

Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/140093/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Gu, Zhongru, Pan, Shengkai, Lin, Zhenzhen, Hu, Li, Dai, Xiaoyang, Chang, Jiang, Xue, Yuanchao, Su, Han, Long, Juan, Sun, Mengru, Ganusevich, Sergey, Sokolov, Vasily, Sokolov, Aleksandr, Pokrovsky, Ivan, Ji, Fen, Bruford, Michael W., Dixon, Andrew and Zhan, Xiangjiang 2021. Climate-driven flyway changes and memory-based long-distance migration. *Nature* 591 (7849) , 259–264. 10.1038/s41586-021-03265-0 file

Publishers page: <http://dx.doi.org/10.1038/s41586-021-03265-0>
<<http://dx.doi.org/10.1038/s41586-021-03265-0>>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies.

See

<http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



1
2 **Climate-driven flyway changes and memory-based long-distance migration**

3 Zhongru Gu^{1,2,3,18}, Shengkai Pan^{1,2,18}, Zhenzhen Lin^{1,2,18}, Li Hu^{1,2,3}, Xiaoyang Dai⁴, Jiang Chang⁵,
4 Yuanchao Xue⁶, Han Su^{1,2,3}, Juan Long^{1,2,3}, Mengru Sun^{3,6}, Sergey Ganusevich⁷, Vasiliy Sokolov⁸,
5 Aleksandr Sokolov⁹, Ivan Pokrovsky^{9,10,11}, Fen Ji¹², Michael W. Bruford^{2,13}, Andrew Dixon^{2,14,15,16},
6 Xiangjiang Zhan^{1,2,3,17*}

- 7
- 8 1. Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese
9 Academy of Sciences, Beijing 100101, China.
 - 10 2. Cardiff University – Institute of Zoology Joint Laboratory for Biocomplexity Research,
11 Chinese Academy of Sciences, Beijing 100101, China.
 - 12 3. University of the Chinese Academy of Sciences, Beijing 100049, China.
 - 13 4. School of Biological Sciences, University of Bristol, BS8 1TQ, UK.
 - 14 5. State Key Laboratory of Environmental Criteria and Risk Assessment, Chinese Research
15 Academy of Environmental Sciences, Beijing 100012, China.
 - 16 6. Key Laboratory of RNA Biology, CAS Center for Excellence in Biomacromolecules, Institute
17 of Biophysics, Chinese Academy of Sciences, Beijing 100101, China.
 - 18 7. Wild Animal Rescue Centre, Krasnostudencheskiy pr., 21-45, Moscow 125422, Russia.
 - 19 8. Institute of Plant and Animal Ecology, Ural Division Russian Academy of Sciences, 202-8
20 Marta Street, Ekaterinburg 620144, Russia.
 - 21 9. Arctic Research Station of the Institute of Plant and Animal Ecology, Ural Division Russian
22 Academy of Sciences, 21 Zelenaya Gorka, Labytnangi, Yamalo-Nenetski District 629400,
23 Russia.
 - 24 10. Department of Migration, Max Planck Institute of Animal Behavior, Am Obstberg 1, D-78315
25 Radolfzell, Germany.
 - 26 11. Laboratory of Ornithology, Institute of Biological Problems of the North FEB RAS, 18
27 Portovaya Street 685000 Magadan, Russia.

- 28 12. State Key Laboratory of Stem Cell and Reproductive Biology, Institute of Zoology, Chinese
29 Academy of Sciences, Beijing 100101, China.
- 30 13. School of Biosciences and Sustainable Places Institute, Cardiff University, Cardiff CF10 3AX,
31 Wales, UK.
- 32 14. Emirates Falconers' Club, P.O. Box, 47716, Al Mamoura Building (A), Muroor Road, Abu
33 Dhabi, United Arab Emirates.
- 34 15. Reneco International Wildlife Consultants, PO Box 61741, Sky Tower, Al Reem, Abu Dhabi,
35 UAE.
- 36 16. International Wildlife Consultants, P.O. Box 19, Carmarthen SA33 5YL, UK.
- 37 17. Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences,
38 Kunming 650223, China.
- 39 18. These authors contributed equally.

40

41 *Corresponding authors:

42 Xiangjiang Zhan

43 Email: zhanxj@ioz.ac.cn

44 **Millions of migratory birds occupy seasonally favourable breeding grounds in the Arctic¹, but**
45 **we know little about the formation, maintenance and future of Arctic bird migration routes**
46 **and genetic determinants of migratory distance. Here, we established a continental-scale**
47 **migration system, satellite tracking 56 peregrine falcons (*Falco peregrinus*) from six Eurasian**
48 **Arctic breeding populations and resequencing 35 genomes from four of these. Different**
49 **breeding populations used five migration routes across Eurasia, likely formed by longitude**
50 **and latitude breeding ground shifts during the LGM-Holocene transition. Contemporary**
51 **inter-route environmental divergence appears to maintain distinct migration routes. We**
52 **found that the novel gene *ADCY8* was associated with population-level migratory distance**
53 **differences. We elucidated its regulatory mechanism and found the most likely selective agent**
54 **for this divergence was long-term memory. Global warming is predicted to influence**
55 **migration strategies and diminish breeding ranges of Eurasian Arctic peregrines. Harnessing**
56 **ecological interactions and evolutionary processes to study climate-driven changes in**
57 **migration can facilitate the conservation of migratory birds.**

58 Global climate change and anthropogenic development are expected to affect the annual adaptive
59 movements of migratory Arctic birds¹⁻³, with potential fitness effects imposed by inhospitable
60 routes and temporally mismatched breeding^{2,4}. Next generation genome sequencing has facilitated
61 studies of the interaction between genomic variation and environment in migratory birds⁵. However,
62 to date there is no published research on the role of climate-driven genomic responses in shaping
63 differences of migratory strategy among bird populations. Here, we combined satellite-tracking of
64 56 peregrine falcons from migratory Arctic populations⁶ (**Fig. 1a, Extended Data Fig. 1,**
65 **Supplementary Table 1**) with genome data to explore their demographic history and the
66 spatiotemporal dynamics of their migration.

67 **Migration patterns of Arctic peregrines**

68 From 41 individuals, we identified 150 completed migration paths (**Supplementary Table 2**).
69 Peregrines initiated autumn migration in September, travelled 2,280-11,002 km, in *ca.* 27 days
70 (95% confidence interval (CI): 14-46) covering 213 km/day (49-420), and arrived at their wintering
71 areas in October. Peregrines migrate solitarily, with those departing from different breeding
72 grounds, except Kola and Kolguev, using different routes and wintering at widely distributed sites
73 across four distinct regions (**Fig. 1a, Extended Data Fig. 2**). Individuals tracked for more than one
74 year exhibited strong path repeatability during migration ($n = 26$; $R_{pt} = 0.45$, $P < 0.001$), complete
75 fidelity to wintering locations and limited breeding dispersal (5.37 km on average; **Fig. 1b,**
76 **Supplementary Table 3**). All populations demonstrated a high degree of migratory connectivity
77 ($R^2 = 0.86$, $P < 0.001$; **Fig. 1c**), suggesting strong selection for long-term memory.

78 Principal component analysis (PCA) identified two main groups with migratory distance being the
79 most significant differentiation (**Figs. 1d, e, Extended Data Fig. 3, Supplementary Table 4**). The
80 Eastern birds flew significantly farther than Western birds (6,134 km *vs* 3,680 km; $P < 1E-6$; **Fig.**
81 **1e**). We therefore classified them as long-distance (LD; Kolyma-Lena-Popigai-Yamal) and short-
82 distance (SD; Kolguev-Kola) migrants.

