Effects of Pleistocene glaciations on population structure of North American chestnut-backed chickadees

THERESA M. BURG,*† ANTHONY J. GASTON,‡ KEVIN WINKER§ and VICKI L. FRIESEN*
*Department of Biology, Queen's University, Kingston, ON, Canada K7L 3N6, †Centre d'Ecologie Fonctionnelle et Evolutive, CNRS, UMR 5175, 1919 route de Mende, 34293 Montpellier, France, ‡Environment Canada, National Wildlife Research Centre, Carleton University, Ottawa, ON, Canada K1A 0H3, §University of Alaska Museum, 907 Yukon Drive, Fairbanks, AK 99775, USA

Abstract

The postglacial recolonization of northern North America was heavily influenced by the Pleistocene glaciation. In the Pacific Northwest, there are two disjunct regions of mesic temperate forest, one coastal and the other interior. The chestnut-backed chickadee is one of the species associated with this distinctive ecosystem. Using seven microsatellite markers we found evidence of population structure among nine populations of chestnut-backed chickadees. High levels of allelic variation were found in each of the populations. Northern British Columbia and central Alaska populations contained a large number of private alleles compared to other populations, including those from unglaciated regions. The disjunct population in the interior was genetically distinct from the coastal population. Genetic and historical records indicate that the interior population originated from postglacial inland dispersal. Population structuring was found within the continuous coastal population, among which the peripheral populations, specifically those on the Queen Charlotte Islands and the central Alaska mainland, were genetically distinct. The pattern of population structure among contemporary chickadee populations is consistent with a pioneer model of recolonization. The persistence of genetic structure in western North American chestnut-backed chickadees may be aided by their sedentary behaviour, linear distribution, and dependence on cedar–hemlock forests.

Keywords: disjunct population, historical biogeography, microsatellite, Pleistocene glaciation, *Poecile rufescens*, postglacial recolonization

Received 12 October 2005; revision received 29 January 2006; accepted 2 March 2006

Introduction

Pleistocene glaciations caused historical range expansions and contractions and were therefore important factors affecting the levels of genetic diversity within and among populations. The retreat of the Cordilleran glacier in western North America was asynchronous (Warner *et al.* 1982; Mann & Hamilton 1995), and as the ice retreated trees rapidly colonized deglaciated areas (Warner *et al.* 1982; Barnosky *et al.* 1987; Mann & Hamilton 1995). Expansion of forests into deglaciated regions was dependent on dispersal from refugia. For mesic temperate forests, the main refugium was south of the ice sheets, with possible additional refugia along the Pacific Coast (Brunsfeld *et al.* 2001; Ritland *et al.*

2001). The two main arboreal components of the mesic temperate forest, western red cedar (*Thuja plicata*) and western hemlock (*Tsuga heterophylla*), were present in northern British Columbia and southeast Alaska by the early to mid-Holocene (Barnosky *et al.* 1987; Mann & Hamilton 1995). The arrival of these species within the British Columbia interior appears to be more recent, 2500–1500 years before present (BP) (Barnosky *et al.* 1987; Brunsfeld *et al.* 2001). The rate and pattern of expansion of mesic forests into northern areas would depend not only on their dispersal abilities, but also on the successional stage of existing forests. The pattern of colonization of species integrally associated with these cedar–hemlock ecosystems, such as the chestnut-backed chickadee (*Poecile rufescens*), would depend on similar factors.

Hewitt (1996) proposed two models explaining the recolonization of previously glaciated areas from refugia.

Correspondence: Theresa M. Burg, Fax: +33 04 67 41 21 38; E-mail: theresaburg@yahoo.com

In the 'phalanx' model, range expansion is slow, resulting in a homogeneous population and no loss of genetic variation. In contrast, the 'pioneer' or 'leading edge' model involves the rapid recolonization of previously glaciated regions through both short-distance and long-distance dispersal. Depending on how colonizing groups of individuals are formed, and the relationship between the number of colonists and the subsequent number of post-colonizing immigrants from the source population, different patterns of differentiation are expected to emerge. Ibrahim *et al.* (1996) modelled spatial patterns of genetic variation generated by different modes of dispersal during a range expansion and found that if dispersal is leptokurtic, i.e. most colonists disperse a short distance and a few disperse a long distance, then heterogeneity is greater and genetic diversity is reduced in the colonized areas. This is partially due to the early arrival of long-distance colonists creating a high-density barrier. When populations have filled available habitat, it is more difficult for later waves of colonists to advance because they must disperse into the already-colonized area and, unlike the original colonists, their reproduction is not exponential (Hewitt 2000). As dispersal is made more leptokurtic and more long-distance dispersal occurs, populations become more heterogeneous, and this pattern will persist and increase with time (Hewitt 1996). In contrast, in the absence of long-distance dispersal, differentiation among demes is lower and decreases with time. The patterns of population structure predicted by the two models of recolonization are very different (see Johansen & Latta 2003). Under the pioneer model, pockets of genetically isolated populations would exist within a larger genetically homogenous population. In contrast, population structure resulting from expansion using the phalanx model would be minimal. The newly founded populations would be genetically similar to the source population.

In several cases, historical range shifts associated with Pleistocene glaciations led to the existence of disjunct populations of plants and animals. In the Pacific Northwest, there are two large, disjunct regions of mesic, temperate coniferous forests: the Cascade/Coast Mountains in the west and the central Rocky Mountains in the interior. Chestnut-backed chickadees are integrally associated with this regionally important forest ecosystem and can thus serve as a model for understanding the effects of Pleistocene glaciation on this ecosystem. Several molecular studies have focused on organisms from these mesic forest areas to determine their origins (Green *et al.* 1996; Soltis *et al.* 1997; Taylor *et al.* 1999; Brunsfeld *et al.* 2001; Carstens *et al.* 2004). Brunsfeld *et al.* (2001) proposed two hypotheses for the origin of mesic forest disjuncts: ancient vicariance and inland dispersal. Under the 'ancient vicariance' hypothesis, the mesic forests of the Pacific coast and western Rocky Mountains were contiguous until the late Pliocene when the rise of the Cascade Mountains created a rain shadow in the intervening Columbia Basin, thereby isolating populations in the interior. In the 'inland dispersal' hypothesis, the interior populations are postglacially derived from either a northern or southern colonization around the Columbia Basin. Genetic studies have shown that the disjunct lineages are ancient in some species (Nielson *et al.* 2001; Carstens *et al.* 2004) and recent in others (Richardson *et al.* 2002; Carstens *et al.* 2004, 2005). We may expect some species to give different signals with respect to these hypotheses because although glacial refugia played an integral role in recolonization, species-specific dispersal patterns, dispersal capabilities, habitat requirements, and colonization rates determine the rate and pattern of recolonization.

