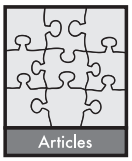

Journal of Avian Biology 43: 198–214, 2012
doi: 10.1111/j.1600-048X.2012.05604.x

Subject Editor: Staffan Bensch. Accepted 3 February 2012

Molecular markers provide insights into contemporary and historic gene flow for a non-migratory species

Brendan A. Graham and Theresa M. Burg

B. A. Graham (b.graham@uleth.ca) and T. M. Burg, Dept of Biological Sciences, Univ. of Lethbridge, 4401 Univ. Drive, Lethbridge, AB T1K 3M4, Canada.

Hairy woodpeckers *Picoides villosus* are a common, year round resident with distinct plumage and morphological variation across North America. We genotyped 335 individuals at six variable microsatellite loci and analyzed 322 mtDNA control region sequences in order to examine the role of contemporary and historical barriers to gene flow. In addition we combined genetic analyses with ecological niche modelling to test if hairy woodpeckers were isolated in northern refugia (Alaska, Newfoundland and the Queen Charlotte Islands) during the last glacial maximum. Genetic analyses revealed that gene flow among North American hairy woodpecker populations is restricted, but not to the extent predicted for a sedentary species. Populations clustered into two main genetic groups, east and west of the Great Plains in the south and the Rocky Mountains in the north. Contact zones between the two main genetic groups exist in central British Columbia and Washington, but are narrow. Within each group we found additional population structure with genetic breaks between subgroups in the geographic west corresponding to breaks in forested habitat and physical barriers like open expanses of water. Population genetic patterns for hairy woodpeckers have resulted from isolation in multiple southern refugia with the current distribution of genetic groups resulting from post-glacial expansion and subsequent reduction in gene flow. While populations in Alaska, Newfoundland and the Queen Charlotte Islands are genetically distinct from other populations, we found no evidence of these areas acting as refugia throughout the Pleistocene. Atlantic Canada populations contained unique haplotypes raising the possibility of a separate colonization from the rest of eastern Canada. The endemic subspecies on the island of Newfoundland is not genetically distinct from their closest mainland population unlike the Queen Charlotte Island subspecies.

Advances in the availability and application of molecular markers have aided in our understanding of how landscapes affect population genetic patterns, variation and gene flow (Manel et al. 2003). Past studies have shown that barriers to gene flow caused by topographical features (Keyghobadi et al. 1999), unsuitable habitat (Piertney et al. 1998) and human-mediated land changes (Gerlach and Musolf 2000) all affect contemporary genetic patterns. Exploring environmental conditions and landscape features will provide further insight into understanding the processes that influence gene flow and how patterns of genetic variation within species are produced.

In North America previous studies have examined the role of past glacial events in shaping population structure and promoting genetic diversification for both plant (reviewed by Jarmillo-Correa et al. 2009) and animal (Avise et al. 1998, Hewitt 2000, Burg et al. 2006) species. During the last glacial maximum (LGM), most of North America north of 48°N was covered by ice sheets (Pielou 1991). The ice sheets were disruptive and fragmented species' ranges, restricting individuals and populations to ice-free areas known as refugia (Klicka and Zink 1999). Known refugia during the LGM included Beringia and southern North America, while the Queen Charlotte Islands (QCI), also known as Haida Gwaii, and Newfoundland are contested to have been ice-free (Pielou 1991). Populations expanded from these refugia following the melting of the ice sheets. How individuals colonized previously glaciated areas following deglaciation remains an important question, as the mode of colonization affects contemporary genetic patterns (Johansen and Latta 2003). Two alternate patterns of colonization have been proposed by Hewitt (1996). Following the pioneer model, individuals disperse long distances establishing pockets of isolated populations ahead of the leading edge, thereby preventing further colonization events. The second model, the phalanx model, is where colonization from the refugium is gradual and continuous resulting in little genetic structure with more recently founded populations being similar genetically to source populations.

The hairy woodpecker *Picoides villosus* is a common year-round resident with very limited short distance dispersal (Jackson et al. 2002) found in most forest and woodland habitats at both high and low elevations. Its range extends from Alaska to the highlands of Panama and from the Pacific to the Atlantic Coast (Fig. 1). Currently seventeen subspecies are classified by plumage, morphological and behavioural

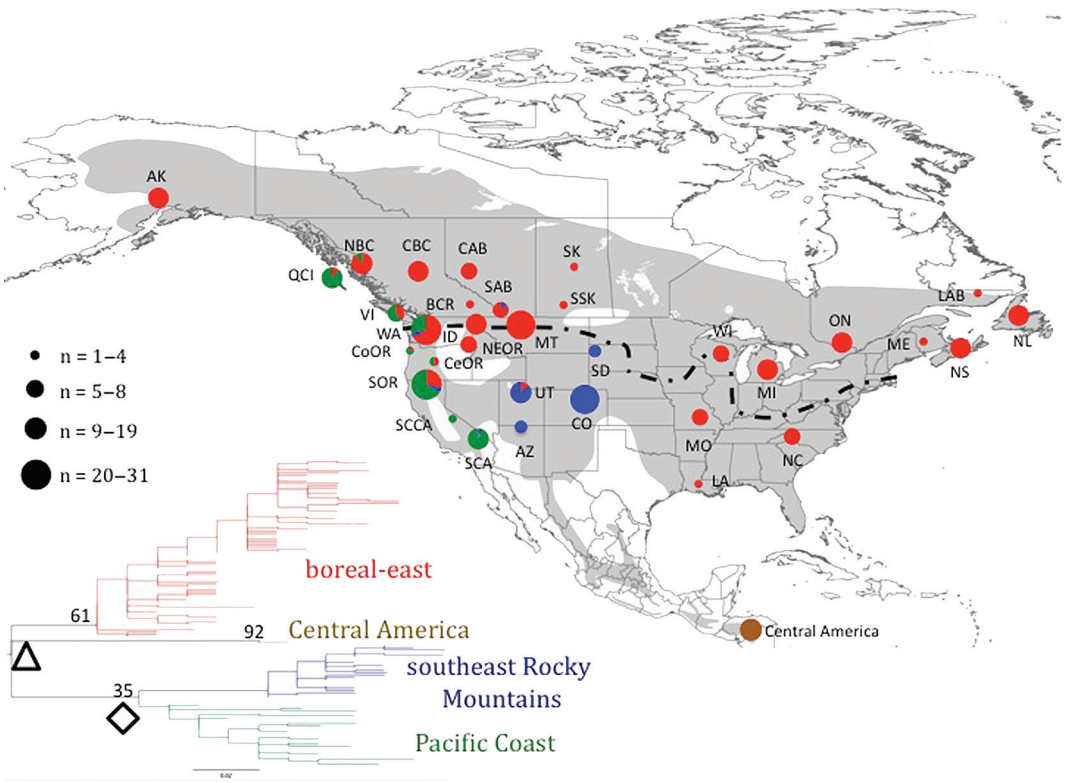

Figure 1. Geographic distribution of sampling sites for mtDNA analyses of hairy woodpeckers. Shaded area represents the current distribution of hairy woodpeckers throughout North and Central America (Ridgely et al. 2007). Pie charts show population assignment to the four mtDNA groups based on TCS (Fig. 2): red = boreal-east group; green = Pacific Coastal group; blue = southeast Rocky Mountain group; brown = Central America group. Dashed black line represents the extent of ice-sheets at the last glacial maximum (~21 ka). Inset shows an unrooted maximum likelihood tree. Numbers at major nodes represent bootstrap values and colours correspond to TCS mtDNA groups above. Triangle denotes haplogroups diverged ~100 ka (see text for exact dates) and diamond denotes that divergence occurred within the western group (Pacific Coast and southeast Rocky Mountains) ~34 ka.

characteristics with eleven subspecies described in North America (Jackson et al. 2002). The presence of distinct plumage and morphological traits across the range suggests limited dispersal or rapid phenotypic adaptation. Recent genetic studies on hairy woodpeckers (Topp and Winker 2008, Klicka et al. 2011) show restricted gene flow in previously unglaciated areas. Topp and Winker (2008) found birds on the Queen Charlotte Islands (n = 3) contained unique cytochrome b haplotypes; though a subsequent study by Klicka et al. (2011) using mtDNA ND2 and examining birds from a wider geographic area showed these birds clustered with mainland populations. More comprehensive sampling of unglaciated areas showed the presence of multiple matrilineal groups in central and southern North America and Central America (Klicka et al. 2011). Despite these studies, the question remains as to how previously glaciated portions of North America were colonized and how populations in these areas are connected.

The hairy woodpecker's range is delineated by a number of physical barriers including mountain ranges, areas of unsuitable habitat, and large bodies of water separating continental subspecies from those found on islands (i.e. QCI and Newfoundland; Jackson et al. 2002). Studies of migratory birds have shown genetic differences between populations on either side of mountains (Clegg et al. 2003, Boulet and Gibbs 2006, Hull et al. 2008), but processes creating phylogeographic patterns remain poorly understood because of the confounding effects of long distance migration and potential mixing on wintering grounds. Studying sedentary species like the hairy woodpecker will provide greater insight into gene flow, as patterns found in resident birds are not affected by long distance, seasonal migration. Many, but not all, of the studies on sedentary species have found evidence of restricted dispersal with two or more western clades and, where ranges extend to the east coast, a single eastern clade (Barrowclough et al. 2004, Spellman and Klicka 2007, Spellman et al. 2007). Hairy woodpeckers remain in their breeding areas year round and do not undergo true migration (Ouellet 1977). Banding data suggest sedentary behavior with 97% of all banding recoveries (n = 813) < 40 km from their original capture location (Jackson et al. 2002). However patterns of gene flow are not always reflected in banding data or migratory patterns. Some migratory species show patterns inconsistent with migratory flyways (Arguedas and Parker 2000), while other species with extremely high philopatry show evidence of gene flow (Burg and Croxall 2001).

To explore patterns of contemporary and historic gene flow in North America, we used mitochondrial and microsatellite loci and samples from across the range to test the hypotheses: a) populations separated by physical barriers are

genetically isolated and b) island endemics are genetically distinct. A second objective was to determine the genetic affinity of birds from previously glaciated regions and to determine if hairy woodpeckers recolonized previously glaciated regions from northern (i.e. Beringia (Alaska), QCI and Newfoundland) or multiple southern refugia.

Methods

Sampling and DNA extraction

Blood and tissue samples were obtained from birds at 34 sampling sites covering the North American range of the hairy woodpecker (Fig. 1). We attempted to restrict sampling to a 30 km radius for each sampling location. We used mist nets to capture birds during the 2007–2010 breeding seasons and collected blood samples (~100 µl of blood) from the brachial vein. All blood samples were preserved in 95% ethanol. Tissue samples were also obtained from collections (Acknowledgements) to supplement our field sampling and add new sampling sites (e.g. Costa Rica). All museum samples were from birds collected over the last 20 yr to ensure that we analyzed contemporary genetic patterns and used the same criteria (i.e. during the breeding season from a small geographic area) as blood sampling. By limiting our samples to birds from the breeding season, we tried to avoid sampling birds atypical of each population, since hairy woodpeckers have been reported to move short distances during the late fall and early winter (Jackson et al. 2002). DNA was extracted from blood and tissue samples using a modified chelex protocol (Walsh et al. 1991) as described in Burg and Croxall (2001).

Microsatellite genotyping

Samples were screened with fifteen microsatellite primer sets isolated from other woodpecker species (Dlu1, Dlu5, Ellegren et al. (1999); DMA2, DMD7, DMD9, DMC111, DMD112, DMC115, DMC118, DMB119, DMA120, Vila et al. (2008); Ptri 3, Välimäki et al. (2008); and RCW05, RCW20, Fike et al. (2009)). All polymerase chain reactions (PCR) were conducted in 10 µl reactions using Goflexi clear buffer (Promega), 0.2 mM dNTP, 2.5 mM $MgCl_2$, 1 µM of each forward, reverse and M13 fluorescently labeled primers and 0.5 U of Crimson Taq (New England BioLabs). The only exceptions were DMD9 and Ptri3 where the concentration of $MgCl_2$ was 1.5 mM and 2 mM respectively. DNA was amplified using the following PCR conditions: one cycle for 60 s at 94°C, 45 s at T_1, 60 s at 72°C; seven cycles of 60 s at 94°C, 30 s at T_1 and 45 s at 72°C; 31 cycles for 30 s at 94°C, 30 s at T_2 and 45 s at 72°C; and ended with 1 cycle of 72°C for 5 min and 20 s at 4°C. Eleven of the fifteen loci (Dlu1, Dlu5, DMD7, DMC111, DMD112, DMC115, DMC118, DMB119, DMA120, Ptri5 and RCW20) were optimized using $T_1 = 50$°C and $T_2 = 52$°C, while the remaining four loci (DMA2, DMD9, Ptri3 and RCW5) were optimized using $T_1 = 45$°C and $T_2 = 48$°C. PCR products were run on a 6% acrylamide gel on a Li-Cor 4300 DNA analyzer. To maintain consistent sizing and scoring of alleles, we ran controls with known size standards on every load. All gels were scored independently by two observers and a subset of samples from each gel was re-amplified and re-run together to ensure consistent amplification and scoring among loads. For all analyses we only used individuals that had genotypic information for at least four of the variable loci.

