Title Page

Title

Phylogenomics reveals the demography of reverse colonization in a rail species

Short running title: Demography of reverse colonization

Authors

Aoki D[1,*], Senzaki M[2], Ando H[3], Odaya Y[4], Heim W[5,6], Kitazawa M[3,7], Wulf T[8], Trense D[9], Bastardot M[10], Fukuda A[11], Takahashi M[12,13], Kondo N[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

OCRID

Daisuke Aoki: 0000-0002-5780-8487

Masayuki Senzaki: 0000-0002-3843-5793

Haruko Ando: 0000-0001-7827-3883

Wieland Heim: 0000-0002-3262-2491

Munehiro Kitazawa: 0000-0003-0167-4464

Daronja Trense: 0000-0003-3598-8699

Masao Takahashi: 0000-0001-8385-6098

Natsuko Kondo: 0000-0002-0883-0506

Acknowledgements

Some tissue and blood samples were kindly provided by Yamashina Institute for Ornithology. We thank many members of our field teams; Toshio Sadakuni at Kushiro, Akio Miya and Jun-ichi Ebina at Hotokenuma, Aomori; Sergei M. Smirenski and Yury Anisimov in Russia. We also thank Nobuyoshi Nakajima for supports on sequencing runs on MiSeq. DNA samples were collected with permission from the Ministry of the Environment Government of Japan. Our field surveys were

conducted under permissions from landowners of each site, including Tomatoh. Transportation of DNA samples were 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 Nagoya Protocol. 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). Bioinformatics was mostly performed on the supercomputer of AFFRIT, MAFF, Japan.

Abstract

Aim:

Islands can become sources of continental biodiversity through reverse colonization. This is contradicted by the classic view of islands as sinks, because microevolution predicts difficulties in continental colonization by small island populations that are genetically depauperate. We fill this gap by examining historical demography that contributed to genetic variation of a continental population after reverse colonization. We exploit a rail species that likely evolved through reverse colonization, challenging the expectation that rails are the representatives of "end-of-colonization" on islands.

Location:

Continental East Asia (Russian Far East), Japanese Archipelago

Taxa:

Swinhoe's Rail (*Coturnicops exquisitus*)

Methods:

To confirm reverse colonization, therein a continental population is derived from an island lineage, we reconstruct a complex phylogenetic relationship considering gene flow and introgression. We apply multiple methods using both high-throughput sequencing and mitochondrial datasets. We reconstruct

a comprehensive demographic history of the species based on historical trajectory inference by Stairway Plot 2, fastsimcoal2 demographic modeling, and species distribution modeling. We assess genetic variation of the continental population and the key demographic process underlying them.

Results:

The continental population is inferred as a derived group within an island lineage, supporting reverse colonization. Demographic reconstruction suggests a significant population size decline through reverse colonization, with intraspecific gene flow from the island population. Introgression from an American sister species into the continental population is supported, which may be responsible for the unexpectedly high genetic variation observed.

Main conclusions:

Using phylogenomics, we reconstruct reverse colonization associated with a complex history of population size changes and gene flow. Introgression, particularly, plays a pivotal role in rapidly restoring advantageous genetic variation for persistence and expansion on the continent post-reverse colonization. In theory, introgression assists adaptive radiation and diversification. Thus, our findings facilitate an understanding of the processes that islands contribute to global biodiversity, thereby bridging a gap between micro- and macroevolution.

Key words

introgression, intraspecific gene flow, Japanese Archipelago, MIG-seq, phylogeography, population bottleneck, reverse colonization, Swinhoe's Rail, Coturnicops exquisitus

Main Text

Introduction

Island biological diversity contributes disproportionately to global biodiversity (Whittaker & Fernández-Palacios, 2007). Islands, comprising 3.5% of the terrestrial land, contribute nearly 20% of global species richness and contribute more than 60% of extinct terrestrial species (Whittaker et al., 2017). It has been assumed that the small and isolated landmass and the unique environment of the island cause small population size with low genetic variation (Cerca et al., 2023) and specialization of genomes (e.g. Burga et al., 2017), likely leading to high risks of extinction (Frankham, 2005) and the evolutionary 'dead-end' (Vamosi et al., 2014). Based on these microevolutionary assumptions, islands have been viewed as sinks where lineages end diversification (Bellemain & Ricklefs, 2008; MacArthur & Wilson, 1963).

Recently, however, "reverse colonization", through which the island becomes the source of continental colonization and radiation, has become increasingly evident (Bellemain & Ricklefs, 2008; Whittaker et al., 2017). It has been reported from a variety of taxa, including birds, reptiles, bats, insects and plants (Bellemain & Ricklefs, 2008; Patiño et al., 2015) across various kinds of islands, including the Proto-Papuan archipelago (Jønsson et al., 2011), Hawaii (O'Grady & DeSalle, 2008), and Japan (Ikeda & Motokawa, 2021). Reverse colonization can be further followed by mainland diversification. Examples include the radiation of passerine species, representing almost 60% of the modern bird species richness (Jønsson et al., 2011; Jønsson & Holt, 2015), and Anolis lizard radiation in North America (Alföldi et al., 2011). Previously, inference of reverse colonization has been done only phenomenologically by detecting a macroevolutionary pattern in a phylogenetic tree therein a continental lineage is derived from an island lineage (Bellemain & Ricklefs, 2008). However, prevalence of reverse colonization suggests that there are overlooked microevolutionary processes by which insular organisms succeed colonization and persist on the continents by overcoming aforementioned population genetics obstacles (Bellemain & Ricklefs, 2008). Therefore, reverse

colonization is needed to be assessed from microevolutionary processes to understand how islands potentially become a source of biodiversity.

Demographic history is one promising microevolutionary aspect that can help us understand how an island population overcomes its potentially low genetic variation. High genetic variation is important for persistence and adaptation to novel environments, such as more competitive environments (Hart et al., 2019; Vellend & Geber, 2005), like continents in comparison with islands. Given the several indications that insular populations may not necessarily be genetically depauperate (e.g. Patiño et al., 2015), there should be important demographic conditions under which reverse colonization is successful with retaining high genetic variation. For instance, island populations can maintain high genetic variation by avoiding severe population size decline due to unexpectedly large founder population size (Clegg et al., 2002; Sendell-Price et al., 2021) or inter-island metapopulation structure (Seki et al., 2007). Alternatively, insular populations might have gained genetic variation through gene flow from other populations or other species (Choi et al., 2021). These past demographic events linked with genetic variation are crucial to reflect its microevolutionary lesson in macroevolution because they critically determine whether speciation and diversification can initiate after population establishment (Harvey et al., 2019; Levin, 1995).

Phylogeography can cover both macro- and microevolutionary arenas (Lomolino et al., 2017), and its application with high-throughput sequencing (HTS) techniques and a species distribution model (SDM) will provide both the evidence for reverse colonization and its associated demography. HTS allows resolving phylogenetic trees while simultaneously considering gene flow and genetic drift (Beichman et al., 2018; Pickrell & Pritchard, 2012). HTS allows further reconstruction of demography beyond the phylogenetic tree, including population genetics structure (Skotte et al., 2013) and historical changes in population size (Liu & Fu, 2020). SDMs allow modeling range limits and shapes at particular periods, and therefore, its integration allows interpretation of the demographic parameters in spatially explicit ways (Harvey et al., 2019), which helps associate demographic processes with reverse colonization.

The objective of this study is therefore to provide evidence of reverse colonization at the population level and reveal its historical demography in relation to genetic variation by applying phylogeography. To achieve this goal, we selected the Swinhoe's Rail (*Coturnicops exquisitus*), the world's smallest Asian rail species, as the ideal model case, because candidate phylogeographic scenarios are easily narrowed down, given that it is only the Eurasian descendent of the New World rail species group ("*Laterallus* group") (Garcia-R & Matzke, 2021). Upon colonization from North America around the Pliocene to the Early Pleistocene (Garcia-R & Matzke, 2021), we expect that colonization by the Swinhoe's Rail occurred either 1) along the Japanese Archipelago through the Kamchatka Peninsula, followed by continental colonization (i.e., reverse colonization scenario, a dotted line in Fig. 1a), or 2) along the continental North Asia to East Asia, followed by colonization of the Japanese Archipelago (i.e., island colonization scenario, a hatched line in Fig. 1a). A reverse colonization scenario is likely in this species because a warm and wet environment that was presumably suitable for rails extended more along the islands than the continental North Asia around this period (Tindall et al., 2022). This is contrasted by many rail species being the representatives of "end-of-colonization" on islands for extensive specialization and small population size (Garcia-R et al., 2017; Garcia-R & Matzke, 2021; Hume & Martill, 2019), making this study further challenging to reinterpret islands as sources of biodiversity.

In this study, we first ask whether a Japanese population is basal of and a continental population nests within Japanese populations in a phylogenetic tree, supporting the reverse colonization scenario, by reconstructing phylogenetic relationships among populations through application of an HTS technique as well as standard mitochondrial analyses. Then, by integrating HTS and SDM, we compare genetic variations of each population and examine underlying historical demography, such as the presence, levels and timings of intraspecific gene flow, introgression and changes in population size. Finally, we evaluate a key demographic process associated with genetic variation of the continental population. We predict that the population under reverse colonization has retained genetic variation through maintaining unexpectedly large population size or through gene flow from other populations or species.

Materials & Methods

Details about the methods and datasets are described in the Supplementary Methods and scripts in data depository (see Data Availability Statement).

Field sampling

The poorly studied Swinhoe's Rail is listed as Vulnerable on the IUCN Red List (B. Taylor & Sharpe, 2020). Almost all of the presently recognized breeding areas were re-discovered recently following the first description of its vocalizations (Wulf et al., 2017), including the Amur region and Lake Baikal in Russia (Anisimova et al., 2019; Heim et al., 2019) and some northern regions of Japan, namely Aomori (Miya et al., 2005) and Hokkaido (Senzaki et al., 2020) (Fig. 1a). Despite its highly cryptic lifestyle inhabiting inaccessible wet meadows (B. Taylor & Sharpe, 2020), we developed our original sampling strategies (Heim et al., 2019; Senzaki et al., 2020) and collected feather, buccal swab or blood samples successfully covering the entire extent of the known breeding area, Amur (n = 6) and Baikal (n = 1) for continental Russia, Kushiro (n = 12) and Tomakomai (n = 8) for Hokkaido Island, and Aomori (n = 4) for Honshu Island (Fig. 1a, Table S1). We also collected samples from wintering rails in the Kanto region of Honshu Island (n = 37), to infer the migratory connectivity of this species and to speculate any historical range shifts. We additionally retrieved mitochondrial data from the NCBI nucleotide database for phylogenetic analyses (Table S1).