The chestnut-backed chickadee is a year-round resident and a near obligate of cedar–hemlock ecosystems. Chestnut-backed chickadees primarily nest in Douglas fir (*Pseudotsuga menziesii*), western hemlock and western red cedar (Lundquist & Mariani 1991), and the species is most abundant in mature and old-growth forests (Anthony *et al.* 1996). The distribution of this species follows closely that of the mesic temperate forests, occupying a narrow band less than 200 km wide along the Pacific Coast from north California to Alaska, with a disjunct population in the western Rocky Mountains (Fig. 1). The coastal and interior populations of these chickadees are separated by the Coast and Cascade mountain ranges. In 1904, the interior population was restricted to several sites in northern Idaho (Grinnell 1904); however, a century later the chestnut-backed chickadee has extended its range into southeastern British Columbia and parts of eastern Washington and Oregon and western Montana and Alberta, while remaining physically isolated from the coastal population (Dahlsten *et al.* 2002). A similar expansion appears to have occurred at the southern end of its distribution, in the Sierra Nevada mountains of California (Grinnell 1904; Root 1964; Brennan & Morrison 1991).

Our study examines spatial patterns of genetic variation in natural populations representative of this important forest ecosystem, following the retreat of the Cordilleran ice sheet from northwestern North America. We studied patterns of genetic structure throughout most of the previously glaciated portions of the range of the chestnut-backed chickadee in British Columbia and southeast Alaska, and unglaciated portions of the species' range in Washington and Oregon. We attempted to answer four central questions: (i) How are contemporary populations of chestnut-backed chickadees structured? (ii) Was long-distance dispersal a factor in recolonization of the northern range? (iii) Is there lower genetic variation in populations from deglaciated areas? (iv) Are interior and coastal populations genetically isolated? By addressing the first three questions, we can determine whether the postglacial range expansion of this species followed the 'phalanx' or 'pioneer' model. A priori,
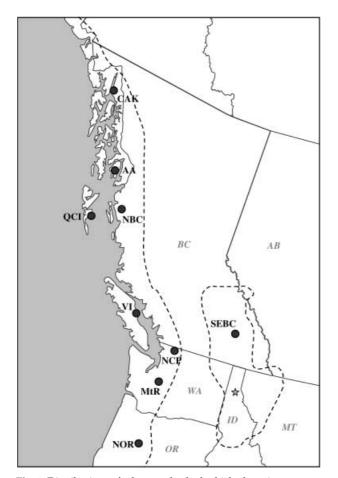
Fig. 1 Distribution of chestnut-backed chickadees in western North America. Species range is indicated by dashed lines. Sampling sites are central, coastal Alaska (CAK), Alexander Archipelago (AA), Queen Charlotte Islands (QCI), northern British Columbia (NBC), Vancouver Island (VI), North Cascades Park (NCP), Mount Rainier (MtR), northern Oregon (NOR) and southeastern British Columbia (SEBC). The star in Idaho (ID) represents 1904 range of chestnut-backed chickadees in the interior (Grinnell 1904).

both models are possible. If no suitable habitat was available north of the leading edge, then recolonization of previously glaciated regions would likely have been a slow process and followed the 'phalanx' model of expansion. Alternatively, if there were pockets of suitable habitat north of the leading edge, then colonization of available habitat could have been rapid and hence followed the 'pioneer' model. The fourth question relates to the disjunct distribution of the interior and coastal populations. If the interior population was recently founded, then it should contain a subset of the alleles present in the coastal population. Alternatively, if the two populations are the result of ancient vicariance, then the interior populations should be genetically distinct from the coastal population and contain a number of private alleles due to mutation.

Materials and methods

Samples and genotyping

The sampling area covered >75% of the contemporary range of the chestnut-backed chickadee, including the disjunct population in the western Rocky Mountains. A total of 249 samples were collected from nine populations in Alaska, British Columbia, Washington and Oregon (Fig. 1). The populations sampled in Washington and Oregon were from areas that were unglaciated during the Pleistocene. Blood samples were collected from 223 individuals during the summer of 2002 and 2003. Birds were caught using mist nets, and blood was taken from the brachial vein, dried on filter paper, and stored in individual bags. Twenty-six samples (central, coastal Alaska, $n = 6$; Alexander Archipelago, $n = 9$; northern Oregon, $n = 1$; and Queen Charlotte Islands, $n = 10$) collected over a similar time period were obtained from the University of Alaska Museum.

DNA was extracted using standard proteinase K/phenol–chloroform extraction followed by ethanol precipitation (Sambrook *et al.* 1989). Seven microsatellite primer pairs isolated from other passerine species were used for genotyping (Table 1). The forward primers were modified by the addition of M13 sequence (CACGACGTTGTAAAAC-GAC) to the 5′ end to allow for direct incorporation of a fluorescently labelled M13 primer. All loci were amplified using a two-step annealing procedure: one cycle for 2 min at 94 °C, 45 s at T_{A1}, 60 s at 72 °C; seven cycles of 60 s at 94 °C, 30 s at T_{A1}, 45 s at 72 °C; 30 cycles of 30 s at 89 °C, 30 s at T_{A2}, 45 s at 72 °C; and one final cycle of 5 min at 72 °C. For locus Pat43 $T_{A1} = 55$ °C and $T_{A2} = 57$ °C and for the other six loci $T_{A1} = 50$ °C and $T_{A2} = 52$ °C. PCR products were run on a 6% acrylamide gel using a LI-COR 4200 IR2 (LI-COR Inc.). Individuals of known allele sizes were included on each gel to ensure that alleles were sized consistently between gels. Alleles were scored using Gene-ImagIR (Scanalytics) and sizing was confirmed by visual inspection.

