

Migration-tracking integrated phylogeography supports long-distance dispersal-driven divergence for a migratory bird species in the Japanese archipelago

Daisuke Aoki¹, Haruna Sakamoto¹, Munehiro Kitazawa¹, Alexey Kryukov², and Masaoki Takagi³

¹Hokkaido University

²Russian Academy of Sciences Far Eastern Branch

³Hokkaido University Faculty of Science Graduate School of Science

October 26, 2020

Abstract

Aim: In theory, long-distance dispersal (LDD) outside a species' range contributes to genetic divergence. However, previous studies have not discriminated this process from vicariant speciation in migratory bird species. We conducted an integrative phylogeographic approach to test the LDD hypothesis, which predicts that a Japanese migratory bird subspecies diverged from a population in the coastal region of the East China Sea (CRECS) via LDD over the East China Sea (ECS). **Location:** East Asia **Taxon:** Brown Shrike (*Lanius cristatus*) **Methods:** Both a haplotype network and a multi-locus gene network of its three subspecies were reconstructed to examine from which continental population the Japanese subspecies diverged. A species distribution model (SDM) for the Japanese subspecies was constructed using bioclimatic variables under the maximum entropy algorithm. It was projected to the climate of the last glacial maximum (LGM) to infer the candidate source area of colonisation. A migratory route of the Japanese subspecies, which possibly reflects a candidate past colonisation route, was tracked by light-level geolocators. **Results:** Molecular phylogenetic networks suggest that the Japanese subspecies diverged from a population in the CRECS. The SDM inferred that the emerged continental shelf of the ECS and the present CRECS were suitable breeding areas for the Japanese subspecies during the LGM. A major migratory route for the Japanese subspecies was inferred between the CRECS and the Japanese archipelago across the ECS. **Main conclusions:** Our integrative approach supported the LDD hypothesis for divergence of the Japanese subspecies of the Brown Shrike. Shrinkage and expansion of the ECS may have been responsible for successful colonisation and isolation of the new population. Vicariance was inferred for divergence of the subspecies in the northeast Asian continent from the Japanese population. Our framework provides a new phylogeographic scenario in this region, and discriminating LDD and vicariance models should improve our understanding of the phylogeographic histories of migratory species.

i. Title:

Migration-tracking integrated phylogeography supports long-distance dispersal-driven divergence for a migratory bird species in the Japanese archipelago

ii. A short running title:

Divergence scenario for a Japanese migrant

iii. Author:

AOKI Daisuke¹, SAKAMOTO Haruna¹, KITAZAWA Munehiro², KRYUKOV Alexey P³ & TAKAGI Masaoki^{4,*}

iv. Author's institutional affiliations

¹Department of Natural History Sciences, Graduate School of Science, Hokkaido University, N10W8, Kita-ku, Sapporo 060-0810, Japan

²Frontiers in Environmental Sciences, Graduate School of Agriculture, Hokkaido University, N9W9, Kita-ku, Sapporo 060-8589, Japan

³Laboratory of Evolutionary Zoology and Genetics, Federal Scientific Center of the East Asia Terrestrial Biodiversity, Far Eastern Branch of the Russian Academy of Sciences, Vladivostok 690022, Russia

⁴Department of Natural History Sciences, Faculty of Science, Hokkaido University, N10W8, Kita-ku, Sapporo 060-0810, Japan

* Correspondence: email address mtakagi@eis.hokudai.ac.jp, telephone number +81(11)706-4464

ORCID: AOKI Daisuke, 0000-0002-5780-8487

TAKAGI Masaoki, 0000-0001-8308-899X

KRYUKOV Alexey P, 0000-0001-7010-7338

v. Acknowledgements

We thank Hiroaki Matsumiya, Seiichi Hara, Sachiko Endo, Junco Nagata, Yaroslav A. Red'kin, Ivan M. Tiunov, and Yamashina Institute for Ornithology for providing us with samples of shrikes from various regions. The manuscript was greatly improved by valuable comments from Professor Itsuro Koizumi, Keisuke Atsumi, and laboratory members of the Biodiversity course, Hokkaido University. Analyses were greatly improved using high throughput computers, which Professor Itsuro Koizumi, Professor Hiroshi Kajihara, Professor Toru Kato and Professor Keiichi Kakui kindly allowed us to use. We are grateful for the considerable help received from many daily field assistants: Noritomo Kawaji, Takayuki Kawahara, Yusuke Nishida, Mari Esashi, Ryotaro Sato and Riku Chiba. This study was partly conducted with a support of a grant-in-aid for Scientific Research (C) to MT (no. 16K14796 and 16H04737) and to DA (no. 19J21406) from the Japanese Society for the Promotion of Science (JSPS).

vi. Abstract and keywords:

Aim : In theory, long-distance dispersal (LDD) outside a species' range contributes to genetic divergence. However, previous studies have not discriminated this process from vicariant speciation in migratory bird species. We conducted an integrative phylogeographic approach to test the LDD hypothesis, which predicts that a Japanese migratory bird subspecies diverged from a population in the coastal region of the East China Sea (CRECS) via LDD over the East China Sea (ECS).

Location : East Asia

Taxon : Brown Shrike (*Lanius cristatus*)

Methods : Both a haplotype network and a multi-locus gene network of its three subspecies were reconstructed to examine from which continental population the Japanese subspecies diverged. A species distribution model (SDM) for the Japanese subspecies was constructed using bioclimatic variables under the maximum entropy algorithm. It was projected to the climate of the last glacial maximum (LGM) to infer the candidate source area of colonisation. A migratory route of the Japanese subspecies, which possibly reflects a candidate past colonisation route, was tracked by light-level geolocators.

Results : Molecular phylogenetic networks suggest that the Japanese subspecies diverged from a population in the CRECS. The SDM inferred that the emerged continental shelf of the ECS and the present CRECS were suitable breeding areas for the Japanese subspecies during the LGM. A major migratory route for the Japanese subspecies was inferred between the CRECS and the Japanese archipelago across the ECS.

Main conclusions : Our integrative approach supported the LDD hypothesis for divergence of the Japanese subspecies of the Brown Shrike. Shrinkage and expansion of the ECS may have been responsible for successful colonisation and isolation of the new population. Vicariance was inferred for divergence of the subspecies in the northeast Asian continent from the Japanese population. Our framework provides a new phylogeographic scenario in this region, and discriminating LDD and vicariance models should improve our understanding of the phylogeographic histories of migratory species.

Keywords: Light-level geocator, long-distance dispersal, migratory route, phylogeography, species distribution modelling, the Brown Shrike, the Japanese archipelago, vicariance

vii. Main text

Introduction:

High dispersal propensities due to animal migratory behaviour, and how they contribute to speciation, is an infrequently studied area of biogeography (Greenberg & Marra, 2005; Winger, Auteri, Pegan, & Weeks, 2019). High dispersal ability often leads to long-distance dispersal (LDD) well outside a species' known range. This phenomenon is called vagrancy and derives mostly from migratory populations that have, for instance in the case of birds, drifted past or overshot their anticipated destinations, influenced by wind during their seasonal migration (Newton, 2008). In theory, the frequency of vagrancy is a key determinant of its effect on a population that receives vagrants (Rose & Polis, 2000). A high frequency of vagrancy may increase the chance of successful establishment of an allopatric population (O'Connor, 1986; but see Lees & Gilroy, 2014), whereas it may prevent isolation of the new population and hence preclude genetic divergence (Harvey et al. 2019). Paleogeographical change is a candidate mechanism that allows divergence of a new population since it changes the frequency of LDD over geological timescales (Weeks & Claramunt, 2014). For example, vagrancy will decrease in frequency when a geographic barrier between the source and new populations expands, because vagrancy decreases with increasing distance (Lees & Gilroy, 2014; O'Connor, 1986). Many phylogeographic studies of migratory species, however, have anecdotally assumed that previously known biogeographical barriers subdivided a formerly continuous population, i.e. vicariance (e.g. Weir & Schluter, 2004; Zink, Pavlova, Rohwer, & Drovetski, 2006). While vicariance has provided sufficient explanations for many cases of genetic divergence of migratory species, an alternative scenario — that a paleogeographical change facilitated LDD-driven divergence — has never been empirically tested. This might be because many currently migratory bird species are continental and widespread (Somveille, Rodrigues, & Manica, 2018), thus their phylogeographic processes of speciation could have hardly been inferred without the assumption of vicariance (Albert, Schoolmaster, Tagliacollo, & Duke-Sylvester, 2017).