83 **Historical formation of migration routes**

84 We sequenced the genomes of 35 peregrines, obtaining 6,328,655 high-quality SNPs
85 (**Supplementary Table 5**). Multiple analytical approaches consistently supported four distinct
86 genetic clusters corresponding to the sequenced populations, with Yamal and Kolyma inferred to
87 have diverged after the separation of their ancestors from that of Kola and Kolguev (**Fig. 2a**).
88 Sequential Markovian Coalescent (SMC++) analysis revealed that the effective population size (N_e)
89 of the ancestral lineage increased from ~100 thousands of years ago (kya) to a peak 20-30 kya (**Fig.**
90 **2b**), around the Last Glacial Maximum (LGM)⁷. To resolve uncertainties in the recent demographic
91 history (**Fig. 2a**), we developed a new Approximate Bayesian Computation (ABC) approach. The
92 ABC-random forest model choice (**Fig. 2c, Supplementary Figs. 1-4, Supplementary Tables 6, 7**)
93 confirmed the divergence pattern of four studied populations and ABC simulations further found
94 that Eastern and Western populations started to separate during the LGM (23.03 kya; 95% CI:
95 17.67-32.94), followed by an eastern split between Yamal and Kolyma 11.30 kya (9.14 -14.29), and
96 between Kola and Kolguev 10.53 kya (9.18-12.90) (**Fig. 2d, Supplementary Table 8, Extended**
97 **Data Fig. 4**).

98 Ecological Niche Modelling (ENM) based on present and paleo-climate datasets showed that
99 potential breeding distribution range positively correlated with N_e fluctuations (**Supplementary**
100 **Figs. 5, 6**). There was a much larger area suitable for breeding in Siberia during the LGM than the
101 last interglacial period (LIG; 120-140 kya)⁸ or Mid-Holocene, coinciding with the largest N_e
102 estimate (**Fig. 2e, f, Supplementary Fig. 5**). Arctic-dwelling peregrines mainly occupy tundra
103 habitat⁹, and we found close coincidence between reconstructed tundra habitat and peregrine
104 breeding distribution in the LGM (**Supplementary Fig. 7**), suggesting that enlargement of tundra
105 habitat underpinned peregrine population expansion during the LGM. Conversely, population
106 declines and gradual divergences after the LGM mirrored large-scale loss and northward
107 contraction of tundra. Recent population declines after the Mid-Holocene (**Fig. 2b**) may have also

108 resulted from anthropogenic factors since habitat distributions have remained relatively stable (**Fig.**
109 **2f, g**).

110 Interestingly, our ENM simulations suggest that peregrines had less potential western wintering
111 area during the LGM, while eastern wintering areas remained stable (**Fig. 2e, Supplementary Fig.**
112 **5**). Thus, during the LGM peregrines likely migrated to a wintering area across India and Southeast
113 Asia, a striking south-eastward migration (**Fig. 2e**), distinct from the current south-westerly
114 migration route formed during the Mid-Holocene (**Fig. 2f, g**). Furthermore, since the Mid-Holocene
115 breeding areas are inferred to have shifted northward compared with the LGM, resulting in a longer
116 migratory route (**Fig. 2e, f**), we conclude that glacial cycles can regulate both migratory orientation
117 and distance.

118 **Present migration route separation**

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

124 The proposition that migration routes are genetically determined, is mostly based on migratory
125 restlessness¹¹, displacement experiments¹² and correlations between genetic background and
126 migration route¹³. We addressed this fundamental question by randomly selecting 93 putatively
127 neutrally-evolved SNP loci and 75 loci under positive selection (**Methods**). With these markers, we
128 genotyped nine and six individuals respectively for the Popigai and Lena populations, where we
129 obtained insufficient DNA from shed feathers for genome resequencing. Combining the genotypes
130 with those from the population genomic data, we measured genetic differentiation (F_{ST}) among five
131 peregrine populations and tested their relationship with the mean Hd for each migration route. We
132 found a non-significant correlation between route Hd and F_{ST} ($R^2 = 0.02$, $P = 0.69$) based on neutral

133 loci, suggesting that demographic isolation is not the major factor in isolating migration routes. In
134 contrast, we found a significant relationship for loci under selection ($R^2 = 0.42$, $P = 0.03$),
135 suggesting a role for adaptive genomic regions in the maintenance of separated migration routes
136 (**Fig. 2h**).

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

145 **Genes for migratory distance differences**

146 We used window-based F_{ST} and extended haplotype homozygosity (XP-EHH) to detect selection
147 signals across the genomes of SD and LD peregrines. The methods combined identified 149
148 selection sweeps (37 genes) between the two groups (**Supplementary Table 9**), and found the most
149 significant outlier occurred at the *ADCY8* locus (**Fig. 2j**). We narrowed the signal down to a 1.8 kb
150 region, containing 14 linked SNPs in the second intron of the gene (**Fig. 2k**). Haplotype frequency
151 analysis demonstrated a positive selection signature of *ADCY8* in the LD group (**Fig. 2l**). The
152 dominant haplotype (Hap2) was at a high frequency (34/38) in the LD group, whereas the SD group
153 had six haplotypes with varying frequencies (Hap2 11/26, Hap1 6/26, with the other four haplotypes
154 occurring 4, 2, 2, 1 times; **Fig. 2l**). We investigated the potential functional significance of Hap2 in
155 LD peregrines and a randomly selected Hap1 sequence from SD peregrines by designing a dual
156 luciferase reporter assay for functional analysis in our cultured chicken hippocampus primary cells
157 (**Methods**). In contrast to Hap2 insert cells, the Hap1 insert cells showed a significantly lower

158 luciferase activity ($P < 0.001$; **Fig. 2m, Supplementary Fig. 9**), suggesting that the peregrine Hap1
159 sequence has a suppressing effect.

160 Of the 14 loci identified within the *ADCY8* locus, the SNP (*C/T*) in the position 5,170,169 of
161 peregrine chromosome 3 produced the largest XP-EHH value (**Fig. 2n**) and the allele *T* was 100%
162 fixed in the LD group, suggesting that this standing variation is under strongest selection and may
163 have a major role in functional differentiation. A search of the focal fragment against the motif
164 database found that the ancestral SD group had a 5'-*CGTCA*-3' motif, a canonical half-site cAMP-
165 responsive element (CRE) that is a binding site for the transcription factor CREB^{14,15}, while the
166 fixed chr3-5170169**T* changes the first nucleotide of the motif. We used ATAC-seq¹⁶ to sequence
167 both hippocampus and cerebral cortex tissues from a SD peregrine. Our ATAC-seq analysis
168 detected a significant peak coinciding with the position of this motif (**Fig. 2n**), experimentally
169 supporting its existence as a functional element.

170 Previous studies suggest that CREB can regulate gene expression by binding to the CRE element
171 through CREB basic region/leucine zipper domain (bZIP), which can be regulated by its DNA
172 methylation level^{17,18}. In peregrines, the identified substitution from *C* to *T* in *ADCY8* gene creates a
173 novel transcriptional binding site, 5'-*TGTCA*-3', which potentially disrupts the DNA methylated
174 site, CpG island, on the canonical motif. Moreover, we found that, *CREB1*, the transcription
175 activator, expressed the bZIP domain in the peregrine brain, but the conditional repressor *CREB2*¹⁹
176 did not (**Supplementary Fig. 10**). Our results suggest that the new CRE motif may be free from
177 DNA methylation and facilitate the binding of CREB1 on *ADCY8*, and will consequently maintain a
178 higher activity of *ADCY8* in LD peregrines. Supporting evidence came from our comparison of
179 expression levels of *ADCY8* in the peregrine brain with two genotypes of the focal SNP ($RPKM =$
180 81.11 in *CT* vs 73.96 in *CC*; $P = 3.73E-5$, hypergeometric test).

181 Empirical evidence indicates that *ADCY8* is involved in long term memory^{20,21}. *ADCY8* encodes
182 Adenylyl Cyclase type 8 that catalyses the conversion of ATP to cAMP, which acts as a secondary

183 messenger and regulates downstream memory-related genes^{22,23}. We found LD peregrines had a
184 significantly higher mean migration path fidelity than SD peregrines ($P < 0.001$; **Supplementary**
185 **Fig. 11**), requiring strong long-term memory. Together with numerical evidence that *CREB* genes
186 determine the development of long-term memory²⁴, our work suggests that the *ADCY8* and *CREB*
187 genes play a key role in influencing migratory flight distance via co-regulating the capacity of long
188 term memory. For the *T* allele that is positively selected in the LD populations, the selection time
189 was estimated to be 18.87 kya (**Supplementary Information**), after the separation of two
190 migratory groups, strengthening our conclusion that the regulation of migratory distances is the
191 result of natural selection.

