Avian diversification across space

van Els, Paul

Chapter III: Extensive gene flow characterizes the phylogeography of a North American migrant bird: Black-headed Grosbeak (*Pheucticus melanocephalus*)

Paul van Els [1], Garth M. Spellman [2], Brian Tilston Smith [1], John Klicka [1]

[1]University of Nevada Las Vegas, Marjorie Barrick Museum of Natural History, 4505 S. Maryland Parkway, Las Vegas, NV 89154, USA [2] Center for the Conservation of Biological Resources, School of Natural Sciences, Black Hills State University, Spearfish, SD 57799, USA

Published in Molecular Phylogenetics and Evolution (2014)

Abstract

We describe range-wide phylogeographic variation in the Black-headed Grosbeak (*Pheucticus melanocephalus*), a songbird that is widely distributed across North American scrublands and forests. Phylogenetic analysis of mitochondrial DNA (mtDNA, n = 424) revealed three geographically structured clades. One widespread clade occurs throughout the Rocky Mountains, Great Basin, and Mexican Plateau, a second clade is found on the Pacific coast and in coastal ranges; and, a third in the Sierra Madre del Sur of Oaxaca and Guerrero. Some geographical structuring occurs in Mexican Plateau and Sierra Madre Oriental mtDNA clade, presumably because these populations have been more stable over time than northern populations. Multiple mitochondrial groups are found sympatrically in the Okanogan River Valley in Washington, the eastern Sierra Nevada, and the Transvolcanic Belt across central Mexico, indicating that there is a potential for introgression. Analyses of 12 nuclear loci did not recover the same geographically structured clades. Population analyses show high levels of gene flow in nucDNA from the Interior into the Sierra Madre del Sur and Pacific population groups, possibly indicating expansion of the Interior population at the expense of peripheral populations.

1. Introduction

The phylogeography of (mostly montane) western North American taxa has received increased attention recently. Western North America is a favored area for phylogeography, because of its relative ease of access to sampling localities, isolation and multitude of biomes, and habitat heterogeneity. These properties have led to studies of a number of widespread species with morphologically, behaviorally or ecologically well-defined subspecies or populations, with several

geographically isolated populations across mountain ranges, deserts, coastal scrub, and grasslands (Demboski and Cook, 2001; Nielson et al., 2001; Burns et al., 2007; Milà et al., 2007a; Spellman and Klicka, 2007; Spellman et al., 2007, Gugger et al., 2011, Manthey et al. 2011; McCormack et al., 2011; Van Els et al., 2012). The examination of phylogeographic structure in these species has resulted in the observation multiple reciprocally monophyletic mitochondrial DNA (mtDNA) groups (in some cases supported by nuclear DNA), isolated from each other by natural barriers.

Species that have continuous distributions across the North American West are, however, underrepresented in phylogeographic studies. Highly vagile taxa (volant organisms including birds) in particular are assumed to show little phylogeographic variation through extensive gene flow in landscapes without major breaks in suitable habitat, such as grasslands, continuous forests, and oceanic environments (Horne et al., 2008; Zink et al., 2002). However, even if habitat is continuous across the landscape and throughout the distribution of the studied taxon, phylogeographic variation may build up through historic demographic processes. In several western North American taxa, Pleistocene refugia played important roles in the buildup of genetic variation (Barrowclough et al., 2004; Demboski and Cook, 2001; Manthey et al., 2011; Nielson et al., 2001; Spellman et al., 2007). These refugia are generally found outside of the current range of the taxa studied (although see Van Els et al., 2012), and in many instances led to isolation of lineages and the resulting accrual of distinct genetic variation. Secondary contact, often involving mechanisms that 'erase' phylogeographic structure, frequently followed the retreat of Pleistocene glaciers, and current continuous distributions across homogeneous habitats may contain the genetically distinct legacy of glaciation that is informative about the historic demography of taxa. Aside from dispersal and resulting gene flow, migratory behavior is an additional factor that may lead to increased gene flow and therefore presumably to decreased phylogeographic structure, particularly in birds (Montgomery, 1896). Most current evidence points towards maintenance of phylogeographic structure within avian species in the face of migratory movements (Ruegg and Smith, 2002; Joseph et al., 2003). However, migration may also lead to gene flow between isolated intraspecific migratory populations, or to increased gene flow from migratory populations into sedentary populations of the same species.

The Black-headed Grosbeak (*Pheucticus melanocephalus*) is a migratory songbird distributed in western North American and found in forested habitats over a wide elevational range, including oak (*Quercus*) scrub, cottonwood (*Populus*) groves, openings in pine forest (*Pinus*), and pinyon-juniper (*Pinus-Juniperus*) belts (Grinnell and Miller, 1944; Hill, 1995). Northern populations of the Black-headed Grosbeak undertake medium to long-distance annual migration, whereas in much of Mexico populations are sedentary (Hill, 1988a). In addition to their annual movements between breeding and wintering grounds, juvenile Black-headed Grosbeaks exhibit low natal philopatry (Hill, 1988a). There are several landscape elements within the species' distribution that may serve as isolating factors over short distances, notably deserts and grass-lands. However, the species is composed of only two subspecies (based on subtle differences in bill size and head pattern), distributed west and east of the Sierra Nevada and Cascades ranges respectively, and

shows no other apparent plumage or vocal variation (Grinnell and Miller, 1944) indicative of population isolation. The pattern of migratory northern populations separated along an east–west axis and sedentary populations in Mexico, is also found in other species complexes, such as in the juncos (*Junco hyemalis/ phaeonotus*), Yellow-rumped Warbler (*Setophaga coronata*) and yellowthroats (*Geothlypis trichas/beldingi/flavovelata/nelsoni/rostrata*), with all of these taxa having morphologically distinct resident lineages in Mexico, as well as multiple migratory morphological groups in the United States and Canada (Escalante et al., 2009; Milà et al., 2007a,b). The widespread distribution without apparent breaks to gene flow, life history peculiarities, and the dichotomy between migratory and resident populations yields interesting questions regarding population structure differences between north and south. Phylogeographic questions that may be formulated in this context often require the use of multiple markers.

To explore these differences further, we developed a range-wide phylogeography of the Black-headed Grosbeak based on mitochondrial and nuclear DNA. The use of multilocus data may allow circumvention of a number of problems associated with studies based solely on mtDNA (Edwards et al., 2005; Edwards and Bensch, 2009), including inaccurate parameter estimation and several issues surrounding the description of evolutionary process rather than pattern. Especially the latter is problematic; phylogeographic studies often have a strong population focus and to adequately explain any interpopulation processes such as gene flow, the estimation of effective population sizes, and population size changes over time requires the use of multiple markers. A phylogenetic study by Pulgarín et al. (2013) showed that nuclear genes in many *Pheucticus* species did not mirror mitochondrial variation.

Using multiple markers will allow us to investigate possible patterns of differentiation between mitochondrial and nuclear genes within the Black-headed Grosbeak. Population genetic statistics and gene flow analyses will shed light on the mechanisms by which this differentiation may have arisen. Phylogeographic studies often focus on the early stages of diversification and to more accurately estimate the parameters associated with the population-level processes operating at this scale, multilocus data is required. We apply an array of phylogeographic analyses to examine the evolutionary history of the Black-headed Grosbeak.

Using both a mitochondrial and nuclear component will allow us to; (1) characterize range-wide phylogeographic variation in the Black-headed Grosbeak; (2) estimate and compare patterns of genetic variation; (3) compare phylogeographic patterns between migratory and non-migratory populations (and with similarly distributed taxa); (4) investigate what ecological or behavioral traits may have led to the genetic patterns found, and (5) assess how potential gene flow can be best characterized in terms of directionality, quantity, and any bias towards clade or sex.

Table 1. Characteristics of 13 loci used in this study, first is mitochondrial locus, following 5 are introns, last 7 are anonymous loci.

