Title Page

Title

Phylogenomics supports island contribution to metapopulation dynamics in a predominantly continental bird species

Short Running Title

Phylogenomics in continent-island system

Authors

AOKI Daisuke[1,*], SENZAKI Masayuki[2], ANDO Haruko[3], ODAYA Yoshiya[4], HEIM Wieland[5,6], KITAZAWA Munehiro[3,7], WULF Tom[8], TRENSE Daronja[9], BASTARDOT Marc[10], FUKUDA Atsunori[11], TAKAHASHI Masao[12,13], KONDO Natsuko[3]

[1.] Department of Wildlife Biology, Forestry and Forest Products Research Institute, Tsukuba, Japan 305-8687

[2.] Faculty of Environmental Earth Science, Hokkaido University, Sapporo, Japan, 060-0810

[3.] Biodiversity Division, National Institute for Environmental Studies, Tsukuba, Japan, 305-8506

[4.] Abiko City Museum of Birds, Abiko, Japan, 270-1145

[5.] Swiss Ornithological Institute, Sempach, Switzerland, 6204

[6.] Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland, 8006

[7.] Graduate School of Agriculture, Hokkaido University, Sapporo, Japan, 060-0809

[8.] Department of Agriculture, Ecotrophology, and Landscape Development, Anhalt University of Applied Science, Bernburg, Germany, 06406

[9.] Department of Biology, Institute of Integrated Natural Sciences, Koblenz-Landau University, Koblenz, Germany, 56070

[10.] Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland, 1015

[11.] The Wild Bird Society of Japan, Ibaraki Branch, Mito, Japan, 310-0055

[12.] Iwate Prefectural Museum, Morioka, Iwate, Japan, 020-0102

[13.] Oosekka-land, Hachinohe, Aomori, Japan, 031-0823

* Corresponding author: Daisuke AOKI, aokid@ffpri.affrc.go.jp, +81-29-829-8258

Current address

HEIM Wieland: Institute of Biology and Environmental Sciences, University of Oldenburg, Oldenburg, Germany, 26129

ODAYA Yoshiya: Chiba Natural History Museum and Institute, Chiba, Japan, 260-0852

OCRID

AOKI Daisuke: 0000-0002-5780-8487

SENZAKI Masayuki: 0000-0002-3843-5793

ANDO Haruko: 0000-0001-7827-3883

ODAYA Yoshiya: 0009-0004-1057-4005

HEIM Wieland: 0000-0002-3262-2491

KITAZAWA Munehiro: 0000-0003-0167-4464

TRENSE Daronja: 0000-0003-3598-8699

FUKUDA Atsunori: 0009-0009-2479-2886

TAKAHASHI Masao: 0000-0001-8385-6098

KONDO Natsuko: 0000-0002-0883-0506

Author Contributions

DA and MS conceived the ideas and designed the study with contributions by WH; MS, YO, WH, MK, and AF led the field work with substantial support by DA and MT; HA, NK, and DA led the laboratory procedures; DA conducted bioinformatics and analyzed the data with supports by HA and NK; DA coordinated the study, prepared the draft, led the writing and revised the manuscript with assistance from MS. All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Acknowledgements

Some tissue and blood samples were kindly provided by the Yamashina Institute for Ornithology. We thank many members of our field teams; Toshio Sadakuni for sampling in Kushiro, Akio Miya and Jun-ichi Ebina for sampling at Hotokenuma, Aomori; Sergei M. Smirenski and Yury Anisimov for sampling in Russia; Shigeo Ozawa, Ryosuke Abe, Koichi Shirakawa, Daisuke Akoshima, Ayumu Hamachi, Hiroko Shirakawa, Mikiya Oikawa, and

Shiomi Hakataya for sampling in the Kanto region. We also thank Nobuyoshi Nakajima for his support on sequencing runs on MiSeq. Our field surveys were conducted with permission from the landowners of each site, including Tomatoh. DNA samples were collected with permission from the Ministry of the Environment Government of Japan (permission # includes 21-31-0172-74, 19-129, 1810254, 2012217-16). We also acknowledge the provision of rings by the Bird Ringing Centre of Russia. Transportation of DNA samples was conducted under a Material Transfer Agreement between the National Institute for Environmental Studies and the Amur Bird Project based on the Access and Benefit-Sharing (ABS) agreement in the Nagoya Protocol. Bioinformatics was mostly performed on the supercomputer of AFFRIT, MAFF, Japan.

Funding

This study was partly conducted with the support of a grant-in-aid for Scientific Research (C) to DA (no. 22K20670) and MS (no. 23H02243) from the Japanese Society for the Promotion of Science (JSPS).

Conflict of Interest Statement

The authors declare no conflicts of interest.

Abstract

Aim:

Islands have recently been recognized as potential sources of biodiversity, challenging the traditional view that their small population sizes and low genetic diversity limit such roles. This raises the question of how insular genetic variation becomes incorporated into continental populations, contrary to expectations of unidirectional colonization. Here, we investigate whether and how island-derived genetic variation has influenced a continental population through population establishment and gene flow in a bird species where frequent trans-ocean dispersal is expected.

Location:

Continental East Asia (Russian Far East), Japanese Archipelago

Taxon:

Swinhoe's Rail (*Coturnicops exquisitus*)

Materials and Methods:

We apply integrative phylogenomics to reconstruct the spatiotemporal history of the species. Colonization sequences and gene flow are inferred by comparing four different phylogenetic reconstruction methods, using mitochondrial sequences obtained by Sanger sequencing and genome-wide data obtained by genotyping by sequencing (MIG-seq). We assess a history of colonization and gene flow based on summary statistics, demographic trajectory inference by Stairway Plot2, demographic modeling by fastsimcoal2, and species distribution modeling.

Results:

Analyses collectively supported asymmetric gene flow from the island to the continental population, following divergence around the Middle Pleistocene. Post-divergence, the island maintained a large and stable population, while the continental population underwent a severe bottleneck, suggesting a significant evolutionary role of the island for the continental population. Additionally, evidence of recent re-establishment of the island by continental individuals indicates dynamic exchange and persistence within a continent-island metapopulation.

Main conclusions:

The maintenance of insular genetic variation within a dynamic continent-island metapopulation may have enabled the island to act as a genetic and demographic reservoir for the continental population. Thus, continent-island metapopulation dynamics may be a key evolutionary pathway through which island populations contribute to continental genetic diversity.

Keywords

Coturnicops exquisitus, genetic variation, gene flow, island biogeography, metapopulation, phylogeography, reverse colonization, Swinhoe's Rail

Main Text

Introduction

Island biological diversity contributes disproportionately to global biodiversity (Whittaker and Fernández-Palacios 2007). Islands, comprising 3.5% of the terrestrial land, contribute nearly 20% of global species richness and more than 60% of terrestrial extinctions (Whittaker et al. 2017). Traditionally, islands were assumed to sustain smaller population sizes than continents due to their small and isolated landmasses (Leroy et al. 2021). Small populations are typically subjected to strong genetic drift, resulting in reduced genetic variation (Leroy et al. 2021; Cerca et al. 2023), which likely increases extinction risks (Frankham 2005) and leads to "evolutionary dead ends" (Vamosi et al. 2014). Consequently, islands have often been viewed as sinks where lineages evolve independently and do not contribute to global diversification (Bellemain and Ricklefs 2008).

Recently, evidence has emerged showing that islands can also serve as sources of continental colonization (Whittaker et al. 2017). Macroevolutionary (i.e., species-level) studies revealed numerous cases of continental species or lineages nested within island radiations–a pattern termed as "reverse colonization" (Bellemain and Ricklefs 2008). A review found that over one-third of the phylogenies across diverse taxa in various continent-island systems exhibited such patterns (Bellemain and Ricklefs 2008). Reverse colonization can also trigger mainland diversification. For example, the passerine radiation–now nearly 60% of global bird species–originated from colonization out of New Guinea (Jønsson and Holt 2015).

Although reverse colonization operates at the species level (i.e., a new continental species arises from an ancestral island species), its microevolutionary (i.e., population-level)

130 mechanisms remain poorly understood. Speciation requires several population-level processes,
131 including the establishment and persistence of populations–processes influenced by
132 spatiotemporal dynamics of migration between the ancestral and derived populations (Harvey et
133 al. 2019). For instance, depending on divergence levels, gene flow from ancestral populations
134 may lead to genetic or demographic rescue (Hufbauer et al. 2015), introduce novel genetic
135 variation (Cerca et al. 2023), or cause incompatibilities (Bank et al. 2012), thereby facilitating or
136 impeding divergence. Given the distinct population genetic processes and environments
137 experienced by island and continental populations (Barton 1996; Weigelt et al. 2013), these
138 population-level processes are likely important at the initial stage of reverse colonization. To our
139 knowledge, there has been only one study that modeled the population genetic process of
140 continental establishment by insular organisms (Patiño et al. 2015). This study, however, used
141 bryophytes, whose genetic variation may not be sensitive to demographic changes due to their
142 ability for asexual reproduction (Ellegren and Galtier 2016), leaving why many sexually
143 reproductive animals exhibit reverse colonization uncovered. Furthermore, gene flow in
144 continent-island systems has mostly been studied in the context of divergence-with-gene-flow–
145 how island populations diverge despite gene flow (Warren et al. 2015)–with a focus on genomic
146 regions resisting homogenization (Feder et al. 2012). As a result, studies have been biased
147 towards the effects of continent-derived gene flow, with limited exploration of migration
148 dynamics and their demographic consequences in continent-island systems as a whole.

149 Phylogeography is a powerful framework to examine population establishment and gene
150 flow directionality in continent-island systems, as it infers historical processes shaping present
151 genetic patterns (Lomolino et al. 2017). When integrated with demographic reconstruction and
152 species distribution modeling (SDM), it becomes especially effective in reconstructing

153 spatiotemporally explicit histories. Phylogenetic trees provide patterns that emerge as a result of
154 demographic processes, while the processes themselves remain unmodeled (Harvey et al. 2019).
155 In contrast, demographic reconstruction using high-throughput sequencing explicitly models
156 these processes (Nadachowska-Brzyska et al. 2022). SDMs estimate the timing and extent of
157 historical range shifts (Harvey et al. 2019). Therefore, these methods aid inference on the
158 direction, intensity, and timing of population establishment or gene flow.

159 Here, we apply integrative phylogeography to examine the direction and demographic
160 consequences of population establishment and gene flow in a continent-island system. We
161 selected Swinhoe's Rail (*Coturnicops exquisitus*), the smallest Asian rail, as the ideal model.
162 Demographic processes shape the present genetic variation of birds more strongly than in many
163 other taxa (Ellegren and Galtier 2016), and therefore, they are expected to underpin important
164 demographic processes in continent-island systems. Rails are known for repeated transoceanic
165 dispersals, as inferred from phylogenetic reconstruction at the species level (Garcia-R and
166 Matzke 2021). Swinhoe's Rail mainly breeds in continental East Asia (Heim et al. 2019), while a
167 limited number of island populations have recently been discovered (Senzaki et al. 2020).
168 Therefore, we expect that Swinhoe's Rail is prone to frequent colonization events across the sea,
169 even at the population level, and it is ideal to examine how a currently small island population
170 has historically contributed to the continental population establishment, persistence, and
171 evolution.