Laboratory procedures & Bioinformatics

The whole genomic DNA extracted from each tissue was used in the two experiments, sanger sequencing of the partial mitochondrial Cytochrome *b* (*Cytb*) region for 49 samples and one of the HTS techniques, MIG-seq (Suyama & Matsuki, 2015), for 52 samples (Table S1). The latter is a type of genotyping by sequencing with conducting polymerase chain reactions at inter-simple sequence repeats. In exchange for smaller genome coverage, it is effective for low quality DNA samples, best adapted to some of our dried feather or swab samples. For sanger sequencing, we used the standard PCR and BigDye terminator protocols to sequence *Cytb* genes. For MIG-seq, we used eight forward

and eight reverse primers originally described by Suyama & Matsuki (2015) to prepare a library mixture containing all the samples. The final mixture was applied to MiSeq Reagent Kit v3 150 cycles (Illumina) to obtain 150bp paired-end reads.

The raw FASTQ reads for each sample were trimmed and filtered using Trimmomatic v. 0.39 (Bolger et al., 2014). Both paired and unpaired trimmed reads were aligned against a reference genome of a closely related species in the '*Laterallus* group' (Garcia-R & Matzke, 2021), the Black Rail *Laterallus jamaicensis coturniculus* (Assembly number GCA_022695575, Table S1), using the BWA-MEM algorithm of Burrows-Wheeler Aligner (Li, 2013). After removal of individuals based on the levels of missingness, 49 samples were retained.

We mainly adopted probabilistic analysis frameworks to account for uncertainties about the true genotype, by calculating genotype likelihoods using a software ANGSD (Korneliussen et al., 2014). Genotypes were hard-called using the '-doGeno' function of ANGSD for softwares not accepting genotype likelihoods as inputs (RAxML-ng and TreeMix). Genotype likelihoods were repeatedly calculated using the 'SAMtools' model for different subsets of the dataset for different purposes. They included the entire dataset without population categorization (DATASET 1), the entire dataset with five populations (Russia, Kushiro, Tomakomai, Aomori and Kanto) (DATASET 2), and the dataset with three genetic breeding populations, namely Kushiro, Tomakomai, and continental (Russia + Aomori), which was delineated based on the results of ADMIXTURE at K = 3 showing clear genetic distinctiveness (DATASET 3). Linkage disequilibrium among variant sites were calculated and pruned using ngsLD (Fox et al., 2019) for analyses that do not recommend linked sites. Genotype likelihoods were similarly calculated for outgroup sequences downloaded through the Sequence Read Archive, which included *Laterallus jamaicensis* (reference sample), *Atlantisia rogersi*, and *Zapornia atra* (Table S1). For demographic analyses, we used unfolded site frequency spectrums (SFS) or two dimensional SFSs (2d-SFS) using 167,120 sites (including both polymorphic and monomorphic sites) for each population or population pair defined by the DATASET 3, calculated by the '-doSAF' option of ANGSD and the 'realSFS' function. For Fst computation and fastsimcoal2 analyses, the 2d-SFSs of unlinked sites were used.

Genetic structure and genetic diversity indices

We examined global genetic structure using the DATASET 1 by principal component analysis (PCA) using PCAngsd (Meisner & Albrechtsen, 2018) and ADMIXTURE analysis using NGSadmix (Skotte et al., 2013) with K ranging from 1 to 6. Using the DATASET 3, we calculated two genetic diversity indices, Watterson's theta θ_W (Watterson, 1975) and nucleotide diversity θ_π (Tajima, 1983), and two neutrality test statistics, Tajima's *D* (Tajima, 1989) and Fu's *Fs* (Fu, 1997), per reference scaffold and population by using the 'do_stat' option of the 'thetaStat' function with the SFSs as inputs. We also calculated pairwise weighted and unweighted Fst on unlinked SFSs for each population pair using the 'fst' function implemented in 'realSFS'. Values below zero for the neutrality test statistics indicate population bottlenecks followed by demographic expansions. Diversity indices and neutrality statistics were compared among populations, using linear mixed models. To set populations as covariates, we constructed two independent models with either Tomakomai or Kushiro set as the base population to be compared with as a dummy variable.

Inference of phylogenetic relationship

Phylogenetic relationships reconstructed using different datasets and methods differ to each other because they may reflect different gene genealogy of a complex population history, such as intraspecific gene flow and introgression (Kapli et al., 2020; Twyford & Ennos, 2012). Therefore, our phylogenetic reconstructions consist of four analyses, a mitochondrial phylogenetic tree, a maximum likelihood (ML) tree using SNP data (ML SNP tree), a population-level tree, and demographic modeling. Then, a population history that can reasonably explain all the results was comprehensively interpreted.

Mitochondrial phylogenetic tree: The time-calibrated Bayesian inference method was applied to an alignment of 772 bps of partial *Cytb* by using BEAST v.2.7.3 (Bouckaert et al., 2019). Samples included 12 haplotypes from 54 Swinhoe's rails and three outgroups (two yellow rails and one black rail). The standard molecular clock rate of 1.05%/lineage/Myr (Weir & Schluter, 2008) was applied to estimate divergence time. A Markov chain Monte Carlo analysis was run for 2e+07 generations with

trees sampled every 1,000, from which a maximum clade credibility tree was drawn with mean heights after 10% of the sampled trees was discarded as burn-in.

ML SNP tree: RaxML-ng (Kozlov et al., 2019) with 'GTGTR4+G+ASC_LEWIS' model, which corrects acquisition biases when using SNP datasets (Leaché et al., 2015), was applied to the unlinked SNPs acquired using the DATASET 2. A black rail was set as an outgroup and standard non-parametric bootstrapping with 100 replicates was simultaneously conducted.

Population-level tree: TreeMix (Pickrell & Pritchard, 2012) was used to infer a population-level tree that simultaneously considered gene flow and genetic drift. A model with a migration edge = 0 was initially estimated, which were used to compare model performance of models with migration edges ranging from one to five. Ten replicates of runs for each number of migration edges were conducted. Changes in log likelihood, total variance in population relatedness, and residual fits were evaluated to determine the best number of migration edges (Fitak, 2021; Pickrell & Pritchard, 2012). Consistency was also checked among different trials at the best number of migration edges. Among the replicates with the best number of edges, a run with the largest model likelihood was selected as the best-fit model.

Demographic modeling: A model-based approach using fastsimcoal2 v. 2.7.0.9 (Excoffier et al., 2021) was applied to reconstruct a population history, which can estimate gene flow and population size changes while simultaneously inferring the population branching history although outgroup sequences from the database cannot be included in the models. Using the 2d-SFSs, six independent models were tested. In each model, the three populations (Tomakomai, Kushiro and continental) were assigned to POP1 to POP3 differently giving three different branching histories, with and without recent population size changes at t_{EXP} (Fig. 3a). In each model, we considered only bidirectional gene flow between adjacent populations (between Kushiro and Tomakomai, and Tomakomai and the continental) at t_0. We assumed a mutation rate of 1.91e-09 (Nam et al., 2010), which agrees well with other rates estimated or used elsewhere (Nadachowska-Brzyska et al., 2015; Smeds et al., 2016). We applied 'bounded reference' to constrain the first split time (t_{DIV2}) between 100,000 and 1,000,000,

whose upper bound was set according to the divergence time estimated for the two statistically supported mitochondrial clades (see results). Model selection was conducted among the six models tested by calculating Akaike Information Criterion (AIC) for each model using the model likelihood and the number of parameters used.

Inference of demographic trajectories

Stairway Plot 2 (SWP2) can estimate demographic trajectories through time, independently of user-defined models (Liu & Fu, 2020). Using the one-dimensional unfolded SFSs, we separately estimated demographic changes for each of the three populations. The same mutation rate was used as the fastsimcoal2 analyses and analyses were run with default settings. In the SWP2 analysis, the timing that a population started to show different demography could be interpreted as the timing of divergence, which can further support population history.

We compared the effective population size (N_e) estimates with the contemporary population size to understand the spatial extent of the populations by converting census population size (N_c) in Kitazawa et al. (2022) to effective population size of a Hokkaido region including Tomakomai by using a mean ratio of N_c:N_e = 1:0.44 for birds (Frankham, 1995).

SDMs

Because the known breeding range of the Swinhoe's Rail is restricted to only a handful localities, we modeled that of its sister species, the Yellow Rail distributed in North America, and we projected the model to the present and past climatic environment conditions of North America and East Asia. The model projection was intended to estimate the breeding range of both the Swinhoe's Rail, the Yellow Rail, and their common ancestor. This benefits from utilizing a large dataset from eBird (https://ebird.org/) from North American countries. Meanwhile, this is strongly based on an assumption of niche conservatism; we assume that these two sister species retain the environment niche of their common ancestor.

The SDM was applied to the eBird Dataset of the Yellow Rail using MaxEnt (Phillips et al., 2006). We generated a 'zero-filled' dataset following the instruction of an R package 'auk' (Strimas-Mackey et al., 2018), and we used presence and 'pseudo-absence' points of this species per observation event and site as occurrence and background inputs, respectively. We used an R package 'pastclim' v. 1.2 (Leonardi et al., 2023) to extract time slices of 22 variables (Krapp et al., 2021). We selected one with the best explanatory power to the response variable per pair of variables correlated >0.7, by using the 'corSelect' function in an R package fuzzySim v. 4.9.2 (Barbosa, 2015), resulting in seven covariates for modeling. We tested 12 models with different parameter settings by using the function 'ENMevaluate' implemented in an R package ENMeval v. 2.0.4 (Kass et al., 2021), and the best model was selected by the lowest delta AIC for small sample size (AICc), which was then projected to seven different time slices of environmental variables, including three interglacial (0 (present), 0.424, and 0.790 Mya) and three glacial (0.029 (LGM), 0.478, and 0.761 Mya) periods. The two oldest periods represented the initial population split inferred by the time-calibrated mitochondrial tree and the SWP2 analyses (split1 in Fig. 3), and the two intermediate periods roughly corresponded to the oldest estimate for the population expansion of the continental population inferred by the SWP2 analyses (split2 in Fig. 3).

Results

Sequencing and reference mapping

MIG-seq resulted in paired-end reads of 4,834,210 overall and 92,966 per-sample average. Reads after quality-filtering were mapped on the reference genome, giving 3,724,169 paired-end reads successfully mapped.