Locus	# bp	Location	Primers	Primer sequence (F/R)	Annealing temp. (LC)	# Inf. sites
ND2	1041	mitochondrion	l5215/hTRPC	TATCGGGCCCATACCCCGAAAAT CGGACTTTAGCAGAAACTAAGAG	54	86
00227	139	chromosome 22	Backström et al. (2008)	ACGAATCTCTCTGCTGGGAT GGGATTTTTGGTGGGCTCTT	60	11
02108	231	chromosome 15	idem	TGTAGTCTCACTGGTTAC GCTGAGGAAACAATTCTG	60	7
17969	212	chromosome 8	idem	CTGCTCACCTACTGCGATTC CGTGGGGCCAATCATCTTCA	60	14
19599	273	chromosome 13	idem	AATCCCTGTGATGACAAACG GAACTGTACCTCCTCAAAGG	60	12
20771	280	chromosome 2	idem	TGGAATCTTGTAGAAGGACG CATATTTCTCTCCATCAGGG	60	14
Pm23	268	chromosome 2	PmeLocus 23	TTGGACACCCCTATCTTTGCA TTACAAATTGGCCCGAAGGT	61	8
Pm24	281	chromosome 2	PmeLocus 24	GGCTTCCAGTGCTCTCCAAA TGCTTCTTTTTCATGTCCTAAATGC	61	12
Pm35	405	chromosome 2	PmeLocus 35	TTGGCTGCAGGGACACATT CAGTGAGACTTGGGCTGCCT	60	16
Pm36	341	chromosome 11	PmeLocus 36	ACTGCTGCCTCACAAGCTCA TTCTTTAAGAGTCCCGACCACAG	60	15
Pm40	116	chromosome 3	PmeLocus 40	GAGCAGGTCTTCCTCAGGCA TGCAGGGTTGGAAGTTGGAC	61	12
Pm52	228	chromosome 6	PmeLocus 52	CTTGAGACCTTCTGCCCCAA AGTCCAAGCACTAACCCAGCA	61	8
Pm55	155	chromosome 15	PmeLocus 55	TGTGGTGGGTTCTCTGTAGCC TTCTGCATATGGCAGTGCTGA	60	21

2. Methods

2.1. Sample data and collection

We obtained 424 tissue samples from 68 localities (Table 1) across the range of the Black-headed Grosbeak during the breeding season (carefully avoiding migrant or wintering birds by sampling within a limited time interval), including representatives from both the nominate subspecies as well as *P. m. maculatus* found along the Pacific coast. We used the sister taxon Rose-breasted Grosbeak (*P. ludovicianus*) as out-group in all analyses (Klicka et al., 2007, Pulgarín et al., 2013). We extracted total genomic DNA using a Qiagen DNeasy tissue extraction kit (QIAGEN, Valencia, California) following manufacturer's protocol. We amplified one mitochondrial gene (NADH dehydrogenase subunit 2 – ND2), five anonymous nuclear loci, and five autosomal introns (Table 1) via polymerase chain reaction (PCR) in 12.5 µl reactions using the following protocol: denaturation at 94 LC for 10 min, 40 cycles of 94 LC for 30 s, 54 LC for 45 s, and 72 LC for 2 min, followed by 10 min elongation at 72 LC and 4 LC soak. We used the primers L5215 (Hackett 1996) and HTrpC (Smithsonian Tropical Research Institution, Balboa, Panamá) for the ND2 gene and we designed primers for ten anonymous loci with annealing temperatures varying from 54 LC to 61 LC. For the introns we used the primer sequences from Backström et

al. (2008). We sequenced ND2 for 424 individuals and the nuclear DNA for 42 individuals, including samples from Guanajuato (n = 6, Interior mtDNA clade), Guerrero (n = 6, Mexican mtDNA clade), Michoacán (n = 6 Interior mtDNA clade), Oregon (n = 8, Pacific mtDNA clade), central California (n = 3, Pacific mtDNA clade), New Mexico (n = 6, Interior mtDNA clade), and South Dakota (n = 6, Interior mtDNA clade), as well as for the out-group taxon (n = 5). We used the program Sequencher (Gene Codes Corporation, Ann Arbor, Michigan) to align complementary DNA strands, detect stop codons, and translate genetic information into amino acids. To detect and interpret insertions and deletions in the nucDNA, we used the program Indelligent (Dmitriev and Rakitov, 2008). We phased sequences in DnaSP using the algorithm provided by PHASE (Stephens and Donnelly, 2003). For sites that had posterior probabilities of <0.70 we specified the nucleotide as ambiguous. We deposited sequences in Genbank under accession numbers KJ714059-KJ715160.

2.2. Phylogenetic tree reconstruction

We identified the best-fit nucleotide substitution model for each locus using jModeltest 2 (Darriba et al., 2012; Guindon and Gascuel, 2003). The HKY model was the best-fit for all loci but one locus, Pm55. We estimated gene trees for all genes using MrBayes 3.1.2 (Huelsenbeck and Ronquist, 2001) for 10 million generations, sampling every 1000 generations, using four chains (4 4 nucleotide model, gamma-distributed rate). We analyzed posterior out-put in TRACER v. 1.5, and determined the burn-into be 10%, using ESS values > 200 to assess convergence. We attempted to recover a species tree in BEAST, a component of BEAST v. 1.7.3 (Drummond and Rambaut, 2007) including the 13 loci used in this study for 41 individuals. BEAST did not recover a species tree (no MCMC chain convergence), achieving ESS values < 200 for a large number of parameter values, even after specification of different prior

Table 2. Black-headed Grosbeak sampling localities (this study n = 68) with mtDNA clade composition (I = Interior, P = Pacific, M = Mexican), population assignment, sample size, country, state, county/region, locality, and coordinates. The Mexican and Interior clades correspond to P. m. melanocephalus, and the Pacific clade to P. m. maculatus. Numeric designations refer to subpopulations in Fig. 1 and Table 3.

Clade	Subpopulation	Sample	Country	State	County/region	Locality	Lat/long
P1.1	Sierra Nevada	12	USA	California	Shasta	Shasta Natl. For.	40L56.72^{0}N, 122L31.06^{0}W
P1.1	Sierra Nevada	3	USA	California	Modoc	Cedarville	41L28.91^{0}N, 120L13.48^{0}W
P1.1	Sierra Nevada	1	USA	California	Modoc	Alturas	41L28.91^{0}N, 41L28.29^{0}W
P1.1	Sierra Nevada	3	USA	California	Eldorado	Lake Tahoe	115L38.35^{0}N, 120L8.51^{0}W
P1.1	Sierra Nevada	10	USA	California	Fresno	Sanger	36L57.50^{0}N, 119L10.90^{0}W
P1.1	Sierra Nevada	4	USA	California	Inyo	Bishop	37L26.17^{0}N, 118L06.25^{0}W
P1.1	Sierra Nevada	2	USA	California	Lassen	Milford	40L07.00^{0}N, 120L21.50^{0}W
P1.2	NW California	11	USA	California	Monterey	Los Padres Natl. For.	36L19.90^{0}N, 121L34.60^{0}W
P1.2	NW California	9	USA	California	Mendocino	Covelo	39L48.90^{0}N, 122L59.60^{0}W
P1.2	SW California	3	MEX	Baja California	Ensenada	Sierra de Juárez	31L52.33^{0}N, 115L38.49^{0}W
P1.2	SW California	10	USA	California	San Diego	Julian	32L51.08^{0}N, 116L25.20^{0}W
P1.2	SW California	4	USA	California	San Bernardino	Big Bear	34L13.29^{0}N, 116L43.85^{0}W
P1.3	Coastal Northwest	6	USA	Oregon	Coos	Coos Bay	43L14.70^{0}N, 124L07.20^{0}W

