedures were under the guidance of the Ethics Committee of the Institute of Zoology, Chinese Academy of Sciences (IoZ, CAS). Studies in this manuscript involving peregrine brain tissue collection and analyses were in full compliance with the Institutional Animal Care and Use Committee at the IoZ, CAS.

Note that full information on the approval of the study protocol must also be provided in the manuscript.