MtDNA amplification and sequencing

The mitochondrial control region (CR) was amplified using LThr (5′ CAT TGG TCT TGT AAR CCA AAG 3′) and H16824 (5′ TGA TGG GAT TTT AGA GGA TGT G 3′) primers. Nested primers were designed to amplify DNA from degraded tissue samples and full sequences were obtained using two primer sets: LThr and H16454 (5′ GAC CAG TAA TGG CCC TGA GA 3′) and L16388 (5′ GCT TCA GGC CCA TAC TTT CC 3′) with H16824. All PCRs were conducted in a 25 µl volume using 5× Crimson Taq (Mg-free) reaction buffer (New England Biolabs), 0.2 mM dNTP, 2.5 mM $MgCl_2$, 1 µM of each an L strand and H strand primer and 1 U of Crimson Taq (New England BioLabs) (Burg et al. 2005). DNA was amplified using the following conditions: one cycle for 2 min at 94°C, 45 s at 54°C, 1 min at 72°C; 37 cycles of 30 s at 94°C, 45 s at 54°C and 1 min at 72°C; ending with 1 cycle of 72°C for 5 min and 20 s at 4°C. PCR products were run on a 1% agarose gel to confirm DNA amplification.

For sequencing, five microlitres of PCR product were incubated with 0.1 units of SAP and 0.1 units of exonuclease at 37°C for 15 min and enzymes were deactivated by heating to 80°C for 15 min. One microlitre of purified PCR product was used for sequencing in a 10 µl reaction using 2.5× McLab BigDye Terminator sequencing buffer, 1.5 units of BigDye Terminator Mix (ver. 3.1) and 0.5 µM of primer (LThr or L16388). Cycle sequencing was done using the following conditions: one cycle at 96°C for 45 s; 25 cycles at 96°C for 30 s; 50°C for 15 s; 4 min at 60°C and 4°C for 20 s. Sequencing reactions were purified using a standard sodium acetate ethanol precipitation protocol (Sambrook et al. 1989) and run on an Applied Biosystems 3130 DNA Analyzer (Univ. of Lethbridge) or 3730xl DNA Analyzer (Génome Québec Innovation Centre). Sequences were aligned using MEGA ver. 4.0 (Tamura et al. 2007) and double-checked by visual inspection of the chromatograms.

Statistical analyses

We measured genetic variation within populations and haplogroups (Nei and Tajima 1981) by calculating haplotype diversity (H_d) and nucleotide diversity (π) in DNASP ver. 5.0 (Librado and Rozas 2009). Deviations from Hardy–Weinberg and linkage equilibrium between loci were analyzed using GENEPOP ver. 4.0.10 (Raymond and Rousset 1995) and genotyping errors, presence of null alleles and allelic dropout were checked using MICROCHECKER ver. 2.23 (Van Oosterhout et al. 2004). Observed heterozygosity (Ho), expected heterozygosity (He), total number of alleles (Na) and number of private alleles (PA) were calculated using GenAlEx ver. 6.41 (Goudet 1995, Peakall and Smouse 2006), while allelic richness was calculated using the program FSTAT ver. 2.9.2.3 (Goudet 1995). As population

bottlenecks result in a loss of alleles and temporarily high levels of observed heterozygosity, they can influence contemporary genetic patterns. To determine if any of the populations have recently undergone a bottleneck, a Wilcoxon sign-rank test was performed with BOTTLENECK ver. 1.2.02. We used the Wilcoxon sign-rank test, as it is the most powerful test to use when analyzing populations with fewer than 20 individuals (Piry et al. 1999).

Genetic variation and population structure

Pairwise F_{ST} values were calculated for both mitochondrial and nuclear data to examine population structure and assess genetic differentiation between populations and mtDNA haplogroups. Pairwise comparisons were conducted for populations with at least ten individuals for both microsatellites (n = 14) and mtDNA (n = 17) (Harding 1996). We included the Queen Charlotte Islands (n = 9) as it contains an endemic subspecies (*P. v. picoideus*). We used pairwise F_{ST} values, as they are capable of detecting differences when sample sizes are low (n = 10; Harding 1996) and genetic differences are small. We calculated pairwise F_{ST} values using ARLEQUIN ver. 3.11 (Excoffier et al. 2005; for mtDNA) and GENETIX ver. 4.0.10 (Belkhir 1999; for microsatellites) and significance was determined using 10 000 permutations. All tests of population differentiation using pairwise F_{ST} were corrected for using the false discovery rate correction method (Benjamini and Hochberg 1995). Recent studies have suggested that Bonferroni corrections are too severe and often underestimate the number of populations that are significantly different leading to an increase in type II errors (Garamszegi 2006). As sequential Bonferroni corrections (Rice 1989) are still widely reported, we have included both types of corrections (Cabin and Mitchell 2000). Because the theoretical maximum of one for F_{ST} is only valid when there are two alleles and microsatellites have high allelic variation, we calculated the global, theoretical maximum F_{ST} for microsatellites (Hedrick 1999).

In addition to calculating pairwise F_{ST} values, we used two Bayesian clustering models: STRUCTURE ver. 2.3.3 (Pritchard et al. 2000) and BAPS ver. 5 (Corander et al. 2003) to assess genetic differentiation for microsatellites. Bayesian models are more conservative than pairwise F_{ST} values. Clustering models perform best when F_{ST} values are ≥ 0.05 and they can overlook weak population structure (F_{ST} values are ≤ 0.03; Latch et al. 2006). One disadvantage of F_{ST} is it uses pre-defined groupings, in this case sampling sites. For these reasons we used the two different methods to properly examine population structure.

As assignments in the program STRUCTURE are based on individual genotypes and not on population allele frequencies, samples from the less well sampled populations (Table 1, Fig. 1) were included. For all STRUCTURE runs, we used correlated alleles, location information, admixture, and a burn in length of 100 000 Monte Carlo Markov chains (MCMC) followed by a run of 500 000. Each K (1–10) was run for 10 iterations. To determine the number of genetic clusters (K) present, we used two methods: Bayes factor (Pritchard et al. 2000) and ΔK (Structure Harvester < http://taylor0.biology.ucla.edu/struct_harvest/ >; Evanno et al. 2005). Following the initial STRUCTURE runs, two main groups were detected. A second set of runs with the same settings was done using individuals from each of the two main groups separately to determine if additional clusters were present within the groups. Those individuals that were assigned between the two main groups (max $Q \leq 0.60$; Fig. 2a) were removed from subsequent runs.

The second Bayesian clustering program, BAPS ver. 5 differs from STRUCTURE in that it originally assumes one panmictic population and assigns individuals to genetic clusters using geographic information arriving at its conclusions based on joint posterior probabilities (Corander et al. 2003). In addition BAPS implements a spatial model (Corander et al. 2008) that utilizes coordinate points (i.e. longitude and latitude of the sample) and plots the spatial pattern of the genetic variation. BAPS has been shown to be conservative in the assignment of individuals (Latch et al. 2006) and so we compared results with STRUCTURE and F_{ST}. We used the no admixture spatial cluster model to plot the spatial genetic patterns using the default settings in BAPS (Corander et al. 2008) and 23 populations (Fig. 2a) with at least five individuals. Preliminary runs showed the program was inconsistent with its assignment to populations with fewer than five individuals.

Finally we used the program SAMOVA (ver. 1.0, Dupanloup et al. 2002) to determine the geographic

Table 1. Repeat motif, primer sequence and allele size ranges for variable microsatellite loci used to genotype hairy woodpecker samples. F primers were modified with M13 sequence (not shown) added to the 5′ end.

Locus	Repeat type	Sequence (5′ to 3′)	Size
DlU1F[1]	$(TG)_{17}$	CAC ACT GAA CAT ACC ATG TG	158–174
DlU1R		TAA AGA CCC TAA ACT TGC ACA	
DlU5F[1]	$(GT)_{10}$*	CTG ACC AAA GTG GAA AAG TAA	167–182
DlU5R		TCC TAC TAC CAT TTC TAG AAC	
DMC111F[2]	$(CATC)_{10}$	CGT ATG GAC CAG AAC ATA ATG	208–268
DMC111R		TGG GCT TTT AAG TCT TGT TG	
DMC115F[3]	$(ATCC)_{10}$	TGT CAG AGA TGG TTC ATG GGT GCA CT	292–332
DMC115R		CCA CTG GTG GCT CAG TTG CAC A	
DMD118F[2]	$(TGGA)_{10}(TAGA)_{13}$	CCC ATA TCC AGA GTT AGT TCT G	180–236
DMD118R		TCC TAG AGT CTT CAA CCT GAT C	
Ptri3F[4]	$(AGAT)_{11}$	GCA AAA GCC AGT TCC TGT GCA TGG	292–348
Ptri3R		GTT TCT TCA CCA TCA TTT TCC AGA CAG AA	

[1]Ellegren et al. (1999); [2]Vila et al. (2008); [3]modified from Vila et al. (2008); [4]modified from Välimäki et al. (2008); *single base pair insertion/deletion in flanking region for hairy woodpecker.

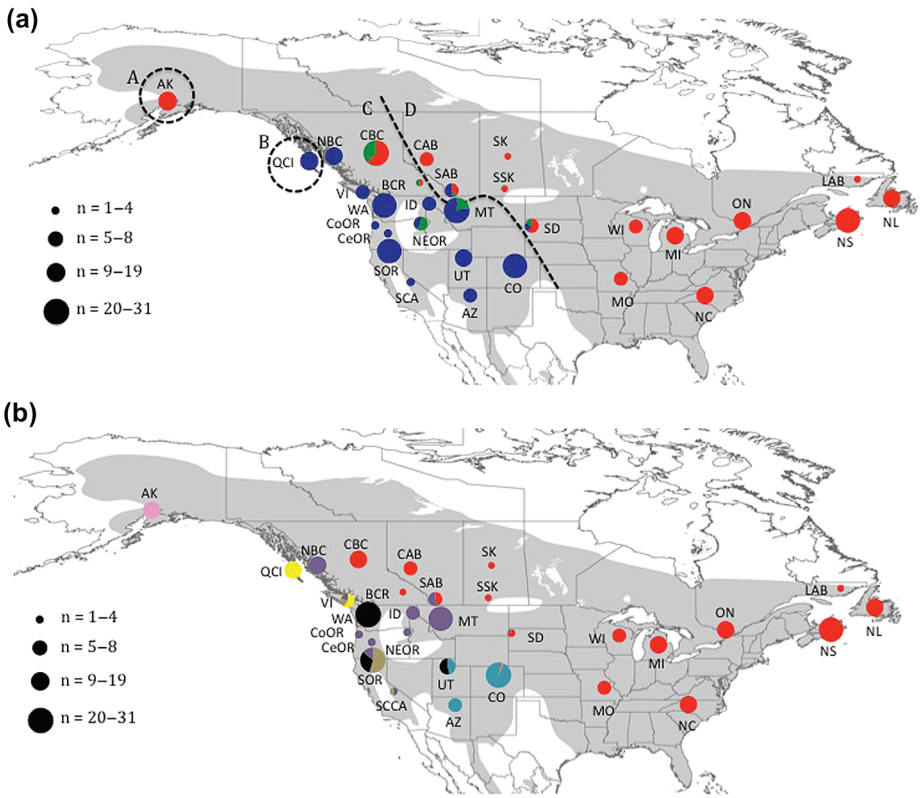

Figure 2. Geographic distribution and population structure of hairy woodpeckers in North America based on microsatellite data. Sample sizes are provided in Table 2. Shaded grey area represents the current range distribution of hairy woodpeckers (Ridgely et al. 2007). Colours show assignment of populations to groups based on: (a) STRUCTURE analysis for K = 2; red = east group; dark blue = west group; green = individuals showing admixture between east and west groups; dashed circles and line indicate population assignment to four groups (A–D) based on BAPS analysis for K = 4 excluding sampling sites with ≤ 4 individuals. (b) Results from separate STRUCTURE analysis for K = 2 within the east group and K = 4 within the west group. Pink = Alaska subgroup, red = boreal-eastern US subgroup; yellow = the Queen Charlotte Islands subgroup, purple = interior west subgroup, black = Pacific Coast subgroup, aqua = southeastern Rocky Mountain subgroup, khaki = admixed individuals between western subgroups. Solid black lines represent barriers to gene flow supported by pairwise F_{ST} values and BARRIER. Black dashed line represents barriers to gene flow identified by BARRIER but not supported by pairwise F_{ST} values.

population structure for mtDNA data. Similar to Bayesian clustering models, SAMOVA analyzes population structure by assigning populations to genetic clusters. SAMOVA uses a simulated annealing procedure to determine genetic structure and the number of genetic clusters (K) is determined by finding the K with the highest F_{CT} value. We ran SAMOVA for K = 1 to 7 for 500 iterations.

