






## RESEARCH PAPER

Journal of  
Biogeography

WILEY

# Great journey of Great Tits (*Parus major* group): Origin, diversification and historical demographics of a broadly distributed bird lineage

Gang Song<sup>1</sup>  | Ruiying Zhang<sup>1</sup> | Fernando Machado-Stredel<sup>2</sup> | Per Alström<sup>1,3</sup>  |  
Ulf S. Johansson<sup>4</sup>  | Martin Irestedt<sup>5</sup> | Herman L. Mays Jr.<sup>6</sup> | Bailey D. McKay<sup>7</sup> |  
Isao Nishiumi<sup>8</sup> | Yalin Cheng<sup>1,9</sup> | Yanhua Qu<sup>1</sup> | Per G. P. Ericson<sup>5</sup>  | Jon Fjeldså<sup>10</sup> |  
Andrew Townsend Peterson<sup>2</sup> | Fumin Lei<sup>1,9,11</sup> 

<sup>1</sup>Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing, China

<sup>2</sup>Biodiversity Institute and Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, KS, USA

<sup>3</sup>Animal Ecology, Department of Ecology and Genetics, Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden

<sup>4</sup>Department of Zoology, Swedish Museum of Natural History, Stockholm, Sweden

<sup>5</sup>Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Stockholm, Sweden

<sup>6</sup>Department of Biological Sciences, Marshall University, Huntington, WV, USA

<sup>7</sup>Department of Ornithology, American Museum of Natural History, New York City, NY, USA

<sup>8</sup>Department of Zoology, National Museum of Nature and Science, Ibaraki, Japan

<sup>9</sup>University of Chinese Academy of Sciences, College of Life Sciences, Beijing, China

<sup>10</sup>Center for Macroecology, Evolution and Climate at the Natural History Museum of Denmark, University of Copenhagen, Copenhagen, Denmark

<sup>11</sup>Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming, China

## Correspondence

Fumin Lei, Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing, China.  
Email: leifm@ioz.ac.cn

## Funding information

National Science and Technology Major Project, Grant/Award Number: 2018ZX10101004; National Natural Science Foundation of China, Grant/Award Number: 31471991, 31572249, 31572291 and 31630069; Vetenskapsrådet, Grant/Award Number: 2015-04402; National Science Foundation (NSF), Grant/Award Number: DBI-0821703

Handling Editor: Aristeidis Parmakelis

## Abstract

**Aim:** The Pleistocene glacial cycles play a prominent role in shaping phylogeographical patterns of organisms, while few studies have focused on the regional difference of glacial effects. By acquiring comprehensive knowledge of the origin, diversification and historical demography of an intensively studied passerine species complex, Great Tit, we aim to test the regional variation of the Late Pleistocene glaciation impacts on this widely distributed bird lineage.

**Location:** Eurasia and associated peninsulas and archipelagos.

**Taxa:** *Parus major* species complex.

**Methods:** Phylogeny, divergence times and demographic dynamics were estimated with Bayesian methods. Population structure, genetic diversity and correlation between genetic and physical distances were estimated based on mtDNA variation. Glacial-to-present distributional changes were assessed via ecological niche modelling (ENM).

**Results:** Five major clades (Central Asia, Eastern Asia, Eastern Himalaya, Northern and Western Eurasia and Southern Asia) were detected, with divergence times ranging 1.57–0.50 million years ago. Genetic diversity values and Bayesian skyline plots suggest that the three eastern clades had a deeper population history. A more

complex geographic structure was observed in East Asia. Demographic expansion during the last glacial cycle was indicated for all five clades. ENM results showed broad conservatism of traits related to climate tolerances, and generally broader and more continuous distributional patterns under glacial conditions.

**Main Conclusions:** The Great Tit complex probably originated in Southeast Asia. Geographic barriers, such as the deserts of Central Asia and the Qinghai-Tibet Plateau appear to be related to the lineage divergence. Late Pleistocene climate cycles influenced both demographic dynamics and divergence, especially in terms of east-west differences in relation to geographic complexity.

#### KEYWORDS

Central Asia, East Asia, Himalaya, historical demography, *Parus*, Qinghai-Tibet Plateau

## 1 | INTRODUCTION

Contemporary biotas on Earth show influences of past climatic changes, particularly from glacial cycles over the past 2.5 million years (Hewitt, 2000; Lovette, 2005). Climate oscillations shape genetic diversity and provide isolation necessary for speciation (Hosner, Liu, Peterson, & Moyle, 2015; Klicka, Campillo, Manthey, & Qu, 2015; Lim, Zou, & Sheldon, 2015; Quan, Qu, & Lei, 2015; Reddy & Nyári, 2015). Impacts of climatic oscillations are variable, depending on many factors, including latitude, elevation, topography and the particular ecological requirements of species (Burbrink et al., 2016; Lovette, 2005; Qiao et al., 2016; Stewart, Lister, Barnes, & Dalén, 2010). Previous studies have indicated that populations of numerous species in Western Europe retreated southward dramatically during glacial periods (Pavlova, Rohwer, Drovetski, & Zink, 2006; Taberlet, Fumagalli, Wust-Saucy, & Cosson, 1998; Wang, Wijk, Braun, & Wink, 2017), whereas East Asian populations apparently showed more modest range shifts (Li et al., 2009; Li, Lin, et al., 2016; Song et al., 2009; Zhao et al., 2012). However, our current understanding of influences of past climatic oscillations on species' history is far from complete, requiring more detailed information for more taxa across broader geographic areas.

Although several studies have described phylogeographical patterns in widely distributed species (Dalén et al., 2005; Hung, Drovetski, & Zink, 2012, 2013a, 2013b; Pavlova, Zink, & Rohwer, 2005; Pavlova et al., 2005), few have assessed regional differences in climate impacts (but see Saitoh et al., 2010). Palaeovegetation studies show significant differences in vegetation dynamics among different areas of Eurasia: a great ice sheet covered much of north-western Europe whereas East Asia was characterized by temperate forest and steppe grassland (Ashastina et al., 2018; Janská et al., 2017; Yu et al., 2000). However, a recent phylogeographical analysis of magpies (*Pica* spp.), distributed across the Northern Hemisphere in open, dry habitats, revealed synchronous dynamic trends between western and eastern Eurasian lineages. East Asia presented a more complex phylogenetic history, but a less deeply differentiated population structure than in northern

Eurasia (Song et al., 2018), counter to the customary interpretations about geographic variation of glacial effects.

The Great Tit (*Parus major*, sensu lato, family Paridae) is a small passerine that ranks among the most intensively studied bird species, having been the focus of over 1,000 studies (e.g. Bosse et al., 2017; Cauchard, Doucet, Boogert, Angers, & Doligez, 2017; Slabbekoorn & Peet, 2003; Templeton, Zollinger, & Brumm, 2016). The complex has the broadest geographic distribution of any species in the family Paridae, covering all of Eurasia from the Atlantic to the Pacific, and from northern Fennoscandia to southern Indonesia (Gosler, Clement, & Christie, 2007). In contrast to the magpies, however, Great Tits inhabit diverse forest types (Gosler et al., 2007/2007).

Relationships within this complex have long been debated, with 43 recognized subspecies (Gosler et al., 2007), and contrasting opinions about species limits. Gosler et al. (2007) treated all populations within a single species, whereas Gill and Donsker (2019) treated the complex as comprising three species: Great Tit (*P. major*), Japanese Tit (*P. minor*) and Cinereous Tit (*P. cinereus*). Phylogeographical analyses have detected three to five clades that diverged up to about 2 million years ago (Ma) (Kvist et al., 2003; Päckert et al., 2005; Zhao et al., 2012), high genetic diversity in a clade in south-west China, and a population expansion time antedating the Last Glacial Maximum (LGM). Molecular and song data (Päckert et al., 2005) refuted the idea that this complex represents a 'ring species' (Mayr, 1942). However, the genetic structure and historical dynamics of the group are not fully understood, as previous studies suffered from insufficient taxon sampling from key areas across the massive geographic distribution of the complex.

Here, with an expanded sampling of populations, we re-analysed lineage diversification, geographic differentiation and historical demography of the Great Tit complex. This study focuses on three questions: (1) what factors drove lineage divergence in the Great Tit complex, (2) do regional differences in genetic structure among Great Tit populations relate to topographic complexity, and (3) do regional differences in Great Tit genetic structure relate to the severity of past climatic changes in different regions?

## 2 | MATERIALS AND METHODS

### 2.1 | Samples collection and laboratory work

We assembled 123 fresh tissue or blood samples, 80 toepads, and 5 feather samples for sequencing. We examined these samples in combination with existing sequences of 132 individuals from Zhao et al. (2012), for a total of 340 samples from 67 geographic populations, covering most of the distribution of the complex (Figure 1, Table S1). We included single samples of *Periparus ater*, *Cyanistes cyanus* and *Parus monticolus* as outgroups.

Total genomic DNA was extracted using the QIAamp DNA Mini Kit (QIAGEN) following manufacturer's protocols. Feather and toepad extractions were conducted in the DNA Laboratory, Department of Bioinformatics and Genetics, Swedish Museum of Natural History. The same kit was used, but with slight modifications as follows: During the initial lysis step in 200 µl AL, 20 µl of 1 M DTT was added, followed with an incubation step at 72°C for 10 min. Subsequent extraction steps followed manufacturer's protocols.