Statistical analyses

Tests for departures from HardyWeinberg equilibrium and for linkage disequilibrium were examined using exact tests (Guo & Thompson 1992) as implemented in GENEPOP version 3.3 (Raymond & Rousset 1995b), and sequential Bonferroni corrections for multiple tests were applied (Rice 1989). As estimates of allelic diversity can be biased due to unequal samples sizes, allelic richness was estimated using FSTAT version 2.9.3.2 (Goudet 2001).

Population differentiation can be estimated using two types of tests: standard statistical methods and clustering methods. Of the standard statistical methods, allelic goodness of fit tests are the most powerful for detecting

Sampling site		Pat14*	Pat43*	BT6*	Pdo5†	Ppi2‡	Pocc1§	Escu6¶	Ave.
CAK	H_E	0.59	0.83	0.57	0.71	0.58	0.60	0.89	0.68
($n = 44$)	H_O	0.56	0.93	0.46	0.71	0.45	0.72	0.88	0.67
	A_N	13	13	6	9	8	6	16	10.1
	A_R	5.1	7.2	3.3	5.9	4.1	3.6	8.9	5.5
	A_P								15.5
AA	H_E	0.59	0.79	0.57	0.75	0.41	0.59	0.83	0.65
($n = 9$)	H_O	0.67	0.78	0.56	0.89	0.25	0.44	1.00	0.65
	A_N	6	9	5	5	4	3	8	5.7
	A_R	5.6	8.5	3.9	5.0	4.9	4.2	9.5	6.1
	A_P								7.5
NBC	H_E	0.63	0.84	0.57	0.71	0.57	0.65	0.91	0.70
($n = 38$)	H_O	0.64	0.87	0.53	0.64	0.52	0.60	0.95	0.68
	A_N	15	21	7	9	9	7	16	12.0
	A_R	5.9	8.5	3.98	5.9	4.9	4.2	9.5	6.1
	A_P								14.3
QCI	H_E	0.72	0.84	0.46	0.82	0.49	0.55	0.85	0.68
($n = 43$)	H_O	0.88	0.81	0.43	*0.63*	0.41	0.48	0.88	0.64
	A_N	14	11	4	8	10	5	12	9.1
	A_R	5.7	7.3	2.4	6.1	4.8	2.9	7.6	5.3
	A_P								4.7
VI	H_E	0.63	0.82	0.40	0.82	0.71	0.56	0.90	0.69
($n = 40$)	H_O	0.75	0.87	0.38	*0.55*	0.66	0.75	0.93	0.70
	A_N	15	15	4	9	12	4	16	10.7
	A_R	5.9	7.2	2.8	6.2	5.9	2.8	9.3	5.7
	A_P								6.7
NCP	H_E	0.74	0.76	0.54	0.76	0.81	0.59	0.90	0.73
($n = 21$)	H_O	0.86	0.91	0.50	0.71	0.65	0.57	1.00	0.74
	A_N	11	12	4	8	8	3	15	8.7
	A_R	7.0	6.7	3.3	6.1	6.2	3.0	9.7	6.0
	A_P								1.6
MtR	H_E	0.68	0.76	0.59	0.78	0.61	0.57	0.89	0.70
($n = 16$)	H_O	0.81	0.67	0.60	0.69	0.64	0.73	0.86	0.72
	A_N	8	10	7	8	6	3	12	7.7
	A_R	5.4	6.8	4.9	6.8	4.6	2.9	9.5	5.8
	A_P								7.4
NOR	H_E	0.42	0.72	0.56	0.66	0.69	0.46	0.83	0.62
($n = 8$)	H_O	0.38	0.75	0.38	0.50	0.57	0.50	0.75	0.55
	A_N	5	6	3	3	6	3	8	4.9
	A_R	5.0	6.0	3.0	3.0	6.0	3.0	8.0	4.9
	A_P								5.9
SEBC	H_E	0.66	0.69	0.42	0.83	0.56	0.62	0.88	0.67
($n = 30$)	H_O	0.59	0.70	0.33	0.73	0.60	0.52	0.93	0.63
	A_N	10	12	6	8	6	6	13	8.7
	A_R	5.6	6.2	3.5	6.8	3.8	4.2	8.6	5.5
	A_P								6.6

*Otter *et al.* (2001), †Griffith *et al.* (1999), ‡Martinez *et al.* (1999), §Bensch *et al.* (1997), ¶Hanotte *et al.* (1994).

population structure when sample sizes are unequal (Goudet *et al.* 1996). TFPGA version 1.3 was used to test for differences in allele frequencies among populations (1000 dememorization steps, 20 batches and 20 000 permutations/ batch, Miller 1997). TFPGA uses a Markov chain Monte Carlo (MCMC) approximation of Fisher's exact test (Raymond & Rousset 1995a) and significance values are combined across all loci (Fisher 1954). In contrast to standard statistical methods that use pre-defined populations, normally sampling sites, clustering methods use multilocus genotypes to

create groups of individuals within which linkage disequilibrium is minimized (Pritchard *et al.* 2000; Manel *et al.* 2005). The program STRUCTURE version 2.1 (Pritchard *et al.* 2000) was used to determine the level of population structure in the data set independent of the geographic origin of the samples. Three independent runs of 10^6 MCMC iterations were performed using uncorrelated allele frequencies and the admixture model to estimate the number of populations (K) for $K = 1$–10. Results from runs at each value of K were averaged.