Phylogeographic structure

To visualize phylogeographic structure and evaluate relationships among haplotypes, we constructed a statistical parsimony network using the program TCS ver. 1.21 (Clement et al. 2000). We created an unrooted maximum likelihood tree to further assess the relationship among individual haplotypes. Before creating the maximum likelihood tree we first ran JModelTest ver. 0.1.1 (Guindon and Gascuel 2003, Posada 2008) to determine the nucleotide evolution model that best fit our data set. The program chose a generalized time-reversible model plus gamma plus invariants (GTR + G + I) as the best fit and we confirmed this model over other suggested models using Bayes factor. The ML tree was created in MEGA 5.0 (Tamura et al. 2011) and FigTree ver. 1.3.1 (Rambaut and Drummond 2006) using 500 bootstrap replicates.

We used two demographic tests, Fu's F_S (Fu 1997) and R_2 (Ramos-Osnins and Rozas 2002), to test whether populations of hairy woodpeckers have been stable or undergone recent demographic expansion. Both statistical measurements have high statistical power and are less conservative than other demographic measurements (Ramos-Osnins and Rozas 2002), with F_S performing best when sample sizes are high (n = 50) and R_2 when sample sizes are limited (n = 10). Demographic tests suggest recent population expansion when values are significant ($p \leq 0.02$; Ramos-Osnins and Rozas 2002) and for F_S if values are negative and significant. To ensure that populations were evolving neutrally and not subjected to background selection or genetic hitchhiking, we performed Fu's F* and Li's D* (Fu and Li 1993). Background selection is deemed to occur when Fu's F* and Li's D* are significant and Fu's F_S is non-significant for populations (Fu 1997). All tests of demographic expansion were performed in DNASP using 10 000 coalescent simulations. To characterize and visualize population demography (in addition to the demographic tests D, F_S and R_2), we plotted mismatch distributions. Mismatch distribution

differs from other neutrality tests as it explores the distribution and frequency of the number of pairwise nucleotide differences between individual sequences. Mismatch distributions can be used to infer population history where a leptokurtic, unimodal distribution indicates a recent, rapid expansion or bottleneck, while a bimodal curve indicates the presence of more than one genetic lineage (Slatkin and Hudson 1991). Raggedness index (r) measures whether mismatch distributions fit a model of population expansion (Harpending 1994), with significant values indicating stable populations.

To determine when populations began expanding we calculated tau (τ). τ is equal to 2ut, where u equals $2\mu k$, μ is the mutation rate (11.1%/Ma, see below), k is the length of the sequence and t the time of expansion (Rogers and Harpending 1992). Estimates for divergence times (T) were calculated using a standard population genetic technique (Morris-Pocock et al. 2010) where $T = \delta/r$, with δ equal to the net number of nucleotide substitutions per site between populations and r equal to the mutation rate (Rogers and Harpending 1992). Mutation rates were estimated from avian mitochondrial control region values calculated for other avian species (domain I = 20%/Ma; domain II = 5%/Ma; domain III = 23%/Ma; Baker and Marshall 1997). The 11.1%/Ma mutation rate was the average based on the proportion of each domain of the mtDNA CR sequenced (domain I = 288 bp; domain II = 496 bp; domain III = 41 bp). All τ and net nucleotide differences were calculated using DNASP ver. 5.0 (Librado and Rozas 2009).

Influence of barriers on genetic structure

To study how barriers affect genetic structure, we used three approaches. First a Mantel test was conducted to test for isolation by distance (IBD; Wright 1946) using GenAlEx ver. 6.41 (Peakall and Smouse 2006). We plotted $F_{ST}/(1\text{-}F_{ST})$ against straight line geographic distances, calculated using the Geographic Distance Matrix Generator <http://biodiversityinformatics.amnh.org/open_source/gdmg/>, between each of the populations used for F_{ST} analysis (Table 3). Tests for IBD were done for all populations and within geographic regions (east vs west). Second BARRIER 2.2 (Manni et al. 2004) was used to detect restrictions in gene flow. BARRIER connects sampling locations based on latitude and longitude (centre of sampling site) using Delaunay triangulations. Barriers are determined using Monmonier's distance algorithm. F_{ST} matrices were generated for each locus and for all six loci combined for 23 populations. The strength of each barrier was assessed based on the number of loci supporting it. We kept all barriers up to the tenth order with at least five loci supporting it. Finally we paired populations separated by contemporary physical barriers (e.g. sampling sites on either side of the Rocky Mountain Range: UT and CO) and compared F_{ST} values between populations.

The program BayesAss+ (Bayesian inference of recent migration using multilocus genotypes) ver. 1.3 uses MCMC to estimate recent migration rates (Wilson and Rannala 2003). We used BayesAss+ to determine rates of migration between populations with sample sizes of nine or more and to further determine how isolated populations are from each other. We ran BayesAss+ for 3 000 000 MCMC chains with a burn in of 999 999 using the default options as advised by Wilson and Rannala (2003). We performed 10 iterations and determined the best fit model using a Bayesian deviance formula as suggested by Faubet et al. (2007).

Ecological niche modelling

We reconstructed past distributions of hairy woodpeckers using MAXENT ver. 3.3.3e (Phillips et al. 2006) to predict potential refugia occupied by hairy woodpeckers in North America during the LGM. Ecological niche modelling (ENM) predictions have been shown to correspond with phylogeographic patterns, suggesting that both methods are complimentary to each other (Waltari et al. 2007). MAXENT combines current distribution records with climate conditions to infer past distributions by identifying areas in the past where these climate conditions were found. To determine if historical distribution models were suitable, we compared AUC (area under the curve) values between the training and testing models and looked at the omission rate. Model suitability is determined using the AUC values where values range between 0 and 1, and predictions are considered strong if differences between testing and training model values are above 0.8 and the omission rate does not show large deviations from the predicted omission (Phillips et al. 2006).

Climate variables were downloaded from the WorldClim dataset (Hijmans et al. 2005). Of the 19 variables, nine were correlated with each other ($r > 0.9$) and we removed these variables as suggested by McCormack et al. (2010) and therefore model predictions are based on 10 climate variables (BIO1, BIO2, BIO3, BIO7, BIO8, BIO12, BIO14, BIO15, BIO18, BIO19). We collected 5185 occurrence records, combining samples used for genetic analyses (n = 335) with occurrence records downloaded from the Global Biodiversity Information Facility (GBIF) data portal <http://data.gbif.org>. Distribution modelling for the LGM used paleoclimate data drawn from the community climate system model (Otto-Bliesner et al. 2006). All duplicate points were omitted by the program for distribution predictions to prevent any sampling bias due to specific locations in the data set. We ran 10 replicates and all simulations used the default convergence threshold, maximum number of iterations (500) and the cross-validation method.

Results

Of the fifteen loci we screened, eight (DMA2, DMD7, DMD9, DMD112, DMB119, DMA120 RCW05, and RCW20) were monomorphic and one (Ptri5) amplified inconsistently even after redesigning the primers, so they were not used. We genotyped 335 hairy woodpeckers from 30 sampling locations across North America for the remaining six variable microsatellite loci (Table 1, Fig. 2b). Two loci (Dlu1 and Dlu5) showed evidence of null alleles. These same two loci showed significant deviations from Hardy–Weinberg equilibrium for twelve sampling sites following corrections for multiple tests ($p < 0.025$). Removing Dlu5 from the analysis resulted in all sampling sites being in Hardy–Weinberg equilibrium with the exception of CO,

MT and NS ($p \leq 0.025$). We included Dlu1 and Dlu5 in all microsatellite analyses as preliminary analyses showed similar results when the two loci were excluded. No tests for linkage disequilibrium were significant and no population showed significant evidence of a recent bottleneck ($p \geq 0.05$ for excess heterozygosity).

Observed and expected heterozygosity ranged from 0.10 to 1.00 depending on the population and locus. We found 8 to 15 alleles/locus and allelic richness ranged from 2.9 to 4.3 (Table 2). No population had more than one private allele and the five private alleles are not restricted to refugial populations (Table 2). The AK and QCI populations showed reduced genetic variability (fewer total alleles and low allelic richness) from nearby mainland populations while genetic variation on Newfoundland was comparable to other populations.

We sequenced an 825 bp sequence of the avian mitochondrial control region for 322 hairy woodpeckers from 34 sampling localities (Fig. 1). We found 50 variable sites, 32 of which were parsimony informative and identified 120 different haplotypes, two of which were unique to Central America (Fig. 3; Supplementary material Appendix 1, 2). Haplogroups were differentiated based on our maximum likelihood tree (Fig. 1 inset) and TCS (Fig. 3), which recognized three distinct lineages: Central America, boreal-east and western North America. Additionally western North America can be further differentiated into two subgroups: Pacific Coast and southeast Rockies. Within the Pacific Northwest, admixture of Pacific Coast and boreal-east haplotypes are present, although Pacific Coast haplotypes are more common within this region (Fig. 1). With the exception of Central America, haplotype diversity was high ($H_d = 0.92–0.94$, total 0.97; Table 2) with moderate levels of nucleotide diversity ($\pi = 0.003$, overall 0.009) within the four groups. Within individual populations patterns for both haplotype and nucleotide diversity were similar (Table 2) with the exception of Alaska where haplotype diversity was low ($H_d = 0.36$) and Central America ($H_d = 0.17$,

Table 2. Summary statistics of mtDNA and microsatellite results for each population; n = sample size; Nh = number of haplotypes; Priv H = private haplotypes; Hd = haplotype diversity; π = nucleotide diversity; F_S = Fu's F_S; R_2 = R_2 test; PA = number of private alleles; Ho = observed heterozygosity; He = expected heterozygosity; AR = allelic richness; and overall = all birds combined. Asterisk (*) denotes significant at $p \leq 0.02$; (–) denotes missing data. Refer to Fig. 1 for location of sampling sites.

	Mitochondrial DNA							Microsatellites				
Locality	n	Nh	Priv H	Hd	π	F_S	R_2	n	PA	Ho	He	AR
AK	10	2	1	0.36	0.0020	3.03	0.20	10	0	0.60	0.56	2.89
NL	13	8	3	0.86	0.0030	−2.85	0.10	14	0	0.76	0.79	4.00
SD	5	4	1	0.90	0.0030	−0.33	0.20	5	0	0.58	0.61	3.43
CO	20	12	6	0.93	0.0030	−5.18*	0.10	29	0	0.58	0.77	3.77
AZ	7	6	4	0.95	0.0050	−1.70	0.19	7	0	0.73	0.62	3.18
UT	15	9	3	0.89	0.0060	−1.36	0.11	15	0	0.68	0.75	3.76
SCCA	3	2	0	0.67	0.0020	–	0.12	2	0	–	–	–
SCA	16	7	2	0.75	0.0030	−1.50	−0.47	–	–	–	–	–
QCI	9	5	2	0.81	0.0050	0.78	0.21	9	0	0.58	0.59	3.03
CeOR	2	2	0	1.00	–	2.56	0.50	2	0	–	–	–
CoOR	4	4	1	1.00	0.0100	0.09	0.26	4	0	–	–	–
SOR	22	17	8	0.97	0.0100	−5.54*	0.17	22	1	0.69	0.79	3.93
VI	5	5	4	1.00	0.0110	−0.57	0.25	5	1	0.65	0.64	3.56
WA	21	10	6	0.78	0.0090	0.48	0.17	21	0	0.67	0.80	4.04
SAB	7	4	2	0.71	0.0050	1.25	0.29	8	0	0.66	0.73	3.78
MT	20	9	2	0.78	0.0020	−5.10*	0.09*	31	1	0.69	0.80	4.05
ID	11	5	0	0.78	0.0020	−1.08	0.16	7	0	0.71	0.74	3.79
NEOR	6	5	2	0.93	0.0030	−1.97	0.19	5	0	0.66	0.73	4.00
BCR	2	1	0	–	–	–	–	2	0	–	–	–
CBC	18	11	5	0.86	0.0030	−5.20*	0.14	21	0	0.83	0.79	3.94
NBC	10	6	2	0.78	0.0050	−0.08	0.20	10	0	0.62	0.76	3.85
CAB	7	4	1	0.81	0.0020	−0.23	0.22	7	0	0.72	0.67	3.59
SK	5	5	3	1.00	0.0050	−1.72	0.18	5	0	–	–	–
LAB	4	3	1	0.83	0.0030	0.25	0.34	4	0	–	–	–
NS	19	16	5	0.98	0.0040	−13.48*	0.12	30	1	0.66	0.81	4.05
ME	1	1	1	–	–	–	–	–	–	–	–	–
MO	3	3	1	1.00	0.0050	0.13	0.14	5	0	0.74	0.72	4.02
NC	7	7	3	1.00	0.0050	−3.71*	0.14	15	1	0.72	0.76	3.77
LA	3	3	1	1.00	–	−0.08	0.25	–	–	–	–	–
WI	6	6	2	1.00	0.0030	−3.39*	0.20	6	0	0.50	0.62	4.25
MI	14	11	6	0.96	0.0030	−6.62*	0.10	15	0	0.59	0.76	3.79
ON	15	10	5	0.86	0.0030	−4.79*	0.08*	19	0	0.53	0.62	3.70
C. America	12	2	2	0.17	0.0002	−0.48	0.28	–	–	–	–	–
Boreal-east	208	70	–	0.94	0.0030	−96.16*	0.04	–	–	–	–	–
Pac Coast	53	23	–	0.92	0.0030	−17.09*	0.07	–	–	–	–	–
SE Rockies	49	25	–	0.94	0.0030	−15.05*	0.06	–	–	–	–	–
Total	322	120	85	0.97	0.0090	−117.98*	0.07	335	5	0.66	0.71	3.75