P1.3	Coastal Northwest	12	USA	Oregon	Josephine	Grant's Pass	42L18.2^{0}N, 123L46.9^{0}W
P1.3	Coastal Northwest	1	USA	Washington	Pacific	Frances	46L31.59^{0}N, 123L25.01^{0}W
P1.3	Coastal Northwest	1	USA	Oregon	Lincoln	Newport	44L28.50^{0}N, 123L55.00^{0}W
P1.3	Coastal Northwest	2	USA	Washington	Clallum	Sequim	47L55.43^{0}N, 123L1.30^{0}W
P1.4	Cascades	6	USA	Washington	King	Renton	47L28.59^{0}N, 122L13.12^{0}W
P1.4	Cascades	4	USA	Oregon	Douglas	Reedsport	43L46.80^{0}N, 124L00.90^{0}W
P1.4	Cascades	1	USA	Oregon	Jackson	Kadoka	42L20.60^{0}N, 122L40.23^{0}W
P1.4	Cascades	2	USA	Washington	Okanogan	Winthrop	48L28.31^{0}N, 120L5.48^{0}W
P1.4	Cascades	1	USA	Washington	Kittitas	Ellensburg	42L20.43^{0}N, 122L64.53^{0}W
I2.1	S. Madre Occidental	3	MEX	Chihuahua	Madera	Ciudad Madera	29L39.22^{0}N, 108L10.24^{0}W
I2.1	S. Madre Occidental	11	MEX	Durango	Durango	Salvador Allende	24L5.13^{0}N, 104L55.58^{0}W
I2.1	S. Madre Occidental	10	USA	Arizona	Cochise	Chiricahua Mts.	31L50.99^{0}N, 109L19.59^{0}W
I2.2	Eastern Sierra Nevada	9	USA	California	Inyo	Bishop	37L26.17^{0}N, 118L06.25^{0}W
I2.2	Eastern Sierra Nevada	5	USA	California	Modoc	Eagleville	41L14.23^{0}N, 41L14.41^{0}W
I2.2	Eastern Sierra Nevada	1	USA	California	Mono	Indian Creek Canyon	37L44.49^{0}N, 118L15.50^{0}W
I2.2	Eastern Sierra Nevada	2	USA	California	Lassen	Milford	40L07.00^{0}N, 120L21.50^{0}W
I2.2	Eastern Sierra Nevada	1	USA	Nevada	Washoe	Verdi	39L28.00^{0}N, 119L57.00^{0}W
I2.3	S. Madre Oriental	12	MEX	Nuevo León	General Zaragoza	Peña Nevada	23L48.23^{0}N, 99L50.49^{0}W
I2.4	E Northwest	8	USA	Washington	Okanogan	Winthrop	48L28.31^{0}N, 120L5.48^{0}W
I2.4	E Northwest	9	USA	Washington	Columbia	Dayton	46L14.27^{0}N, 46L14.36^{0}W
I2.4	E Northwest	7	USA	Oregon	Grant	Dayville	44L8.30^{0}N, 119L29.53^{0}W
I2.4	E Northwest	2	USA	Oregon	Crook	Dayville	44L8.30^{0}N, 119L29.53^{0}W
I2.4	E Northwest	2	USA	Washington	King	Renton	47L28.588^{0}N, 122L13.12^{0}W
I2.4	E Northwest	3	USA	Oregon	Wallowa	Enterprise	45L13.00^{0}N, 116L51.70^{0}W
I2.4	E Northwest	3	USA	Oregon	Josephine	Grant's Pass	42L18.20^{0}N, 123L46.90^{0}W
I2.4	E Northwest	1	USA	Washington	Kittitas	Ellensburg	47L1.60^{0}N, 47L1.12^{0}W
I2.5	Black Hills	28	USA	South Dakota	Lawrence	Spearfish	44L28.30^{0}N, 103L52.12^{0}W
I2.5	Black Hills	8	USA	South Dakota	Lyman	Oacoma	43L47.46^{0}N, 99L23.46^{0}W
I2.5	Black Hills	8	USA	South Dakota	Lyman	Kennebec	43L54.15^{0}N, 99L51.43^{0}W
I2.5	Black Hills	6	USA	South Dakota	Lyman	Presho	43L54.32^{0}N, 100L3.32^{0}W
I2.6	Great Basin	2	USA	Nevada	Esmeralda	Dyer	37L46.35^{0}N, 118L13.20^{0}W
I2.6	Great Basin	3	USA	Nevada	Elko	Jarbridge	41L54.00^{0}N, 115L2.05^{0}W
I2.6	Great Basin	3	USA	Nevada	Lincoln	Pioche	37L55.00^{0}N, 114L33.00^{0}W
I2.6	Great Basin	20	USA	Nevada	Clark	Spring Mts.	36L18.13^{0}N, 36L18.45^{0}W
I2.6	Great Basin	2	USA	Nevada	Clark	Virgin Mts.	36L36.30^{0}N, 114L5.29^{0}W
I2.6	Great Basin	5	USA	California	San Bernardino	Big Bear	34L13.29^{0}N, 116L43.85^{0}W
I2.6	Great Basin	1	USA	Arizona	Mohave	Hualapai Mts.	35L03.00^{0}N, 113L53.00^{0}W
I2.6	Great Basin	1	USA	Arizona	Coconino	Flagstaff	35L19.00^{0}N, 111L45.00^{0}W
I2.7	S. Rocky Mountains	9	USA	Colorado	Fremont	Canon City	38L56.15^{0}N, 105L44.55^{0}W
I2.7	S. Rocky Mountains	10	USA	New Mexico	Eddy	Guadalupe Mts.	32L02.02^{0}N, 104L47.92^{0}W
I2.7	S. Rocky Mountains	10	USA	Colorado	La Plata	Durango	37L25.77^{0}N, 107L45.04^{0}W
I2.7	S. Rocky Mountains	1	USA	New Mexico	Lincoln	Carrizozo	33L30.20^{0}N, 105L46.80^{0}W
I2.7	S. Rocky Mountains	1	USA	New Mexico	Taos	Taos	36L13.30^{0}N, 105L35.33^{0}W
I2.7	S. Rocky Mountains	1	USA	Colorado	Pueblo	San Isabel Natl. For.	38L07.440^{0}N, 105L07.553^{0}W
I2.8	Western Transvolcanic	11	MEX	Michoacán	Nuevo San Juan	San Juan	19L25.86^{0}N, 102L15.87^{0}W
I2.8	Western Transvolcanic	17	MEX	Jalisco	Bolaños	Sierra de Bolaños	21L52.90^{0}N, 103L51.90^{0}W
I2.9	Eastern Transvolcanic	12	MEX	Guanajuato	Moroleón	Moroleón	20L04.31^{0}N, 101L19.45^{0}W
I2.9	Eastern Transvolcanic	10	MEX	Morelos	Huitzilac	Huitzilac	19L05.21^{0}N, 99L11.63^{0}W
I2.9	Eastern Transvolcanic	4	MEX	Querétaro	Pinal de Amoles	Peñamiller	21L3.72^{0}N, 99L48.54^{0}W
I2.9	Eastern Transvolcanic	1	MEX	Guerrero	Chilpancingo	Omiltemi	17L32.73^{0}N, 99L38.02^{0}W
I2.9	Eastern Transvolcanic	2	MEX	Distrito Federal	Ajusco	Ajusco	19L12.25^{0}N, 99L15.29^{0}W
I2.10	C Rocky Mountains	2	USA	Idaho	Caribou	Soda Springs	42L25.57^{0}N, 111L33.53^{0}W
I2.10	C Rocky Mountains	1	USA	Idaho	Bear Lake	Montpelier	42L24.43^{0}N, 42L24.58^{0}W