172 In this study, we reconstruct the phylogeographic history of Swinhoe's Rail to test
173 whether island genetic variation has contributed to continental populations via population
174 establishment and/or gene flow. We infer the direction of establishment by examining 1) whether
175 the insular population forms an early-branching clade with nested continental populations, 2) the

176 demographic stability of each population, and 3) habitat suitability at the time of divergence. We
177 assume that derived populations arose through bottlenecks from stable ancestral ones (e.g.,
178 Yeung et al. 2011). We determine the direction of gene flow by 1) analyzing phylogenetic
179 signals of gene flow, 2) estimating gene flow by demographic modeling, and 3) predicting
180 historical range shifts that could enable migration by the SDM. Finally, we discuss how these
181 events impacted population size and genetic variation on the continent.

182 Materials and Methods

183 *Field sampling*

184 Swinhoe's Rail is a poorly studied species. Previously known breeding ranges were limited to
185 the continent (Fig. 1a), while evidence of breeding is now absent from most of the region (Heim
186 et al. 2019). The currently recognized breeding areas were re-discovered recently, including the
187 Amur region (Heim et al. 2019) and Lake Baikal (Anisimova et al. 2019) in Russia, as well as
188 parts of Japan, namely Aomori of Honshu (Miya et al. 2005) and Hokkaido (Senzaki et al. 2020)
189 (Fig. 1a). Despite the species' cryptic behavior and preference for inaccessible wetlands, we
190 successfully developed sampling strategies (Heim et al. 2019; Senzaki et al. 2020) and collected
191 feather, buccal swab, or blood samples covering the entire extent of the currently known
192 breeding range: Amur (n = 6) and Baikal (n = 1) for continental Russia, Kushiro (n = 12) and
193 Tomakomai (n = 8) for Hokkaido, and Aomori (n = 4) for Honshu during their breeding season
194 (Fig. 1a, Table S1). We also collected samples of wintering birds in the Kanto region of Honshu
195 (n = 37) and used a genetic sample from a migratory bird in Ishikari, Hokkaido (Supplementary
196 Methods), to assess migratory connectivity and historical range shifts. Additionally, we retrieved
197 mitochondrial data from GenBank for phylogenetic analyses (Table S1).

198 *Laboratory procedures & Bioinformatics*

199 The whole genomic DNA was extracted from each tissue and used in two experiments: Sanger
200 sequencing of the partial mitochondrial Cytochrome *b* (*Cytb*) region for 49 samples, and
201 multiplexed inter-simple sequence repeats genotyping by sequencing (MIG-seq; Suyama and
202 Matsuki 2015) for 52 samples (Table S1). MIG-seq, a polymerase chain reaction-based
203 genotyping by sequencing approach, is effective for low-quality DNA at the cost of reduced
204 genome coverage and is appropriate for swab and dried feather samples. It yielded 4,834,210
205 read pairs (9,668,420 reads), which were trimmed and aligned to a reference genome of a closely
206 related species, *Laterallus jamaicensis coturniculus* (WGS Accession ID, JAKCOX01, Table
207 S1). Three individuals with over 30% missing sites in ANGSD pre-runs (one from Amur and two
208 from Kanto) were excluded (Fig. S1–S4), yielding a final dataset of 7,244,578 paired and 37,186
209 unpaired reads for 49 samples. See Supplementary Methods for parameter settings.

210 We accounted for genotype uncertainty by calculating genotype likelihoods (GLs) with
211 ANGSD (Korneliussen et al. 2014) (see Supplementary Methods, Fig. S1–S2). GLs were first
212 calculated separately for five regional populations, using half the sample size of each population
213 as '-minInd' and 2 as '-setMinDepthInd', retaining sites with depth ≥2 in over half the samples.
214 For downstream analyses, GLs were recalculated on intersecting sites among sample subsets
215 using different filtering parameters, yielding varying site counts (Table 1, S2, Fig. S4). For
216 analyses of variant sites, we set 1e-06 for '-SNP-pval' and 0.05 for '-minMaf'. For SNAPP
217 analyses, genotypes were called using '-doGeno' with a posterior cutoff of 0.9. Populations were
218 defined either by five sampling regions (Fig. 1a) or by two genetically supported breeding
219 groups–"island" (Kushiro + Tomakomai) and "continental" (Russia + Aomori) (Table 1). This
220 grouping was robustly supported by multiple analyses (Fig. 1b–c, 2b, S5–S10). The 'SAMtools'

221 model was used in the entire GL calculations. Linkage disequilibrium (LD) was assessed and
222 pruned using ngsLD (Fox et al. 2019) for some analyses (Table 1). Demographic inference used
223 folded site frequency spectra (SFS). Full workflows and parameters are provided in
224 Supplementary Methods and Fig. S4.

225 *Genetic structure*

226 We examined global genetic structure via principal component analysis (PCA) using PCAngsd
227 (Meisner and Albrechtsen 2018), ADMIXTURE analysis with NGSadmix (Skotte et al. 2013)
228 with K=1–6, NeighborNet (Bryant and Moulton 2004) based on pairwise genetic distance from
229 ngsDist (Vieira et al. 2015), and the D-statistic using 'doabbababa2' in ANGSD (Soraggi et al.
230 2018) (see Supplementary Methods for details). Two Kanto samples, closely related to one
231 Tomakomai and one Kushiro sample, showing kinship coefficients (KING, Waples et al. 2019) >
232 0.088 based on estimates from ngsRelate (Hanghøj et al. 2019), were excluded to meet the
233 assumptions of these analyses. We included two American *Laterallus* group species (*Laterallus*
234 *jamaicensis,* Accession SRX14572910; *Atlantisia rogersi,* Accession SRX6608087) as
235 outgroups in ngsDist and supplementary PCA/ADMIXTURE analyses (Fig. S5–S7). After
236 confirming their monophyly, only *L. jamaicensis* was used as an outgroup for phylogenetic
237 reconstruction.

238 *Inference of phylogenetic relationship*

239 Phylogenetic trees from different datasets and methods can reflect gene flow and incomplete
240 lineage sorting (ILS) differently (Twyford and Ennos 2012). Thus, we inferred a phylogenetic
241 relationship through four approaches: a mitochondrial phylogenetic tree (mtDNA tree) using
242 *Cytb* data, a maximum likelihood tree (RAxML tree), a graph-based population tree based on

243 allele frequency (TreeMix), and a coalescent-based population tree based on single nucleotide
244 polymorphisms (SNPs) (SNAPP tree) using MIG-seq genome-wide nuclear data.

245 **mtDNA tree:** The time-calibrated Bayesian inference was performed on 772 bps of partial *Cytb,*
246 a widely used marker suitable for avian phylogeography (e.g., Weir and Schluter 2004), using an
247 uncorrelated lognormal relaxed clock model and coalescent exponential population prior in
248 BEAST v.2.7.3 (Bouckaert et al. 2019). Samples included 12 haplotypes from 54 Swinhoe's rails
249 and three outgroups (two *C. noveboracensis*, one *L. jamaicensis*; Table S1). The standard
250 mitochondrial molecular clock rate for birds, 1.05%/lineage/Myr (Weir and Schluter 2008), was
251 applied to estimate divergence times. Markov chain Monte Carlo (MCMC) was run for 2 million
252 generations, sampling every 1,000. Convergence was checked by following Aoki et al. (2018),
253 and a maximum clade credibility tree with mean heights was drawn after discarding the first 10%
254 as burn-in. We confirmed the inferred tree by a haplotype network. See Supplementary Methods
255 for settings.

256 **RAxML tree:** A maximum likelihood tree was reconstructed from both variant and invariant
257 sites using RAxML-ng (Kozlov et al. 2019) under the 'GTGTR4+FO+G' model. Genotype
258 probabilities (GPs) calculated in a VCF file output of ANGSD were used to prepare a CATG
259 input file. For missing sites, a low, uninformative probability of 0.01 was assigned. With '--prob-
260 msa' enabled, the tree was reconstructed considering genotype uncertainty. Standard non-
261 parametric bootstrapping with 100 replicates was simultaneously conducted.

262 **TreeMix:** We used TreeMix (Pickrell and Pritchard 2012) to infer a population-level tree
263 incorporating gene flow and drift. Allele frequencies were estimated based on GLs using
264 glactools (Renaud 2017), and an ACF file was converted to TreeMix input by the 'acf2treemix'

265 program with a '--noprivate' option, to exclude population-specific alleles. Migration edges
266 ranging from zero to five were tested across ten replicates with a block size of 20 (based on the
267 LD decay; Supplementary Methods) for jackknife sampling. The best number of migration edges
268 scoring the highest Δm score was determined by an R package 'OptM' v. 0.1.6 (Fitak 2021),
269 selecting the run with the highest model likelihood as the best-fit model. We repeated the same
270 analyses without grouping (i.e., individual-level tree) to confirm that inferred gene flow reflects
271 actual migration rather than population grouping artifacts.

272 **SNAPP tree**: SNAPP infers a population tree based on a coalescent model using unlinked SNPs
273 without missing data (Bryant et al. 2012). We selected three or four samples scoring the lowest
274 missing proportion from each breeding population (three from Aomori due to the small sample
275 size), whose missing proportions ranged between 2% and 12%. Unlinked hard-called SNPs
276 without missing data were obtained from these samples by ANGSD. We estimated tree topology
277 while simultaneously estimating divergence times (Stange et al. 2018). We calibrated the split
278 between Swinhoe's Rail and *L. jamaicensis* using two log-normal priors (lognormal(5.11, 0.09)
279 and lognormal(8.95, 0.0225) in million years ago [Mya]), which corresponded to its divergence
280 time estimates by our clock-calibrated mtDNA tree (Fig. 2a) and fossil-calibrated mtDNA tree
281 by Garcia-R et al. (2020), respectively. We ran 500,000 MCMC iterations to get posterior trees
282 and analyzed the posterior trees using 'TreeSetAnalyser' to derive a 95% credible set.

283 *Demographic inferences*

284 We examined the historical demography of the continental and island populations via summary
285 statistics, demographic trajectory inferences (Stairway Plot 2, SWP2), and parameter estimations
286 in demographic models (fastsimcoal2, FSC2). These methods are complementary to each other

287 because SWP2 is a nonparametric model, flexibly inferring complex effective population size
288 (N_e) trajectories yet without modeling gene flow, while FSC2 models both gene flow and N_e
289 changes but in user-specified simpler models (Nadachowska-Brzyska et al. 2022). We also
290 estimated the contemporary N_e of the island population by converting the estimated census
291 population size (N_c) in a plain of Hokkaido in the early nineteenth century (Kitazawa et al.
292 2022), using N_c:N_e = 1:0.44, a mean ratio for birds (Frankham 1995).

293 **Summary statistics**: We calculated Watterson's theta θ_W (Watterson 1975) and Tajima's *D*
294 (Tajima 1983) per reference scaffold and population. Tajima's *D* values around zero suggest
295 demographic stability, and values below zero (especially <-1) indicate a strong ancient
296 bottleneck (Gattepaille et al. 2013). Furthermore, the large and small variance of Tajima's *D*
297 among scaffolds reflects moderate bottleneck and population expansion, respectively (Gattepaille
298 et al. 2013). Per-scaffold estimates of these indices were compared between populations using
299 linear mixed models. We also estimated the fixation index (F_{ST}) using two-dimensional SFS
300 between four breeding populations. Details in Supplementary Methods.

