

Molecular Markers Reveal Limited Population Genetic Structure in a North American Corvid, Clark's Nutcracker (*Nucifraga columbiana*)

Kimberly M. Dohms*, Theresa M. Burg

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

Abstract

The genetic impact of barriers and Pleistocene glaciations on high latitude resident species has not been widely investigated. The Clark's nutcracker is an endemic North American corvid closely associated with *Pinus*-dominated forests. The nutcracker's encompasses known barriers to dispersal for other species, and glaciated and unglaciated areas. Clark's nutcrackers also irruptively disperse long distances in search of pine seed crops, creating the potential for gene flow among populations. Using the highly variable mitochondrial DNA control region, seven microsatellite loci, and species distribution modeling, we examined the effects of glaciations and dispersal barriers on population genetic patterns and population structure of nutcrackers. We sequenced 900 bp of mitochondrial control region for 169 individuals from 15 populations and analysed seven polymorphic microsatellite loci for 13 populations across the Clark's nutcracker range. We used species distribution modeling and a range of phylogeographic analyses to examine evolutionary history. Clark's nutcracker populations are not highly differentiated throughout their range, suggesting high levels of gene flow among populations, though we did find some evidence of isolation by distance and peripheral isolation. Our analyses suggested expansion from a single refugium after the last glacial maximum, but patterns of genetic diversity and paleodistribution modeling of suitable habitat were inconclusive as to the location of this refugium. Potential barriers to dispersal (e.g. mountain ranges) do not appear to restrict gene flow in Clark's nutcracker, and postglacial expansion likely occurred quickly from a single refugium located south of the ice sheets.

Citation: Dohms KM, Burg TM (2013) Molecular Markers Reveal Limited Population Genetic Structure in a North American Corvid, Clark's Nutcracker (*Nucifraga columbiana*). PLoS ONE 8(11): e79621. doi:10.1371/journal.pone.0079621

Editor: Wolfgang Arthofer, University of Innsbruck, Austria

Received: July 3, 2013; **Accepted:** October 3, 2013; **Published:** November 4, 2013

Copyright: © 2013 Dohms, Burg. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The work was supported by the following: Natural Science and Engineering Research Council (NSERC) Discovery Grant (TMB) and Post-Graduate Scholarship D (KMD), <http://www.nserc-crsng.gc.ca/>; Alberta Innovates (AI) New Faculty Award (TMB) and Graduate Scholarship (KMD), <http://www.albertatechfutures.ca/>. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

* E-mail: kim.dohms@uleth.ca

Introduction

Phylogeography is the study of processes underlying spatial and temporal dimensions of genetic variation [1]. Patterns of genetic variation attained through phylogeographic methods can provide insights into population-level response to disturbance and the processes responsible for historical dispersal and colonization [2,3]. Examining patterns of genetic variation can help us to evaluate the roles of gene flow, bottlenecks, and historical or geological barriers in explaining contemporary patterns of diversity across geographical regions [4]. Contemporary genetic patterns are strongly influenced by postglacial expansion from refugia [5,6], historical and contemporary barriers to dispersal [7,8], and dispersal potential [9].

The repeated glaciations and climate oscillations of the Pleistocene epoch provide a natural tool to address population responses to large-scale landscape changes. Multiple

expansions and contractions of ice sheets created a dynamic landscape that repeatedly fragmented historical populations, alternately creating barriers to dispersal and creating new habitats for colonization [10,11]. During the Pleistocene glacial maxima in North America, many populations were confined to refugia (ice-free areas) [10]. Plant and animal species expanded from several known refugia following the retreat of the ice sheets, including Beringia (present-day Bering Sea and parts of Alaska) and three main areas south of the ice sheets (Pacific Coast, Rockies, and Taiga) [10,12]. Additional areas along the periphery of the ice sheets are contested to have been ice-free, such as an Atlantic shelf near present-day Newfoundland [5,10,13].

High latitude resident bird species provide a unique opportunity to investigate patterns of postglacial and barrier-mediated dispersal; historical events shaping current population structure should be particularly evident in these species. Non-migratory species have the potential to retain

patterns of genetic variation longer due to limited dispersal, allowing researchers to make inferences about past historic events [14-16]. For example, species of trees show distinct patterns of population genetic structure and have retained information on historic environmental changes [13,15]. As the number of studies on resident species increase, similar patterns are emerging in vertebrate taxa [16-18]. However, other resident species show limited differentiation despite potential barriers to dispersal [19,20].

Clark's nutcracker (*Nucifraga columbiana*) is a high latitude resident corvid species (Family Corvidae) that inhabits North American coniferous forests. Much of the published work on this species has focused on spatial cognition exhibited by their extensive food caching behaviours [21]. Additional research has concentrated on Clark's nutcracker's role as essential seed dispersers for many conifers, particularly *Pinus* species [22-24]. To date, no published phylogeographic work exists for this species, though genetic work has been identified as a priority [21]. Clark's nutcracker prefers higher altitude montane forests dominated by one or more *Pinus* species. Their close association with pine-dominated forests is due to a specialised diet of pine seeds, which influences dispersal patterns of both nutcrackers and seeds [21,25]. After nestlings fledge, Clark's nutcrackers have been known to undergo seasonal altitudinal movement to nearby *Pinus*-dominated subalpine areas [21]. In times of food shortage, and occasionally during the non-breeding season, large numbers of birds have been shown to irruptively disperse, leaving home ranges, and travelling more than 100 km in search of large crops of pine seeds [21,26]. Range-wide morphological and plumage variation is well documented for some high latitude resident species [27,28], but there is little evidence of morphological variation in Clark's nutcracker [21]. Some suggest that this lack of variation is a reflection of high gene flow between populations due to irruptive dispersal, but this has not been investigated to date [21,29].

To address these gaps in phylogeographic knowledge, our goals for this study were to investigate patterns of dispersal and postglacial colonization in Clark's nutcracker as reflected in genetic structure. We predict that contemporary Clark's nutcracker populations will exhibit low levels of genetic differentiation between populations due to irruptive food-related dispersal patterns. Mountain ranges and unsuitable habitats will not act as strong barriers to gene flow due to this species' preference for high altitude sub-alpine habitats [21] where barriers are limited. Rather, dispersal will be restricted by availability of pine-dominated coniferous forests and seed crops. We expect that postglacial colonization likely occurred from a single refugium south of the ice sheets, similar to patterns of postglacial colonization in several North America *Pinus* species [13].

Materials and Methods

Ethics statement

All animals captured in the field were handled following animal welfare protocols (#0614 and #1028) approved by the University of Lethbridge Animal Welfare Committee using

guidelines set by the Canadian Council on Animal Care (CCAC). Banding in Canada was performed under Canadian Wildlife Service banding permit #10425W (2007) and #10804 (2008 onward) and in the US under US Fish and Wildlife banding permit #23522. All netting and blood sampling was conducted under the appropriate state, provincial, federal, and institutional authorities; detailed permit information is available upon request.

Sample collection and preparation

From 2009-2012, we captured up to 30 individual Clark's nutcrackers at each sampling site (Table 1) using standard mistnetting techniques with call playback. For call playback, we used a medley of group interaction vocalizations sourced from multiple tracks found on xeno-canto.org. Sample collection locations for each defined population (i.e. sampling site) were limited to within a 50 km of radius and were not separated by any obvious barriers to dispersal. We chose sites on either side of possible barriers to dispersal (e.g., Rocky Mountains, Great Basin) and from areas that were previously glaciated and unglaciated during the last glacial maximum (e.g., Central Alberta and New Mexico, respectively; Figure 1). We collected less than 100 μ L of blood from each bird by pricking the brachial vein with a sterile needle and collecting blood in a capillary tube. Blood was stored in 95% ethanol and archived at -80°C upon return to the University of Lethbridge. Each bird was banded with a US Fish & Wildlife aluminum band, aged and sexed (if possible), and mass, tarsus, and other morphological measurements taken. Additional tissue and feather samples were obtained from museum collections taken from birds during the breeding season within the past 20 years to ensure all samples were from contemporary populations (Table S1).

Laboratory procedures

DNA extraction. DNA was extracted from blood, feather, and tissue samples using a modified Chelex protocol [30]. Once extracted, DNA was stored at -20°C in low TE buffer.

Mitochondrial DNA. We amplified a section of the mitochondrial DNA control region (CR) using primers L46 SJ (5'-TTT GGC TAT GTA TTT CTT TGC-3'; Birt & Lemmen, unpublished data) and H1030 JCR 18 [31] corresponding to positions 46 (Domain I) and 1030 (Domain III) of the corvid mitochondrial control region, respectively. Where the complete fragment would not amplify, we used internal primers designed in-house, H560 clnuCR (5'-GCA AAG GGA GGA GTA TGC AG-3') or L530 corvidCR (5'-CGC CTC TGG TTC CTA TTT CA-3'), with L46 SJ or H1030 JCR 18, respectively, to amplify two overlapping fragments.

Polymerase chain reactions (PCR) were performed on a Master gradient thermocycler (Eppendorf: Hauppauge, NY) in a 25 μ L volume containing 1x goTaq Flexi buffer (Promega: Madison, WI), 2.5 mM MgCl₂, 200 μ M dNTP, 0.4 μ M of each primer, and 0.5 units goTaq Flexi taq polymerase (Promega) under the following conditions: one cycle of 94°C for 120 s, 52°C for 45 s, and 72°C for 60 s, 37 cycles of 94°C for 30 s, 52°C for 45 s and 72°C for 60 s and one cycle of 72°C for five

Table 1. Summary table of and mitochondrial DNA information for populations included in this study.

Population	Lat (N)	Long (W)	<i>n</i>	<i>H_n</i>	<i>H_d</i>	π
CAB	52.687	-118.056	8	6	0.893	0.342
SAB	49.774	-114.397	9	5	0.583	0.127
MT	46.592	-112.158	20	18	0.964	0.281
WY	43.756	-110.581	19	13	0.901	0.295
UT	40.703	-110.612	19	10	0.673	0.226
CO	39.813	-106.195	13	10	0.859	0.221
NM	35.815	-106.852	8	7	0.964	0.396
SCA	34.279	-117.551	14	9	0.912	0.312
CCA	37.771	-118.926	11	9	0.964	0.375
NECA	41.151	-120.205	2	2	1.000	0.456
SOR	42.745	-122.144	12	6	0.849	0.354
NEOR	45.243	-117.524	20	15	0.953	0.402
WA	47.554	-120.958	7	6	0.857	0.329
NEWA	48.845	-119.625	6	4	0.800	0.222
SEBC	49.162	-119.270	1	1	-	-
Overall			169		0.854	0.324

Population abbreviations are as follows: CAB = central Alberta, SAB = southern Alberta, MT = Montana, WY = Wyoming, UT = Utah, CO = Colorado, NM = New Mexico, SCA = Southern California, CCA = central California, NECA = northeast California, SOR = southern Oregon, NEOR = northeast Oregon, WA = central Washington, NEWA = northeast Washington, and SEBC = southeast British Columbia. Column headers are: Lat (N) = latitude and Long (W) = longitude of central locations; *n* = number of samples used for mitochondrial DNA analyses; *H_n* = number of haplotypes; *H_d* = haplotype diversity; and π = nucleotide diversity multiplied by 100 for ease of interpretation.

doi: 10.1371/journal.pone.0079621.t001

min. PCR products were run on a 0.8% agarose gel to confirm DNA amplification.