Owing to its insular nature, well-documented paleogeography, and relatively high endemism of breeding migratory passerine lineages (Saitoh et al., 2010, 2015), the Japanese archipelago provides a suitable system within which to identify a biogeographic mode of speciation of a migratory lineage. During glacial periods of the Quaternary (approximately 2.7 million years ago to the present), due to lower sea levels, some parts of the Japanese archipelago were connected to the continent via land bridges, while other parts remained separated (Gallagher et al., 2015; Matsuzaki, Itaki, & Tada, 2019; Ohshima, 1990; Ota, 1998). Most terrestrial animal species currently occurring in Japan colonized the archipelago from the East Asian continent during glacial periods, and they subsequently diverged as Japanese endemics (McKay, 2012; Motokawa, 2017). Moreover, source continental populations could be either on the coastal region of the East China Sea (CRECS) or northeast Asian continent (the Korean Peninsula, northern China to Far East Russia), where different species or populations occupy their breeding ranges (e.g. Dong et al., 2015; Päckert et al., 2011; Saitoh et al., 2015; Zhao et al., 2017). This system, predicting from which regional population and along which route the Japanese archipelago was colonized, will help indicate which of the two following hypotheses is more plausible. 1) Colonisation across the open East China Sea (ECS) from the CRECS, implying LDD-driven divergence; hereafter called the LDD Hypothesis (Figure 1a). 2) Colonisation by means of a range shift over a land bridge from the Korean Peninsula to Japan (where there is currently the Tsushima Strait), implies that the disappearance of the land bridge led to vicariant speciation; hereafter called the Vicariance Hypothesis (Figure 1b).

We conducted integrative phylogeography by means of phylogenetic networks, a species distribution model (SDM), and migration tracking using light-level geolocators, to infer the past colonisation of the Japanese archipelago. Phylogenetic networks are informative for disentangling relationships of intraspecific genetic variations both in haplotype and multi-locus data (Huson & Bryant, 2006; Mardulyn, 2012). SDMs can infer suitable breeding distributions of species during the last glacial maximum (LGM), ca. 20,000 years ago (20 kya) from present species distributions (Elith et al., 2011). Any continental range that was also suitable during a glacial period may be interpreted as a potential area from which colonisation occurred (Zink & Gardner, 2017). An avian migratory route over a sea barrier between the Japanese archipelago and the Asian continent may be interpreted as a candidate for the past colonisation route to the Japanese archipelago for the following reasons. First, an existing migration route over a large geographic barrier may reflect LDD in the past when the barrier was smaller or absent (Newton, 2008; Winger et al., 2019). Second, the migratory route also partly retraces past shifts in the breeding range (Alvarado, Fuller, & Smith, 2014; Newton, 2008; Ruegg, Hijmans, & Moritz, 2006; Winger et al., 2019). Therefore, both modes of past colonisation may be reflected in the migratory route over a present sea barrier in this system.

Results of the three components can be differently predicted from the LDD and vicariance hypotheses in this framework. (1) The LDD hypothesis predicts that a genetic split between the archipelagic population and a population in the CRECS will be inferred by phylogenetic networks. For the archipelagic population, CRECS is predicted to have been suitable during the glacial period by the SDM, and a migratory route across the ECS between the Japanese archipelago and the CRECS is expected to have been likely (Figure 1a). (2) The vicariance hypothesis predicts that the archipelagic population genetically nests within, or is not well divergent from, the population in the northeast Asian continent, because the land bridge between the Korean Peninsula and the Japanese archipelago repeatedly formed and disappeared. For the archipelagic population, either current northern China or the Korean Peninsula are predicted to have been suitable during a glacial period by the SDM, and a migratory route is expected over the Korean Peninsula from and to the Japanese archipelago (Figure 1b).

We studied a long-distance migratory passerine, the Brown Shrike (*Lanius cristatus*), as our first attempt to test these hypotheses given the breeding distribution of its three subspecies. *L. c. superciliosus* breeds in the Japanese archipelago, *L. c. lucionensis* breeds in the CRECS, and *L. c. cristatus* breeds in the northeast Asian continent and most of Sakhalin Island (Figure 1c; Lefranc & Worfolk, 1997). With a dated phylogenetic tree, we provided a phylogeographic scenario of the divergence of the Brown Shrike based on the results obtained from our integrative approach. From this scenario, we discuss the paleogeographical context in which LDD could have contributed to the genetic divergence of a migratory species.

Materials and Methods:

Study species and field procedure

To track the migratory routes of *L. c. superciliosus*, we captured 25 adult male shrikes in Hokkaido and marked them with leg rings. We attached MK5740 geocator tags (BioTrack) with leg-loop harness strings (Rappole & Tipton, 1991). We ensured that the total mass of the system (< 1.1 g) did not exceed 5% of the birds' mass (Bridge et al., 2013). Eight birds (32%) returned to the breeding study site in 2018. Deployment of tags did not significantly change the return rate when compared with a previous report by Takagi (2003) (chi-square test for independence; $\chi^2 = 0.11$, $df = 1$, $p = 0.74$). Four males lost their tags before recapture. One male marked with leg rings disappeared before breeding in 2018 and did not return to the site in 2019, so it remained uncaptured (the presence of this bird's geocator was not confirmed).

Laboratory procedure and haplotyping

We collected tissue and blood samples from the Russian Far East for *L. c. cristatus*, from Hokkaido and Nagano prefectures for *L. c. superciliosus* and from the Ryukyus for *L. c. lucionensis*. A total of 95 genetic samples was obtained. We also used database sequences of mitochondrial genes from GenBank and BOLD (Ratnasingham & Hebert, 2007) for a coverage of its wider distribution range (Figure 1c). Since the Ryukyus and the Philippines are occupied only by non-breeding (i.e. migrating and wintering) *L.*

c. lucionensis (Lefranc & Worfolk, 1997; Severinghaus, 1996), they were used as representatives of *L. c. lucionensis*. See Table S1.1 for details about samples and sequences.

Total genomic DNA was extracted from blood and tissue samples using DNeasy Blood & Tissue Kit (Qiagen). Two mitochondrial genes (complete cytochrome b (*cytb*), and partial cytochrome oxidase c subunit I (COI)) and two autosomal introns (myoglobin intron-2 (MB), transforming growth factor beta 2 intron-5 (TGFb2)) were amplified by polymerase chain reaction (PCR). The autosomal introns were sequenced only for a subset of the samples (Table S1.1). Automated sequencing was run on either an ABI7130 or ABI3130 Genetic Analyzer (ABI). Sequences were aligned using ProSeq v. 3.5 software (Filatov, 2009). To determine haplotypes of nuclear introns with multiple heterozygous sites, we used PHASE v. 2.1.1 (Stephens et al., 2001; Stephens & Donnelly, 2003). Haplotypes inferred with support values below 0.5 were not used in the following analyses. See Appendix S1 for details about primers, PCR reactions, and PHASE analysis.

Phylogenetic analyses

We constructed a median-joining (MJ) haplotype network (using NETWORK 5.0.0.3; Bandelt, Forster, & Rohl, 1999) using partial COI (521 bp) from 108 samples. We also constructed a multi-locus network of four loci (*cytb*, COI, MB and TGFb2) under the NeighborNet algorithm (using Splits Tree 4.14.4; Huson & Bryant, 2006). We only included individuals with sequences of all four loci and calculated uncorrected pairwise (p) distances among sequences in each locus using MEGA v. 7 (Kumar, Stecher, & Tamura, 2016). Then, we combined the four p -distance matrices to produce a standardized genetic distance matrix among individual samples using POFAD v. 1.03 (Joly & Bruneau, 2006), which was used as an input file for the NeighborNet analysis.

We reconstructed mitochondrial phylogeny by using the two mitochondrial sequences. A dataset consisting of unique haplotypes of concatenated sequences of *cytb* and COI was prepared for tree reconstruction by the Bayesian inference (BI) method using BEAST v. 2.4.8 (Bouckaert et al., 2014). In the BEAST analysis, two genes were independently assigned to different partitions with the best substitution models determined by the lowest Bayesian Information Criterion on MEGA v. 7; HKY + G was selected for each gene ($\alpha = 0.10$ for *cytb* and 0.19 for COI). Application of the standard avian clock rate for *cytb*, 0.0105/lineage/million years was justified since shrikes are passerines (Nabholz, Lanfear, & Fuchs, 2016; Weir & Schluter, 2008). Markov chain Monte Carlo (MCMC) sampling proceeded for 50 million states with 100 thousand pre-burn-in chains. Convergence for each parameter was checked through Tracer v. 1.7.1 software (Rambaut, Drummond, Xie, Baele, & Suchard, 2018). We followed Aoki et al. (2018) for other settings.