I2.10	C Rocky Mountains	7	USA	Idaho	Franklin	Wasatch Mts.	42L21.25^{0}N, 111L41.37^{0}N
I2.10	C Rocky Mountains	4	USA	Idaho	Flathead	Choteau	47L54.02^{0}N, 113L31.52^{0}W
I2.10	C. Rocky Mountains	3	USA	Wyoming	Sheridan	Dayton	44L52.30^{0}N, 107L15.43.^{0}W
I2.10	C Rocky Mountains	5	USA	Montana	Lake	Swan Lake	47L42.82^{0}N, 113L52.12^{0}W
I2.x	–	1	USA	California	Mendocino	Covelo	39L48.9^{0}N, 122L59.6^{0}W
M3	–	9	MEX	Guerrero	Chilpancingo	Omiltemi	17L32.73^{0}N, 99L38.02^{0}W
M3	–	2	MEX	Guerrero	Chilpancingo	Carrizal de Bravo	17L37.010N, 99L50.460W
M3	–	2	MEX	Morelos	Huitzilac	Huitzilac	19L05.210N, 99L11.630W
M3	–	6	MEX	Oaxaca	Ixtlán de Juárez	Ixtlán de Juárez	17L33.180N, 99L41.930W

characteristics and run lengths. We then produced a concatenated tree using BEAST v. 1.7.3 from the full nucDNA set and corresponding ND2 data. We unlinked each locus and we specified the corresponding nucleotide substitution model. We used 'coalescent: constant size' for the tree prior and we ran the analysis for 25 million generations, sampling every 1000. We analyzed posterior output in TRACER v. 1.5 and the determined burn-into be 10%. ND2 data were determined to be clocklike in MEGA5.0 (AIC = 2692.016). We used a substitution rate prior of 1.25 10^{-8} - substitutions/site/year (2.5% divergence per million years, Smith and Klicka, 2010) for ND2 and 1.35 10^{-9} substitutions/site/year for nuclear loci, applying a strict clock and lognormal prior distributions in all cases (standardized intron rate, Ellegren and Parsch, 2007). To verify the topology estimated in BEAST, we also constructed a ML tree in Garli 2.0 (Zwickl, 2006) using 1000 bootstrap replicates and the same nucleotide substitution model settings as used for the Beast analysis.

2.3. Population-level analyses

To further examine phylogeographic structure in the mtDNA, we constructed median-joining haplotype networks using the pro-gram NETWORK v. 4.5.1.6 (Bandelt et al., 1999) for each of the mtDNA clades identified from our phylogenetic analysis. However, the nucDNA had low phylogenetic information content (7–21 informative sites compared to 86 in ND2) and haplotypes in net-works constructed with nucDNA were unsorted. We calculated Fst values, Fu's Fs, Tajima's D, and haplotype and nucleotide diversity (DNASp v. 5.10.01, Rozas, 2009) based on 16 mtDNA populations (pooled from geographically proximate locations that are not separated by barriers to gene flow). We calculated nucleotide and haplotype diversity, and Tajima's D for seven populations (n = 42) based on all loci (Supplement 1). Finally, we used the online tool Isolation by Distance v. 3.23 (Jensen et al., 2005) to calculate isolation by distance effects across loci, analyzing populations based on mitochondrial clades separately. IBD allows genetic data to be regressed against distances between populations while simultaneously producing a number of population differentiation statistics, including Fst.

To identify the number of populations based on nuclear loci without a priori input of population assignments, we used the pro-gram Structure 2.3.4 (Pritchard et al., 2000; Falush et al., 2003; Hubisz et al., 2009). We used single nucleotide polymorphism (SNP) data from nuclear loci,

including an admixture model with correlated allele frequencies between populations. Lambda was inferred by estimating the log likelihood of k = 1 and allowing lambda to converge. Subsequent runs kept lambda constant using the obtained value from the initial run. We tested values of k = 1– 10, running 10 iterations for each k. Each run consisted of a burn-in period of 150,000 steps followed by 1 million iterations. To determine the appropriate value of k for our data, we used the highest log likelihood of the data (before reaching an asymptote). Additionally, we used DK (rate of change of the likelihood of k) as a metric to estimate k, calculated using Structure Harvester Web v. 0.6.93 (Evanno et al., 2005). We used Distruct 1.1 (Rosenberg, 2004) to display Structure output.

To estimate the effect of gene flow between populations on the lack of structure in nuclear genes, we used the Isolation-with-Migration (IMa) software (Hey and Nielsen, 2007). We analyzed pair-wise comparisons of populations based on geographically proximate mtDNA clades. We independently analyzed the mtDNA and nucDNA to assess the significance of sex-biased dispersal. We applied a mutation rate prior of $1.25\ 10^{-8}$ substitutions/site/year (Smith and Klicka, 2010) to the ND2 gene, and of $1.35\ 10^{-9}$ substitutions/site/year to introns and anonymous loci (Ellegren and Parsch, 2007). To convert these rates to substitution rates per locus, we multiplied rates by the length of each locus expressed in number of nucleotides. We ran several (~15) trials to identify appropriately calibrated model parameter priors, after which we ran the program for a burn-in period of 500,000 steps followed by 1– 2.5 million iterations (>150 effective sample size for each parameter). Using results from IMa, we calculated the effective migration rate between populations (2 Nm). We were unable to calculate migration rates between each pair of subpopulations (Table 2) used in our study due to widespread haplotype sharing, which violates the minimum amount of divergence required by IMa to produce interpretable parameter distribution values.

Finally, we performed Extended Bayesian Skyline Plot (Heled and Drummond, 2008) analysis using multilocus data in BEAST v. 1.6.2 to estimate population size changes over time. We used the same parameter settings as in our tree reconstruction, except we used a chain length of 10,000,000 generations and 15% burn-in, and we assumed a lognormal shape for most priors, except the 'frequencies' and 'kappa' parameters. Because there was no significant genetic structure within the nucDNA, we combined all the populations to achieve a higher sample size and increase the likelihood of detecting demographic changes. We also performed independent Skyline analyses on our two major mitochondrial clades, using the same parameter values, but with an increased chain length of 20,000,000 generations.

Table 3. Mitochondrial genetic diversity indices and mismatch statistics for clades and 15 populations within *Pheucticus melanocephalus*. Shown are: sample size (N), number of haplotypes (#H), number of private haplotypes (H_{priv}), haplotype diversity (H_{div}), nucleotide diversity (p), Fu's F_s (F_s), Tajima's D, mismatch sum of squared deviation (SSD), Harpending's raggedness index (r). Significance is indicated by asterisks = <0.02, = <0.01 for Fu's F_s; = <0.05, = <0.01 for SSD, and Harpending's raggedness index). Numeric designations refer to subpopulations identified in Fig. 1 and Table 2.

Sample	N	#H	H_{priv}	H_{div}	p	F_s	D	SSD	r
PACIFIC									
1.1 Sierra Nevada	3510	7		0.519	0.0010	0.4057	2.0637	0.0102	0.0594
1.2 NW California	329	3		0.753	0.0010	5.1238	1.9072	0.0115	0.1237
1.3 SW California	1712	9		0.941	0.0026	3.2930	2.2232	0.0074	0.0363
1.4 Coastal Northwest	2213	3		0.913	0.0016	1.2238	2.4667	0.0144	0.1131
1.5 W Cascades	146	1		0.727	0.0009	0.4099	1.1285	0.0176	0.4380
INTERIOR									
2.1 S.M. Occidental	2415	9		0.924	0.0023	12.5593	1.7361	0.0125	0.2900
2.2 Sierra Nevada E	1711	4		0.882	0.0018	4.4542	1.5856	0.0050	0.0257
2.3 S.M. Oriental	125	1		0.667	0.0010	1.2388	0.8419	0.0088	0.8700
2.4 E Northwest	3617	12		0.930	0.0040	9.3373	2.2284	0.0023	0.0418
2.5 Black Hills	5021	10		0.947	0.0026	11.6279	0.7067	0.0159	0.0694
2.6 Great Basin	3720	7		0.914	0.0021	13.3633	1.8871	0.0248	0.1139
2.7 S Rocky Mts	3316	10		0.926	0.0021	10.6734	1.8591	0.0063	0.0622
2.8 W Transvolcanic	2814	8		0.907	0.0028	21.4526	2.2287	0.0070	0.0783
2.9 E Transvolcanic	2816	9		0.934	0.0023	17.3289	2.3821	0.0079	0.0679
2.10 C Rocky Mts	2211	2		0.913	0.0020	4.3239	0.9831	0.0214	0.1053
MEXICAN	1712			0.956	0.0040	7.1173	1.0694	0.0014	0.0188

3. Results

3.1. Population genetic summary statistics

In the mtDNA dataset there were 86 polymorphic sites (Table 1). The populations with highest genetic diversity (p, haplotype diversity, Table 2) were in southern California (Pacific mtDNA clade) and the southern part of the Interior population, e.g. Mexican Plateau and Transvolcanic Belt. The Sierra Madre de Sur population also showed high genetic diversity indices. Significant Fu's Fs and Tajima's D furthermore showed that Interior populations likely expanded recently. The Sierra Madre Oriental population was consistently different from other populations in the Interior clade by various statistics; it showed lower haplotype and nucleotide diversity, did not show signatures of expansion, although lower sample size of this population relative to others may partially account for this pattern.