DNA sequencing was performed at McGill University and Génome Québec Innovation Centre on a 3730xl DNA Analyzer (Applied Biosystems: Carlsbad, CA) or at the University of Lethbridge on a 3130 DNA Analyzer (Applied Biosystems). For in-house sequencing we used a shrimp alkaline phosphatase-exonuclease clean up followed by sequencing and sodium acetate precipitation [28] prior to gel electrophoresis.

Microsatellite DNA. Three individuals from geographically distant populations (Montana, southern Oregon, and Colorado) were initially screened for 32 microsatellite primer pairs developed for and used in other corvids. Thirteen additional nutcracker-specific microsatellite loci have since been developed by another independent research group [32]. If screened loci were suspected to be polymorphic and/or amplified clearly, two additional individuals from each of the initial test populations plus three individuals from another population potentially separated by dispersal barriers were screened. Eight loci were polymorphic (Table 2). All forward primers were modified by adding an M13 sequence (5' – CAC

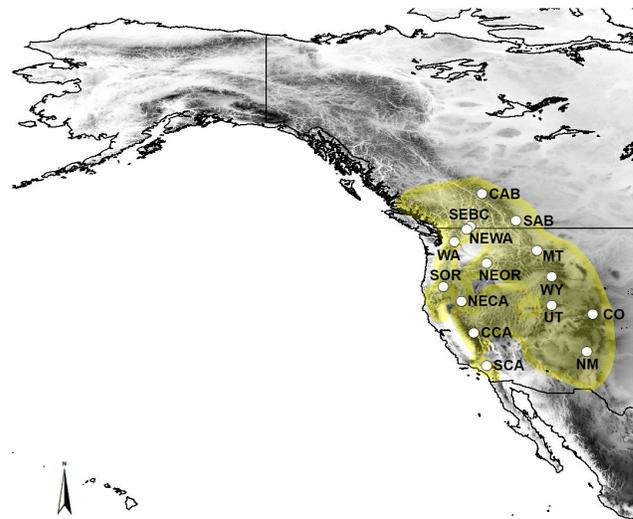


Figure 1. Clark's nutcracker population sampling locations. Current range overlaid in yellow [76]. Darker shades imply higher elevation areas. See Table 1 for coordinates, sample sizes, and abbreviations per population.

doi: 10.1371/journal.pone.0079621.g001

GAC GTT GTA AAA CGA C - 3') to the 5' end, which allowed the integration of a fluorescently labeled primer (700 nm or 800 nm) directly into the PCR product. DNA was amplified in a 10 μ L reaction with 1x buffer, 1 mM MgCl₂, 200 μ M dNTP (Fisher Scientific: Hampton, NH), 1 μ M of each primer (forward and reverse), 0.05 μ M of the fluorescent primer (Eurofins MWG Operon: Huntsville, AL) and 0.5 units taq polymerase under the following conditions: one cycle of 94°C for 120 s, T₁ for 45 s, and 72°C for 60 s, seven cycles of 94°C for 60 s, T₁ for 30 s and 72°C for 45 s, 31 cycles of 94°C for 30 s, T₂ for 30 s, and 72°C for 45 s, and one final elongation cycle at 72°C for five minutes (Table 2).

PCR products were mixed with a stop solution (95% formamide, 20 mM EDTA and bromophenol blue), denatured for 3 min at 94°C, and run on a 6% polyacrylamide gel using a LI-COR 4300 DNA Analyzer (LI-COR Inc.: Lincoln, NE). Alleles were scored via visual inspection, and genotypes were independently confirmed by a second person. Three controls of known allele sizes (pre-screened individuals) plus a size standard were included on each gel to ensure consistent scoring along with a negative control to ensure no contamination was present.

Analyses of genetic structure

Mitochondrial DNA. Chromatograms were edited and sequences aligned manually in MEGA v5.0 [33]. Shared haplotypes were determined using DnaSP v5.10 [34]. We assessed population structure and relationships among haplotypes using a maximum parsimony network constructed using Network v4.611, which is designed to construct the shortest, least complex network [35]. The median-joining (MJ) tree algorithm was used as this allows for multi-state data (e.g., four nucleotide base options at each variable site), with settings

Table 2. Microsatellite primer pairs used in this study.

Primer	Sequence (5' to 3')	Buffer	taq	T ₁	T ₂
ApCo19F	CAG ACT GCA GTC TTG CTA TAG C	Flexi	crimson	45	48
ApCo19R	GCC TTG GAT GCT TTT ACG				
ApCo30F	GCC CTG ATG CTG TTG ATG GT	Flexi	Flexi	45	48
ApCo30R	CTG GAG CCT GGT TTA GAG TTA TGC				
ApCo37F	TGC CAA ATG CAA CCA AAT CTT	Flexi	Flexi	50	52
ApCo37R	CAT CAC TTG CAG AGA GGG CA				
ApCo41F	CCT ACT CTG GGC ACT GTT ATT ATC	crimson	crimson	50	52
ApCo41R	CCC ATT ATC AGC ATG TCG TAC A				
ApCo46F	GGG AGC CTA GTA TGT TAA GAT GCT	Flexi	crimson	50	52
ApCo46R	TTC CAG GTG AGG TGA TTC AGC				
ApCo91F	GTA GTC CCA ATG GTT TCT CTG TC	crimson	crimson	45	48
ApCo91R	GAT GAA GTA ATG TGA AAC CTG G				
PnuA106WF	GTA TTT TGG GAT GTC TTA GGG TTG	Flexi	Flexi	50	52
PnuA106WR	CAC ACT CGA GCT ACA ATA AAC CTG				

Primer names, sequence, focal species, source, and reaction conditions for this study (buffer and taq polymerase, and annealing temperatures, T₁ and T₂ (°C)). Forward primers (shown with an 'F' suffix) were modified to include a short sequence at the 5' end allowing for incorporation of the fluorescent tag. All loci except PnuA106W are sourced from Florida Scrub Jay (*Aphelocoma coerulescens*) [77]. PnuA106W is sourced from Yellow-billed Magpie (*Pica nuttalli*) [78].

doi: 10.1371/journal.pone.0079621.t002

as follows: weight = 10 for each site, epsilon = 0, and the MJ squared option turned off. After the tree was constructed, the MP (maximum parsimony) option was used to remove superfluous links and unnecessary median vectors from the network, which can be produced during the initial network calculations [36].

Genetic variation within populations was measured by calculating haplotype (H_d) and nucleotide (π) diversity using DnaSP v5.10 [34]. We calculated pairwise Φ_{ST} values (an analogue of Wright's fixation index (F_{ST})) using Arlequin v3.5.1.3 [37] to examine population structure and test for genetic differentiation among populations and haplogroups. Significance values were corrected using a Benjamini-Hochberg correction [38], to control for false discovery rate (FDR), the expected proportion of falsely rejected hypotheses in multiple significance testing situations where multiplicity might occur.

A Mantel's test was performed to examine the correlation between genetic and geographic distances in GenAlEx v6.0 [39]. We calculated geographic distances using weighted central coordinates for each population (Table 1) and the Geographic Distance Matrix Generator v1.2.3 [40]. Linearised Φ_{ST} values ($\Phi_{ST} / 1 - \Phi_{ST}$) were used for genetic distances and

significance was tested using 9,999 permutations. A mismatch distribution was used to help visualize signatures of demographic expansion, and to test the null hypothesis of historic population growth and expansion [41].

Genetic variation allocated within and among populations was examined using an analysis of molecular variance (AMOVA) in Arlequin v3.5.1.3 [37]. We investigated the potential impact of barriers on genetic variation by performing a spatial analysis of molecular variance (SAMOVA; $K = 2-12$; 100 iterations) [42]; however, the process cannot test for range-wide panmixia as K cannot be set to $K = 1$. Alternatively, BAPS v6.0 [43], a Bayesian method for analysing population structure, can detect range-wide panmixia. A preliminary population mixture analysis for spatial clustering of individuals in BAPS used no *a priori* population information and $K_{max} = 13$. A subsequent BAPS run was then conducted using *a priori* population information.

Microsatellite DNA. Allele frequencies, deviation from Hardy-Weinberg equilibrium (observed (H_o) versus expected (H_e) heterozygosity), and pairwise F_{ST} values [44] were calculated with 999 permutations using GenAlEx v6.0 [39] and P values were corrected for multiple tests using a Benjamini-Hochberg correction for FDR [38]. Allelic richness was calculated in FSTAT v2.9.3 [45]. A Mantel's test was performed to test for isolation by distance using the same procedure as for mitochondrial DNA.

Bayesian clustering analyses were conducted using the programs STRUCTURE v2.3.3 [46,47] and BAPS v6.0 [48]. STRUCTURE was run with the following settings: $K = 1-15$, a burn-in of 100,000 followed by 500,000 runs, admixture assumed, correlated allele frequencies and including *a priori* population information 10 replicates were performed for each value of K . In STRUCTURE, it can be difficult to decide when K captures major structure in the data due to similar $\ln P(X|K)$ values, thus two additional analyses were conducted using STRUCTURE results: Bayes factor calculations [46] and ΔK [49] using STRUCTURE HARVESTER [50]. One limitation with using STRUCTURE HARVESTER is that it cannot detect when $K = 1$ is the true number of clusters. However, Bayes factors and BAPS are able to detect if $K = 1$, so the addition of these methods is useful in species where panmixia is a possibility. A preliminary population mixture analysis for spatial clustering of individuals in BAPS used no *a priori* population information and $K_{max} = 13$. A subsequent BAPS run was then conducted using *a priori* population information.