192 **Predicted effect of global warming**

193 We used ENM simulations to project future (2070) breeding and wintering distributions for each
194 Arctic peregrine population under Representative Concentration Pathway (RCP) 8.5. The breeding
195 and wintering distributions of all populations would shift poleward by 2.08° (95% CI: 0.31-3.44)
196 and 1.47° (0.11-11.06) latitude, respectively (**Extended Data Fig. 6, Supplementary Fig. 12**),
197 which is consistent with observations for most Arctic shorebirds²⁵ and congruent with the climatic
198 envelope corresponding to tundra habitats. Greatest reduction is predicted to occur in Kolguev and
199 Kola, losing 100% and 93% of their suitable breeding habitats, respectively (**Fig. 3a**). We also
200 found Western peregrines may have a much shorter migration route (655 km; 442-868), while
201 Eastern peregrines may have a longer route (286 km; 56-515) (**Fig. 3a, Supplementary Fig. 13**). If
202 the climate warms at the same rate as over recent decades, peregrines in Western Eurasia may stop
203 migrating altogether, while Eastern peregrines may face greater risks since mortality is positively
204 associated with migratory distance²⁶.

205 Recent population declines in migratory Arctic birds have been attributed to the amplification of
206 global warming in the High Arctic^{1,4,25}. Climate change may have already impacted peregrine
207 populations, so we compared N_e changes of each population with local temperature during breeding

208 periods (May-July) since 1840. Our *SNeP* analysis showed that each population has undergone
209 declines during the past 25 generations (*ca.* 150 years). However, *Ne* slope (*NeS*) analysis further
210 revealed some variation, with recovery detected during relatively cooler summers (**Fig. 3b**), and
211 four populations showed the largest negative *NeS* 8~9 generations ago (1960s), coinciding
212 widespread use of organochlorine pesticides²⁷. Importantly our generalized linear modelling (GLM)
213 analysis found that *Ne* change was negatively correlated with mean breeding season temperature in
214 Kola ($R^2 = 0.46$, $P = 0.03$) and Yamal ($R^2 = 0.39$, $P = 0.02$), with mean breeding season temperature
215 and duration of extreme warm days in Kolguev ($R^2 = 0.61$, $P = 0.01$), and with duration of extreme
216 cold/warm days in Kolyma ($R^2 = 0.68$, $P = 0.02$; **Fig. 3c**, **Supplementary Table 10**). Our predicted
217 *NeS* in the future (i.e. from 2020 to 2100) showed a continuing decrease trend, but the SD migrants
218 in Kola and Kolguev will suffer the highest probability of population decline (**Fig. 3d**,
219 **Supplementary Fig. 14**).

220 Discussion

221 Spatiotemporal changes in animal migration behaviour are thought to be related to climatic changes,
222 anthropogenic impacts²⁸ and evolutionary responses of migrants. Since these dynamic processes can
223 leave a footprint on the genome, by combining animal movement and population genomic data, we
224 were able to identify a major role of climate in the formation and maintenance of peregrine
225 migration patterns (**Fig. 2**, **Extended Data Fig. 5**).

226 Previous studies have identified several candidate genes that may regulate migration^{29,30}. The higher
227 activity of ADCY8 we identified in LD peregrine migrants (**Fig. 2**) may increase their long-term
228 memory. Our analysis reveals a unique mutation that facilitates the binding of its transcription
229 factor CREB1, and fixation of this variation happened after the divergence of LD and SD
230 populations. Our work thus not only reveals a novel causative gene to explain migratory
231 differences, but also provides a mechanistic basis.

232 In a changing global climate, peregrines may move to new wintering areas and adjust their
233 migration routes. However, our prediction of dramatic shrinkage in Arctic breeding areas, together
234 with a predicted population collapse in the European Arctic, represents a clear threat to peregrines
235 and possibly many other migratory Arctic species. Our study demonstrates the value of an
236 integrated approach, combining satellite telemetry, population and functional genomics and
237 modelling, to untangle intriguing scientific questions related to migration, laying a cornerstone for
238 conserving migratory species in conjunction with ecological interactions and evolutionary
239 processes.

240 **References**

- 241 1. McRae, L. et al. *Arctic Species Trend Index 2010. Tracking Trends in Arctic Wildlife. CAFF*
242 *International Secretariat* (2010).
- 243 2. Lameris, T. K. et al. Potential for an Arctic-breeding migratory bird to adjust spring migration
244 phenology to Arctic amplification. *Global Change Biology* **23**, 4058-4067 (2017).
- 245 3. Trautmann, S. Climate Change Impacts on Bird Species. In *Bird Species* (pp. 217-234).
246 Springer, Cham (2018).
- 247 4. Zurell, D., Graham, C. H., Gallien, L., Thuiller, W. & Zimmermann, N. E. Long-distance
248 migratory birds threatened by multiple independent risks from global change. *Nature Climate*
249 *Change* **8**, 992 (2018).
- 250 5. Bay, R. A. et al. Genomic signals of selection predict climate-driven population declines in a
251 migratory bird. *Science* **359**, 83-86 (2018).
- 252 6. White, C. M., Cade, T. J. & Enderson, J. H. *Peregrine Falcons of the World* (Lynx Edicions,
253 Barcelona, 2013).
- 254 7. Clark, P. U. et al. The last glacial maximum. *Science* **325**, 710-714 (2009).
- 255 8. Otto-Bliesner, B. L., Marshall, S. J., Overpeck, J. T., Miller, G. H. & Hu, A. Simulating Arctic
256 climate warmth and icefield retreat in the last interglaciation. *Science* **311**, 1751-1753 (2006).

- 257 9. Brambilla, M., Rubolini, D. & Guidali, F. Factors affecting breeding habitat selection in a cliff-
258 nesting peregrine *Falco peregrinus* population. *Journal of Ornithology* **147**, 428-435 (2006).
- 259 10. Hausdorff, F. Bemerkung über den Inhalt von Punktmengen. *Mathematische Annalen* **75**, 428-
260 433 (1914).
- 261 11. Pulido, F. The genetics and evolution of avian migration. *Bioscience* **57**, 165-174 (2007).
- 262 12. Perdeck, A. C. An experiment on the ending of autumn migration in starlings. *Ardea* **52**, 133-
263 139 (1964).
- 264 13. Delmore, K. E., Toews, D. P., Germain, R. R., Owens, G. L. & Irwin, D. E. The genetics of
265 seasonal migration and plumage color. *Current Biology* **26**, 2167-2173 (2016).
- 266 14. Impey, S. et al. Stimulation of cAMP response element (CRE)-mediated transcription during
267 contextual learning. *Nature Neuroscience* **1**, 595 (1998).
- 268 15. Bourtchuladze, R. et al. Deficient long-term memory in mice with a targeted mutation of the
269 cAMP-responsive element-binding protein. *Cell* **79**, 59-68 (1994).
- 270 16. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of
271 native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding
272 proteins and nucleosome position. *Nature Methods* **10**, 1213 (2013).
- 273 17. Mayr, B. & Montminy, M. Transcriptional regulation by the phosphorylation-dependent factor
274 CREB. *Nature Reviews Molecular Cell Biology* **2**, 599 (2001).
- 275 18. Iguchi-Arigo, S. M. & Schaffner, W. CpG methylation of the cAMP-responsive
276 enhancer/promoter sequence *TGACGTCA* abolishes specific factor binding as well as
277 transcriptional activation. *Genes & Development* **3**, 612-619 (1989).
- 278 19. Bartsch, D. et al. *Aplysia* CREB2 represses long-term facilitation: relief of repression converts
279 transient facilitation into long-term functional and structural change. *Cell* **83**, 979-992 (1995).
- 280 20. Wieczorek, L. et al. Absence of Ca²⁺-stimulated adenylyl cyclases leads to reduced synaptic
281 plasticity and impaired experience-dependent fear memory. *Translational Psychiatry* **2**, e126
282 (2012).