Table 4. IMa statistics (highest point estimates) and 95% HPD for two pair-wise population comparisons, population designations based on mtDNA clades. Values indicated with asterisk based on run that did not converge, and resulted in a bimodal distribution.

Loci	Populations	Parameter	q1	q2	qA	m1	m2	t
All	Pacific-Interior	high.point	3560.89	5111.03	3950.64	0.01	2.42	447.32
		95%HPDlow	2745.96	3180.00	3312.87	0.00	0.33	306.06
		95%HPDhigh	29496.92	14465.01	34882.55	0.17	2.85	41835.94
All	Interior-Mexican	high.point	5918.13	2024.75	171953.06	1.43	0.00	15685.53
		95%HPDlow	3819.50	1433.78	25485.72	0.93	0.00	10094.07
		95%HPDhigh	13036.67	2723.18	201367.41	2.13	0.15	28104.47
ND2	Pacific-Interior	high.point	123.69	93.47	1887.43	0.00	0.25	4.22
		95%HPDlow	95.36	67.04	1677.82	0.00	0.14	3.01
		95%HPDhigh	165.23	136.91	1887.43	0.00	0.46	8.80
All	Interior-Mexican	high.point	98.78	50.23	1556.83	0.22	0.00	2.89
		95%HPDlow	78.66	34.67	1380.12	0.13	0.00	2.55
		95%HPDhigh	122.13	78.23	1775.36	0.27	0.00	5.12

Table 5. Structure output data. #K is inferred number of populations.

# K	Reps	Mean LnP(K)	Stdev LnP(K)	$Ln^{0}(K)$	$\|Ln^{00}(K)\|$	Delta K
1	5	1603.8800	0.8228	NA	NA	NA
2	5	1209.2600	0.5367	394.620000	284.960000	530.991609
3	5	1099.6000	3.1401	109.660000	12.600000	4.012657
4	5	1002.5400	1.7126	97.060000	27.040000	15.788856
5	5	932.5200	15.2529	70.020000	7.500000	0.491709
6	5	855.0000	27.2720	77.520000	13.540000	0.496481
7	5	791.0200	60.3916	63.980000	NA	NA

3.2. Phylogenetic tree reconstruction

Bayesian and ML concatenated phylogenetic tree reconstruction resulted in three clades within the Black-headed Grosbeak, congruent with the findings from previous analyses that had sparser sampling (Pulgarín et al., 2013). The most widespread clade ('Interior') occurs in interior North America, from the Transvolcanic Belt in Mexico, to the Black Hills in South Dakota, and the eastern Sierra Nevada in the west (Fig. 1a). This clade overlaps extensively with the distribution of the subspecies *P. melanocephalus melanocephalus*, except for the Sierra Madre del Sur of Oaxaca and Guerrero, where a second clade occurs ('Mexican'). Our third clade ('Pacific') is found in California west of the Sierra Nevada and north to south-ern British Columbia, and corresponds with the known geographic limits of the subspecies *P. melanocephalus maculatus*. No

geographically defined structuring was found in either nucDNA gene trees or individual mtDNA haplotype networks. The three clades present in the tree from the concatenated dataset were also recovered in the mtDNA tree, indicating that the topology was driven by the mtDNA (Fig. 1b). Posterior values on all major nodes were 1.0, whereas maximum likelihood bootstrap values ranged from 97% to 100%. Divergence time estimates based on mtDNA were 2.5 Mya (95% highest posterior density (HPD): 1.87–3.23 Mya) for the split between P. m. melanocephalus and out-group *P. ludovicianus* (agreeing with previous estimates of 2.2 Mya based on cytochrome b, Klicka and Zink, 1997), just over 1 Mya for the divergence between the Pacific and Interior clades (95% HPD: 0.87–1.70 Mya), around 0.6 Mya (95% HPA: 0.37–0.86 Mya) between the Mexican and Interior clades, and around 0.3 Mya for a non-geographically defined split within the Interior clade (95% HPD: 0.14–0.36). Divergence times based on the concatenated tree were similar.

There was geographic overlap between the Pacific and Interior (mitochondrial) clades in at least 7 localities along the eastern slope of the Sierra Nevada, the Oregon Cascades, and the central Rocky Mountains of southern British Columbia. Additionally, we found overlap between the Mexican and Interior clades in the Transvolcanic Belt of Morelos and in Guerrero. Haplotype networks of mtDNA for the Pacific and Interior clades revealed fair amounts of diversity, mostly not structured geographically (Fig. 2). The Interior and Pacific clades are composed of a few common and closely related haplotypes, with both clades showing a pattern of less frequent, but mostly closely related, haplotypes emanating from the dominant haplotypes. The Mexican clade seems to show more structure, although it is difficult to discern any geographic patterns due to low sample size. Extended Bayesian Skyline Plot analysis yielded a signature of constant population size over time, with population expansion and wider HPD towards the present (Fig. 3).

3.3. Population genetic structure

We found no consistent patterns of high or low diversity by geographic location across nuclear loci. Locus Pm55 showed higher (0.9236) than average (0.6599) haplotype diversity. Loci Pm23, Pm24, and Pm35 showed significant positive values of Tajima's D across most populations, indicative of population decrease or balancing selection. Isolation by Distance plots revealed no significant isolation-by-distance effects within both the Pacific ($r = 0.03$, $Z = 1171.86$, $p = 0.18$) and Interior ($r = 0.12$, $Z = 5850.06$, $p = 0.14$) mtDNA clades. We extracted a total of 136 SNPs per individual from the total nuclear data set, with a range of 8–21 SNPs per sequence. From this data set, we inferred $k = 3$ populations from the STRUCTURE analysis using L(K) and $k = 2$ using Dk (Fig. 3, Table 5, Supplement 2). Using Dk, all populations clustered together, except for three individuals from Eddy, New Mexico. According to L(K), we found three groups, but without geographically defined structure. Most individuals had assignment scores of $Q > 0.9$, indicating strong sorting to populations.