Species distribution modelling

We modelled the breeding range of *L. c. superciliosus* and projected it into the LGM. Data on occurrence in Japan were obtained from a summary of breeding localities based on the National Surveys conducted by the Ministry of Environment of Japan and literature (M Kitazawa and his colleagues (in prep.) hereafter referred to as ‘review data’). Additionally, on 12 May 2019 we downloaded data from Global Biodiversity Information Facility (GBIF; <https://www.gbif.org/>, <http://doi.org/10.15468/dl.igj5uw>). A polygon bounding the presumed distribution of *L. c. superciliosus* according to Lefranc & Worfolk (1997) was used to clip the data. Records were limited only to the breeding season (May to August) to avoid integration of migrating birds. After a few quality-checking processes, final coordinates were thinned to include only one occurrence per 20 km square covering the Japanese archipelago (Figure S1.1).

We constructed a model with the final coordinate data. Our purpose is to infer the continental area that might have been suitable for *L. c. superciliosus* as its breeding habitat during the LGM. Therefore, any ecological differences between the current archipelago and the continental area need to be integrated in the model (Elith et al., 2011). We clipped the environmental data to include the present distribution of the Brown Shrike regardless of subspecies and this environmental data was defined as the modelling domain. Of the available WorldClim data (version 1.4; Hijmans et al. 2005), only one variable from each pair of correlated variables ($|r| > 0.7$) was used, resulting in seven variables: bio1, 2, 3, 5, 12, 14 and 15.

We used the R package ‘ENMeval’ v. 0.3.0 (Muscarella et al., 2014) to execute the maximum entropy (MaxEnt) algorithm (Phillips, Anderson, & Schapire, 2006) across a range of several settings. Combinations of three feature classes (linear, a combination of linear and quadratic, and hinge) and regularization multipliers ranging from 0.5 to 3 (increments of 0.5) were tested. The ‘maxent.jar’ option was used as the model algorithm. From the geographical modelling domain defined above, 10,000 background points were sampled randomly. The ‘block’ method was used to partition data for cross-validation since it assesses model transferability better than other methods (Muscarella et al., 2014). The hinge feature, with a regularization multiplier of 1.5, was selected by the lowest AICc. The present distribution was predicted, and then it was projected to the LGM environmental data using cloglog transformation. The LGM environmental data were drawn by Hijmans et al. (2005) from three different climate models: the community climate system model (CCSM, Gent et al., 2011), the model for interdisciplinary research on climate model (MIROC-ESM, Hasumi & Emori, 2004) and the Max-Planck-Inst. Für Meteorologie climate model (MPI-ESM-P, Liepert & Lo, 2013). Multivariate environmental similarity surface (MESS) was also calculated and extrapolation was restricted to regions indicated by negative values from this analysis (See Appendix S1 for the detailed procedures on the SDM analysis).

Migratory route analysis

Data retrieved from light-level geolocators were analysed in the R package ‘GeoLight’ version 2.0.0 (Lisovski & Hahn, 2012) and ‘SGAT’ version 0.1.3 (Wotherspoon, Sumner, & Lisovski, 2015). We prepared an R script by following the manual provided by Lisovski et al. (2019). We determined sunset and sunrise transitions for specific dates by light intensity thresholds of 0.5. We incorporated a statistical modelling approach, called the ‘group model’ in a Bayesian framework implemented in ‘SGAT’. This model refines the estimated locations by means of three components: the twilight model, a movement model, and spatial masks. The twilight model incorporates the deviation between determined and actual twilight times in calibration periods of each tag. We used twilight data of 26 days after geocator deployment as the calibration period. We fitted gamma distribution as an error distribution of detection of twilight. To define a movement model and spatial masks separately for stationary and movement periods, the ChangeLight function in ‘GeoLight’ was applied to discriminate between these periods from the twilight data. A probability threshold for discrimination was set to 0.9. For the movement model, we set a flight speed prior with a gamma distribution for periods of movement (shape = 13, rate = 0.3), determined allowing for the average air speed of 12.9 m/s for *Lanius collurio* (Bruderer & Boldt, 2001). We conducted a sensitivity analysis for the gamma distribution, and we found that priors did not greatly influence our interpretation of the results (Table S1.2). In the group model, twilight times grouped for a stationary period are treated as records at a single location, thus no movement model was defined. For a probability mask, birds must cross the sea between the Japanese archipelago and the continent and must stopover on land. Therefore, a spatial mask for stationary periods was designed only to include terrestrial land while both sea and land areas were allowed for movement periods. We slightly modified our analysis for bird #V5604-025 because its geocator stopped recording during the autumn equinox period (see Appendix S1 for details).

We interpolated latitudinal estimations around the equinox periods by setting a ‘tol’ value between 0.08 and 0.1. As burn-in and fine tuning, we ran 120,000 chains with three iterations, and the final paths were determined by the last 15,000 samples with three iterations. Regarding large uncertainties involved in migration estimations by geolocators (Lisovski et al., 2019), we standardized the posterior probabilities of location estimates on each raster cell, calculated their means among the three individuals, and we presented them as a probability distribution in a region of interest where we expect birds to retrace the route of past colonisation. We also inferred median paths of the three migratory tracks. Because birds spend more time at stopover sites than at an instantaneous point along a migratory path, the former will be indicated with relatively higher probabilities than the latter.

Results:

Phylogenetic analyses

An MJ haplotype network of COI was constructed using 108 sequences from the species' range (Figure 2a). This network inferred three genetic groups, which were supported as distinct clades in the BI phylogenetic tree (Figure S1.2); the archipelagic, northern, and southern clades possibly reflected the three subspecies, *L. c. superciliosus*, *L. c. cristatus*, and *L. c. lucionensis*, respectively (Figure 1c). In this network, a split between the southern and archipelagic clades was observed with nine mutational gaps. The northern clade was genetically most distant from the southern clade and its major haplotype (H1) was separated only by two mutational positions from one of the archipelagic major haplotypes. A similar structure was supported by the multi-locus network including both the mitochondrial and nuclear loci. Two individuals appeared to be intermediate between the southern and northern clusters (Figure 2b).

The southern clade was inferred as the basal lineage of the Brown Shrike by the BI tree (Figure S1.2). The first split occurred between the southern and the archipelagic/northern clades, around 0.80 [0.55, 1.10] million years ago (Mya), corresponding to the late Early to Middle Pleistocene while the northern and the archipelagic clade diverged around 0.32 [0.20, 0.47] Mya, during the Middle Pleistocene (numbers within square brackets indicate 95% highest posterior density).

Species distribution modelling

The present breeding distribution of *L. c. superciliosus* was modelled based on presence-only data and seven non-correlated bioclimatic variables using the MaxEnt algorithm (Figure 3). Minimum training presence omission rates for test localities were 1.9% and the model was evaluated to estimate suitable habitat without problematic overfitting. The present prediction of suitable areas (Figure 3a) reflected the taxon's known breeding range according to Lefranc & Worfolk (1997).

Projection of the model to three climatic models of the LGM (Figure 3b-d) inferred that the coastal area of the currently submerged ECS continental shelf, and an inner continental range corresponding to the present CRECS, were predicted to be suitable breeding habitat for *L. c. superciliosus*. Suitability for these regions was estimated to be relatively lower than in the Japanese archipelago, within which suitable breeding habitat receded southwards but largely remained during the LGM. Most of the Ryukyu islands were not predicted to be suitable for the species during the LGM, based on our model.

Migratory route analysis

Under the Bayesian framework, autumn and spring migration routes were inferred for three individual *L. c. superciliosus* (Figure 4, Figure S1.4). Full annual data were retrieved from two tags, and similarity was observed between their autumn and spring routes. One tag stopped recording on the continent due to cell failure three months after deployment. Despite uncertainties involved with location estimates, higher mean probabilities for migratory locations were observed among the three individuals both in the western part of the Japanese archipelago and along the CRECS to northern Taiwan. The probability map and the median migratory tracks suggest that birds migrated across the ECS between these two regions and did not detour over the Korean Peninsula. The data from the two successfully retrieved geolocators revealed that both shrikes had wintered in the Sunda Islands, specifically on Java and Flores.