Species distribution modeling

Sample points. In addition to our sample (field and museum) locations, geo-referenced Clark's nutcracker locations were obtained from the Global Biodiversity Information Facility (GBIF; <http://data.gbif.org/>, accessed on 22 January 2013). We excluded occurrences outside of North America, without geo-references, or recorded before 1950 from modeling and further inspected accuracy by plotting points using ArcMap v10.1 (ESRI: Redlands, CA). We removed duplicate records prior to model-building.

Bioclimatic data. Current bioclimatic data were extracted from the WORLDCLIM dataset (v1.4, [PLOS ONE | www.plosone.org](http://</p>
</div>
<div data-bbox=)

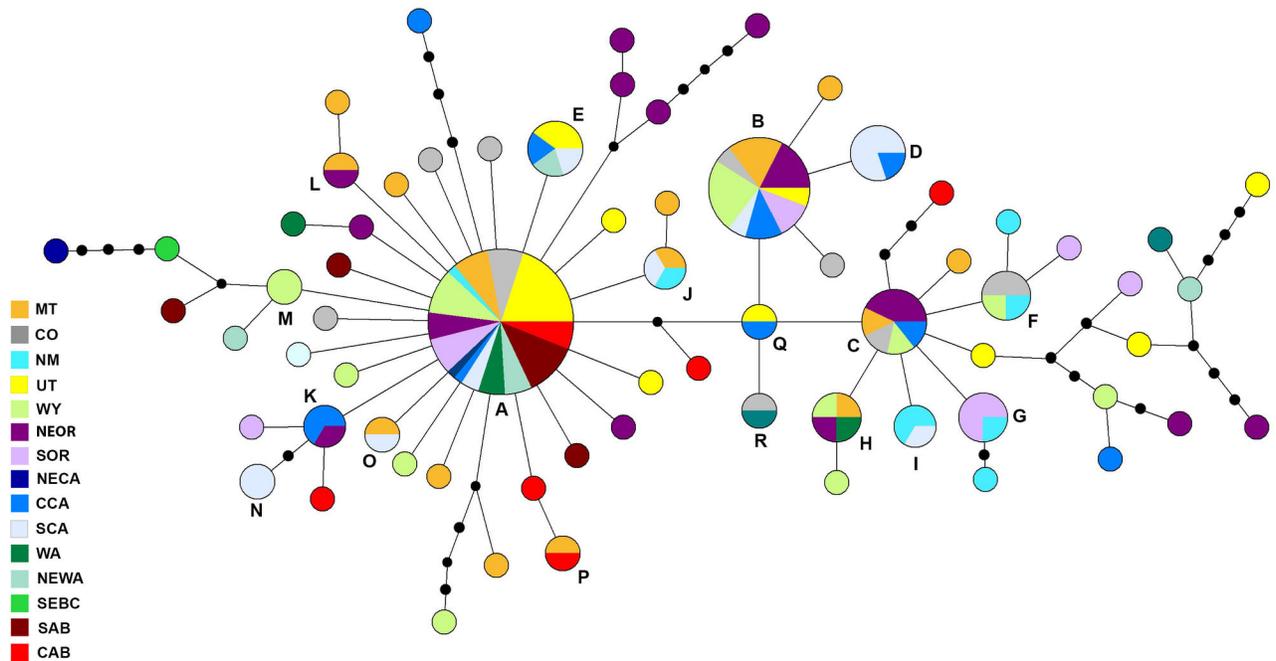


Figure 2. Clark's nutcracker mitochondrial DNA haplotype network. Legend: Maximum parsimony network constructed using a median joining algorithm and post-processed with a maximum parsimony procedure for a 900 base pair fragment of mitochondrial control region from 169 Clark's nutcrackers across North America. Each colour represents a different population ($n = 15$). Circle size is proportional to number of individuals with that haplotype. Haplotypes represented as pie charts include individuals from multiple populations with pie segments proportional to the number of individuals from each population. Capital letters are those assigned to shared haplotypes (see Table 3). Small black circles represent inferred nodes. See Table 1 for population abbreviations.

doi: 10.1371/journal.pone.0079621.g002

www.worldclim.org/) and LGM bioclimatic data from the Model for Interdisciplinary Research on Climate (MIROC) database [51] at 2.5 min resolution (~4 x 4 km tiles). The current bioclimatic dataset ranges over a 50 year period (1950 - 2000), hence exclusion of observations prior to 1950. Nineteen bioclimatic variables are included in the WORLDCLIM current and LGM dataset [52]. ArcMap v10.1 (ESRI) was used to clip climatic variable layers to include only North America as using smaller geographic areas can improve the predictive power of Maxent models [53]. Prior to constructing SDM, ENMTools v1.3 [54] was used to determine which bioclimatic variables were correlated using $R > 0.90$ as a cutoff. Nine variables were correlated with at least one other variable and all but one from each set of correlated variables was removed.

Maximum entropy (Maxent) distribution modeling. Maxent v3.3.3 [55] was used to model suitable habitat distributions for Clark's nutcracker at present and during the LGM using hinge features only, regularization multiplier = 1, max number of background points = 10,000, replicate run type of 10 cross-validations, 500 maximum iterations, and 0.00001 convergence threshold. Resulting layers were then exported to ArcMap v10.1, overlaid on a digital elevation map, and processed into images.

Results

Genetic analyses

Mitochondrial DNA. Mitochondrial DNA control region sequences ($n = 169$; Table S1) from 15 populations range-wide revealed 48 variable sites in a 900 base pair (bp) region. In total, 68 haplotypes were recognized in the median-joining analysis (GenBank accession nos. KF687612-KF687679; Figure 2; Table 3). Eighteen haplotypes accounted for 119 individuals; 50 singleton haplotypes were found (Figure 2; Table 3). The largest haplotype group contained a total of 50 individuals from 14 of the 15 populations analysed. Most haplotypes were a single mutational step removed from each other. Overall, limited geographic clustering was observed in the maximum parsimony network (Figure 2).

Haplotype diversity (H_d) values ranged from 0.583 (southern Alberta (SAB)) to 0.964 (Montana (MT), New Mexico (NM), and central California (CCA)) and nucleotide diversity (π) ranged from 0.00127 (SAB) to 0.00402 (northeastern Oregon) for populations with greater than five samples (Table 1). When all samples were combined, overall H_d was 0.854, and overall π was 0.00324 indicating high levels of genetic diversity in Clark's nutcracker. In 78 population pairwise comparisons of genetic differentiation, Φ_{ST} values ranged from -0.050 ($P_{corrected} = 0.779$) for Utah (UT) and northeast Washington (NEWA) to

Table 3. Geographic distribution of shared haplotypes.

Haplotype	CAB	SAB	MT	WY	UT	CO	NM	SCA	CCA	NECA	SOR	NEOR	WA	NEWA	SEBC	Total
A	3	6	4	5	10	4	1	2	1	1	4	3	3	3	-	50
B	-	-	3	4	1	1	-	1	2	-	2	3	-	-	-	17
C	-	-	1	1	-	1	-	-	1	-	-	3	-	-	-	7
D	-	-	-	-	-	-	-	4	1	-	-	-	-	-	-	5
E	-	-	-	-	2	-	-	1	1	-	-	-	-	1	-	5
F	-	-	-	1	-	2	1	-	-	-	-	-	-	-	-	4
G	-	-	-	-	-	-	1	-	-	-	3	-	-	-	-	4
H	-	-	1	1	-	-	-	-	-	-	-	1	1	-	-	4
I	-	-	-	-	-	-	2	1	-	-	-	-	-	-	-	3
J	-	-	1	-	-	-	1	1	-	-	-	-	-	-	-	3
K	-	-	-	-	-	-	-	-	2	-	-	1	-	-	-	3
L	-	-	1	-	-	-	-	-	-	-	-	1	-	-	-	2
M	-	-	-	2	-	-	-	-	-	-	-	-	-	-	-	2
N	-	-	-	-	-	-	-	2	-	-	-	-	-	-	-	2
O	-	-	1	-	-	-	-	1	-	-	-	-	-	-	-	2
P	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	2
Q	-	-	-	-	1	-	-	-	1	-	-	-	-	-	-	2
R	-	-	-	-	-	1	-	-	-	-	-	-	1	-	-	2
Total	4	6	13	14	14	9	6	13	9	1	9	12	5	4	0	119
Unique	4	3	7	5	5	4	2	1	2	1	3	8	2	2	1	50

Haplotype codes correspond to those found in Figure 2. Population abbreviations and locations are found in Table 1.

doi: 10.1371/journal.pone.0079621.t003

0.359 ($P_{\text{corrected}} = 0.100$) for NM and SAB (Table 4). Most population pairs (57/78) had Φ_{ST} values less than 0.050, though 14 pairs had Φ_{ST} values between 0.050 and 0.150, four pairs had Φ_{ST} values between 0.150 and 0.250, and three pairs had Φ_{ST} values above 0.250. Only four of those Φ_{ST} values were significant after the FDR correction and all occurred between NM and other populations (UT, Wyoming (WY), southern California (SCA), and NEWA). The mismatch distribution did not deviate significantly from the expected signature, so we cannot reject the null hypothesis of historical population growth and expansion [41]. Geographic distance in 78 paired populations ranged from 339 km for WY and UT to 2074 km for NM and central Alberta (CAB). A low, but significant correlation between genetic and geographic distance was detected for mitochondrial DNA ($R^2 = 0.082$, $P = 0.040$; Figure 3). When the NM population was removed from the analyses, the relationship between genetic and geographic distance still trended towards isolation by distance, though it was no longer significantly correlated ($R^2 = 0.064$, $P = 0.058$; not shown).

Both AMOVA and SAMOVA showed that most genetic variation is contained within rather than among population groups. The highest value of between-group variation was observed when $K = 1$ for AMOVA analyses (3.4% between groups, 96.6% within groups) and for $K = 2$ in SAMOVA (17.8% between groups, 80.7% within groups), with one group consisting of NM and the other of all other populations pooled. The preliminary (no *a priori* population information) and secondary (using *a priori* population information) BAPS analyses for spatial clustering of individuals found no significant differentiation between individuals or populations (i.e., $K = 1$).

Microsatellite DNA. A total of seven polymorphic loci (Table 5) were used for analyses; an eighth locus (ApCo37) was removed due to limited amplification success and subsequent missing data (~50%). We used the thirteen populations with greater than five samples for microsatellite analyses (Table 1; Table S1). The total number of alleles detected ranged from four in PnuA106W to 10 in ApCo40 with an average of 6.43. Only 13 of the 91 loci-population comparisons showed significant deviations from Hardy-Weinberg and all showed lower heterozygosity than expected. ApCo41 has a significant heterozygote deficit in five of the 13 populations sampled and while two of these (NEWA and CAB) had fewer samples ($n \leq 8$), the others had a larger number of samples ($n \geq 14$; Table 5).