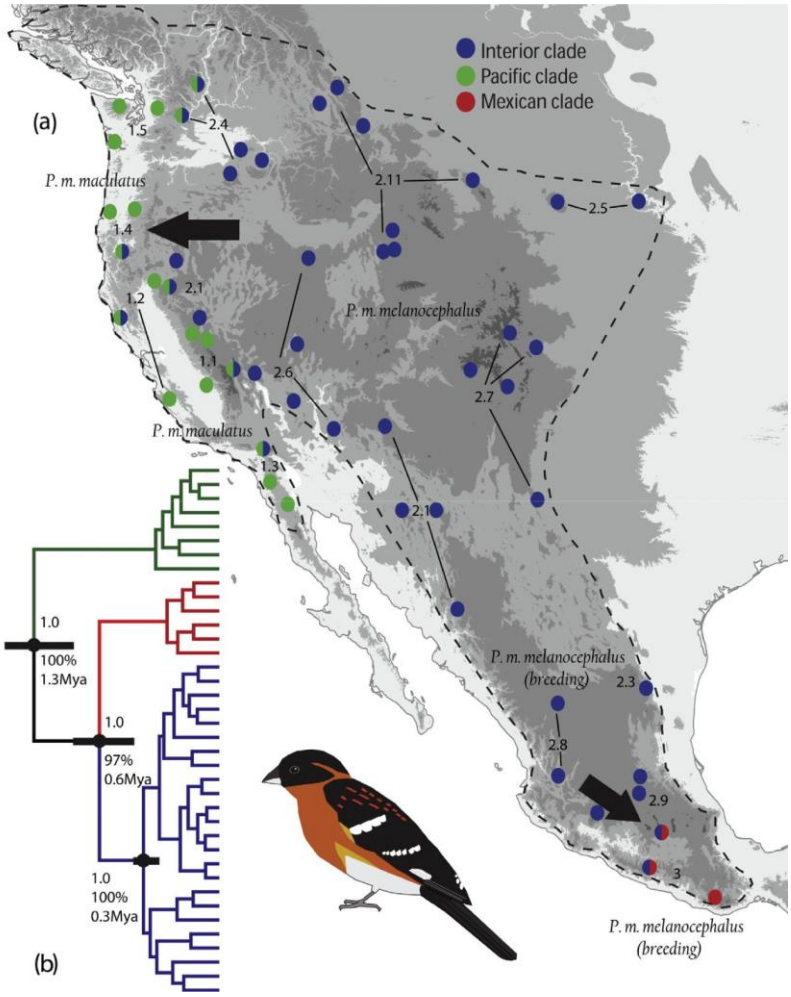


Fig. 1. (a) sampling locations used in this study. In blue samples from the Interior mtDNA clade, in green the Pacific clade, red the Mexican clade. Bicolored circles represent two sympatric clades. Arrows indicate gene flow calculated by IMa (Hey and Nielsen, 2007). (b) Phylogenetic tree based on concatenation of ND2 and 12 nuclear loci. Branch colors represent mtDNA clades as in 1a. Major nodes are indicated with black circles. Upper numbers are posterior probabilities values for closest node, middle numbers are percentage likelihood bootstrap values, lower numbers are divergence times based on mtDNA.

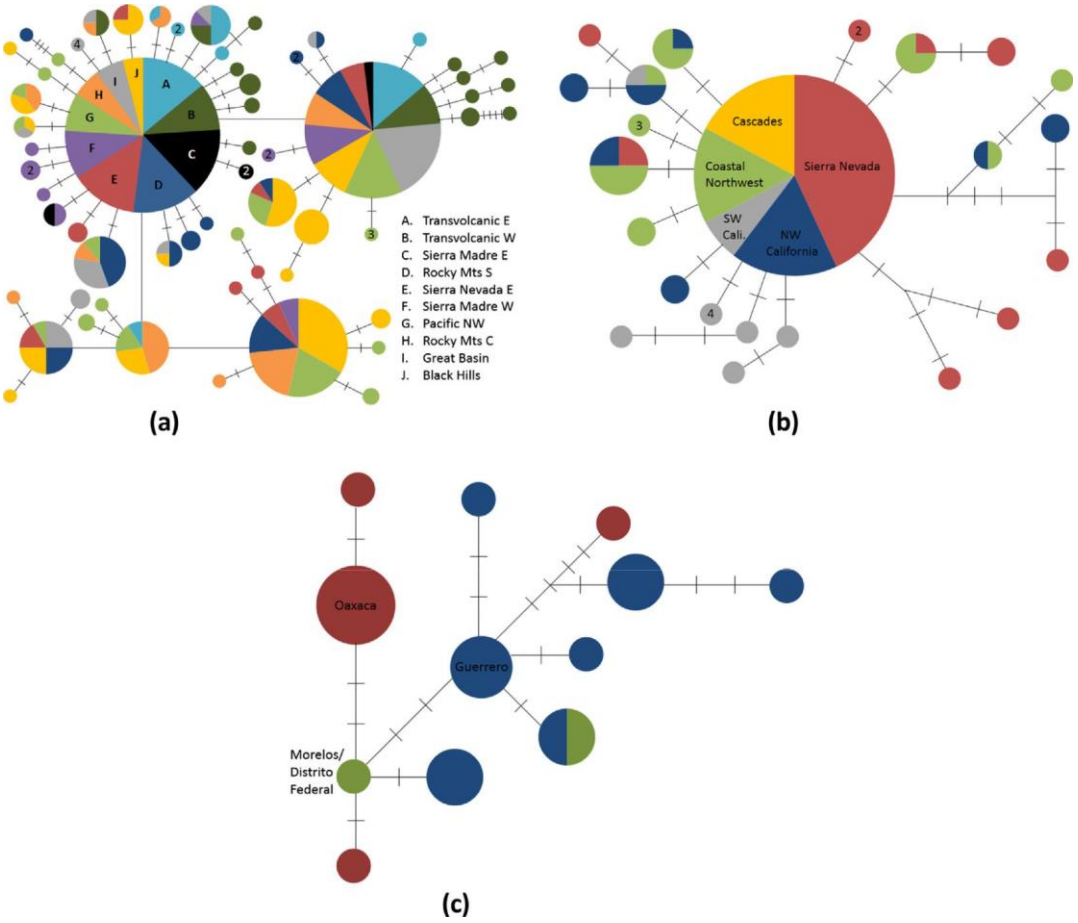
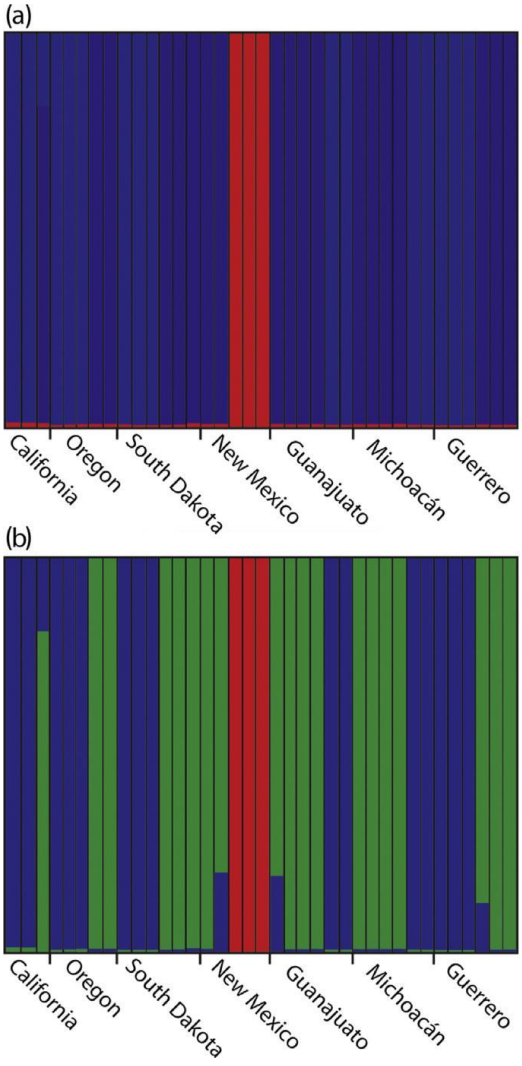


Fig. 2. Median-joining haplotype networks indicating geographic structure in mtDNA in (a) the Interior clade, (b) the Pacific clade, and (c) the Mexican clade. Tick marks between haplotype circles indicate 1 base pair change, numbers represent the amount of samples represented by circle.

In the IMa pair-wise population comparisons (Fig. 1., Table 4), the Interior clade had increased size relative to the ancestral population shared with the Pacific clade (see also Fig. 4), all others were smaller than the ancestral population. Effective population migration (2 Nm) was high from the Interior population into the Pacific population (2 Nm > 2), low from the Interior into the Mexican population (2 Nm < 2), and non-existent from the Pacific and Mexican populations into the Interior (2 Nm = 0). In the mtDNA only IMa analysis, migration was very low (2 Nm = 0.3) from the Interior into the Pacific population, and non-existent (2 Nm = 0) from the Pacific and Mexican populations into the Interior.