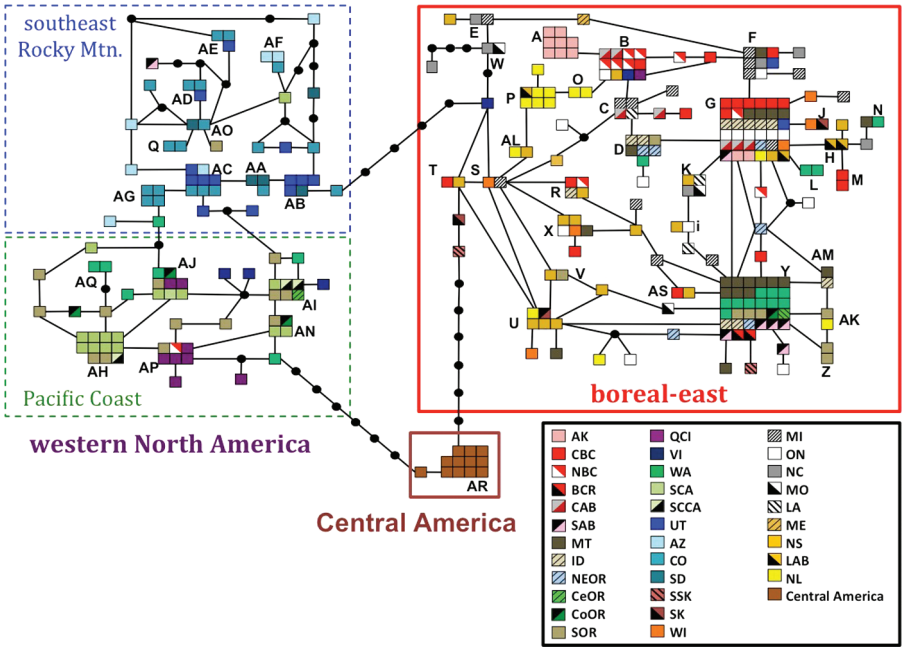

Figure 3. Haplotype network of the 120 mtDNA control region haplotypes (Supplementary material Appendix 1, 2) for hairy woodpeckers across their range. Squares represent a single individual and black dots represent inferred/missing haplotypes. Letter codes indicate the name of each haplotype as found in Supplementary material Appendix 1. Solid boxes show the three genetic lineages (western North America, boreal-east and Central America), with dashed boxes around the Pacific Coast and southeast Rocky Mountain groups showing subdivision within the western North America lineage. Locations of sampling sites are shown in Fig. 1.

$\pi = 0.0002$) where 11 of 12 individuals shared the same haplotype. Neither QCI or NL showed higher H_d or π as levels were comparable to neighbouring mainland populations.

Pairwise F_{ST} values revealed high levels of genetic structure across the range (Table 3, 4). Of the 120 pairwise F_{ST} values for mtDNA, 110 were significant, while 68 of the 91 pairwise F_{ST} values for microsatellites were significant after corrections for multiple tests using the false discovery rate. Sequential Bonferroni corrections showed similar patterns for mtDNA (90 of 120 are significant), but fewer F_{ST} values were significant for microsatellite data (44 of 91). Due to the strict nature of the Bonferroni correction many p-values were non-significant (e.g. QCI and NBC $F_{ST} = 0.07$, $p = 0.004$) because our critical p-value was close to zero ($p = 0.001$). Alaska is significantly different from all other populations for both mtDNA and microsatellites, while NL is significantly different from all populations for mtDNA and most populations (except NC, ON and NS) for microsatellites. QCI is significantly different from all populations for microsatellites and most populations (except SCA, SOR and VI) for mtDNA and SCA from all populations for mtDNA except QCI. For mtDNA Central America is significantly different from all North American groups and populations and pairwise F_{ST} values were high ranging from 0.57 to 0.93 (Table 3). Pairwise F_{ST} values reveal all four mtDNA groups are significantly different from each other ($p < 0.01$; Table 4). The global F_{ST} for microsatellites is 0.04, which is high given a theoretical global maximum F_{ST} of 0.05.

Plotting ΔK for STRUCTURE runs at each K using the full dataset produced three peaks at $K = 2$, $K = 6$ and $K = 8$. $K = 2$ produced the highest peak and is indicative of the overall hierarchical structure (Pr ln (K|D) = −7326.2). At both $K = 6$ and $K = 8$ the posterior probability (range = −7265.8 to −7800.3, Bayes factor ≤ 0.01) became more variable and had a lower ΔK value than at $K = 2$. At $K = 2$ individuals clustered into distinct east-west groups following the Great Plains in the south and approximately the Rocky Mountains in the north (all populations west of this line are hereby referred to as the west group and all populations east of the line are referred to as the east group; Fig. 2a). Several sampling sites (BCR, CBC, MT, NEOR, SAB and SD, Fig. 2a) showed high admixture between the east and west groups (< 60% assignment to either the east or west genetic cluster) suggesting a zone of contact between the two groups in British Columbia, the Pacific Northwest and further south on the Great Plains (Fig. 2a).

Within both the east and west groups we found distinct population substructure (Fig. 2b). In the east group at $K = 2$, Alaska was genetically distinct from all other populations in the boreal-eastern US group (Pr ln(K|D) = −3450.3; Bayes factor = 0.99). Additional runs excluding AK revealed no further population substructure. Within the west group we found more population substructure than in the east group. Posterior probability was highest at $K = 4$ (Pr ln(K|D) = −3776.96) and a Bayes factor of 0.92 confirmed four as the true K. The four distinct west subgroups are: the Queen Charlotte Islands, a Pacific Coast group west of the Cascade Mountains; and two interior groups, one in the interior west and a second in the southeast Rocky Mountains (Fig. 2b). The majority of individuals from a single population were assigned to one genetic cluster, however individuals in SOR, SCCA, CO and VI populations showed high admixture (Fig. 2b) and were assigned equally to more than one genetic cluster.

Table 3. Population pairwise F_{ST} values for 17 hairy woodpecker populations based on mtDNA (below diagonal) and microsatellites (above diagonal) data (only populations with $n \geq 9$ individuals were used). *indicates that pairwise comparisons are significant ($P_{critmtDNA} < 0.040$ and $P_{critmsat} < 0.036$) following Benjamini–Hochberg corrections. **indicates that pairwise comparisons are significant ($P_{critmtDNA} < 0.002$ and $P_{critmsat} \leq 0.001$) following both sequential Bonferroni and Benjamini–Hochberg corrections.

	AK	NL	CO	UT	SCA	QCI	SOR	WA	MT	ID	CBC	NBC	NS	NC	MI	ON	C Am
AK		0.12**	0.17**	0.13**	–	0.24**	0.17**	0.14*	0.12**	–	0.08**	0.12**	0.11**	0.13*	0.13**	0.12**	–
NL	0.44**		0.08**	0.06**	–	0.16**	0.05*	0.08**	0.04**	–	0.03*	0.06*	0.01	0.01	0.04*	0.03*	–
CO	0.81**	0.77**		0.01	–	0.12**	0.03*	0.06*	0.02*	–	0.04**	0.04*	0.07**	0.06**	0.09**	0.11**	–
UT	0.70**	0.65**	0.06*		–	0.10**	0.01	0.04	0.01	–	0.03*	0.01	0.05**	0.06**	0.09**	0.10**	–
SCA	0.86**	0.82**	0.52**	0.48**		–	–	–	–	–	–	–	–	–	–	–	–
QCI	0.78**	0.74**	0.52**	0.43**	0.07		0.07*	0.09**	0.10**	–	0.12**	0.07*	0.12**	0.15**	0.17**	0.12**	–
SOR	0.48**	0.42**	0.33**	0.21*	0.20*	0.10		0.02	0.02*	–	0.03*	0.00	0.03**	0.04**	0.06**	0.07**	–
WA	0.36**	0.27**	0.49**	0.34**	0.49**	0.38**	0.09		0.03*	–	0.03*	−0.01	0.04**	0.07*	0.09**	0.11**	–
MT	0.65**	0.45**	0.82**	0.72**	0.86**	0.81**	0.47**	0.19*		–	0.01	0.01	0.04**	0.02*	0.05**	0.05**	–
ID	0.57**	0.33**	0.80**	0.69**	0.85**	0.78**	0.42**	0.16*	0.02		–	–	–	–	–	–	–
CBC	0.40**	0.28**	0.77**	0.66**	0.81**	0.74**	0.43**	0.18*	0.11**	0.01		0.00	0.01	0.01	0.03*	0.03*	–
NBC	0.17*	0.20**	0.71**	0.56**	0.75**	0.62**	0.32**	0.18*	0.42**	0.28**	0.17*		0.02	0.04*	0.04*	0.03	–
NS	0.37**	0.17**	0.73**	0.61**	0.78**	0.70**	0.38**	0.15*	0.16*	0.10*	0.07*	0.13*		0.01	0.01	0.01	–
NC	–	–	–	–	–	–	–	–	–	–	–	–	–		0.01	0.01	–
MI	0.31**	0.21**	0.75**	0.63**	0.80**	0.71**	0.39**	0.18*	0.22**	0.11*	0.02	0.09*	0.05	–		0.02	–
ON	0.42**	0.28**	0.77**	0.66**	0.81**	0.74**	0.42**	0.18*	0.14**	0.00	−0.01	0.20*	0.11*	–	0.03		–
C Am	0.93**	0.86**	0.82**	0.75**	0.88**	0.82**	0.57**	0.62**	0.92**	0.93**	0.85**	0.80**	0.81**	–	0.85**	0.87**	

Table 4. Pairwise F_{ST} values (below diagonal) for the four mtDNA groups identified in TCS network (Fig. 2). All values are significant at $p = 0.05$ (p values above diagonal).

	Boreal-east	Pacific Coast	Southeast Rockies	Central America
Boreal-east		< 0.01	< 0.01	< 0.01
Pacific Coast	0.79		< 0.01	< 0.01
Southeast Rockies	0.76	0.54		< 0.01
Central America	0.78	0.82	0.78	

Results from BAPS differed slightly from those obtained using STRUCTURE, recognizing only four distinct clusters (Fig. 2a). While BAPS results were concordant with the subgroups identified by STRUCTURE in the east group (Alaska and boreal-eastern US), it only found two subgroups in the west group (QCI and all other west populations). BAPS also assigned CBC to the west group, whereas this population was assigned to the east group in STRUCTURE. Several individuals from CBC showed admixture between the east and west group (at $K = 2$, Fig. 2a) in our STRUCTURE runs and so it is not surprising that BAPS assigned CBC to the west group. CBC was included in the west for microsatellite IBD analyses (discussed below) as it was more similar to western populations than eastern populations based on pairwise F_{ST} values (Table 3).

Both Bayesian clustering analyses failed to separate NL from other boreal-east populations, despite F_{ST} values between NL and some boreal-east populations being significant. Bayesian models likely failed to distinguish Newfoundland from other populations as they perform best when $F_{ST} \geq 0.05$ (Latch et al. 2006) and F_{ST} values for Newfoundland and boreal-east populations were smaller (0.01 to 0.12).

Results for SAMOVA (Table 5) analyses revealed significant population structure within mtDNA groups. F_{CT} was highest at $K = 5$ ($F_{CT} = 60.62\%$; $p < 0.001$) although this value was only marginally higher than at $K = 4$ ($F_{CT} = 59.99\%$; $p < 0.001$). At $K = 5$, SAMOVA recognized two western groups (Pacific Coast and southeast Rockies; matching our ML tree (Fig. 1) and TCS network (Fig. 3)), Central America, the boreal-east and grouped CeOR, CoOR, SOR and VI (populations with both eastern and western haplotypes; Fig. 1) into a fifth group. In contrast to Bayesian clustering models used for microsatellite data, AK and QCI were not distinct clusters at $K = 5$ but grouped with the boreal-east and Pacific Coast groups respectively.

Demographic tests suggest the Central American group has been stable ($F_S = -0.48$, $p = 0.13$) while the three North American groups have undergone recent demographic expansion ($F_S = -15.05$ to -96.16, $p < 0.01$; Table 2). Within the boreal-east haplogroup, multiple populations (MT, ON, CBC, NS, NC, WI and MI) show recent demographic expansion, while two western populations, CO and SOR, show evidence of recent demographic expansion. F* and D* (not shown) were non-significant for all populations, indicating that recent population expansion and not background selection are responsible for the observed genetic patterns. Mismatch distributions (not shown) support demographic tests and suggest recent population expansion. We found a unimodal distribution for all

three North American groups and all but two populations (WA and SOR were both bimodal), and as probabilities for mismatch distributions and raggedness indices were also non-significant ($p > 0.05$) we cannot reject the null hypothesis of recent demographic expansion for these groups or populations.