Phylogenetic relationship and gene flow

The multiple phylogenetic trees based on different methods/datasets exhibited different populations as the basal population of the Swinhoe's Rail, suggesting complex histories of intra- and interspecific gene flow. The Swinhoe's Rail diverged from the Yellow Rail around the Late Pliocene to the Early

Pleistocene (median = 2.92 Mya), and two clades diverged within species around the Middle Pleistocene (the Chibanian epoch) (median = 0.5 Mya) but showing no apparent genetic structure between them, inferred by mitochondrial phylogenetic analyses (Fig. 2a, Fig. S1). Meanwhile, the Tomakomai population was inferred as basal in both the population-level tree by TreeMix (Fig. 2c) and demographic modeling by fastsimcoal2 (Fig. 3a, b), and its divergence ('split1') was estimated around the Middle Pleistocene (0.981 Mya, Fig. 4b), suggesting that the two mitochondrial clades possibly correspond to this split. Results from different dataset suggested multiple gene flow events. Of the two mitochondrial clades, the Tomakomai samples were contained exclusively in one clade while the others were included in both clades (Fig. 2a, Fig. S1), suggesting post-divergence intraspecific gene flow from the common ancestor of the Tomakomai population to the other populations. This is contradicted by the best fastsimcoal2 model reconstructed using MIG-seq data (Fig. 3a, b; AIC = 14019.1, Akaike weight = 0.515, Table S2) and ADMIXTURE analysis (red clusters in Tomakomai samples at K = 3, Fig. 1c) that suggested that the direction of contemporary migrants are biased from both the Kushiro and continental populations to the Tomakomai population. Together, historical but post-divergence gene flow from Tomakomai to the other populations (or their common ancestors) and contemporary gene flow in the other direction were inferred, given that mitochondrial DNA reflects older events than autosomal SNPs (de Jong et al., 2023). Complex intraspecific gene flow was also supported by Fst which showed considerable difference among pairs (Table S3).

Interspecific gene flow (i.e., introgression) from the *Laterallus* lineage to the continental population was influential, supported by the best TreeMix model that well explained the intraspecific genetic covariance when considering one migration edge from an outgroup (migration weight = 0.080, Fig. 2c, S2-S4, Supplementary Discussion). Introgression was most likely from the Yellow Rail, given that the analysis inferred introgression from an unsampled *Laterallus* group species that split from the ancestor of *Laterallus* group simultaneously with the Swinhoe's Rail (see Supplementary Discussion), and this reasonably explained that introgression caused the continental population to be basal in the ML SNP tree without considering gene flow (Fig. 2b, Fig. S5). Introgression occurred exclusively to

the entire continental population, supported by genetic homogeneity in the PCA (PC1, 99% variance explained), ADMIXTURE analysis (K = 2–5), and the ML SNP tree (Fig. 2b).

Overall, the Tomakomai population would most likely be the basal population of the Swinhoe's Rail, and the common ancestor of the Kushiro and continental population diverged from it around the Middle Pleistocene (Fig. 6a-b). The continental population was its most derived population, which received both gene flow from the insular population and introgression from the *Laterallus* lineage on the continent (Fig. 6c-d), supporting the reverse colonization scenario.

Historical demography of reverse colonization and contemporary genetic variation

The reverse colonization scenario inferred by the tree topology was well corroborated by the demographic history. The common ancestor of the Kushiro and continental populations (ANC1) occurred through a founder event, supported by the split1 associated with a drastic population decline around the early Middle Pleistocene (0.9–0.8 Mya, SWP2), while leaving the Tomakomai population demographically stable (Fig. 3b, c). This caused a one- (SWP2) to two-fold (fastsimcoal2) decline in effective population size of the ANC1 in comparison with the Tomakomai or the common ancestor of the entire populations (ANC2), resulting in N_e = 2,400 (fastsimcoal2) to 151,000–159,000 (SWP2). Historical stability of the Tomakomai population was also supported by the highest neutrality test values around zero (Fig. 4c, d), suggesting that the Tomakomai population was likely the source population from which a founder event occurred. Because the large and stable effective population size of the Tomakomai population was contradicted by its small contemporary effective population size (roughly two orders of magnitude smaller than estimates by SWP2 or fastsimcoal2, Table 1), it indicated that the Tomakomai population could have experienced drastic declines recently ~20,000 years ago which cannot be detected by our demographic reconstruction (Nadachowska-Brzyska et al., 2022).

Then, another founder event from the ANC1 and its subsequent population expansion to the continent gave a rise to the continental population, supported by the 'split2' associated with an increased population size of the continental population towards the late Middle Pleistocene (0.3–0.6

Mya, SWP2) to the Late Pleistocene (0.0538 Mya, fastsimcoal2), while leaving the Kushiro population demographically stable (Fig. 3b, c). Again, demographic stability of the Kushiro population at the split2 suggests that the continental population was derived by a founder event from it. A genetic bottleneck specific to the continental population was supported by the highest level of drift parameter in the TreeMix analysis (drift parameter = 0.107, Fig. 2c) in a model considering introgression, while it was lowest among populations without considering introgression (drift parameter = 0.100 , Fig. S5). This indicated that the additional founder event causing genetic bottleneck was possibly associated specifically with continental colonization and that introgression was responsible for counteracting genetic drift.

The overall high genetic variation of the continental population in comparison with the two Japanese populations (Fig 4a, b, Table S4) contradicted the repeated population bottlenecks inferred for the continental population. Finally, the population expansion on the continent caused an increase in effective population size by 1.5× (SWP2) to 20× (fastsimcoal2), resulting in their significantly negative neutrality test values (Fig. 4c-d).

SDMs

The best-fit model (response curve 'LQ' + regularization multiplier 1.0; AICc = 2446.725, model weight = 0.54, AUC = 92.5%) based on the present Yellow Rail occurrence data was projected to the environment layers of both North America and East Asia during different glacial and interglacial periods. We evaluated our best-fit model to be well reflecting the present distribution of both the Yellow Rail and the Swinhoe's Rail, and our assumptions of niche conservatism are thought to be met (see Discussion and Supplementary Results, Fig. S6).

During both glacial and interglacial periods, Alaska was suitable for this species, which was possibly the potential source of Asian colonization. The Japanese Archipelago was more likely to have been colonized by rails in Alaska than continental East Asia around the onset of the Middle Pleistocene (~0.8 Mya). This was supported by the abundant and stable species' suitable habitat of the Kuril and the Japanese Archipelagoes being more exposed to Alaska (Fig. 5e, f) than the continental

East Asian (e.g., current breeding ranges, such as Amur) that was rather isolated by potentially unsuitable habitats in North Asia (colored in black in Fig. 5) during both the interglacial and glacial periods. Meanwhile, secondary contacts and introgression between North American and Asian continental populations could have been facilitated towards the Late Pleistocene, supported by suitable regions gradually expanded in North Asia along the coastline of the Sea of Okhotsk (arrows in Fig. 5), towards Amur. At the same time, suitable habitats became increasingly less available along the Japanese Archipelago except for the LGM. Therefore, SDMs supported our phylogenomic scenarios that the Japanese Archipelago was more likely to be the source area for the Swinhoe's Rail before the Middle Pleistocene, and introgression was likely to have occurred on the continent towards the Late Pleistocene.

Discussion

Despite the importance of reverse colonization for tackling the traditional assumption of islands as sinks of biodiversity (Whittaker et al., 2017), how potentially genetically depauperate island populations (Frankham, 1997) can successfully colonize the continent remained unexamined (Bellemain & Ricklefs, 2008). To answer this question, using an integrative phylogenomics approach, we examined reverse colonization not only by means of phylogenetic patterns but also from demographic perspectives, including temporal population size changes, intraspecific gene flow, and introgression. We disentangled a complex population history of the Swinhoe's Rail; an American ancestry colonized the Japanese Archipelago first, and its island descendent back-colonized the Eurasian continent, representing reverse colonization. We found that the back-colonized continental population represented high contemporary genetic variation even after severe population bottlenecks. Demographic analysis suggested that introgression from an American sister lineage and intraspecific gene flow might explain its unexpectedly high genetic variation.

Demography affecting genetic variation under reverse colonization

It has long been assumed that small effective population size of island populations (Frankham, 1997, 2005) lead to low genetic variation and hence a low chance of reverse colonization (Patiño et al., 2015). Our results showed that the ancestral insular population retained a large effective population size, corroborating the recent arguments that the insular population is not necessarily genetically depauperate (García-Verdugo et al., 2015; Patiño et al., 2015). This is inferred by a comparison with previous pairwise sequentially Markovian coalescent (PSMC) analyses in birds that could be roughly comparable to our SWP2 analyses; our mean N_e (Table 1) was similar to that for large island populations, New Guinea (mean N_e = 325,000) (Pujolar et al., 2022) or continental populations (mean N_e = 362,456) (Leroy et al., 2021). Rather, an important decline in population size was associated with colonization along the Japanese Archipelago (Fig. 2c, Fig. 3b-c, Fig. 6b), which resulted in as small population size as oceanic island populations (mean N_e = 94,944) (Leroy et al., 2021). This may have strongly lowered genetic variation, because colonization along island chains usually causes severe genetic bottlenecks and facilitates non-adaptive genomic evolution (Clegg et al., 2002; Sendell-Price et al., 2021). Given that reverse colonization has been frequently reported along archipelagic chains (e.g. Jønsson et al., 2011; Patiño et al., 2015), reverse colonization processes themselves may be responsible for reducing genetic variability.

Meanwhile, unexpectedly as high contemporary genetic variation (especially pairwise nucleotide diversity) of the continental population as the stable Tomakomai population contradicted many other cases experiencing population bottlenecks towards the late Middle to the Late Pleistocene (Excoffier et al., 2009; e.g. Boria et al., 2021; Jezkova et al., 2015; Saitoh et al., 2010). Introgression might have introduced a profound effect on restoring genetic variation of the continental population as has been supported by our TreeMix analysis (Fig. 2c, Fig. S5) to a larger extent than intra-specific gene flow and the other candidate process, mutation. Our results showed that genetic variation was not heightened in the Kushiro population despite a signature of historical intraspecific gene flow similar to the continental population in our mitochondrial analyses. This suggested that intraspecific gene flow had a minor effect for heightening genetic variation. The latter was supported by genetic

homogeneity of the continental population despite its widespread distribution (Fig. 1a-c), indicating that genetic variation accumulated more rapidly through introgression than through genetic mutation that should have been led by spatial and population expansion (Excoffier et al., 2009).