Discussion:

The results from the three approaches supported the LDD hypothesis for genetic divergence of the Brown Shrike in the Japanese archipelago (Figure 1a, 5a-b). A phylogenetic split between the southern and archipelagic clades was inferred in the gene networks (Figure 2), suitable glacial breeding habitat for *L. c. superciliosus* was estimated to have existed around the CRECS by the SDM (Figure 3), and its current migratory route was observed between the CRECS and the Japanese archipelago across the ECS (Figure 4). Our results cannot reflect gradual expansion of a population via normal dispersal over a terrestrial landscape. This is because, during the glacial periods around which the divergence was estimated, the gulf separating the Japanese archipelago, the Korean Peninsula and the CRECS was relatively wide (Figure 5; Matsuzaki et al., 2019; Ota, 1998), compared to the LGM to which we projected our SDM (Figure 3). LDD colonisation in a stepping-stone manner along the Ryukyu islands from the CRECS or Taiwan may also be less likely. This

is supported by our results showing that most of the Ryukyu islands were predicted to have been unsuitable during a glacial period (Figure 3b-d), and the migratory route was not inferred along them (Figure 4), unlike the routes reported for other species (Shiu et al., 2006; Koike, Hijikata, Higuchi, 2016). Meanwhile, since northern Taiwan was connected to the CRECS during this period (Figure 5), the inferred migration aligns with our prediction. Therefore, within our hypothesis-testing framework, we imply that our results reflect the LDD-driven divergence of a Japanese migratory subspecies across the ECS.

Hypothetical phylogeographic scenario of the Brown Shrike

The southern clade represented the basal lineage of the Brown Shrike (Figure S1.2), suggesting that the ancestral breeding area of this species may have been located in the eastern part of the Asian continent, including CRECS. Given the LDD hypothesis, the archipelagic population may have diverged from this population around the late Early Pleistocene. The northern clade was found to be most distant from the southern clade and genetically close to the archipelagic clade (Figure 2), possibly suggesting divergence of the northern clade from the archipelagic population. Therefore, our interpretation of the results is that divergence between the archipelago and the continent occurred at least twice. Such an interpretation can be explained by the following hypothetical scenario that involves different divergence processes across different geographic barriers.

The ECS has remained open during a glacial period throughout the species' history. Meanwhile, the ECS shrank and became less effective as a biogeographic barrier during glacial periods because the lower sea level led to the emergence of the continental shelf in the late Early to Middle Pleistocene (Matsuzaki et al., 2019; Ota, 1998). Paleopalynological records are in line with our SDM analysis; the emergent continental shelf of the ECS was covered by grassland with sparse forest (Xu, Lu, Wu & Liu, 2010), which provides suitable breeding habitat for the Brown Shrike (Lefranc & Worfolk, 1997). Present vagrancy records of many migratory species in the Japanese archipelago during spring are probably due to overshooting (birds traveling too far due to favourable winds and anticyclonic conditions; Newton, 2008) from the CRECS (Senzaki et al., 2019). Increased proximity of the two landmasses, therefore, may have resulted in sufficient propagule numbers (accidental migrants), which arrived in the Japanese archipelago from the CRECS (Figure 5a; Gillespie et al., 2012; MacArthur & Wilson, 1967). During a subsequent interglacial period, the breadth of the ECS expanded to almost 1,000 km (Figure 5b). This could have facilitated isolation of the archipelagic shrike population from that in the CRECS since vagrancy to islands becomes increasingly rare as the distance from the continental source area to the island increases (Lees & Gilroy, 2014). Thus, changes in the size of a geographic barrier may have been responsible for LDD-driven divergence between these regions.

In contrast, the following description supports a hypothesis for the northern clade diverging from the archipelagic clade by vicariance. During the LGM, the present Korean peninsula and current northern China, with the Japanese archipelago connected to the continent at the time, were predicted to be highly suitable for *L. c. cristatus* (Figure S1.3). The Japanese archipelago was predicted to be more suitable than either the east or southern continental coast (current China). This implies that part of the archipelagic population colonized the Korean Peninsula and current northern China. Moreover, a major COI haplotype of the northern clade (H1) was shared among samples from a wide range of localities in Russia, and a star-like pattern of this clade was found (Figure 2a). The structure of this haplotype network suggests a small population during the glacial period that expanded its range, probably northwards, during the subsequent interglacial period (Avise et al., 1987). Therefore, we suggest that a population may have colonised the Korean Peninsula from the Japanese archipelago over a land bridge during a glacial period, and subsequently expanded its range during an interglacial period (Figure 5c), as the reverse colonisation hypothesis predicts (Hikida 2003; Nishiumi & Kim, 2015). Disappearance of the land bridge and northward expansion of the archipelagic and northern clades could have resulted in the present vicariant distribution (Figure 5d; Aoki et al., 2018; Saitoh et al., 2010). The genetically intermediate individuals between the most distant clade pair, shown in our multi-locus network, possibly reflect their secondary contact with presumed introgression on the continent (Lefranc & Worfolk, 1997). Although their past population dynamics and gene flow needs to be directly estimated to refine our hypothetical scenario, our integrative approach provides, at least, a new scenario for

this region (Motokawa, 2017) and a better phylogeographic interpretation of the Brown Shrike.

Justification for the use of a migratory route as a proxy for the past colonisation pathway

Our data on the migratory route of the Brown Shrike across the ECS is far too limited for population-level inference, thus the universality of the route for the entire *L. c. superciliosus* remains unclear. Careful interpretation of our tracking data is required because individual variation in migratory routes is recognized (Stanley, MacPherson, Fraser, McKinnon, & Stutchbury, 2012). Migratory routes inferred from the analysis also included large uncertainties (Figure 4) because of the methodological limitations of light-level geolocators (Lisovski et al., 2019). However, Lefranc & Worfolk (1997) stated that, based on direct observation during migration, many *L. c. superciliosus* fly at least 700 km across the sea from Japan to the coast of China’s Jiangsu Province or the Zhoushan Islands. Our geocator data concur with their statement. Although further deployment of geolocators is needed, their statement guarantees that the focal sea crossing route is at least one of the major routes for the archipelagic population.

We have taken the recent view that idiosyncratic routes of obligate long-distance migratory species are highly persistent traits throughout their evolutionary histories (Winger et al., 2019) in order to interpret part of our tracking data as the route of past colonisation. However, the dominant view is to consider migratory routes as highly labile traits that are easily subjected to natural selection (Pulido, 2007; Winger et al., 2019). Thus, one could claim that the observed migratory route across the ECS is a result of optimization to the present environmental conditions and does not reflect past colonisation (Alerstam, 2011; Berthold, Helbig, Mohr, & Querner, 1992; Sutherland, 1998; Zink & Gardner, 2017).

For the following two reasons, we believe Winger et al.’s (2019) and our interpretation to be more plausible than the counterview. First, our system does not include any region that experienced severe glaciation. Previous studies have used species that inhabit regions previously covered by ice sheets during the LGM (Haché et al., 2017; Milá, Smith, & Wayne, 2006; Ruegg et al., 2006; Sokolovskis et al., 2018). Reverting from being migratory to becoming sedentary residents was probably a major response to glaciation, which was inferred by SDM analyses (Zink & Gardner, 2017). This implied that migration needed to be labile throughout a species’ evolutionary history and migratory routes should trace the post-LGM expansion to the present breeding range. By contrast, we have shown that suitable breeding habitat for the archipelagic population remained in the Japanese archipelago even during the LGM (Figure 3b-d). This implies that the focal migratory route traces a distribution change before the LGM, supporting our view that idiosyncratic routes over a large barrier have been conserved. Second, similarity between autumn and spring migration across the ECS (Figure 4) may support the conservatism of the focal route. If a migratory route is optimized with the present environment, different selection pressures upon passages in different seasons should independently shape the route (Stanley et al., 2012; Tøttrup et al., 2012). Generally, loop migration, in which spring and autumn migration routes are divergent, is thought to have evolved by retaining one ancestral route retracing colonisation while deriving the other to adjust to current environmental conditions (Newton, 2008). Wind is an important determinant (Alerstam 2001). This loop migration phenomenon has been shown for a population of the Oriental Honey Buzzard *Pernis ptilorhynchus* breeding in Japan. Its autumn migration involves crossing the ECS whereas its spring migration route detours around the Korean Peninsula to cross a narrow sea channel to enter the Japanese archipelago (Nourani, Yamaguchi, Manda, & Higuchi, 2016). It has been shown that a route over the Tsushima Strait was specifically suitable for spring wind conditions, when wind directions and strengths over the ECS were unstable (Nourani et al., 2016; Yamaguchi, Arisawa, Shimada, & Higuchi, 2012). These conditions would be unfavourable also for passerines such as shrikes (Alerstam 2001; Tøttrup, Pederson, Onrubia, Klassen, & Thorup, 2017).

