ORIGINAL ARTICLE When east meets west: population structure of a high-latitude resident species, the boreal chickadee (*Poecile hudsonicus*)

LA Lait and TM Burg

The population genetic structure of northern boreal species has been strongly influenced both by the Quaternary glaciations and the presence of contemporary barriers, such as mountain ranges and rivers. We used a combination of mitochondrial DNA (mtDNA), nuclear microsatellites and spatial distribution modelling to study the population genetic structure of the boreal chickadee (*Poecile hudsonicus*), a resident passerine, and to investigate whether historical or contemporary barriers have influenced this northern species. MtDNA data showed evidence of eastern and western groups, with secondary admixture occurring in central Canada. This suggests that the boreal chickadee probably persisted in multiple glacial refugia, one in Beringia and at least one in the east. Palaeo-distribution modelling identified suitable habitat in Beringia (Alaska), Atlantic Canada and the southern United States, and correspond to divergence dates of 60–96 kya. Pairwise F_{ST} values for both mtDNA and microsatellites were significant for all comparisons involving Newfoundland, though mtDNA data suggest a more recent separation. Furthermore, unlike mtDNA data, nuclear data support population connectivity among the continental populations, possibly due to male-biased dispersal. Although both are significant, the isolation-by-distance signal is much stronger for mtDNA ($r^2 = 0.51$) than for microsatellites ($r^2 = 0.05$), supporting the hypothesis of male-biased dispersal. The population structure of the boreal chickadee was influenced by isolation in multiple refugia and contemporary barriers. In addition to geographical distance, physical barriers such as the Strait of Belle Isle and northern mountains in Alaska are restricting gene flow, whereas the Rocky Mountains in the west are a porous barrier. Heredity (2013) 111, 321-329; doi:10.1038/hdy.2013.54; published online 12 June 2013

Keywords: population genetic structure; glacial refugia; ecological niche modelling; mitochondrial DNA; microsatellites; boreal chickadee

INTRODUCTION

The major climate fluctuations of the Quaternary period, most recently the Pleistocene glaciations, played a major role in shaping today's populations (Hewitt, 2000; 2004). Temperate species were unable to survive in their pre-glaciation range and persisted instead in ice-free regions known as refugia. A combination of fossil evidence, pollen data, molecular data and sediment cores support the presence of multiple refugia during the last glacial maximum (LGM; the Wisconsin glaciation) in the mid-latitude United States and in Beringia (Pielou, 1991; Provan and Bennett, 2008; Shafer *et al.*, 2010b). A number of smaller refugia are purported to have existed along the periphery of the ice sheets. The now submerged Atlantic Shelf refugium in eastern North America (Pielou, 1991) probably supported numerous tree (Walter and Epperson, 2001; Jaramillo-Correa *et al.*, 2004) and bird (Gill *et al.*, 1993; Zink *et al.*, 2003) species.

The boreal chickadee (*Poecile hudsonicus*) is a small, resident, boreal songbird found throughout Canada and the northern United States (Figure 1; Ficken *et al.*, 1996). Banding records (Canadian Bird Banding Office) are sparse, with all recaptured birds found within 40 km from the initial banding site (n = 21; recaptures occurred 4–60 months after initial banding). The boreal chickadee specialises on soft

heartwood such as black and white spruce (*Picea mariana* and *Picea glauca*) and balsam fir (*Abies balsamea*) and is found in northern boreal forests (Ficken *et al.*, 1996). The current distribution was almost entirely glaciated during the LGM and includes both potential refugia: Beringia (Alaska), the Atlantic shelf and areas south of the ice sheets; and physical barriers: the Rocky Mountains, the Wrangell and Chugach Mountains (Alaska), the Cabot Strait (between Newfound-land and Nova Scotia) and the Strait of Belle Isle (between Newfoundland and Quebec/Labrador). Barriers such as water may restrict dispersal as boreal chickadees are absent from the large western islands (that is, Vancouver Island, Haida Gwaii and the Alexander Archipelago), and mountains may be restricting their western distribution as they are not found west of the Coast Range in British Columbia (Ficken *et al.*, 1996).

Gill *et al.* (1993) examined populations of North American chickadees using mitochondrial DNA (mtDNA) restriction fragment-length polymorphisms; in the boreal chickadee, they looked at 37 birds from seven populations. Little genetic differentiation was seen in the continental populations (n = 25), which roughly correspond to the previously glaciated region, and a unique haplotype was shared among all Atlantic Canada birds (n = 12). The present study

Department of Biological Sciences, University of Lethbridge, Lethbridge, Alberta, Canada

E-mail: linda.lait@uleth.ca

Received 6 August 2012; revised 16 April 2013; accepted 30 April 2013; published online 12 June 2013

Correspondence: LA Lait, Department of Biology, Memorial University of Newfoundland, Science Building, Elizabeth Avenue, St John's, Newfoundland and Labrador A1B 3X9, Canada.



Figure 1 The distribution of the boreal chickadee across North America (shaded), and the geographical distribution of the three clusters found in BAPS v5.2: eastern (black), western (white) and continent-wide (grey) groups. Sampling sites are as follows: AKA (Alaska Anchorage), AKF (Alaska Fairbanks), AKW (Alaska Wrangell St Elias), NBC (northern British Columbia), CBC (central British Columbia), CAB (central Alberta), SAB (southern Alberta), SK (Saskatchewan), NON (northern Ontario), NQC (northern Quebec), NY (New York), NSNB (Nova Scotia and New Brunswick), NL (Newfoundland), and LAB (Labrador). Physical barriers are shown by triangles (mountains) and dashes (bodies of water). Distribution modified from Birds of North America online (Ficken *et al.*, 1996).

aims to expand upon these findings using an intensive, range-wide sampling regime and more sensitive mtDNA sequencing and microsatellite analyses in combination with spatio-geographic modelling. The combination of nuclear and mitochondrial markers will allow for a more complete picture of the chickadee's history than is currently available and avoids the limitations of single gene inferences. Population genetic and phylogeographic studies are increasingly employing niche modelling as an additional tool to assess historical areas of suitable habitat (for example, Graham and Burg, 2012; Ralston and Kirchman, 2012; van Els *et al.*, 2012).

Although a number of studies have examined the genetic structure of boreal forest trees (Walter and Epperson, 2001; Jaramillo-Correa *et al.*, 2004; Anderson *et al.*, 2006), few have looked at their avian inhabitants, especially at the continental scale (Ralston and Kirchman, 2012; van Els *et al.*, 2012). This study aims to assess the population genetic structure of the boreal chickadee across its range and to understand what role the Pleistocene glaciations and contemporary physical barriers have in this wholly boreal species. For this study we combined both genetic analyses and spatial modelling to ask:

- (1) What is the population genetic structure of the boreal chickadee?
- (2) Did the boreal chickadee survive in a single or multiple glacial refugia, and can the location of the refugia be identified? As the current distribution of boreal chickadees is widespread and includes both Alaska (Beringia) and Nova Scotia and Newfoundland (the putative Atlantic Shelf refugium), we predict that the populations expanded from multiple glacial refugia.
- (3) Do contemporary physical barriers impede dispersal and gene flow in the boreal chickadee? Mountains and large expanses of water have been shown to prevent dispersal in many species (for example, Cronin *et al.*, 2005; Hearn *et al.*, 2006), and different

Heredity

subspecies of boreal chickadees are found on either side of several barriers (for example, Rocky Mountains and Strait of Belle Isle; see Ficken *et al.*, 1996). We predict reduced gene flow across these physical barriers.