- 283 21. Rosenegger, D., Wright, C. & Lukowiak, K. A quantitative proteomic analysis of long-term
284 memory. *Molecular Brain* **3**, 9 (2010).
- 285 22. Ferguson, G. D. & Storm, D. R. Why calcium-stimulated adenylyl cyclases? *Physiology* **19**,
286 271-276 (2004).
- 287 23. Zhang, M. et al. Ca-stimulated type 8 adenylyl cyclase is required for rapid acquisition of novel
288 spatial information and for working/episodic-like memory. *Journal of Neuroscience* **28**, 4736-
289 4744 (2008).
- 290 24. Yin, J. C. & Tully, T. CREB and the formation of long-term memory. *Current Opinion in*
291 *Neurobiology* **6**, 264-268 (1996).
- 292 25. Wauchope, H. S. et al. Rapid climate-driven loss of breeding habitat for Arctic migratory
293 birds. *Global Change Biology* **23**, 1085-1094 (2017).
- 294 26. Lok, T., Overdijk, O. & Piersma, T. The cost of migration: spoonbills suffer higher mortality
295 during trans-Saharan spring migrations only. *Biology Letters* **11**, 20140944 (2015).
- 296 27. Brown, J. W. et al. Appraisal of the consequences of the DDT-induced bottleneck on the level
297 and geographic distribution of neutral genetic variation in Canadian peregrine falcons, *Falco*
298 *peregrinus*. *Molecular Ecology* **16**, 327-343 (2007).
- 299 28. Wilcove, D. S. & Wikelski, M. Going, going, gone: is animal migration disappearing. *PLoS*
300 *Biology* **6**: e188 (2008).
- 301 29. Mueller, J. C., Pulido, F. & Kempenaers, B. Identification of a gene associated with avian
302 migratory behaviour. *Proceedings of the Royal Society B: Biological Sciences* **278**, 2848-2856
303 (2011).
- 304 30. Peterson, M. P. et al. Variation in candidate genes *CLOCK* and *ADCYAP1* does not consistently
305 predict differences in migratory behavior in the songbird genus *Junco*. *F1000Research* **2**, 115
306 (2013).
- 307

308 **Figure Legends**

309 **Figure 1. Migration system.** **a**, Five migration routes for 56 satellite tracked peregrine falcons in
310 Old World. Only complete migration path is shown ($n = 41$). **b**, Individual migration path fidelity of
311 one representative individual (Lena Delta) tracked for four years (2010-2014). **c**, Migration
312 connectivity at the population level as shown in a linear regression analysis. Significance was
313 calculated using F -test ($P = 2.2E-16$). **d**, PCA analysis of migratory strategy. Small and large dots
314 represent individuals ($n = 41$) and centroids of the minimum convex polygon for each population (n
315 $= 6$). PC1 is the mix of autumn departure and arrival dates and PC2 is the migration distance. **e**,
316 Migratory distance comparison ($P < 1E-6$, two-sided t -tests, effect size = 1.43) between short-
317 distance (SD) ($n = 12$ individuals; 3,680 km: 95% CI: 2,443-5,018) and long-distance (LD) groups
318 ($n = 32$ individuals; 6,134 km: 95% CI: 3,282-8,828). In the box plots, the center line represents the
319 median, whiskers represent maximum and minimum values, and box boundaries represent 75th and
320 25th percentiles. The P values for the comparisons for any two populations within SD and LD were
321 not significant (ns; $P > 0.05$, two-sided t -tests). The tracking data of three additional Kola
322 peregrines from a previous study (**Methods**) were also used for this analysis.

323 **Figure 2. Past formation, present maintenance of migration routes and genetic basis for**
324 **migration distance differences.** **a**, PCA, Neighbor-Joining tree and *frappe* showing the
325 evolutionary relationship of the four genome-enabled populations. **b**, Demographic history
326 reconstruction for each population using SMC++. **c**, Four candidate models for model choice in
327 ABC. 84% of the total of 313 chunks (**Methods**) support Model 1. T_1 , T_2 , T_3 are divergence times.
328 **d**, Posterior distribution of divergence time estimates for the Model-1 in ABC. **e-g**, Species
329 distributions predicted during the LGM, Mid-Holocene and present. **h**, The relationship between the
330 route distance, Hd , and neutral genetic distance (Left), $F_{ST}/(1-F_{ST})$, and genetic distance based on
331 selected loci (Right). The dashed line is the linear regression line. Significance levels were
332 calculated using F -test. **i**, Proportion of grids ($0.083^\circ \times 0.083^\circ$) with different Köppen-Geiger
333 climate zones within each migration route. Full names of climate zones can be seen in

334 **Supplementary Information. j**, Spline-window based F_{ST} and XP-EHH to detect selective sweeps.
335 Red points indicate windows containing selected genes. **k**, *ADCY8* haplotype heatmap. The dashed
336 rectangle marks the focal 1.8kb fragment in *ADCY8*. Red and light blue squares symbol different
337 alleles in each column (SNP). **l**, Haplotype frequency in the identified segment. **m**, Results of dual-
338 luciferase reporter assay in chicken hippocampus primary cells. Data are mean \pm s.e.m. Significance
339 levels were calculated using two-sided *t*-test ($n = 14$ replicates for each of the first three groups and
340 $n = 6$ for the *pGL3*-basic group). **n**, XP-EHH results for every selected SNPs within 1.8 kb flanking
341 regions (Upper) and ATAC-seq results confirming the existence of CRE-motif (Lower).

342 **Figure 3. Shortened migration route and population decline in Europe populations due to**
343 **global warming. a**, Area changes in breeding and wintering areas (Upper) between present and
344 future (2070; RCP 8.5) and migratory distance comparisons (Lower) in the six peregrine
345 populations. N/A: no predicted future breeding areas. Δd is the mean change of migration
346 distances. In the box plots, the center line represents the median, whiskers represent maximum and
347 minimum values, and box boundaries represent 75th and 25th percentiles. $n = 200$ for each
348 comparison. Significance and effect size were calculated using two-sided *t*-tests and Cohen's *d*,
349 respectively (Kola: $P = 3.4E-9$, effect size = 0.580; Yamal: $P = 0.090$, effect size = 0.170; Popigai:
350 $P = 0.676$, effect size = 0.042; Lena: $P = 0.868$, effect size = 0.017; Kolyma: $P = 0.015$, effect size
351 = 0.243). **b**, LD_{Ne} and NeS estimates in recent 25 generations. **c**, The linear regression between NeS
352 and the most significant environmental variable identified in the GLM analysis. Data are median \pm
353 95% CI. Dashed lines represent the linear regression lines. **d**, NeS changes of each population
354 predicted in the future. Data are median \pm 95% CI.

355 **Methods**

356 **Tracking peregrine migration**

357 We used satellite-received Argos Platform Transmitter Terminals (PTTs) and GSM-received GPS
358 transmitters to track 56 peregrines from six breeding regions in Arctic Eurasia (**Supplementary**
359 **Table 1**). From which, we obtained the data of at least one full migration for 41 peregrines: Kola
360 Peninsula ($n = 1$ autumn/0 spring migration), Kolguev Island ($n = 8/4$), Yamal Peninsula ($n = 9/5$),
361 Popigai River ($n = 10/9$), Lena Delta ($n = 6/6$) and Lower Kolyma River ($n = 7/5$) (**Supplementary**
362 **Information**). Permits to trap, collect blood samples and deploy satellite transmitters on peregrines
363 were provided by the relevant authorities in Russia. All lab experiment procedures were under the
364 guidance of the Ethics Committee of the Institute of Zoology, Chinese Academy of Sciences (IoZ,
365 CAS). Studies in this manuscript involving peregrine brain tissue collection and analyses were in
366 full compliance with the Institutional Animal Care and Use Committee at the IoZ, CAS.

367 For tracking data processed by the Argos system, we removed duplicate timestamps and used the
368 Douglas Argos Filter algorithm designed to retains points (**Supplementary Fig. 15**), which
369 correspond to a realistic rate of movement (≤ 100 km/h) and which do not form tight angles
370 between successive locations ($\leq 15^\circ$)³¹. Migration strategy^{32,33} was quantified using
371 departure/arrival date, duration and migratory distance (migration path distance) for all individuals,
372 where possible (**Supplementary Fig. 16**). We defined the start and end of migration as the day that
373 birds moved more than 40 km from their breeding (natal) range, or arrived at wintering range.