Introgression can potentially result not only in a microevolutionary consequence but also in a macroevolutionary outcome, diversification (Seehausen, 2013). Like our results showed, introgression can introduce genetic variation without disrupting within-species (population) divergence more likely than intraspecific gene flow (Seehausen, 2004, 2013). Such a population has a potential to undergo adaptive evolution in a novel environment (Seehausen, 2013; S. A. Taylor & Larson, 2019) and repeated niche expansion, potentially resulting in diversification (e.g. Cerca et al., 2023; Choi et al., 2021). Our case was unique in that both the back-colonized population and the source of introgression were produced through transcontinental colonization (Fig. 6a, c). A chance of introgression through reverse colonization is probably much higher in a conceivable situation where a source continental species persists on the continent where its insular descendent back-colonizes after completion of speciation. Therefore, introgression is one hopeful candidate to answer why diversification can be often led by reverse colonization. In this study, we could not quantify the degree of a gain in genetic variations through introgression by specifying introgressed genomic regions, because of a low genomic resolution of our dataset and not including genomic data of its possible source, the Yellow Rail. Nevertheless, it may underestimate but not overestimate the level of introgression, supported by a remnant of a positive residual between the continental population and the *Laterallus* group even after adding a migration edge between them (Fig. S3; Pickrell & Pritchard, 2012). Future investigations need to include these samples to evaluate the relative contribution of introgression to successful reverse colonization.

Reverse colonization in phylogeographic contexts

The island ancestry of the Swinhoe's Rail was possibly linked with higher availability of their suitable habitats along the Japanese Archipelago than the continent. This scenario may also be adequate to the period of species divergence (the Late Pliocene to the Early Pleistocene; Fig 2a) because the climate

of the Early Pleistocene resembles the early Middle Pleistocene with low amplitudes of glacial cycles more than later Pleistocene periods with severer cycles (Head, 2021) (see Supplementary Discussion). The island ancestry might have occupied not only a part of Hokkaido but also Honshu Island (Fig. 6a). This is supported by 1) the two orders of magnitude larger population size for the ancestral and Tomakomai populations than its contemporary size, 2) suitable areas in Honshu during glacial and older interglacial periods but not in the present in our SDMs, and 3) our Kanto wintering samples showing an affinity towards Japanese breeding birds (Fig. 1b-c, 2b), since wintering regions of short-distance migratory birds may reflect the past breeding locality (Zink & Gardner, 2017). The Japanese Archipelago served as important Quaternary refugia from which many organisms underwent subsequent population expansions along the archipelago (e.g. Aoki et al., 2018; Saitoh et al., 2010). Therefore, our rail system reflects a special case of this refugial scenario, therein birds persisting in the archipelago colonized the continent in the wave of northward expansion. Reverse colonization in the wave of northward expansion was also inferred in different taxonomic groups (bats, Ikeda & Motokawa, 2021; bryophytes, Patiño et al., 2015), suggesting a potentially higher prevalence of reverse colonization from islands that served as refugia.

Limitations in demographic estimates

Despite several limitations in our demographic inference, these effects should be minor in our conclusions. First, effective population size estimates by different approaches reflect different historical events, making important differences in the level of population size changes. For example, the population size change for the split1 differed between the two analyses by more than one order of magnitude (Fig. 3, Table 1), which blurs the strength of the population bottleneck and hence possibly compromises the importance of introgression. Simplified demographic models of the fastsimcoal2 analysis may have concealed important demographic parameters, or the effect of gene flow cannot be taken into account by SWP2 resulting in overestimation of the population size (Nadachowska-Brzyska et al., 2022). However, the largest estimate of the population size after bottleneck ranged N_e = 151,000–159,000 (SWP2) which is still small in comparison with small oceanic island populations

(mean N_e = 94,944) (Leroy et al., 2021), not precluding the importance of introgression in reverse colonization.

Second, an assumption of niche conservatism for the SDM potentially causes overestimation of the breeding range of the Swinhoe's Rail, because island colonization may lead to specialization, narrowing its niche width (Wilson, 1961). However, this only confines us to draw the breeding limit of the Swinhoe's Rail itself, but the total range occupied by the two sister species may resemble our estimation. Hence, this does not affect our inference of their secondary contact. Rather, we suggest a possibility that introgression from the Yellow Rail to the continental Swinhoe's Rail may have introduced genetic variants preferred in the continent, resulting in the continental range expansion.

Conclusion

Demographic processes after population establishment are fundamental to bridge a gap between micro- and macroevolution since they determine population genetics structure that is significantly linked with subsequent speciation and diversification (Harvey et al., 2019). This study reconstructed a dynamic demographic history of reverse colonization by a species belonging to a group that repeatedly underwent island colonization; the Japanese ancestral population that was established from Alaska colonized the Eurasian continent with severe population bottlenecks, and introgression from its American sister species has possibly assisted colonization by providing genetic variability. Islands have unique adaptive landscapes different from the continent, providing ecological opportunities of island diversification, but its consequence on the continental biodiversity has not been discussed. We argue that reverse colonization creates an important opportunity for an insular species to hybridize with its continental relatives, and it may generate high genetic variation and novel allelic combinations that detect unoccupied adaptive peaks on the continent (Stroud & Losos, 2016), leading to another ecological opportunity for diversification on the continent. Interpreting islands not as an "evolutionary dead-end" but as an "evolutionary turning-point" will facilitate our understanding of evolutionary mechanisms underlying global biodiversity.

Tables

Table 1 Comparisons of effective population size estimates for the common ancestor and present Tomakomai populations derived by different methods

	Stairway Plot 2	Fastsimcoal2	Nc-based [referred to Kitazawa et al. (2022)]
N_e (common ancestor)	607,000 - 818,000	56,200	-
N_e (Tomakomai)	867,000	1,010,000	4,840

Figures

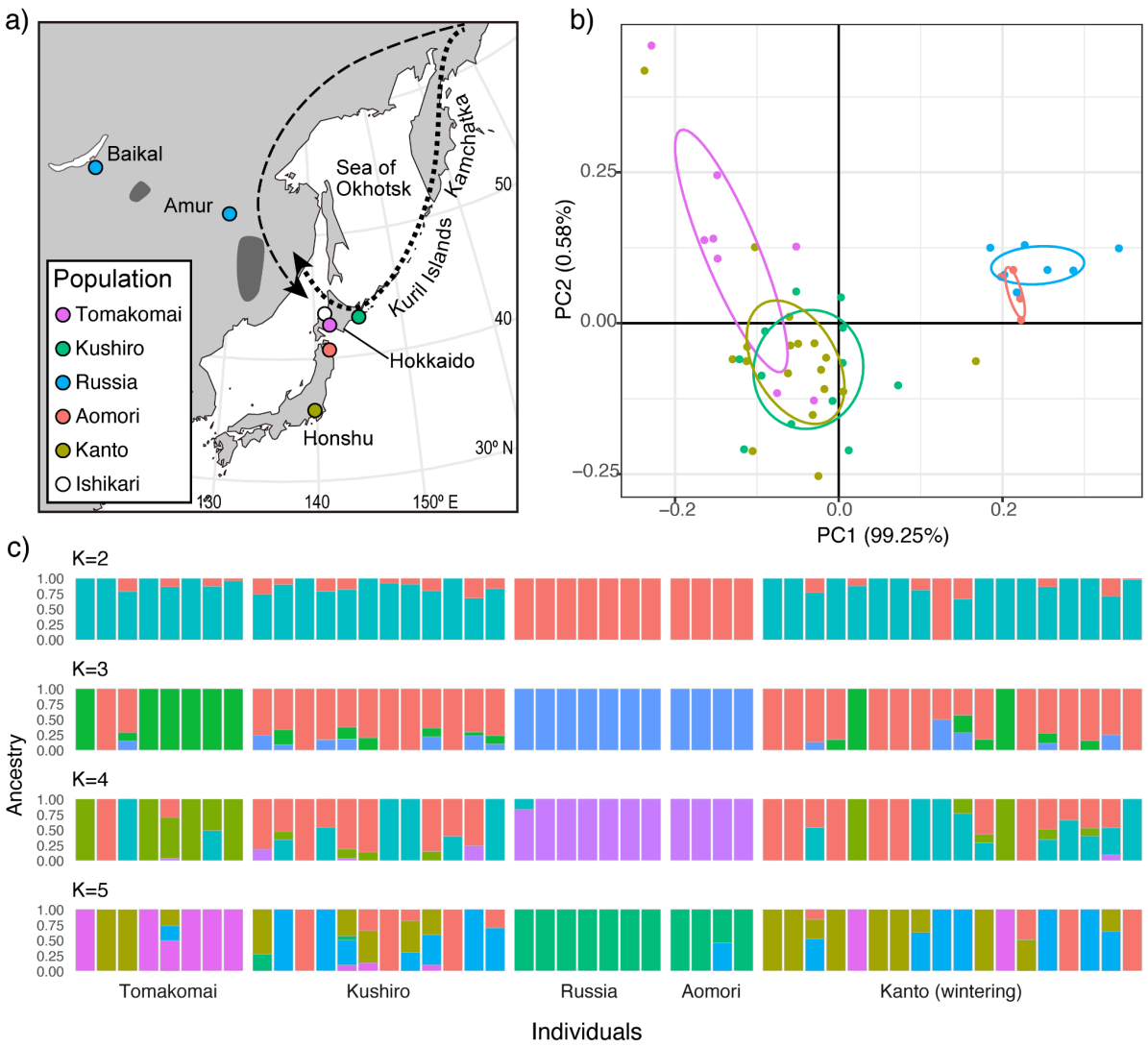

Figure 1 A map of sampling localities of Swinohoe's rails used in our genetic analyses (a), and their population genetic structure inferred by (b) a principal component analysis (PCA) and (c) an ADMIXTURE analysis conducted using genotype likelihoods of 3,016 SNPs. (a, b) Dotted and hatched arrows indicate the reverse colonization and island colonization scenarios, respectively. Colors of dots and inertia ellipses represent different regional populations. Dark grayed regions are breeding areas following BirdLife International (2023) but where present breeding evidence is scarce or lacking (BirdLife International, 2023). (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.