F_{ST} and R_{ST} are two commonly used estimators of population divergence. R_{ST} was developed specifically for microsatellites and incorporates microsatellite-specific mutation models, yet simulation studies show that F_{ST} has lower variance (Paetkau *et al.* 1997) and performs better when sample sizes (<50/population) and/or number of loci (<20) are small (Gaggiotti *et al.* 1999). For populations sharing a recent history, mutation probably plays a minor role in differentiation (Hardy *et al.* 2003); however, to determine whether allele sizes were informative, and whether F_{ST} or R_{ST} was a better estimator for our data, allele size permutation tests were performed in SPAGEDI (Hardy & Vekemans 2002). Permutation tests showed that allele sizes were not informative ($P = 0.80$) and therefore, F_{ST}, not R_{ST}, was used. Weir and Cockerham's estimator of F_{ST} (1984) was used to measure population variation. Both global and pairwise F_{ST} estimates were calculated in GENETIX 4.02 (Belkhir *et al.* 2000). Relative measures of differentiation can be difficult to compare directly (Hedrick 1999). The genotype likelihood ratio distance (D_{LR}, Paetkau *et al.* 1997) is well suited to studying fine-scale population structure and has lower variance than other distance measures (Paetkau *et al.* 1997). D_{LR} is the likelihood of a genotype from one population being identical to a genotype in another population. When $D_{LR} = 1$, the genotypes of individuals from the two populations being compared are one order of magnitude more likely to occur in the individuals' own population than in the other population (Paetkau *et al.* 1997). D_{LR} was calculated in Doh (http://biodb.biology.ualberta.ca/jbrzusto).

A factorial correspondence analysis (FCA) was performed to aid in visualization of the patterns of genetic structure using GENETIX 4.02 (Belkhir *et al.* 2000). FCA uses individual genotypic data to quantify the amount of inertia among populations. The global inertia is proportionally weighted relative to sample size and the total number of alleles present in each sample. The 'centre of gravity' for each population was plotted in 3D space.

To estimate the relative importance of drift and migration in determining population structure we used isolation by distance (IBD) and Bayesian approaches. Tests for isolation by distance allow us to evaluate the relative historical roles of gene flow and drift on population structure by comparing expected pairwise genetic and geographic distances

with those expected under a stepping-stone model of population structure (Hutchison & Templeton 1999). They also allow us to determine if gene flow is affected by geographic distance. IBD tests were performed in TFPGA version 1.3 (Miller 1997) and significance was determined using 999 permutations. Geographic distances to the southeastern British Columbia population were calculated as the shortest distance through mesic forest habitat. The program 2MOD (Ciofi *et al.* 1999) uses a Bayesian approach to estimate the relative importance of drift and migration in determining population structure by comparing two models of population history. The gene flow–drift model assumes that allele frequencies in the populations are determined by a balance of genetic drift and immigration. The drift model assumes one panmictic population was fragmented into multiple subpopulations and that subpopulations are diverging in the absence of migration. Both models assume that mutation is negligible. The program uses a coalescent-based MCMC approach with Metropolis–Hastings sampling to compare the likelihood of the two models. 2MOD also simulates the posterior density of F, i.e., the probability that two alleles chosen at random in a population are identical by descent rather than due to immigration or a founder event. Two independent runs of 10^6 iterations were performed and results compared to determine whether the posterior probabilities had converged. The first 10% of the runs were discarded to remove effects of initial starting parameters.

Founder effects can cause a reduction in the number of alleles in a new population, but similar decreases in allelic variation can also result from a population bottleneck. To test for recent reductions in population size, we used the program BOTTLENECK version 1.2.02 (Cornuet & Luikart 1996; Piry *et al.* 1999). During a bottleneck, alleles will be lost from the population and levels of heterozygosity will be temporarily higher than expected under mutation–drift equilibrium. The one-tailed Wilcoxon sign test is the most powerful and robust of the three tests in BOTTLENECK for studies using fewer than 20 loci (Piry *et al.* 1999). We therefore used this test with a two-phase mutation model (TPM), with 95% single-stepwise mutations and 1000 iterations, as recommended by Piry *et al.* (1999), to determine if a bottleneck occurred in the last $2N_e$–$4N_e$ generations.

Results

Levels of expected heterozygosity for the seven microsatellite loci ranged from 0.40 to 0.91 in the nine populations (Table 1). The total number of alleles at each locus was high (9–31) and each population contained 3–16 alleles/locus (Table 1). The allele frequencies in the two less extensively sampled populations (Alexander Archipelago and northern Oregon) at Pat14 and Escu6 may not accurately reflect the alleles present in each population due to the small sample

	CAK	AA	NBC	QCI	VI	NCP	MtR	NOR	SEBC
CAK									
AA	**0.002**								
NBC	**0.003**	0.266							
QCI	**< 0.001**	**0.003**	**< 0.001**						
VI	**< 0.001**	0.055	0.008	**< 0.001**					
NCP	**< 0.001**	0.013	0.014	**< 0.001**	0.155				
MtR	**< 0.001**	0.157	0.470	**< 0.001**	0.014	0.405			
NOR	**0.002**	0.025	0.192	**0.003**	**0.004**	**0.006**	0.059		
SEBC	**< 0.001**	**0.003**	**< 0.001**	**< 0.001**	**< 0.001**	**< 0.001**	**0.001**	0.025	

Table 2 Results of exact tests for population differentiation. Significant pairwise tests are indicated in bold. Refer to Fig. 1 for abbreviations of sampling sites

	CAK	AA	NBC	QCI	VI	NCP	MtR	NOR	SEBC
CAK		1.64	0.74	0.82	2.01	2.69	0.70	1.21	1.21
AA	0.026		0.49	1.63	1.41	1.30	0.90	0.97	1.32
NBC	−0.002	0.011		1.47	0.75	0.76	−0.05	0.90	0.54
QCI	**0.021**	0.017	0.019		2.22	2.67	1.58	1.46	1.48
VI	**0.022**	0.010	0.010	**0.017**		0.55	0.99	1.16	1.40
NCP	**0.026**	0.026	0.017	**0.041**	0.008		0.38	2.69	1.94
MtR	0.016	0.003	0.004	0.016	0.008	0.007		0.96	0.62
NOR	0.019	0.033	0.012	0.029	0.029	0.037	0.010		1.11
SEBC	**0.022**	0.034	0.013	**0.025**	**0.019**	**0.035**	0.014	0.020	