MATERIALS AND METHODS Sample collection

Two hundred and thirty blood samples were collected from 11 sampling locations over four breeding seasons (2007–2010; Figure 1, Supplementary Table S1). Sampling locations, hereafter referred to as populations, were limited to a 50-km radius, where possible, with no obvious barriers to dispersal within a sampling site. Fifty-three museum samples (tissue and toe-pads) augmented sample sizes from field sites and added an additional three sampling locations (Figure 1, Supplementary Tables S1 and S2). DNA was extracted using a modified chelex procedure (Walsh *et al.*, 1991).

Lab protocols

MtDNA. Two mtDNA fragments were amplified: the control region partial domain I and II (CR) and the ATPase coding region (ATP), which contains ATPase6, ATPase 8 and the tRNA lysine. The control region was amplified using the primers LmochCR1/H1015chCR (Lait *et al.*, 2012). Some museum samples (toe-pads) required a semi-nested PCR using the primers L26chCR/H1015chCR, followed by LmochCR2/H1015chCR (Supplementary Table S3). The ATP gene fragment was amplified using universal avian primers L8929COII/H9855ATP6 (Sorenson *et al.*, 1999) and toe-pad samples were amplified in two shorter, overlapping fragments using H534chATP/L8929COII and L298chATP/H9855ATP6 (Supplementary Table S3). Samples were sequenced on a 3130 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA), and museum samples were sent to Genome Quebec for sequencing (McGill University, Montreal, QC, Canada).

Microsatellites. A small number of individuals from geographically distant populations (AKA and NL) were screened using 20 microsatellite primer pairs developed in a number of other passerines with a M13-tailed forward primer (see Burg et al., 2005). Of those that yielded PCR products, five loci were monomorphic (BT6, Escu1, Escu3, Pocc2 and Titgata88; Hanotte et al., 1994; Bensch et al., 1997; Otter et al., 2001; Wang et al., 2005), and eight were polymorphic (Escu4, Escu6, Pat14, Pat43, Pdo5, Ppi2, Titgata02 and Titgata39; Hanotte et al., 1994; Otter et al., 1998; Griffith et al., 1999; Martinez et al., 1999; Wang et al., 2005; Supplementary Table S3). The PCR used a two-step annealing process with eight cycles at T1 and 31 cycles at T2. For four of the loci (Escu6, Pat14, Ppi2 and Titgata39), the second step was reduced from 31 to 25 cycles (Supplementary Table S3). The PCR products were run on a 6% polyacrylamide gel using a LI-COR 4300 DNA Analyzer (LI-COR Inc., Lincoln, NE, USA). Alleles were scored manually by visual inspection and were independently confirmed by a second person. Four controls of known size were included on each gel, and a 'mystery plate' of 44 individuals was run as an additional test for consistency.

Genetic analyses

MtDNA. All analyses were done on a concatenated sequence of the CR and ATP fragments and on the two fragments separately; the results were similar and only the concatenated results are discussed further. The sequences were aligned in MEGA v5.05 (Tamura *et al.*, 2011), haplotypes were assigned manually and confirmed using TCS v1.21 (Clement *et al.*, 2000). Nucleotide and haplotype diversity were calculated in DNASP v5 (Librado and Rozas, 2009). Pairwise genetic differences (*p* distances; Φ_{ST} ; 10000 permutations) were calculated in ARLEQUIN v3.11 (Excoffier *et al.*, 2005), and a modified false discovery rate (FDR) correction (Benjamini and Yekutieli, 2001) was applied. A Mantel's test was performed in GENEPOP v4.0.10 to test for isolation-by-distance (Raymond and Rousset, 1995; Rousset, 2008) using geographical distances calculated in GEOGRAPHIC DISTANCE MATRIX GENERATOR v1.2.3 (Ersts, 2010) and linearised Φ_{ST} values. Significance was tested using 10 000 permutations.

A statistical parsimony network was constructed in TCS v1.21 (Clement *et al.*, 2000) with a 95% connection limit. A spatial analysis of molecular variance (SAMOVA) was run to identify genetic clusters independent of sampling locations using both geographical and genetic data (K=2-13; 100 iterations; Dupanloup *et al.*, 2002). To test if any structure detected was due to historical separation in refugia or to contemporary barriers, a series of hierarchical AMOVAs (10 000 permutations) were run in ARLEQUIN v3.11 (Excoffier *et al.*, 1992, 2005). The scenarios tested included two or three refugia (east and west; east, central and west; and east, southern (SAB) and west), one, two or three contemporary barriers (mountain ranges and water barriers as shown in Figure 1) and combinations of refugia and barriers.

A principal coordinates analysis (PCoA) was performed in GENALEX v6.3 (Peakall and Smouse, 2006) on both individuals and population pairwise Φ_{ST} values. BAPS (Bayesian Analysis of Population Structure) v5.2 (Corander *et al.*, 2008) was used to assign individuals to *K* clusters based on genetic data, with no *a priori* population information. As per Corander *et al.* (2008), the analyses used the clustering with linked loci option (Corander and Tang, 2007) and varied *K* (K=1–14).

Approximate divergence dates and relative migration rate estimates were calculated using the isolation with migration model as implemented in IMA v3.5 (Hey and Nielsen, 2007). IMA allows two populations to be compared; as such, the eastern and western groups identified by SAMOVA were tested. Multiple initial runs were performed to identify upper parameter bounds and an ideal heating scheme. The three final runs were performed using the HKY model, 500 000 burn-in and 1 000 000 post burn-in MCMC (Markov Chain Monte Carlo) steps, geometric heating, different random number seeds and a weighted average divergence rate of 1.7% per Myr (see below) with a low (1.0% per Myr) and a high (5.0% Myr) rate.

The control region and ATP divergence rates were estimated using an average cytochrome b (cyt b) rate of 2.1% per Myr (Weir and Schluter, 2008) and a combination of *Poecile* cytb, ATP and CR sequences (Uimaniemi *et al.*, 2003; Gill *et al.*, 2005; Lait *et al.*, 2012). The calculated divergence rate was 1.78% per Myr for ATP, 2.62% per Myr for control region domain I and 0.95% per Myr for domain II. Although these values are considerably lower than traditionally

considered for the control region, similar rates have recently been calibrated in a number of parids (Kvist *et al.*, 2001; Päckert *et al.*, 2007) and other birds (Zink *et al.*, 2003; Lerner *et al.*, 2011). Although using sequences to calculate divergence times is not as ideal as using fossil calibrations, the calculations do show that cyt *b*, ATP and the control region have similar divergence rates in *Poecile*.

Microsatellites. Only individuals with genotypes from at least six loci were included in the analyses with the exception of museum samples from NON, NQC and NY where samples with up to four (50%) missing loci were included in some analyses. The SAB population was not genotyped. MICRO-CHECKER v2.2.3 was used to detect input errors, allelic dropout, slippage stutter or null alleles (Van Oosterhout *et al.*, 2004). Exact tests were run in GENEPOP v4.0.10 (Raymond and Rousset, 1995; Rousset, 2008) to check for linkage disequilibrium and deviations from Hardy–Weinberg equilibrium using modified Markov chain parameters (1000 batches, 10 000 iterations and 10 000 dememorisation steps). All *P*-values were corrected for multiple tests using the modified FDR method (Benjamini and Yekutieli, 2001).