Limited population differentiation was detected with a global F_{ST} value of 0.070. Paired F_{ST} values ranged from 0.000 (14/78 comparisons) to 0.123 (CCA and central Alberta (CAB); Table 6). After Benjamini-Hochberg FDR correction, 21 of 78 paired comparisons were significant (Table 6). A weak, but significant, correlation was found between genetic and geographic distance using Mantel's test ($R^2 = 0.078$, $P = 0.029$; Figure 4).

STRUCTURE results suggested that the optimal number of groups was $K = 1$ (Figure 5A). Further analyses produced peaks at $K = 1$, $K = 4$, and $K = 12$ (Figure 5B), though $K = 1$ produced the highest peak and highest posterior probabilities (mean (Pr ($K=1$)) = -2625.45) and is supported by Bayes factor values (Pr ($K=1$) = 1) and BAPS results.

Species distribution modeling. A total of 1563 presence records were used for training and 174 for testing the distribution models. After removing correlated bioclimatic layers, 10 layers were used for model construction (bio1-4, 8,

Table 4. Population pairwise Φ_{ST} values for Clark's nutcracker mitochondrial DNA.

MT	CO	NM	UT	WY	NEOR	SOR	CCA	SCA	WA	NEWA	SAB	CAB	
MT		0.832	0.100	0.509	0.714	0.373	0.316	0.509	0.373	0.825	0.530	0.468	0.574
CO	-0.026		0.270	0.414	0.883	0.546	0.714	0.771	0.414	0.771	0.373	0.373	0.414
NM	0.186	0.147		<0.001	<0.001	0.234	0.509	0.289	<0.001	0.373	<0.001	0.100	0.234
UT	0.005	0.017	0.281		0.589	0.296	0.289	0.509	0.296	0.714	0.779	0.473	0.270
WY	-0.012	-0.031	0.156	0.001		0.462	0.509	0.771	0.373	0.771	0.654	0.479	0.414
NEOR	0.023	0.003	0.099	0.042	0.007		0.383	0.756	0.270	0.756	0.414	0.289	0.414
SOR	0.039	-0.016	0.009	0.072	0.012	0.013		0.779	0.373	0.714	0.373	0.100	0.366
CCA	0.008	-0.025	0.102	0.011	-0.014	-0.015	-0.031		0.654	0.771	0.513	0.270	0.456
SCA	0.022	0.023	0.202	0.043	0.034	0.045	0.063	-0.013		0.691	0.499	0.334	0.366
WA	-0.041	-0.035	0.102	-0.019	-0.032	-0.020	-0.025	-0.032	0.003		0.771	0.373	0.714
NEWA	0.005	0.037	0.282	-0.050	-0.005	0.031	0.080	0.010	0.023	-0.046		0.876	0.509
SAB	0.017	0.077	0.359	0.017	0.019	0.066	0.153	0.076	0.059	0.032	-0.034		0.289
CAB	0.003	0.028	0.180	0.055	0.028	0.029	0.061	0.021	0.048	-0.018	0.029	0.066	

Φ_{ST} values are below diagonal, significance values corrected for false discovery rate above diagonal. Values are based on 110 permutations for mitochondrial DNA control region in 166 Clark's nutcrackers from 13 populations ($n \geq 5$) in North America. Bolded values indicate significance. See Table 1 and Figure 1 for sampling locations and abbreviations.

doi: 10.1371/journal.pone.0079621.t004

12, 14-15, 18-19). Mean area under the receiver operating curve (AUC) was 0.907 (range = 0.906 - 0.908), suggesting that the current model was highly suitable for backcasting to the paleodistribution model. Isothermality (mean diurnal temperature range/mean annual temperature range (bio3)), temperature seasonality (bio4), and mean annual temperature (bio1) were the most important variables contributing to the model at 44.3%, 37.7%, and 11.1% respectively.

The current distribution predicted by the Maxent model (Figure 6A) closely approximated the known Clark's nutcracker range in North America (Figure 1). The model returned a maximum 66% probability of suitable habitat throughout the range. The paleodistribution model returned a maximum 25% probability of predicted suitable habitat, mostly concentrated in the western mountainous regions and along the southeast edge of the ice sheets during the last glacial maximum (Figure 6B).

Discussion

Population genetic structure

Using a combination of measures of population differentiation and clustering analyses, our study suggests high levels of population connectivity between most Clark's nutcracker populations in North America. We found limited differentiation in mitochondrial DNA (Table 4), with only the New Mexico population showing significant pairwise differences from four other populations. Microsatellite data indicate higher levels of population differentiation between Clark's nutcracker populations, though this predominantly occurs between populations at the edges of the species' range (e.g. CAB and SCA; Table 6). Geographic distance rather than barriers appear to influence mitochondrial and nuclear gene flow, as there is a weak, but significant effect of isolation by distance on nutcracker populations for both molecular markers

(Figures 4 and 5). However, for mitochondrial DNA, this pattern seems to be driven by populations in New Mexico, suggesting limited peripheral isolation for this marker. For microsatellite data, additional peripheral populations (CAB, SAB, and SCA) are significantly differentiated in multiple comparisons (Table 6), suggesting peripheral isolation in addition to isolation by distance. Clustering analyses do not support strong population differentiation for either nuclear or mitochondrial markers, with a single population being the best fit for the data in both cases. In the case of microsatellite data, these results should be interpreted with caution as both STRUCTURE and BAPS may struggle with situations where differentiation is low ($F_{ST} < 0.030$) [56], though global values ($F_{ST} = 0.070$) suggest this is likely not the case.

High levels of haplotype, nucleotide, and allelic diversity throughout most nutcracker populations suggest large population sizes and few founder effects or population bottlenecks. The exception to this is the southern Alberta population, which exhibits reduced haplotype and nucleotide diversity for mitochondrial markers relative to the rest of the populations, though microsatellite loci have high levels of allelic richness and heterozygosity (Table 5). This difference between marker genetic diversity may be due to a historical founder effect in previously glaciated southern Alberta (mtDNA), with nuclear gene flow being less limited (microsatellites).

Patterns of potential panmixia and limited population differentiation are not unique to Clark's nutcracker. Many other species that undergo irruptive and seasonal dispersal also exhibit limited population differentiation despite the presence of proposed barriers to dispersal. Studies of *Pinus* seed specialists, such as pygmy nuthatch (*Sitta pygmaea*) [57], crossbill species (*Loxia spp.*) [19], and Clark's nutcracker's sister species, Eurasian nutcracker (*Nucifraga caryocatactes*) [58], found no significant genetic differentiation among populations despite potential barriers to dispersal. Limited

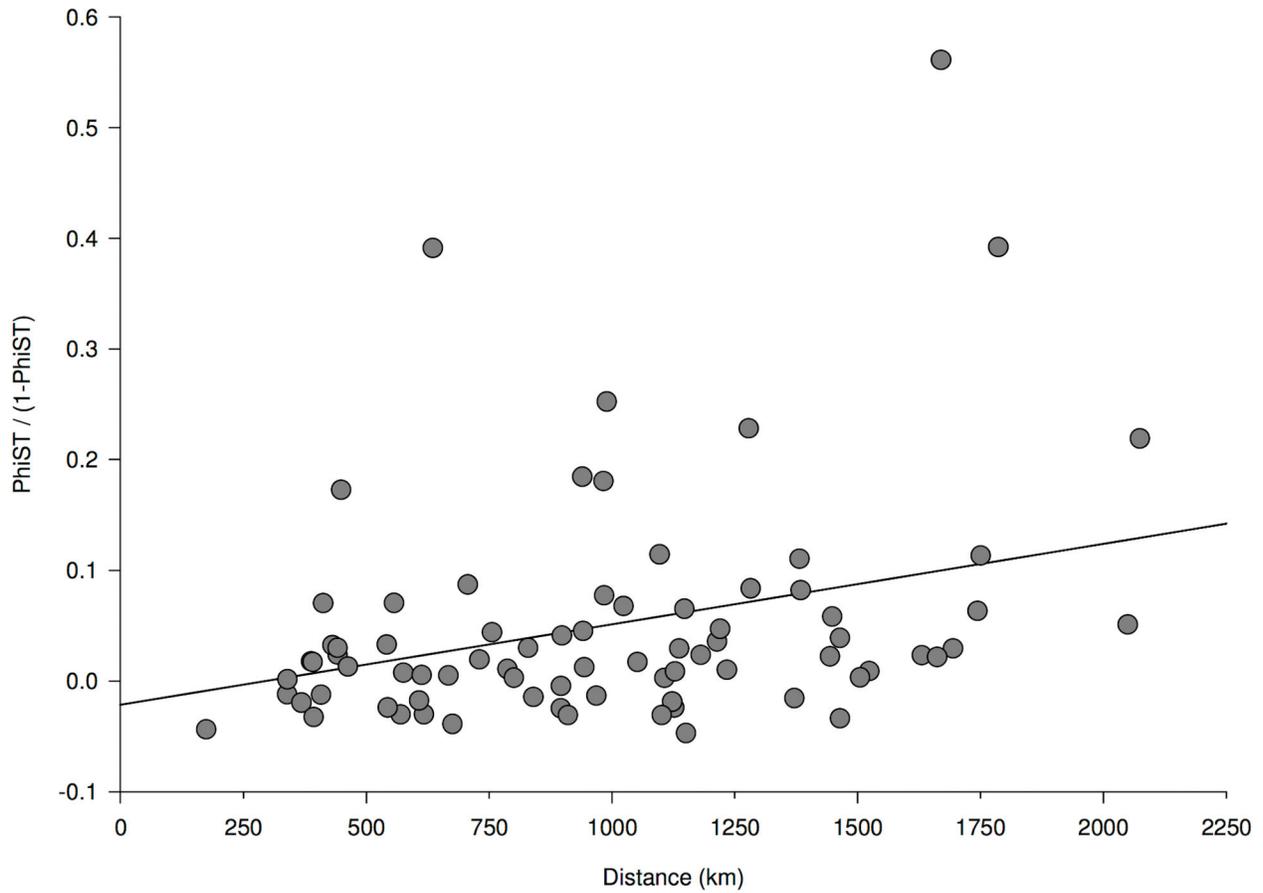


Figure 3. Mantel test of isolation by distance for mitochondrial DNA. Legend: $R^2 = 0.082$, $P = 0.040$. Each point represents one population pairwise $\Phi_{ST} / (1 - \Phi_{ST})$ plotted against geographic distance between paired populations.

doi: 10.1371/journal.pone.0079621.g003

Table 6. Population pairwise F_{ST} values for Clark's nutcracker.