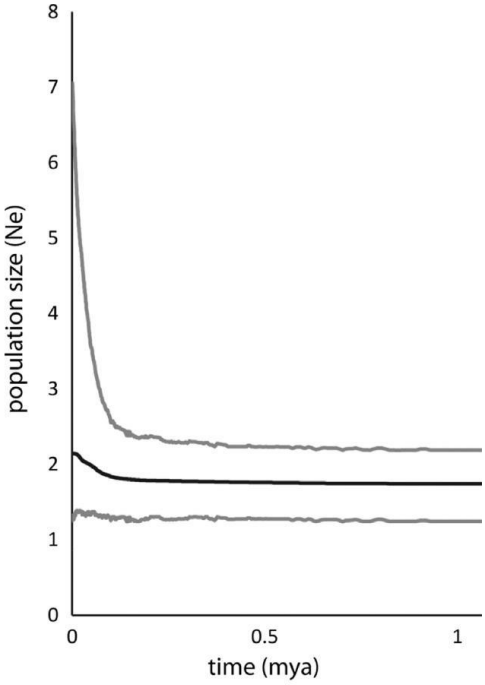


Fig. 3. STRUCTURE plot, (a) k = 2 and (b) k = 3, different k represented by red, green, and blue. Tick marks on x axis indicate separation of different populations.

Fig. 4. Demographic history of the Black-headed Grosbeak estimated using Extended Bayesian Skyline Plot. Black line represents median values of population size (Ne), gray lines are 95% lower and upper posterior density intervals. A Bayesian Skyline Plot for only the Interior mtDNA clade was also

4. Discussion

4.1. Phylogeographic structure in mtDNA

We found three distinct mtDNA lineages in the Black-headed Grosbeak. The geographic regions occupied by these lineages mirror those recovered for a number of other montane western North American taxa. The Sierra Nevada and Cascades ranges are a barrier to dispersal that has led to divergence between one or more Pacific group(s) and one or more Interior group(s) in a number

of western North American bird species, e.g. Brown Creeper (*Certhia americana*, Manthey et al., 2011), Gray Jay (*Perisoreus canadensis*, Van Els et al., 2012), White-breasted Nuthatch (*Sitta carolinensis*, Spellman and Klicka, 2007), Mountain Chickadee (*Poecile gambeli*, Spellman et al., 2007) and Western Scrub-Jays (*Aphelocoma californica*, McCormack et al., 2011). Even if coastal and interior groups both contain multiple clades, the deepest intraspecific divergence in montane species of western North America is generally between the coast and the interior (Shafer et al., 2010). This holds true not only for birds, but for a wide range of organisms, including mammals (Dusky Shrew [*Sorex monticolus*] Demboski and Cook, 2001, Gray Wolf [*Canis lupus*] Muñoz-Fuentes et al., 2009), amphibians (Tailed Frog [*Ascaphus truei*] Nielson et al., 2001, Pacific Chorus Frog [*Pseudacris regilla*] Recuero et al., 2006), plants (Douglas Fir [*Pseudotsuga menziesii*] Gugger et al., 2011, Dusky Willow [*Salix melanopsis*] Carstens and Richards, 2007), nematodes (*Soboliphyme baturini*, Koehler et al., 2009), fungi (*Armillaria ostoyae*, Hanna et al., 2007), and lichens (*Lobaria pulmonea*, Walser et al., 2004). In spite of some of these species having different demographic histories (Carstens and Richards, 2007) or being isolated by virtue of their unique behavior or ecology (Muñoz-Fuentes et al., 2009), many are characterized by having distinct coastal and inland refugia during the Last Glacial Maximum (LGM). These vicariant patterns in western North America are largely due to orogeny (Brunsfeld et al., 2001), although past climatic fluctuations must have played a substantial role, given current sympatry of different clades in multiple species (McCormack et al., 2011; Van Els et al., 2012), including Black-headed Grosbeak.

Unlike many other North American birds (*Junco* complex, Milà et al., 2007a; *Setophaga coronata* complex Milà et al., 2007b; Brown Creeper *Certhia americana*, Manthey et al., 2011), there is no genetic discontinuity between the Mexican and U.S. highlands in the Black-headed Grosbeak. For a few species for which this is also the case (Chipping Sparrow *Spizella passerina*, Milà et al., 2006, White-breasted Nuthatch *Sitta carolinensis*, Spellman and Klicka, 2007), birds moved north out of a Pleistocene refugium in Mexico after the LGM. Mexican Central Plateau populations of the species are non-migratory (and perhaps refugial relics), whereas more northern populations are strictly migratory (although movements are poorly known, and northern birds may merely replace Central Plateau populations in wintertime). Unlike in other species (e.g. Broad-tailed Hummingbird [*Selasphorus platycercus*, Malpica and Ornelas, 2014], Chipping Sparrow [Milà et al., 2006]), southern, presumably non-migratory populations of the Black-headed Grosbeak do not show differentiation in genetic composition in terms of haplotype composition or genetic diversity indices from northern migratory birds. This lack of dichotomy may indicate that northern populations did not differentiate from southern populations in spite of the development of distinct life history characteristics.

Several populations contain haplotypes of two clades, indicating potential contact zones and (mitochondrial) introgression. These populations are located mainly along the eastern slope of the Californian Sierra Nevada and San Bernardino Mountains, as well as in the Okanogan Valley, which are both known suture zones (Remington, 1968; Spellman et al., 2007; Barrowclough et al., 2004; Swenson and Howard, 2004; Van Els et al., 2012). Sympatry between clades is not

restricted to narrow bands along major natural barriers, but appears to extend to coastal areas, including Mendocino County, California, and Josephine County, Oregon. Dispersal into these locations likely occurred along the northern edge of the Sierra Nevada. We did not investigate phenotypic differences between sympatric birds from different clades and we cannot make further inferences regarding this contact zone. However, the incursion of Interior mtDNA haplotypes towards the Pacific coast may indicate that competitive displacement has taken place, much as in the Townsend's/Hermit Warbler (*Setophaga townsendi/occidentalis*) contact zone located just a little farther north (Krosby and Rohwer, 2009). This indicates that the potential for introgression and nucDNA admixture is sub-stantial over a fairly broad geographic area.

A similar situation arises in southern Mexico, where Interior clade birds occur alongside Mexican clade birds in Guerrero, as well as in the southern Transvolcanic Belt. IMa analysis indicated unidirectional gene flow into the Sierra Madre del Sur/Transvolcanic populations, possibly a result of the effects of competitive exclusion of one clade by another or perhaps due to dispersal constraints. The Sierra Madre del Sur has been recognized as a 'genetic hotspot' (Blancas-Calva et al., 2010), due to its isolation from other mountain ranges by the dry Balsas Drainage. Several lineages are restricted to the Sierra Madre del Sur in Oaxaca and Guerrero (e.g. Common Bush-Tanager [*Chlorospingus ophthalmicus*, García-Moreno, 2004], Unicolored Jay [*Aphelocoma unicolor guerrerensis/oaxacae*, McCormack et al., 2010], the deer mouse *Peromyscus aztecus* [Sullivan et al., 1997], and the rattlesnake *Crotalus triseriatus* [Bryson et al., 2011]), with the latter three examples having multiple lineages restricted to the range. In spite of the Sierra Madre del Sur representing a major hotspot partly because of its isolation, gene flow does occur between birds from the Sierra Madre del Sur and the Transvolcanic Belt for some montane birds (e.g. Red Warbler [*Ergaticus ruber*, Barrera-Guzmán et al., 2012]). The Sierra Madre de Oaxaca, a southern extension of the Sierra Madre Oriental, potentially provides a pathway for gene flow between the Sierra Madre del Sur and the Transvolcanic Belt. Haplotypes of the Sierra Madre del Sur co-occur in the eastern Sierra Madre del Sur with haplotypes from the Sierra Madre Occidental or the Transvolcanic Belt. The Black-headed Grosbeak shows a very similar pattern, with co-occurrence of clades both in the Transvolcanic Belt and the Sierra Madre del Sur of Guerrero. Given these patterns, post-Pleistocene population expansion of Interior birds probably led to dispersal into the eastern Sierra Madre del Sur and potentially the western portions of the range. Finer-scale sampling in southern Mexico is desirable to examine these patterns further.