Population genetic diversity was compared using expected heterozygosity and allelic richness. Observed ($H_{\rm O}$) and expected ($H_{\rm E}$) heterozygosities were calculated in ARLEQUIN v3.11 (Excoffier *et al.*, 2005). Allelic richness ($A_{\rm R}$) and private allelic richness ($PA_{\rm R}$) were calculated in HP-RARE v1.1 using rarefaction (Kalinowski, 2005). Both global and population pairwise $F_{\rm ST}$ and $R_{\rm ST}$ values were calculated in ARLEQUIN v3.11 (100 000 permutations; Excoffier *et al.*, 2005); Jost's $D_{\rm est}$ (Jost, 2008) was calculated in GENALEX v6.5b3 (Peakall and Smouse, 2012). All comparisons involving NON and NQC were run using four and five loci, respectively (excluding Escu6, Ppi2 and Pdo5 in both, as well as Escu4 in NON), and all other populations were analysed using both six (excluding Ppi2 and Pdo5; see Results) and eight loci. A modified FDR correction was applied (Benjamini and Yekutieli, 2001).

Clustering analyses were run to delineate groupings based on individual genotypes using STRUCTURE V2.3 (Pritchard *et al.*, 2000). STRUCTURE was run using uncorrelated allele frequencies and the admixture model, both with and without sampling locations as priors. The program was run from K=1-13 (10 runs each) for 100 000 burn-in and 300 000 post burn-in MCMC steps. The results of the 10 runs were averaged using STRUCTURE HARVESTER v0.6.6 (Earl and vonHoldt, 2012), and the most probable value of *K* was calculated using both the highest penalised log likelihood (Pritchard *et al.*, 2000), and ΔK (Evanno *et al.*, 2005). All runs were repeated for the 12 mainland populations excluding NL to test for additional substructure (see Results). In order to visualise the pattern of population genetic structure, a PCoA on population summary statistics ($F_{\rm ST}$) was performed in GENALEX v6.3 (Peakall and Smouse, 2006), with significance tested in PCA-GEN v1.2.1 (Goudet, 1999).

Ecological niche modelling

Both contemporary (ecological niche modelling) and historical (palaeodistributional modelling) patterns of species distribution were estimated in MAXENT V3.3.3 using a maximum entropy statistical model on presence-only occurrence data (Phillips *et al.*, 2006; Phillips and Dudík, 2008). The model estimates the potential distribution of a species by assuming that the species is found in its preferred environmental conditions and that the current niche requirements are reflective of those in the past and/or future (that is, niche fidelity; Phillips *et al.*, 2006; Richards *et al.*, 2007).

Occurrence records were comprised of sampling locations from this study (n = 279) and sightings downloaded from the Global Biodiversity Information Facility data portal (n = 9624; GBIF data portal, 2011) and the Avian Knowledge Network (n = 22252; Avian Knowledge Network, 2009). The models were based on 19 WorldClim climatic variables (see Supplementary Table S4; Hijmans *et al.*, 2005) extrapolated from GIS (Geographical Information Systems) layers as described in Carstens *et al.* (2007). The variables included averages, extremes and ranges of temperature and precipitation and were available for both the contemporary timescale and estimates for the LGM (ca. 21 kya). The Model for Interdisciplinary Research on Climate (MIROC) climate layers used as the past climate estimates were provided by the Palaeoclimate Modelling Intercomparison Project Phase II (Braconnot *et al.*, 2007). Ten replicates were run using the cross-validation method.

324

RESULTS

MtDNA

Two fragments of mtDNA were successfully amplified for 281 samples from 14 populations (GenBank accession numbers: JN654564-JN654638 (control region); JN654639-JN654699 (ATP)). A 766-bp fragment of the control region contained 56 variable sites, almost two-thirds of which were found in domain I (34 of the 56). A 923-bp fragment of coding DNA (ATP) contained 53 variable sites with 32 synonymous and 18 non-synonymous substitutions (and three in the tRNA). There were no unexpected stop codons. Both the CR and ATP fragments contained fixed differences between eastern (LAB, NL, NSNB and NY) and western (AKA, AKF, AKW, NBC, CBC, CAB and SAB) populations (with few exceptions). The concatenated fragment contained 105 transitions, three transversions and one insertion/ deletion (in the control region). No nucleotide position had more than two base variants.

One hundred and twenty-seven haplotypes were identified, 26 shared between two or more individuals and 101 unique (Supplementary Table S5). No haplotypes were shared among the most eastern populations (LAB and NL) and western populations (AK, BC and AB). Haplotype diversity was high in all populations (0.600-1.000), with the highest values seen in AKF (0.986), SAB (1.000), NON (0.990) and LAB (0.935), and the lowest in NY (0.600) and NL (0.800; Supplementary Table S5). Nucleotide diversity ranged from 0.001 to 0.003 (Supplementary Table S5). Although both haplotype and nucleotide diversities were lower in NL, it had the second highest number of private haplotypes. Pairwise Φ_{ST} values (Table 1) showed NL to be significantly different from all other populations, and AKA and AKW were significantly different from all but each other. Following the modified FDR correction ($P_{crit} = 0.01$), 70 of 91 pairwise differences were significant, with most nonsignificant values found between neighbouring populations. A significant correlation between genetic and geographical distance (isolation-by-distance) was found ($r^2 = 0.51$; P < 0.001).

The statistical parsimony network showed very little overlap among the eastern (haplotypes P, R–Z) and western (haplotypes A–O) populations, whereas the central populations (SK, NON and NQC) were present throughout the network (Figure 2). Similar patterns were seen when the ATP and CR fragments were analysed separately

(not shown). The SAMOVA detected the presence of two groups $(\Phi_{\rm CT} = 0.305, P < 0.001)$: a western group (AKA, AKF, AKW, NBC, CBC, CAB, SAB, SK and NON) and an eastern group (NQC, NSNB, LAB, NY and NL). In the hierarchical AMOVA, the largest Φ_{CT} value was obtained when populations were clustered into the eastern and western groups as above ($\Phi_{\rm CT} = 0.307$, P<0.001), with the next highest when SAB was included as a third group ($\Phi_{\rm CT} = 0.299$, P < 0.001; Supplementary Table S6). Separation of the samples into the eastern and western groups plus NL as a third group was also supported ($\Phi_{\rm CT} = 0.281$, P<0.001). The pattern could not be explained by contemporary barriers alone (highest $\Phi_{\rm CT} = 0.194$, P = 0.008; Supplementary Table S6). Both the PCoA on the individual samples (not shown) and pairwise $\Phi_{\rm ST}$ values (Figure 3a) supported the separation of the eastern and western populations, with central populations (SK, NON and NQC) falling intermediately. BAPS separated the samples into three clusters: one primarily in the east (CAB, SK, NON, NQC, NSNB, LAB, NY and NL), one in the west (AKA, AKF, AKW, NBC, CBC, CAB, SK and NON), and a continentwide group (all populations; see Figures 1 and 2).