Tests for expansion times indicate the southeastern Rocky Mountain group began expanding ~7 ka around the same time as the Pacific Coast and boreal-east groups ~6.6 and 8 ka, respectively. By comparison expansion in the central Rocky Mountains (ID and MT; ~3.2 and 3.7 ka, respectively), Washington (~4.7 ka), southern California (~2.6 ka), Utah (~2 ka) and Newfoundland (~4.5 ka) has occurred more recently. While demographic tests indicate recent population expansion, divergence times indicate the three main genetic lineages (west, boreal-east and Central America) separated from each other prior to or at the onset of Wisconsin glaciation. Central American populations diverged from North America ~99 to 104 ka, while the boreal-east group split from the western group ~96 (southeast Rocky Mountain group) to 113 ka (Pacific Coast group). Divisions within the western group occurred more recently with the southeast Rockies and Pacific Coast groups having split ~34 ka during the last glacial period.

Similar patterns are found at the population level as the closest Pacific Coast (SCA), southeast Rocky Mountain (AZ) and boreal-east (MO) populations diverged from Central America ~106, 99 and 84 ka respectively. Colorado has been isolated from the closest eastern populations (MO, MT and WI) between ~80 and 105 ka , while neighbouring Pacific Coast (SCA and SOR) and southeast Rockies (UT) populations separated more recently ~34 ka. Estimates of divergence between Pacific Coast and boreal-east populations (WA vs ID and SOR vs NEOR) suggest recent divergence ~12 to 46 ka, but lower divergence times are likely due to secondary contact between Pacific Coast and boreal-east haplogroups in this region. The Queen Charlotte Islands birds have been isolated from NBC ~80 ka, while Alaska and Newfoundland diverged more recently from their closest populations (CBC and NS) ~16 ka and 6 ka, respectively.

Effects of barriers on genetic variation

Testing for isolation by distance (IBD) across the whole North American range showed a weak but significant relationship for microsatellite data only ($r^2 = 0.11$, $p = 0.02$; mtDNA: $r^2 = 0.006$, $p = 0.22$). As QCI and AK are both significantly different from most of the other populations (Table 3) and located at the edge of the range (i.e. large geographic distances to other sampling sites), we removed both populations from the analysis to see if these two populations influence IBD for the whole range. Patterns of IBD strengthened and remained significant when we did this for microsatellites ($r^2 = 0.20$, $p < 0.01$; mtDNA: $r^2 = 0.002$, $p = 0.22$), but the relationship between genetic and geographic distance is still weak. We also tested for IBD within the east and west groups and found contrasting patterns for microsatellites and mtDNA. There is a strong and significant relationship for IBD within the east for microsatellites ($r^2 = 0.92$, $p = 0.02$ and weak but non-significant pattern for mtDNA: $r^2 = 0.05$, $p = 0.07$), while results are significant within the west for mtDNA ($r^2 = 0.23$, $p = 0.01$) but not for microsatellites ($r^2 = 0.11$, $p = 0.11$). We performed the test for the east and west separately removing AK and QCI (respectively) to see if these populations influence IBD patterns. When Alaska was removed, we found a weak and non-significant relationship in the east group for both microsatellites ($r^2 = 0.19$, $p = 0.12$) and mtDNA ($r^2 = 0.005$, $p = 0.29$). Removing QCI did not change results within the west for mtDNA ($r^2 = 0.27$, $p = 0.003$), but we did find a significant relationship for microsatellites ($r^2 = 0.32$, $p = 0.01$).

BARRIER identified six main barriers to gene flow (Fig. 2b). When we compared pairwise F_{ST} values between populations on either side of these barriers, they were significantly different ($p = 0.001$ to 0.010) with the exception of populations on either side of the barrier separating North Carolina from Missouri ($p = 0.328$). Comparisons between Alaska and the nearest continental sampling sites (CAB, CBC and NBC) show high F_{ST} values ($F_{ST} = 0.08$–0.12), but these areas are separated by large geographic distances. Without additional sampling sites between Alaska and these areas, it is difficult to tell if Alaska is completely isolated due to physical barriers like the Alaskan Mountain Ranges.

Estimates for migration using BAYESASS+ showed little to no gene flow between either AK or QCI and any other sampling site supporting pairwise F_{ST} values which identified these two populations as being significantly different from all other populations. Gene flow was detected (Table 6) between sites separated by mountains with asymmetrical gene flow from MT to CBC and NBC; WA to SOR and UT; and from CO to UT. Of the migration rate estimates > 0.10, all involved NS, WA, MT or CO (Table 6). The confidence intervals for these estimates are large, so caution should be taken not to over interpret the values.

Table 5. Results from SAMOVA analyses estimated for mtDNA analyses. Presented are the three groups with the highest F_{CT} as suggested by Dupanloup et al. (2002). At K = 5 four populations with individuals from both eastern and western haplogroups (Fig. 3) were recognized as a separate group. WA which also includes individuals from both eastern and western haplogroups grouped with the boreal-east at K = 3 to 5.

		Source of variation		
Group	K	Among groups (F_{CT})	Among pop. within groups	Within pop.
1) Boreal-east 2) Western North America 3) Central America	3	57.07%	9.76%	33.17%
1) Boreal-east 2) Central America 3) CoOR, QCI, SCA, SCCA, SOR, VI 4) AZ, CO, SD, UT	4	59.99%	6.06%	33.95%
1) Boreal-east 2) Central America 3) QCI, SCA, SCCA 4) AZ, CO, SD, UT 5) CeOR, CoOR, SOR, VI	5	60.62%	4.87%	34.51%

Table 6. Results from BAYESASS+. Proportion of migrants in the each population (Pop) and their source (migrants). Populations that are not listed have no migrants. Where the total for a row is <1.00, the remaining individuals were not assigned to a population (NA).

Pop	Non-migrants (95% CI)	Migrants (95% CI) and source population
AK	0.94 (0.67, 1.00)	NA
NL	0.71 (0.67, 0.79)	NS 0.16 (0.05, 0.28)
CO	0.97 (0.90, 1.00)	NA
UT	0.69 (0.67, 0.74)	CO 0.11 (0.02, 0.23); WA 0.08 (0.00, 0.22); MT 0.04 (0.00, 0.21)
QCI	0.96 (0.87, 1.00)	NA
SOR	0.70 (0.67, 0.81)	WA 0.10 (0.00, 0.29); CO 0.06 (0.00, 023); MT 0.05 (0.00, 0.24)
WA	0.91 (0.67, 1.00)	MT 0.04 (0.00, 0.26)
MT	0.80 (0.67, 1.00)	WA 0.07 (0.00, 0.23); CO 0.05 (0.00, 0.18); NS 0.04 (0.00, 0.12)
CBC	0.69 (0.67, 0.73)	NS 0.17 (0.02, 0.29); MT 0.05 (0.00, 0.25); WA 0.02 (0.00, 0.12)
NBC	0.71 (0.67, 0.81)	MT 0.05 (0.00, 0.20); CO 0.04 (0.00, 0.14); WA 0.04 (0.00, 0.16); NS 0.04 (0.00, 0.14)
NS	0.97 (0.90, 1.00)	NA
NC	0.69 (0.67, 0.73)	NS 0.21 (0.09, 0.31)
MI	0.74 (0.67, 0.87)	NS 0.16 (0.04, 0.29)
ON	0.71 (0.67, 0.75)	NS 0.19 (0.06, 0.29)

Ecological niche modelling

Ecological niche models were constructed using 3125 training points and 347 testing points. The current distribution predicted by MAXENT (Fig. 4a) matched the known range of hairy woodpeckers (Fig. 1) and the model performed well as it had a high AUC value (0.863) and curves for both the training and test sample omission were close to the predicted omission curve. Predicted distributions 21 ka showed large range retraction with hairy woodpeckers in North America being restricted to three main areas south of the ice sheets (Fig. 4b). Along the Pacific Coast potential habitat extended primarily from southern California to Washington with smaller amounts of habitat along the British Columbian Coast. In the southeast Rockies, predicted habitat was centralized in New Mexico and Arizona and extended south into Mexico through the Sierra Madre Occidental Mountain Range. In the southeast United States potential habitat was present from eastern Texas in the southeast to Virginia in the northeast (Fig. 4b). The model also predicted potential habitat in Alaska (≤ 0.40 probability), the Hecate Strait (QCI putative refugium; ≤ 0.50) and Newfoundland (putative refugium ≤ 0.40) but probabilities were lower in comparison to predictions for areas south of the ice sheets (0.50 to 0.90). We did not include paleodata for Central America and are therefore unable to suggest where hairy woodpecker populations were potentially located in Central America 21 kya.

Discussion

MtDNA and microsatellite analyses show multiple geographically structured groups and high levels of genetic variation for hairy woodpeckers. We found a main east-west division as observed in other widely distributed North American species (Milot et al. 2000, Ruegg and Smith 2002, Clegg et al. 2003, Godbout et al. 2005, Boulet and Gibbs 2006, Turmelle et al. 2011). Microsatellite analyses revealed greater population structure within both groups, contrary to other widely distributed bird species like the red-tailed hawk, which only shows reduced gene flow in the western part of its range (Hull et al. 2008), and the yellow warbler, which shows unrestricted gene flow across its range (Gibbs et al. 2000). MtDNA analyses revealed three divergent haplogroups (west (Pacific and SE Rockies), boreal-east and Central America). Greater population structure for both mtDNA and microsatellite data is present in the west, which reflects barriers like mountains and large expanses of open water being more prevalent than in the east, where relatively few barriers exist. Data suggest hairy woodpeckers were

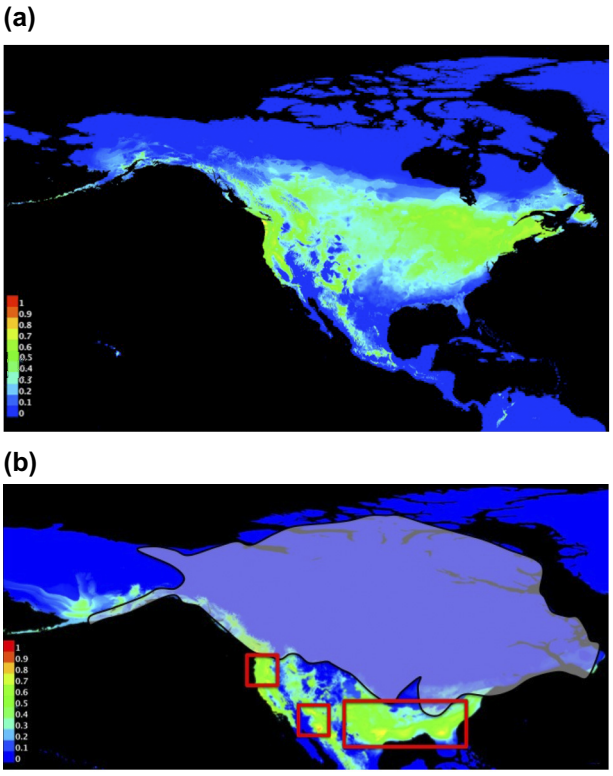

Figure 4. Ecological niche modelling predictions from MAXENT based on paleoclimate data. Current (a) and historical (b) distribution (~21 ka) for hairy woodpeckers are indicated where cooler colours (dark blue) indicate low prediction rates for presence and warmer colours (yellow to red) indicate higher prediction rates for presence of hairy woodpeckers. Locations of forested refugia from Williams (2003) are outlined in red boxes (b) and white overlay represents ice sheets at last glacial maximum (~21 ka).

isolated in multiple southern refugia and recolonized northern areas through multiple colonization routes with subsequent reduction in gene flow caused by barriers and behaviour. While Pleistocene glacial cycles have promoted high genetic structure in other sedentary bird species (Barrowclough et al. 2004, Manthey et al. 2011), the high levels of genetic structure for hairy woodpeckers contrasts patterns seen in other sedentary woodpecker species (Ball et al. 1998, Zink et al. 2002, Pulgarin-Restrepo 2011).

Refugia and historical isolation

The presence of three major genetic lineages (boreal-east, west and Central America) coupled with ENM results suggest hairy woodpeckers were isolated south of the ice sheets during the last glacial in multiple refugia. These three lineages diverged during the Pleistocene ~98 to 113 ka with subsequent divergence in the west (Pacific Coast and southeast Rockies) occurring ~34 ka towards the end of the last glacial. While the timing of the splits differs from those found by Klicka et al. (2011), the order of the splits between the three main lineages is similar. Within the east, genetic patterns suggest hairy woodpeckers were potentially isolated in multiple areas (Fig. 4b), as has been suggested for eastern tree and animal species (reviewed by Godbout et al. 2005, 2010, Soltis et al. 2006, de Lafontaine et al. 2010). Hairy woodpeckers from Atlantic Canada are distinct from other populations in eastern Canada suggestive of multiple colonizations as discussed below.