301 **SWP2**: We inferred demographic trajectories of the continental and island populations using
302 one-dimensional folded SFS in Stairway Plot 2 v. 2.1.1 (Liu and Fu 2020). We assumed a
303 mutation rate of 1.91e-09 bp/year (Nam et al. 2010) and converted it to bp/generation by using a
304 generation time of 1 year (a value is for its congener, *C. noveboracensis*; Alvo and Robert 1999).
305 Default settings were used. The timing at which populations started to show diverging
306 demographic trends was interpreted as a sign of population divergence.

307 **FSC2**: We tested five models differing in the presence and timing of migration matrices (Fig.
308 3c), using two-dimensional SFS. We estimated t_{MIG} at which both the continental and island

309 populations experienced demographic changes and migration matrices may have switched, by
310 fixing their divergence time at 0.5 Mya, which was roughly but collectively agreed by multiple
311 approaches (Fig. 2a, 2d, 3f). Between the two populations, we modeled no migration throughout
312 ("no migration"), migration only after ("recent migration") or before ("ancient migration") t_{MIG},
313 and migration throughout at the same rates ("constant migration") or different rates ("different
314 migration") across t_{MIG} (Fig. 3c). We conducted 100 replicates of 1 million coalescent
315 simulations across 40–100 optimization ECM cycles for parameter estimates, in fastsimcoal2 v.
316 2.7.0.9 (Excoffier et al. 2021). The same mutation rate as the SWP2 analyses was used. A
317 simulation with the smallest difference between 'MaxEstLhood' and 'MaxObsLhood' was
318 selected from each model, and models with the ΔAIC (delta Akaike Information Criterion, the
319 difference between the lowest AIC and an AIC in focus) <2 were considered best-fit. The ranges
320 of parameter values for simulation are in Table S5.

321 *Species distribution modeling*

322 SDMs were run using MaxEnt (Phillips et al. 2006) to model the species' breeding extent. We
323 compiled 48 breeding and 111 non-breeding records from previously published sources (Table
324 S6), including both previously known and recently discovered sites (Fig. S20). After spatial
325 thinning (one record per 0.5º grid cell), 33 breeding points were retained as an occurrence input
326 (Fig. 4b). Background points were sampled to cover one-fourth of the total number of girds in a
327 1,000-km buffered convex hull around the full dataset (both breeding and non-breeding records),
328 resulting in 2,104 points (Fig. S20). We extracted time slices of 22 environmental variables
329 (Krapp et al. 2021) using an R package 'pastclim' v. 1.2.3 (Leonardi et al. 2023) and selected
330 eight covariates with low collinearity, satisfying that any pair scores variance inflation factor
331 <10. Twelve models with different parameter settings were tested via 'ENMevaluate' in an R

332 package ENMeval v. 2.0.4 (Kass et al. 2021), and the best model was selected based on the
333 lowest ΔAIC for a small sample size (ΔAICc). The best-fit model was projected across time
334 slices from the present and 0.5 Mya at 1,000-year intervals (Fig. 4c). We compared the historical
335 changes in the potential habitat of this species on the island and the continent (Fig. S23), by
336 calculating three indices, habitat suitability (sum of values of the predicted presence grids),
337 availability (the total number of predicted presence grids), and quality (suitability divided by
338 availability). We also calculated centroid distances (weighted by suitability) between island and
339 continental habitats to track historical changes in the level of geographic isolation. Effects of
340 negative multivariate environmental similarity surface (MESS) were controlled for these
341 calculations. See Supplementary Methods for details.

342 Results

343 *Phylogenetic relationships and gene flow*

344 Multiple phylogenetic methods collectively inferred population divergence. A bifurcation
345 between an island clade (Tomakomai and Kushiro) and a continental clade (Russia and Aomori)
346 was supported by TreeMix (Fig. 2c), SNAPP tree (Fig. 2d), and the D-statistic (Fig. S10). This
347 structure was further supported by ADMIXTURE at K = 2 that scored the highest ΔK value (Fig.
348 1c, S6–7), the first axis of PCA explaining >99% of variation (Fig. 1b, S5), NeighborNet (Fig.
349 S8), and F_{ST} (Fig. S9). On the other hand, some phylogenies indicated ILS and/or post-
350 divergence gene flow. In the mtDNA tree, divergence between the two clades was supported,
351 while Russian samples appeared in both clades, and most Tomakomai and Kushiro samples
352 clustered in the island clade (except one from Kushiro) (Fig. 2a, S11). This pattern suggests
353 either ILS due to the island population being derived from a large continental population or post-

354 divergence gene flow from the island to the continental clade. Low bootstrap supports in the
355 RAxML tree (Fig. 2b, S12–13) and low posterior probabilities for a bifurcation tree (41.5%) in
356 the SNAPP tree (Fig. 2d) can also reflect this uncertainty. The divergence time of the two clades
357 was estimated to be relatively old, around 0.5 Mya (ranging between 0.2–1.0 depending on
358 analyses), corresponding to the Middle Pleistocene (the Chibanian epoch), collectively supported
359 by different calibration methods, data, and algorithms, including the mtDNA tree (Fig. 2a), the
360 SNAPP tree (Fig. 2d), and SWP2 (Fig. 3f).

361 Gene flow was also inferred by the best-fit population-level TreeMix tree with one
362 migration edge scoring the highest Δm value (Fig. 2c, S14, S16), which was also supported by
363 our individual-level tree (Fig. 2c, S15), indicating that gene flow inferred in our population-level
364 analyses is not the artifact of population grouping (see Supplementary Discussion, Fig. S16–
365 S19). Meanwhile, gene flow was directed from the common ancestor of this species to Russia.
366 On our individual-level tree, the Tomakomai samples were located at the basal position. This
367 suggests that ancient variation may have been introduced to the Russian population via gene flow
368 from a population preserving ancient genetic variation, such as the Tomakomai population
369 (Supplementary Discussion, Fig. S17–S19). This gene flow possibly occurred relatively recently
370 (e.g., post-expansion of the continental population), as migration was not towards the ancestor of
371 the continental cluster (Fig. 2c), but to Russia (at the population-level) or to an ancestor of one
372 Aomori and one Russian sample (at the individual-level). One Russian individual clustered with
373 island samples (Fig. 2c, inset), which was inferred to have the island cluster in our
374 ADMIXTURE analysis at K = 3 (Fig. 1c), also supporting recent gene flow.

375 Aomori samples consistently clustered with and showed low genetic differentiation from
376 Russian samples in all methods–phylogenetic trees (Fig. 2), PCA, ADMIXTURE (Fig. 1b–c),

377 NeighborNet (Fig. S8), F_{ST} (Fig. S9), and the D-statistic (Fig. S10)–suggesting recolonization of
378 Aomori by the continental population. This was unlikely to be a sample size artifact, as
379 supported patterns held across genetic distance (NeighborNet) and F_{ST} that consider sample size
380 and were robust to excluding Kanto wintering samples (i.e., smaller effect of sample size) (Fig.
381 1c, S5–S7). Additionally, the Baikal sample, collected ~1,500 km from Amur, showed no major
382 differentiation, despite the limited sample size. Together with the genetic position of Aomori, the
383 present range of the continental population is a result of a recent expansion.

384 *Historical demography and contemporary genetic variation*

385 The continental population had marginally higher genetic diversity than the island population (p
386 = 0.089; Fig. 3a, Table S3). FSC2 results suggested that gene flow from the island contributed
387 significantly to contemporary continental diversity. The best-fit FSC2 model (ΔAIC <2.0) was
388 "different migration" model (ΔAIC = 0.0, Akaike model weight = 0.747), with "constant
389 migration" as the second-best yet with a marginal ΔAIC value (ΔAIC = 2.17, Akaike model
390 weight = 0.253). Both supported post-last-glacial-maximum (LGM, ~0.02 Mya) gene flow (t_{MIG}
391 = 0.014–0.019 Mya) from the island to the continent, at rates 1.8–2.7 times higher than in the
392 opposite direction (Fig. 3c–e, Table S4). In contrast, gene flow before the LGM was inconclusive
393 due to significant differences between the two models.

394 Our demographic analyses suggested that the island population remained relatively
395 stable, or, if any, experienced a modest level of expansion since the population split until the
396 LGM, supported by Tajima's D values close to zero (Fig. 3b, Table S3), and FSC2 (0.5–2.3-fold
397 change), and SWP2 (~1.9-fold change) (Fig. 3d–e, Table S4). Meanwhile, different methods
398 showed discordant results for the demographic history of the continental population. Expansion

399 after a strong ancient population bottleneck was suggested by the significantly low value of
400 Tajima's *D* close to -1 (median value = -0.845) with a smaller inter-scaffold variance than the
401 island population (Fig. 3b) and the best-fit FSC2 models (nearly 98% population decline, Fig.
402 3d–e), while expansion soon after the population split around 0.5 Mya without a population
403 bottleneck was suggested by SWP2.

404 A sharp post-LGM bottleneck was suggested for the island population (97–99%
405 reduction; Fig. 3d–e) by the best-fit FSC2 models. Our SWP2 results did not contradict it despite
406 a large uncertainty towards the present (Fig. 3f). This was also supported by the small
407 contemporary effective population size of Ishikari Plain in Hokkaido (including Tomakomai) (N_e
408 = 4,840) calculated from the census population size estimates around the early nineteenth
409 century, which well corroborated the results of the FSC2.

410 *SDMs*

411 The best-fit model (response curve 'LQH' + regularization multiplier 2.0, AICc = 515.86,
412 Akaike model weight = 0.77, the Area Under the Curve (AUC) = 95.2%) accurately predicted
413 the present breeding ranges of Swinhoe's Rail (Fig. 4a–b, S21). During historical periods, some
414 areas did not resemble any climate in the modeling domain of this species, indicated by the
415 negative values in the MESS analyses (dark gray-shaded grids in Fig. 4a, S22, see
416 Supplementary Methods). We restricted our interpretations to results robust to different
417 treatments (inclusion or removal) of negative MESS grids.

418 The island served as a relatively stable glacial refugium, albeit smaller in extent than the
419 continent. Continental habitats underwent drastic and repeated changes in habitat suitability and
420 availability since 0.5 Mya (Fig. 4d). Habitat quality remained consistently higher on the island

421 than on the continent throughout this period (Fig. S24). After the LGM, the island habitat
422 contracted, corresponding to both SWP2 and FSC2, whereas the continental habitat expanded,
423 only supporting the post-LGM demography of FSC2 (Fig. 4d, S22, S24). The temporal change in
424 the level of habitat isolation between the island and the continent well supported the times of
425 divergence and gene flow estimated by our genetic analyses. The 0.38–0.43 Mya interval showed
426 stable isolation without periods of extreme proximity between the two regions, corresponding to
427 the divergence period (Fig. 4e, S22, S25). The LGM was the most recent period of reduced
428 isolation, consistent with post-LGM gene flow. Overall trends in these measures were similar
429 between different treatments of negative MESS grids, although the amplitude of their changes
430 differed between each other (Supplementary Results and Discussion, Fig. S24–S25).