Approximate divergence times were calculated between the western (AKA, AKF, AKW, NBC, CBC, CAB, SAB, SK and NON) and eastern (LAB, NL, NSNB, NY and NQC) groups identified by SAMOVA. Marginal distribution curves were unimodal and approached 0 on each side. The mean separation date between the east and west was 78.6 kya (60.4–96.0 kya 90% HPD). The high (5% per Myr) and low (1% per Myr) rates gave dates of 26.7 kya (20.5–32.6 kya 90% HPD) and 133.3 kya (102.3–162.8 kya 90% HPD), respectively. Divergence estimates increased when the central (SK, NON and NQC) populations were excluded (104.3 kya; 65.6–141.8 kya 90% HPD). Gene flow (west to east = 0.0021; east to west = 0.0033 migrations per 1000 generations per gene copy); estimates dropped considerably when the three central populations were removed from the analysis (west to east = 0.0012; east to west = 0.0006).

Microsatellites

After excluding individuals due to missing data, 260 individuals from 13 populations were successfully genotyped. Genotyping of Pdo5 and Ppi2 was not possible for museum samples from NON, NQC and

Table 1 Population pairwise differences based on 100 000 permutations for mtDNA (Φ_{ST} , above diagonal) and microsatellites (F_{ST} , below diagonal)

	AKA	AKF	AKW	NBC	CBC	CAB	SAB	SK	NON	NQC	NSNB	LAB	NY	NL
AKA	*	0.108**	0.021	0.157**	0.292**	0.238**	0.250*	0.164**	0.118**	0.271**	0.391**	0.361**	0.469**	0.506**
AKF	0.002	*	0.078*	0.020	0.101*	0.066	0.051	0.077	0.121**	0.238**	0.378**	0.348**	0.436**	0.489**
AKW	0.000	0.000	*	0.123**	0.237**	0.179*	0.212*	0.131*	0.101**	0.241**	0.361**	0.335**	0.417**	0.491**
NBC	0.015*	0.012*	0.008	*	0.222**	0.185**	0.090	0.158**	0.154**	0.298**	0.418**	0.391**	0.493**	0.563**
CBC	0.007	0.000	0.005	0.014	*	0.000	0.235	0.129	0.252**	0.408**	0.494**	0.483**	0.545**	0.631**
CAB	0.000	0.000	0.000	0.000	0.000	*	0.194	0.045	0.166*	0.320**	0.427**	0.410**	0.462**	0.579**
SAB	NA	NA	NA	NA	NA	NA	*	0.186	0.153*	0.342*	0.450**	0.437**	0.512*	0.645**
SK	0.017*	0.003	0.007	0.010	0.011	0.000	NA	*	0.035	0.101	0.228**	0.193**	0.295*	0.349**
NON	0.022	0.008	0.011	0.021	0.019	0.000	NA	0.003	*	0.018	0.141**	0.111*	0.183*	0.255**
NQC	0.000	0.000	0.000	0.019	0.021	0.000	NA	0.003	NA	*	0.047	0.000	0.153	0.117**
NSNB	0.007	0.007	0.003	0.016*	0.021*	0.007	NA	0.008	0.016	0.001	*	0.000	0.030	0.130**
LAB	0.009	0.005	0.002	0.018*	0.020	0.006	NA	0.012	0.000	0.006	0.005	*	0.070	0.090**
NY	0.022	0.020	0.009	0.038	0.015	0.015	NA	0.037	0.002	0.042	0.017	0.000	*	0.327**
NL	0.043**	0.034**	0.040**	0.046**	0.046**	0.039**	NA	0.030**	0.030*	0.023*	0.025**	0.049**	0.074**	*

Abbreviation: mtDNA, mitochondrial DNA.

For values are based on data from six loci for 11 populations, five loci for NQC and four for NON. Significant values after modified false discovery rate correction are shown in bold ($P_{crit} = 0.01$); *P < 0.01, *P < 0.001. All negative values were treated as zero.



Figure 2 Statistical parsimony network of the concatenated mtDNA sequences. Each square represents an individual, the groups of squares (A–Z) represent shared haplotypes, each individual is colour-coded by population and the circles represent inferred haplotypes. The inset (top right) shows the network with only the two putative refugia (AKF and NL) coloured. The haplotypes which make up the BAPS clusters (Figure 1) are shown by the dashed (western—white), grey (continental—grey) and black (eastern—black) outlines.

many from NY, most probably due to the degraded nature of the toe pad samples. All other populations had little or no missing data. MICRO-CHECKER v2.2.3 suggested the presence of null alleles in at least one of Pdo5 or Ppi2 for 7 of the 10 populations tested. Exact tests showed departures from Hardy–Weinberg equilibrium, after correction for multiple tests ($P_{\rm crit}$ = 0.01), at two loci: Pdo5 (AKA, AKF, NL; P < 0.01) and Ppi2 (AKA, AKW, CBC, NSNB, NL; P < 0.01). When testing for linkage, only 4 of 338 tests were significant ($P_{\rm crit}$ = 0.008; P < 0.001). Pdo5 and Ppi2 were excluded from final analyses.

Observed heterozygosity ($H_{\rm O}$) was similar to expected heterozygosity ($H_{\rm E}$) across all populations for five of the six loci (Supplementary Table S7). Pat14 had significantly lower heterozygosity than expected (P<0.01). Observed heterozygosity ranged from 0.400 to 1.000 across all loci. Allelic richness was similar across populations (3.01–3.37, average 3.20; Supplementary Table S7), with private allelic richness highest in NSNB (0.51) and NL (0.50) relative to an average of 0.34.

Global F_{ST} , based on six loci, was 0.015 and highly significant (P < 0.001). Population pairwise F_{ST} values showed significant differences between NL and all other populations (Table 1); otherwise, pairwise values were generally not significant (only 6 of the 66 non-NL comparisons were significant after modified FDR correction). A similar pattern was seen with both R_{ST} values and Jost's D_{est} values (Supplementary Table S8) and when F_{ST} was calculated using all eight loci (global $F_{ST} = 0.013$, P < 0.001). To ensure that the missing data in NON, NQC and the small sample size of NY were not affecting the overall results, pairwise comparisons were also run excluding these populations. Results were similar and are not shown. A weak, but significant, isolation-by-distance pattern was present in the nuclear data ($r^2 = 0.05$, P = 0.002).

STRUCTURE identified two clusters: NL and all other populations (Figure 4). The two groups were supported by both log likelihood penalised tests (Bayes factor = 1.00, ln Pr (X | K) = -7162) and ΔK . All *Q* values (ancestry coefficients) were >0.50, and 95% of individuals were assigned with *Q*>0.75, the exceptions being mostly NON samples (Figure 4). No additional substructure was seen when NL was removed from the analysis (not shown). The PCoA (Figure 3b) showed little genetic structure among the populations with the exception of NL. The broken-stick method employed in PCA-GEN v1.2.1 (Goudet, 1999) found only coordinate 1 to be significant (*P*<0.05).

Ecological niche modelling

The maximum entropy model performed significantly better than random, as shown by both the binomial test of omission (tested rate close to predicted rate; P < 0.001) and the receiver operating characteristic analysis (area under the curve = 0.829 ± 0.004). The contribution of each of the environmental layers varied considerably (Supplementary Table S4). The present distribution predicted by the model (Figure 5a) resembles the current range of the boreal chickadee (Figure 1), with the northern extent well predicted and the southern boundary moved to the south. The predicted LGM distribution (ca. 21 kya; Figure 5b) shows that suitable habitat was probably present in Alaska, along the western coast of North America, in the southern United States and in Atlantic Canada.