Table 3 Matrix of genetic distances: pairwise F_{ST} values (below diagonal) and D_{LR} (above diagonal) for nine chestnut-backed chickadee populations. F_{ST} values significant after correction for multiple tests are in bold. Refer to Fig. 1 for abbreviations of sampling sites
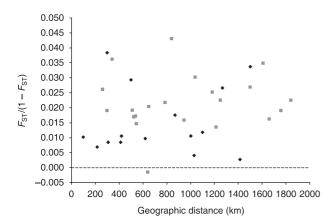
size and the large number of alleles at these two loci. Average allelic richness, adjusting for unequal sample sizes, ranged from 4.9 to 6.1 alleles/locus (Table 1). Northern and southeastern British Columbia populations contained similar levels of genetic diversity to populations from unglaciated areas in Oregon and Washington (Table 1). Private alleles were detected at each locus and were present in all nine populations (Table 1). The central Alaska and northern British Columbia populations contained a disproportionately large number of private alleles (11 and 12, respectively) compared to the other populations (1–5 alleles). The distribution of private alleles is not significantly heterogeneous (G_{corr} = 14.30, d.f. = 8, P = 0.07); however, a Freeman–Tukey test showed that the central Alaska and northern British Columbia populations had more private alleles than the other populations and that the North Cascades Park population had fewer private alleles.

Prior to Bonferroni corrections (Rice 1989), 6 of the 63 locus-population comparisons showed significant departures from Hardy–Weinberg equilibrium (Pdo5 in Queen Charlotte Islands and Vancouver Island); however after corrections for multiple tests none were significant. No tests for linkage disequilibrium were significant.

Population differentiation was detected in the western North American populations of the chestnut-backed chickadee. Exact tests showed significant differences for all but one of the comparisons involving the central Alaska, southeastern British Columbia and Queen Charlotte Islands populations (Table 2). The northern Oregon–Vancouver Island and northern Oregon–North Cascades

Park comparisons also had significantly different allele frequencies.

The results of STRUCTURE were not clear. Clustering algorithms are known to work well when population divergence is high, but it is not known how well they perform when divergence is low (Manel *et al.* 2005). The highest probability was for K = 5 [Pr(K = 5) = 0.83]; however, values of K > 3 resulted in further division of sampling sites within the existing inferred clusters. The clustering at K = 3 consisted of Queen Charlotte Islands and central Alaska, southeastern British Columbia, and the remaining populations.

The global F_{ST} value was low (F_{ST} = 0.018 compared to the theoretical maximum F_{ST} = 0.291; Hedrick 1999), but highly significant ($P < 0.001$). Pairwise F_{ST} values between populations ranged from −0.002 to 0.041 (Table 3). Based on the results from the exact tests and cluster analysis, we examined the F_{ST} values more closely. The F_{ST} values were higher for comparisons involving the central Alaska, southeastern British Columbia and Queen Charlotte Islands populations (Table 3 and Fig. 2). The exception was the central Alaska–northern British Columbia comparison. Five of the F_{ST} values for the other population comparisons were relatively high (0.017–0.037, Table 3). These comparisons involved the two populations with smaller samples sizes (northern Oregon and Alexander Archipelago) and therefore higher variance; none of these five values was significant after correction for multiple tests. D_{LR} distances also suggested differentiation within western North America. The average D_{LR} was 1.24. The smallest distance
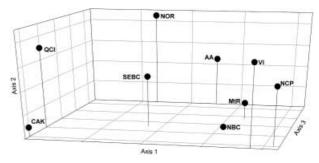
Fig. 2 Graph of genetic versus geographic distances between populations of chestnut-backed chickadees. Comparisons involving central Alaska, Queen Charlotte Islands and southeastern British Columbia populations are represented by grey squares, and all other populations are represented by black diamonds.

was between Mount Rainier and northern British Columbia, where the likelihood of genotypes occurring in either population was identical. In comparison, most of the other values were above 0.50, indicating genotypes had a higher likelihood of occurring in the populations from which they were sampled (Table 3).

The first three axes of the FCA explained 47.6% of the inertia (Fig. 3): 19.6%, 14.9% and 13.1%, on the first, second and third axes, respectively. The individuals from the central Alaska, Queen Charlotte Islands, southeastern British Columbia, northern Oregon and Alexander Archipelago populations showed little overlap with each other or with individuals from the four remaining populations. However, individuals from the Vancouver Island, Mount Rainier, North Cascades Park and northern British Columbia populations, while forming loose clusters with other individuals from the same population, overlapped extensively with individuals from other populations (data not shown).

No significant isolation by distance was found ($r = 0.157$, $P = 0.18$, Fig. 2). The high level of variance suggests that the populations are not in gene flow–drift equilibrium and that the effects of gene flow are greater than drift (Hutchison & Templeton 1999). Bayesian tests using 2MOD confirmed this. The pattern of population structure is best explained by a gene flow–drift model. The likelihood of the gene flow–drift model was 1 (Bayes factor = 999; a Bayes factor > 3 is considered substantial support, Kass & Raftery 1995). Under the gene flow–drift model, the relative effects of drift and immigration can be examined using F (probability of identical alleles by descent). F values range from 0 to 1, and smaller F values imply that the effects of immigration are stronger than the effects of drift. F values from the chickadees ranged from 0.007 to 0.069. The effects of drift are highest in the northern Oregon population and lowest in the northern British Columbia population.

Fig. 3 FCA of chestnut-backed chickadee populations in western North America. Each point represents the centre of the distribution of genotypes from all of the individuals in the population.

No evidence of a recent population bottleneck was found. None of the populations tested showed a significant excess of heterozygosity ($P > 0.81$).

Discussion

Genetic analyses of population structure in chestnut-backed chickadees revealed significant structure within the portions of their range thought to have been glaciated, with four main groupings: Queen Charlotte Islands, central Alaska, southeastern British Columbia and a large coastal group (Table 2). The southernmost sampling site in northern Oregon also showed evidence of population differentiation, but this may be due to small sample sizes.