	MT	CO	NM	UT	WY	NEOR	SOR	CCA	SCA	WA	NEWA	SAB	CAB
MT		0.185	0.454	0.205	0.100	0.454	0.210	0.032	0.028	0.454	0.311	0.093	0.196
CO	0.016		0.238	0.189	0.035	0.189	0.205	0.017	0.035	0.454	0.189	0.268	0.016
NM	0.001	0.017		0.084	0.189	0.454	0.378	0.196	0.454	0.454	0.454	0.414	0.189
UT	0.009	0.016	0.033		0.454	0.189	0.454	0.023	0.016	0.454	0.142	0.017	0.093
WY	0.012	0.028	0.017	0.000		0.081	0.189	0.016	0.016	0.454	0.072	0.017	0.035
NEOR	0.000	0.015	0.000	0.012	0.016		0.167	0.024	0.050	0.454	0.454	0.189	0.210
SOR	0.012	0.017	0.007	0.001	0.013	0.020		0.058	0.024	0.467	0.293	0.031	0.032
CCA	0.040	0.072	0.021	0.063	0.059	0.054	0.036		0.234	0.093	0.394	0.017	0.016
SCA	0.038	0.044	0.000	0.071	0.066	0.039	0.062	0.014		0.196	0.454	0.093	0.024
WA	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.043	0.026		0.454	0.358	0.267
NEWA	0.014	0.033	0.000	0.036	0.040	0.000	0.017	0.010	0.000	0.000		0.196	0.072
SAB	0.026	0.011	0.004	0.045	0.046	0.017	0.052	0.090	0.034	0.008	0.027		0.037
CAB	0.020	0.075	0.032	0.034	0.041	0.020	0.066	0.123	0.096	0.021	0.071	0.057	

F_{ST} values are below diagonal and significance values corrected for false discovery rate are above diagonal. Values are based on 110 permutations using seven polymorphic microsatellite loci in 187 Clark's nutcrackers from 13 populations in North America. See Table 2 and Figure 1 for population abbreviations.

doi: 10.1371/journal.pone.0079621.t006

Table 5. Summary table of seven microsatellite loci used to analyze 13 Clark's nutcracker populations.

	ApCo19	ApCo30	ApCo40	ApCo41	ApCo46	ApCo91	PnuA106W
Central Alberta (n=8)							
A _n	3	3	5	2	4	4	2
A _r	2.359	2.361	3.160	1.692	2.967	2.963	1.429
Ho	0.571	0.625	0.750	0.000	0.429	0.625	0.143
He	0.500	0.555	0.625	0.245	0.602	0.648	0.133
P	ns	ns	ns	**	ns	ns	ns
Southern Alberta (n=12)							
A _n	3	3	5	2	6	2	2
A _r	2.190	2.274	3.039	2.000	4.018	1.846	1.908
Ho	0.500	0.400	0.583	0.667	0.750	0.143	0.417
He	0.497	0.535	0.611	0.444	0.806	0.337	0.413
P	ns						
Montana (n=25)							
A _n	3	3	6	4	5	5	3
A _r	2.142	2.566	3.430	1.835	3.446	3.062	1.878
Ho	0.640	0.440	0.840	0.320	0.783	0.750	0.320
He	0.486	0.586	0.734	0.282	0.738	0.684	0.351
P	ns	*	ns	Ns	ns	ns	ns
Wyoming (n=30)							
A _n	3	4	9	3	5	3	3
A _r	2.022	2.510	3.937	1.620	2.686	2.314	1.800
Ho	0.483	0.357	0.724	0.136	0.550	0.400	0.241
He	0.463	0.543	0.793	0.208	0.556	0.460	0.313
P	ns	ns	**	***	*	*	ns
Utah (n=20)							
A _n	2	4	6	3	5	3	3
A _r	1.962	2.558	3.247	1.891	2.926	2.426	1.862
Ho	0.579	0.294	0.650	0.316	0.471	0.571	0.278
He	0.478	0.528	0.700	0.342	0.619	0.538	0.323
P	ns	***	ns	Ns	*	ns	ns
Colorado (n=13)							
A _n	3	3	6	2	5	4	3
A _r	2.247	2.353	3.509	1.273	3.643	2.632	2.062
Ho	0.385	0.417	0.769	0.091	0.750	0.444	0.385
He	0.462	0.517	0.728	0.087	0.760	0.512	0.411
P	ns						
New Mexico (n=9)							
A _n	3	3	5	2	4	3	2
A _r	2.241	2.317	3.452	1.867	3.084	2.467	1.975
Ho	0.667	0.778	0.750	0.400	0.429	0.400	0.111
He	0.475	0.549	0.719	0.320	0.663	0.460	0.475
P	ns	ns	ns	Ns	ns	ns	*
Southern California (n=14)							
A _n	3	3	7	3	5	4	2
A _r	2.111	2.375	3.615	1.646	3.265	2.645	1.976
Ho	0.429	0.571	1.000	0.077	0.571	0.636	0.286
He	0.457	0.487	0.737	0.210	0.694	0.566	0.490
P	ns	ns	ns	**	ns	ns	ns
Central California (n=11)							
A _n	3	3	7	2	4	4	3
A _r	1.909	2.231	4.233	1.834	3.173	2.961	2.247
Ho	0.364	0.545	0.636	0.273	0.800	0.636	0.636
He	0.310	0.517	0.818	0.351	0.685	0.665	0.533

Table 5 (continued).

	ApCo19	ApCo30	ApCo40	ApCo41	ApCo46	ApCo91	PnuA106W
P	ns	ns	ns	Ns	ns	ns	ns
Southern Oregon (n=12)							
A _n	2	5	6	3	4	4	2
A _r	1.940	2.894	3.759	2.202	2.843	2.645	1.908
Ho	0.500	0.750	1.000	0.300	0.750	0.417	0.417
He	0.444	0.597	0.760	0.405	0.618	0.549	0.413
P	ns	ns	ns	Ns	ns	ns	ns
Northeast Oregon (n=20)							
A _n	2	3	8	3	6	4	2
A _r	1.970	2.681	3.294	1.891	3.798	2.406	1.764
Ho	0.450	0.600	0.700	0.105	0.750	0.500	0.100
He	0.489	0.625	0.704	0.342	0.790	0.510	0.320
P	ns	ns	ns	**	ns	ns	**
Washington (n=7)							
A _n	2	3	7	2	3	3	3
A _r	1.992	2.622	4.045	1.500	2.758	2.682	2.359
Ho	0.500	0.571	0.857	0.167	0.500	0.500	0.714
He	0.486	0.571	0.765	0.153	0.625	0.569	0.500
P	ns	ns	ns	Ns	ns	ns	ns
Northeast Washington (n=6)							
A _n	2	3	4	2	5	3	2
A _r	1.998	2.545	2.969	1.773	3.924	2.545	1.992
Ho	0.333	0.667	0.667	0.000	0.833	0.500	0.833
He	0.500	0.500	0.625	0.278	0.764	0.500	0.486
P	ns	ns	ns	*	ns	ns	ns
Overall	A _n = 10	A _n = 5	A _n = 10	A _n = 7	A _n = 7	A _n = 7	A _n = 4

Only populations with greater than five samples were used; *n* = number of samples used in genotyping and analyses; A_n = number of alleles; A_r = allelic richness; Ho = observed and He = expected heterozygosity; P = departures from Hardy-Weinberg equilibrium (ns = not significant, **P* < 0.05, ***P* < 0.01, ****P* < 0.001. See Table 1 for population locations.

doi: 10.1371/journal.pone.0079621.t005

differentiation has been found in other species with irruptive or seasonal dispersal, such as boreal owl (*Aegolius funereus*) [59], and European redpoll species (*Carduelis spp.*) [60], yet Steller's jay (*Cyanocitta stelleri*) [14] and white-breasted nuthatch (*Sitta carolinensis*) [61] exhibit significant differentiation between populations and limited gene flow corresponding with barriers to dispersal. In other taxa, peripheral populations are more likely to be isolated due to reduced gene flow [16,62,63], while maintaining panmixia in the majority of the range [64]. Overall, Clark's nutcracker population connectivity seems to be primarily limited by distance with some peripheral isolation at the southern and northern edges of the range, which is reasonable for a high elevation species with irruptive and seasonal altitudinal dispersal.

Postglacial colonization

After the LGM, species that rapidly expanded from a single refugium with high levels of gene flow are expected to exhibit three characteristics: 1) a star-like phylogeny with many low