DISCUSSION

Significant population differentiation was found in boreal chickadees. At the level of populations, most of the pairwise $\Phi_{\rm ST}$ values were significant, suggesting restricted maternal gene flow (Table 1). The non-significant comparisons involved adjacent populations and

support an isolation-by-distance pattern whereby gene flow is restricted by geographical distance. By contrast, nuclear data support the genetic isolation of NL from mainland populations; however, few of the other populations are significantly differentiated from one another (Table 1, Figures 3 and 4). Uncovering different patterns with mtDNA and microsatellite data is not uncommon in phylogeographic and population genetic studies (see Zink and Barrowclough, 2008). Several factors could explain the dissimilar patterns between the two types of molecular markers, including the age of the population and mode of inheritance (female vs biparental inheritance).



Coordinate 1 (42.5%)

Figure 3 PCoA based on the population pairwise (a) Φ_{ST} values and (b) F_{ST} values. With the mtDNA, most of the variation was explained by the first coordinate (74.0% coordinate 1, 10.0% coordinate 2 and 8.8% coordinate 3). With the microsatellite data, the partition of NL from the rest of the populations was along coordinate 1 (42.5%), while coordinate 2 (24.8%) separated the eastern and western populations. Populations are colour-coded as per Figure 2.

Several species show evidence of sex-biased dispersal (Gibbs et al., 2000; Milot et al., 2000; Helbig et al., 2001; Austin et al., 2003; Bensch et al., 2006), including the yellow warbler which, while a long distance migrant, showed similar patterns of population differentiation, at comparable sampling sites, to the boreal chickadee (Gibbs et al., 2000; Milot et al., 2000). Yellow warbler populations also showed higher levels of population differentiation at mtDNA genes than microsatellite loci. Gibbs et al. (2000) attributed the patterns to male-biased dispersal but also examined the possibility that populations may be too voung to be in equilibrium and therefore differences between markers are due to recent ancestry. Hutchison and Templeton (1999) evaluated the relative influence of drift and migration on population differentiation and concluded that once isolated populations are in migration-drift equilibrium they may show isolation-by-distance. If either genetic drift or gene flow is higher relative to the other, then at large geographical distances, no isolation-by-distance pattern will be



Figure 5 The predicted distribution of the boreal chickadee using maximum entropy in MAXENT v3.3.3 for the (a) present and (b) LGM. The darker shading represents the most probable locations. S.d. was low for both timeframes.



Figure 4 Bayesian clustering analysis run in STRUCTURE v2.3 with six microsatellite loci for K=2. Each vertical line represents an individual, and the y axis denotes the cluster membership (Q).

326

present. Boreal chickadee populations show significant isolation-bydistance patterns at both nuclear and organellar markers, suggesting the populations are in equilibrium, leaving the possibility of malebiased dispersal to explain the observed patterns. Higher gene flow in males could also explain the disappearance of an east/west split detected in the mainland populations with mtDNA for nuclear data (Table 1, Figures 3 and 4). The significant differentiation of NL from the mainland may be due to other factors such as the presence of physical barriers (see below) and therefore affects both male- and female-mediated dispersal.

The mtDNA data (haplotype distribution, high Φ_{ST} values, hierarchical AMOVA, significant SAMOVA, IMA and PCoA) support the separation of the eastern (Atlantic Canada and NY) and western (AK, BC and AB) groups of boreal chickadees with a central cline. The gradient seen is probably the result of a secondary contact zone with admixture of multiple lineages (Taberlet *et al.*, 1998; Petit *et al.*, 2003). Although boreal tree species commonly show evidence of a phylogenetic break in the Great Lakes region similar to that found in the boreal chickadee (Jaramillo-Correa *et al.*, 2009), it is less commonly seen in avian species. North American birds tend to show genetic breaks further west across the Rocky Mountains between British Columbia and Alberta (Burg *et al.*, 2005; Toews and Irwin, 2008) or on the east coast (Gill *et al.*, 1993; Holder *et al.*, 1999). The mtDNA pattern seen in the boreal chickadee more closely resembles that of the black and white spruce, two species with which it is closely associated.

Impact of glaciation

Pleistocene glaciations strongly influenced the population structure and distribution of many temperate species. In the case of the boreal chickadee, palaeo-modelling indicates suitable habitat was present in multiple refugia (Figure 5b) and divergence estimates suggest the eastern and western mtDNA lineages diverged during the last glaciation. Specifically, isolation with migration estimates (60–96 kya) place the divergence of the two lineages within the Wisconsin glaciation. When the ice sheets formed over 100 kya (Gillespie *et al.*, 2004), the physical separation of the boreal chickadee populations occurred, thereby initiating genetic divergence in this species.

The location of suitable boreal chickadee habitat at the LGM in Beringia, the west coast of North America, the southern United States and Newfoundland (Figure 5b) is consistent with putative refugia described in other species (Chappell et al., 2004; Jaramillo-Correa et al., 2004; Colbeck et al., 2008; de Lafontaine et al., 2010; Graham and Burg, 2012; Pulgarín-R and Burg, 2012). Patterns of genetic diversity in the boreal chickadee were as expected if multiple refugial populations merged following prolonged isolation, with higher diversity and private haplotypes in or near putative refugia (AKF and Atlantic Canada; Supplementary Table S5) and evidence of secondary mixing in the centre portion of their range (that is, SK, NON and NQC; Taberlet et al., 1998; Petit et al., 2003). Unlike populations in the west and east, which contained only western or eastern haplotypes, respectively, central populations contained a subset of both haplotypes (Figure 2, Supplementary Table S5). If boreal chickadees survived in a single central refugium and populations expanded outwards, levels of genetic diversity would be higher in the centre of the range and would decrease with increasing distance (for example, diversity would be lower in AKF and NL). Newly colonised populations would contain a subset of haplotypes found in the central area, a pattern seen in blackpoll warblers (Ralston and Kirchman, 2012). Several lines of evidence support the presence of multiple Pleistocene refugia, namely Alaska and the Atlantic Coast, for the boreal chickadee.

AKF, a population in the area of Beringia, contains high levels of mtDNA diversity, a large number of private haplotypes and almost all shared haplotypes in the west are present in AKF (Figure 2, Supplementary Table S5). Although the majority of species thought to have persisted in Beringia are either restricted to the Northwest or Arctic regions (for example, Holder *et al.*, 1999), evidence supports the use of this refugium by several widespread species, including the wolverine (Chappell *et al.*, 2004) and red fox (Aubry *et al.*, 2009), as well as black spruce and white spruce, two boreal trees upon which the boreal chickadee relies heavily (Jaramillo-Correa *et al.*, 2004; Anderson *et al.*, 2006; Gerardi *et al.*, 2010).