Pattern of recolonization

The pioneer model of recolonization best describes the contemporary population structure for chestnut-backed chickadees. The factorial correspondence analysis showed that genetically similar populations were not the most geographically proximal populations (Fig. 3). A similar pattern was found in Steller's jay (*Cyanocitta stelleri*) and yellow cedar (*Chamaecyparis nootkatensis*) in the Pacific Northwest (Ritland *et al*. 2001; Burg *et al*. 2005). In addition, all of the chickadee populations inhabiting previously glaciated regions south of the main ice sheet contained private alleles. If colonization had occurred continuously along the Pacific coast (i.e., the 'phalanx' model), adjacent populations should contain similar allelic composition or allelic variation should decrease progressively as colonization proceeded northward (Pruett & Winker 2005). Furthermore, two of the most northerly sites (central, coastal Alaska and northern British Columbia) contained a large number of alleles that were absent in the other populations. The large number of private alleles in these two populations raises a question as to the origin of these populations.

Grinnell (1904) hypothesized that the chestnut-backed chickadee may have arisen in Alaska or northern British Columbia from the boreal chickadee (*P. hudsonicus*), another northern chickadee, and expanded southwards. Mitochondrial data suggest the two closely related species diverged during the Pleistocene (Gill *et al*. 1993), and the explanation of a northern origin is consistent with the large number of private alleles found in two northern populations. Indeed, these two species may have diverged in the north during an interglacial. It is possible that these populations persisted in a northern refugium; however few of the studies of mesic temperate forest species have included samples from this far north. Genetic studies on western red cedar and yellow cedar, components of mesic temperate forests with a similar distribution to the chestnut-backed chickadee, provide conflicting results. Ritland *et al*. (2001) raise the possibility of a northern refugium for yellow cedar, while Glaubitz *et al*. (2000) suggest a single southern refugium for western red cedar populations. If a northern refugium was present, it is not clear whether or not it was large enough to support a population of chestnut-backed chickadees. The paleoecological data for other species of cedar and hemlock is inconclusive, but shows that western red cedar and western hemlock were present along the central coast in the mid- Holocene (Mann & Hamilton 1995). Further sampling of southern populations, especially populations in California, is required to determine if these private alleles are also present in populations from unglaciated areas.

Contemporary population structure

In addition to coastal and interior population subdivision, population structuring exists west of the Cascade and Coast mountain ranges. Within the coastal population, three genetic groups of chestnut-backed chickadees were detected: Queen Charlotte Islands, central Alaska, and a central coastal population. Both the Queen Charlotte Islands and central Alaska chestnut-backed chickadee populations differ in their geographical position relative to the core coastal population. Like the interior population, the Queen Charlotte Island population is physically isolated from the main coastal population. In contrast, the central Alaska population, near the northern periphery of the range, is contiguous with the coastal population.

Peripheral populations are more likely to be genetically isolated than central populations (Jump *et al*. 2003; Wisely *et al*. 2004). Individuals at the centre of a species' range can disperse in many directions, while those at the range edge can disperse in fewer directions. This effect is even more pronounced in linearly distributed species which are more likely to exhibit reduced dispersal at the ends of their ranges due to the reduced number of pathways for gene flow (Kimura & Weiss 1964). Although the central Alaska population is not at the extreme northern end of the contemporary range, it may be considered as a peripheral population. The mesic temperate forest between the Kenai Peninsula, at the far northwestern end of the chestnut-backed chickadee distribution, and northern Southeast Alaska is reduced to a very narrow strip between the Pacific ocean and the St Elias and Chugach mountain ranges, which remain heavily glaciated (Little 1971; Mann & Hamilton 1995). A 100 km break in the distribution of western hemlock, one of the main winter food sources for chestnut-backed chickadees (Price *et al*. 1995), also exists near the British Columbia–Yukon–Alaska border. In addition, many of the cedar and hemlock species that comprise the mesic temperate forest reach their northern limits in Southeast Alaska and northern British Columbia (Little 1971). The combination of these factors suggests that the central Alaska population may be a peripheral population, at least with respect to optimal habitat. Given the large number of private alleles in the central Alaska population it is unlikely that the genetic differentiation is due solely to a reduction in gene flow between it and other populations, but rather a combination of a different founding population and subsequently reduced gene flow occurring in this peripheral portion of the species' range.

Similarly, the peripheral location of the Queen Charlotte Islands, situated 80 km off the British Columbia coast and isolated by the waters of Hecate Strait, may be contributing to the reduction in gene flow between this population and the main coastal populations. Unlike the two other island populations sampled in this study (Alexander Archipelago and Vancouver Island), the Queen Charlotte Islands are located a substantial distance from the mainland. Barriers are often species-specific and while dispersal of other species from the Queen Charlotte Islands to the mainland may be high (e.g. seasonal migrants), for the chestnut-backed chickadee dispersal appears to be reduced.

Inland dispersal or ancient vicariance

The interior and coastal populations appear to be geographically and genetically distinct. The age and origin of the interior population are not known but the patterns of genetic differentiation and private alleles, together with levels of allelic variation, suggest that the origin was due to historical, inland dispersal and included numerous individuals. It is possible that chestnut-backed chickadees survived part of the Pleistocene in the Clearwater refugium in north central Idaho (Brunsfeld *et al*. 2001; Carstens *et al*. 2004); but it is doubtful that they were isolated from the coastal populations for a prolonged period of time. Several factors suggest that the interior population has a more recent origin and is the result of inland dispersal. First, levels of divergence, while high, are comparable to populations

from deglaciated regions (e.g. central Alaska and the Queen Charlotte Islands). Second, if the populations were the result of ancient vicariance, then the interior population should harbour a large number of private alleles. The number of private alleles present in the interior population is no higher than in other coastal populations (Table 1). Third, surveys from the early 20th century showed a paucity of chestnut-backed chickadees in the interior. Grinnell (1904) found chestnut-backed chickadees were present only near Coeur d'Alene, Idaho (Fig. 1). While he may have failed to identify some isolated interior populations, it is unlikely that the 1904 chestnut-backed chickadee distribution was as extensive as it is today. Other researchers have also reported the scarcity of chestnut-backed chickadees east of the Cascade mountains (Bowles 1909). Chestnut-backed chickadees were present in the British Columbia interior in 1974 (Brennan & Morrison 1991), but no data have been published as to their arrival date. If the interior population was of pre-Holocene origin, then the 1904 range should have been more widespread throughout the interior temperate mesic forest. No estimates of the 1904 population size are known, but levels of contemporary genetic variation suggest that it was sufficiently large to retain a large amount of genetic variation and not experience any bottlenecks. Finally, if the interior population persisted in the Clearwater refugium in southern Idaho for a prolonged period of time and was isolated from the coastal populations, then mutation should have played a role in divergence, and there is no genetic evidence for this having occurred.