High genetic diversity in southern populations of the Black-headed Grosbeak may indicate that these populations were more stable over time than northern populations, which may have expanded northward from Pleistocene refugia in southern California/Baja California or perhaps the west coast of the Mexican mainland (Pacific clade) and the Mexican Plateau and Transvolcanic Belt (Interior clade). The non-migratory population in the Sierra Madre de Sur showed equally high diversity. Most of the within-clade mtDNA variation is not structured geographically. Both the Pacific and Interior clades seem to show signs of recent expansion ('star-shaped'), and hold relatively more variation in some regions (Transvolcanic, S California) than others (Cascades).

4.2. mtDNA versus nucDNA phylogeographic patterns

No consistent geographically structured patterns could be uncovered from the construction of phylogenetic trees, from Structure analysis, and from haplotype networks based on nucDNA. Additionally, IMa analysis revealed that there is extensive gene flow from Interior populations into both Pacific populations and the Sierra Madre del Sur and virtually none vice versa. Focusing exclusively on IMa-deduced mtDNA gene flow, we found (nearly) non-existent levels of gene flow between all clades. However, sympatry of Interior and Pacific mtDNA clades in northern California, southern Oregon, and central Washington indicate that there is at least potential for mtDNA introgression. Isolation-and-migration analyses suffer from a number of limitations related to estimating gene flow using a single locus, including problems associated with estimation between long branches and among non-sampled ('ghost') populations (Nielsen and Wakeley, 2001; Marko and Hart, 2011). This makes it likely that mtDNA-based IMa analyses are biased. Although sample size from the contact zone is small, haplotypes from different mitochondrial groups are maintained after secondary contact, indicating that mitochondrial gene flow is probably more limited than IMa analyses revealed. NucDNA showed approximately four to tenfold lower diversity than mtDNA, and the only geographically defined structure according to Structure analysis was found in New Mexico, where three individuals were assigned to a separate group. The other two groups were found across all other populations, and given the low diversity per nuclear locus, any structure in the nuclear data set likely is defined by a small number of SNPs, possibly caused by close relatedness of these individuals. A larger nuclear DNA data set is desirable to establish geographic structure more accurately and to allow for meaningful comparisons with mtDNA.

Differences between nucDNA and mtDNA phylogeographic patterns can be explained through a number of population processes (Toews and Brelsford, 2012). In cases where diversification is relatively recent, structure in nucDNA is predicted to evolve slower than in mtDNA (Funk and Omland, 2003; McKay and Zink, 2010), which is the most frequently invoked explanation for mitonuclear discordance. Because nucDNA evolves more slowly than mtDNA, time since divergence may not have been long enough to allow the accrual of significant nucDNA variation. With incomplete lineage sorting, any geographically defined nucDNA patterns may be absent altogether, leading to incongruent patterns between mtDNA and nucDNA. Large effective population sizes and recent population expansions may lead to similar patterns of low nuclear structure coupled with well-defined mitochondrial structure, due to lower substitution rates in nuclear DNA. The lack of phylogeographic structure in the nucDNA can possibly be attributed to gene flow. Runs to estimate ancestral population sizes in IMa did not converge. However, our Extended Bayesian Skyline Plot and Bayesian Skyline analyses of the two major mtDNA clades revealed a mostly constant population size over time. Expansion of populations according to Skyline analyses was recent (within the last 0.5 My). Additionally, population genetic statistics based on mtDNA show patterns of recent population expansion (low Fu's Fs and Tajima's D). Historical population size changes thus did not likely play a major role in shaping the

homogeneity of nucDNA data. However, according to neutral theory (Kimura, 1983), nucDNA effective population size (Ne) is four times larger than the mitochondrial Ne, resulting in decreased levels of genetic drift and consequently a reduced mutation rate relative to mitochondrial DNA. MtDNA in birds is maternally inherited and haploid and thus has a fourfold smaller effective population size relative to bi-parentally inherited diploid autosomal nucDNA (Zink and Barrowclough, 2008).

Another possible explanation of low information content in nucDNA is male-mediated dispersal. Under this scenario, the number of female migration events has not been large enough to overcome the effects of drift and resulting buildup of genetic diversity in the mitochondrial genome. The observation that introgression of mtDNA haplotypes may be occurring (e.g. northern California, southern Oregon) could be a result of limited dispersal through low natal philopatry (Hill, 1988a). Young male Black-headed Grosbeaks also display delayed plumage maturation (Hill, 1988b) and frequently move around extensively, occasionally setting up territories in suboptimal habitat, to avoid competition from more successful adult-plumaged males. Delayed plumage maturation and resulting male-driven dispersal may thus be responsible for the lack of structure in nuclear genes in the Black-headed Grosbeak. Delayed plumage maturation and consequent differentiation in the quality of occupied habitats occurs at least in Rose-breasted Grosbeak (*Pheucticus ludovicianus*, Francis and Cooke, 1990), and in the more distantly related *Passerina* buntings (Rohwer, 1986; Thompson, 1991). However in birds, and especially passerines, male-mediated dispersal is infrequent compared to female-driven dispersal (Pusey, 1987). Birds in which male-mediated dispersal is common generally have unique courtship systems. Among birds with frequent male-mediated dispersal are polyandrous sandpipers (Scolopacidae), in which males take on roles usually employed by females in other avian families, and waterfowl (Anatidae), in which pairs are often formed on the wintering grounds (Pusey, 1987). Furthermore, given that our gene flow analyses based on mtDNA are likely biased, the presence of co-occurring mtDNA haplotypes from different clades, and the aforementioned inherently differentiated demographic properties between mtDNA and nucDNA, this explanation is not likely as a main driver of nuclear gene flow. Gene flow in both mtDNA and nucDNA indicates that there is no sex bias, but that a general expansion of interior birds to the periphery of the range of the species seems to be occurring. This pattern may be part of a post-Pleistocene expansion event, producing secondary contact between formerly isolated populations.

Migratory events are potentially an additional source of gene flow. The genetic study of wintering or migrating individuals requires extensive sampling during the non-breeding season and is outside the scope of this study. However, the known wintering range of Black-headed Grosbeaks breeding in northern Mexico, the United States and southern Canada overlaps extensively with sedentary populations in central and southern Mexico (Howell, 1995). Although uncommon in birds (Martínez, 1983; Petracci and Delhey, 2004), there is potential for wintering birds to linger year-round and start breeding within the winter distribution, or for birds overshooting the usual

breeding range after migration, leading in both cases to co-occurrence of different lineages and the possibility of introgression.

4.3. Taxonomic implications

Our data set revealed three distinct mtDNA clades, two of which agree broadly with current subspecific designations (agreeing with Pulgarín et al., 2013). None of these clades were supported by nuclear data, which was characterized by low diversity overall and limited geographic structure. The two main groups, the Interior and Pacific clades, correspond with the nominate subspecies and *P. m. maculatus*, respectively. It is not known if the incursion of Interior haplotypes into the Pacific clade in northern California and southern Oregon is accompanied by introgression and corresponding phenotypic shifts (Hill, 1995). Because differences between subspecies are subtle, it will be challenging to establish whether or not suture zones are represented by morphologically intermediate phenotypes. Although mitochondrial monophyly is not an indicator of morphological differentiation (let alone of subspecific status (Patten, 2010; Remsen, 2010)), further work will also be needed to verify the existence of phenotypic differentiation in birds from the Sierra Madre del Sur, which pertain to our Mexican mtDNA clade. This will require careful morphological analysis of full breeding-plumage birds from geographically disparate locations within the Sierra Madre del Sur and Transvolcanic Belt, accompanied by genetic analyses of these individuals.

Acknowledgments

We are grateful to Sharon Birks and Sievert Rohwer at the Burke Museum of Natural History of the University of Washington and Patricia Escalante-Pliego in the Zoology Department of the Universidad Autónoma Nacional de México for specimens. We would like to thank Joseph Manthey for helpful insights in gene flow analyses and Robb T. Brumfield and James V. Remsen, Jr. for useful comments on the manuscript. This work was funded in part by NSF DEB 0815057 (to J.K.) and the Barrick Museum Foundation.