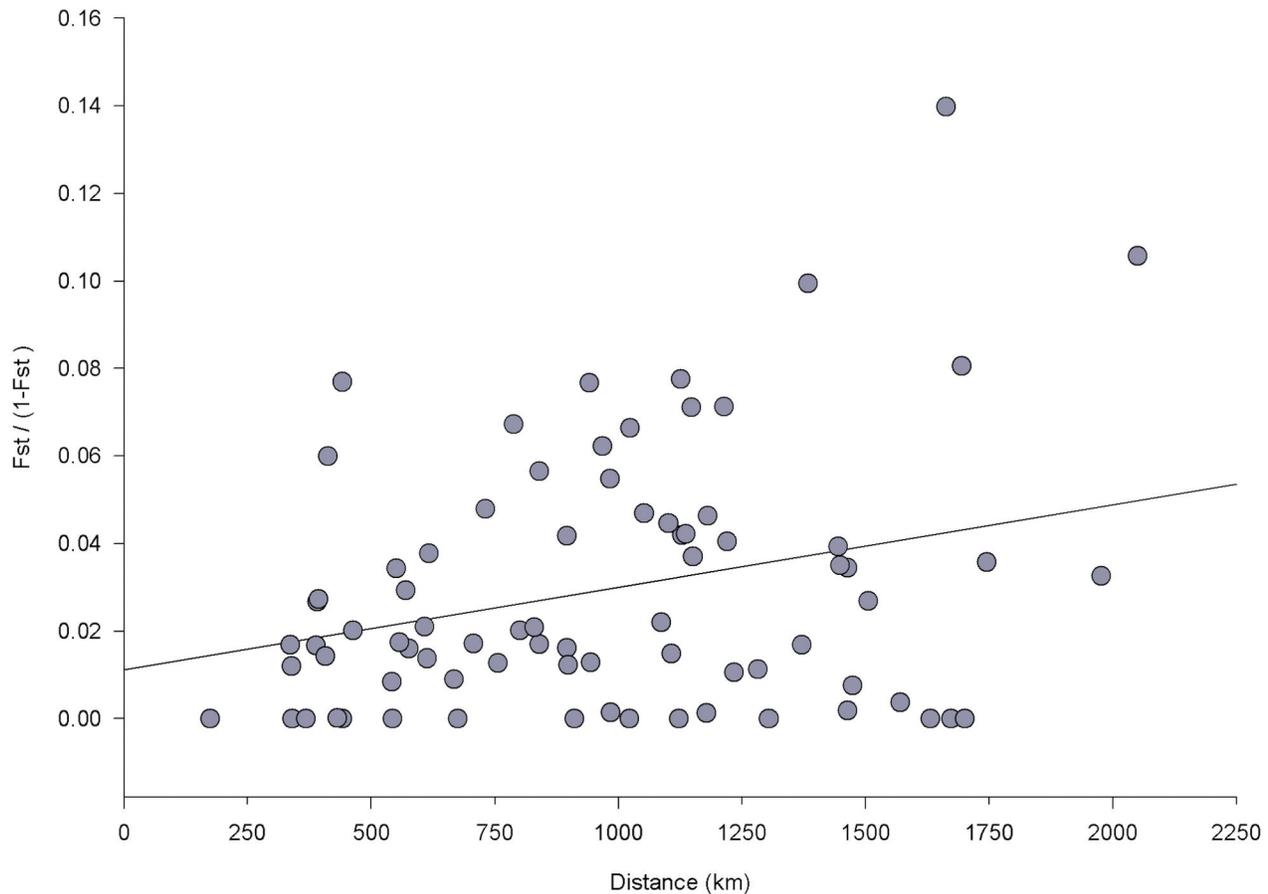


Figure 4. Mantel test of isolation by distance for seven microsatellite loci. Legend: $R^2 = 0.078$, $P = 0.029$. Each point represents one population pairwise $F_{ST} / (1 - F_{ST})$ plotted against straight line geographic distance between paired populations.

doi: 10.1371/journal.pone.0079621.g004

frequency single haplotypes separated from high frequency central ancestral haplotypes by few mutational steps [65]; 2) low levels of genetic subdivision between and within populations [4,66]; and 3) a mismatch distribution of pairwise differences among haplotypes indicating a sudden increase in expansion from a single population [67]. These signals of expansion from a single refugium have been reported in several North American bird species (e.g., multiple species reviewed in 68, downy woodpecker (*Picoides pubescens*) [20]) as well as many plant species, including several *Pinus* species that Clark's nutcracker specializes upon (reviewed in 13). Clark's nutcracker mitochondrial DNA exhibits a star-like phylogeny with one major shared haplotype, and short connections between this haplotype and others (Figure 2), in addition to a modeled paleodistribution that suggests that this species expanded from a single refugium after the LGM. This pattern is less complex than that found in whitebark pine postglacial expansion: colonization is thought to have occurred from three refugia (Oregon, Idaho, and Colorado) [25]. Given

that Clark's nutcrackers range extends farther south than that of whitebark pine, nutcrackers may have been isolated in a more southern refugium and expanded into whitebark pine refugia from there, subsequently assisting with seed distribution northward.

In addition to the aforementioned characteristics, temperate species that expanded north from a southern refugium have been shown to exhibit genetic differentiation between northern and southern populations corresponding to previously unglaciated and glaciated areas [12]. Southern populations were less impacted by glaciation and should show higher levels of nucleotide diversity and phylogeographic structure compared to northern populations [69]. This pattern occurs to varying degrees in *Pinus* species on which Clark's nutcracker commonly feed: whitebark pine (*Pinus albicaulis*) [25], white pine (*P. monticola*) [70], limber pine (*P. flexilis*) [71,72], and Ponderosa pine (*P. ponderosa*) [73,74]. However, this pattern does not distinctly appear in Clark's nutcracker. All southern and northern populations sampled exhibit high levels of

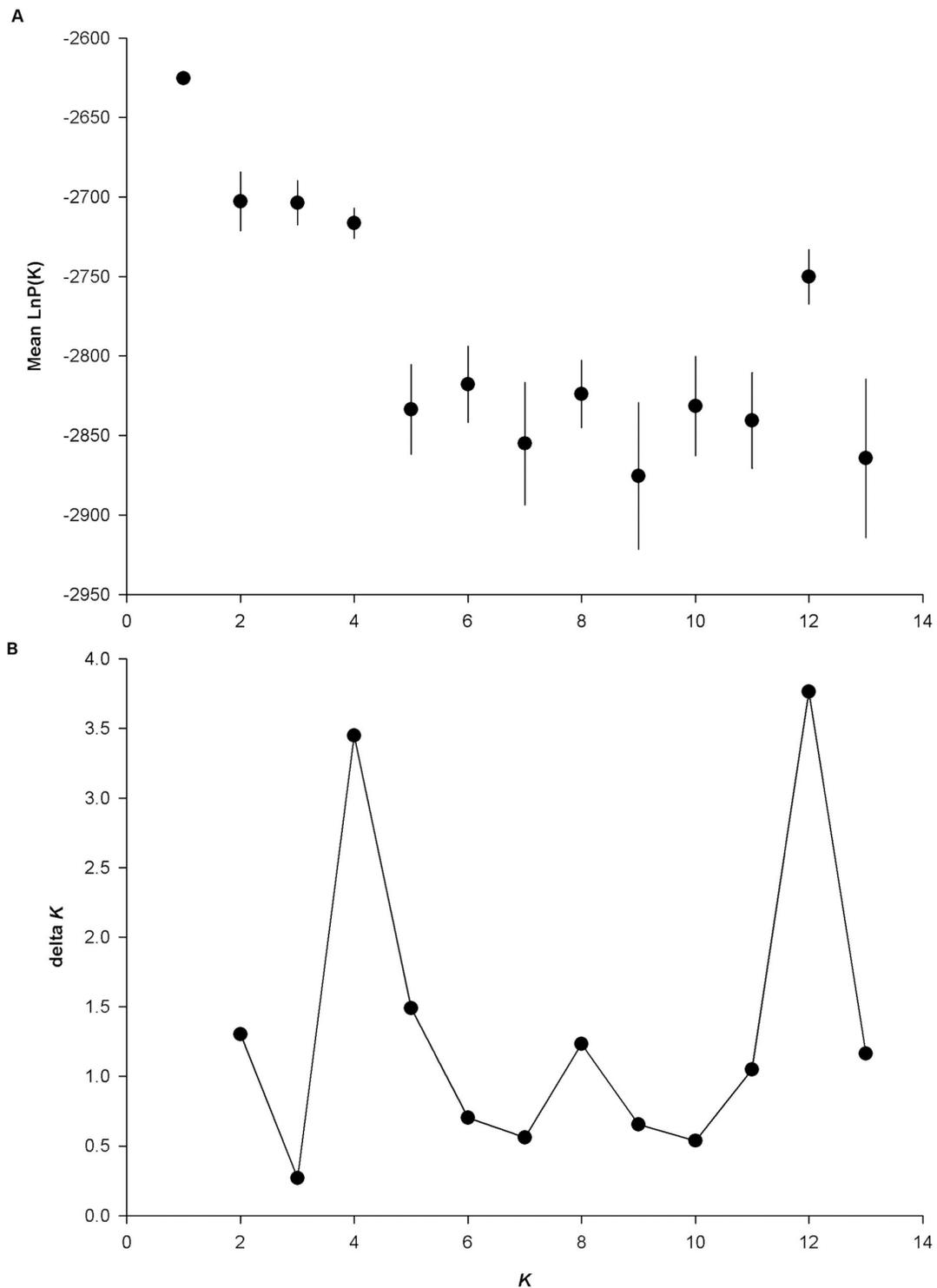


Figure 5. Measures of optimal K values from STRUCTURE analyses. Legend: A) Penalised log likelihood test (mean \pm SE) and B) ΔK depicting clusters (K). The penalised log likelihood test takes the maximum $\ln \Pr(X | K)$ as the correct number of clusters and ΔK infers the number of clusters from the difference between $\ln \Pr(X | K)$.

