Phylogeography and population history of Northern Goshawks across North America

S. Bayard de Volo†,††, R. Reynolds††, S. Sonsthagen‡, S. Talbot±, M. Antolin†

Graduate Degree Program in Ecology · Department of Biology · Colorado State University · Rocky Mountain Research Station · Forest Service U.S. Department of Agriculture · National Science Foundation

BACKGROUND

Understanding biogeographic history is important for effective species conservation and management. Such knowledge puts perspective on contemporary relationships among populations, and helps inform predictions about how changes in current climate and habitat may affect the distributions of future populations. Historically, North American forests were distributed differently from what we see today. Large glaciers covered Canada and parts of the intermountain West, and the climate was drier and colder than today causing latitudinal and elevational shifts in forest distributions.

One way to infer the phylogeographic history of a species is to determine the geographic pattern of its genetic lineages (mtDNA haplotypes). Of particular interest is differentiating the effects of demographic processes like dispersal and range expansion, from vicariant events (barriers to dispersal) that result in population divergence.

Little is known about the paleohistory and genetic relationships of North American goshawks (*Accipiter gentilis*). Major questions of interest include:

1. **Are goshawks in the Eastern US genetically different from those in the West ?**
2. **Are goshawks in the West further subdivided ?**
3. **Are goshawks in the Southwest genetically distinct from other western populations?**

Forest Refugia and Predictions about Historical Distributions

Paleo-forest history predicts that goshawks likely occupied at least four forest refugia during the last glacial period. Forest refugia were regions that were "islands of habitat" isolated from other forested regions by glaciers and non-forested habitat. Forest dependent fauna took "refuge" in these regions during the last glaciation 120-18 TYA.

Because glaciers covered northern dispersal routes, goshawks occupying Eastern forests were isolated from those in the West. Forested regions in the West were somewhat more connected, which may have facilitated more dispersal among refugia. As the glaciers melted, Eastern forests expanded northward into Canada, and Pacific Coastal and Rocky Mountain forests expanded into British Columbia. Goshawks likely experienced demographic growth as they occupied the newly formed northern habitats. Alternatively, Great Basin and Southwestern forests retreated to higher elevations, fragmenting forests across this region and goshawks there may have experienced demographic contraction.

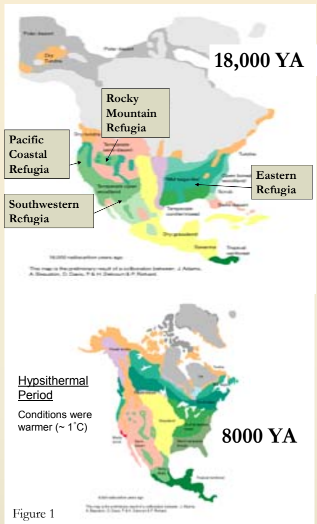

Figure 1

Current Distribution and Status

Two North American subspecies are formally recognized (American Ornithologists' Union 1983): the continental type, *A. g. atricapillus,* which occurs in most North American forests, and the Queen Charlotte goshawk, *A .g. laingi*, which is restricted to insular British Columbia and Southeast Alaska. A third putative subspecies, the Apache goshawk (*A. g. apache*) inhabits the Sky Islands of Southeast Arizona and the Sierra Madre Occidental of northwest Mexico . It is described as darker in plumage, almost black dorsally, larger in size and longer winged. While the recognition of the Apache subspecies is still debated, the US Fish and Wildlife Service (USFWS) currently considers its status unresolved (USFWS 1998).

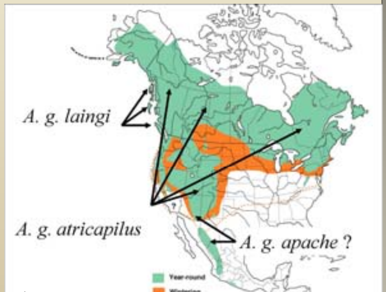

Figure 2

HYPOTHESES AND PREDICTIONS

1. If goshawks were historically isolated within glacial period refugia, the geographic structure of major mitochondrial lineages will support a pattern of population expansion from the locations of proposed forest refugia.
2. If Eastern and Western goshawks experienced long term isolation, contemporary populations in these two regions will be genetically diverged.
3. If Southwestern goshawks were historically isolated from other western populations, they should possess unique genetic lineages.