Suitable habitat also existed east of the ice sheets in north-eastern North America (Figure 5b). The persistence of the boreal chickadee in an Atlantic Shelf refugium is concordant with patterns seen in a number of plants (Walter and Epperson, 2001; Jaramillo-Correa et al., 2004; de Lafontaine et al., 2010) and birds (Zink and Dittmann, 1993; Zink et al., 2003; Colbeck et al., 2008; van Els et al., 2012). Although both mtDNA and microsatellite analyses support the separation of NL as a distinct population (Table 1, Figure 4), it is unlikely to have been the sole refugium on the east coast. In contrast to AKF, NL has lower nucleotide and haplotype diversities (Supplementary Table S5); therefore, while we cannot exclude the possibility that individuals from NL may have emigrated following the retreat of the Laurentide ice sheet, it is unlikely that the island alone acted as a source population. Boreal chickadees probably persisted on the Atlantic shelf or the north-eastern edge of the United States and, as the ice sheets melted, birds on NL became isolated and have remained so. MtDNA patterns in multiple boreal chickadee populations from Atlantic Canada, namely NSNB, LAB and NL, are consistent with refugial populations and it is probable that all served as, or were located near, the eastern refugia. For example, NSNB and LAB both have high genetic diversity, NSNB and NL populations contain a large number of private alleles and NL has a large number of private haplotypes.

Following the retreat of the ice sheets, individuals would have expanded into the previously glaciated regions. For the boreal chickadee, this includes areas of central Canada. Populations in these areas contain a mixture of haplotypes found in Alaska and Atlantic Canada, suggestive of secondary contact. A similar pattern exists in the yellow warbler with western (BC and AK) and eastern (MB to NL) haplotypes occurring in the Great Plains and Canadian Prairies (Milot *et al.*, 2000).

Physical barriers

A number of physical barriers restrict dispersal in boreal chickadees, including mountains, large bodies of water and geographical distance. The Rocky Mountains are often the site of phylogenetic breaks (Toews and Irwin, 2008) or suture zones (Brelsford and Irwin, 2009; Flockhart and Wiebe, 2009) in North American species. However, they do not appear to restrict gene flow in boreal chickadees. Chickadee populations on either side of the Rocky Mountains (for example, CAB and CBC, Table 1) are not genetically differentiated from one another. The prevalence of suitable habitat and treed corridors through the multitude of valleys and passes of the Northern Rockies prevent genetic isolation. However, further north, the Alaskan mountain ranges may be restricting gene flow (for example, between AKA and AKF, Table 1).

On the east coast, both the Cabot Strait and the Strait of Belle Isle are acting as barriers to gene flow in the boreal chickadee, isolating NL from the mainland populations. As with other island populations, Newfoundland hosts a number of endemic species and subspecies (for example, Cronin *et al.*, 2005; Hearn *et al.*, 2006), and genetically distinct populations are found in a number of species, including birds (Gill et al., 1993; Zink and Dittmann, 1993), mammals (Cronin et al., 2005) and plants (Gerardi et al., 2010). A combination of the physical water barrier and inhospitable conditions (harsh weather and lack of suitable habitat) probably contribute to the maintained genetic isolation of boreal chickadees.

The third type of physical barrier restricting gene flow in boreal chickadees is geographical distance. Both mtDNA and microsatellite data show significant isolation-by-distance, and short distance (<40 km) dispersal is supported by banding records. Reduced dispersal is consistent with other non-migratory, resident species, including both birds (Burg et al., 2005; Graham and Burg, 2012; Lait et al., 2012) and mammals (Chappell et al., 2004; Aubry et al., 2009; Shafer et al., 2010a).

Comparing genetic data to morphological taxonomy

Three to five subspecies of boreal chickadees have been described based on morphological and plumage differences (Ridgway, 1904; American Ornithologists' Union, 1957). The most widely accepted designation includes five described subspecies: P. h. hudsonicus (AK to ON), P. h. columbianus (west of Rocky Mountains), P. h. cascadensis (northern Cascade Mountains), P. h. littoralis (QC to Atlantic coast) and P. h. rabbittsi (NL) (American Ornithologists' Union, 1957). Our genetic data do not support the separation of populations on the western and eastern edges of the Rocky Mountains (P. h. columbianus and P. h. hudsonicus) but do support both the separation of populations west and east of Hudson's Bay (P. h. hudsonicus and P. h. littoralis) and a distinct Newfoundland subspecies (P. h. rabbittsi). No samples were collected from P. h. cascadensis.

CONCLUSIONS

The genetic patterns seen in the boreal chickadee support the historical separation in Beringia and Atlantic coastal refugia with subsequent expansion into central Canada. Colonisation of previously glaciated areas probably followed the spread of spruce species as the patterns are congruent with black and white spruce, supporting niche conservatism in this songbird. The close association between these species suggests that they may be useful bioindicators for each other. A combination of mountains, water barriers and physical distance are influencing dispersal among contemporary chickadee populations. Within the mainland populations, male-biased dispersal is acting to homogenise populations while mtDNA differences are retained. The exception to this is Newfoundland where both sets of molecular markers support genetic isolation of this island population. The various impacts of physical barriers in different areas (for example, mountains in Alaska vs western Canada) highlights the importance of including matrix quality as well as habitat features when looking at dispersal barriers.

DATA ARCHIVING

Sequence data have been submitted to GenBank: JN654564-JN654638 (control region); JN654639-JN654699 (ATP).

Genotype data have been submitted to Dryad: doi:10.5061/ dryad.82hs7. Data files: Microsatellite raw data of six loci for boreal chickadees across North America.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank everyone who assisted with sample collection: Heather Bird, Kim Dohms, Natalie Freeman, Brendan Graham, Heather Hansen, John Hindley, Elysha Koran, Christine Mishra, and Paulo Pulgarin-Restrepo, as well as both

Provincial and National Parks. We also thank everyone who helped with the analyses, especially Christin Pruett and Scott Taylor for their help with IMa. We acknowledge the hard work of all of the collaborators involved in the Avian Knowledge Network, Global Biodiversity Information Facility (GBIF) data portal and the Palaeoclimate Modelling Intercomparison Project Phase II (PMIP2). A special thanks to all of the museums who kindly provided tissue samples: American Natural History Museum, Burke Museum of Natural History and Culture, Canadian Museum of Nature, New Brunswick Museum, New York State Museum, Royal Alberta Museum, Royal British Columbia Museum, Royal Ontario Museum, and the Smithsonian Institution National Museum of Natural History. This project was supported by the Natural Science and Engineering Research Council (TMB). Funding for the laboratory and field components was provided by NSERC (Discovery Grant to TMB, Alexander Graham Bell Canada Graduate Scholarship to LAL), Alberta Innovates Technology Futures (TMB, Graduate Scholarship to LAL) and Alberta Scholarships Program (Ralph Steinhauer Award to LAL). Last but not least, we thank the reviewers and editor for their constructive comments.