Multiple refugia are also hypothesized for species in the western US (Byun et al. 1997, Arbogast et al. 2001, Brunsfeld et al. 2001, Gugger et al. 2010). Population genetic structure in the west (Table 3, Fig. 1; Klicka et al. 2011) and ENM support hairy woodpeckers being isolated in two separate western refugia (Pacific Coast and southeast Rockies; Fig. 4b). Predictions are supported by fossil data as fossils for hairy woodpeckers have been found in California (Miller and de May 1942), Alabama (Parmalee 1992) and Florida (Emslie 1998) areas included in our niche modelling predictions. While we do not have ENM predictions for Central America, the distinct and divergent mtDNA group (this study; Klicka et al. 2011) strongly suggests this population has been isolated from North American populations for a prolonged period.

ENM predicted potential habitat in Alaska (Anchorage region), the Queen Charlotte Islands and Newfoundland during the LGM (Fig. 4b). However, genetic patterns do not support any of these areas acting as a refugium throughout the entire Pleistocene. If any of these areas had acted as a long-term refugium we would predict: 1) divergent haplotypes restricted to Alaska, Queen Charlotte Islands or Newfoundland (refugia with no subsequent dispersal) or 2) higher haplotype diversity in refugial populations with subsets of AK, NL or QCI haplotypes found in previously glaciated areas (refugia with subsequent gene flow). While haplotype diversities are high in QCI and NL ($H_d = 0.81$ and 0.86, respectively), they are comparable to most of the other sampling sites, including those from previously glaciated areas. The large number of unique haplotypes and clusters of shared haplotypes (e.g. P, O, U and AL) restricted to Atlantic Canada (LAB, NL and NS) support their genetic isolation. However, these populations have not been isolated long enough for lineage sorting to be complete. It is possible that birds dispersed to this area from an eastern refugium some time during the Pleistocene and remained isolated since, raising the idea of multiple expansions out of an eastern refugium. Both Newfoundland and Queen Charlotte Islands acted as refugia for other species (Gill and Mostrom 1993, Byun et al. 1997, Burg et al. 2005, 2006, Godbout et al. 2010) and may have been part of larger refugia for hairy woodpeckers either on the Pacific or Atlantic Coast during the latter part of the Pleistocene.

The reduced genetic variation in AK ($H_d = 0.36$) is likely the result of a recent bottleneck or founder effect as eight of the ten birds sampled contain the same haplotype. If the reduced variation was the result of prolonged isolation in a refugium; the haplotypes should be more divergent (Fig. 3). Similarly, allelic patterns of birds from these areas have similar or slightly lower genetic diversity (Table 2) than other populations in their respective groups and contain a subset of alleles found within the larger groups. Pairwise F_{ST} values do support reduced gene flow between these three putative refugial populations; however, this is likely the result of recent rather than historical isolation.

Patterns of colonization following glaciation

Patterns of genetic differentiation suggest recent expansion from ice-free refugia followed the phalanx model (Hewitt 1996). While we found significant genetic differences between populations of hairy woodpeckers, we did not find pockets of highly divergent, isolated populations, which is what we would have expected if hairy woodpeckers had expanded under the pioneer model. Admixture of boreal-east and Pacific Coast haplotypes in the Cascade Mountains (Fig. 1) is similar to the pattern found by Klicka et al. (2011). Even with North America being colonized following the phalanx model, we found distinct genetic structure across the hairy woodpecker's range. The presence of geographically distinct clades across the range suggests decreased gene flow following colonization, as with other high latitude bird species (Holder et al. 1999, Pruett and Winker 2008).

Inclusion of samples from previously glaciated areas suggests expansion from multiple refugia and does not support a single, widespread expansion from an eastern refugium, as suggested by Klicka et al. (2011). In addition, haplotype composition and the number of private haplotypes in Atlantic Canada (LAB, NL and NS; Table 2, Fig. 3) suggest two separate movements out of the eastern refugium. Furthermore two common haplotypes (G and Y; Fig. 3; Supplementary material Appendix 2) sort north (CAB, CBC and ON) and south (MT, SAB and WA) geographically within the west, suggesting alternative colonization routes of previously glaciated areas in this region.

Central American birds are distinct (Table 3, 4, Fig. 1; Klicka et al. 2011) from those in North America. Genetic diversity is low in Central America, but this may be the result of sample size. More detailed sampling of Mexico and Central America by Klicka et al. (2011) found more haplotypes in Mexico; however, only a single ND2 haplotype in Costa Rica and Panama birds and an additional two

haplotypes in Guatemala. Post-glacial expansion in North America was exclusively from North American refugia. We found no evidence for post-glacial expansion into North America from Central America, as has been shown for other widely distributed bird species like the chipping sparrow *Spizzela passerina* (Milá et al. 2006) and dark-eyed junco *Junco hyemalis* (Milá et al. 2007b).

Contemporary and historic gene flow

Microsatellite and mtDNA data offer different interpretations of the evolutionary history and gene flow. Microsatellites are able to show recent events while mtDNA are often used to examine historical processes. Differences between these markers are present in the west, at contact zones and in patterns of gene flow.

Microsatellite data show the west group (Fig. 2b) extends further into BC and the eastern Rockies (NBC and MT) than suggested by western haplotypes which are restricted to the Pacific Coast (as far north as QCI) and absent from MT (Fig. 1). Despite the high levels of population structure and sedentary nature of hairy woodpeckers, we see several examples of female gene flow (Fig. 3). A single bird from SAB contains a haplotype typical of the SE Rockies group, two Oregon birds have haplotypes from the boreal-east and a NBC bird has a haplotype from the Pacific Coast group. STRUCTURE analyses do not show evidence of mixed ancestry for these birds and as mtDNA is uniparentally inherited these could be due to historical rather than contemporary movements. Indeed if gene flow was high, we would not see the high levels of population structure between groups and high levels of population differentiation within groups. A second example of admixture occurs in the Pacific Northwest where populations contain both boreal-east and Pacific Coast haplotypes (Fig. 1). This contact zone between east and west genetic groups of hairy woodpeckers resembles those in other species (Rohwer and Wood 1998, Brelsford and Irwin 2009, Irwin et al. 2009, Rush et al. 2009, Toews et al. 2011).

Patterns with mtDNA and microsatellite data are different for both Alaska and Queen Charlotte Islands populations. Whereas SAMOVA analyses on mtDNA data group these two populations with the boreal-east and Pacific Coast haplogroups respectively, they are distinct based on STRUCTURE and BAPS microsatellite analyses. Reduced gene flow within the last few generations likely explains some contrasting patterns for mtDNA and microsatellites. Microsatellite markers show higher levels of structure in the west; the Pacific Coast mtDNA group is divided into three microsatellite groups (QCI, Interior West and Pacific).

Overall major patterns in unglaciated parts of North and Central America are similar to a recent study (Klicka et al. 2011); however, a few differences exist. Both mtDNA datasets show significant pairwise F_{ST} values in the west, but only the control region data show evidence of restricted gene flow in the east. The control region data from our study do not show a starburst pattern in the eastern US. Lower mutation rates in coding genes compared to non-coding genes (Milá et al. 2007a) may be one reason for differences between studies. The starburst pattern in the ND2 data could be a relic of historical patterns, which is no longer evident in the control region network due to mutations.

Along the contact zones, several populations with eastern mtDNA populations (ID, NBC, MT, and WA) are more similar to western populations based on microsatellite data. Whereas pairwise F_{ST} values are large between these populations for mtDNA, they are small for microsatellites (Table 3). One explanation for this is short distance dispersal by males. Any movements of males would go undetected with mtDNA, unless first generation migrant, but would be detected with microsatellite data and would homogenize nuclear loci.

Barriers

The hairy woodpecker's range encompasses a number of dispersal barriers including physical distance, large bodies of water, mountains and breaks in continuous habitat. Genetic patterns based on mtDNA and in particular microsatellite data correspond to a number of these features.

Genetic differences over large geographic distances are likely more pronounced, as the sedentary nature of the hairy woodpecker has likely aided in restricting gene flow throughout large portions of the range. Rangewide patterns of IBD are weak but significant for nuclear loci and suggest that dispersal is restricted by physical distance in the absence of barriers (e.g. within the east group) with some exceptions, most notably NS (Table 6). In the west maternal gene flow is restricted by geographic distance as is nuclear gene flow when QCI birds are excluded. Differences in patterns between the two types of markers where microsatellites show stronger patterns are likely the result of contemporary patterns of gene flow which are better detected with microsatellite loci.

Three pairs of populations are separated by large bodies of water from other nearby populations. QCI is isolated from northern BC by a large expanse of open water, Hecate Strait (~80 km), while the Strait of Georgia separates Vancouver Island (VI) from the mainland coast (<35 km) and the Strait of Belle Isle separates the island of Newfoundland from Labrador (<20 km). QCI individuals have reduced genetic diversity compared to populations with similar sample sizes (Table 2) and patterns of genetic variation between the Queen Charlotte Islands and northern BC populations on the mainland are distinct for both mtDNA and microsatellites (Fig. 1, 2b, Table 3). While sample sizes for VI ($n=5$) are limited, based on mtDNA F_{ST} (data not shown) and conservative Bayesian clustering programs (Table 3, Fig. 2a, 3), it is not genetically distinct from the nearest sampled mainland population of WA. A number of forested islands present between Vancouver Island and the mainland may aid in dispersal. In contrast no small islands are present between NL and the mainland, yet F_{ST}, and Bayesian clustering analyses group it with NS despite low gene flow relative to other sites. Our results suggest large open expanses of water like that between QCI and the mainland act as complete barriers to gene flow, whereas smaller expanses of open water do not.

The hairy woodpecker's range is delineated by a number of prominent mountain ranges, including the younger and

taller Rocky and Cascade Mountain ranges in the west and the older and lower Appalachian Mountain range in the east. While our results do suggest that distributions of subgroups in the west coincide with or are specific to mountainous regions, there is no evidence to suggest that mountains themselves are acting as complete barriers to gene flow. Mountains do restrict gene flow between some northern populations (NBC and CBC; ID and CAB; Fig. 2a, b), yet gene flow is occurring between other populations (CBC and CAB; UT and CO; AZ and CO) separated by tall mountains. If mountains were acting as complete barriers to gene flow, we would expect to see populations west of the mountains being significantly different from populations east of the mountains. Hairy woodpeckers are viewed as generalists, as they have the ability to live in a wide range of forested habitats and at a range of elevations. Their ability to survive at high elevations (up to 1900 m in British Columbia and 3500 m in New Mexico) and disperse short distances (Jackson et al. 2002) may enable them to move through valleys and mountain passes within mountainous regions.

Genetic clusters within North America are concordant with discontinuities in contemporary forested habitat. The east-west division follows the Great Plains, where grassland species are dominant and contiguous forest is rare (Barker and Whitman 1988). Furthermore, populations with admixture between the east and west groups occur in the Pacific Northwest (for both mtDNA and microsatellites) and central British Columbia (microsatellites only) where boreal forest tree species meet western montane forest tree species. Areas like the Columbia and Harney Basins in Washington and Oregon separate the Cascades from the central Rocky Mountains, and hairy woodpeckers, due to the absence of forest vegetation, inhabit neither of these basins. Significant genetic differences occur between WA and MT ($F_{STmsat} = 0.03$, $p = 0.004$; $F_{STmtDNA} = 0.19$, $p = 0.004$), WA and ID ($F_{STmtDNA} = 0.16$, $p = 0.03$) and SOR and NEOR ($F_{STmtDNA} = 0.35$, $p = 0.013$) over a relatively short distance (200–500 km) suggesting the Columbia and Hearny Basins may be acting as barriers to gene flow. Similar patterns are present for mesic forest species (Carstens et al. 2005, Burg et al. 2006) where interior populations are genetically distinct from coastal populations. Discontinuities in forested habitat are believed to be the reason for genetic patterns within this region (Nielson et al. 2001).

The Wyoming Basin, an area characterized by shrub steppe and grasslands (Driese et al. 1997), separates the southeastern Rocky Mountains from Montana and the northern Rocky Mountains. To the west of the Wyoming Basin are the Rocky Mountains and to the east are the Great Plains. The Wyoming Basin may act as a barrier to gene flow as populations of hairy woodpeckers on either side of the Wyoming Basin (MT and CO) are genetically distinct from each other (Table 3). Currently contiguous forest is fragmented by the Wyoming Basin (Driese et al. 1997), while during the last glacial, forested habitat would have been fragmented by paleogeographic barriers (Porter et al. 1983, Hafner and Sullivan 1995). Patterns of genetic differentiation between the central Rocky Mountain and southeastern Rocky Mountain populations are observed for both forest trees (Aagaard et al. 1995) and forest dwelling species (Arbogast and Kenagy 2001), as well as for other habitat specialists (Demboski and Cook 2001, DeChaine and Martin 2005, Galbreath et al. 2010) and isolation of forested habitats has been used to explain genetic patterns within this region (Hafner and Sullivan 1995, Arbogast 1999).