RESULTS

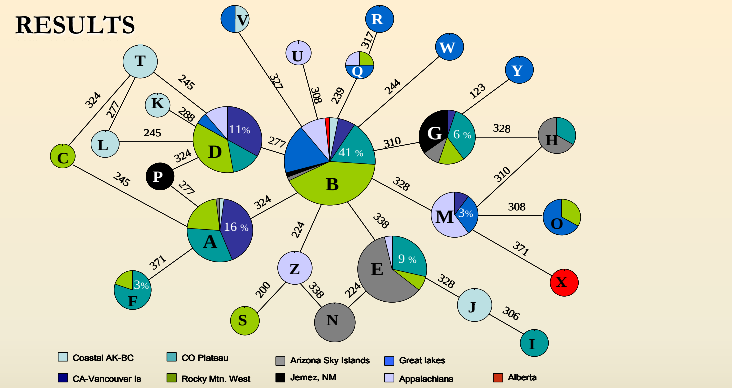

Figure 3

- **Minimum spanning parsimony tree of genetic lineages**. Circles represent haplotypes (mtDNA lineages), and numbered lines represent the sequence site where a nucleotide mutation occurred resulting in the new haplotype. Percent of individuals sampled having that haplotype is indicated.
- Because haplotype-B was found in every sample site, its assumed to be the most ancestral lineage and all other lineages have since derived. Six other haplotypes were geographically structured – haplotypes A, D, E, F, G and M. The remaining haplotypes occurred in less than 3% of individuals, mostly of whom were in Eastern sites.
- The star-like pattern of the tree is indicative of recent demographic expansion and shallow phylogenetic divergence.

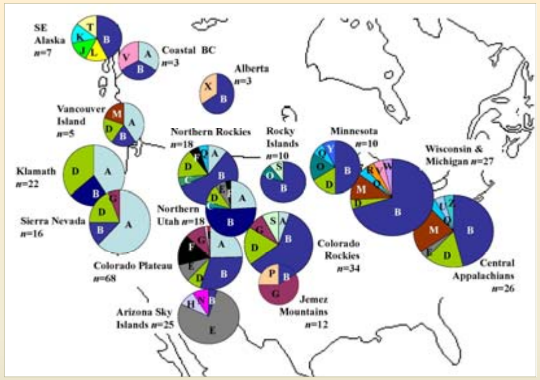

Figure 4

The geographic distribution and relative frequencies of mtDNA control region haplotypes for North American Goshawks. Site names and sample sizes are indicated.

Patterns to note include:

1. Relative lack of diversity in California, and dominance of haplotypes A & D.
2. Relative lack of diversity in Arizona Sky Islands and dominance of haplotype-E.
3. High haplotype diversity in Colorado Plateau.
4. Unique haplotypes in Southeast Alaska.
5. Dominance of haplotype-B across Rocky Mountains, Great Lakes and Central Appalachia.

Table 1. Estimates of genetic distance (Φ_{ST}) among population pairs based on 450 bp mtDNA control region sequence.

H_O: The genetic distance between two population pairs is ≤ 0, indicating gene flow connects the populations. Statistically significant estimates (in **bold**; $P \leq 0.05$) indicate limited gene flow, resulting in genetic divergence. Estimates of $\Phi_{ST} \geq 0.05$ indicate moderate to high differences.

	1	2	3	4	5	6	7	8	9	10	11	12	13
1: SE AK - BC	0.000												
2: Vancouver Island	-0.073	0.000											
3: Alberta	-0.061	-0.029	0.000										
4: Great Lakes	0.013	0.057	-0.026	0.000									
5: Appalachian	-0.022	0.035	-0.035	-0.005	0.000								
6: Northern Rockies	0.002	0.015	0.036	-0.004	0.015	0.000							
7: Rocky Islands	0.007	0.068	-0.023	-0.025	-0.002	-0.007	0.000						
8: Colorado Rockies	0.003	0.064	0.128	**0.043**	**0.029**	0.016	0.034	0.000					
9: Northern Utah	-0.015	-0.013	0.139	0.042	**0.051**	-0.018	**0.051**	0.021	0.000				
10: Colorado Plateau	0.044	-0.040	0.034	**0.122**	**0.121**	**0.061**	**0.105**	**0.100**	0.029	0.000			
11: Jemez NM	**0.155**	0.135	0.232	**0.233**	**0.253**	**0.248**	**0.254**	**0.211**	**0.243**	**0.124**	0.000		
12: Sky Islands, AZ	**0.348**	**0.468**	**0.461**	**0.404**	**0.412**	**0.458**	**0.440**	**0.470**	**0.454**	**0.307**	**0.508**	0.000	
13: Sierra-Cascade	0.088	-0.064	**0.251**	**0.203**	**0.193**	**0.134**	**0.225**	**0.136**	**0.083**	**0.062**	**0.212**	**0.508**	0.000