Possible LDD-driven divergence process of migratory species

Previous studies have illustrated a divergence scenario for migratory species that assumed vicariance in generating isolated populations (Hewitt, 2004; Weir & Schluter, 2004). While this scenario has provided suitable explanations for many cases, some divergence events may not adequately be explained by it. For example, the traditional vicariance scenario suggests that the 100 km wide strait between the Korean Peninsula and

the Japanese archipelago caused deep genetic divergence of archipelagic populations of mammals from the continent (McKay, 2012; Sato, 2017). From an ecological perspective, such a small barrier would ineffectively restrict the movement of a highly volant bird species (e.g. Gotelli & Graves, 1990), and establishment of a new population from a more distant area as a consequence of LDD is likely to be its alternative explanation. However, traditional frameworks based on the continental distribution for most migratory species have probably hindered the opportunities to test it (Warren et al., 2015).

Our integrative approach using a Japanese migratory bird species provided a suitable system within which to discriminate the mode of genetic divergence under a hypothesis testing framework. We successfully provided a hypothetical scenario for how LDD and paleogeography have contributed to divergence of a migratory species. We also showed that discriminating LDD from vicariance in our framework benefited a more comprehensive understanding of diversification mechanisms of migratory species in this region. We believe that consideration of LDD is also important for continental regions where most migratory species have diversified. Attention has recently been paid to LDD and colonisation as an important phenomenon for distribution and speciation of continental migratory species (Rushing, Dudash, Studds, & Marra, 2015; Winkler et al., 2017). Like the ECS in our system, some geographic barriers (such as seas, deserts, and mountain ranges) in continental areas have also changed size and shape during their geological history (Hewitt, 2004; Sun et al., 2018). Some of the genetic divergence in migratory species over these barriers should be revisited taking the LDD hypothesis into account (e.g. Reeves, Drovetski, & Fadeev, 2008). Further verification of the usability of our integrative approach among several species will facilitate our understanding of the relative importance of LDD to vicariance in the speciation of migratory birds and other animals.

ix. Figures embedded in the text:

Figure 1 Predicted colonisation route (black arrows) and expected results from integrative approaches in (a) long distance dispersal (LDD) and (b) vicariance hypotheses in the Brown Shrike. Predicted gene networks (upper left panes), migratory routes (lines with open circles), and suitable glacial regions (coloured circles) under the two hypotheses are shown. Colours correspond to subspecies in Figure 1c. We predicted that *L. c. lucionensis* represents the basal lineage of the Brown Shrike since close relatives of this species are distributed around South-east Asia (Zhang et al., 2007). Because there are various possibilities for the genetic position of *L. c. cristatus* in the LDD hypothesis, we did not present it in Figure 1a. Grey coloured regions represent terrestrial areas during the Early to the Middle Pleistocene, and dotted lines indicate the present coastlines. (c) DNA sampling localities. Circles and triangles represent samples from breeding and non-breeding localities, respectively, and their colours indicate haplogroups inferred in the median-joining haplotype network of COI (Figure 2a). Land colours indicate breeding ranges of presumed subspecies following Worfolk (2000). The striped region indicates a presumed hybrid zone between subspecies.

Figure 2 (a) A median joining haplotype network constructed using 521 bp of the partial COI gene. Each circle indicates a unique haplotype and its size is proportional to the number of samples. Colours of circles indicate regions where samples were collected. One mutational gap was indicated by one bar, and a black node indicates a median vector. Clusters analogous to clades found in the Bayesian inference tree (Figure S1.2) are indicated. (b) The multi-locus network constructed for the Brown Shrike using the four loci under the NeighborNet algorithm. Coloured shades correspond to clades to which their mitochondrial haplotypes belong.

Figure 3 Estimated suitable distribution for *L. c. superciliosus* for different time periods; a constructed model was projected to (a) the present climate, and (b-d) three different climate models for the Last Glacial Maximum (LGM, 20 kya). Suitability for a 2.5 min grid increases from yellow, green to indigo, whereas the grey area was determined as absence. The black area is that of strict extrapolation indicated by multivariate environmental similarity surface analysis. Black lines indicate present coastlines. Note that the terrestrial landmass expanded during the LGM due to the lowered sea level.

Figure 4 Migratory routes of three individual *L. c. superciliosus* around the East China Sea. Mean posterior probabilities of location estimates of the three birds for each raster cell are indicated by colour gradation from

yellow (high) to indigo (low). Dots with line segments indicate median tracks of their inferred migratory routes (white and red dots indicate migrating and stationary sites, respectively). Autumn and spring migrations are indicated by solid and hatched lines, respectively.

Figure 5 Schematic presentation of the hypothetical phylogeographic scenario of divergence in the Brown Shrike. (a) The ancestral population colonised the Japanese archipelago via LDD across the East China Sea (ECS) in a glacial period around the late Early Pleistocene. (b) Increased isolation of the archipelagic population through expansion of the ECS facilitated its divergence. Its migratory route was shaped to retrace the colonisation route. (c) In a later glacial period, part of the archipelagic population colonised the Korean Peninsula over a land bridge. (d) Northward expansions caused vicariance of both the archipelagic and northern clades and resulted in their present geographical distributions. The hatched lines in each pane indicate the predicted coastlines in the previous inter/glacial periods, following Ota (1998). Coloured circles indicate ancestral populations of each lineage. Lines with open circles indicate the migration route and arrows indicate movements of populations.