Endemic subspecies and island populations

Both the Queen Charlotte Islands and Newfoundland are home to endemic subspecies of hairy woodpeckers (*P. v. picoideus* and *P. v. terranovae*, respectively). As is the case with other species (Questiau et al. 1999), the distinct morphological differences likely arose more recently. Genetic differentiation of hairy woodpeckers on the Queen Charlotte Islands supports previous work documenting morphology, plumage and genetic differences (Miller et al. 1999, Topp and Winker 2008) and previous work showing distinct genetic differentiation for other bird species found in QCI (Burg et al. 2005, 2006, Topp and Winker 2008). Our study is the first genetic study on the Newfoundland subspecies of hairy woodpeckers. The results suggest that while these birds are genetically different from most populations (Table 3), genetic differentiation is unlikely to be the result of long-term isolation. First individuals from Newfoundland are not significantly different from nearby continental populations and second Newfoundland does not contain any private alleles (Table 2). Previous studies have demonstrated similar patterns for other plant and animal species, where Newfoundland populations are genetically similar to continental populations despite reduced gene flow (Paetkau and Strobeck 1996, Rajora et al. 1998), suggesting more recent colonization of this region.

Acknowledgements – We would like to thank Jenny Woolf, the American Museum of Natural History, Burke Museum, Field Museum of Natural History, Louisiana Museum of Zoology, New Brunswick Museum, Royal British Columbia Museum, Royal Alberta Museum, Royal Ontario Museum, Royal Saskatchewan Museum and Smithsonian for blood and tissue samples. In addition, field assistance was provided by Kim Dohms, John Hindley, Elysha Koran, Linda Lait, Paulo Pulgarin-Restrepo, numerous field assistants and participants of Project FeederWatch in both Canada and the United States. Staffan Bensch provided helpful comments on previous drafts. Funding was provided by Alberta Innovates Technologies Futures, NSERC DG and the Univ. of Lethbridge.

References

Aagaard, J. E., Vollmer, S. S., Sorensen, F. C. and Strauss, S. H. 1995. Mitochondrial DNA products among rapd profiles are frequent and strongly differentiated between races of Douglas-fir. – Mol. Ecol. 4: 441–446.

Arbogast, B. S. 1999. Mitochondrial DNA phylogeography of the New World flying squirrels (*Glaucomys*): implications for Pleistocene biogeography. – J. Mammal. 80: 142–155.

Arbogast, B. and Kenagy, G. J. 2001. Comparative phylogeography as an integrative approach to historical biogeography. – J. Biogeogr. 28: 819–825.

Arbogast, B. S., Brown, R. A. and Weigl, P. D. 2001. Genetics and Pleistocene biogeography of North American tree squirrels (*Tamiasciurius*). – J. Mammal. 82: 302–319.

Arguedas, N. and Parker, P. G. 2000. Seasonal migration and genetic population structure in house wrens. – Condor 102: 517–528.
Avise, J. C., Walker, D. and Johns, G. C. 1998. Speciation durations and Pleistocene effects on vertebrate phylogeography. – Proc. R. Soc. B 265: 1707–1712.
Baker, A. and Marshall, H. 1997. Mitochondrial control region sequences as tools for understanding evolution. – In: Mindell, D. (ed.), Avian molecular evolution and systematics. Academic Press, pp. 49–80.
Ball, R. M., Freeman, S., James, F. C., Bermingham, E. and Avise, J. C. 1988. Phylogeographic population structure of red-winged blackbirds assessed by mitochondrial DNA. – Proc. Natl Acad. Sci. USA 85: 1558–1562.
Barker, W. T. and Whitman, W. C. 1988. Vegetation of the northern Great Plains. – Rangelands 10: 266–272.
Barrowclough, G. F., Groth, J. G., Mertz, L. A. and Gutiérrez, R. J. 2004. Phylogeographic structure, gene flow and species status in blue grouse (*Dendragapus obscurus*). – Mol. Ecol. 13: 1911–1922.
Belkhir, K. 1999. genetix, version 4.0. A windows program for population genetic analysis. – Laboratoire Génome, Populations, Interactions, Univ. Montpellier, Montpellier, France.
Benjamini, Y. and Hochberg, Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. – J. R. Stat. Soc. B 57: 289–300.
Boulet, M. and Gibbs, H. L. 2006. Lineage origin and expansion of a neotropical migrant songbird after recent glaciation events. – Mol. Ecol. 15: 2505–2525.
Brelsford, A. and Irwin, D. E. 2009. Incipient speciation despite little assortative mating: the yellow-rumped warbler hybrid zone. – Evolution 63: 3050–3060.
Brunsfeld, S. J., Sullivan, J., Soltis, D. E. and Soltis, P. S. 2001. A comparative phylogeography of northwestern North America: a synthesis. – In: Silvertown, J. and Antonovics, J. (eds), Integrating ecology and evolution in a spatial context. Blackwell, pp. 319–344.
Burg, T. M. and Croxall, J. P. 2001. Global relationships amongst black-browed and grey-headed albatrosses: analysis of population structure using mitochondrial DNA and microsatellites. – Mol. Ecol. 10: 2647–2660.
Burg, T. M., Gaston, A. J., Winker, K. and Friesen, V. L. 2005. Rapid divergence and postglacial colonization in western North American Steller's jays (*Cyanocitta stelleri*). – Mol. Ecol. 14: 3745–3755.
Burg, T. M., Gaston, A. J., Winker, K. and Friesen, V. L. 2006. Effects of Pleistocene glaciations on population structure of North American chestnut-backed chickadees. – Mol. Ecol. 15: 2409–2419.
Byun, S. A., Koop, B. F. and Reimchen, T. E. 1997. North American black bear mtDNA phylogeography: implications for morphology and the Haida Gwaii glacial refugium controversy. – Evolution 51: 1647–1653.
Cabin, R. J. and Mitchell, R. J. 2000. To Bonferroni or not to Bonferroni? When and how are the questions. – Bull. Ecol. Soc. Am. 81: 246–248.
Carstens, B. C., Brunsfield, S. J., Demboski, J. R., Good, J. M. and Sullivan, J. 2005. Investigating the evolutionary history of the Pacific Northwest forest ecosystem: hypothesis testing within a comparative phylogeographic framework. – Evolution 59: 1639–1652.
Clegg, S. M., Kelly, J. F., Kimura, M. and Smith, T. B. 2003. Combining genetic markers and stable isotopes to reveal population connectivity and migration patterns in a Neotropical migrant Wilson's warbler (*Wilsonia pusilla*). – Mol. Ecol. 12: 819–830.
Clement, M., Posada, D. and Crandall, K. A. 2000. TCS: a computer program to estimate gene genealogies. – Mol. Ecol. 9: 1657–1659.
Corander, J., Waldmann, P. and Sillanpaa, M. 2003. Bayesian analysis of genetic differentiation between populations. – Genetics 163: 367–374.
Corander, J., Siren, J. and Arjas, E. 2008. Bayesian spatial modelling of genetic population structure. – Comput. Stat. 23: 111–129.
de Lafontaine, G., Turgeon, J. and Payette, S. 2010. Phylogeography of white spruce *Picea glauca* in eastern North America reveals contrasting ecological trajectories. – J. Biogeogr. 37: 741–751.
DeChaine, E. G. and Martin, A. P. 2005. Marked genetic divergence among sky island populations of *Sedium lanceolatum* Crassulaceae in the Rocky Mountains. – Am. J. Bot. 92: 477–486.
Demboski, J. R. and Cook, J. A. 2001. Phylogeography of the dusky shrew, *Sorex monticolus* (Insectivora, Soricidae): insight into deep and shallow history in northwestern North America. – Mol. Ecol. 10: 1227–1240.
Driese, K. L., Reiners, W. A., Merrill, E. H. and Gerow, K. G. 1997. A digital land cover map of Wyoming, USA: a tool for vegetation analysis. – J. Veg. Sci. 8: 133–146.
Dupanloup, I., Schneider, S. and Excoffier, L. 2002. A simulated annealing approach to define the genetic structure of populations. – Mol. Ecol. 11: 2571–2581.
Ellegren, H., Carlson, A. and Stenberg, I. 1999. Genetic structure and variability of white-backed woodpecker (*Dendrocopos leucotos*) populations in northern Europe. – Hereditas 130: 291–299.
Emslie, S. D. 1998. Avian community, climate and sea-level changes in the Plio-Pleistocene of the Florida Peninsula. – Ornithol. Monogr. 50.
Evanno, G., Regnaut, S. and 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. and Schneider, S. 2005. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. – Evol. Bioinform. Online 1: 47–50.
Faubet, P., Waples, R. S. and Gaggiotti, O. E. 2007. Evaluating the performance of a multilocus Bayesian method for the estimation of migration rates. – Mol. Ecol. 16: 1149–1166.
Fike, J. A., Athrey, G., Bowman, R., Leberg, P. L. and Rhodes, Jr O. E. 2009. Development of twenty-five polymorphic microsatellite markers for the endangered red-cockaded woodpecker *Picoides borealis*. – Conserv. Genet. 10: 1021–1023.
Fu, Y. X. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. – Genetics 147: 915–925.
Fu, Y. X. and Li, W. H. 1993. Statistical tests of neutrality of mutations. – Genetics 133: 693–709.
Galbreath, K. E., Hafner, D. J., Zamudio, K. R. and Agnew, K. 2010. Isolation and introgression in the Intermountain West: contrasting gene genealogies reveal the complex biogeographic history of the American pika (*Ochotona princeps*). – J. Biogeogr. 37: 344–362.
Garamszegi, L. Z. 2006. Comparing effect sizes across variables: generalization without the need for Bonferroni correction. – Behav. Ecol. 17: 682–687.
Gerlach, G. and Musolf, K. 2000. Fragmentation of landscape as a cause for genetic subdivision in bank voles. – Conserv. Biol. 14: 1066–1074.
Gibbs, H. L., Dawson, R. J. G. and Hobson, K. A. 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, F. B. and Mostrom, A. M. 1993. Speciation in North American chickadess: I. Patterns of mtDNA divergence. – Evolution 47: 195–212.
Godbout, J., Jaramillo-Correa, J. P., Beaulieu, J. and Bousquet, J. 2005. A mitochondrial DNA minisatellite reveals the