374 **Population migration connectivity and wintering distribution pattern**

375 To quantify the degree of migratory connectivity³⁴, we extracted the longitudes of breeding and
376 wintering sites for each tracked individual and used a linear regression to explore the correlation
377 between breeding sites and wintering sites as used in a previous study³³. The coefficient of
378 determination (R^2) was used to proxy the migratory connectivity. For the spatial distribution pattern

379 of wintering sites, we obtained the minimum convex polygon (MCP) of wintering sites and used a
380 G -function from R package *spatstat*³⁵ to conduct a point pattern analysis. A greater empirical $\hat{G}(x)$
381 than the theoretical function suggests that the sites tend to be closer than expected, in contrast with a
382 dispersed pattern. More details please see **Supplementary Information**.

383 **Individual migration path fidelity**

384 We estimated the path fidelity of each bird by assessing individual repeatability³⁶ of migration paths
385 across multiple years ($n \geq 2$). To quantify the consistency of these migratory paths, we first
386 calculated the track deviation from the great circle distance for each path in each year and then
387 evaluated the repeatability of deviation after standardized measurement based on latitude.
388 Specifically, we measured the total migratory distance (sum of all distances between successive
389 positions along a migration path). Then, the straightness was calculated as the distance between the
390 start and end locations of the path divided by total migratory distance³⁷. The repeatability of path
391 straightness was calculated using a linear mixed model implemented in the R package *rptR*³⁸,
392 followed by a significance testing through a permutation test.

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

400 **Migratory strategy comparison**

401 Principal component analysis was used to cluster individuals based on their migratory strategy (i.e.
402 migratory distance, duration and departure/arrival date during autumn migration). Noted that we did

403 not use the spring departure/arrival date because we did not obtain these data for the Kola
404 population. In our study, only individuals that completed at least one migration route were included,
405 and for individuals that were tracked for multiple years, we used mean migratory values to control
406 pseudo-replication. We removed one of the variables if they were highly related ($|r| > 0.75$). For the
407 comparison of migratory strategy between SD and LD groups, we used a random forest model in *R*
408 and *t*-test to detect the most significantly different strategy parameter. We calculated the effect size
409 (Cohen's d^{39}) for the *t*-test. In the comparisons, as we obtained the data of only one complete
410 autumn migration in the Kola population (**Supplementary Table 2**), we also used the tracking data
411 of three Kola peregrines from a previous study⁴⁰.

412 **Sample collection and genomic DNA extraction and sequencing**

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

420 **Sequencing data filtering and SNP calling**

421 An average of 68.75 Gb clean data (55.79×) were generated for 35 individuals. We used FASTQC
422 and trimmomatic⁴¹ to remove reads with low quality as previously described⁴². We used the
423 chromosome-level peregrine genome assembly⁴³ as the reference genome, where our original
424 assembly⁴⁴ was upgraded to chromosomal fragments. We then used the Burrows-Wheeler
425 Alignment⁴⁵ to map the filtered reads from each individual onto the autosomal reference genome
426 with Z-chromosome fragments excluded (**Supplementary Information**). Finally, we used the
427 pipeline in Genome Analysis Toolkit⁴⁶ (version 3.5) to call SNPs.

428 **Population genomic analysis**

429 Since close relatedness can bias population assignment⁴⁷, pairwise Identity-By-State scores among
430 all individuals were estimated using PLINK⁴⁸ (version 1.9) and a threshold of 0.1 was applied,
431 resulting in the removal of three closely related individuals for the following analysis
432 (**Supplementary Fig. 17**). PCA was conducted on the autosomal biallelic SNPs for the remaining
433 32 individuals using our in-house scripts. To reconstruct the phylogenetic tree, we used PLINK to
434 calculate genetic distances among the studied peregrines based on the identified SNPs with default
435 settings. A neighbor joining unrooted tree was then obtained using the *R* package ‘*phangorn*’ with
436 *upgma* function⁴⁹. Analysis of genetic structure was implemented in *frappe*⁵⁰ which employs an
437 expectation maximization algorithm. The number of genetic clusters *K* was set to range from 2 to 5.

438 **Demographic history reconstruction**

439 We used SMC++ (version 1.10.0)⁵¹ to model historical effective population sizes for each peregrine
440 population with the mutation rate and generation time derived from our previous estimates⁴⁴. To
441 date the divergence among the peregrine populations, we developed a new ABC approach (details
442 in **Supplementary Information**). Briefly, we established four candidate historical demographic
443 models for model choice according to our phylogenetic results. The prior *N_e* distributions and
444 divergence time parameters were set from a range of 1,000 to 100,000 and 1,000 to 10,000
445 generations ago, respectively, according to the SMC++ results. The peregrine genome was divided
446 into chunks with a size of 2 Mb and gene number in each chunk was counted and ranked. We
447 conducted 100,000 simulations for each candidate model using the rapid coalescent *scrm*
448 simulator⁵² and summarized them using 95 different summary statistics (e.g. the total number of
449 segregating sites, summarized site frequency spectrum). A machine learning tool, ABC random
450 forest⁵³ was employed to conduct model choice. For the selected model with the highest
451 approximated posterior probability (Model-1), we further simulated 1,000,000 datasets for
452 parameter inferences. The neural network methods were applied for the inference in *R* package

453 *abc*⁵⁴. To evaluate the ABC performance, we applied the cross validation on model choice and
454 parameter inference.

455 **Ecological niche modeling**

456 We used MaxEnt (version 3.3.3k), in the *R dismo*⁵⁵ package, to predict breeding and wintering
457 distributions under present environment conditions. Based on the satellite tracking data, we
458 randomly selected presence data within the MCPs of individuals' summer (June to August) and
459 wintering areas (December to January). The 90% MCP was calculated for each individual using the
460 *MCP* function of the *adehabitatHR*⁵⁶ package in *R*. For climate data, we downloaded 19 present and
461 paleo bioclimatic variables (**Supplementary Table 11**) from WorldClim⁵⁷. To reconstruct breeding
462 and wintering distributions in the past, we projected the ENMs built under present climate to
463 paleoclimates during the Mid-Holocene (*ca.* 6 kya), LGM (22 kya) and LIG (120-140 kya),
464 respectively.

465 **Paleo vegetation data analysis**

466 To examine whether the predicted LGM breeding areas mostly consisted of tundra biome, we
467 obtained Eurasian paleo pollen data from a previous study⁵⁸. Paleo pollen data classified as tundra
468 biome by the *biome_2000* model were extracted for LGM and then mapped to our predicted
469 breeding areas to estimate the overlapping extent.

470 **Quantification of inter-route distances**

471 *Hausdorff distances* were used to quantify the dissimilarity between migratory paths of individual
472 pairs of peregrines. The approach measures how far apart two subsets of a space metric are from
473 each other¹⁰. In our study, a migratory path was treated as positional distribution of bird movement
474 points in time and space and the distance between the migration path *A* and *B* was obtained using
475 '*hausdorff_dist*' function in *R* package *pracma*⁵⁹ (**Supplementary Information**). The mean *Hd*
476 estimates were compared within and between neighboring populations using a *t*-test. The generated

477 *Hd* matrix was then used for the route clustering analysis using a ‘*hclust*’ method in *R* package
478 ‘*heatmap*’. The inter-route distance was finally calculated as the mean *Hd* between pairs of
479 individual paths from different migration routes.

480 **Maintenance mechanisms of present migration routes**

481 To investigate maintenance mechanisms of present migration routes, we first checked the influence
482 of neutral genetic and selective genetic distance on the route distances estimated above. For the
483 estimation of genetic distance, we randomly selected 93 putatively neutrally-evolving and 75
484 selected SNP loci based on the selection analysis of 32 resequenced peregrine genomes described
485 below. Then, we genotyped nine and six individuals (shed feathers) sampled in Popigai and Lena,
486 respectively. The detailed PCR amplifications and sequencing of these feather DNA extracts are
487 described in **Supplementary Information**. With the combined genotypes from blood and feather
488 samples, we calculated the F_{ST} for each locus among five populations (Kolguev, Yamal, Popigai,
489 Lena and Kolyma) using *vcftools*⁶⁰. The genetic distance was then calculated as $F_{ST}/(1 - F_{ST})^5$. We
490 fitted the relationship between migration route separation (*Hd*) and neutral and selective genetic
491 distance, respectively, using a linear regression model.