Figure 1

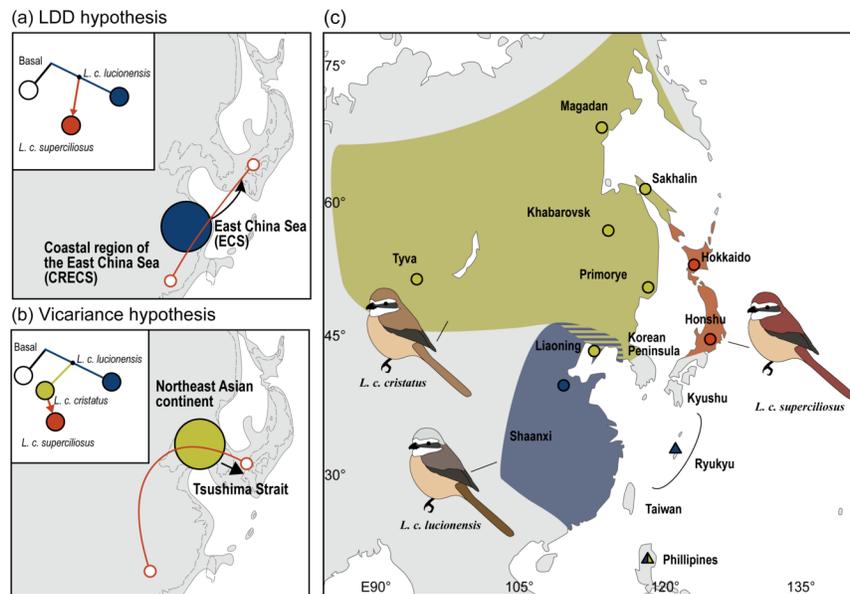


Figure 2

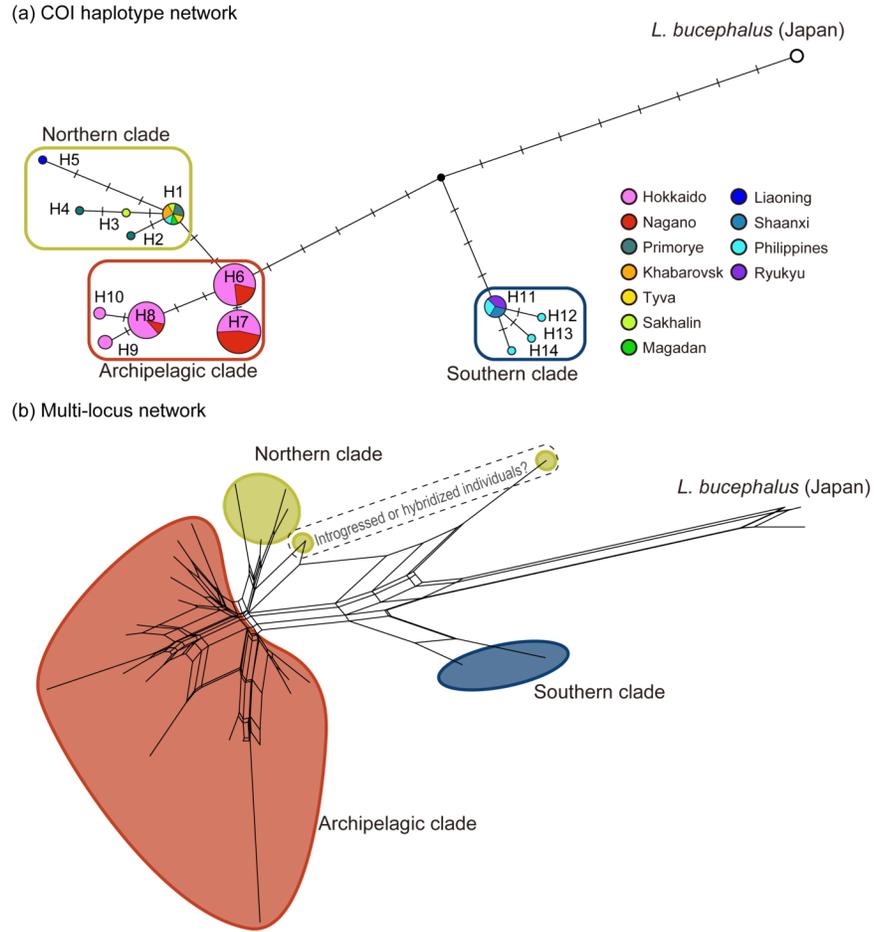


Figure 3

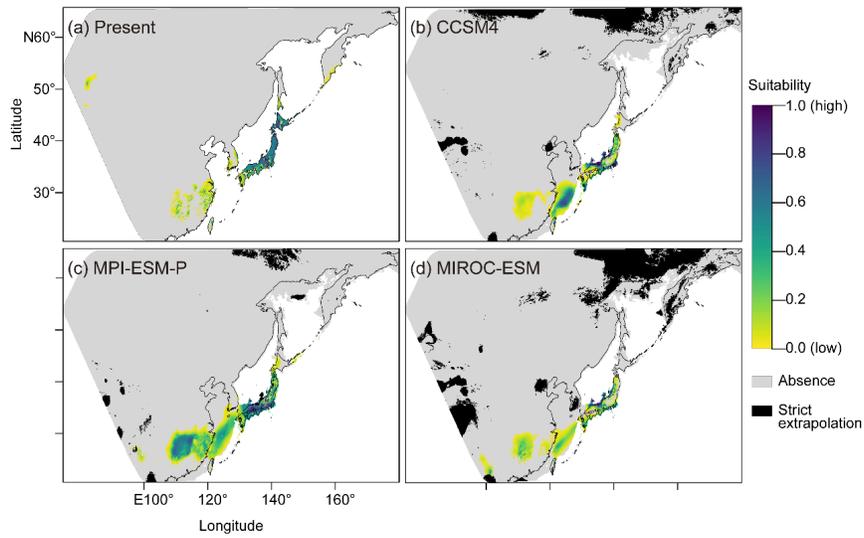


Figure 4

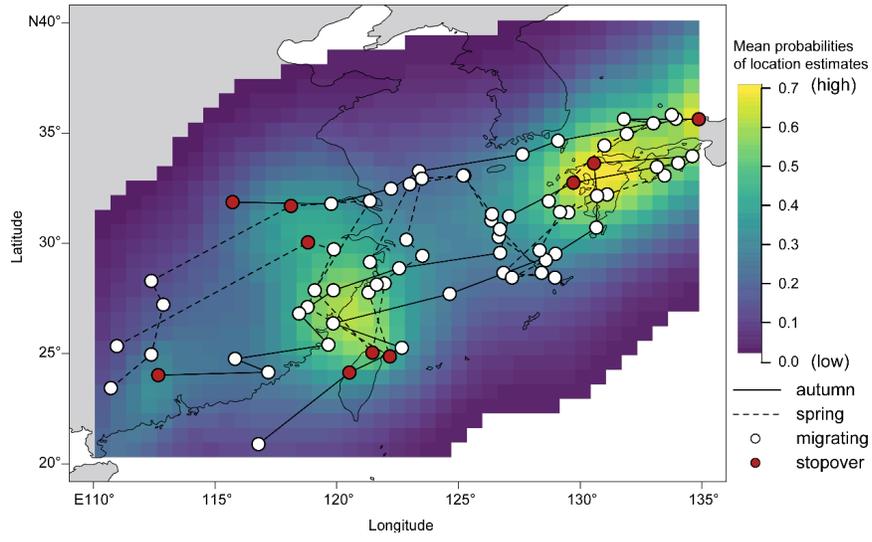
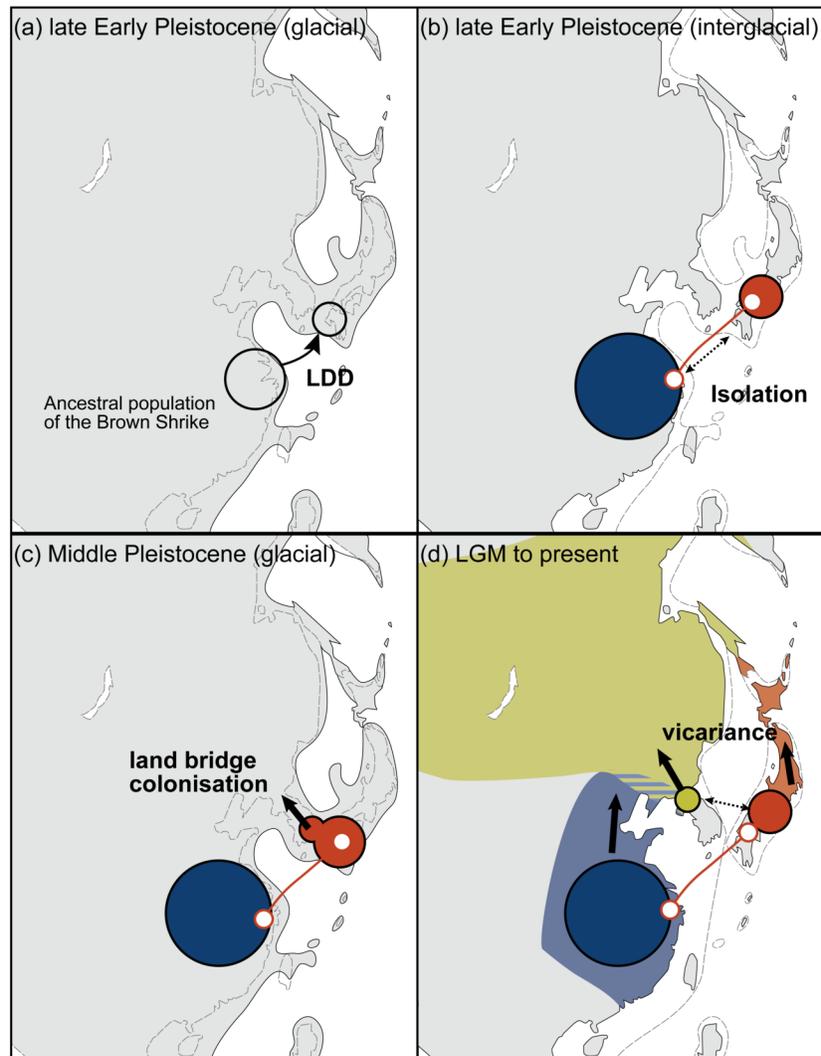


Figure 5



x. Data availability statement:

Genetic data is available at DDBJ (accession numbers are listed in Table S1.1). R scripts and the raw data for migratory route analyses and species distribution modelling are available at data repository (submitted to DRYAD with DOI <https://doi.org/10.5061/dryad.cfxpnvx2j>).

xi. References:

- Albert, J. S., Schoolmaster, D. R., Tagliacollo, V., & Duke-Sylvester, S. M. (2017). Barrier displacement on a neutral landscape: Toward a theory of continental biogeography. *Systematic Biology*, 66(2), 167-182.
- Alerstam, T. (2001). Detours in bird migration. *Journal of Theoretical Biology*, 209(3), 319-331.
- Alerstam, T. (2011). Optimal bird migration revisited. *Journal of Ornithology*, 152, S5-S23.
- Alvarado, A. H., Fuller, T. L., & Smith, T. B. (2014). Integrative tracking methods elucidate the evolutionary dynamics of a migratory divide. *Ecology and Evolution*, 4(17), 3456-3469.