postglacial history of jack pine *Pinus banksiana*, a broad-range North American conifer. – Mol. Ecol. 14: 3497–3512.
Godbout, J., Beaulieu, J. and Bousquet, J. 2010. Phylogeographic structure of jack pine *Pinus banksiana*; Pinaceae supports the existence of a coastal glacial refugium in northeastern North America. – Am. J. Bot. 97: 1903–1912.
Goudet, J. 1995. FSTAT version 1.2: a computer program to calculate F-statistics. – J. Hered. 86: 485–486.
Gugger, P. F., Sugita, S. and Cavender-Bares, J. 2010. Phylogeography of Douglas-fir based on mitochondrial and chloroplast DNA sequences: testing hypotheses from the fossil record. – Mol. Ecol. 19: 1877–1897.
Guindon, S. and Gascuel, O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum liklihood. – Syst. Biol. 52: 696–704.
Hafner, D. J. and Sullivan, R. M. 1995. Historical and ecological biogeography of Nearctic pikas (Lagomorpha: Ochotonidae). – J. Mammal. 76: 302–321.
Harding, R. M. 1996. Lines of descent from mitochondrial Eve: an evolutionary look at coalescence. – In: Donnelly, P. and Tavaré, S. (eds), Progress in population genetics and human evolution. Springer, pp. 15–32.
Harpending, H. C. 1994. Signature of ancient populations growth in a low-resolution DNA mismatch distribution. – Hum. Biol. 66: 591–600.
Hedrick, P. W. 1999. Highly variable loci and their interpretation in evolution and conservation. – Evolution 53: 313–318.
Hewitt, G. M. 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. – Biol. J. Linn. Soc. 58: 247–276.
Hewitt, G. M. 2000. The genetic legacy of the Quaternary ice ages. – Nature 405: 907–913.
Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. and Jarvis, A. 2005. Very high resolution interpolated climate surfaces for global land areas. – Int. J. Climatol. 25: 1965–1978.
Holder, K., Montgomerie, R. and Friesen, V. L. 1999. A tests of the glacial refugium hypothesis using patterns of mitochondrial and nuclear DNA sequence variation in rock ptarmigan (*Lagopus mutus*). – Evolution 53: 1936–1950.
Hull, J. M., Hull, A. C., Sacks, B. N., Smith, J. P. and Ernest, H. B. 2008. Landscape characteristics influence morphological and genetic differentiation in a widespread raptor (*Buteo jamaicensis*). – Mol. Ecol. 17: 810–824.
Irwin, D. E., Brelsford, A., Toews, D. P. L., MacDonald, C. and Phinney, M. 2009. Extensive hybridization in a contact zone between MacGillivray's warblers *Oporornis tolmiei* and mourning warblers *O. philadelphia* detected using molecular and morphological analyses. – J. Avian Biol. 40: 539–552.
Jackson, J. A., Ouellet, H. R. and Jackson, B. J. S. 2002. Hairy woodpecker. – In: Poole, A. and Gill, F. (eds), The birds of North America. The Academy of Natural Sciences/The American Ornithologists' Union, Philadelphia/Washington, DC.
Jarmillo-Correa, J. P., Beaulieu, J., Khasa, D. P. and Bousquet, J. 2009. Inferring the past from the present phylogeographic structure of North American forest trees: seeing the forest for the genes. – Can. J. For. Res. 39: 286–307.
Johansen, A. D. and Latta, R. G. 2003. Mitochondrial haplotype distribution, seed dispersal and patterns of postglacial expansion of ponderosa pine. – Mol. Ecol. 12: 293–298.
Keyghobadi, N., Roland, J. and Strobeck, C. 1999. Influence of landscape on the population genetic structure of the alpine butterfly (*Parnassius smintheus* Papilionidae). – Mol. Ecol. 8: 1481–1495.
Klicka, J. and Zink, R. M. 1999. Pleistocene effects on North American songbird evolution. – Proc. R. Soc. B 266: 695–700.
Klicka, J., Spellman, G. M., Winker, K., Chua, V. and Smith, B. T. 2011. A phylogeographic and population genetic analysis of a widespread, sedentary North American bird: the hairy woodpecker (*Picoides villosus*). – Auk 128: 346–362.
Latch, E., Dharmarajan, G., Glaubitz, J. and Rhodes, O. 2006. Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. – Conserv. Genet. 7: 295–302.
Librado, P. and Rozas, J. 2009. DNASP v5: a software for comprehensive analysis of DNA polymorphism data. – Bioinformatics 25: 1451–1452.
Manel, S., Schwartz, M. K., Luikart, G. and Taberlet, P. 2003. Landscape genetics: combining landscape ecology and population genetics. – Trends Ecol. Evol. 18: 189–197.
Manni, F., Guéard, E. and Heyer, E. 2004. Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected using Monmomier's algorithm. – Hum. Biol. 76: 173–190.
Manthey, J. D., Klicka, J. and Spellman, G. M. 2011. Cryptic diversity in a widespread North American songbird: phylogeography of the brown creeper (*Certhia americana*). – Mol. Phylogenet. Evol. 58: 502–512.
McCormack, J. E., Zellmer, A. J. and Knowles, L. L. 2010. Does niche divergence accompany allopatric divergence in *Aphelocoma* jays as predicted under ecological speciation? Insights from tests with niche models. – Evolution 64: 1231–1244.
Milá, B., Smith, T. B. and Wayne, R. K. 2006. Postglacial population expansion drives the evolution of long-distance migration in a songbird. – Evolution 60: 2403–2409.
Milá, B., Smith, T. B. and Wayne, R. K. 2007a. Speciation and rapid phenotypic differentiation in the yellow-rumped warbler *Dendroica coronata* complex. – Mol. Ecol. 16: 159–173.
Milá, B., McCormack, J. E., Castañeda, G., Wayne, R. K. and Smith, T. B. 2007b. Recent postglacial range expansion drive the rapid diversification of a songbird lineage in the genus *Junco*. – Proc. Natl Acad. Sci. USA 274: 2653–2660.
Miller, E. H., Walters, E. L. and Ouellet, H. 1999. Plumage, size, and sexual dimorphism in the Queen Charlotte Islands hairy woodpecker. – Condor 101: 86–95.
Miller, L. H. and de May, I. S. 1942. The fossil birds of California, an avifauna and bibliography with annotations. – Univ. Calif. Publ. Zool. 47: 47–142.
Milot, E., Gibbs, H. K. and Hobson, K. A. 2000. Phylogeography and genetic structure of northern populations of the yellow warbler (*Dendroica petechia*). – Mol. Ecol. 9: 667–681.
Morris-Pocock, J. A., Steeves, T. E., Estela, F. A., Anderson, D. J. and Friesen, V. L. 2010. Comparative phylogeography of brown (*Sula leucogaster*) and red-footed boobies (*S. sula*): the influence of physical barriers and habitat preference on gene flow in pelagic seabirds. – Mol. Phylogenet. Evol. 54: 883–896.
Nei, M. and Tajima, F. 1981. DNA polymorphism detectable by restriction endonucleases. – Genetics 97: 145–163.
Nielson, M., Lohman, K. and Sullivan, J. 2001. Phylogeography of the tailed frog *Ascaphus truei*: implications for the biogeography of the Pacific Northwest. – Evolution 55: 147–160.
Otto-Bliesner, B. L., Brady, E. C., Clauzet, G., Tomas, R., Levis, S. and Kothavala, Z. 2006. Last glacial maximum and Holocene climate in ccsm3. – J. Clim. 19: 2526–2544.
Ouellet, H. R. 1977. Biosystematics and ecology of *Picoides villosus* (L.) and *Picoides pubescens* (L.) (Aves: Picidae). – PhD thesis, McGill Univ.
Paetkau, D. and Strobeck, C. 1996. Mitochondrial DNA and the phylogeography of Newfoundland black bears. – Can. J. Zool. 74: 192–196.
Parmalee, P. W. 1992. A late Pleistocene avifauna form northwestern Alabama. – Nat. Hist. Mus. Los Angeles Co. Sci. Ser. 36: 307–318.
Peakall, R. and Smouse, P. E. 2006. GenAlEx 6: genetic analysis in excel. Population genetic software for teaching and research. – Mol. Ecol. Not. 6: 288–295.

Phillips, S. J., Anderson, R. P. and Schapire, R. E. 2006. Maximum entropy modelling of species geographic distributions. – Ecol. Model. 190: 231–259.
Pielou, E. C. 1991. After the ice age: the return of life to glaciated North America. – Univ. of Chicago Press.
Piertney, S. B., MacColl, A. D. C., Bacon, P. J. and Dallas, J. F. 1998. Local genetic structure in red grouse (*Lagopus lagopus scoticus*): evidence from microsatellite DNA markers. – Mol. Ecol. 7: 1645–1654.
Piry, S., Luikart, G. and Cornuet, J.-M. 1999. BOTTLENECK: a computer program for detecting recent reductions in the effective allele size using allele frequency data. – J. Hered. 90: 502–503.
Porter, S. C., Pierce, K. L. and Hamilton, T. D. 1983. Late Wisconsin mountain glaciation in the western United States. – In: Wright, Jr H. E. (ed.), Late-Quaternary environments of the United States. Univ. of Minnesota Press, pp. 71–111.
Posada, D. 2008. JModelTest: phylogenetic model averaging. – Mol. Biol. Evol. 25: 1253–1256.
Pritchard, J., Stephens, M. and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. – Genetics 155: 945–959.
Pruett, C. and Winker, K. 2008. Evidence for criptic northern refugia among high- and temperate latitude species in Beringia. – Clim. Change 86: 23–27.
Pulgarin-Restrepo, P. C. 2011. The population history of the downy woodpecker *Picoides pubescens* in North America: insights from genetics, ecological niche-modelling and bioacoustics. – MSc thesis, Univ. of Lethbridge.
Questiau, S., Gielly, L., Clouet, M. and Taberlet, P. 1999. Phylogeographical evidence of gene flow among common crossbills *Loxia curvirostra*, (Aves, Fringillidae) populations at the continental level. – Heredity 83: 196–205.
Rajora, O. P., DeVerno, L., Mosseler, A. and Innes, D. J. 1998. Genetic diversity and population structure of disjunct Newfoundland and central Ontario populations of eastern white pine (*Pinus strobus*). – Can. J. Bot. 76: 500–508.
Rambaut, A. and Drummond, A. 2006. Figtree: tree drawing tool, version 1.3.1. – Inst. Evolutionary Biology, Univ. of Edinburgh, <http://tree.bio.ed.ac.ik/software/figtree/>.
Ramos-Osnins, S. E. and Rozas, J. 2002. Statistical properties of new neutrality tests against population growth. – Mol. Biol. Evol. 19: 2092–2100.
Raymond, M. and Rousset, F. 1995. GENEPOP version 1.2: population genetics software for exact tests and ecumenicism. – J. Hered. 86: 248–249.
Rice, W. R. 1989. Analyzing tables of statistical tests. – Evolution 43: 223–225.
Ridgely, R., Alinutt, T., Brooks, T., McNicol, D., Mehlman, D., Young, B. and Zook, J. 2007. Digital distribution maps of the birds of the western hemisphere, version 3.0. – NatureServe, Arlington, VA, USA.
Rogers, A. R. and Harpending, H. 1992. Population growth makes waves in the distribution of pairwise genetic differences. – Mol. Biol. Evol. 9: 552–569.
Rohwer, S. and Wood, C. 1998. Three hybrid zones between hermit and Townsend's warblers in Washington and Oregon. – Auk 115: 284–310.
Ruegg, K. C. and Smith, T. B. 2002. Not as the crow flies: a historical explanation for circuitous migration in Swainsosn's thrush (*Catharus ustulatus*). – Proc. R. Soc. B 269: 1375–1381.
Rush, A. C., Cannings, R. J. and Irwin, D. E. 2009. Analysis of multilocus DNA reveals hybridization in a contact zone between *Empidonax* flycatchers. – J. Avian Biol. 40: 614–624.
Sambrook, J., Fritsch, E. F. and Maniatis, T. 1989. Molecular cloning: a laboratory manual. – Cold Spring Harbor, NY.
Slatkin, M. and Hudson, R. R. 1991. Pairwise comparison of mitochondrial DNA sequences in stable and exponentially growing populations. – Genetics 129: 555–562.
Soltis, D. E., Morris, A. B., McLachlan, J. S., Manos, P. S. and Soltis, P. S. 2006. Comparative phylogeography of unglaciated eastern North America. – Mol. Ecol. 15: 4261–4293.
Spellman, G. M. and Klicka, J. 2007. Phylogeography of the white-breasted nuthatch (*Sitta carolinensis*): diversification in North American pine and oak woodlands. – Mol. Ecol. 16: 1729–1740.
Spellman, G., Riddle, B. and Klicka, J. 2007. Phylogeography of the mountain chickadee (*Poecile gambeli*): diversification, introgression, and expansion in response to Quaternary climate change. – Mol. Ecol. 16: 1055–1068.
Tamura, K., Dudley, J., Nei, M. and Kumar, S. 2007. MEGA4: molecular evolutionary genetics analysis MEGA, software version 4.0. – Mol. Biol. Evol. 24: 1596–1599.
Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. – Mol. Biol. Evol. 28: 2731–2739.
Toews, D. P. L., Brelsford, A. and Irwin, D. E. 2011. Hybridization between Townsend's *Dendroica townsendi* and black-throated green warblers *D. virens* in an avian suture zone. – J. Avian Biol. in press.
Topp, C. M. and Winker, K. 2008. Genetic patterns of differentiation among five landbird species from the Queen Charlotte Islands, British Columbia. – Auk 125: 461–472.
Turmelle, A. S., Kunz, T. H. and Sorenson, M. D. 2011. A tale of two genomes: contrasting patterns of phylogeographic structure in a widely distributed bat. – Mol. Ecol. 20: 357–375.
Välimäki, K., Jaari, S., Piha, M., Pakkala, T. and Merilä, J. 2008. Isolation and characterization of 17 polymorphic microsatellite loci for the three-toed woodpecker (*Picoides tridactylus*). – Mol. Ecol. Resour. 8: 1152–1154.
Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M. and Shipley, P. 2004. Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. – Mol. Ecol. Not. 4: 535–538.
Vila, M., Robles, H., Ciudad, C., Olea, P. P. and Baglione, V. 2008. Isolation and characterization of 12 microsatellite markers in the middle-spotted woodpecker *Dendrocopos medius*. – Mol. Ecol. Resour. 8: 415–417.
Waltari, E., Hijmans, R. J., Peterson, A. T., Nyári, Á. S., Perkins, S. L. and Guralnick, R. P. 2007. Locating Pleistocene refugia: comparing phylogeographic and ecological niche model predictions. – PLoS One 2: e563.
Walsh, P. S., Metzger, D. A. and Higuchi, R. 1991. Chelex 100 as a medium for PCR based typing from forensic material. – Biotechniques 10: 506–513.
Williams, J. W. 2003. Variations in tree cover in North America since the last glacial maximum. – Global Planet. Change 35: 1–23.
Wilson, G. A. and Rannala, B. 2003. Bayesian inference of recent migration rates using multilocus genotypes. – Genetics 163: 1177–1191.
Wright, S. 1946. Isolation by distance under diverse systems of mating. – Genetics 31: 39–59.
Zink, R. M., Rohwer, S., Drovetski, S., Blackwell-Rago, R. C. and Farrell, S. L. 2002. Holarctic phylogeography and species limits of three-toed woodpeckers. – Condor 104: 167–170.

Supplementary material (Appendix J5604 at <www.oikosoffice.lu.se/appendix>). Appendix 1–3.