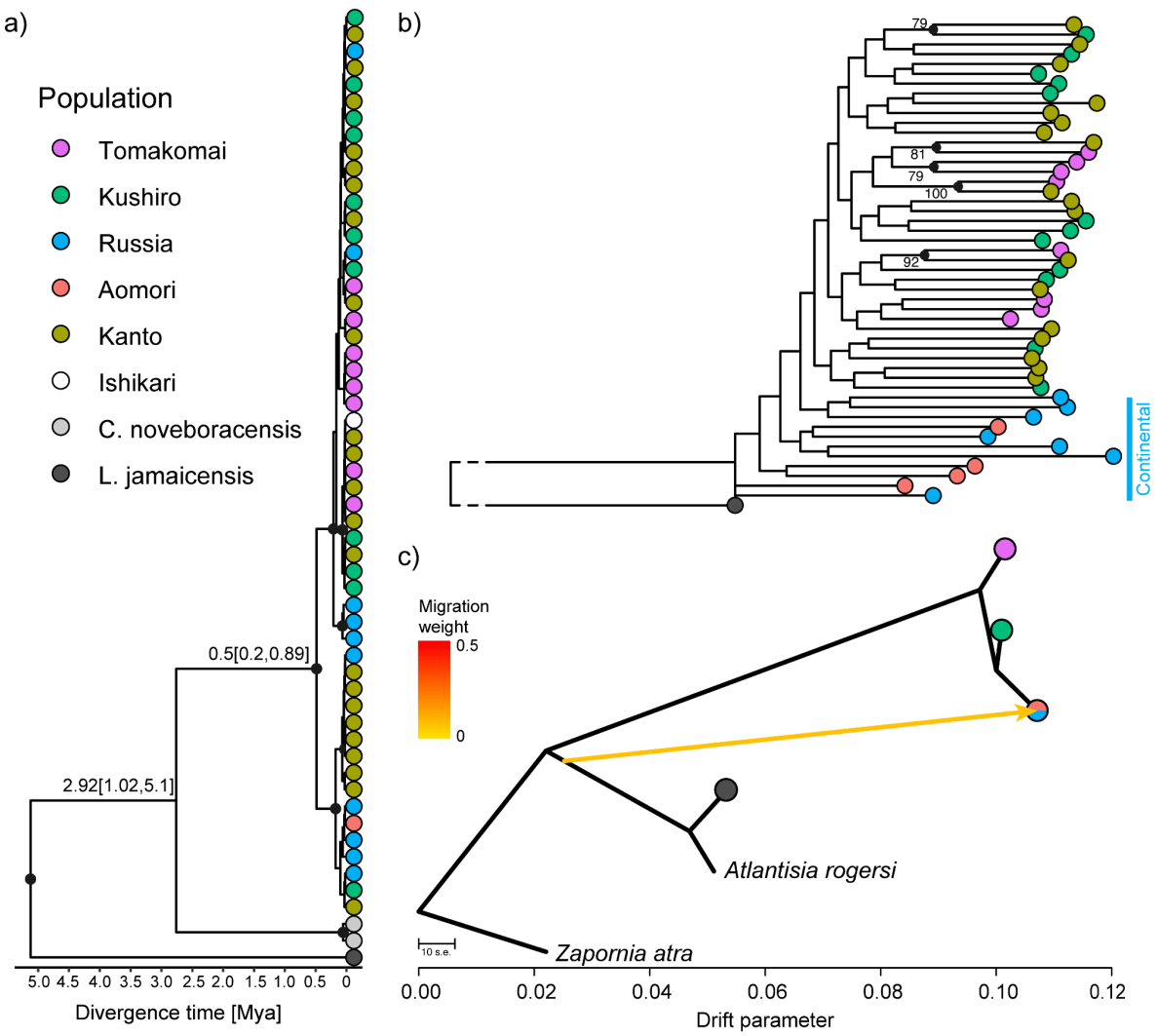

Figure 2 Phylogenetic trees of the Swinhoe's Rail reconstructed using different methods and genetic markers. (a) A mitochondrial phylogenetic tree using 772 bp of the *Cytb* gene inferred by Bayesian inference method. Nodes with black dots are supported nodes with posterior probabilities >0.9. Numbers at the nodes indicate median divergence time estimates with 95% highest posterior density within brackets. (b) A maximum likelihood (ML) tree reconstructed using 3,288 unlinked SNPs. Only minor nodes were supported with bootstrap values >70 (black nodes with bootstrap values alongside). Branch lengths are shortened at the hatched branches for drawing. (c) A population-level tree reconstructed by TreeMix using 1,912 unlinked polymorphic loci. Drift parameters represented by the branch length indicate strength of drift specific to each branch. The warm colored arrow indicates estimated gene flow. *C. noveboracensis* = *Coturnicops noveboracensis*, Yellow Rail; *L. jamaicensis* = *Laterallus jamaicensis*, Black Rail.

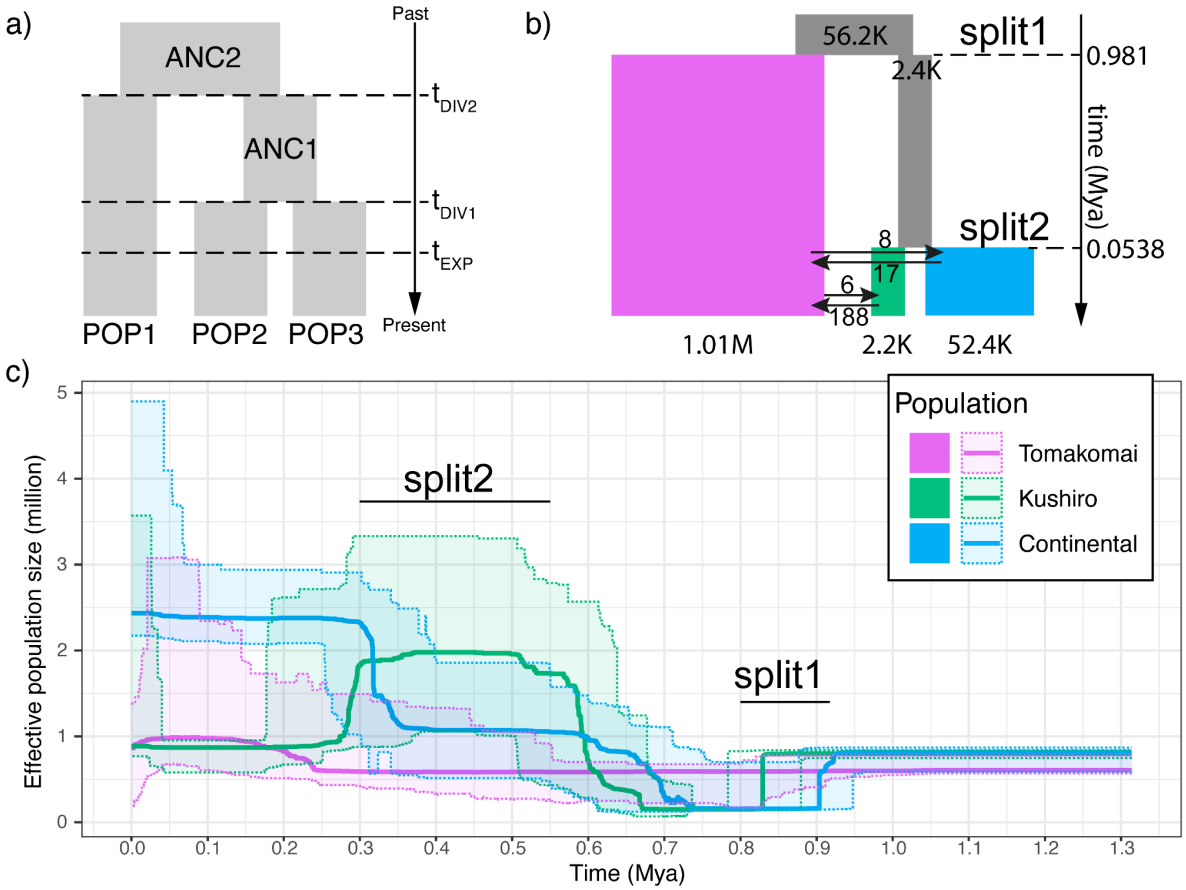

Figure 3 Demographic inference of the three genetic populations of the Swinhoe's Rail based on site frequency spectra, using (a, b) the fastsimcoal2 demographic modeling and (c) the Stairway Plot 2 analysis. (a) Generalized schematic representation of the six fastsimcoal2 models tested. Parameters with upper cases represent effective population size at specific event times t_{EVENTS}. Effective population size at the time t_{EXP} and migration are not shown. (b) The best-fit model with parameter estimates. Arrows with numbers between populations represent the number of migrants/year. Effective population size is shown in either million (M) or kilo (K) individuals scaled by mutation rate (1.91e-09) and generation time (1 year). (c) Changes in effective population size (in million) over time (in million years) are shown for each genetic population, using the same mutation rate and generation year as the fastsimcoal2 analysis. Bold solid lines and ribbons around them show median and 95% confidence intervals, respectively.