431 Discussion

432 Using an integrative phylogeography in Swinhoe's Rail, we found that the contemporary
433 continental population exhibited larger effective population size than the island population,
434 consistent with general expectations (Leroy et al. 2021). However, multiple lines of evidence
435 suggested that, historically, the island population was more stable and possibly larger than the
436 continent, and that gene flow occurred from the island to the continent. These findings suggest
437 that the island served as an important refugium for this species. In contrast, the direction of
438 population establishment at the time of divergence remains unclear. Some phylogenetic trees
439 supported bifurcating topologies at the population level, but individual-level analyses did not
440 support reciprocal monophyly, leaving colonization direction unresolved.

441 *Gene flow and incomplete lineage sorting in the continent-island system*

442 Incongruent phylogenetic patterns across analyses suggest a complex history of this
443 species in the continent-island system, involving two non-mutually exclusive mechanisms, ILS
444 and post-divergence gene flow. In recently diverged populations like Swinhoe's Rail, gene
445 genealogies may remain shared between two geographic populations (Twyford and Ennos 2012).
446 For example, higher uncertainties were involved in the inference of the ancestral divergence and
447 gene flow, including low posterior supports for bifurcation in the SNAPP tree and lack of
448 reciprocal monophyly in the individual-level trees. These may reflect ILS, low marker
449 resolution, or a combination of the two (Maddison and Knowles 2006), where ILS, derived from
450 ancestral genetic variation, can obscure ancient evolutionary events (Twyford and Ennos 2012).
451 Therefore, ancient gene flow inferred by FSC2 may be confounded with ILS, and thus its
452 presence remains inconclusive. Nevertheless, the absence of monophyly due to ILS does not
453 negate the existence of the two genetic clusters (i.e., continental and island populations),
454 especially given multiple lines of support, and ILS is commonly retained to various degrees
455 across populations and species (Maddison and Knowles 2006).

456 By contrast, recent gene flow is more readily distinguished from ILS (Twyford and
457 Ennos 2012), and multiple results are consistent with this notion, supporting the presence of
458 post-LGM gene flow from the island to the continent. First, the continental population showed
459 signatures of strong historical bottlenecks–e.g., low Tajima's *D*, FSC2 results, and SDM-inferred
460 habitat instability–reducing the likelihood of retained ancestral variation. This is because a strong
461 reduction in population size likely causes a fixation of genetic variants. Given the historically
462 stable and relatively large island population, the island population may have preserved ancient
463 genetic variations, causing ILS (see Supplementary Results & Discussion). Second, two-

464 dimensional-SFS-based approaches, such as FSC2, and TreeMix– both of which explicitly
465 account for ILS by estimating the level of genetic drift (Pickrell and Pritchard 2012; Murray et
466 al. 2025)–consistently inferred gene flow to the continent. Therefore, while the relative
467 contribution of gene flow versus ILS to the continental genetic variation remains unclear, our
468 findings at least support the presence of recent gene flow from the island to the continent.

469 Discrepancies in population size histories between FSC2 and SWP2 may also reflect this
470 gene flow. In demographic trajectory inference methods, such as SWP2, migration can inflate
471 effective population size estimates, even without true expansion (Nadachowska-Brzyska et al.
472 2022). Therefore, the genetic bottleneck inferred in FSC2 might have been concealed in SWP2
473 by gene flow to the continental population. Alternatively, our simplified FSC2 models may have
474 failed to capture complex fluctuating histories and rather showed averaged trends (long-term N_e,
475 Nadachowska-Brzyska et al. 2022), which may have exaggerated the impact of a bottleneck.
476 However, because a long-term N_e reflects a harmonic mean of a demographic trajectory
477 (Nadachowska-Brzyska et al. 2022), the effect of past bottlenecks on the continental genetic
478 variation should not be undervalued.

479 *Population persistence through a continent-island metapopulation*

480 Historically large island populations and gene flow from islands have rarely been expected in
481 continent-island systems. This is because small island size is typically assumed to limit
482 population size (Leroy et al. 2021), leading to low dispersal potential and resultant gene flow
483 (Poethke and Hovestadt 2001; Matthysen 2005). However, based on our FSC2 estimates,
484 ancestral island population retained a large N_e that was comparable to that of avian continental
485 populations (mean N_e = 362,456) (Leroy et al. 2021) and large tropical island populations, such

486 as New Guinean birds (mean N_e = 325,000) (Pujolar et al. 2022), whereas the N_e of the ancestral
487 continental population was as small as that of oceanic island populations (mean N_e = 94,944)
488 (Leroy et al. 2021). These differences may reflect contrasting habitat stability; our SDMs
489 suggested more stable conditions on the island throughout glacial cycles than on the continent
490 (Fig. 4d, S24), which could have allowed the persistence of a stable population and promoted
491 dispersal (Ronce et al. 2000), leading to gene flow to the continent. In a large population,
492 adaptive or ancient variation can be more likely to be preserved than in a small population where
493 weakly deleterious mutations accumulate (Leroy et al. 2021; Lai et al. 2019). Asymmetric gene
494 flow from such a large population may restore genetic variation and result in demographic and
495 evolutionary consequences for a small population (Fitzpatrick et al. 2016). Therefore, the island
496 population could have potentially contributed to the evolution and persistence of the continental
497 population in this species. However, it remains unknown to what extent gene flow has
498 contributed to the continental expansion, as the signature of gene flow was not found originating
499 from the entire continental population (Fig. 2c).

500 Aomori of Honshu Island was inferred to have been recolonized recently in the wave of
501 continental expansion, suggesting the possibility that the island population has also been
502 sustained by migrants from the continent. Honshu might have been occupied previously by the
503 present Hokkaido ("island") population during a glacial period, supported by 1) the potential
504 presence of suitable habitats on Honshu but not on Hokkaido during some glacial maxima
505 inferred by the SDM (Fig. S22), and 2) a high affinity of Kanto samples towards Hokkaido
506 breeding birds (Fig. 1b–c), given that wintering regions of short-distance seasonal migrants may
507 reflect the past breeding range (Zink and Gardner 2017). Indeed, the population contraction of
508 Swinhoe's Rail in Japan reflects not only the post-LGM environmental changes but also

509 anthropogenic factors, such as habitat loss and fragmentation due to land use changes (Kitazawa
510 et al. 2022). This was also supported by our SDMs that predicted present suitable areas larger
511 than currently recognized (Fig. 4a–b) and that the current Aomori habitat originated from an
512 abandoned rice paddy (Ministry of the Environment, 2023). Hence, continental immigration
513 resulting in re-establishment rather than gene flow may reflect the extinction of local island
514 populations under anthropogenic impacts.

515 Altogether, our findings suggest that Swinhoe's Rail has persisted as a dynamic
516 continent-island metapopulation, where the Japanese Archipelago was not merely a one-way sink
517 but played a role as pivotal as the continent, functioning as potentially a demographic and
518 genetic source. Following divergence during a period of prolonged geographic isolation (~0.5
519 Mya; Fig. 5a), spatiotemporal climatic heterogeneity likely led the island and the continent to
520 shift roles within the metapopulation: the island supported relatively a large, stable population
521 that retained ancient genetic variation, while the continent underwent drastic and repeated
522 demographic changes that resulted in smaller long-term population size (Fig. 5b–c). Source-sink
523 roles may have shifted over time via gene flow and population re-establishment, potentially
524 promoting persistence of the metapopulation in the evolutionary timescale (Fig. 5b–c). Our
525 findings were different from many previous studies that found islands to harbor small genetic
526 variation and N_e and act only as recipients of continental migrants. This is possibly because these
527 patterns were often observed in systems with strong phenotypic or genomic divergence on
528 islands (e.g., Wang et al. 2016; Delmore et al. 2020; Recuerda et al. 2021), where post-
529 divergence gene flow from islands may be prevented or overlooked. In contrast, studies have
530 reported that multiple island populations are maintained with gene flow on archipelagic or land-
531 bridge islands (Seki et al. 2007; Mapel et al. 2021), and our results may be an expansion of such

532 systems to the continent-island setting. Some Japanese endemic lineages have extended their
533 ranges to the continental Russian Far East (e.g., *Gallinago hardwickii*, Nechaev and Fujimaki
534 1998) or Korean Peninsula (e.g., *Cettia diphone*, Clement and Kirwan 2020). Gene flow from
535 Japanese lineages to their continental counterparts has also been suspected, yet without explicit
536 inferences of the underlying historical processes (*Locustella amnicola* and *L. fasicolata*,
537 Drovetski et al. 2015; *Lanius cristatus*, Aoki et al. 2021). Similar continent-island systems are
538 recognized in other regions (Bellemain et al. 2008; Garcia-R et al. 2017), suggesting the
539 universality of such a system across the globe.

540 A recent review emphasizes the importance of long-term population persistence in
541 dynamic metapopulations for promoting speciation and adaptive radiation (Harvey et al. 2019).
542 Our view also emphasizes the long-term consequences of metapopulation between continental
543 and island populations, where source and sink relationships dynamically change, differently from
544 the traditional perspective that focuses on the persistence of a metapopulation in the ecological
545 timescale (Hanski 1998). A metapopulation structure across the continent-island system can be
546 perceived as a transitional state for insular populations to undergo reverse colonization, thereby
547 promoting long-term persistence and adaptation to the continental environment. Therefore, we
548 highlight that our insights obtained from the population-level study may shed light on the origin
549 of reverse colonization at the species level. Future research needs to incorporate more
550 polymorphic sites to resolve ILS and gene flow and estimate the number, direction, and intensity
551 of past migration, to understand the relative impact of these elements on the present genetic
552 variation. This would disentangle the evolutionary importance of how a continent-island
553 metapopulation may facilitate expansion, adaptation, and diversification on the continent,
554 extending our microevolutionary idea to a macroevolutionary framework.

555 *Conclusion*

556 Historical demography is fundamental to understanding a determinant of population genetic
557 structure that is significantly linked with subsequent speciation and diversification (Harvey et al.
558 2019). In this study, we reconstructed the demographic history of a rail species in a continent-
559 island system. We found strong evidence of asymmetrical gene flow from the Japanese
560 Archipelago to the continental population, which may have influenced the continental genetic
561 variation and demography. Furthermore, our results suggested that this species has persisted as a
562 metapopulation spanning both the islands and the continent. These findings challenge
563 conventional views of continent-island dynamics, particularly given that breeding populations of
564 this species in Japan were only recently documented (Senzaki et al. 2020). To fully understand
565 the processes that have given rise to the present biodiversity, evolutionary studies on the current
566 and apparent species distribution are insufficient. Rather, reconstructing the long-term, spatially
567 explicit population histories of organisms can shed light on the previously unfilled gap between
568 micro- and macroevolutionary concepts and facilitate our understanding of global biodiversity
569 patterns.

References

Alvo R., and M. Robert. 1999. COSEWIC status report on the yellow rail *Coturnicops noveboracensis* in Canada. Committee on the Status of Endangered Wildlife in Canada, Ottawa.

Anisimova, V., Y. Anisimov, M. Bastardot, and W. Heim. 2019. First observation of Swinhoe's Rail *Coturnicops exquisitus* in the delta of Selenga river. Baikal Zoological Journal 2:112–113.

Aoki, D., G. Kinoshita, A. P. Kryukov, I. Nishiumi, S.-I. Lee, and H. Suzuki. 2018. Quaternary-related genetic differentiation and parallel population dynamics of the Eurasian Jay (*Garrulus glandarius*) in the circum-Japan Sea region. J. Ornithol. 159:1087–1097.