We obtained partial sequences of the mitochondrial cytochrome b (Cytb) and NADH dehydrogenase subunit 2 (ND2) genes, as well as the nuclear  $\beta$ -fibrinogen intron 5 (Fib5) and transforming growth factor beta 2 intron 5 (TGFB2), via polymerase chain reaction (PCR) amplification. For toepad and feather samples, multiple primer pairs were designed to amplify shorter fragments (180–200 bp) of each segment. All primers and amplifying conditions are summarized in Table S2. PCR products were purified using the QIAquick™ PCR Purification Kit (QIAGEN). Sanger sequencing was carried out on ABI

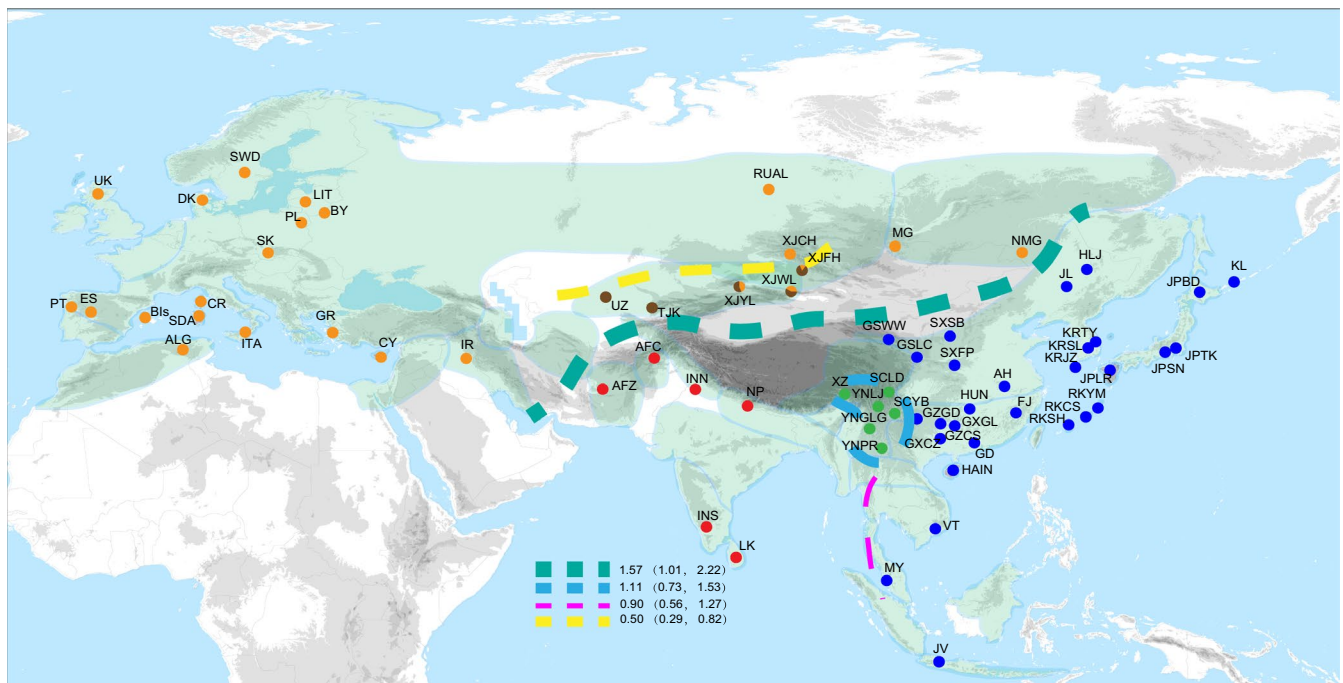
377 automatic sequencer, following the ABI PRISM BigDye Terminator Cycle Sequencing protocol. Sequences were aligned and checked in MEGA5 (Tamura et al., 2011). For blood samples, we ruled out the possibility of Numt amplification based on (a) no stop codons were found in the sequences based on the vertebrate mitochondrial DNA coding table; (b) few or no sites showed ambiguous double peaks in the sequence trace file; (c) no 'strange' sequence with exceptional genetic variation was noted; and (d) the mtDNA phylogeny of the present study was generally consistent with previously published works. All sequences are accessible at GenBank (MT212810–MT213943).

### 2.2 | Genetic diversity

Two parameters of genetic diversity, haplotype diversity ( $H_d$ ) and nucleotide diversity ( $\pi$ ), were calculated in DNASP 5.0 (Librado & Rozas, 2009). To test population expansion, Fu's  $F_s$  and Tajima's  $D$  were also computed in DNASP. For heterozygous nuclear sequences, haplotype phases were reconstructed via the PHASE algorithm (Stephens & Donnelly, 2003; Stephens, Smith, & Donnelly, 2001) in DNASP 5.0. Sequences with uncertain resolution probabilities (i.e. <80%) were deleted from subsequent analyses.

### 2.3 | Phylogeny and population structure

We used MrModelTest2 (Nylander, 2004) to select the most appropriate model of molecular sequence evolution for each gene via



**FIGURE 1** Distributions and sampling sites of the Great Tit species complex. Sampling sites are marked by coloured dots with acronyms referring to locality names (see Table S1). The green shading indicates the probable extent of the range of Great Tits; colours of the dots represent the major clades: blue, Eastern Asia; brown, Central Asia; green, Eastern Himalaya; red, Southern Asia; and orange, Northern and Western Eurasia. Dashed lines in different colours and sizes show the major divergences between clades and associated time estimates



the Akaike Information Criterion (AIC). The maximum likelihood (ML) algorithm implemented in PhyML 3.0 (Guindon & Gascuel, 2003) and the Bayesian inference (BI) algorithm in BEAST 1.8 (Drummond, Suchard, Xie, & Rambaut, 2012) were used to reconstruct phylogenetic relationship. Nonparametric bootstrapping (1,000 replicates) on the ML tree was used to evaluate nodal support among branches. For Bayesian analyses, we applied a constant-growth coalescent tree prior for phylogenetic reconstruction, and ran Markov Chain Monte Carlo (MCMC) simulations with 100 million generations with a relaxed log-normal clock model. Convergence of posterior distributions of parameter estimates was evaluated by monitoring for effective sample sizes (ESS) over 200 and trace plots were made in Tracer 1.6 (Rambaut, Suchard, Xie, & Drummond, 2014). The same BEAUTI file was run three times to ensure the consistency of results.

For nuclear DNA, we applied coalescent-based species tree estimation using \*BEAST in BEAST 2 (Bouckaert et al., 2014). We unlinked the substitution models of the two nuclear introns, and set substitution parameters for each according to MrModeltest2 results. We implemented a piecewise linear and constant root model as the species tree prior, and used default molecular clock settings in major lineage identification. MCMC chains ran for 500 million generations, sampling every 5,000 generations. The first 10% of samples was discarded as 'burn-in', and convergence of the MCMC chains was examined in Tracer 1.6.

We used Structure version 2.3 (Pritchard, Stephens, & Donnelly, 2000) to detect population structure based on information in the frequency of genotypes. The admixture model with correlated allele frequencies was applied, and sampling locality was set as a prior to augment signal in the results (Hubisz, Falush, Stephens, & Pritchard, 2009). Numbers of populations ( $K$ ) were set at 2–5, and each  $K$  value was run 10 times. The most appropriate  $K$  was identified according to the  $\Delta K$  method (Evanno, Regnaut, & Goudet, 2005). CLUMPP (Jakobsson & Rosenberg, 2007) was used to merge results from replicate analyses for each  $K$ , and DISTRUCT (Rosenberg, 2004) was used to output the bar plot for combined results.

With mtDNA, we computed mean distances within populations as a measure of nucleotide diversity (Nei & Kumar, 2000), and net average distance between populations in MEGA 5. Geographic distances between populations were calculated in the Geosphere R package. Isolation by distance was tested in each major lineage using Mantel tests in the ape R package (Paradis, Claude, & Strimmer, 2004), each with 10,000 permutations. We applied Wilcoxon non-parametric tests to evaluate significance of differences in pairwise distances between major geographic clades.

## 2.4 | Estimation of divergence times and historical demography

We estimated divergence times between clades in BEAST 1.8. As no fossil record for Paridae or its relatives is available, the substitute rate adjustment would be a better way for divergence

time estimation (Li et al., 2009). We used a substitution rate of 0.0105 substitutions/site/million years, the standard Cytb molecular clock (Weir & Schluter, 2008), and modified the value for combined mtDNA sequences as 0.0115 substitutions/site/million years following Song et al. (2015). MCMC chains were run for more than 100 million generations with a relaxed log-normal molecular clock model and a coalescent exponential growth tree prior. We used Tracer 1.6 to check the posterior distribution and ESSs, and TreeAnnotator 1.8 in the BEAST package to summarize trees with 'mean height'. Trees and divergence times were visualized in Figtree 1.4.

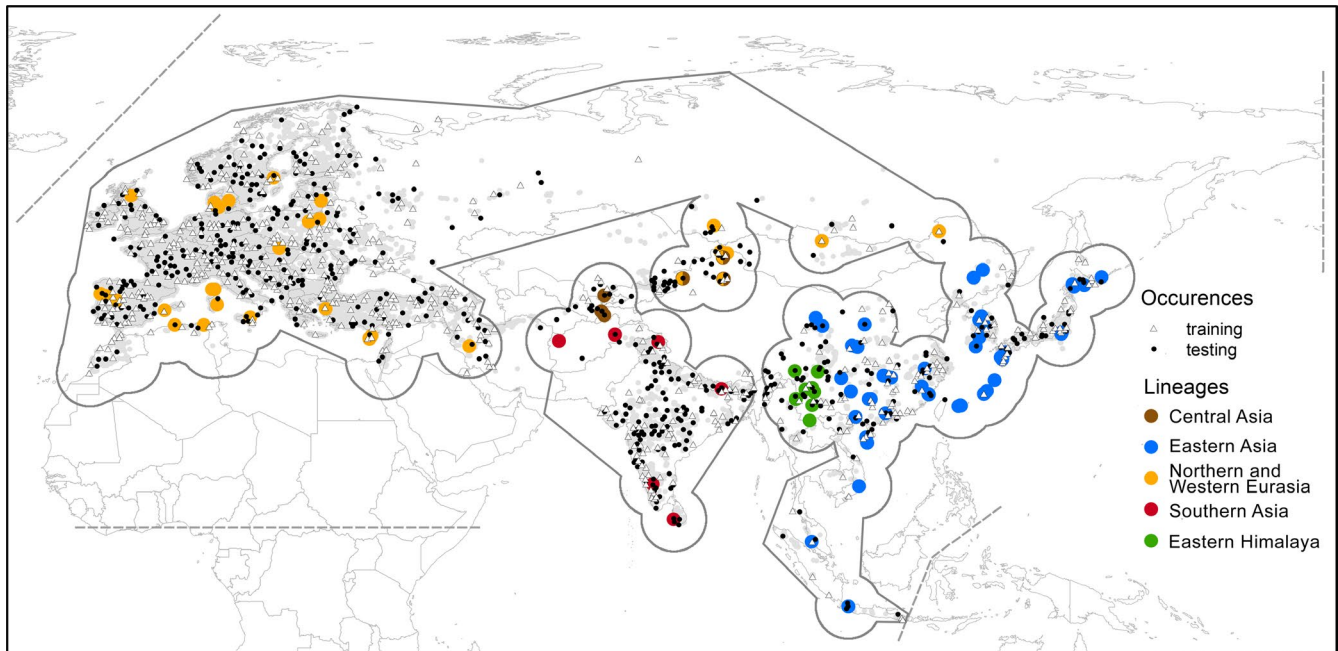
Bayesian skyline plots (BSPs), an MCMC integration and coalescent-based method (Drummond, Rambaut, Shapiro, & Pybus, 2005), were used to estimate the changes in effective population size since the time to the most recent common ancestor (TMRCA). BSP analyses were carried out for each of the major regional clades in BEAST, based on mtDNA sequences. We applied a strict molecular clock model and a Bayesian skyline tree prior, as sequence substitutions were expected to be constant for each intraspecific lineage. We ran chains for 100 million generations or more until the ESS was >200, discarding the first 10% as 'burn-in'. Results were summarized and displayed in Tracer 1.6.