doi: 10.1371/journal.pone.0079621.g005

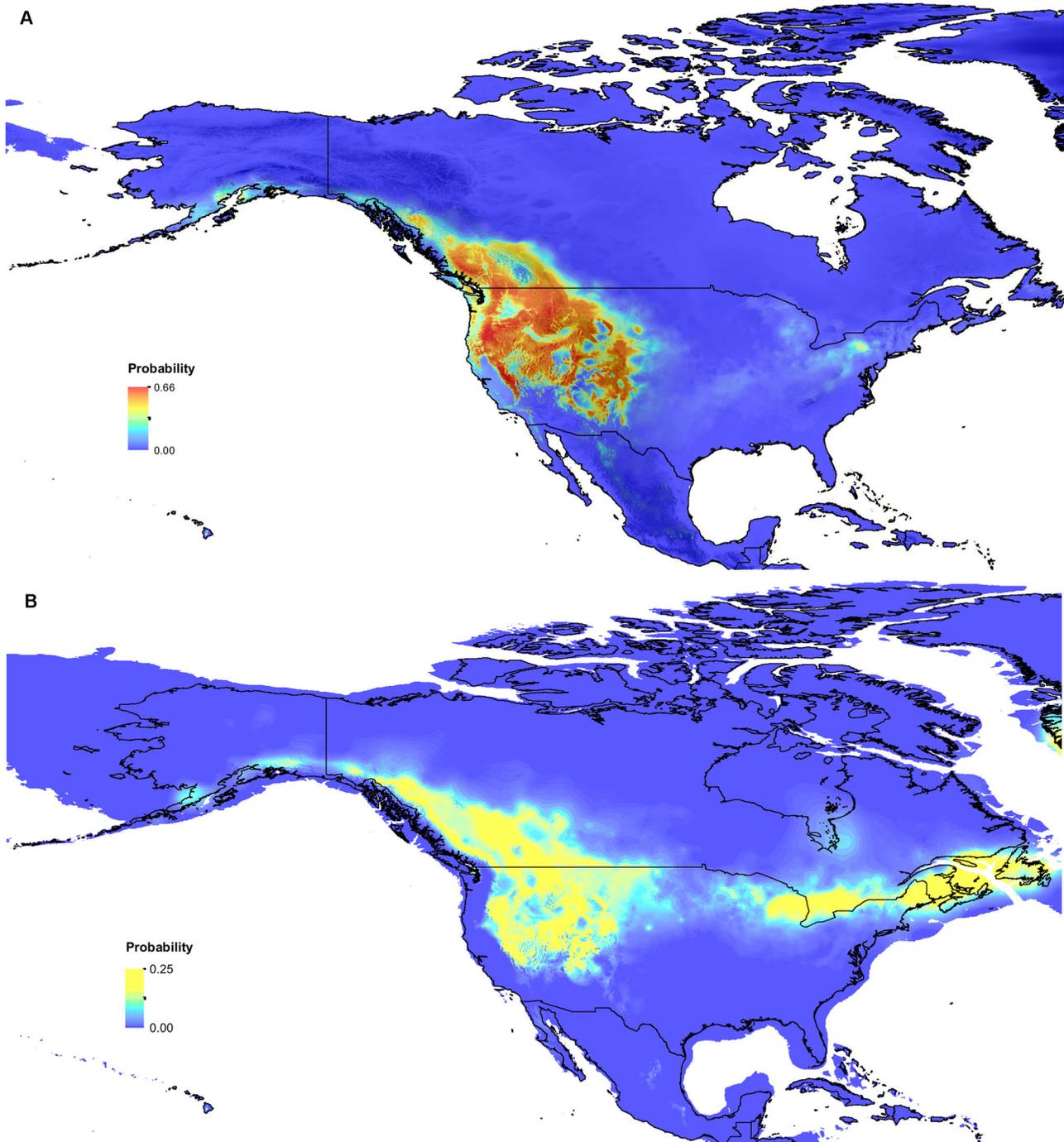


Figure 6. Current and paleodistribution models of suitable Clark's nutcracker habitat in North America. Legend: Probability of suitable Clark's nutcracker habitat at A) present (overlaid on digital elevation model and B) during last glacial maximum (~21 kya), overlaid on current North American boundaries generated by species distribution modeling. In B), blue areas outside of current coastlines represent historical extent of terrestrial landmass.

doi: 10.1371/journal.pone.0079621.g006

heterozygosity, nucleotide and allelic diversity, and limited differentiation from other populations, making it very

challenging to pinpoint a single refugium for this species. In addition, while species distribution models for current suitable

nutcracker habitat performed very well (high AUC values), paleodistribution models returned reduced probabilities (< 30%) of suitable LGM habitat throughout unglaciated areas. Taken together, these results point to a single refugium or multiple small, connected refugia, the location of which is not obvious through genetic signatures or paleodistribution modeling.

Conclusions and future research

Throughout its range, Clark's nutcracker exhibits low levels of population differentiation and does not appear to be limited by potential barriers to dispersal, though peripheral populations may be slightly isolated. Future research could integrate additional mitochondrial markers to elucidate the location of a glacial refugium or refugia for this species. Mitochondrial markers with lower rates of mutation in corvids (e.g. cytochrome B) [31] may retain historic genetic signals longer and thus provide additional historic information to this study. Samples at the northern and isolated southern extremes of the range may also increase our understanding of peripheral isolation and postglacial colonization for this species. Given the importance of Clark's nutcracker to dispersal of *Pinus* species in North America [21,22] and current conservation and management concerns for many *Pinus* species [75], further understanding of nutcracker dispersal may also assist in future forest management and recovery plans.

Supporting Information

Table S1. Summary table of Clark's Nutcracker samples used in analyses. Sequence (mtDNA): Y = successfully sequenced. N = not sequenced. Genotyped (SSR): Y =

successfully genotyped. N = not genotyped. Sources include Burg lab (wild), TS = T. Schaming, The Field Museum (TFM), Burke Museum at the University of Washington (UWBM), Museum of Southwest Biology (MSB), Louisiana State University Museum of Natural Sciences (LSU), American Museum of Natural History (AMNH), Smithsonian National Museum of Natural History (USNM), Royal Saskatchewan Museum (RSM) and Royal Alberta Museum (RAB). (XLS)

Acknowledgements

We thank R. Adams, N. Freeman, K. Gohdes, B. Graham, H. Hansen, J. Hindley, E. Koran, L. Lait, C. Macfarlane, K. Nielson, A. Martin, H. Pirot, P. Pulgarin-Restrepo and others for their invaluable and enthusiastic help in the field. We are grateful for samples provided by American Museum of Natural History, Burke Museum of Natural History and Culture – University of Washington, Smithsonian Museum of Natural History, Louisiana State University Museum of Natural Science, Museum of Southwest Biology, Royal Alberta Museum, Royal Saskatchewan Museum, and the Field Museum. T. Schaming of Cornell University graciously provided samples from Wyoming. A. Iwaniuk provided valuable comments during preparation of this manuscript.

Author Contributions

Conceived and designed the experiments: KMD TMB. Performed the experiments: KMD. Analyzed the data: KMD. Contributed reagents/materials/analysis tools: TMB. Wrote the manuscript: KMD TMB.

References

1. Avise JC (2004) Molecular Markers, Natural History, and Evolution. Sunderland, MA: Sinauer Associates, Inc.. p. 684.
2. Milot E, Gibbs HL, Hobson KA (2000) Phylogeography and genetic structure of northern populations of the yellow warbler (*Dendroica petechia*). Mol Ecol 9: 667-681. doi:10.1046/j.1365-294x.2000.00897.x. PubMed: 10849283.
3. Moritz C (2002) Strategies to protect biological diversity and the evolutionary processes that sustain it. Syst Biol 51: 238-254. doi: 10.1080/10635150252899752. PubMed: 12028731.
4. Zink RM (1996) Comparative phylogeography in North American birds. Evolution 50: 308-317. doi:10.2307/2410802.
5. Weir JT, Schluter D (2004) Ice sheets promote speciation in boreal birds. Proc R Soc Lond B Biol Sci 271: 1881-1887. doi:10.1098/rspb.2004.2803. PubMed: 15347509.
6. Williams JW (2003) Variations in tree cover in North America since the last glacial maximum. Global Planet Change 35: 1-23. doi:10.1016/S0921-8181(02)00088-7.
7. Brunsfeld SJ, Sullivan J, Soltis DE, Soltis PS (2001) Comparative geography of northwestern North America: a synthesis. In: J Silvertown J Antonovics. Integrating Ecology and Evolution in a Spatial Context. Williston, VT: Blackwell Publishing. pp. 319-339.
8. Keyghobadi N (2007) The genetic implications of habitat fragmentation for animals. Can J Zool 85: 1049-1064. doi:10.1139/Z07-095.
9. Burg TM, Lomax J, Almond R, de L, Brooke M, Amox W (2003) Unravelling dispersal patterns in an expanding population of a highly mobile seabird, the northern fulmar (*Fulmarus glacialis*). Proc R Soc Lond B Biol Sci 270: 979-984. doi:10.1098/rspb.2002.2322.
10. Pielou EC (1991) After the ice age: the return of life to glaciated North America. Chicago, Illinois: The University of Chicago Press. 366 pp.
11. Dynesius M, Jansson R (2000) Evolutionary consequences of changes in species' geographical distributions driven by Milankovitch climate oscillations. Proc Natl Acad Sci U S A 97: 9115-9120. doi:10.1073/pnas.97.16.9115. PubMed: 10922067.
12. Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. Philos Trans R Soc Lond B: Biol Sci 359: 183-195. doi: 10.1098/rstb.2003.1388. PubMed: 15101575.
13. Jaramillo-Correa JP, Beaulieu J, Khasa DP, Bousquet J (2009) Inferring the past from the present phylogeographic structure of North American forest trees: seeing the forest for the genes. Can J Forest Res 39: 286-307. doi:10.1139/X08-181.
14. Burg TM, Gaston AJ, Winker K, Friesen VL (2005) Rapid divergence and postglacial colonization in western North American Steller's jays (*Cyanocitta stelleri*). Mol Ecol 14: 3745-3755. doi:10.1111/j.1365-294X.2005.02710.x. PubMed: 16202093.
15. Petit RJ, Duminil J, Fineschi S, Hampe A, Salvini D et al. (2005) Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. Mol Ecol 14: 689-701. PubMed: 15723661.
16. Burg TM, Gaston AJ, Winker K, Friesen VL (2006) Effects of Pleistocene glaciations on population structure of North American chestnut-backed chickadees. Mol Ecol 15: 2409-2419. doi:10.1111/j.1365-294X.2006.02957.x. PubMed: 16842415.
17. Barrowclough GF, Groth JG, Mertz LA, Gutiérrez RJ (2004) Phylogeographic structure, gene flow and species status in blue grouse (*Dendragapus obscurus*). Mol Ecol 13: 1911-1922. doi:10.1111/j.1365-294X.2004.02215.x. PubMed: 15189213.
18. Galbreath KE, Hafner DJ, Zamudio KR, 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. doi:10.1111/j.1365-2699.2009.02201.x.
19. Piertney SB, Summers R (2001) Microsatellite and mitochondrial DNA homogeneity among phenotypically diverse crossbill taxa in the UK.