Aoki, D., H. Sakamoto, M. Kitazawa, A. P. Kryukov, and M. Takagi. 2021. Migration-tracking integrated phylogeography supports long-distance dispersal-driven divergence for a migratory bird species in the Japanese archipelago. Ecol. Evol. 11:6066–6079.

Bank C., R. Bűrger, and J. Hermisson. (2012) The limits to parapatric speciation: Dobzhansky-Muller incompatibilities in a continent-island model. Genetics, 191:845–863.

Barton N. H. 1996. Natural selection and random genetic drift as causes of evolution on islands. Phil. Trans. R. Soc. Lond. B, 351:785–795.

Bellemain, E., E. Bermingham, and R. E. Ricklefs. 2008. The dynamic evolutionary history of the bananaquit (*Coereba flaveola*) in the Caribbean revealed by a multigene analysis. BMC Evol. Biol. 8:240.

Bellemain, E., and R. E. Ricklefs. 2008. Are islands the end of the colonization road? Trends Ecol. Evol. 23:461–468.

592 BirdLife International, and Handbook of the Birds of the World. 2024. Bird species distribution
593 maps of the world. Version 2024.2. Available at
594 http://datazone.birdlife.org/species/requestdis.
595 Bouckaert, R., T. G. Vaughan, J. Barido-Sottani, S. Duchêne, M. Fourment, A. Gavryushkina, J.
596 Heled, G. Jones, D. Kühnert, N. De Maio, M. Matschiner, F. K. Mendes, N. F. Müller, H. A.
597 Ogilvie, L. du Plessis, A. Popinga, A. Rambaut, D. Rasmussen, I. Siveroni, M. A. Suchard,
598 C.-H. Wu, D. Xie, C. Zhang, T. Stadler, and A. J. Drummond. 2019. BEAST 2.5: An
599 advanced software platform for Bayesian evolutionary analysis. PLoS Comput. Biol.
600 15:e1006650.
601 Bryant, D., R. Bouckaert, J. Felsenstein, N. A. Rosenberg, and A. RoyChoudhury. 2012.
602 Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full
603 coalescent analysis. Mol. Biol. Evol. 29:1917–1932.
604 Bryant, D., and V. Moulton. 2004. Neighbor-net: an agglomerative method for the construction
605 of phylogenetic networks. Mol. Biol. Evol. 21:255–265.
606 Cerca, J., D. D. Cotoras, V. C. Bieker, R. De-Kayne, P. Vargas, M. Fernández-Mazuecos, J.
607 López-Delgado, O. White, M. Stervander, A. J. Geneva, J. E. Guevara Andino, J. I. Meier,
608 L. Roeble, B. Brée, J. Patiño, J. M. Guayasamin, M. de L. Torres, H. Valdebenito, M. D. R.
609 Castañeda, J. A. Chaves, P. J. Díaz, L. Valente, M. L. Knope, J. P. Price, L. H. Rieseberg, B.
610 G. Baldwin, B. C. Emerson, G. Rivas-Torres, R. Gillespie, and M. D. Martin. 2023.
611 Evolutionary genomics of oceanic island radiations. Trends Ecol. Evol. 38:631–642.
612 Clement, P. and G. M. Kirwan. 2020. Japanese Bush Warbler (*Horornis diphone*), version 1.0. In
613 del Hoyo, J., A. Elliott, J. Sargatal, D. A. Christie, and E. de Juana eds. Birds of the World.
614 Cornell Lab of Ornithology, Ithaca, NY, USA.

615 Delmore, K., J. C. Illera, J. Pérez-Tris, G. Segelbacher, J. S. Lugo Ramos, G. Durieux, J.
616 Ishigohoka, and M. Liedvogel. 2020. The evolutionary history and genomics of European
617 blackcap migration. Elife 9:e54462.
618 Drovetski, S. V., G. Semenov, Y. A. Red'kin, V. N. Sotnikov, I. V. Fadeev, and E. A. Koblik.
619 2015. Effects of asymmetric nuclear introgression, introgressive mitochondrial sweep, and
620 purifying selection on phylogenetic reconstruction and divergence estimates in the pacific
621 clade of *Locustella* warblers. PLoS One. 10:e0122590.
622 Ellegren, H., and N. Galtier. 2016. Determinants of genetic diversity. Nat. Rev. Genet. 17:422–
623 433.
624 Excoffier, L., N. Marchi, D. A. Marques, R. Matthey-Doret, A. Gouy, and V. C. Sousa. 2021.
625 fastsimcoal2: demographic inference under complex evolutionary scenarios. Bioinformatics
626 37:4882–4885.
627 Feder, J. L., S. P. Egan, and P. Nosil. 2012. The genomics of speciation-with-gene-flow. Trends
628 Genet. 28:342–350.
629 Fitak, R. R. 2021. OptM: estimating the optimal number of migration edges on population trees
630 using Treemix. Biol Methods Protoc 6:bpab017.
631 Fitzpatrick, S. W., J. C. Gerberich, L. M. Angeloni, L. L. Bailey, E. D. Broder, J. Torres-
632 Dowdall, C. A. Handelsman, A. López-Sepulcre, D. N. Reznick, C. K. Ghalambor, and W.
633 Chris Funk. 2016. Gene flow from an adaptively divergent source causes rescue through
634 genetic and demographic factors in two wild populations of Trinidadian guppies. Evol.
635 Appl. 9:879–891.
636 Fox, E. A., A. E. Wright, M. Fumagalli, and F. G. Vieira. 2019. ngsLD: evaluating linkage
637 disequilibrium using genotype likelihoods. Bioinformatics 35:3855–3856.

638 Frankham, R. 1995. Effective population size/adult population size ratios in wildlife: a review.
639 Genet. Res. 66:95–107.
640 Frankham, R. 2005. Genetics and extinction. Biol. Conserv. 126:131–140.
641 Garcia-R, J. C., L. Joseph, G. Adcock, J. Reid, and S. A. Trewick. 2017. Interisland gene flow
642 among populations of the buff-banded rail (Aves: Rallidae) and its implications for insular
643 endemism in Oceania. J. Avian Biol. 48:679–690.
644 Garcia-R, J. C., E. M. Lemmon, A. R. Lemmon, and N. French. 2020. Phylogenomic
645 reconstruction sheds light on new relationships and timescale of rails (Aves: Rallidae)
646 evolution. Diversity 12:70.
647 Garcia-R, J. C., and N. J. Matzke. 2021. Trait-dependent dispersal in rails (Aves: Rallidae):
648 Historical biogeography of a cosmopolitan bird clade. Mol. Phylogenet. Evol. 159:107106.
649 Gattepaille, L. M., M. Jakobsson, and M. G. B. Blum. 2013. Inferring population size changes
650 with sequence and SNP data: lessons from human bottlenecks. Heredity 110:409–419.
651 Hanghøj K., I. Moltke, P. A. Andersen, A. Manica, T. S. Korneliussen. 2019. Fast and accurate
652 relatedness estimation from high-throughput sequencing data in the presence of inbreeding.
653 GigaScience 8:giz034
654 Hanski, I. 1998. Metapopulation dynamics. Nature 396:41–49.
655 Harvey, M. G., S. Singhal, and D. L. Rabosky. 2019. Beyond reproductive isolation:
656 demographic controls on the speciation process. Annu. Rev. Ecol. Evol. Syst. 50:75–95.
657 Heim, W., D. Trense, A. Heim, and J. Kamp. 2019. Discovery of a new breeding population of
658 the vulnerable Swinhoe's Rail *Coturnicops exquisitus* confirmed by genetic analysis. Bird
659 Conserv. Int. 29:454–462.
660 Hufbauer, R. A., M. Szűcs, E. Kasyon, C. Youngberg, M. J. Koontz, C. Richards, T. Tuff and B.

661 A. Melbourne. 2015. Three types of rescue can avert extinction in a changing environment.
662 Proc. Natl. Acad. Sci. U.S.A. 112:10557–10562.
663 Jønsson, K. A., and B. G. Holt. 2015. Islands contribute disproportionately high amounts of
664 evolutionary diversity in passerine birds. Nat. Commun. 6:8538.
665 Kass, J. M., R. Muscarella, P. J. Galante, C. L. Bohl, G. E. Pinilla-Buitrago, R. A. Boria, M.
666 Soley-Guardia, and R. P. Anderson. 2021. ENMeval 2.0: Redesigned for customizable and
667 reproducible modeling of species' niches and distributions. Methods Ecol. Evol. 12:1602–
668 1608.
669 Kitazawa, M., Y. Yamaura, M. Senzaki, M. Hanioka, H. Ohashi, M. Oguro, T. Matsui, and F.
670 Nakamura. 2022. Quantifying the impacts of 166 years of land cover change on lowland
671 bird communities. Proc. Biol. Sci. 289:20220338.
672 Korneliussen, T. S., A. Albrechtsen, and R. Nielsen. 2014. ANGSD: Analysis of Next
673 Generation Sequencing Data. BMC Bioinformatics 15:356.
674 Kozlov, A. M., D. Darriba, T. Flouri, B. Morel, and A. Stamatakis. 2019. RAxML-NG: a fast,
675 scalable and user-friendly tool for maximum likelihood phylogenetic inference.
676 Bioinformatics 35:4453–4455.
677 Krapp, M., R. M. Beyer, S. L. Edmundson, P. J. Valdes, and A. Manica. 2021. A statistics-based
678 reconstruction of high-resolution global terrestrial climate for the last 800,000 years. Sci
679 Data 8:228.
680 Lai, Y., C. K. L. Yeung, K. E. Omland, E. Pang, Y. Hao, B. Liao, H. Cao, B. Zhang, C. Yeh, C.
681 Hung, H. Hung, M. Yang, W. Liang, Y. Hsu, C. Yao, L. Dong, K. Lin, & S. Li. 2019.
682 Standing genetic variation as the predominant source for adaptation of a songbird, Proc.
683 Natl. Acad. Sci. U.S.A. 116:2152–2157.

Leonardi, M., E. Y. Hallett, R. Beyer, M. Krapp, and A. Manica. 2023. pastclim 1.2: an R package to easily access and use paleoclimatic reconstructions. Ecography 2023:e06481.

Leroy, T., M. Rousselle, M.-K. Tilak, A. E. Caizergues, C. Scornavacca, M. Recuerda, J. Fuchs, J. C. Illera, D. H. De Swardt, G. Blanco, C. Thébaud, B. Milá, and B. Nabholz. 2021. Island songbirds as windows into evolution in small populations. Curr. Biol. 31:1303–1310.e4.

Lisiecki, L. E., and M. E. Raymo. 2005. A Pliocene-Pleistocene stack of 57 globally distributed benthic $\delta^{18}O$ records. Paleoceanography 20:PA1003. American Geophysical Union (AGU).

Liu, X., and Y.-X. Fu. 2020. Stairway Plot 2: demographic history inference with folded SNP frequency spectra. Genome Biol. 21:280.

Lomolino, M. V., B. R. Riddle, and R. J. Whittaker. 2017. Biogeography: Biological Diversity across Space and Time. 5th Edition. Sinauer Associates, Maryland.

Maddison W. P., and L. L. Knowles. 2006. Inferring phylogeny despite incomplete lineage sorting. System. Biol. 55:21–30.