## 2.5 | Distributional changes

To explore whether lineage divergence, genetic differentiation and demographic dynamics responded to distribution dynamics of suitable habitat for the complex during Pleistocene glacial–interglacial cycles, we assessed those dynamics using ecological niche modelling (ENM). For development of detailed ENMs, crucial data inputs include occurrence data, environmental data and a priori hypotheses of areas that have been accessible to species over relevant time periods (Barve et al., 2011). For occurrence data, we drew 121,041 primary occurrence records from the eBird Basic Dataset (eBird, 2017), which we subjected to detailed procedures of data cleaning and filtering. First, we inspected data records for obvious geographic outliers, and compared maps of these records to range summaries available for the species complex (Gosler et al., 2007). To reduce the effects of spatial autocorrelation, we reduced data densities such that no pairs of points would be <50 km apart, using the 'Spatially Rarefy Occurrence Data' tool in SDM Toolbox v2.2b in ArcGIS v10.5.1. Additionally, since the data were heavily biased towards European records (Peterson & Anamza, 2015), we reduced data from that region still further, via random selection, striving for equal densities for each of the major clades, leaving a total of 1,112 records for actual use in analysis (Figure 2). A randomly selected half of these points was used for model calibration, and the other half was used for evaluating the predictions of the resulting models.

As an initial hypothesis of the accessible area ( $M$ ) for Great Tit complex populations, we generated a 5° (~500 km) buffer around the 67 genotyped populations of the complex. We modified this initial hypothesis via careful consideration of finer-scale barriers, such as sea channels (Figure 2). To avoid analysis of so-called 'Wallacean species', which have distributional limits set by dispersal barriers rather than





**FIGURE 2** Distribution of Great Tit occurrences and geographic extents for ecological niche models. Light grey points represent eBird records for the species complex; white triangles and black circles are the training and testing datasets used in the ENM, respectively. The continuous grey line depicts the limits of the accessible area hypothesis (M) that was used as our ENM calibration area; the limits of the area over which models were projected are shown as lighter dashed lines. ENM, ecological niche modelling

by climatic or other environmental characteristics (Saupe et al. 2012), we analysed the entire complex, assuming negligible differentiation in niches between clades, which offered the important advantage of having a broader and more environmentally diverse accessible area.

To characterize environmental landscapes, we chose 15 'bioclimatic' variables at a spatial resolution of 10' (~17 km) from the WorldClim climate data archive (four variables including combinations of precipitation and temperature were excluded owing to known spatial artefacts: Bio8, Bio9, Bio18 and Bio19; Escobar, Lira-Noriega, Medina-Vogel, & Peterson, 2014). We used a principal components analysis (PCA) to reduce dimensionality of the environmental data (prcomp function; RStudio v 1.1.383), and selected the first eight components for analysis, as they explained >99% of the overall variance. We transferred the equations from the PCA onto a LGM (22,000 yr ago) climate scenario (CCSM4) to permit effective model transfers in a space with fewer environmental dimensions (Owens et al., 2013). These variables were projected onto a geographic extent using the ENMGadgets R package (Barve & Barve, 2013), considering broader LGM coastlines, and including North Africa, and all of Eurasia south and east to Wallace's Line (Figure 2).

In a model selection exercise (Warren & Seifert, 2011), we assessed models including all 595 possible combinations of 5 response types (i.e. linear, quadratic, product, threshold and hinge), 17 regularization multiplier values (i.e. 0.1, 0.2, ... 1, 2, ... 6, 8, 10), and 7 sets of environmental data (i.e. PC 1–2, PC 1–3, and so on up to PC 1–8; Table S3). We used the kuenm R package (Cobos, Peterson, Barve, & Osorio-Olvera, 2019) to select as best models those model parameterizations that were statistically significant (partial ROC; Peterson,

Papeş, & Soberón, 2008), showed <7% omission rates, and presented the lowest AICc values.

Resulting models were thresholded in ArcGIS to a 20% calibration omission threshold (5% and 10% thresholds were also explored, but only the 20% results are presented herein because results were closely similar). We combined binary versions of present and LGM distributional summaries to identify areas that have changed in suitability since LGM as suitable at LGM but not presently (range loss), or suitable now but not at LGM (range gain).

### 3 | RESULTS

#### 3.1 | Nucleotide polymorphism and evolutionary properties

We amplified 722 bp of Cytb and 737 bp of ND2 from 340 individuals, 281 bp of Fib5 from 275 individuals, and 334 bp of TGFB2 from 279 individuals. For the concatenated mitochondrial DNA (mtDNA) dataset, 160 haplotypes in total were defined, haplotype diversity (Hd) was 0.955, and nucleotide diversity (Pi) was 0.0197. Hd and Pi for Fib5 (0.722, 0.00493) and TGFB2 (0.941, 0.00989) were lower than those in the mtDNA. Tajima's D and Fu's Fs showed negative values for the mtDNA sequences, and the Fu's Fs statistics were significant ( $p < .05$ ) in ND2 and in mtDNA overall, indicating significant deviation from neutrality. For the two nuclear introns, we recorded a significant negative Tajima's D value only for Fib5, implying less deviation from neutral evolution (Table 1).

**TABLE 1** Nucleotide polymorphism and results of neutrality tests for mitochondrial and nuclear genes

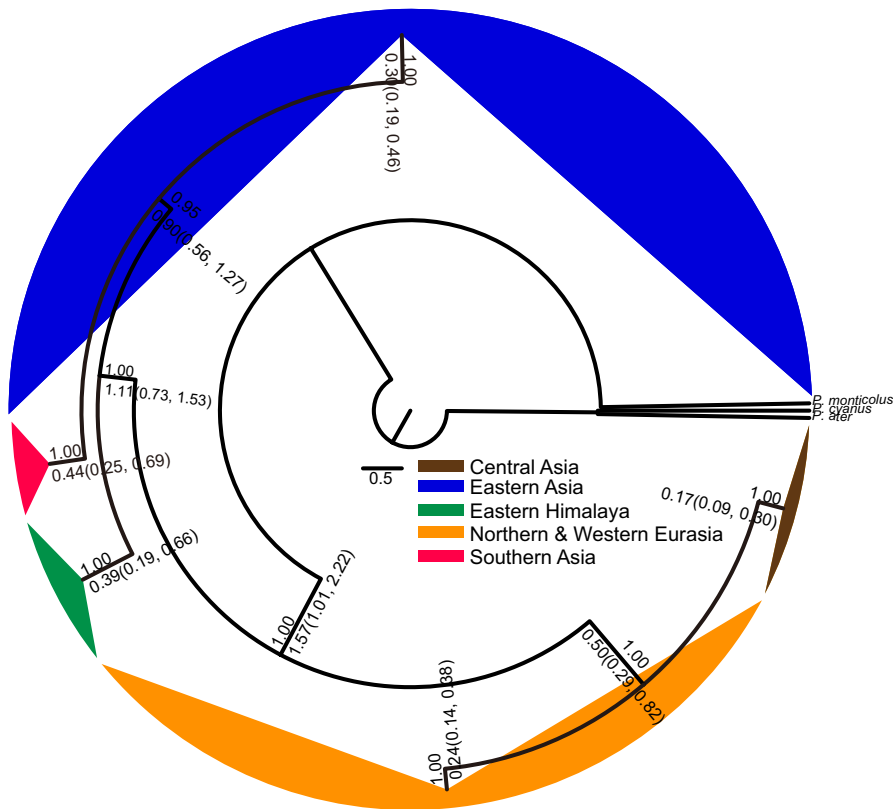
Gene	N	L	H	Hd	Pi	Tajima's D	Fu's Fs
Cytb	340	722	96	0.912	0.0179	-0.736	0.10 > p > .05
ND2	340	737	94	0.874	0.0214	-0.512	p < .05
Combined mtDNA	340	1,459	160	0.955	0.0197	-0.631	p < .05
TGFB2	279	334	101	0.941	0.0099	-1.375	p > .10
Fib5	275	281	53	0.722	0.0049	-2.149	p > .10

Abbreviations: H, number of haplotypes; Hd, haplotype diversity; L, gene fragment length; N, sample size; Pi, nucleotide diversity.