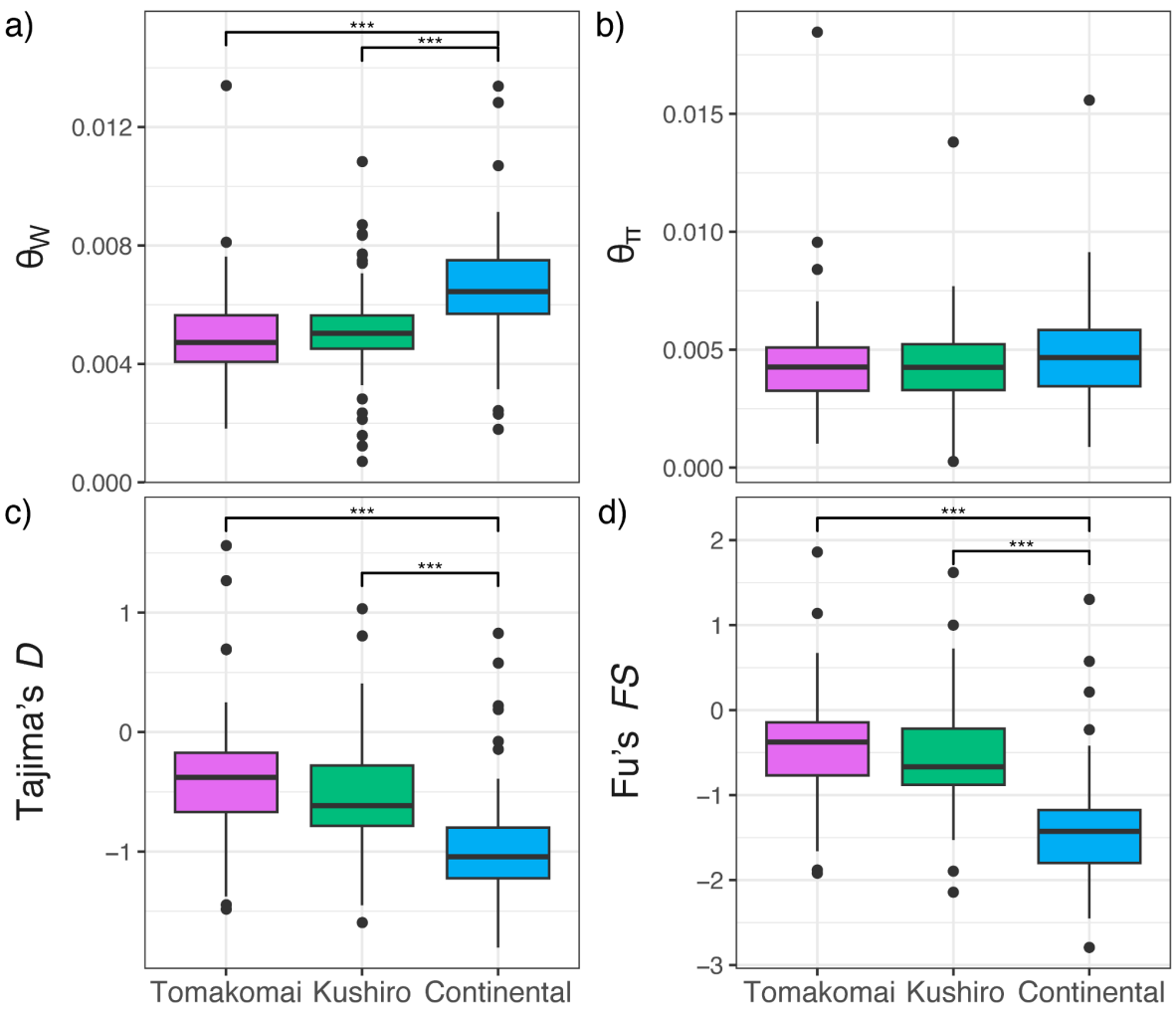

Figure 4 Comparison of scaffold-level genetic diversity indices across three genetic populations, including two genetic diversity indices, (a) Watterson's theta and (b) nucleotide diversity, and two neutrality test statistics, (c) Tajima's *D* and (d) Fu's *FS*. Population pairs with significant differences in the values of each index were annotated with asterisks (*** $p < 0.001$, no signature for $p > 0.05$ after Bonferroni correction), while those without annotation showed no significant differences.

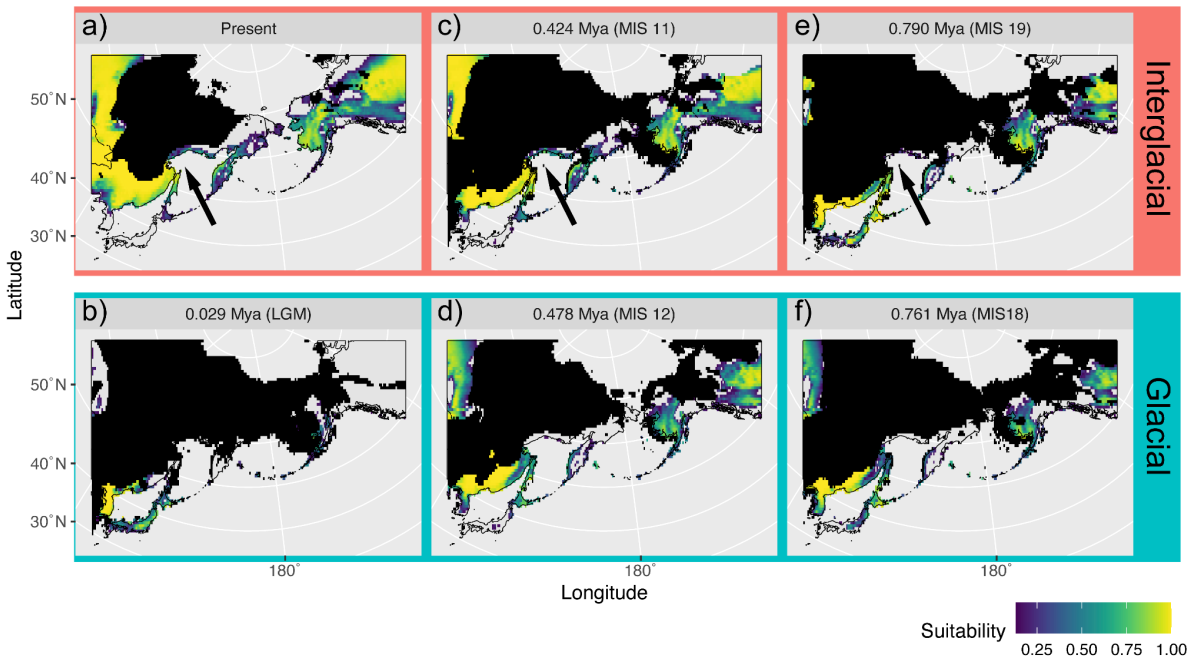

Figure 5 Species distribution models (SDMs) constructed using the present occurrence records of the Yellow Rail in North America were projected to Eurasia and North America in different time slices, including interglacial periods (a, c, e) as well as glacial periods (b, d, f). Color gradients from indigo to yellow represent environmental suitability for rails. Regions with black are where model extrapolation was restricted due to absence of the climate in the modeling domain. Maps are displayed using the Lambert Azimuthal Equal Area projection centered at the Pacific.

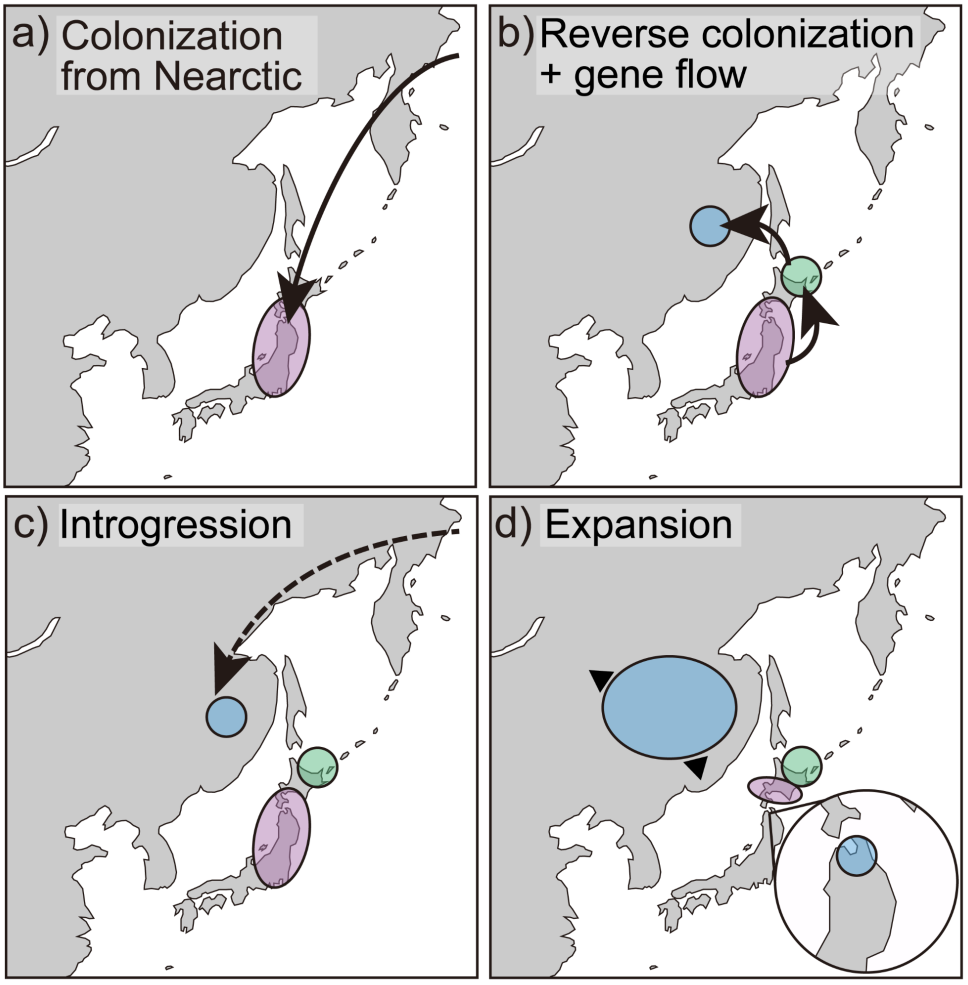

Figure 6 Schematic representation of our reverse colonization scenario of the Swinhoe's Rail. In this scenario, (a) the ancestral population of rails colonized the Japanese Archipelago from North America, then (b) the Asian continent was back colonized from this insular population along the island chain, and (c) the continental population received introgression from its close relative on the continent. Finally, (d) the continental population expanded while the insular population shrank its population size. Aomori was re-colonized by the continental population (inset).

References

Alföldi, J., Di Palma, F., Grabherr, M., Williams, C., Kong, L., Mauceli, E., Russell, P., Lowe, C. B., Glor, R. E., Jaffe, J. D., Ray, D. A., Boissinot, S., Shedlock, A. M., Botka, C., Castoe, T. A., Colbourne, J. K., Fujita, M. K., Moreno, R. G., ten Hallers, B. F., … Lindblad-Toh, K. (2011). The genome of the green anole lizard and a comparative analysis with birds and mammals. *Nature*, *477*(7366), Article 7366. https://doi.org/10.1038/nature10390

Anisimova, V., Anisimov, Y., Bastardot, M., & Heim, W. (2019). First observation of Swinhoe's Rail Coturnicops exquisitus in the delta of Selenga river. *Baikal Zoological Journal*, *2*(25), 112–113.