The exact tests for population differentiation showed that the southeastern British Columbia and northern Oregon populations were homogeneous, but this could be due to the small number of samples from northern Oregon. If these two populations are connected, or were in the recent past, the connection would have been around the southern end of the Columbia Basin, as populations closest to the northern Columbia Basin in Washington and coastal British Columbia are genetically isolated from the interior population.

In summary, the pattern of population structure among chestnut-backed chickadees in western North America suggests that the 'pioneer' model of colonization was the main factor shaping contemporary patterns. Post-colonization dispersal between coastal and interior populations was restricted by a large area of unsuitable habitat, and gene flow was reduced between peripheral populations near the northern end of the distribution and central populations. The high levels of population differentiation in the chestnut-backed chickadee may also be attributed to their historical isolation in multiple refugia, their sedentary nature, their effectively linear distribution and their dependence on cedar–hemlock forests, which together act to decrease dispersal.

References

Anthony RG, Green GA, Forsman ED, Nelson SK (1996) Avian abundance in riparian zones of three forest types in the Cascade Mountains, Oregon. *Wilson Bulletin*, **108**, 280–291.

Barnosky CW, Anderson PM, Bartlein PJ (1987) The northwestern U.S. during deglaciation; vegetational history and paleoclimatic implications. In: *North American and Adjacent Oceans During the Last Deglaciation* (eds Ruddiman WF, Wright HE Jr), pp. 289–321. Geological Society of America, Boulder, Colorado.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2000) GENETIX 4.02, *logiciel sous WindowsTM pour la génétique des populations*. Laboratoire Génome et Populations, CNRS UPR 9060, Université Montpellier II, Montpellier, France.

Bowles JH (1909) Notes on *Parus rufescens* in western Washington. *Condor*, **11**, 55–57.

Bensch S, Price T, Kohn J (1997) Isolation and characterization of microsatellite loci in a *Phylloscopus* warbler. *Molecular Ecology*, **6**, 91–92.

Brennan LA, Morrison ML (1991) Long-term trends of chickadee populations in western North America. *Condor*, **93**, 130–137.

Brunsfeld SJ, Sullivan J, Soltis DE, Soltis PS (2001) Comparative phylogeography of northwestern North America: a synthesis. In: *Integrating Ecological and Evolutionary Processes in a Spatial Context* (eds Silverton J, Antonovics J), pp. 319–339. Blackwell Science, Oxford.

Burg TM, Gaston AJ, Winker K, Friesen VL (2005) Rapid divergence and postglacial colonization in western North American Steller's jays (*Cyanocitta stelleri*). *Molecular Ecology*, **14**, 3745–3755.

Carstens BC, Stevenson AL, Degenhardt JD, Sullivan J (2004) Testing nested phylogenetic and phylogeographic hypothesis in the *Plethodon vandykei* species group. *Systematic Biology*, **53**, 781–792.

Carstens BC, Degenhardt JD, Stevenson AL, Sullivan J (2005) Accounting for coalescent stochasticity in testing phylogeographical structure in the Idaho giant salamander *Dicamptodon aterrimus*. *Molecular Ecology*, **14**, 255–265.

Ciofi C, Beaumont MA, Swingland IR, Bruford MW (1999) Genetic divergence and units for conservation in the Komodo dragon *Varanus komodoensis*. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **266**, 2269–2274.

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144**, 2001–2014.

Dahlsten DL, Brennan LA, McCallum DA, Gaunt SL (2002) Chestnut-backed chickadee (*Poecile rufescens*). In: *The Birds of North America* (eds Poole A, Gill F). The Birds of North America, Philadelphia, Pennsylvania.

Fisher RA (1954) *Statistical Methods for Research Workers*. Oliver and Boyd, Edinburgh.

Gaggiotti OE, Lange O, Rassmann K, Gliddon C (1999) A comparison of two indirect methods for estimating average levels of gene flow using microsatellite data. *Molecular Ecology*, **8**, 1513–1520.

Gill FB, Mostrom AM, Mack AL (1993) Speciation in North American chickadees: I. Patterns of mtDNA genetic divergence. *Evolution*, **47**, 195–212.

Glaubitz JC, El-Kassaby YA, Carlson JE (2000) Nuclear restriction fragment length polymorphism analysis of genetic diversity in western red cedar. *Canadian Journal of Forestry Research*, **30**, 379––389.

Goudet J (2001) FSTAT. *A program to estimate and test gene diversities and fixation indices (version 2.9.3)*. Available from http://www.unil.ch/izea/softwares/fstat.html. Updated from Goudet (1995).

Goudet J, Raymond M, de Meeus T, Rousset F (1996) Testing differentiation in diploid populations. *Genetics*, **144**, 1933–1940.

Green DM, Sharbel TF, Kearsley J, Kaiser H (1996) Postglacial range fluctuation, genetic subdivision and speciation in the western North American spotted frog complex, *Rana pretiosa*. *Evolution*, **50**, 374–390.

Griffith SC, Stewart IRK, Dawson DA, Owens IP, Burke T (1999) Contrasting levels of extra-pair paternity in mainland and island populations of the house sparrow (*Passer domesticus*): is there an 'island effect'? *Biological Journal of the Linnean Society*, **68**, 303–316.

Grinnell J (1904) The origin and distribution of the chestnut-backed chickadee. *Auk*, **21**, 364–382.