492 To investigate the environmental divergence among migration routes, we randomly sampled 200
493 grids ($0.083^\circ \times 0.083^\circ$) from the 90% MCP of each route and extracted the variable of climate
494 zones referring to the Köppen-Geiger classification system⁶¹ from each grid. Chi-squared tests were
495 applied for the testing of differences in climate zones between adjacent migration routes (**Extended**
496 **Data Fig. 5b**). To check the environmental boundaries, we further divided the Eurasian continent
497 into geographic bands (width in 2° longitude based on the estimated migration distance per day in
498 the studied peregrines) with direction parallel to the mean migration angle of all individuals and
499 calculated the median value of climate zones of the grids ($0.083^\circ \times 0.083^\circ$) at regular intervals (1°
500 in latitude) from neighbouring bands for paired comparisons (illustrated in **Extended Data Fig. 5c**).
501 Paired *t*-tests were used to check the abrupt change of climate zone (boundary) (**Extended Data**

502 **Fig. 5d)** between adjacent bands. It is noted that during comparison, we used the same latitude
503 range between the pairwise routes.

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

513 **Identification of selective sweeps and detection of selected SNPs between SD and LD** 514 **peregrines**

515 We used two methods, a window-based F_{ST} and XP-EHH⁶² to identify selective sweeps between the
516 SD and LD groups. The F_{ST} of each locus was calculated using *vcftools*. A smoothed spline
517 technique in *R* package *GenWin*⁶³ (version 1.0.1) was implemented to determine the window
518 boundary and *w*-statistic was used as a proxy of windowed F_{ST} . The XP-EHH value of each locus
519 was calculated using *selscan*⁶⁴ (version 1.1) with BEAGLE-phased⁶⁵ SNPs (*n iterations*=100).
520 Outlier regions (top 1%) detected by both methods were considered as selective sweeps. For the
521 sweeps identified on the focal gene, we calculated the nucleotide diversity (θ_{π}) of each locus using
522 *vcftools* and narrowed the sweep down to a specific region. In addition, to verify the phased
523 haplotype of *ADCY8* gene, we chose three individuals from each population for 10x genomics
524 sequencing, followed by linked-reads phasing using the Long Ranger⁶⁶ (version 2.2.2)
525 (**Supplementary Table 12**).

526 We then used an integrative method⁶⁷ to detect the selected SNPs between the pairwise groups
527 (Kolguev vs LD and Yamal vs Kolyma), which integrated the results obtained from three selection
528 tests: the *FLK*⁶⁸ based on comparing different patterns of allele frequencies among populations to
529 the values expected under a scenario of neutral evolution⁶⁹, latent factor mixed models (LFMM) for
530 which the environment is used as a fixed effect and latent factors used to infer environmental
531 associations⁷⁰, and *pcadapt* (version 3.03) analysis based on Bayesian factor model. Detailed
532 settings were described in our previous work⁷¹. The adjusted z -scores were calculated for each of
533 three above tests, and the calibrated P values were obtained as previously reported⁷⁰. The candidate
534 SNP loci were ultimately determined using Benjamini-Hochberg FDR (false discovery rate) control.
535 The level of FDR was set to 0.05.

536 **Luciferase reporter assay for the focal *ADCY8* haplotype**

537 To investigate the potential functional significance of different *ADCY8* haplotypes, a randomly
538 selected Hap1 in SD peregrines and the dominant Haplotype (Hap2) in LD peregrines (**Fig. 2l**) were
539 fully synthesized in SinoGenoMax Co., Ltd and inserted into the *pGL3*-promoter backbone
540 according to KpnI/XhoI restriction sites. Positive (*pGL3*-promoter) and negative (*pGL3*-basic)
541 controls were also constructed. The activity of Hap1 or Hap2 was examined in primary cells
542 cultured from chicken embryonic hippocampus tissues. We isolated and cultured neurons from
543 fertilized chicken eggs (Boehringer Ingelheim; details in **Supplementary Information**) for
544 lipofection and luciferase reporter assays referring to previous studies⁷². *pGL3*-Hap1 or *pGL3*-Hap2
545 were co-transfected with pRL-TK into the neurons using Lipofectamine 2000 (Invitrogen). After 48
546 h incubation, the dual-luciferase activity was measured using Dual-Glo[®] Luciferase Assay kit
547 (Promega). A t -test was used to compare the fluorescence intensity among experimental groups. At
548 least three independent experiments of each assay were performed with a minimum of six
549 replicates. To predict the motif in this focal region, we extracted the sequences (10 bp) around each
550 SNP and searched the sequences against the TFBS motif database⁷³.

551 **ATAC-seq and RNA-seq analysis**

552 Hippocampus and cortex tissues were collected from a peregrine that died of natural causes in the
553 Chongqing Zoo. For the ATAC experiment (**Supplementary Information**), the samples were
554 prepared according to the manual⁷⁴ in Shanghai Jiayin Biotechnology Ltd, followed by library
555 construction and subjected to sequencing on an Illumina NovaSeq 6000 (Novogene). The raw reads
556 generated were further quality controlled. The clean reads were mapped to our reference peregrine
557 genome and sequencing depth around the target region was evaluated.

558 For the RNA-seq, brain tissues were collected from two humanely euthanized peregrines from the
559 Beijing Raptor Rescue Center. Total RNA was extracted from the samples of two peregrines (chr3-
560 5170169SNP genotype *CC* and *CT*) using the TRIzol reagent according to the user guide
561 (Invitrogen). For each brain RNA sample, library with insert-size of 350 bp was constructed and
562 then sequenced on a HiSeq 2500 (Novogene). *Reads Per Kilobases per Million reads (RPKM)* of
563 *ADCY8* with different genotypes were calculated by mapping RNA-seq reads to the peregrine gene
564 set⁴⁴ using SOAP⁷⁵ (version 2.22). Expression difference between these genotypes was compared
565 using a hypergeometric test.

566 Detailed descriptions of the ATAC-seq please see **Supplementary Information**.

567 **Global warming impacts and prediction of migratory distance in the future**

568 Based on the inferred present peregrine distribution, we predicted future (2070) potential breeding
569 and wintering distributions under RCP 8.5, a scenario where emissions continue to rise through the
570 21st century. Occurrence probabilities were transformed into binary maps using true skill statistic-
571 maximizing values as thresholds. Differences between present and future distributions were
572 investigated using two parameters: area change and latitude shift. We counted the grids ($0.083^\circ \times$
573 0.083°) within the non-overlap regions between present and future, and the shifted latitude was
574 represented as the ratio of area and longitude (per degree). To compare the migratory distance

575 between the present and future, we randomly selected 50 sites per population within breeding or
576 wintering ranges and calculated great circle distances between corresponding sites. The confidential
577 interval of these distance estimates were calculated as the minimum and maximum of great circle
578 distances among all the sites in each of breeding and wintering ranges. For comparison, we
579 conducted the same analyses under RCP 4.5, a mitigation scenario where emissions peak around
580 2040 (**Supplementary Information**).

581 **Prediction of effective population size changes in the future**

582 To study how effective population size will change under future climate change, we initially
583 investigated the association between most recent changes of N_e and climate variables since the
584 industrial revolution (1840). We used a linkage disequilibrium based method $SNeP$ ⁷⁶ (version 1.11)
585 to reconstruct recent N_e changes in generations for each population. For this analysis, we randomly
586 selected 5,000 loci per chromosome and used a NeS to investigate the rate of N_e changes⁷⁷. It is
587 noted that $SNeP$ could reliably examine the changes or trends of N_e , rather than the actual N_e ⁷⁶.

588 For the climate data, the monthly temperature of the Community Climate System Model (version 4)
589 output was downloaded from Coupled Model Intercomparison Project 5. The yearly mean
590 temperature was calculated as the average of monthly temperature in the breeding season (May to
591 July), a vital period for the breeding success of peregrines^{78,79}. To quantify the historical weather
592 extremes, we downloaded the gridded daily minimum and maximum temperature data from
593 Berkeley Earth Surface Temperature. The number of days exceeding the 95% upper threshold of
594 average temperature was considered as extreme hot days and below the 5% lower threshold as
595 extreme cold days. The simulated future daily temperature (2020 to 2100) was downloaded from
596 the NASA Earth Exchange Global Daily Downscaled Projections. The inferring method for future
597 weather extremes was the same as historical extremes.