Mapel, X. M., E. F. Gyllenhaal, T. H. Modak, L. H. DeCicco, A. Naikatini, R. B. Utzurrum, J. O. Seamon, A. Cibois, J.-C. Thibault, M. D. Sorenson, R. G. Moyle, L. N. Barrow, and M. J. Andersen. 2021. Inter- and intra-archipelago dynamics of population structure and gene flow in a Polynesian bird. Mol. Phylogenet. Evol. 156:107034.

Matthysen, E. 2005. Density-dependent dispersal in birds and mammals. Ecography 28:403–416.

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

Ministry of the Environment, Government of Japan. 2023. Hotokenuma. Available at https://www.env.go.jp/nature/ramsar/conv/ramsarleaflet/14_Hotokenuma.pdf. Accessed on June 27th, 2025.

Miya, A., S. Sannohe, J. Ebina, and H. Sekishita. 2005. A survey report on Yellow Rail (*Coturnicops noveboracensis exquisitus*) observed during the breeding season at Hotokenuma Reclaimed Land in Misawa, Aomori Prefecture. Jap. J. Ornithol. 54:104–107.

Murray, C. S., M. Karram, D. J. Bass, M. Doceti, D. Becker, J. C. B. Nunez, A. Ratan, A. O. Bergland. 2025. Trans-specific polymorphisms between cryptic *Daphnia* species affect fitness and behavior. 34:e17632

Nadachowska-Brzyska, K., M. Konczal, and W. Babik. 2022. Navigating the temporal continuum of effective population size. Methods Ecol. Evol. 13:22–41.

Nam, K., C. Mugal, B. Nabholz, H. Schielzeth, J. B. W. Wolf, N. Backström, A. Künstner, C. N. Balakrishnan, A. Heger, C. P. Ponting, D. F. Clayton, and H. Ellegren. 2010. Molecular evolution of genes in avian genomes. Genome Biol. 11:R68.

Nechaev, V. A., and Y. Fujimaki. 1998. Present status of the Latham's snipe *Gallinago hardwickii* on Sakhalin. Res. Bull. Obihiro. Univ. 21:61–65.

Patiño, J., M. Carine, P. Mardulyn, N. Devos, R. G. Mateo, J. M. González-Mancebo, A. J. Shaw, and A. Vanderpoorten. 2015. Approximate Bayesian computation reveals the crucial role of oceanic islands for the assembly of continental biodiversity. Syst. Biol. 64:579–589.

Poethke H. J., and T. Hovestadt. 2001. Evolution of density- and patch-size-dependent dispersal rates. Proc. R. Soc. Lond. B. 269:637–645.

Phillips, S. J., R. P. Anderson, and R. E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. Ecol. Modell. 190:231–259.

Pickrell, J. K., and J. K. Pritchard. 2012. Inference of population splits and mixtures from genome-wide allele frequency data. PLoS Genet. 8:e1002967.

Pujolar, J. M., M. P. K. Blom, A. H. Reeve, J. D. Kennedy, P. Z. Marki, T. S. Korneliussen, B.

G. Freeman, K. Sam, E. Linck, T. Haryoko, B. Iova, B. Koane, G. Maiah, L. Paul, M. Irestedt, and K. A. Jønsson. 2022. The formation of avian montane diversity across barriers and along elevational gradients. Nat. Commun. 13:268.

Recuerda, M., J. C. Illera, G. Blanco, R. Zardoya, and B. Milá. 2021. Sequential colonization of oceanic archipelagos led to a species-level radiation in the common chaffinch complex (Aves: *Fringilla coelebs*). Mol. Phylogenet. Evol. 164:107291.

Renaud, G. 2017. glactools: a command-line toolset for the management of genotype likelihoods and allele counts. Bioinformatics 34:1398–1400.

Ronce, O., F. Perret, and I. Olivieri. 2000. Evolutionarily stable dispersal rates do not always increase with local extinction rates. Am. Nat. 155:485‑96.

Seki, S.-I., H. Takano, K. Kawakami, N. Kotaka, A. Endo, and K. Takehara. 2007. Distribution and genetic structure of the Japanese wood pigeon (*Columba janthina*) endemic to the islands of East Asia. Conserv. Genet. 8:1109–1121.

Senzaki, M., M. Kitazawa, T. Sadakuni, and M. Takahashi. 2020. Breeding evidence of the vulnerable Swinhoe's Rail (*Coturnicops exquisitus*) in Japan. Wilson J. Ornithol. 132:711–717.

Skotte, L., T. S. Korneliussen, and A. Albrechtsen. 2013. Estimating individual admixture proportions from next generation sequencing data. Genetics 195:693–702.

Soraggi, S., C. Wiuf, and A. Albrechtsen. 2018. Powerful inference with the D-statistic on low-coverage whole-genome data. G3 8:551–566.

Spratt, R. M., and L. E. Lisiecki. 2016. A Late Pleistocene sea level stack. Clim. Past 12:1079–1092.

Stange, M., M. R. Sánchez-Villagra, W. Salzburger, and M. Matschiner. 2018. Bayesian

753 divergence-time estimation with genome-wide single-nucleotide polymorphism data of sea
754 catfishes (Ariidae) supports Miocene closure of the Panamanian Isthmus. Syst. Biol.
755 67:681–699.
756 Suyama, Y., and Y. Matsuki. 2015. MIG-seq: an effective PCR-based method for genome-wide
757 single-nucleotide polymorphism genotyping using the next-generation sequencing platform.
758 Sci. Rep. 5:16963.
759 Tajima, F. 1983. Evolutionary relationship of DNA sequences in finite populations. Genetics
760 105:437–460.
761 Taylor, B., and C. J. Sharpe. 2020. Swinhoe's Rail (*Coturnicops exquisitus*), version 1.0. *in* J. del
762 Hoyo, A. Elliott, J. Sargatal, D. A. Christie, and E. de Juana, eds. Birds of the World.
763 Cornell Lab of Ornithology, Ithaca, NY.
764 Twyford, A. D., and R. A. Ennos. 2012. Next-generation hybridization and introgression.
765 Heredity 108:179–189.
766 Vamosi, J. C., W. S. Armbruster, and S. S. Renner. 2014. Evolutionary ecology of specialization:
767 insights from phylogenetic analysis. Proc. R. Soc. B 281:20142004.
768 Vieira, F. G., F. Lassalle, T. S. Korneliussen, and M. Fumagalli. 2015. Improving the estimation
769 of genetic distances from Next-Generation Sequencing data. Biol. J. Linnean Soc. 117:139–
770 149.
771 Wang, N., B. Liang, J. Wang, C.-F. Yeh, Y. Liu, Y. Liu, W. Liang, C.-T. Yao, and S.-H. Li.
772 2016. Incipient speciation with gene flow on a continental island: Species delimitation of the
773 Hainan Hwamei (*Leucodioptron canorum owstoni*, Passeriformes, Aves). Mol. Phylogenet.
774 Evol. 102:62–73.
775 Waples R. K., A. Albrechtsen, and I. Motke. 2019. Allele frequency-free inference of close

familial relationships from genotypes or low-depth sequencing data. Mol. Ecol. 28:35–48

Warren, B. H., D. Simberloff, R. E. Ricklefs, R. Aguilée, F. L. Condamine, D. Gravel, H. Morlon, N. Mouquet, J. Rosindell, J. Casquet, E. Conti, J. Cornuault, J. M. Fernández-Palacios, T. Hengl, S. J. Norder, K. F. Rijsdijk, I. Sanmartín, D. Strasberg, K. A. Triantis, L. M. Valente, R. J. Whittaker, R. G. Gillespie, B. C. Emerson, and C. Thébaud. 2015. Islands as model systems in ecology and evolution: prospects fifty years after MacArthur-Wilson. Ecol. Lett. 18:200–217.

Watterson, G. A. 1975. On the number of segregating sites in genetical models without recombination. Theor. Popul. Biol. 7:256–276.

Weigelt P., W. Jetz, and H. Kreft. 2013. Bioclimatic and physical characterization of the world's islands. Proc. Natl. Acad. Sci. U.S.A. 110:15307–15312.

Weir J. T., and D. Schluter. 2004. Ice sheets promote speciation in boreal birds. Proc. R. Soc. Lond. B. 271:1881–1887.

Weir, J. T., and D. Schluter. 2008. Calibrating the avian molecular clock. Mol. Ecol. 17:2321–2328.

Whittaker, R. J., and J. M. Fernández-Palacios. 2007. Island Biogeography: Ecology, Evolution, and Conservation. Oxford University Press, Oxford.

Whittaker, R. J., J. M. Fernández-Palacios, T. J. Matthews, M. K. Borregaard, and K. A. Triantis. 2017. Island biogeography: Taking the long view of nature's laboratories. Science 357:eaam8326.

Yeung, C. K. L., P.-W. Tsai, R. T. Chesser, R.-C. Lin, C.-T. Yao, X.-H. Tian, and S.-H. Li. 2011. Testing founder effect speciation: divergence population genetics of the spoonbills *Platalea regia* and *Pl. minor* (Threskiornithidae, Aves). Mol. Biol. Evol. 28:473–482.

Zink, R. M., and A. S. Gardner. 2017. Glaciation as a migratory switch. Sci Adv 3:e1603133.

Data Availability Statement

Scripts and data for analysis are available on Dryad: 10.5061/dryad.37pvmcvs1 (private link: https://datadryad.org/stash/share/2PRy64eM1UIjsjCJlNLx4B9ownPLtYVob4aFp_1MXys). Raw sequence reads are deposited to Sequence Read Archives (see Table S1). Settings and procedures for these processes are fully described in the main and supplementary texts and the original scripts.

806 # Tables

807 **Table 1** Dataset and the number of sites used in each analysis

Analyses	Linkage	Site type	Outgroup	Populations	# Sites	# Samples	Notes
PCA, ADMIXTURE	unlinked	V	Non	4 breeding[1], 1 wintering[2]	1,147	47	-
NeighborNet (ngsDist)	linked	V	Laja/Atro	4 breeding	1,958	32	
RAxML, D-statistic	linked	V/I	Laja	4 breeding	108,095	31	-
TreeMix	linked	V	Laja	4 breeding	1,669 (683)	31	Population-specific private alleles were removed by glactools. The number of sites inside the bracket is for the individual-level tree.
SNAPP	unlinked	V	Laja	4 breeding	206	16	The top three/four samples with the lowest missing allele proportions in each population were selected. Excess heterozygous sites were removed by BCFtools
Diversity indices, FSC2, SWP2	linked	V/I	Non	2 genetic[3]	130,368	30	-

808 Abbreviation: PCA, principal component analysis; FSC2, fastsimcoal2; SWP2, Stairway Plot 2;

809 V, variants; V/I, both variants and invariants; Laja, *Laterallus jamaicensis*; Atro, *Atlantisia*

810 *rogersi*; [1] Tomakomai, Kushiro, Russia, and Aomori; [2] Kanto; [3] Island, and Continental.