3.2 | Phylogenetic structure and divergence times

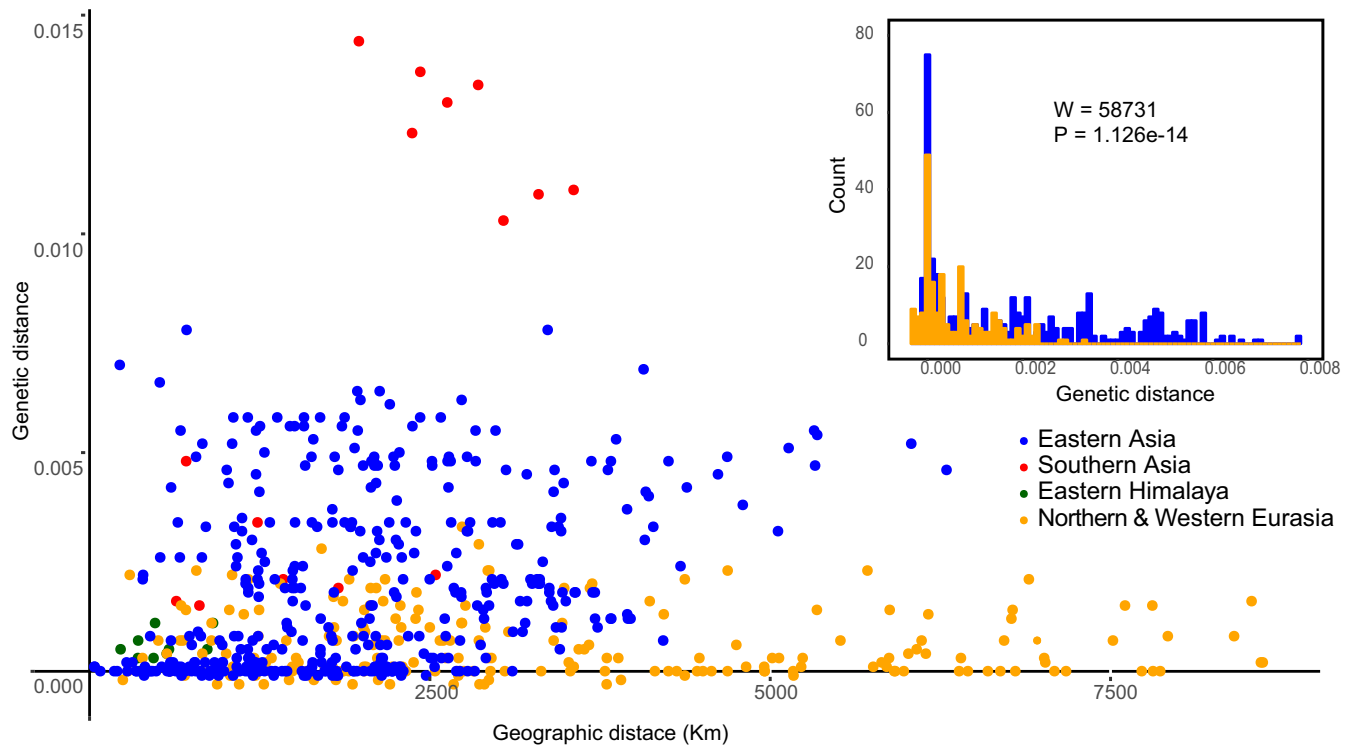
The best fitting models were GTR + I + G for the combined mtDNA sequences, K80 + G for Fib5 and SYM + I + G for TGFB2. ML and BI analyses based on combined mitochondrial sequences produced closely similar tree topologies (Figure 3, Figure S1). Here, we present the BI-based phylogeny (Figure 3), which included five well-supported clades: the most broadly distributed clade ('Northern and Western Eurasia', NWE) was composed of individuals from western Europe, the Iberian Peninsula and North Africa, and populations from across the north part of Eurasia all the way to the Russian Far East, corresponding to the green-backed 'nominat*e major group*' of Vaurie (1959). The Central Asia clade (CA), corresponding to the sometimes-recognized *Parus bokharensis*, included individuals from Uzbekistan, Tajikistan and the Xinjiang Uygur Autonomous Region of western China. NWE was most closely related to the CA clade, and the two were parapatric in Xinjiang (Figure 1).

The other widely distributed clade, Eastern Asia (EA), consisted of individuals from East Asia and Southeast Asia, including China (except Xinjiang, the Qinghai-Tibet Plateau, QTP, and neighbouring mountain regions; see below), South Korea, the Japanese Archipelago, the Indochinese Peninsula, Malaysia and Java. Vaurie (1959) placed the Chinese, Korean and Japanese populations in the '*minor group*', whereas he placed the Indochinese, Malaysian and Javan populations in the '*cinereus group*'. The Eastern Himalaya (EH) clade was composed of individuals from north-western Sichuan, western Yunnan and south-eastern Tibet, spanning the eastern portion of the Himalayas; this clade corresponds to the subspecies *tibetanus* and *subtibetanus* within (Vaurie's 1959) '*minor group*'. The Southern Asia (SA) clade contained individuals from Sri Lanka, the Indian Subcontinent, and parts of Afghanistan, corresponding to Vaurie's (1959) '*cinereus group*' (together with the Southeast Asian populations in our EA clade). The EA clade was sister to the SA clade, and EH was sister to those two clades.



**FIGURE 3** Bayesian tree based on mtDNA sequence data. Values above branches at nodes indicate Bayesian posterior probabilities (PP), and values below the branches at nodes indicate the median value of the branch heights with 95% highest posterior distribution (HPD). The colours represent the major geographic clades

**FIGURE 4** STRUCTURE results based on genotypes of the two nuclear introns, with different values for  $K$  (2–5). The acronyms under the barlots refer to sampling localities (see Table S1).



**FIGURE 5** Genetic pairwise distance for populations within geographic lineages. The populations of Central Asia were excluded from analysis, as only two populations remained in this lineage when the contact zone populations were discarded. The subset shows bin counts of pairwise values for East Asia and Northern and Western Eurasia clades, with a bin size of pairwise distances of 0.0001. Wilcoxon tests of difference between the two pairwise distance matrices of the East Asia and Northern and Western Eurasia are shown

**TABLE 2** Mantel test results of isolation by distance for geographic clades

Clade	Population numbers	Z statistics	p
Eastern Asia	30	10419.41	.071
Eastern Himalaya	5	58.05	.383
Northern and Western Eurasia	22	3702.22	.089
Southern Asia	6	420.31	.199

models with the lowest AICc values among the models that met the significance and omission criteria (Table S3). The model with the lowest AICc had an unacceptably high omission rate (9.4%), and so was not included. The 11 models with lowest AICc values were similar in

their structure (i.e. identifying a more restricted suitable area); varying the number of variables or inclusion of different features did not change their results markedly (Figure S4). These 11 models identified most of Western Europe, South Asia and Southeast Asia, and the Mediterranean coast of North Africa as suitable. However, one of these parametrizations (model 114, see Table S3) identified parts of the Sahara Desert as suitable, and all of them excluded most of Russia. The other two models (1qp6 and 1p3, both with regularization multiplier values of 1) were broader and more consistent, identifying most of Eurasia as suitable for the complex, but excluding the Arabian Peninsula, south-eastern India, and north-eastern Siberia; one model excluded the QTP as well (model 1qp6, Table S3, Figure S3).

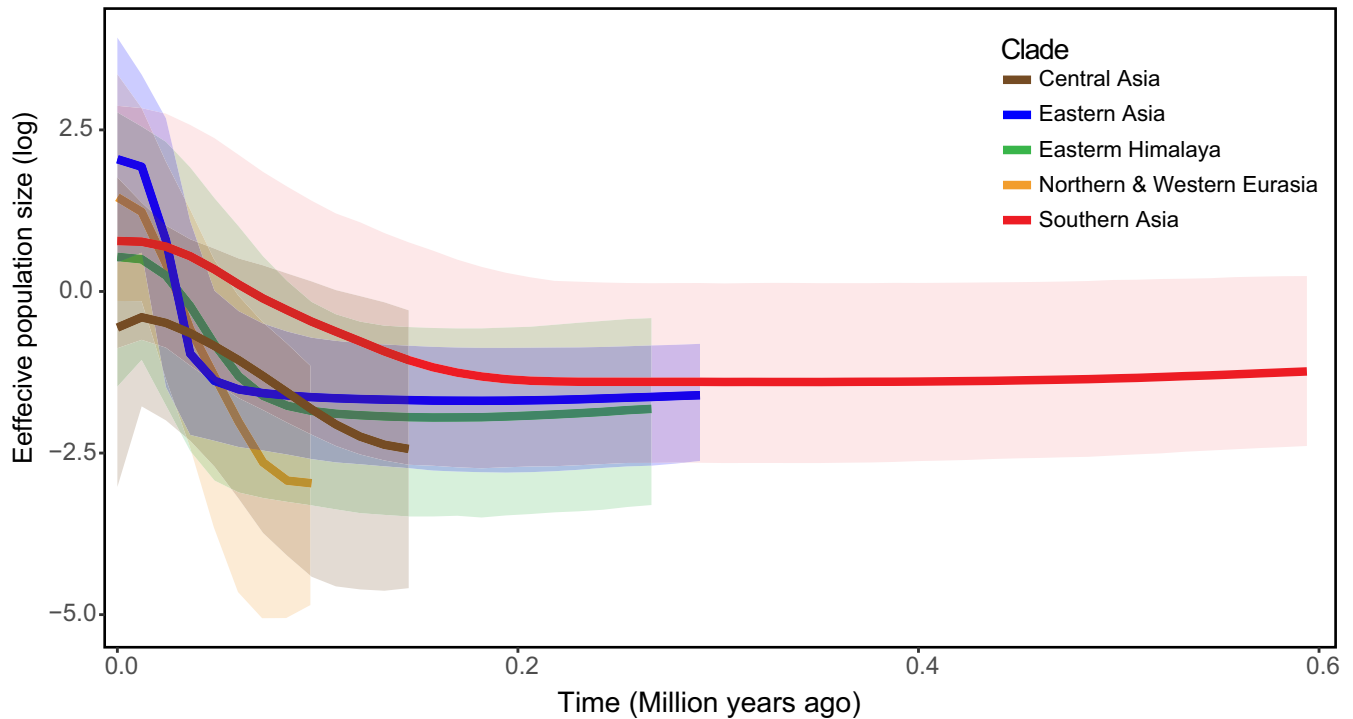
Transferring models to LGM climate conditions showed greater environmental continuity and broader suitable areas for the complex across much of its range (Figure 7 and Figure S5).