598 We constructed a GLM to model the relationship between *NeS* and changes of climate variables
599 (**Supplementary Information**) and finally predicted the future *NeS* under future climate using the
600 fitted GLM model.

601 **Statistical analysis**

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

- 611 31. Douglas, D. C., Weinzierl, R. C., Davidson, S., Kays, R., Wikelski, M. & Bohrer, G.
612 Moderating Argos location errors in animal tracking data. *Methods in Ecology and Evolution* **3**,
613 999-1007 (2012).
- 614 32. Mueller, T., O'Hara, R. B., Converse, S. J., Urbanek, R. P. & Fagan, W. F. Social learning of
615 migratory performance. *Science* **341**, 999-1002 (2013).
- 616 33. Trierweiler, C. et al. Migratory connectivity and population-specific migration routes in a long-
617 distance migratory bird. *Proceedings of the Royal Society B: Biological Sciences* **281**,
618 20132897 (2014).
- 619 34. Ambrosini, R., Møller, A. P. & Saino, N. A quantitative measure of migratory
620 connectivity. *Journal of Theoretical Biology* **257**, 203-211 (2009).
- 621 35. Baddeley, A., Rubak, E. & Turner, R. *Spatial Point Patterns: Methodology and Applications*
622 *with R*. Chapman and Hall/CRC Press (2015).

- 623 36. López-López, D. P., García-Ripollés, C. & Urios, V. Individual repeatability in timing and
624 spatial flexibility of migration routes of trans-Saharan migratory raptors. *Current Zoology* **60**,
625 642-652 (2014).
- 626 37. Benhamou, S. How to reliably estimate the tortuosity of an animal's path: straightness,
627 sinuosity, or fractal dimension? *Journal of Theoretical Biology*, **229**, 209-220 (2004).
- 628 38. Stoffel, M. A., Nakagawa, S. & Schielzeth, H. rptR: Repeatability estimation and variance
629 decomposition by generalized linear mixed-effects models. *Methods in Ecology and*
630 *Evolution* **8**, 1639-1644 (2017).
- 631 39. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*. New York, Routledge
632 Academic (1988).
- 633 40. Ganusevich, S. A. et al. Autumn migration and wintering areas of peregrine falcons *Falco*
634 *peregrinus* nesting on the Kola Peninsula, northern Russia. *Ibis* **146**, 291-297 (2004).
- 635 41. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence
636 data. *Bioinformatics* **30**, 2114-2120 (2014).
- 637 42. Zhao, S. et al. Whole-genome sequencing of giant pandas provides insights into demographic
638 history and local adaptation. *Nature Genetics* **45**, 67-71 (2013).
- 639 43. Damas, J. et al. Upgrading short-read animal genome assemblies to chromosome level using
640 comparative genomics and a universal probe set. *Genome Research* **27**, 875-884 (2017).
- 641 44. Zhan, X. et al. Peregrine and saker falcon genome sequences provide insights into evolution of
642 a predatory lifestyle. *Nature Genetics* **45**, 563-566 (2013).
- 643 45. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler
644 transform. *Bioinformatics* **25**, 1754-1760 (2009).
- 645 46. DePristo, M. A. et al. A framework for variation discovery and genotyping using next-
646 generation DNA sequencing data. *Nature Genetics* **43**, 491-498 (2011).

- 647 47. Rodriguez-Ramilo, S. T. & Wang, J. The effect of close relatives on unsupervised Bayesian
648 clustering algorithms in population genetic structure analysis. *Molecular Ecology*
649 *Resources* **12**, 873-884 (2012).
- 650 48. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage
651 analyses. *The American Journal of Human Genetics* **81**, 559-575 (2007).
- 652 49. Schliep, K. P. phangorn: phylogenetic analysis in R. *Bioinformatics* **27**, 592-593 (2010).
- 653 50. Tang, H. et al. Genetic structure, self-identified race/ethnicity, and confounding in case-control
654 association studies. *The American Journal of Human Genetics* **76**, 268-275 (2005).
- 655 51. Terhorst, J., Kamm, J. A. & Song, Y. S. Robust and scalable inference of population history
656 from hundreds of unphased whole genomes. *Nature Genetics* **49**, 303-309 (2017).
- 657 52. Staab, P. R., Zhu, S., Metzler, D. & Lunter, G. scrm: Efficiently simulating long sequences
658 using the approximated coalescent with recombination. *Bioinformatics* **31**, 1680-1682 (2015).
- 659 53. Pudlo, P., Marin, J. M., Estoup, A., Cornuet, J. M., Gautier, M., & Robert, C. P. Reliable ABC
660 model choice via random forests. *Bioinformatics*, **32**, 859-866 (2016).
- 661 54. Csilléry, K., François, O. & Blum, M. G. abc: an R package for approximate Bayesian
662 computation (ABC). *Methods in Ecology and Evolution* **3**, 475-479 (2012).
- 663 55. Hijmans, R. J., Phillips, S., Leathwick, J., Elith, J. & Hijmans, M. R. J. Package
664 ‘dismo’. *Circles* **9**, 1-68 (2017).
- 665 56. Calenge, C. Home range estimation in R: the adehabitatHR package. *Office national de la*
666 *classe et de la faune sauvage: Saint Benoist, Auffargis, France* (2011).
- 667 57. Fick, S. E. & Hijmans, R. J. WorldClim2: new 1-km spatial resolution climate surfaces for
668 global land areas. *International Journal of Climatology* **37**, 4302-4315 (2017).
- 669 58. Cao, X., Tian, F., Dallmeyer, A. & Herzschuh, U. Northern Hemisphere biome changes (> 30°
670 N) since 40 cal ka BP and their driving factors inferred from model-data
671 comparisons. *Quaternary Science Reviews* **220**, 291-309 (2019).
- 672 59. Borchers, H. W. & Borchers, M. H. W. R Package ‘pracma’. (2015).

- 673 60. Danecek, P. et al. The variant call format and VCFtools. *Bioinformatics* **27**, 2156-2158 (2011).
- 674 61. Beck, H. E. et al. Present and future Köppen-Geiger climate classification maps at 1-km
675 resolution. *Scientific Data* **5**, 180214 (2018)
- 676 62. Sabeti, P. C. et al. Genome-wide detection and characterization of positive selection in human
677 populations. *Nature* **449**, 913-918 (2007).
- 678 63. Beissinger, T. M., Rosa, G. J., Kaeppler, S. M., Gianola, D. & De Leon, N. Defining window-
679 boundaries for genomic analyses using smoothing spline techniques. *Genetics Selection
680 Evolution* **47**, 30 (2015).
- 681 64. Szpiech, Z. A. & Hernandez, R. D. selscan: an efficient multithreaded program to perform
682 EHH-based scans for positive selection. *Molecular Biology and Evolution* **31**, 2824-2827
683 (2014).
- 684 65. Browning, S. R. & Browning, B. L. Rapid and accurate haplotype phasing and missing-data
685 inference for whole-genome association studies by use of localized haplotype clustering. *The
686 American Journal of Human Genetics* **81**, 1084-1097 (2007).
- 687 66. Zheng, G. X., Lau, B. T., Levin, S. M., Jarosz, M., Bell, J. M., Hindson, C. M., Kyriazopoulou-
688 Panagiotopoulou, S., Masquelier, D. A., Merrill, L., Terry, J.
689 M. Haplotyping germline and cancer genomes with high-throughput linked-
690 read sequencing. *Nature Biotechnology*, **34**, 303 (2016).
- 691 67. François, O., Martins, H., Caye, K. & Schoville, S. D. Controlling false discoveries in genome
692 scans for selection. *Molecular Ecology*, **252**, 454-469 (2016).
- 693 68. Fariello, M. I., Boitard, S., Naya, H., SanCristobal, M. & Servin, B. (2013). Detecting
694 signatures of selection through haplotype differentiation among hierarchically structured
695 populations. *Genetics*, **193**, 929-941 (2013).
- 696 69. Bonhomme, M., Chevalet, C., Servin, B., Boitard, S., Abdallah, J., Blott, S. & SanCristobal, M.
697 (2010). Detecting selection in population trees: the Lewontin and Krakauer test extended.
698 *Genetics*, **186**, 241-262 (2010).