- Proc R Soc Lond B 268: 1511-1517. doi:10.1098/rspb.2001.1015. PubMed: 11454296.
20. Pulgarin R, Burg TM (2012) Genetic signals of demographic expansion in downy woodpecker (*Picoides pubescens*) after the last North American glacial maximum. PLOS ONE 7: e40412. doi:10.1371/journal.pone.0040412. PubMed: 22792306.
 21. Tomback DF (1998) Clark's nutcracker (*Nucifraga columbiana*). Birds N Am 24.
 22. Tomback DF (1982) Dispersal of whitebark pine seeds by Clark's nutcracker: a mutualism hypothesis. J Anim Ecol 51: 451-467. doi: 10.2307/3976.
 23. Lorenz TJ, Sullivan KA (2009) Seasonal differences in space use by Clark's Nutcrackers in the Cascade Range. Condor 111: 326-340. doi: 10.1525/cond.2009.080070.
 24. Lorenz TJ, Sullivan KA (2010) Comparison of survey methods for monitoring Clark's Nutcrackers and predicting dispersal of whitebark pine seeds. J Field Ornithol 81: 430-441. doi:10.1111/j.1557-9263.2010.00299.x.
 25. Richardson BA, Brunfeldt 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*). Mol Ecol 11: 215-227. doi:10.1046/j.1365-294X.2002.01435.x. PubMed: 11856423.
 26. Newton I (2006) Advances in the study of irruptive migration. Ardea 94: 433-460.
 27. Klicka J, Spellman GM, Winker K, Chua V, Smith BT (2011) A phylogeographic and population genetic analysis of a widespread, sedentary North American bird: The hairy woodpecker (*Picoides villosus*). Auk 128: 346-362. doi:10.1525/auk.2011.10264.
 28. Graham BA, Burg TM (2012) Molecular markers provide insight into contemporary and historic gene flow in a non-migratory species. J Avian Biol 43: 198-214. doi:10.1111/j.1600-048X.2012.05604.x.
 29. Pyle P (1997) Identification guide to North American Birds. Part I: Columbidae to Ploceidae. Bolinas, CA: Slate Creek Press. 732 pp.
 30. Burg TM, Croxall JP (2001) Global relationships amongst black-browed and grey-headed albatrosses: analysis of population structure using mitochondrial DNA and microsatellites. Mol Ecol 10: 2647-2660. doi: 10.1046/j.0962-1083.2001.01392.x. PubMed: 11883879.
 31. Saunders MA, Edwards SV (2000) Dynamics and phylogenetic implications of mtDNA control region sequences in New World jays (Aves: Corvidae). J Mol Evol 51: 97-109. PubMed: 10948266.
 32. Oyler-McCance SJ, Fike JA, Castoe TA, Tomback DF, Wunder MB, Schaming TD (2013) Development and characterization of thirteen microsatellite loci in Clark's nutcracker (*Nucifraga columbiana*). Conserv Genet Resources 5: 303-305. doi:10.1007/s12686-012-9789-0.
 33. Tamura K, Peterson D, Peterson N, Stecher G, Nei M et al. (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28: 2731-2739. doi:10.1093/molbev/msr121. PubMed: 21546353.
 34. Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19: 2496-2497. doi:10.1093/bioinformatics/btg359. PubMed: 14668244.
 35. Bandelt H-J, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16: 37-48. doi:10.1093/oxfordjournals.molbev.a026036. PubMed: 10331250.
 36. Polzin T, Daneshmand SV (2003) On Steiner trees and minimum spanning trees in hypergraphs. Oper Res Lett 31: 12-20. doi:10.1016/S0167-6377(02)00185-2.
 37. Excoffier L, Lischer HE (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10: 564-567. doi:10.1111/j.1755-0998.2010.02847.x. PubMed: 21565059.
 38. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B Stat Methodol 57: 289-300.
 39. Peakall ROD, Smouse PE (2006) GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6: 288-295. doi:10.1111/j.1471-8286.2005.01155.x.
 40. Ersts PJ (2011) Geographic Distance Matrix Generator (version 1.2.3) American Museum of Natural History. Center for Biodiversity and Conservation Available from http://biodiversityinformatics.amnh.org/open_source/gdmg (Accessed on 17 September 2011).
 41. Rogers AR (1995) Genetic evidence for a Pleistocene population explosion. Evolution 49: 608-615. doi:10.2307/2410314.
 42. Dupanloup I, Schneider S, Excoffier L (2002) A simulated annealing approach to define the genetic structure of populations. Mol Ecol 11: 2571-2581. doi:10.1046/j.1365-294X.2002.01650.x. PubMed: 12453240.
 43. Cheng L, Connor TR, Sirén J, Aanensen DM, Corander J (2013) Hierarchical and spatially explicit clustering of DNA sequences with BAPS software. Mol Biol Evol 30: 1224-1228. doi:10.1093/molbev/mst028. PubMed: 23408797.
 44. Wright S (1978) Evolution and the genetics of populations. Variability within and among natural populations. Vol. 4. Chicago, IL, USA: University of Chicago Press.
 45. Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9. 3). Lausanne, Switzerland <http://www.unil.ch/izea/software/fstat.html>
 46. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155: 945-959. PubMed: 10835412.
 47. Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164: 1567-1587. PubMed: 12930761.
 48. Corander J, Marttinen P, Sirén J, Tang J (2008) Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. BMC Bioinformatics 9: 539. doi:10.1186/1471-2105-9-539. PubMed: 19087322.
 49. Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14: 2611-2620. doi:10.1111/j.1365-294X.2005.02553.x. PubMed: 15969739.
 50. Earl DA, von Holdt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resources 4: 1-3. doi:10.1007/s12686-011-9460-1.
 51. Hasumi H, Emori S (2004). p. K-1. coupled GCM (MIROC) description. Tech. Report 1. Tokyo: Center for Climate System Research, University of Tokyo.
 52. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. Int J Climatol 25: 1965-1978. doi:10.1002/joc.1276.
 53. Anderson RP, Raza A (2010) The effect of the extent of the study region on GIS models of species geographic distributions and estimates of niche evolution: preliminary tests with montane rodents (genus *Nephelomys*) in Venezuela. J Biogeogr 37: 1378-1393. doi: 10.1111/j.1365-2699.2010.02290.x.
 54. Warren DL, Glor RE, Turelli M (2010) ENMTools: a toolbox for comparative studies of environmental niche models. Ecography 33: 607-611.
 55. Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. Ecol Modell 190: 231-259. doi:10.1016/j.ecolmodel.2005.03.026.
 56. Latch EK, Dharmarajan G, Glaubitz JC, Rhodes OE (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. doi:10.1007/s10592-005-9098-1.
 57. Spellman GM, Klicka J (2006) Testing hypotheses of Pleistocene population history using coalescent simulations: phylogeography of the pygmy nuthatch (*Sitta pygmaea*). Proc R Soc Lond B 273: 3057-3063. doi:10.1098/rspb.2006.3682. PubMed: 17015345.
 58. Haring E, Gamauf A, Kryukov A (2007) Phylogeographic patterns in widespread corvid birds. Mol Phylogenet Evol 45: 840-862. doi: 10.1016/j.ympev.2007.06.016. PubMed: 17920300.
 59. Koopman ME, Hayward GD, McDonald DB, Winker K (2007) High connectivity and minimal genetic structure among North American boreal owl (*Aegolius funereus*) populations, regardless of habitat matrix. Auk 124: 690-704. doi: 10.1642/0004-8038(2007)124[690:HCAMGS]2.0.CO;2.
 60. Marthinsen G, Wennerberg L, Lifjeld JT (2008) Low support for separate species within the redpoll complex (*Carduelis flammea-hornemanni-cabaret*) from analyses of mtDNA and microsatellite markers. Molecular Phylogenetics Evolution 47: 1005-1017. doi: 10.1016/j.ympev.2008.03.027. PubMed: 18439841.
 61. Spellman GM, Klicka J (2007) Phylogeography of the white-breasted nuthatch (*Sitta carolinensis*): diversification in North American pine and oak woodlands. Mol Ecol 16: 1729-1740. doi:10.1111/j.1365-294X.2007.03237.x. PubMed: 17402986.
 62. Eckert CG, Samis KE, Loughheed SC (2008) Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond. Mol Ecol 17: 1170-1188. doi:10.1111/j.1365-294X.2007.03659.x. PubMed: 18302683.
 63. Lait LA, Friesen VL, Gaston AJ, Burg TM (2012) The post-Pleistocene population genetic structure of a western North American passerine:

- the chestnut-backed chickadee. *J Avian Biol* 43: 541-552. doi: 10.1111/j.1600-048X.2012.05761.x.
64. Stenzler LM, Makarewicz CA, Coulon A, Ardia DR, Lovette IJ et al. (2009) Subtle edge-of-range genetic structuring in transcontinentally distributed North American tree swallows. *Condor* 111: 470-478. doi: 10.1525/cond.2009.080052.
 65. Slatkin M, Hudson RR (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* 129: 555-562. PubMed: 1743491.
 66. 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*: 1898-1914.
 67. Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Mol Biol Evol* 9: 552-569. PubMed: 1316531.
 68. Zink RM (1997) Phylogeographic studies of North American birds. In: DP Mindell. *Avian Molecular Evolution and Systematics*. San Diego, CA: Academic Press. pp. 301-324.
 69. Milá B, Girman DJ, Kimura M, Smith TB, Milá B et al. (2000) Genetic evidence for the effect of a postglacial population expansion on the phylogeography of a North American songbird. *Proc R Soc Lond B Biol Sci* 267: 1033-1040. doi:10.1098/rspb.2000.1107. PubMed: 10874754.
 70. Kim M-S, Richardson BA, McDonald GI, Klopfenstein NB (2011) Genetic diversity and structure of western white pine (*Pinus monticola*) in North America: a baseline study for conservation, restoration, and addressing impacts of climate change. *Tree Genet Genomes* 7: 11-21. doi:10.1007/s11295-010-0311-0.
 71. Mitton JB, Kreiser BR, Latta RG (2000) Glacial refugia of limber pine (*Pinus flexilis* James) inferred from the population structure of mitochondrial DNA. *Mol Ecol* 9: 91-97. doi:10.1046/j.1365-294x.2000.00840.x. PubMed: 10652078.
 72. Jørgensen S, Hamrick JL, Wells PV (2002) Regional patterns of genetic diversity in *Pinus flexilis* (Pinaceae) reveal complex species history. *Am J Bot* 89: 792-800. doi:10.3732/ajb.89.5.792. PubMed: 21665679.
 73. Latta RG, Mitton JB (1999) Historical separation and present gene flow through a zone of secondary contact in Ponderosa pine. *Evolution*: 769-776.
 74. Johansen AD, Latta RG (2003) Mitochondrial haplotype distribution, seed dispersal and patterns of postglacial expansion of Ponderosa pine. *Mol Ecol* 12: 293-298. PubMed: 12492897.
 75. McKinney ST, Tomback DF, Fiedler CE (2011) Altered species interactions and implications for natural regeneration in whitebark pine communities. In: RE KeaneDF TombackMP MurrayCM Smith. *The future of high-elevation, five-needle white pines in Western North America: Proceedings of the High Five Symposium*. Fort Collins, CO: Rocky Mountain Research Station. pp. 376
 76. Ridgely RS, Allnutt TF, Brooks T, McNicol DK, Mehlman D et al. (2007) Digital distribution maps of the birds of the western hemisphere, version 3.0. Arlington, VA, USA: NatureServe.
 77. Stenzler LM, Fitzpatrick JW (2002) Isolation of microsatellite loci in the Florida scrub jay *Aphelocoma coerulescens*. *Mol Ecol Notes* 2: 547-550. doi:10.1046/j.1471-8286.2002.00312.x.
 78. Ernest HB, Well JA, Kurushima JD (2008) Development of 10 microsatellite loci for yellow-billed magpies (*Pica nuttalli*) and corvid ecology and West Nile virus studies. *Mol Ecol Resour* 8: 196-198. doi: 10.1111/j.1471-8286.2007.01921.x. PubMed: 21585754.