Aoki, D., Kinoshita, G., Kryukov, A. P., Nishiumi, I., Lee, S., & Suzuki, H. (2018). Quaternary-related genetic differentiation and parallel population dynamics of the Eurasian Jay (Garrulus glandarius) in the circum-Japan Sea region. *Journal of Ornithology*, *159*(4), 1087–1097. https://doi.org/10.1007/s10336-018-1573-9

Barbosa, A. M. (2015). fuzzySim: Applying fuzzy logic to binary similarity indices in ecology. *Methods in Ecology and Evolution*, *6*(7), 853–858. https://doi.org/10.1111/2041-210X.12372

Beichman, A. C., Huerta-Sanchez, E., & Lohmueller, K. E. (2018). Using genomic data to infer historic population dynamics of nonmodel organisms. *Annual Review of Ecology, Evolution, and Systematics*, *49*(1), 433–456. https://doi.org/10.1146/annurev-ecolsys-110617-062431

Bellemain, E., & Ricklefs, R. (2008). Are islands the end of the colonization road? *Trends in Ecology & Evolution*, *23*(8), 461–468. https://doi.org/10.1016/j.tree.2008.05.001

BirdLife International. (2023). *Species factsheet: Coturnicops exquisitus*. http://datazone.birdlife.org/species/factsheet/swinhoes-rail-coturnicops-exquisitus

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Boria, R. A., Brown, S. K., Matocq, M. D., & Blois, J. L. (2021). Genome-wide genetic variation coupled with demographic and ecological niche modeling of the dusky-footed woodrat

(Neotoma fuscipes) reveal patterns of deep divergence and widespread Holocene expansion across northern California. *Heredity*, *126*(3), Article 3. https://doi.org/10.1038/s41437-020-00393-7

Bouckaert, R., Vaughan, T. G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., Heled, J., Jones, G., Kühnert, D., De Maio, N., Matschiner, M., Mendes, F. K., Müller, N. F., Ogilvie, H. A., du Plessis, L., Popinga, A., Rambaut, A., Rasmussen, D., Siveroni, I., … Drummond, A. J. (2019). BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLOS Computational Biology*, *15*(4), e1006650. https://doi.org/10.1371/journal.pcbi.1006650

Burga, A., Wang, W., Ben-David, E., Wolf, P. C., Ramey, A. M., Verdugo, C., Lyons, K., Parker, P. G., & Kruglyak, L. (2017). A genetic signature of the evolution of loss of flight in the Galapagos cormorant. *Science*, *356*(6341), eaal3345. https://doi.org/10.1126/science.aal3345

Cerca, J., Cotoras, D. D., Bieker, V. C., De-Kayne, R., Vargas, P., Fernández-Mazuecos, M., López-Delgado, J., White, O., Stervander, M., Geneva, A. J., Andino, J. E. G., Meier, J. I., Roeble, L., Brée, B., Patiño, J., Guayasamin, J. M., Torres, M. de L., Valdebenito, H., Castañeda, M. del R., … Martin, M. D. (2023). Evolutionary genomics of oceanic island radiations. *Trends in Ecology & Evolution*, *38*(7), 631–642. https://doi.org/10.1016/j.tree.2023.02.003

Choi, J. Y., Dai, X., Alam, O., Peng, J. Z., Rughani, P., Hickey, S., Harrington, E., Juul, S., Ayroles, J. F., Purugganan, M. D., & Stacy, E. A. (2021). Ancestral polymorphisms shape the adaptive radiation of *Metrosideros* across the Hawaiian Islands. *Proceedings of the National Academy of Sciences*, *118*(37), e2023801118. https://doi.org/10.1073/pnas.2023801118

Clegg, S. M., Degnan, S. M., Kikkawa, J., Moritz, C., Estoup, A., & Owens, I. P. F. (2002). Genetic consequences of sequential founder events by an island-colonizing bird. *Proceedings of the National Academy of Sciences*, *99*(12), 8127–8132. https://doi.org/10.1073/pnas.102583399

de Jong, M. J., Niamir, A., Wolf, M., Kitchener, A. C., Lecomte, N., Seryodkin, I. V., Fain, S. R., Hagen, S. B., Saarma, U., & Janke, A. (2023). Range-wide whole-genome resequencing of the brown bear reveals drivers of intraspecies divergence. *Communications Biology*, *6*(1),

Article 1. https://doi.org/10.1038/s42003-023-04514-w

Excoffier, L., Foll, M., & Petit, R. J. (2009). Genetic consequences of range expansions. *Annual Review of Ecology, Evolution, and Systematics*, *40*(1), 481–501. https://doi.org/10.1146/annurev.ecolsys.39.110707.173414

Excoffier, L., Marchi, N., Marques, D. A., Matthey-Doret, R., Gouy, A., & Sousa, V. C. (2021). fastsimcoal2: Demographic inference under complex evolutionary scenarios. *Bioinformatics*, *37*(24), 4882–4885. https://doi.org/10.1093/bioinformatics/btab468

Fitak, R. R. (2021). OptM: Estimating the optimal number of migration edges on population trees using Treemix. *Biology Methods and Protocols*, *6*(1). https://doi.org/10.1093/biomethods/bpab017

Fox, E. A., Wright, A. E., Fumagalli, M., & Vieira, F. G. (2019). ngsLD: Evaluating linkage disequilibrium using genotype likelihoods. *Bioinformatics*, *35*(19), 3855–3856. https://doi.org/10.1093/bioinformatics/btz200

Frankham, R. (1995). Effective population size/adult population size ratios in wildlife: A review. *Genetics Research*, *66*(2), 95–107. https://doi.org/10.1017/S0016672300034455

Frankham, R. (1997). Do island populations have less genetic variation than mainland populations? *Heredity*, *78*(3), Article 3. https://doi.org/10.1038/hdy.1997.46

Frankham, R. (2005). Genetics and extinction. *Biological Conservation*, *126*(2), 131–140. https://doi.org/10.1016/j.biocon.2005.05.002

Fu, Y.-X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, *147*(2), 915–925. https://doi.org/10.1093/genetics/147.2.915

Garcia-R, J. C., Joseph, L., Adcock, G., Reid, J., & Trewick, S. A. (2017). Interisland gene flow among populations of the buff-banded rail (Aves: Rallidae) and its implications for insular endemism in Oceania. *Journal of Avian Biology*, *48*(5), 679–690. https://doi.org/10.1111/jav.01201

Garcia-R, J. C., & Matzke, N. J. (2021). Trait-dependent dispersal in rails (Aves: Rallidae): Historical biogeography of a cosmopolitan bird clade. *Molecular Phylogenetics and Evolution*, *159*, 107106. https://doi.org/10.1016/j.ympev.2021.107106

García-Verdugo, C., Sajeva, M., La Mantia, T., Harrouni, C., Msanda, F., & Caujapé-Castells, J. (2015). Do island plant populations really have lower genetic variation than mainland populations? Effects of selection and distribution range on genetic diversity estimates. *Molecular Ecology*, *24*(4), 726–741. https://doi.org/10.1111/mec.13060

Hart, S. P., Turcotte, M. M., & Levine, J. M. (2019). Effects of rapid evolution on species coexistence. *Proceedings of the National Academy of Sciences*, *116*(6), 2112–2117. https://doi.org/10.1073/pnas.1816298116

Harvey, M. G., Singhal, S., & Rabosky, D. L. (2019). Beyond reproductive isolation: Demographic controls on the speciation process. *Annual Review of Ecology, Evolution, and Systematics*, *50*(1), 75–95. https://doi.org/10.1146/annurev-ecolsys-110218-024701

Head, M. J. (2021). Review of the Early–Middle Pleistocene boundary and Marine Isotope Stage 19. *Progress in Earth and Planetary Science*, *8*(1), 50. https://doi.org/10.1186/s40645-021-00439-2

Heim, W., Trense, D., Heim, A., Kamp, J., Smirenski, S. M., Wink, M., & Wulf, T. (2019). Discovery of a new breeding population of the Vulnerable Swinhoe's Rail Coturnicops exquisitus confirmed by genetic analysis. *Bird Conservation International*, *29*(03), 454–462. https://doi.org/10.1017/S0959270918000138

Hume, J. P., & Martill, D. (2019). Repeated evolution of flightlessness in Dryolimnas rails (Aves: Rallidae) after extinction and recolonization on Aldabra. *Zoological Journal of the Linnean Society*, *186*(3), 666–672. https://doi.org/10.1093/zoolinnean/zlz018

Ikeda, Y., & Motokawa, M. (2021). Phylogeography of the Japanese greater horseshoe bat Rhinolophus nippon (Mammalia: Chiroptera) in Northeast Asia: New insight into the monophyly of the Japanese populations. *Ecology and Evolution*, *11*(24), 18181–18195. https://doi.org/10.1002/ece3.8414

Jezkova, T., Riddle, B. R., Card, D. C., Schield, D. R., Eckstut, M. E., & Castoe, T. A. (2015). Genetic consequences of postglacial range expansion in two codistributed rodents (genus Dipodomys) depend on ecology and genetic locus. *Molecular Ecology*, *24*(1), 83–97. https://doi.org/10.1111/mec.13012

Jønsson, K. A., Fabre, P.-H., Ricklefs, R. E., & Fjeldsa, J. (2011). Major global radiation of corvoid birds originated in the proto-Papuan archipelago. *Proceedings of the National Academy of Sciences*, *108*(6), 2328–2333. https://doi.org/10.1073/pnas.1018956108

Jønsson, K. A., & Holt, B. G. (2015). Islands contribute disproportionately high amounts of evolutionary diversity in passerine birds. *Nature Communications*, *6*(1), 8538. https://doi.org/10.1038/ncomms9538

Kapli, P., Yang, Z., & Telford, M. J. (2020). Phylogenetic tree building in the genomic age. *Nature Reviews Genetics*, *21*(7), Article 7. https://doi.org/10.1038/s41576-020-0233-0

Kass, J. M., Muscarella, R., Galante, P. J., Bohl, C. L., Pinilla‐Buitrago, G. E., Boria, R. A., Soley‐Guardia, M., & Anderson, R. P. (2021). ENMeval 2.0: Redesigned for customizable and reproducible modeling of species' niches and distributions. *Methods in Ecology and Evolution*, *12*(9), 1602–1608. https://doi.org/10.1111/2041-210X.13628

Kitazawa, M., Yamaura, Y., Senzaki, M., Hanioka, M., Ohashi, H., Oguro, M., Matsui, T., & Nakamura, F. (2022). Quantifying the impacts of 166 years of land cover change on lowland bird communities. *Proceedings of the Royal Society B: Biological Sciences*, *289*(1975), 20220338. https://doi.org/10.1098/rspb.2022.0338

Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics*, *15*(1), 356. https://doi.org/10.1186/s12859-014-0356-4

Kozlov, A. M., Darriba, D., Flouri, T., Morel, B., & Stamatakis, A. (2019). RAxML-NG: A fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics*, *35*(21), 4453–4455. https://doi.org/10.1093/bioinformatics/btz305

Krapp, M., Beyer, R. M., Edmundson, S. L., Valdes, P. J., & Manica, A. (2021). A statistics-based reconstruction of high-resolution global terrestrial climate for the last 800,000 years. *Scientific Data*, *8*(1), Article 1. https://doi.org/10.1038/s41597-021-01009-3

Leaché, A. D., Banbury, B. L., Felsenstein, J., de Oca, A. nieto-Montes, & Stamatakis, A. (2015). Short tree, long tree, right tree, wrong tree: New acquisition bias corrections for Inferring SNP phylogenies. *Systematic Biology*, *64*(6), 1032–1047.

https://doi.org/10.1093/sysbio/syv053

Leonardi, M., Hallett, E. Y., Beyer, R., Krapp, M., & Manica, A. (2023). pastclim 1.2: An R package to easily access and use paleoclimatic reconstructions. *Ecography*, *n/a*(n/a), e06481. https://doi.org/10.1111/ecog.06481

Leroy, T., Rousselle, M., Tilak, M.-K., Caizergues, A. E., Scornavacca, C., Recuerda, M., Fuchs, J., Illera, J. C., De Swardt, D. H., Blanco, G., Thébaud, C., Milá, B., & Nabholz, B. (2021). Island songbirds as windows into evolution in small populations. *Current Biology*, *31*(6), 1303-1310.e4. https://doi.org/10.1016/j.cub.2020.12.040

Levin, D. A. (1995). Metapopulations: An arena for local speciation. *Journal of Evolutionary Biology*, *8*(5), 635–644. https://doi.org/10.1046/j.1420-9101.1995.8050635.x

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv Preprint arXiv:1303.3997*.