- 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.
- Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., Reeb, C. A., & Saunders, N. C. (1987). Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology, Evolution, and Systematics*, 18, 489–552.
- Bandelt, H. J., Forster, P., & Rohl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, 16(1), 37–48.
- Berthold, P., Helbig, A. J., Mohr, G., & Querner, U. (1992). Rapid microevolution of migratory behaviour in a wild bird species. *Nature*, 360, 668–670.
- Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., ... Drummond, A. J. (2014). BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, 10(4), 1–6.
- Bridge, E. S., Kelly, J. F., Contina, A., Gabrielson, R. M., MacCurdy, R. B., & Winkler, D. W. (2013). Advances in tracking small migratory birds: A technical review of light-level geolocation. *Journal of Field Ornithology*, 84(2), 121–137.
- Bruderer, B., & Boldt, A. (2001). Flight characteristics of birds: I. radar measurements of speeds. *Ibis*, 143, 178–204.
- Dong, L., Wei, M., Alström, P., Huang, X., Olsson, U., Shigeta, Y., ... Zheng, G. (2015). Taxonomy of the Narcissus Flycatcher *Ficedula narcissina* complex: An integrative approach using morphological, bioacoustic and multilocus DNA data. *Ibis*, 157(2), 312–325.
- Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E., & Yates, C. J. (2011). A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, 17(1), 43–57.
- Filatov, D. A. (2009). Processing and population genetic analysis of multigenic datasets with ProSeq3 software. *Bioinformatics*, 25(23), 3189–3190.
- Gallagher, S. J., Kitamura, A., Iryu, Y., Itaki, T., Koizumi, I., & Hoiles, P. W. (2015). The Pliocene to recent history of the Kuroshio and Tsushima Currents: a multi-proxy approach. *Progress in Earth and Planetary Science*, 2(1), 1–23.
- Gent, P. R., Danabasoglu, G., Donner, L. J., Holland, M. M., Hunke, E. C., Jayne, S. R., ... Zhang, M. (2011). The community climate system model version 4. *Journal of Climate*, 24(19), 4973–4991.
- Gillespie, R. G., Baldwin, B. G., Waters, J. M., Fraser, C. I., Nikula, R., Roderick, G. K. (2012). Long-distance dispersal: a framework for hypothesis testing. *Trends in Ecology and Evolution*, 27(1), 47–56.
- Gotelli, N. J., & Graves, G. R. (1990). Body size and the occurrence of avian species on land-bridge islands. *Journal of Biogeography*, 17(3), 315–325.
- Greenberg, R., & Marra, P. P. (2005). The renaissance of migratory bird biology: A synthesis. In R. Greenberg & P. P. Marra (Eds.), *Birds of two worlds: the ecology and evolution of migration* (pp. 437–443). Maryland: The Johns Hopkins University Press.
- Haché, S., Villard, E. M. B. M., Proctor, H., Hallworth, M. T., Davis, C. S., Stralberg, D., ... Alexandra, E. C. (2017). Phylogeography of a migratory songbird across its Canadian breeding range: Implications for conservation units. *Ecology and Evolution*, 7, 6078–6088.
- Harrison, S. P., Yu, G., Takahara, H., & Prentice, I. C. (2001). Palaeovegetation. Diversity of temperate plants in east Asia. *Nature*, 413(6852), 129–130.
- 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*, 50:75–95.

- Hasumi, H., & Emori, S. (2004). K-1 coupled model (MIROC) description. K-1 Technical Report, 1–34.
- Hewitt, G. M. (2004). Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 359(1442), 183–195.
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25(15), 1965–1978.
- Hikida, T. (2003). Biogeography of reptiles in islands in East Asia – Geographic distribution from views of molecular and morphological studies. *Seibutsu Kagaku*, 54, 205–220. (in Japanese)
- Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, 23, 254–267.
- Joly, S., & Bruneau A. (2006) Incorporating allelic variation for reconstructing the evolutionary history of organisms from multiple genes: an example from *Rosa* in North America. *Systematic Biology*, 55(4), 623–636.
- Koike, S., Hijikata, N., & Higuchi, H. (2016). Migration and wintering of Chestnut-cheeked Starlings *Agropsar philippensis*. *Ornithological Science*, 15, 63–74.
- Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, 33(7), 1870–1874.
- Lees, A. C., & Gilroy, J. J. (2014). Vagrancy fails to predict colonisation of oceanic islands. *Global Ecology and Biogeography*, 23(4), 405–413.
- Lefranc, N., & Worfolk, T. (1997). *Shrikes: a guide to the shrikes of the world*. East Sussex: Pica Press.
- Liepert, B. G., & Lo, F. (2013). CMIP5 update of “Inter-model variability and biases of the global water cycle in CMIP3 coupled climate models.” *Environmental Research Letters*, 8(2), 8–11.
- Lisovski, S., & Hahn, S. (2012). GeoLight – processing and analysing light-based geolocator data in R. *Methods in Ecology and Evolution*, 3, 1055–1059.
- Lisovski, S., Bauer, S., Briedis, M., Davidson, S. C., Dhanjal-Adams, K. L., Hallworth, M. T., ... Bridge, E. S. (2019) Light-level geolocator analyses: A user’s guide. *Journal of Animal Ecology*, 89(1), 221–236.
- MacArthur, R. H., & Wilson, E. O. (1967). *The theory of island biogeography*. New Jersey: Princeton University Press.
- Mardulyn, P. (2012). Trees and/or networks to display intraspecific DNA sequence variation? *Molecular Ecology*, 21, 3385–3390.
- Matsuzaki, K. M., Itaki, T., & Tada, R. (2019). Paleooceanographic changes in the Northern East China Sea during the last 400 kyr as inferred from radiolarian assemblages (IODP Site U1429). *Progress in Earth and Planetary Science*, 6(1).
- McKay, B. D. (2012). A new timeframe for the diversification of Japan’s mammals. *Journal of Biogeography*, 39(6), 1134–1143.
- Milá, B., Smith, T. B., & Wayne, R. K. (2006). Postglacial population expansion drives the evolution of long-distance migration in a songbird. *Evolution*, 60(11), 2403–2409.
- Motokawa, M. (2017). “Land emergence” and “elevation shift” affect diversification: A new perspective toward understanding the high species diversity of terrestrial animals in Japan. In M. Motokawa & H. Kajihara (Eds.), *Species Diversity of Animals in Japan* (pp. 3–23).
- Muscarella, R., Galante, P. J., Soley-Guardia, M., Boria, R. A., Kass, J. M., Uriarte, M., & Anderson, R. P. (2014). ENMeval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for Maxent ecological niche models. *Methods in Ecology and Evolution*, 5(11), 1198–1205.