- 699 70. Frichot, E. & François, O. LEA: An R package for landscape and ecological association
700 studies. *Methods in Ecology and Evolution*, **6**, 925-929 (2015).
- 701 71. Pan, S. et al. Population transcriptomes reveal synergistic responses of DNA polymorphism
702 and RNA expression to extreme environments on the Qinghai-Tibetan Plateau in a predatory
703 bird. *Molecular Ecology*, **26**, 2993-3010 (2017).
- 704 72. Zhang, Y. et al. An RNA-sequencing transcriptome and splicing database of glia, neurons, and
705 vascular cells of the cerebral cortex. *Journal of Neuroscience*, **34**, 11929-11947 (2014).
- 706 73. Yang, L. et al. TFBSshape: a motif database for DNA shape features of transcription factor
707 binding sites. *Nucleic Acids Research* **42**, D148-D155 (2013).
- 708 74. Buenrostro, J. D., Wu, B., Chang, H. Y. & Greenleaf, W. J. ATAC-seq: a method for assaying
709 chromatin accessibility genome-wide. *Current Protocols in Molecular Biology* **109**, 21-29
710 (2015).
- 711 75. Li, R. et al. *De novo* assembly of human genomes with massively parallel short read
712 sequencing. *Genome Research* **20**, 265-272 (2010).
- 713 76. Barbato, M., Orozco-terWengel, P., Tapio, M. & Bruford, M. W. SNeP: a tool to estimate
714 trends in recent effective population size trajectories using genome-wide SNP data. *Frontiers in*
715 *Genetics* **6**, 109 (2015).
- 716 77. Pitt, D. et al. Demography and rapid local adaptation shape Creole cattle genome diversity in
717 the tropics. *Evolutionary Applications* **12**, 105-122 (2019).
- 718 78. Carlzon, L., Karlsson, A., Falk, K., Liess, A. & Møller, S. Extreme weather affects peregrine
719 falcon (*Falco peregrinus tundrius*) breeding success in South Greenland. *Ornis Hungarica* **26**,
720 38-50 (2018).
- 721 79. Franke, A. et al. Status and trends of circumpolar peregrine falcon and gyrfalcon
722 populations. *Ambio* **49**, 762-783 (2020).

723 **Acknowledgements**

724 This study was supported by National Natural Science Foundation of China (No. 31821001,
725 31930013, 31911530186, 91740201), Strategic Priority Program of Chinese Academy of Sciences
726 (XDB31000000), the National Key Programme of Research and Development, Ministry of Science
727 and Technology (2016YFC0503200), Youth Innovation Promotion Association of Chinese
728 Academy of Sciences (2020086) to S.P., Second Tibetan Plateau Scientific Expedition and
729 Research Program (STEP) (No. 2019QZKK0501), Water Ecological Security Assessment, the
730 Major Research Strategy for Middle and Lower Yangtze River (Grant No. ZDRW-ZS-2017-3),
731 Biodiversity Survey, Monitoring and Assessment Project (2019-2023) of Ministry of Ecology and
732 Environment, China, the Royal Society to X.Z. and M.W.B., and the CAS President's International
733 Fellowship Initiative for Visiting Scientists to M.W.B. Funding for satellite tracking, sampling and
734 partial genome resequencing was provided by the Environment Agency-Abu Dhabi. Additional
735 tracking and sampling from Kolguev and Kola was funded by the Max Planck Institute of Animal
736 Behavior to I.P. and Earthspan (www.earthspan.foundation). We thank Dr. J.M. Graves for her
737 suggestions, M. Al Bowardi and M. Al Mansouri for their support, Drs. X. Li, Q. Dai, X. Liu, Y.
738 Chen, J. Jiao, G. Wang and M. Barbato for their help in data analysis, W. Wu, X. Hou, Y. Wang
739 and Y. Zhan for their advice on figure drawing, and O. Kulikova, V. Pozdnyakov, S. Troyev, L.
740 Zhou, C. Dai, Y. Yin, C. Li and N. C. Fox during fieldwork.

741 **Author contributions**

742 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.,
743 J.L. and Z.L. conducted the fieldwork and sample collection. X.Z. and A.D. examined migration
744 paths, migration connectivity and genetic structure of peregrines across Eurasia. X.Z. and M.W.B.
745 supervised the population genomic research. Z.G., S.P., L.H., J.C., and X.D. performed the data
746 analyses. Z.L., Y.X., M.S., H.S. and F. J. conducted the molecular experiments. X.Z. and Z.G.
747 wrote the manuscript with contributions from M.W.B., S.K. and A.D.

748 **Competing interests**

749 The authors declare no competing interests.

750 **Additional information**

751 **Supplementary Information** is available in the online version of the paper.

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

753 **Reporting summary**

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

756 **Data availability**

757 All of the sequenced genome data have been deposited in the GenBank under accession number
758 PRJNA686418. The tracking data are included in the Arctic Animal Movement Archive and in
759 Movebank under ID numbers 103426553 and 934079034.

760 **Extended Data Figure Legends**

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

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

771 **Extended Data Fig. 3. The migration strategy comparisons between SD and LD. a,** Variable
772 importance estimated by random forest modelling. **b,** Migratory strategy comparisons between the
773 short-distance (SD) and long-distance (LD) peregrine groups. Significance was determined by a
774 two-sided *t*-test and sample size (n) for each comparison is shown. In the box plots, the center line
775 represents the median, whiskers represent maximum and minimum values, and box boundaries
776 represent 75th and 25th percentiles.

777 **Extended Data Fig. 4. ABC simulation and parameter inference. a,** Linear discriminant
778 summary statistics values of the simulated datasets and the observations given four ABC candidate
779 models. Based on the three statistics (LD1-3), Model-1 is best supported because the targets (dark)
780 fit simulated data (shadow) well. **b,** Distribution of divergence times estimated using the chunks
781 supporting the Model-1. One column represents one chunk and we only show 100 chunks. The
782 density bar symbols posterior distribution of inferred divergence time in each chunk.

783 **Extended Data Fig. 5 Maintenance mechanisms of present migration routes. a,** Route cluster
784 analysis based on *Hd*. **b,** χ^2 testing results of climate zones between adjacent migration routes at the
785 whole route level. **c,** The schematic diagram of environment comparisons between neighboring
786 geographic bands. Each route was divided into geographic bands parallel to the main migration
787 direction. Grids at regular intervals were chosen from neighboring bands for comparisons. **d,**
788 Environmental boundaries coinciding with migration route boundaries. The Eurasia continent was
789 divided into geographic bands (at 2° longitude). The *P* values of paired *t*-tests between compared
790 bands are shown and the dashed line equals to 0.05 (Upper). The bar is scaled as the number of
791 space between two targeted bands in a paired comparison. The MCPs (90%) of five migration
792 routes are shaded (Lower). Arrows point the coincidence between environmental and migration
793 route boundaries. Noted that the distinct environment difference within the Popigai route may result
794 from the inclusion of large “barrier islands” of unsuitable region in comparison. **e,** Illustration of the
795 model simulating the least-cost migration path. For a typical migration route, we simulated the
796 potential migration path (dashed lines) that a peregrine depart from its actual breeding site (e.g. B1
797 in Route1), fly along a least-cost path, but winter in a wintering site of the neighboring route (e.g.
798 W2 in Route2). B1-3 means breeding areas and W1-3 means wintering areas. Solid lines are the
799 actual tracked migration path. **f,** Migration cost comparison between within-route and across-route
800 paths ($P = 0.01$, $t = -2.58$, $df = 101.68$), respectively. Significance was calculated using a two-sided
801 *t*-test ($n = 45$ and 64 for within- and cross-route, respectively). In the box plots, the center line
802 represents the median, whiskers represent maximum and minimum values, and box boundaries
803 represent 75th and 25th percentiles.

804 **Extended Data Fig. 6 Differences in breeding and wintering areas ($\Delta_{\text{Future-Present}}$) between**
805 **present and future (2070).** Predicted breeding (Upper) and wintering (Lower) area changes under
806 RCP 8.5 scenario, and zoomed in Kola and Europe.