Figures

Figure legends

Figure 1 A map of sampling localities of Swinhoe's rails (a), and their population genetic structure inferred by (b) a principal component analysis (PCA) and (c) an ADMIXTURE analysis conducted using genotype likelihoods of 1,147 single nucleotide polymorphisms (SNPs). (a, b) The colors of dots and inertia ellipses represent different regional populations. Dark-grayed regions are the expert range map of this species in BirdLife International and Handbook of the Birds of the World (2024), which also includes previously known ranges. (c) Genetic assignment of individuals to each number of genetic clusters (K = 2 to 5) is shown. Colors do not correspond to the sampling localities. Lambert Azimuthal Equal Area projection was used for the map. The inset photo by Marc Bastardot.

Figure 2 Phylogenetic trees of Swinhoe's Rail reconstructed using different methods and genetic markers. (a) A mitochondrial phylogenetic tree (mtDNA tree) using 772 bp of the *Cytb* gene. Posterior probabilities >0.9 were indicated by black dots on nodes. Numbers at the nodes indicate median divergence time estimates with 95% highest posterior density within brackets (bold) and posterior probabilities (italic). (b) A maximum likelihood (ML) tree (RAxML tree) was reconstructed using 108,095 sites including both variant and invariant sites. Only nodes with bootstrap values >70 are annotated with black circles. (c) A population-level tree reconstructed by TreeMix using the allele frequency of 1,669 polymorphic loci, excluding population-specific private alleles. Branch lengths indicate drift parameters, the strength of drift specific to each branch. The inset indicates an individual-level tree reconstructed using non-private 683 polymorphic loci. The colored arrows indicate the estimated gene flow. (d) Posterior population-

833 level trees reconstructed by SNAPP analyses using 206 single nucleotide polymorphisms (SNPs)
834 with no missing data. Colors of trees indicate percentages of each posterior tree in the credible
835 set obtained by the 'TreeSetAnalyser' program. Values at each node indicate their divergence
836 time estimates with credible intervals within brackets. Two different divergence time estimates
837 were obtained for different calibration constraints. Values beneath each node indicate posterior
838 probabilities.

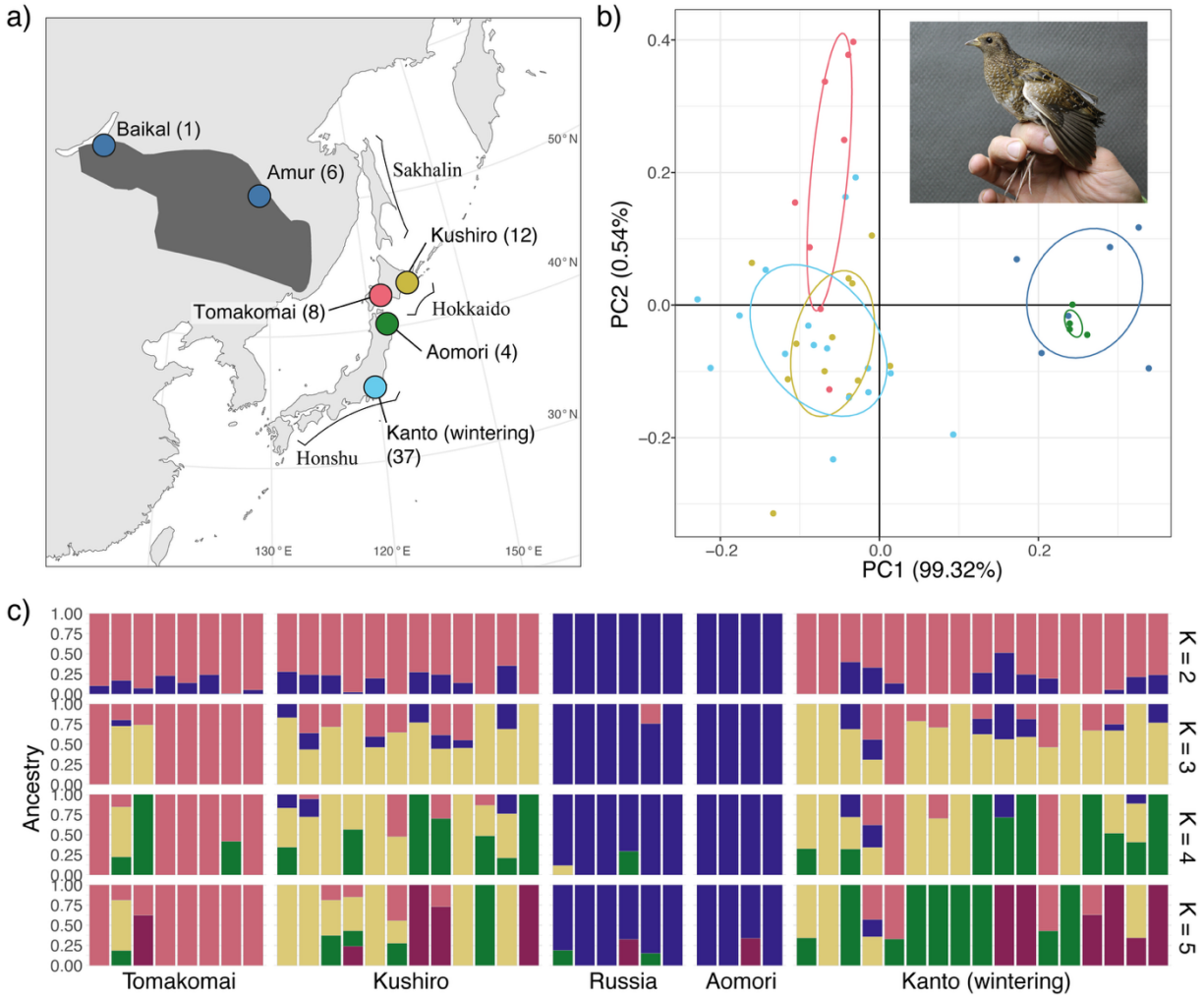
839 **Figure 3** Demographic inference of the two genetic populations of Swinhoe's Rail based on site
840 frequency spectra, using (a, b) the summary statistics, (c–e) fastsimcoal2 (FSC2), and (f)
841 Stairway Plot 2. Fixed values in each model are italicized. Effective population size and time are
842 scaled by mutation rate (1.91e-09 bp/generation) and generation time (1 year). Effective
843 population size shows the number of diploid individuals. (a, b) Comparison of scaffold-level
844 statistics between the island and continental populations for (a) Watterson's theta and (b)
845 Tajima's *D*. Statistically significant differences in the values of each index were annotated with
846 asterisks (* $p < 0.05$). (c) Generalized schematic representation of the five models tested.
847 Parameters with uppercase represent the effective population size we modeled. Hatched and
848 dotted arrows represent different migration matrices within the model. (d, e) The two best-fit
849 FSC2 models with estimated parameters. Numbers by or on branches and arrows represent
850 effective population size and the number of migrants per year, respectively. The size of each box
851 is not scaled to the corresponding population size. Divergence time is shown in million years. (f)
852 Changes in effective population size (in millions) over time (in million years) are shown for the
853 island and continental populations. Bold lines and thin ribbons around them show median and
854 95% confidence intervals, respectively. The gray range indicates the approximate period of
855 population divergence.

856 **Figure 4** A predicted present distribution of Swinhoe's Rail (a) by the species distribution model
857 using its present occurrence records (b). (a) Color gradients from indigo to yellow represent
858 environmental suitability for rails. Regions with dark gray shades are where model extrapolation
859 was restricted due to the absence of the climate in the modeling domain (grids with negative
860 MESS values). (b) The red grids indicate the used occurrence records at the resolution of the
861 environmental layers, and gray areas indicate the expert range map of this species. (c) A
862 temporal plot for a proxy for paleoclimatic temperature (δ^{18}O; lower values indicate higher
863 temperatures) (Lisiecki and Raymo 2005; Spratt and Lisiecki 2016). (d–e) Historical changes in
864 the total suitability of the habitat in the continent (blue) and the island (yellow) (d) and the
865 suitability-weighted centroid distance between the two (e). Black solid and dotted lines indicate
866 the overall median and standard deviation of the distance, and a red bar indicates the period of
867 the smallest amplitude in the temporal change. Grids covered by negative MESS values were
868 also included in the calculation. Gray bars correspond to the estimated period of the population
869 split (left) and gene flow (LGM, right). Lambert Azimuthal Equal Area projection was used for
870 maps.

871 **Figure 5** Schematic representation of our continent-island metapopulation scenario of Swinhoe's
872 Rail; geographic isolation occurred around 0.5Mya (a), and source-sink switching dynamics
873 occurred in the continent-island system between the glacial (b) and interglacial period (c). Gene
874 flow (a dotted arrow) and re-establishment or expansion of population (solid arrows) may have
875 facilitated persistence and evolution of the metapopulation. Filled circles represent stable (solid)
876 and unstable (hatched) populations.
877