- Anderson LL, Hu FS, Nelson DM, Petit RJ, Paige KN (2006). Ice-age endurance: DNA evidence of a white spruce refugium in Alaska. Proc Natl Acad Sci USA 103: 12447-12450.
- Aubry KB, Statham MJ, Sacks BN, Perrine JD, Wisely SM (2009). Phylogeography of the North American red fox: vicariance in Pleistocene forest refugia. Mol Ecol 18: 2668-2686
- Austin JD, Dávila JA, Lougheed SC, Boag PT (2003). Genetic evidence for female-biased dispersal in the bullfrog, Rana catesbeiana (Ranidae). Mol Ecol 12: 3165-3172.
- Avian Knowledge Network (2009). Ithaca, New York, USA. www.avianknowledge.net
- Benjamini Y, Yekutieli D (2001). The control of the false discovery rate in multiple testing under dependency. Ann Stat 29: 1165-1188.
- Bensch S, Irwin DE, Irwin JH, Kvist L, Akesson S (2006). Conflicting patterns of mitochondrial and nuclear DNA diversity in Phylloscopus warblers. Mol Ecol 15: 161-171
- Bensch S, Price T, Kohn J (1997). Isolation and characterization of microsatellite loci in a Phylloscopus warbler. Mol Ecol 6: 91-92.
- Braconnot P. Otto-Bliesner B. Harrison S. Joussaume S. Peterchmitt JY. Abe-Ouchi A et al. (2007). Results of PMIP2 coupled simulations of the Mid-Holocene and Last Glacial Maximum-Part 1: experiments and large-scale features. Climate Past 3: 261-277.
- Brelsford A, Irwin DE (2009). Incipient speciation despite little assortative mating: the yellow-rumped warbler hybrid zone. Evolution 63: 3050-3060.
- Burg TM, Gaston AJ, Winker K, Friesen VL (2005), Rapid divergence and postglacial colonization in western North American Steller's jays (Cvanocitta steller). Mol Ecol 14: 3745-3755
- Carstens BC, Richards CL, Crandall K (2007). Integrating coalescent and ecological niche modeling in comparative phylogeography. Evolution 61: 1439-1454.
- Chappell DE, Van Den Bussche RA, Krizan J, Patterson B (2004). Contrasting levels of genetic differentiation among populations of wolverines (Gulo gulo) from northern Canada revealed by nuclear and mitochondrial loci. Conserv Genet 5: 759-767.
- Clement M, Posada D, Crandall KA (2000). TCS: a computer program to estimate gene genealogies. Mol Ecol 9: 1657-1659.
- Colbeck GJ, Gibbs HL, Marra PP, Hobson K, Webster MS (2008). Phylogeography of a widespread North American migratory songbird (Setophaga ruticilla). J Hered 99: 453-463.
- Corander J, Marttinen P, Siren J, Tang J (2008). Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. BMC Bioinform 9: 539.
- Corander J, Tang J (2007). Bayesian analysis of population structure based on linked molecular information. Math Biosci 205: 19-31.
- Cronin MA, MacNeil MD, Patton JC (2005), Variation in mitochondrial DNA and microsatellite DNA in caribou (Rangifer tarandus) in North America. J Mammal 86: 495-505.
- De Lafontaine G, Turgeon J, Payette S (2010). Phylogeography of white spruce (Picea glauca) in eastern North America reveals contrasting ecological trajectories. J Biogeogr **37** · 741_751
- Dupanloup I, Schneider S, Excoffier L (2002). A simulated annealing approach to define the genetic structure of populations. Mol Ecol 11: 2571-2581
- Earl DA, vonHoldt BM (2012), STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4: 359-361.
- Ersts PJ (2010). Geographic Distance Matrix Generator (version 1.2.3). American Museum of Natural History, Center for Biodiversity and Conservation. Available from http://biodiversityinformatics.amnh.org/open_source/gdmg/
- Evanno G, Regnaut S, Goudet J (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14: 2611-2620.
- Excoffier L, Laval G, Schneider S (2005). Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol Bioinform 1: 47-50
- Excoffier L, Smouse PE, Quattro JM (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131: 479-491.

American Ornithologists' Union (1957). Check-list of North American Birds 5th edn. American Ornithologists' Union: Washington, DC, USA.

- Ficken MS, McLaren MA, Hailman JP (1996). *Birds of North America, No. 254*. In: Poole A, Gill F (eds) The Academy of Natural Sciences: Philadelphia, PA, USA;The American Ornithologists' Union: Washington DC, USA.
- Flockhart TD, Wiebe KL (2009). Absence of reproductive consequences of hybridization in the northern flicker (*Colaptes auratus*) hybrid zone. *Auk* **126**: 351–358.
- GBIF data portal (2011). Royal British Columbia Museum; Ontario Breeding Bird Atlas 2001–2005; Macaulay Library—Audio Data, Bird Collection; Peabody Ornithology DiGIR Service, Bird Tissue Collection; University of Alberta Museums, Ornithology Collection; Project FeederWatch; eBird Bird Observation Checklist Database; DMNS Bird Collection; Ontario Breeding Bird Atlas 1981–1985. http://data.gbif.org/datasets/ resource/x
- Gerardi S, Jaramillo-Correa JP, Beaulieu J, Bousquet J (2010). From glacial refugia to modern populations: new assemblages of organelle genomes generated by differential cytoplasmic gene flow in transcontinental black spruce. *Mol Ecol* **19**: 5265–5280.
- Gibbs HL, Dawson RJG, Hobson KA (2000). Limited differentiation in microsatellite DNA variation among northern populations of the yellow warbler: evidence for male-biased gene flow? *Mol Ecol* **9**: 2137–2147.
- Gill FB, Mostrom AM, Mack AL (1993). Speciation in North-American chickadees: I. Patterns of mtDNA genetic-divergence. *Evolution* 47: 195–212.
- Gill FB, Slikas B, Sheldon FH (2005). Phylogeny of titmice (Paridae): II. Species relationships based on sequences of the mitochondrial cytochrome-b gene. Auk 122: 121–143.
- Gillespie AR, Porter SC, Atwater BF (2004). *The Quaternary Period in the United States*. Elsevier: Amsterdam, The Netherlands.
- Goudet J (1999). PCA-GEN Version 1.2. Lausanne, Switzerland. www.unil.ch/izea/ softwares/pcegen.html.
- Graham BA, Burg TM (2012). Molecular markers provide insight into contemporary and historic gene flow for a non-migratory species. J Avian Biol 43: 1–17.
- Griffith SC, Stewart IRK, Dawson DA, Owens IPF, Burke T (1999). Contrasting levels of extra-pair paternity in mainland and island populations of the house sparrow (*Passer domesticus*): is there an 'island effect'? *Biol J Linn Soc Lond* **68**: 303–316.
- Hanotte O, Zanon C, Pugh A, Greig C, Dixon A, Burke T (1994). Isolation and characterization of microsatellite loci in a passerine bird: the reed bunting *Emberiza* schoeniclus. Mol Ecol 3: 529–530.
- Hearn BJ, Neville JT, Curran WJ, Snow DP (2006). First record of the southern red-backed vole, *Clethrionomys gapperi*, in Newfoundland: implications for the endangered Newfoundland marten, *Martes americana atrata. Can Field Nat* **120**: 50–56.
- Helbig AJ, Salomon M, Bensch S, Seibold I (2001). Male-biased gene flow across an avian hybrid zone: evidence from mitochondrial and microsatellite DNA. J Evol Biol 14: 277–287.
- Hewitt GM (2000). The genetic legacy of the Quaternary ice ages. Nature 405: 907–913. Hewitt GM (2004). Genetic consequences of climatic oscillations in the Quaternary. Phil Trans R Soc Lond B Biol Sci 359: 183–195.
- Hey J, Nielsen R (2007). Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proc Natl Acad Sci USA* **104**: 2785–2790.
- Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005). Very high resolution interpolated climate surfaces for global land areas. Int J Climatol 25: 1965–1978.
- Holder K, Montgomerie R, Friesen VL (1999). A test of the glacial refugium hypothesis using patterns of mitochondrial and nuclear DNA sequence variation in rock ptarmigan (*Lagopus mutus*). *Evolution* **53**: 1936–1950.
- Hutchison DW, Templeton AR (1999). Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution* **53**: 1898–1914.
- Jaramillo-Correa JP, Beaulieu J, Bousquet J (2004). Variation in mitochondrial DNA reveals multiple distant glacial refugia in black spruce (*Picea mariana*), a transcontinental North American conifer. *Mol Ecol* **13**: 2735–2747.
- Jaramillo-Correa JP, Beaulieu J, Khasa DP, Bousquet J (2009). Inferring the past from the present phylogeographic structure of North American forest trees: seeing the forest for the genes. *Can J Forest Res* **39**: 286–307.
- Jost L (2008). G_{ST} and its relatives do not measure differentiation. Mol Ecol $17\colon$ 4015–4026.
- Kalinowski ST (2005). HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. *Mol Ecol Notes* 5: 187–189.
- Kvist L, Martens J, Ahola A, Orell M (2001). Phylogeography of a Palaearctic sedentary passerine, the willow tit (*Parus montanus*). J Evol Biol 14: 930–941.
- Lait LA, Friesen VL, Gaston AJ, Burg TM (2012). The post-Pleistocene population genetic structure of a western North American passerine: the chestnut-backed chickadee (*Poecile rufescens*). J Avian Biol 43: 541–552.
- Lerner HRL, Meyer M, James HF, Hofreiter M, Fleischer RC (2011). Multilocus resolution of phylogeny and timescale in the extant adaptive radiation of Hawaiian honeycreepers. *Curr Biol* **21**: 1838–1844.
- Librado P, Rozas J (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451–1452.
- Martinez JG, Soler JJ, Soler M, Møller AP, Burke T (1999). Comparative population structure and gene flow of a brood parasite, the great spotted cuckoo (*Clamator* glandarius), and its primary host, the magpie (*Pica pica*). Evolution 53: 269–278.
- Milot E, Gibbs HL, Hobson KA (2000). Phylogeography and genetic structure of northern populations of the yellow warbler (*Dendroica petechia*). *Mol Ecol* **9**: 667–681.