Guo SW, Thompson EA (1992) Performing the exact test of Hardy–Weinberg proportion for multiple alleles. *Biometrics*, **48**, 361–372.

Hanotte O, Zanon C, Pugh A, Greig C, Dixon A, Burke T (1994) Isolation and characterization of microsatellite loci in a passerine bird: the reed bunting *Emberiza schoeniclus*. *Molecular Ecology*, **3**, 529–530.

Hardy OJ, Vekemans X (2002) SPAGEDI: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, **2**, 618–620.

Hardy OJ, Charbonnel N, Freville H, Heuretz M (2003) Microsatellite allele sizes: a simple test to assess their significance on genetic differentiation. *Genetics*, **163**, 1467–1482.

Hedrick PW (1999) Perspective: Highly variable loci and their interpretation in evolution and conservation. *Evolution*, **53**, 313–318.

Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, **58**, 247–276.

Hewitt GM (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.

Hutchison DW, Templeton AR (1999) Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution*, **53**, 1898–1914.

Ibrahim KM, Nichols RA, Hewitt GM (1996) Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. *Heredity*, **77**, 282–291.

Johansen AD, Latta RG (2003) Mitochondrial haplotype distribution, seed dispersal and patterns of postglacial expansion of ponderosa pine. *Molecular Ecology*, **12**, 293–298.

Jump AS, Woodward FI, Burke T (2003) *Cirsium* species show disparity in patterns of genetic variation at their range-edge, despite similar patterns of reproduction and isolation. *New Phytologist*, **160**, 359–370.

Kass RE, Raftery AE (1995) Bayes factors. *Journal of the American Statistical Association*, **90**, 773–795.

Kimura M, Weiss GH (1964) The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics*, **49**, 561–576.

Little EL (1971) *Atlas of United States Trees*. United States Department of Agriculture Misc. Publication 1146. Washington, DC.

Lundquist RW, Mariani JM (1991) Nesting habitat and abundance of snag-dependent birds in the Southern Washington Cascade Range. In: *Wildlife and Vegetation of Unmanaged Douglas-fir Forests. General Technical Report PNW-GTR-285* (eds Ruggiero LF, Aubry KB, Carey AB, Huff MH), pp. 221–240. US Department of Agriculture, Forest Service, Portland, Oregon.

Manel S, Gaggiotti OE, Waples RS (2005) Assignment methods: matching biological questions with appropriate techniques. *Trends in Ecology and Evolution*, **20**, 136–142.

Mann DH, Hamilton TD (1995) Late Pleistocene and Holocene paleoenvironments of the North Pacific coast. *Quaternary Science Reviews*, **14**, 449–471.

Martinez JG, Soler JJ, Soler M, Møller AP, Burke T (1999) Comparative population structure and gene flow of a brood parasite, the great spotted cuckoo (*Clamator glandarius*), and its primary host, the magpie (*Pica pica*). *Evolution*, **53**, 269–278.

Miller MP (1997) *Tools for Population Genetic Analysis (TFPGA)*, *version 1.3*. Northern Arizona University, Flagstaff, Arizona.

Nielson M, Lohman K, Sullivan J (2001) Phylogeography of the tailed frog (*Ascaphus truei*): implications for the biogeography of the Pacific Northwest. *Evolution*, **55**, 147–160.

Otter KA, Stewart IRK, McGregor PK, Terry AMR, Dabelsteen T, Burke T (2001) Extra-pair paternity among great tits *Parus major* following manipulation of male signals. *Journal of Avian Biology*, **32**, 338–344.

Paetkau D, Waits LP, Clarkson PL, Craighead L, Strobeck C (1997) An empirical evaluation of genetic distance statistics using microsatellite data from bear (Ursidae) populations. *Genetics*, **147**, 1943–1957.

Piry S, Luikart G, Cornuet J-M (1999) BOTTLENECK: a computer program for detecting reductions in the effective population size using allele frequency data. *Journal of Heredity*, **90**, 502–503.

Price J, Droege S, Price A (1995) *The Summer Atlas of North American Breeding Birds*. Academic Press, San Diego, California.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Pruett CL, Winker K (2005) Northwestern song sparrow populations show genetic effects of sequential colonization. *Molecular Ecology*, **14**, 1421–1434.

Raymond M, Rousset F (1995a) An exact test for population differentiation. *Evolution*, **49**, 1280–1283.

Raymond M, Rousset F (1995b) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Rice WR (1989) Analyzing tables of statistical tests. *Evolution*, **43**, 223–225.

Richardson BA, Brunsfeld SJ, Klopfenstein NB (2002) DNA from bird-dispersed seed and wind-disseminated pollen provides insights into postglacial colonization and population genetic structure of whitebark pine (*Pinus albicaulis*). *Molecular Ecology*, **11**, 215–227.

Ritland C, Pape T, Ritland K (2001) Genetic structure of yellow cedar (*Chamaecyparis nootkatensis*). *Canadian Journal of Botany*, **79**, 822–828.

Root RB (1964) Ecological interactions of the chestnut-backed chickadee following a range expansion. *Condor*, **66**, 229–238.

Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Soltis DE, Gitzendanner MA, Strenge DD, Soltis PS (1997) Chloroplast DNA intraspecific phylogeography of plants from the Pacific Northwest of North America. *Plant Systematics and Evolution*, **206**, 353–373.

Taylor EB, Pollard S, Louie D (1999) Mitochondrial DNA variation in bull trout (*Salvelinus confluentus*) from northwestern North America: implications for zoogeography and conservation. *Molecular Ecology*, **8**, 1155–1170.

Warner BG, Mathewes RW, Clague JJ (1982) Ice-free conditions on the Queen Charlotte Islands, British Columbia, at the height of the late Wisconsin glaciation. *Science*, **218**, 675–677.

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Wisely SM, Buskirk SW, Russell GA, Aubry KB, Zielinski WJ (2004) Genetic diversity and structure of the fisher (*Martes pennanti*) in a peninsular and peripheral metapopulations. *Journal of Mammalogy*, **85**, 640–648.