CONCLUSIONS

1. Goshawks exhibited significant genetic structure at broad geographic scales. The geographic pattern of major mtDNA lineages indicated goshawks were historically isolated within at least three glacial period refugia (Pacific Coastal, Southwestern and Eastern), and have subsequently experienced limited post-glacial range expansion from these regions.
2. Goshawks in California, Arizona Sky Islands, and Jemez Mountains, New Mexico, had less genetic diversity compared to other populations, and were predominately represented by region-specific haplotypes, indicating historical and some degree of contemporary isolation.
3. Goshawks in the Northern Rocky Mountains and the East had relatively high diversity and lacked significant genetic structure. We hypothesized that rapid post-glacial gene flow of a single haplotype (haplotype-B) from eastern populations into the Rocky Mountains has partially homogenized the genetic structure among these geographically distant populations.
4. Gene flow south into the Arizona Sky Islands from northern populations has been limited. We hypothesized that the dominance of a single region specific haplotype (haplotype-E) is a result of long-term isolation of this lineage in the Southwest, and/or the Sky Island population has been more demographically influenced by northward dispersal of goshawks from Mexico.
5. We found support for six major regional populations based on haplotype differences: Southeast Alaska, California, Colorado Plateau, Arizona Sky Islands, New Mexico and a large group including the Rocky Mountains, Great Lakes and Central Appalachian Mountains.

NEW DATA: PRELIMINARY RESULTS FROM MICROSATELLITES

	1	2	3	4	5	6	7	8	9	10
1: Great Lakes	0.000									
2: Appalachia	0.035	0.000								
3: Northern Rockies	0.003	0.036	0.000							
4: No. Colorado	0.017	**0.048**	0.016	0.000						
5: So. Colorado	0.060	0.099	0.012	0.036	0.000					
6: Kaibab Plateau	0.032	**0.065**	-0.003	**0.046**	0.001	0.000				
Jemez, NM	0.043	0.097	0.002	0.080	0.005	-0.017	0.000			
Sky Islands, AZ	**0.110**	**0.144**	**0.058**	**0.099**	0.007	**0.034**	0.045	0.000		
Sierra Nevada, CA	0.005	0.060	-0.005	0.037	0.048	0.016	0.013	**0.088**	0.000	
Klamath, CA	0.017	**0.087**	0.018	**0.047**	**0.081**	**0.043**	0.048	**0.137**	-0.007	0.000

Table 2. Estimates variation in allele frequency (F_{ST}) among population pairs based on three microsatellite loci.

H_O: Variation in allele frequencies between two populations pairs is ≤ 0, indicating gene flow connects the populations. Statistically significant estimates (in bold; $P \leq 0.05$) indicate limited gene flow. Estimates of $F_{ST} \geq 0.05$ indicated moderate to high differences.

Table 3: Estimates of F_{IS}, an indicator of inbreeding and heterozygote deficiency. Negative values indicate heterozygote excess as compared to Hardy-Weinberg expectations. The only population exhibiting a substantial trend for inbreeding was seen in the Jemez Mountains of New Mexico. F_{IS} 0.05 to 0.15 is considered to be indicative of moderate levels of inbreeding..

Genetic Markers	Kaibab Plateau, AZ	Sky Islands, AZ	Jemez Mtn., NM	Klamath, CA	Sierra Nevada	SO. Colorado	NO. Colorado	NO. Rockies	Great Lakes	Central Appalachia
Age1a	-0.017	-0.335	0.023	-0.364	-0.168	0.316	0.194	0.16	-0.063	-0.132
Age2	0.064	-0.195	0.667	0.083	0.153	0.191	0.059	0.156	0.057	-0.102
Age4	-0.014	0.036	-0.077	-0.06	-0.129	-0.228	-0.01	-0.054	0.022	0.044
All loci combined	0.012	-0.144	0.216	-0.086	-0.039	0.092	0.075	0.083	0.011	-0.045

Author contact: sbayard_64@yahoo.com

† Department of Biology & Graduate Degree Program in Ecology, Colorado State University, Fort Collins, CO 80523. USA

†† Rocky Mountain Research Station, 2150 Centre Ave., Bldg. A, Ste. 350., Fort Collins CO 80526. USA

‡ Brigham Young University, Provo, Utah. USA

± Alaska Science Center, USGS, Anchorage, Alaska. USA