Liu, X., & Fu, Y.-X. (2020). Stairway Plot 2: Demographic history inference with folded SNP frequency spectra. *Genome Biology*, *21*(1), 280. https://doi.org/10.1186/s13059-020-02196-9

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

MacArthur, R. H., & Wilson, E. O. (1963). An equilibrium theory of insular zoogeography. *Evolution*, *17*(4), 373–387.

Meisner, J., & Albrechtsen, A. (2018). Inferring Population Structure and Admixture Proportions in Low-Depth NGS Data. *Genetics*, *210*(2), 719–731. https://doi.org/10.1534/genetics.118.301336

Miya, A., Sannohe, S., Ebina, J., & Sekishita, H. (2005). A survey report on Yellow Rail (Coturnicops noveboracensis exquisitus) observed during the breeding season at Hotokenuma Reclaimed Land in Misawa, Aomori Prefecture. *Japanese Journal of Ornithology*, *54*(2), 104–107. https://doi.org/10.3838/jjo.54.104

Nadachowska-Brzyska, K., Konczal, M., & Babik, W. (2022). Navigating the temporal continuum of effective population size. *Methods in Ecology and Evolution*, *13*(1), 22–41. https://doi.org/10.1111/2041-210X.13740

Nadachowska-Brzyska, K., Li, C., Smeds, L., Zhang, G., & Ellegren, H. (2015). Temporal dynamics of avian populations during Pleistocene revealed by whole-genome sequences. *Current Biology*, *25*(10), 1375–1380. https://doi.org/10.1016/j.cub.2015.03.047

Nam, K., Mugal, C., Nabholz, B., Schielzeth, H., Wolf, J. B., Backström, N., Künstner, A., Balakrishnan, C. N., Heger, A., Ponting, C. P., Clayton, D. F., & Ellegren, H. (2010). Molecular evolution of genes in avian genomes. *Genome Biology*, *11*(6), R68. https://doi.org/10.1186/gb-2010-11-6-r68

O'Grady, P., & DeSalle, R. (2008). Out of Hawaii: The origin and biogeography of the genus Scaptomyza (Diptera: Drosophilidae). *Biology Letters*, *4*(2), 195–199. https://doi.org/10.1098/rsbl.2007.0575

Patiño, J., Carine, M., Mardulyn, P., Devos, N., Mateo, R. G., González-Mancebo, J. M., Shaw, A. J., & Vanderpoorten, A. (2015). Approximate bayesian computation reveals the crucial role of oceanic islands for the assembly of continental biodiversity. *Systematic Biology*, *64*(4), 579–589. https://doi.org/10.1093/sysbio/syv013

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, *190*(3–4), 231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026

Pickrell, J. K., & Pritchard, J. K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. *PLOS Genetics*, *8*(11), e1002967. https://doi.org/10.1371/journal.pgen.1002967

Pujolar, J. M., Blom, M. P. K., Reeve, A. H., Kennedy, J. D., Marki, P. Z., Korneliussen, T. S., Freeman, B. G., Sam, K., Linck, E., Haryoko, T., Iova, B., Koane, B., Maiah, G., Paul, L., Irestedt, M., & Jønsson, K. A. (2022). The formation of avian montane diversity across barriers and along elevational gradients. *Nature Communications*, *13*(1), Article 1. https://doi.org/10.1038/s41467-021-27858-5

Saitoh, T., Alström, P., Nishiumi, I., Shigeta, Y., Williams, D., Olsson, U., & Ueda, K. (2010). Old divergences in a boreal bird supports long-term survival through the Ice Ages. *BMC Evolutionary Biology*, *10*(1), 35. https://doi.org/10.1186/1471-2148-10-35

Seehausen, O. (2004). Hybridization and adaptive radiation. *Trends in Ecology & Evolution*, *19*(4), 198–207. https://doi.org/10.1016/j.tree.2004.01.003

Seehausen, O. (2013). Conditions when hybridization might predispose populations for adaptive radiation. *Journal of Evolutionary Biology*, *26*(2), 279–281. https://doi.org/10.1111/jeb.12026

Seki, S.-I., Takano, H., Kawakami, K., Kotaka, N., Endo, A., & Takehara, K. (2007). Distribution and genetic structure of the Japanese wood pigeon (Columba janthina) endemic to the islands of East Asia. *Conservation Genetics*, *8*(5), 1109–1121. https://doi.org/10.1007/s10592-006-9267-x

Sendell-Price, A. T., Ruegg, K. C., Robertson, B. C., & Clegg, S. M. (2021). An island-hopping bird reveals how founder events shape genome-wide divergence. *Molecular Ecology*, *30*(11), 2495–2510. https://doi.org/10.1111/mec.15898

Senzaki, M., Kitazawa, M., Sadakuni, T., & Takahashi, M. (2020). Breeding evidence of the vulnerable Swinhoe's Rail (Coturnicops exquisitus) in Japan. *The Wilson Journal of Ornithology*, *132*(3), 711–717. https://doi.org/10.1676/19-45

Skotte, L., Korneliussen, T. S., & Albrechtsen, A. (2013). Estimating individual admixture proportions from next generation sequencing data. *Genetics*, *195*(3), 693–702. https://doi.org/10.1534/genetics.113.154138

Smeds, L., Qvarnström, A., & Ellegren, H. (2016). Direct estimate of the rate of germline mutation in a bird. *Genome Research*, *26*(9), 1211–1218. https://doi.org/10.1101/gr.204669.116

Strimas-Mackey, M., Miller, E., & Hochachka, W. M. (2018). *auk: eBird Data Extraction and Processing with AWK. R package version 0.3.0* (0.3.0) [R]. https://cornelllabofornithology.github.io/auk/

Stroud, J. T., & Losos, J. B. (2016). Ecological opportunity and adaptive radiation. *Annual Review of Ecology, Evolution, and Systematics*, *47*(1), 507–532. https://doi.org/10.1146/annurev-ecolsys-121415-032254

Suyama, Y., & Matsuki, Y. (2015). MIG-seq: An effective PCR-based method for genome-wide single-nucleotide polymorphism genotyping using the next-generation sequencing platform. *Scientific Reports*, *5*(1), 16963. https://doi.org/10.1038/srep16963

Tajima, F. (1983). Evolutionary relationship of DNA sequences in finite populations. *Genetics*, *105*(2), 437–460. https://doi.org/10.1093/genetics/105.2.437

Tajima, F. (1989). The effect of change in population size on DNA polymorphism. *Genetics*, *123*(3), 597–601. https://doi.org/10.1093/genetics/123.3.597

Taylor, B., & Sharpe, C. J. (2020). Swinhoe's Rail (Coturnicops exquisitus), version 1.0. *Birds of the World*. https://doi.org/10.2173/bow.swirai1.01

Taylor, S. A., & Larson, E. L. (2019). Insights from genomes into the evolutionary importance and prevalence of hybridization in nature. *Nature Ecology & Evolution*, *3*(2), Article 2. https://doi.org/10.1038/s41559-018-0777-y

Tindall, J. C., Haywood, A. M., Salzmann, U., Dolan, A. M., & Fletcher, T. (2022). The warm winter paradox in the Pliocene northern high latitudes. *Climate of the Past*, *18*(6), 1385–1405. https://doi.org/10.5194/cp-18-1385-2022

Twyford, A. D., & Ennos, R. A. (2012). Next-generation hybridization and introgression. *Heredity*, *108*(3), Article 3. https://doi.org/10.1038/hdy.2011.68

Vamosi, J. C., Armbruster, W. S., & Renner, S. S. (2014). Evolutionary ecology of specialization: Insights from phylogenetic analysis. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1795), 20142004. https://doi.org/10.1098/rspb.2014.2004

Vellend, M., & Geber, M. A. (2005). Connections between species diversity and genetic diversity. *Ecology Letters*, *8*(7), 767–781. https://doi.org/10.1111/j.1461-0248.2005.00775.x

Watterson, G. A. (1975). On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, *7*(2), 256–276. https://doi.org/10.1016/0040-5809(75)90020-9

Weir, J. T., & Schluter, D. (2008). Calibrating the avian molecular clock. *Molecular Ecology*, *17*(10), 2321–2328. https://doi.org/10.1111/j.1365-294X.2008.03742.x

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

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

357(6354), eaam8326. https://doi.org/10.1126/science.aam8326

Wilson, E. O. (1961). The nature of the taxon cycle in the Melanesian ant fauna. *The American Naturalist*, *95*(882), 169–193.

Wulf, T., Heim, W., & Thomas, A. (2017). Mysterious calls from the swamp: The song of the Vulnerable Swinhoe's Rail Coturnicops exquisitus. *BirdingAsia*, *27*, 49–53.

Zink, R. M., & Gardner, A. S. (2017). Glaciation as a migratory switch. *Science Advances*, *3*(9), e1603133. https://doi.org/10.1126/sciadv.1603133

Conflict of Interest

The authors have no conflict of interest to declare.

Biosketch

Daisuke Aoki is an evolutionary ecologist, who is specialized in the fields of phylogeography, population genetics and avian migration. His research aim is to bridge a gap between microevolution and macroevolution by asking how biogeographic histories of organisms interacted with natural selection, stochastic processes, and evolutionary constraints. He achieves this aim through multifaceted approaches, including genetics and genomics, spatial modeling, mathematical simulation, and field biologging approaches.

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 supports 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.