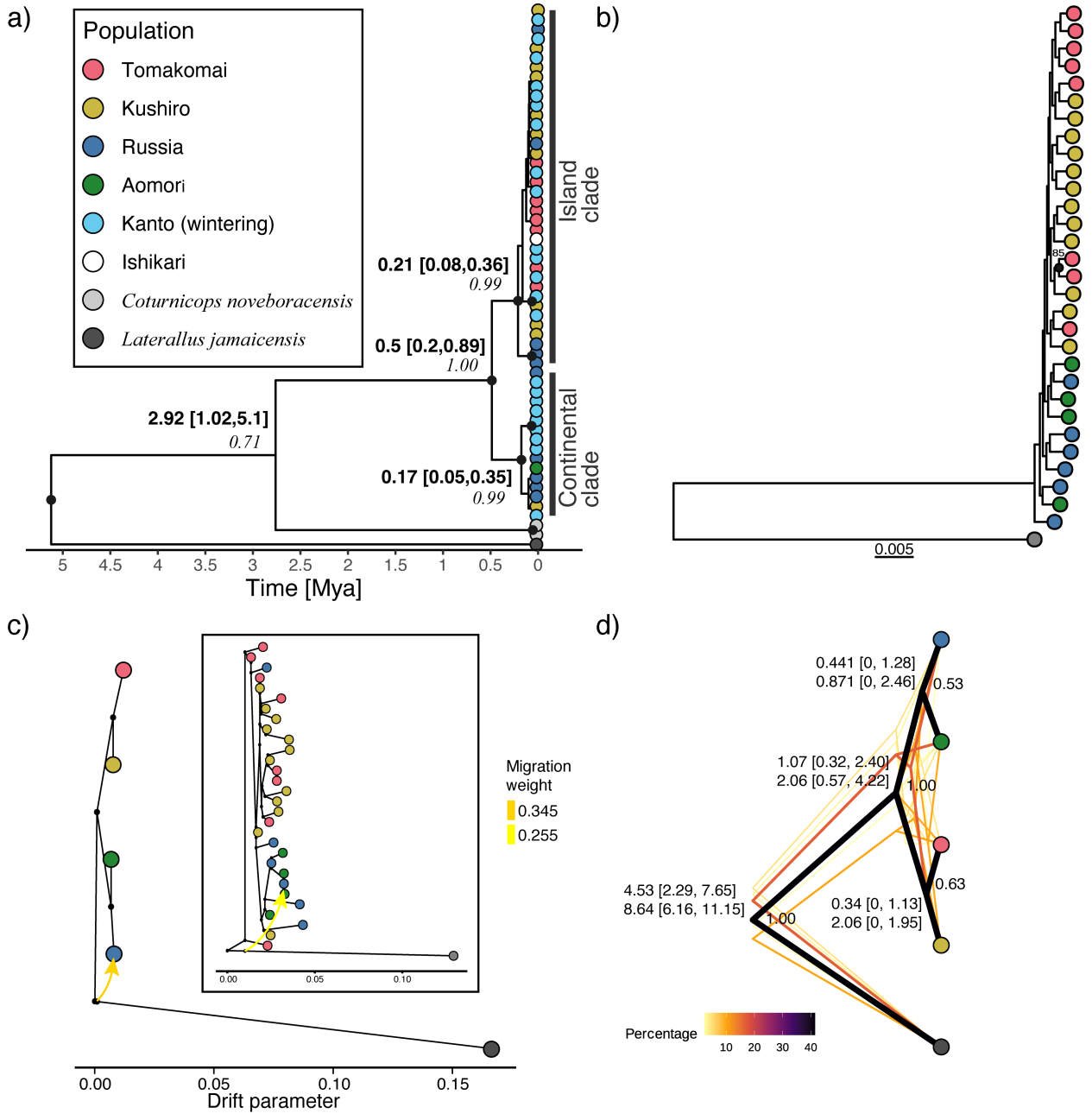
878 Figures

879 **Figure 1**



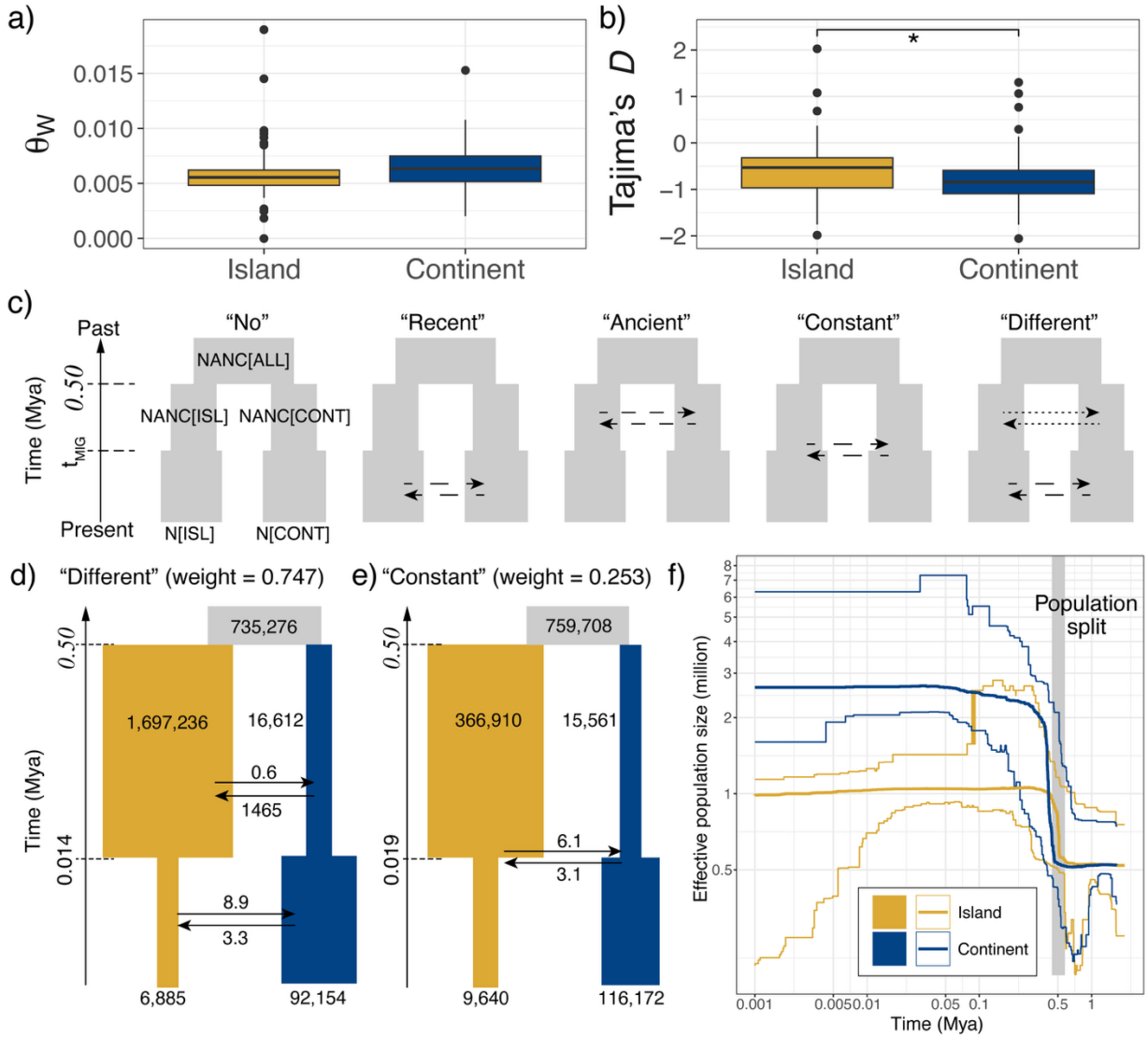
880

881 **Figure 2**



882

883 **Figure 3**



884

Figure 4

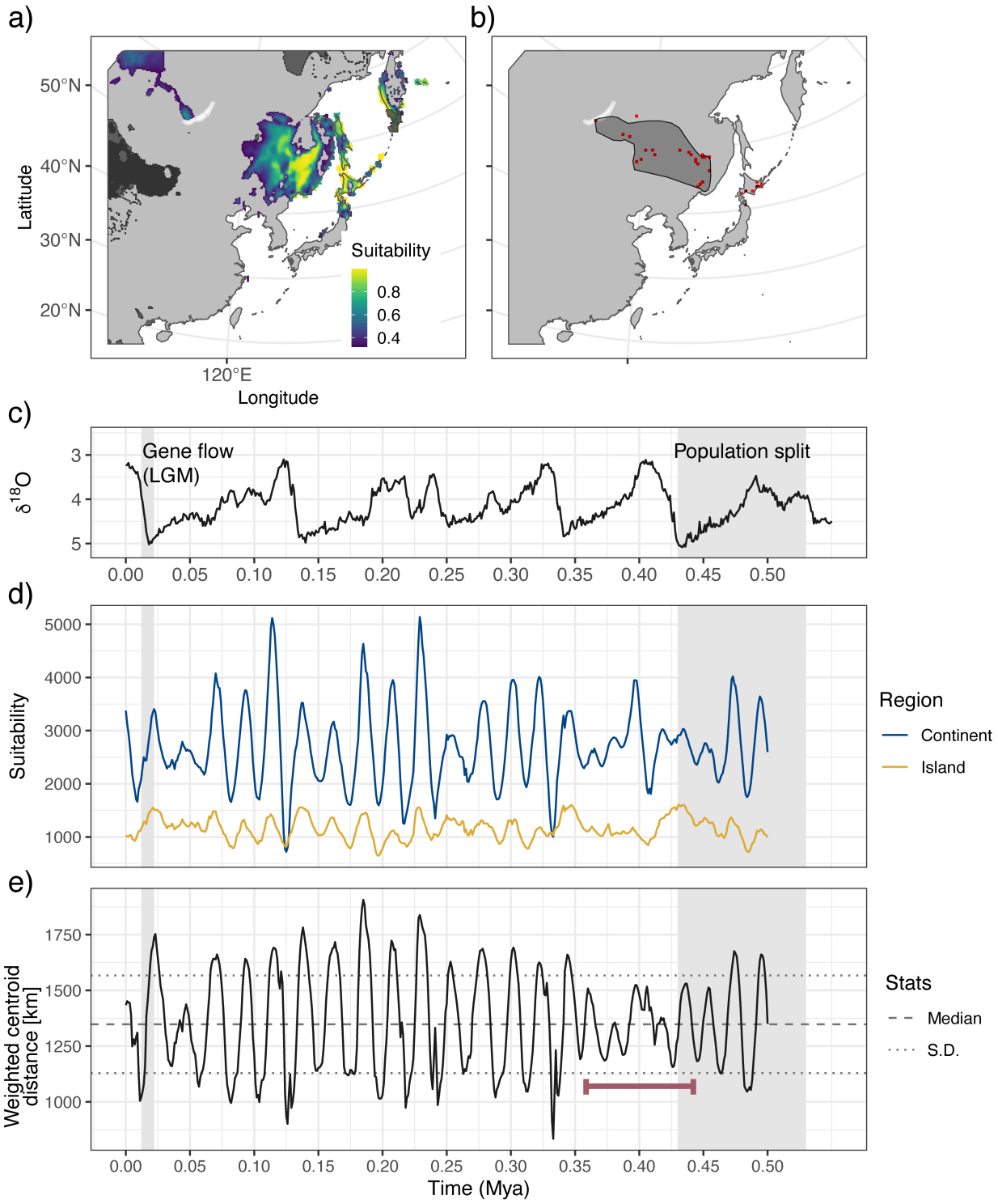
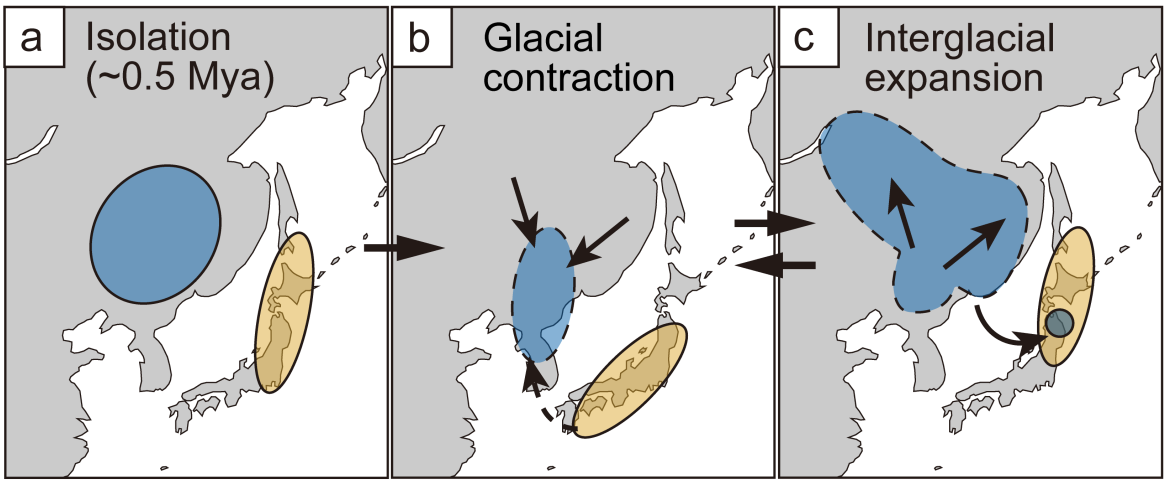

a)
b)
50°N
40°N
30°N
20°N
Latitude
120°E
Longitude
Suitability
0.8
0.6
0.4
c)
Gene flow
(LGM)
Population split
δ18O
3
4
5
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45
0.50
d)
Suitability
5000
4000
3000
2000
1000
Region
Continent
Island
e)
Weighted centroid distance [km]
1750
1500
1250
1000
Stats
Median
S.D.
Time (Mya)

Figure 5


a
Isolation
(~0.5 Mya)
b
Glacial
contraction
c
Interglacial
expansion