- Nabholz, B., Lanfear, R., & Fuchs, J. (2016). Body mass-corrected molecular rate for bird mitochondrial DNA. *Molecular Ecology*, 25(18), 4438–4449.
- Newton, I. (2008). *The migration ecology of birds* (1st edition). London: Elsevier.
- Nishiumi, I., & Kim, C. H. (2015). Assessing the potential for reverse colonization among Japanese birds by mining DNA barcode data. *Journal of Ornithology*, 156(Suppl 1), S325–S331.
- Nourani, E., Yamaguchi, N. M., Manda, A., & Higuchi, H. (2016). Wind conditions facilitate the seasonal water-crossing behaviour of Oriental Honey-buzzards *Pernis ptilorhynchus* over the East China Sea. *Ibis*, 168, 506–518.
- O'Connor, R. J. (1986). Biological characteristics of invaders among bird species in Britain. *Philosophical Transactions of the Royal Society of London. B*, 314, 583–598.
- Ohshima, K. (1990). The History of straits around the Japanese in the Late-Quaternary. *The Quaternary Research*, 29(3), 193–208.
- Ota, H. (1998). Geographic patterns of endemism and speciation in amphibians and reptiles of the Ryukyu archipelago, Japan, with special reference to their paleogeographical implications. *Research on Population Ecology*, 40(2), 189–204.
- Päckert, M., Martens, J., Sun, Y. -H., Severinghaus, L. L., Nazarenko, A. A., Ting, J., Topfer, T. and Tietze, D. T. (2012). Horizontal and elevational phylogeographic patterns of Himalayan and Southeast Asian forest passerines (Aves: Passeriformes). *Journal of Biogeography*, 39, 556–573.
- Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190, 231–259.
- Pulido, F. (2007). The Genetics and Evolution of Avian Migration. *Bio Science*, 57(2), 165–174.
- Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, 67(5), 901–904.
- Rappole, J. H., & Tipton, A. R. (1991). New harness design for attachment of radio transmitters to small passerines. *Journal of Field Ornithology*, 62(3), 335–337.
- Ratnasingham S, Hebert PDN (2007) The barcode of life data system. *Molecular Ecology Notes* 7:355–364.
- Reeves, A. B., Drovetski, S. V., & Fadeev, I. V. (2008). Mitochondrial DNA data imply a stepping-stone colonisation of Beringia by arctic warbler *Phylloscopus borealis*. *Journal of Avian Biology*, 39, 567–575.
- Rose, M. D., & Polis, G. A. (2000). On the insularity of islands. *Ecography*, 23, 693–701.
- Ruegg, K. C., Hijmans, R. J., & Moritz, C. (2006). Climate change and the origin of migratory pathways in the Swainson's thrush, *Catharus ustulatus*. *Journal of Biogeography*, 33(7), 1172–1182.
- Rushing, C. S., Dudash, M. R., Studds, C. E., & Marra, P. P. (2015). Annual variation in long-distance dispersal driven by breeding and non-breeding season climatic conditions in a migratory bird. *Ecography*, 38(10), 1006–1014
- Saitoh, T., Alstrom, 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, 35.
- Saitoh, T., Sugita, N., Someya, S., Iwami, Y., Kobayashi, S., Kamigaichi, H., ... Nishiumi, I. (2015). DNA barcoding reveals 24 distinct lineages as cryptic bird species candidates in and around the Japanese Archipelago. *Molecular Ecology Resources*, 15(1), 177–186.
- Sato, J. J. (2017). A review of the process of mammalian faunal assembly in Japan: Insights from molecular phylogenetics. In M. Motokawa & H. Kajihara (Eds.), *Species Diversity of Animals in Japan*.

- Senzaki, M., Umegaki, Y., Odaya, Y., Senzaki, H., Takagi, S., Nishizawa, B., & Hara, S. (2019). Migratory birds of Japan: a field guide. Tokyo: Bun-ichi Sogo Shuppan (in Japanese).
- Severinghaus, L. (1996). Territory strategy of the migratory brown shrike *Lanius cristatus*. *Ibis*, 138(3), 460–465.
- Shiu, H. J., Tokita, K., Morishita, E., Hiraoka, E., Wu, Y., Nakamura, H., Higuchi, H. (2006). Route and site fidelity of two migratory raptors: Grey-faced Buzzards *Butastur indicus* and Honey-buzzards *Pernis apivorus*. *Ornithological Science*, 5, 151–156.
- Sokolovskis, K., Bianco, G., Willemoes, M., Solovyeva, D., Bensch, S., & Akesson, S. (2018). Ten grams and 13,000km on the wing - route choice in willow warblers *Phylloscopus trochilus yakutensis* migrating from Far East Russia to East Africa. *Movement Ecology*, 6(1), 1–10.
- Somveille, M., Rodrigues, A. S. L., & Manica, A. (2018). Energy efficiency drives the global seasonal distribution of birds. *Nature Ecology and Evolution*, 2(6), 962–969.
- Stanley, C. Q., MacPherson, M., Fraser, K. C., McKinnon, E. A., & Stutchbury, B. J. M. (2012). Repeat tracking of individual songbirds reveals consistent migration timing but flexibility in route. *PLoS ONE*, 7(7), 5–10.
- Stephens, M., Smith, N., & Donnelly, P. (2001). A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics*, 68, 978–989.
- Stephens, M., & Donnelly, P. (2003). A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *American Journal of Human Genetics*, 73, 1162–1169.
- Sun, M., Zhang, X., Tian, M., Liu, R., He, Z., Qi, L., & Qiao, Y. (2018). Loess deposits since early Pleistocene in northeast China and implications for desert evolution in east China. *Journal of Asian Earth Sciences*, 155, 164–173.
- Sutherland, W. J. (1998). Evidence for flexibility and constraint in migration systems. *Journal of Avian Biology*, 29(4), 441–446.
- Takagi, M. (2003). Philopatry and habitat selection in Bull-headed and Brown shrikes. *Journal of Field Ornithology*, 74(1), 45–52.
- Tottrup, A. P., Klaassen, R. H. G., Strandberg, R., Thorup, K., Kristensen, M. W., Jorgensen, P. S., ... Alerstam, T. (2012). The annual cycle of a trans-equatorial Eurasian – African passerine migrant: Different spatio-temporal strategies for autumn and spring migration. *Proceedings of the Royal Society B: Biological Sciences*, 279, 1008–1016.
- Tottrup, A. P., Pederson, L., Onrubia, A., Klassen, R. H. G., & Thorup, K. (2017). Migration of red-backed shrikes from the Iberian Peninsula: optimal or sub-optimal detour? *Journal of Avian Biology*, 48, 149–154.
- Warren, B. H., Simberloff, D., Ricklefs, R. E., Aguilée, R., Condamine, F. L., Gravel, D., ... Thebaud, C. (2015). Islands as model systems in ecology and evolution: Prospects fifty years after MacArthur-Wilson. *Ecology Letters*, 18(2), 200–217.
- Weeks, B. C., & Claramunt, S. (2014) Dispersal has inhibited avian diversification in Australasian archipelagos. *Proceedings of the Royal Society B*, 281, 20141257.
- Weir, J. T., & Schluter, D. (2004). Ice sheets promote speciation in boreal birds. *Proceedings of the Royal Society of London. B*, 271, 1881–1887.
- Weir, J. T., & Schluter, D. (2008). Calibrating the avian molecular clock. *Molecular Ecology*, 17(10), 2321–2328.
- Winger, B. M., Auteri, G. G., Pegan, T. M., & Weeks, B. C. (2019). A long winter for the Red Queen: rethinking the evolution of seasonal migration. *Biological Reviews*, 94(3), 737–752.

Winkler, D. W., Gandoy, F. A., Areta, J. I., Iloff, M. J., Rakhimberdiev, E., Kardynal, K. J., & Hobson, K. A. (2017). Long-distance range expansion and rapid adjustment of migration in a newly established population of Barn Swallows breeding in Argentina. *Current Biology*, 27(7), 1080–1084.

Worfolk, T. (2000). Identification of red-backed, isabelline and brown shrikes. *Dutch Birding*, 22, 323–362.

Wotherspoon, S., Sumner, M., & Lisovski, S. (2015). SGAT: Solar/satellite geolocation for animal tracking. Retrieved from <https://github.com/SWotherspoon/SGAT>

Xu, D., Lu, H., Wu, N., & Liu Z. (2010). 30 000-Year vegetation and climate change around the East China Sea shelf inferred from a high-resolution pollen record. *Quaternary International*, 227, 53–60.

Yamaguchi, N. M., Arisawa, Y., Shimada, Y., & Higuchi, H. (2011). Real-time weather analysis reveals the adaptability of direct sea-crossing by raptors. *Journal of Ethology*, 30, 1–10.

Zhang, W., Lei, F. M., Liang, G., Yin, Z. H. Zhao, H. F. Wang, H. J., & Krištín, A. (2007). Taxonomic status of eight Asian shrike species (*Lanius*): Phylogenetic analysis based on Cyt *b* and CoI gene sequences. 42(2), 173–180.

Zhao, M., Alström, P., Hu, R., Zhao, C., Hao, Y., Lei, F., & Qu, Y. (2017). Phylogenetic relationships, song and distribution of the endangered Rufous-headed *Robin Larvivora ruficeps*. *Ibis*, 159(1), 204–216.

Zink, R. M., & Gardner, A. S. (2017). Glaciation as a migratory switch. *Science Advances*, 3(9), 1–9.

Zink, R. M., Pavlova, A., Rohwer, S., & Drovetski, S. V. (2006). Barn swallows before barns: population histories and intercontinental colonisation. *Proceedings of Royal Society B*, 273, 1245–1251.

xii. Biosketch:

Daisuke Aoki is interested in the relationship between distribution and speciation of taxa with higher dispersal abilities, especially avian species. This work is based on his thesis for a master's course at Hokkaido University. He and other authors collaborate on evolutionary and conservation ecology of the genus *Lanius*.

Author contributions: DA and MT conceived the ideas; DA, HS, MK and MT conducted the fieldwork and APK collected samples in Russia; DA and APK conducted laboratory procedures; DA analysed the data; DA led the writing with assistance from MK, MT and APK.