- Otter K, Ratcliffe L, Michaud D, Boag PT (1998). Do female black-capped chickadees prefer high-ranking males as extra-pair partners? *Behav Ecol Sociobiol* **43**: 25–36.
- Otter KA, Stewart IRK, McGregor PK, Terry AMR, Dabelsteen T, Burke T (2001). Extra-pair paternity among great tits *Parus major* following manipulation of male signals. *J Avian Biol* **32**: 338–344.
- Päckert M, Martens J, Tietze DT, Dietzen C, Wink M, Kvist L (2007). Calibration of a molecular clock in tits (Paridae)—do nucleotide substitution rates of mitochondrial genes deviate from the 2% rule? *Mol Phylogenet Evol* **44**: 1–14.
- Peakall R, Smouse PE (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28: 2537–2539.
- Peakall R, Smouse PE (2006). GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* 6: 288–295.
- Petit RJ, Aguinagalde I, de Beaulieu JL, Bittkau C, Brewer S, Cheddadi R *et al.* (2003). Glacial refugia: hotspots but not melting pots of genetic diversity. *Science* **300**: 1563–1565.
- Phillips SJ, Anderson RP, Schapire RE (2006). Maximum entropy modeling of species geographic distributions. *Ecol Model* **190**: 231–259.
- Phillips SJ, Dudík M (2008). Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography* 31: 161–175.
- Pielou EC (1991). After the Ice Age: The Return of Life to Glaciated North America. University of Chicago Press: Chicago, IL, USA.
- Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.
- Provan J, Bennett K (2008). Phylogeographic insights into cryptic glacial refugia. Trends Ecol Evol 23: 564–571.
- Pulgarín-R PC, Burg TM (2012). Genetic signals of demographic expansion in downy woodpecker (*Picoides pubescens*) after the last North American glacial maximum. *PLoS One* 7: e40412.
- Ralston J, Kirchman JJ (2012). Continent-scale genetic structure in a boreal forest migrant, the blackpoll warbler (Setophaga striata). Auk 129: 467–478.
- Raymond M, Rousset F (1995). GENEPOP (Version 1.2): population genetics software for exact tests and ecumenicism. J Hered 86: 248–249.
- Richards CL, Carstens BC, Knowles L (2007). Distribution modelling and statistical phylogeography: an integrative framework for generating and testing alternative biogeographical hypotheses. *J Biogeogr* **34**: 1833–1845.
- Ridgway R (1904). The Birds of North and Middle America: A Descriptive Catalogue. Smithsonian Museum, United States National Museum: Washington, DC, USA.
- Rousset F (2008). Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Mol Ecol Resour* 8: 103–106.
- Shafer ABA, Côté SD, Coltman DW (2010a). Hot spots of genetic diversity descended from multiple Pleistocene refugia in an alpine ungulate. *Evolution* 65: 125–138.
- Shafer ABA, Cullingham CI, Côté SD, Coltman DW (2010b). Of glaciers and refugia: a decade of study sheds new light on the phylogeography of northwestern North America. *Mol Ecol* **19**: 4589–4621.
- Sorenson MD, Ast JC, Dimcheff DE, Yuri T, Mindell DP (1999). Primers for a PCR-based approach to mitochondrial genome sequencing in birds and other vertebrates. *Mol Phylogenet Evol* **12**: 105–114.
- Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF (1998). Comparative phylogeography and postglacial colonization routes in Europe. *Mol Ecol* 7: 453–464.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011). MEGA 5: Molecular Evolutionary Genetics Analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28: 2731–2739.
- Toews DPL, Irwin DE (2008). Cryptic speciation in a Holarctic passerine revealed by genetic and bioacoustic analyses. *Mol Ecol* **17**: 2691–2705.
- Uimaniemi L, Orell M, Kvist L, Jokimaki J, Lumme J (2003). Genetic variation of the Siberian tit *Parus cinctus* populations at the regional level: a mitochondrial sequence analysis. *Ecography* 26: 98–106.
- van Els P, Cicero C, Klicka J (2012). High latitudes and high genetic diversity: phylogeography of a widespread boreal bird, the gray jay (*Perisoreus canadensis*). *Mol Phylogenet Evol* **63**: 456–465.
- Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004). MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol Notes* 4: 535–538.
- Walsh PS, Metzger DA, Higuchi R (1991). Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques* 10: 506–513.
- Walter R, Epperson BK (2001). Geographic pattern of genetic variation in *Pinus resinosa*: area of greatest diversity is not the origin of postglacial populations. *Mol Ecol* **10**: 103–111.
- Wang M-T, Hsu Y-C, Yao C-T, Li S-H (2005). Isolation and characterization of 12 tetranucleotide repeat microsatellite loci from the green-backed tit (*Parus monticolus*). *Mol Ecol Notes* 5: 439–442.
- Weir JT, Schluter D (2008). Calibrating the avian molecular clock. *Mol Ecol* 17: 2321–2328.
- Zink RM, Barrowclough GF (2008). Mitochondrial DNA under siege in avian phylogeography. Mol Ecol 17: 2107–2121.
- Zink RM, Dittmann DL (1993). Gene flow, refugia, and evolution of geographic-variation in the song sparrow (*Melospiza melodia*). *Evolution* **47**: 717–729.
- Zink RM, Weckstein JD, Hackett S (2003). Recent evolutionary history of the fox sparrows (Genus: *Passerella*). Auk **120**: 522–527.

Supplementary Information accompanies this paper on Heredity website (http://www.nature.com/hdy)