Clade	N	H	Hd	Pi (%)	Tajima's D	Fu's Fs
Central Asia	25	13	0.917	0.221	-1.529	-1.956
Eastern Asia	171	69	0.879	0.261	-2.258**	-4.661**
Eastern Himalaya	22	15	0.948	0.268	-1.715	-1.776
Northern and Western Eurasia	108	49	0.859	0.175	-2.346**	-4.756**
Southern Asia	14	14	1.000	0.598	-0.794	-0.071

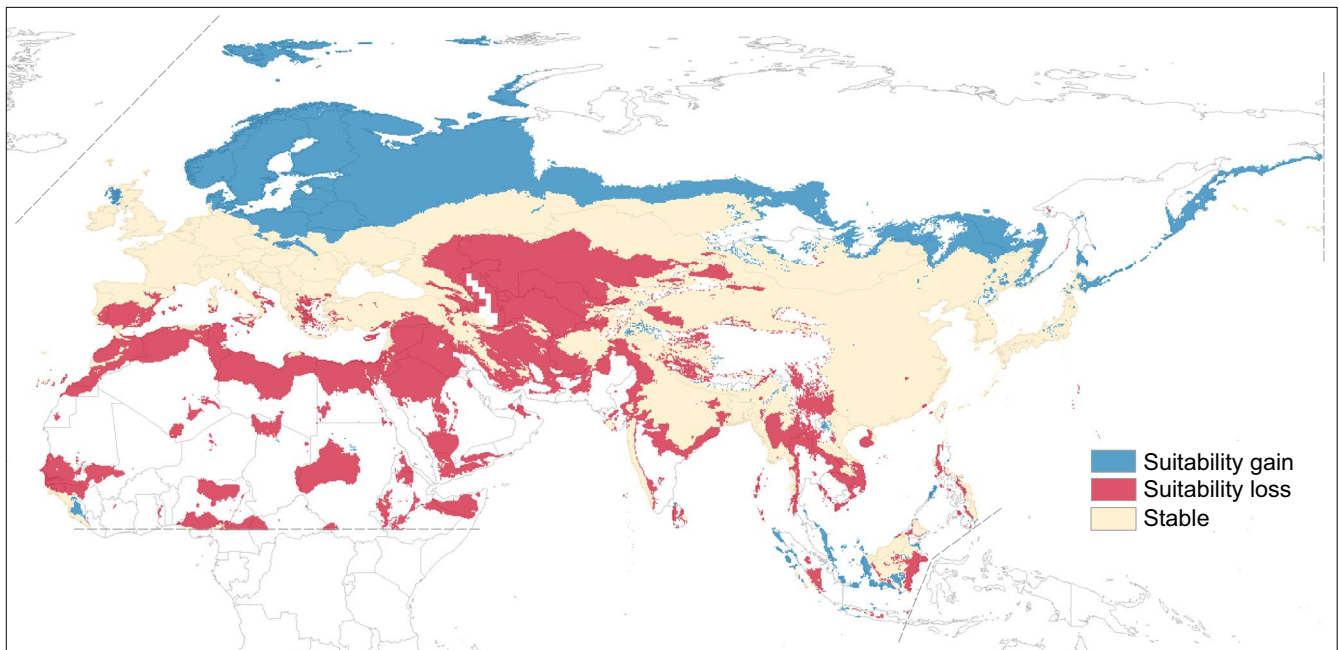
\*\*indicates the  $p < .01$

**TABLE 3** Genetic parameters for regional clades





**FIGURE 6** Bayesian skyline plots of historical demographic trends in regional clades. Estimated medians are joined by a solid line, while the shaded regions delineate 95% posterior density limits



**FIGURE 7** Change in suitability for the Great Tit complex across Eurasia and North Africa from the Last Glacial Maximum to the present. The suitability turnover is obtained by the addition of two binary models, based on the most suitable model projection, 1qp6 in each time period, with a 20% omission threshold (see Table S3 for details). Dashed grey lines represent the limits of the models projection area (G)

Some suitability loss in the transition from LGM to present was noted along the southern fringes of the study area (i.e. in North Africa, Central Asia and India); in contrast, northern Eurasia showed gains in suitability over that time period, particularly in

western Eurasia (Figure S6). We noted that suitability turnover varied among models as regards North Africa (e.g. in Morocco) and southern Spain, with some models showing loss and others showing stability.

## 4 | DISCUSSION

### 4.1 | Lineage divergence and systematics of the Great Tits

Despite a wealth of documentation of geographic variation in the form of recognized subspecies (Gosler et al., 2007) in the Great Tit complex, most of these described taxa present subtle and perhaps clinal variants in pigmentation, size and proportions, which may be artefacts of small sample sizes, or represent arbitrary delimitation or adaptive responses to local environmental conditions (Dufva & Allander, 1995; Hórak, Ots, Vellau, Spottiswoode, & Pape Møller, 2001; Slagsvold & Lifjeld, 1985; Zink & Vázquez-Miranda, 2019). Our results, based on more comprehensive geographic sampling than previous studies, showed five major clades with high support in mtDNA sequences, generally in line with previous results (Kvist et al., 2003; Päckert et al., 2005; Zhao et al., 2012).

Although the Great Tit complex is rather young, it apparently includes some lineages that are at the early stages of the speciation process, and recently diverged taxa are generally difficult to classify or rank objectively. Our data do not support recognition of *P. bokharensis* as a distinct species (corroborating results presented by Päckert et al., 2005 based on analyses of mtDNA and songs). Although the 'bokharensis group' is clearly different from the adjacent 'major group' in plumage and structure (Gosler et al., 2007), the two groups apparently hybridize frequently, as suggested by our mtDNA data and also by field observations of assumed hybrids from several places (Harrap & Quinn 1995; Päckert et al., 2005; P.A. pers. obs.; Zhao et al., 2012). The 'major' and 'minor groups' come into contact in the Amur River area in the Russian Far East and adjacent parts of China; although they do hybridize, they appear to be at least partly reproductively isolated (see Päckert et al., 2005 and references therein). However, splitting the Great Tit complex into two species, corresponding to the two primary clades in the mtDNA tree, would be unsatisfactory in light of the conflict between the mtDNA tree and our Structure analyses of the nuclear introns regarding the SA clade. In other words, it would be unclear to which of the two species the taxa in the SA clade should belong. Genomic data and more intensive geographic sampling, especially from the southern part of the distribution of the complex, are needed to illuminate the relationships and evolutionary history of this complex.

Our results offer detailed scenarios about the Great Tit lineage divergence and the geographic distributions of those lineages. First, we confirmed the unique mtDNA clade first noted by Zhao et al. (2012), referred to here as the Eastern Himalaya clade. This clade is narrowly distributed along the eastern portion of the Himalaya range, and corresponding to subspecies *tibetanus* and *subtibetanus* in Vaurie's (1959) 'minor group'.

Second, we clarified the geographic range of the EA clade, which comprises individuals from China (except Xinjiang, the QTP, and the mountains along the eastern rim of the QTP), South Korea, the Japanese Archipelago, Indochina, Malaysia and Java.

All could be included in the 'minor group' of Vaurie (1959). The phylogenetic position of populations of Vietnam, Malaysia and Java, indicates the need for a taxonomic revision of *P. minor* and *P. cinereus* as circumscribed by for example, Gill and Donsker (2019), and suggests a close relationship of the populations across all East Asia, despite differences in plumage and vocalizations (Gosler et al., 2007).

### 4.2 | Historical biogeography

The mtDNA phylogeny recovered five clades in the Great Tit complex, of which three are distributed in the east from South Asia through the eastern Himalaya to a broad area of the East Asian mainland. BSP results also showed longer TMRCA timescales for eastern clades than for western clades. Considering that the sister species (Green-backed Tit, *Parus monticolus*) is distributed in the eastern Himalayas, southern China and Taiwan (Johansson et al., 2013), Great Tits might have originated somewhere in the eastern Himalayas or southern parts of Asia, consistent with previous hypotheses (Johansson, Nylinder, Ohlson, & Tietze, 2018; Tietze & Borthakur, 2012).

The oldest divergence among clades (ca. 1.6 Ma) is between eastern/southern (EA, EH, SA) and western/northern (CA and NWE) clades. Similar divergence patterns have been noted in other avian taxa, such as Barn Swallows (*Hirundo rustica*; Zink, Pavlova, Rohwer, & Drovetski, 2006), White Wagtails (*Motacilla alba*; Li, Dong, et al., 2016) and magpies (*Pica* spp.; Song et al., 2018). This coincidence suggests a dominant geographic feature shaping phylogeographical structure. The arid desert region in Central Asia is a likely candidate barrier that could account for this dominant pattern of east-west divergence (Abramowski et al., 2006; Cai, Fang, Wu, Miao, & Appel, 2012; Wang et al., 2016; Zhang et al., 2012).

Subsequent divergence events for eastern clades (1.1 and 0.9 Ma) were dated as occurring earlier than divergence for western clades (ca. 0.5 Ma). The distribution of three lineages along the east to south edge of the Himalayas indicates that eastern Great Tit populations may have been structured by the topographic complexity of this area, which presents highly diversified habitats and environments (Lei, Wei, Zhao, Yin, & Lu, 2007), and is well-known as ranking prominently among global biodiversity hotspots (Myers, Mittermeier, Mittermeier, Fonseca, & Kent, 2000). Lineage divergence in western populations seems affected instead by later climatic changes, especially the marked temperature oscillations during the Pleistocene. The comparatively narrower distribution and younger population history imply that the CA clade is recently derived. Great Tit populations likely persisted throughout the Pleistocene in isolated riparian woodlands in the montane foothill zone that forms the southern boundary of the arid lowlands from Turkestan to Kazakhstan (Pavlova et al., 2006) and later expanded in response to favourable climatic conditions.

### 4.3 | Geographic variation of glacial effects

In accordance with estimated divergence times, the BSP curves showed the longest population history in the SA clade, and second longest in the EA clade, whereas the NWE and CA clades displayed comparatively more shallow population histories (Figure 5). The same pattern was evident in inspections of pairwise genetic distances. Top pairwise distances within an area were related to island populations in East Asia (e.g. Japan and the Ryukyu Islands), indicating another important source of population structure and differentiation (Peterson et al., 2015).

Results of palaeovegetation studies indicate that glacial cover extended unevenly, and that vegetation types shrank southward during the LGM variably across Eurasia (Ashastina et al., 2018; Janská et al., 2017; Qian & Ricklefs, 2001; Svendsen et al., 2004; Yu et al., 2000). However, ENMs indicated greater environmental continuity and broader suitable areas of climate conditions for the Great Tit complex during the LGM (Figure 7, Figure S5). These findings are in line with scenarios for the sister species, the Green-backed Tit (*Parus monticolus*), for which the LGM distribution was estimated to be broader than during the Last Interglacial and under present climate conditions (Wang et al., 2013). Our results also suggest a northward shift of the suitable areas for Great Tits since the LGM. Demographic expansions since the last glacial cycle were recorded in all five clades, and the effective population sizes in southern and central Asia have been more stable than the East Asian relatives, yet the East Asian clade showed the steepest curve of recent population size increase.

One possible explanation for the significant population changes in East Asia is that the distribution of the EA clade is vast, covering an area from north-eastern China, South Korea and the Japanese Archipelago, to Malaysia, Java and Indochina. The marginal populations would shrink both in LGM (northern edge) and after the LGM (southern edge), according to our ENM results. Also, post-LGM sea-level shifts would result in greater isolation and therefore genetic structuring of East Asian island populations (Voris, 2000). Great Tits inhabit diverse forest types, from conifer forests and boreal taiga in Siberia, to dry deciduous forest in Myanmar and Thailand, and are common in gardens and parks in large parts of its range. This broad set of habitat associations assure that Great Tit populations would be stable or even expand, if forests developed or persisted during glacial periods. On the other hand, if vegetation changed from forest to grassland or *vice versa* (Harrison, Yu, Takahara, & Prentice, 2001; Yu et al., 2000), populations might respond more dramatically, either in population size or geographic range, which in turn would lead to more dramatic demographic signatures.

Consistent with these interpretations, genetic diversity was lower in populations of Mongolia, Siberia and the eastern periphery of the Chinese mainland, and higher in populations from the eastern edge of the Himalayas and Central Asia. It is not surprising that genetic diversity has been preserved in populations of the Indian Subcontinent, but the highest genetic diversity occurring in Xinjiang was startling. Some prominent geographic features in Central Asia, such as the Pamir Plateau, Taklimakan and Junggar Basin deserts, where suitable habitats are few and scattered for Great Tits both now and during the LGM, therefore would lead to a low genetic

diversity in local populations. We propose that geographic complexity, such as river valleys alongside major mountain ranges, may have offered refugia or micro-refugia. Later secondary contact between the Central Asia and Northern and Western Eurasia clades have further contributed to elevating the genetic diversity in Xinjiang. On the other hand, for populations from Siberia, suitable areas appear to have shifted and reduced more dramatically, leading to more significant population changes over the end of the Pleistocene.

## 5 | CONCLUSIONS

With extensive geographic sampling across the range of the Great Tit complex, our results confirmed previous findings of five distinct mitochondrial lineages, a southern Asia origin, and subsequent expansion across the whole of Eurasia. We found a generally earlier set of divergence times between eastern lineages than between western lineages, yet our results showed a consistent increase in effective population size among lineages since the LGM, implying habitat changes as indicated by the ENM results. Geographic variation of glacial effects, related to geographic complexity and global climate dynamics, has functioned to create refugia in which ancient genetic diversity was able to persist in some regions.

## ACKNOWLEDGEMENTS

We thank Anton Kristin at the Institute of Forest Ecology, Slovakia; Alexander M. Shestopalov at Novosibirsk State University, Russia; and Jae Chun Choe at Ewha Womans University, South Korea, for kindly providing regional samples. We also thank Na Zhao, Zuohua Yin, Haitao Wang, Kaifeng Wang, Xuebin Gao, Chuanyin Dai and Ming Ma for their assistance during field collections. We are grateful to the chief editors Jon P. Sadler and Peter Linder, the editor Parmakelis Aristeidis and the anonymous reviewers for their important comments and suggestions to the manuscript. The research work of G.S. is funded by NSFC (31471991, 31572291). The project was also funded by NSFC (31630069, 31572249) and the National Science and Technology Major Project (2018ZX10101004). P.A. acknowledges support from the Swedish Research Council (grant No. 2015-04402), Jorvall Foundation, and Mark and Mo Constantine. Contributions made by HLM were supported by the National Science Foundation (NSF, DBI-0821703).

## DATA AVAILABILITY STATEMENT

Sample localities are given in Table S1, and sequences included in this study are deposited in GenBank (accession numbers MT212810–MT213943).

## ORCID

Gang Song  <https://orcid.org/0000-0002-0190-8315>

Per Alström  <https://orcid.org/0000-0001-7182-2763>

Ulf S. Johansson  <https://orcid.org/0000-0002-2544-8528>

Per G. P. Ericson  <https://orcid.org/0000-0002-4143-9998>

Fumin Lei  <https://orcid.org/0000-0001-9920-8167>

## REFERENCES

- Abramowski, U., Bergau, A., Seebach, D., Zech, R., Glaser, B., Sosin, P., ... Zech, W. (2006). Pleistocene glaciations of Central Asia: Results from Be-10 surface exposure ages of erratic boulders from the Pamir (Tajikistan), and the Alay-Turkestan range (Kyrgyzstan). *Quaternary Science Reviews*, 25, 1080–1096. <https://doi.org/10.1016/j.quascirev.2005.10.003>
- Ashastina, K., Kuzmina, S., Rudaya, N., Troeva, E., Schoch, W. H., Römermann, C., ... Kienast, F. (2018). Woodlands and steppes: Pleistocene vegetation in Yakutia's most continental part recorded in the Batagay permafrost sequence. *Quaternary Science Reviews*, 196, 38–61. <https://doi.org/10.1016/j.quascirev.2018.07.032>
- Barve, N., & Barve, V. (2013). ENMGadgets: tools for pre and post processing in ENM workflows. Retrieved from <https://github.com/vijaybarve/ENMGadgets>
- Barve, N., Barve, V., Jimenez-Valverde, A., Lira-Noriega, A., Maher, S. P., Peterson, A. T., ... Villalobos, F. (2011). The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling*, 222, 1810–1819. <https://doi.org/10.1016/j.ecolmodel.2011.02.011>
- Bosse, M., Spurgin, L. G., Laine, V. N., Cole, E. F., Firth, J. A., Gienapp, P., ... Slate, J. (2017). Recent natural selection causes adaptive evolution of an avian polygenic trait. *Science*, 358, 365–368. <https://doi.org/10.1126/science.aal3298>
- Bouckaert, R., Heled, J., Kuhnert, 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, e1003537. <https://doi.org/10.1371/journal.pcbi.1003537>
- Burbrink, F. T., Chan, Y. L., Myers, E. A., Ruane, S., Smith, B. T., & Hickerson, M. J. (2016). Asynchronous demographic responses to Pleistocene climate change in eastern Nearctic vertebrates. *Ecology Letters*, 19, 1457–1467. <https://doi.org/10.1111/ele.12695>
- Cai, M. T., Fang, X. M., Wu, F. L., Miao, Y. F., & Appel, E. (2012). Pliocene-Pleistocene stepwise drying of Central Asia: Evidence from paleomagnetism and sporopollen record of the deep borehole SG-3 in the western Qaidam Basin, NE Tibetan Plateau. *Global and Planetary Change*, 94–95, 72–81. <https://doi.org/10.1016/j.gloplacha.2012.07.002>
- Cauchard, L., Doucet, S. M., Boogert, N. J., Angers, B., & Doligez, B. (2017). The relationship between plumage colouration, problem-solving and learning performance in great tits *Parus major*. *Journal of Avian Biology*, 48, 1246–1253. <https://doi.org/10.1111/jav.00953>
- Cobos, M. E., Peterson, A. T., Barve, N., & Osorio-Olvera, L. (2019). kuenm: An R package for detailed development of ecological niche models using Maxent. *PeerJ*, 7, e6281. <https://doi.org/10.7717/peerj.6281>
- Dalén, L., Fuglei, E., Hersteinsson, P., Kapel, C. M. O., Roth, J. D., Samelius, G., ... Angerbjörn, A. (2005). Population history and genetic structure of a circumpolar species: The arctic fox. *Biological Journal of the Linnean Society*, 84, 79–89. <https://doi.org/10.1111/j.1095-8312.2005.00415.x>
- Drummond, A. J., Rambaut, A., Shapiro, B., & Pybus, O. G. (2005). Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution*, 22, 1185–1192. <https://doi.org/10.1093/molbev/msi103>
- Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, 29, 1969–1973. <https://doi.org/10.1093/molbev/mss075>
- Dufva, R., & Allander, K. (1995). Intraspecific variation in plumage coloration reflects immune-response in Great Tit (*Parus major*) males. *Functional Ecology*, 9, 785–789. <https://doi.org/10.2307/2390253>
- EBIRD. (2017). eBird basic dataset. Version: EBD\_relNov-2017. Ithaca, New York: Cornell Lab of Ornithology; 2017.
- Escobar, L. E., Lira-Noriega, A., Medina-Vogel, G., & Peterson, A. T. (2014). Potential for spread of the white-nose fungus (*Pseudogymnoascus destructans*) in the Americas: Use of Maxent and NicheA to assure strict model transference. *Geospatial Health*, 9, 221–229. <https://doi.org/10.4081/gh.2014.19>
- Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, 14, 2611–2620. <https://doi.org/10.1111/j.1365-294x.2005.02553.x>
- Gill, F., & Donsker, D. (2019). IOC World Bird List (v9.1). <https://doi.org/10.14344/IOC.ML.9.1>
- Gosler, A., Clement, P., & Christie, D. A. (2007). Great Tit (*Parus major*). In J. del Hoyo, A. Elliott, J. Sargatal, D. A. Christie, & E. de Juana (Eds.), *Handbook of the birds of the world alive*. Barcelona, Spain: Lynx Edicions. Retrieved from <https://www.hbw.com/node/59900>
- Guindon, S., & Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, 52, 696–704. <https://doi.org/10.1080/10635150390235520>
- Harrap, S., & Quinn, D. (1995). *Tits, nuthatches & treecreepers*. Russel Friedman Books.
- Harrison, S. P., Yu, G., Takahara, H., & Prentice, I. C. (2001). Palaeovegetation - Diversity of temperate plants in east Asia. *Nature*, 413, 129–130.
- Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, 405, 907–913. <https://doi.org/10.1038/35016000>
- Hörak, P., Ots, I., Vellau, H., Spottiswoode, C., & Pape Møller, A. (2001). Carotenoid-based plumage coloration reflects hemoparasite infection and local survival in breeding great tits. *Oecologia*, 126, 166–173. <https://doi.org/10.1007/s004420000513>
- Hosner, P. A., Liu, H., Peterson, A. T., & Moyle, R. G. (2015). Rethinking phylogeographic structure and historical refugia in the rufous-capped babbler *Cyanoderma ruficeps* in light of range-wide genetic sampling and paleodistributional reconstructions. *Current Zoology*, 61, 901–909. <https://doi.org/10.1093/czoolo/61.5.901>
- Hubisz, M. J., Falush, D., Stephens, M., & Pritchard, J. K. (2009). Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, 9, 1322–1332. <https://doi.org/10.1111/j.1755-0998.2009.02591.x>
- Hung, C. M., Drovetski, S. V., & Zink, R. M. (2012). Multilocus coalescence analyses support a mtDNA-based phylogeographic history for a widespread Palearctic passerine bird, *Sitta europea*. *Evolution*, 66, 2850–2864. <https://doi.org/10.1111/j.1558-5646.2012.01657.x>
- Hung, C. M., Drovetski, S. V., & Zink, R. M. (2013a). Multilocus test of the absence of mtDNA phylogeographic structure in a widespread wader, the Common Sandpiper (*Actitis hypoleucos*). *Journal of Ornithology*, 154, 1105–1113. <https://doi.org/10.1007/s10336-013-0980-1>
- Hung, C. M., Drovetski, S. V., & Zink, R. M. (2013b). Recent allopatric divergence and niche evolution in a widespread Palearctic bird, the common rosefinch (*Carpodacus erythrinus*). *Molecular Phylogenetics and Evolution*, 66, 103–111. <https://doi.org/10.1016/j.ympev.2012.09.012>
- Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, 23, 1801–1806. <https://doi.org/10.1093/bioinformatics/btm233>
- Janská, V., Jiménez-Alfaro, B., Chytrý, M., Divíšek, J., Anenkhonov, O., Korolyuk, A., ... Culek, M. (2017). Palaeodistribution modelling of European vegetation types at the Last Glacial Maximum using modern analogues from Siberia: Prospects and limitations. *Quaternary Science Reviews*, 159, 103–115. <https://doi.org/10.1016/j.quascirev.2017.01.011>
- Johansson, U. S., Ekman, J., Bowie, R. C. K., Halvarsson, P., Ohlson, J. I., Price, T. D., & Ericson, P. G. P. (2013). A complete multilocus species phylogeny of the tits and chickadees (Aves: Paridae). *Molecular Phylogenetics and Evolution*, 69, 852–860. <https://doi.org/10.1016/j.ympev.2013.06.019>
- Johansson, U. S., Nylander, S., Ohlson, J. I., & Tietze, D. T. (2018). Reconstruction of the late Miocene biogeographical history of tits and chickadees (Aves: Passeriformes: Paridae): A comparison between discrete area analyses and probabilistic diffusion approach. *Journal of Biogeography*, 45, 14–25. <https://doi.org/10.1111/jbi.13095>



- Klicka, L. B., Campillo, L. C., Manthey, J. D., & Qu, Y. H. (2015). Comparative phylogeography and niche modeling for three species complexes of SE China (*Paradoxornis* spp., *Pycnonotus* spp., *Spizixos* spp.). *Current Zoology*, 61, 943–950. <https://doi.org/10.1093/czoolo/61.5.943>
- Kvist, L., Martens, J., Higuchi, H., Nazarenko, A. A., Valchuk, O. P., & Orell, M. (2003). Evolution and genetic structure of the great tit (*Parus major*) complex. *Proceedings of the Royal Society B*, 270, 1447–1454. <https://doi.org/10.1098/rspb.2002.2321>
- Lei, F. M., Wei, G. A., Zhao, H. F., Yin, Z. H., & Lu, J. L. (2007). China sub-regional avian endemism and biodiversity conservation. *Biodiversity and Conservation*, 16, 1119–1130. <https://doi.org/10.1007/s10531-006-9080-3>
- Li, S. H., Yeung, C. K. L., Feinstein, J., Han, L. X., Manh, H. L., Wang, C. X., & Ding, P. (2009). Sailing through the Late Pleistocene: Unusual historical demography of an East Asian endemic, the Chinese Hwamei (*Leucodioptron canorum canorum*), during the last glacial period. *Molecular Ecology*, 18, 622–633. <https://doi.org/10.1111/j.1365-294x.2008.04028.x>
- Li, X. L., Dong, F., Lei, F. M., Alström, P., Zhang, R. Y., Odeen, A., ... Yang, X. J. (2016). Shaped by uneven Pleistocene climate: Mitochondrial phylogeographic pattern and population history of white wagtail *Motacilla alba* (Aves: Passeriformes). *Journal of Avian Biology*, 47, 263–274. <https://doi.org/10.1111/jav.00826>
- Li, X. J., Lin, L. L., Cui, A. M., Bai, J., Wang, X. Y., Xin, C., ... Lei, F. M. (2016). Taxonomic status and phylogenetic relationship of tits based on mitogenomes and nuclear segments. *Molecular Phylogenetics and Evolution*, 104, 14–20. <https://doi.org/10.1016/j.ympev.2016.07.022>
- Librado, P., & Rozas, J. (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25, 1451–1452. <https://doi.org/10.1093/bioinformatics/btp187>
- Lim, H. C., Zou, F. S., & Sheldon, F. H. (2015). Genetic differentiation in two widespread, open-forest bird species of Southeast Asia (*Copsychus saularis* and *Megalaima haemacephala*): Insights from ecological niche modeling. *Current Zoology*, 61, 922–934. <https://doi.org/10.1093/czoolo/61.5.922>
- Lovette, I. J. (2005). Glacial cycles and the tempo of avian speciation. *Trends in Ecology & Evolution*, 20, 57–59. <https://doi.org/10.1016/j.tree.2004.11.011>
- Mayr, E. (1942). *Systematics and the origin of species*. Cambridge, Massachusetts: Harvard University Press.
- Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, 403, 853–858. <https://doi.org/10.1038/35002501>
- Nei, M., & Kumar, S. (2000). *Molecular evolution and phylogenetics*. New York, NY: Oxford University Press.
- Nylander, J. A. A. (2004). *MrModeltest v2. Program distributed by the author*. Uppsala, Sweden: Evolutionary Biology Centre, Uppsala University.
- Owens, H. L., Campbell, L. P., Dornak, L. L., Saupe, E. E., Barve, N., Soberón, J., ... Peterson, A. T. (2013). Constraints on interpretation of ecological niche models by limited environmental ranges on calibration areas. *Ecological Modelling*, 263, 10–18. <https://doi.org/10.1016/j.ecolmodel.2013.04.011>
- Päckert, M., Martens, J., Eck, S., Nazarenko, A. A., Valchuk, O. P., Petri, B., & Veith, M. (2005). The great tit (*Parus major*) - a misclassified ring species. *Biological Journal of the Linnean Society*, 86, 153–174. <https://doi.org/10.1111/j.1095-8312.2005.00529.x>
- Paradis, E., Claude, J., & Strimmer, K. (2004). APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics*, 20, 289–290. <https://doi.org/10.1093/bioinformatics/btg412>
- Pavlova, A., Rohwer, S., Drovetski, S. V., & Zink, R. M. (2006). Different post-pleistocene histories of Eurasian parids. *Journal of Heredity*, 97, 389–402. <https://doi.org/10.1093/jhered/esl011>
- Pavlova, A., Zink, R. M., & Rohwer, S. (2005). Evolutionary history, population genetics, and gene flow in the Common Rosefinch (*Carpodacus erythrinus*). *Molecular Phylogenetics and Evolution*, 36, 669–681. <https://doi.org/10.1016/j.ympev.2005.02.010>
- Pavlova, A., Zink, R. M., Rohwer, S., Koblik, E. A., Red'kin, Y. A., Fadeev, I. V., & Nesterov, E. V. (2005). Mitochondrial DNA and plumage evolution in the white wagtail *Motacilla alba*. *Journal of Avian Biology*, 36, 322–336.
- Peterson, A. T., & Anamza, T. (2015). Ecological niches and present and historical geographic distributions of species: A 15-year review of frameworks, results, pitfalls, and promises. *Folia Zoologica*, 64, 207–217. <https://doi.org/10.25225/fozo.v64.i3.a3.2015>
- Peterson, A. T., Moyle, R. G., Lei, F. M., Campillo, L. C., Hosner, P. A., Klicka, L. B., ... Zou, F. S. (2015). Avian evolution and speciation in the Southeast Asian tropics. *Current Zoology*, 61, 898–900. <https://doi.org/10.1093/czoolo/61.5.898>
- Peterson, A. T., Papeş, M., & Soberón, J. (2008). Rethinking receiver operating characteristic analysis applications in ecological niche modeling. *Ecological Modelling*, 213, 63–72. <https://doi.org/10.1016/j.ecolmodel.2007.11.008>
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155, 945–959.
- Qian, H., & Ricklefs, R. E. (2001). Palaeovegetation—Diversity of temperate plants in east Asia—Reply. *Nature*, 413, 130–130.
- Qiao, H., Saupe, E. E., Soberón, J., Peterson, A. T., Myers, C. E., Collar, D. C., & Bronstein, J. L. (2016). Impacts of niche breadth and dispersal ability on macroevolutionary patterns. *American Naturalist*, 188, 149–162. <https://doi.org/10.1086/687201>
- Quan, Q., Qu, Y. H., & Lei, F. M. (2015). Genetic diversification in the East Himalayas as revealed by comparative phylogeography of the black-throated bushtit and Elliot's laughing thrush. *Current Zoology*, 61, 935–942. <https://doi.org/10.1093/czoolo/61.5.935>
- Rambaut, A., Suchard, M. A., Xie, D., & Drummond, A. J. (2014). Tracer v1.6. Computer program and documentation distributed by the author.
- Reddy, S., & Nyári, Á. S. (2015). Novel insights into the historical biogeography of the streak-breasted scimitar babbler complex (Aves: Timaliidae: *Pomatorhinus ruficollis* complex). *Current Zoology*, 61, 910–921. <https://doi.org/10.1093/czoolo/61.5.910>
- Rosenberg, N. A. (2004). DISTRUCT: A program for the graphical display of population structure. *Molecular Ecology Notes*, 4, 137–138. <https://doi.org/10.1046/j.1471-8286.2003.00566.x>
- 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, 13.
- Saupe, E. E., Barve, V., Myers, C. E., Soberón, J., Barve, N., Hensz, C. M., ... Lira-Noriega, A. (2012). Variation in niche and distribution model performance: The need for a priori assessment of key causal factors. *Ecological Modelling*, 237, 11–22.
- Slabbekoorn, H., & Peet, M. (2003). Ecology: Birds sing at a higher pitch in urban noise - Great tits hit the high notes to ensure that their mating calls are heard above the city's din. *Nature*, 424, 267–267.
- Slagsvold, T., & Lifjeld, J. T. (1985). Variation in plumage color of the Great Tit *Parus major* in relation to habitat, season and food. *Journal of Zoology*, 206, 321–328. <https://doi.org/10.1111/j.1469-7998.1985.tb05661.x>
- Song, G., Qu, Y. H., Yin, Z. H., Li, S. S., Liu, N. F., & Lei, F. M. (2009). Phylogeography of the *Alcippe morrissonia* (Aves: Timaliidae): Long population history beyond late Pleistocene glaciations. *BMC Evolutionary Biology*, 9, 143. <https://doi.org/10.1186/1471-2148-9-143>
- Song, G., Zhang, R. Y., Alström, P., Irestedt, M., Cai, T. L., Qu, Y. H., ... Lei, F. M. (2018). Complete taxon sampling of the avian genus *Pica* (magpies) reveals ancient relictual populations and synchronous Late-Pleistocene demographic expansion across the Northern



- Hemisphere. *Journal of Avian Biology*, 49, e01612. <https://doi.org/10.1111/jav.01612>
- Song, G., Zhang, R., DuBay, S. G., Qu, Y., Dong, L., Wang, W., ... Lei, F. (2015). East Asian allopatry and north Eurasian sympatry in Long-tailed Tit lineages despite similar population dynamics during the late Pleistocene. *Zoologica Scripta*, 45, 115–126. <https://doi.org/10.1111/zsc.12148>
- 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. <https://doi.org/10.1086/379378>
- Stephens, M., Smith, N. J., & Donnelly, P. (2001). A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics*, 68, 978–989. <https://doi.org/10.1086/319501>
- Stewart, J. R., Lister, A. M., Barnes, I., & Dalén, L. (2010). Refugia revisited: Individualistic responses of species in space and time. *Proceedings of the Royal Society B*, 277, 661–671. <https://doi.org/10.1098/rspb.2009.1272>
- Svendsen, J. I., Alexanderson, H., Astakhov, V. I., Demidov, I., Dowdeswell, J. A., Funder, S., ... Stein, R. (2004). Late quaternary ice sheet history of northern Eurasia. *Quaternary Science Reviews*, 23, 1229–1271. <https://doi.org/10.1016/j.quascirev.2003.12.008>
- Taberlet, P., Fumagalli, L., Wust-Saucy, A. G., & Cosson, J. F. (1998). Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, 7, 453–464. <https://doi.org/10.1046/j.1365-294x.1998.00289.x>
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., & Kumar, S. (2011). MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, 28, 2731–2739. <https://doi.org/10.1093/molbev/msr121>
- Templeton, C. N., Zollinger, S. A., & Brumm, H. (2016). Traffic noise drowns out great tit alarm calls. *Current Biology*, 26, R1173–R1174. <https://doi.org/10.1016/j.cub.2016.09.058>
- Tietze, D. T., & Borthakur, U. (2012). Historical biogeography of tits (Aves: Paridae, Remizidae). *Organisms Diversity & Evolution*, 12, 433–444. <https://doi.org/10.1007/s13127-012-0101-7>
- Vaurie, C. (1959). *The birds of the palearctic fauna, A systematic reference: Order passeriformes*. London, UK: Witherby Ltd.
- Voris, H. K. (2000). Maps of Pleistocene sea levels in Southeast Asia: Shorelines, river systems and time durations. *Journal of Biogeography*, 27, 1153–1167. <https://doi.org/10.1046/j.1365-2699.2000.00489.x>
- Wang, E., Van Wijk, R. E., Braun, M. S., & Wink, M. (2017). Gene flow and genetic drift contribute to high genetic diversity with low phylogeographical structure in European hoopoes (*Upupa epops*). *Molecular Phylogenetics and Evolution*, 113, 113–125. <https://doi.org/10.1016/j.ympev.2017.05.018>
- Wang, W., McKay, B. D., Dai, C., Zhao, N., Zhang, R., Qu, Y., ... Lei, F. (2013). Glacial expansion and diversification of an East Asian montane bird, the green-backed tit (*Parus monticolus*). *Journal of Biogeography*, 40, 1156–1169. <https://doi.org/10.1111/jbi.12055>
- Wang, X., Wei, H. T., Taheri, M., Khormali, F., Danukalova, G., & Chen, F. H. (2016). Early Pleistocene climate in western arid central Asia inferred from loess-palaeosol sequences. *Scientific Reports*, 6, 9.
- Warren, D. L., & Seifert, S. N. (2011). Ecological niche modeling in Maxent: The importance of model complexity and the performance of model selection criteria. *Ecological Applications*, 21, 335–342. <https://doi.org/10.1890/10.1171.1>
- Weir, J. T., & Schluter, D. (2008). Calibrating the avian molecular clock. *Molecular Ecology*, 17, 2321–2328. <https://doi.org/10.1111/j.1365-294x.2008.03742.x>
- Yu, G., Chen, X., Ni, J., Cheddadi, R., Guiot, J., Han, H., ... Zheng, Z. (2000). Palaeovegetation of China: A pollen data-based synthesis for the mid-Holocene and Last Glacial Maximum. *Journal of Biogeography*, 27, 635–664. <https://doi.org/10.1046/j.1365-2699.2000.00431.x>
- Zhang, R., Song, G., Qu, Y., Alström, P., Ramos, R., Xing, X., ... Lei, F. (2012). Comparative phylogeography of two widespread magpies: Importance of habitat preference and breeding behavior on genetic structure in China. *Molecular Phylogenetics and Evolution*, 65, 562–572. <https://doi.org/10.1046/j.1365-2699.2000.00431.x>
- Zhao, N., Dai, C., Wang, W., Zhang, R., Qu, Y., Song, G., ... Lei, F. (2012). Pleistocene climate changes shaped the divergence and demography of Asian populations of the great tit *Parus major*: Evidence from phylogeographic analysis and ecological niche models. *Journal of Avian Biology*, 43, 297–310. <https://doi.org/10.1111/j.1600-048x.2012.05474.x>
- Zink, R. M., Pavlova, A., Rohwer, S., & Drovetski, S. V. (2006). Barn swallows before barns: Population histories and intercontinental colonization. *Proceedings of the Royal Society B*, 273, 1245–1251. <https://doi.org/10.1098/rspb.2005.3414>
- Zink, R. M., & Vazquez-Miranda, H. (2019). Species limits and phylogenetic relationships of Darwin's finches remain unresolved: Potential consequences of a volatile ecological setting. *Systematic Biology*, 68, 347–357. <https://doi.org/10.1093/sysbio/syy073>

## BIOSKETCH

**Gang Song** is a researcher in the Ornithological Research Group, Institute of Zoology, Chinese Academy of Sciences. His major research interests are in biogeography, phylogeography and population genetics of Eurasian birds. He is involved in a series of phylogeographical studies on Eurasian passerines. The present paper is part of the results of a comparative phylogeographical study on Eurasian birds, which is a joint international cooperation research project led by F.L., Y.Q., P.A., J.F. and P.G.P.E. Website: [http://sourcedb.ioz.cas.cn/yw/scs/pi/200907/t20090716\\_2088440.html](http://sourcedb.ioz.cas.cn/yw/scs/pi/200907/t20090716_2088440.html).

**Author contributions:** F.L., Y.Q., P.A., J.F. and P.G.P.E designed the joint international research project. F.L. and G.S conceived the idea for this study. R.Z., G.S., P.A., H.L.M.J., B.D.M. and I.N collected samples, and R.Z., U.S.J, M.I., H.L.M.J., and B.D.M. completed the laboratory works. G.S., R.Z. and Y.C performed the molecular data analyses, and F.M.S. and A.T.P. conducted the ecological modelling analyses. G.S. developed the first draft. Y.Q., P.A., A.T.P., J.F. and F.L. discussed results and composed the discussion section of the study. All authors assisted in the interpretation of the results and commented on the manuscript.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Song G, Zhang R, Machado-Stredel F, et al. Great journey of Great Tits (*Parus major* group): Origin, diversification and historical demographics of a broadly distributed bird lineage. *J Biogeogr*. 2020;00:1–14. <https://doi.org/10.1111/jbi.13863>