Int. J. Plant Sci. 177(2):132–144. 2016.

1058-5893/2016/17702-0003$15.00 DOI: 10.1086/684180

PRESENT-DAY GENETIC STRUCTURE OF THE HOLOPARASITE *CONOPHOLIS AMERICANA* (OROBANCHACEAE) IN EASTERN NORTH AMERICA AND THE LOCATION OF ITS REFUGIA DURING THE LAST GLACIAL CYCLE

Anuar Rodrigues[1,*] and Saša Stefanović*

*Department of Biology, University of Toronto, Mississauga, Ontario L5L 1C6, Canada

Editor: Félix Forest

Premise of research. Understanding how various organisms respond to previous changes in climate could provide insight into how they may respond or adapt to the current changes. *Conopholis americana* has a broad distribution across eastern North America, covering both previously glaciated and unglaciated regions. In this study, we investigated the postglacial history and phylogeographic structure of this parasitic plant species to characterize its genetic variation and structure and to identify the number and locations of refugia.

Methodology. Molecular data from 10 microsatellite markers and DNA sequences from the plastid gene/introns (*clpP*) were obtained for 281 individuals sampled from 75 populations spanning the current range of the species in eastern North America and analyzed using a variety of phylogeographic methods. Distribution modeling was carried out to determine regions with relatively suitable climate niches for populations at the Last Glacial Maximum (LGM) and present.

Pivotal results. We inferred the persistence of a minimum of two glacial refugia for *C. americana* at the LGM, one in north-central Florida and southern Alabama and another in the Appalachian Mountains near the southern tip of the Blue Ridge Mountains. High levels of genetic diversity were observed across the southern Appalachian Mountains, the region where populations from two refugia come together following recolonization northward.

Conclusions. The genetic and geographic patterns revealed by our results provide further evidence of the dynamic nature and phylogeographical history of eastern North American taxa. The discovery of a distinct southern lineage is in agreement with the location of a previously proposed southern glacial refugium spanning across Florida, southern Georgia and Alabama, and the Lower Mississippi Valley. The second lineage is dominant across the present northern range of the species and is hypothesized to have been located in the southern extent of the Blue Ridge mountain range of the Appalachian Mountains at the LGM.

Keywords: eastern North America, Last Glacial Maximum, parasitic plant, plastid DNA, phylogeography, microsatellites.

Introduction

The Pleistocene epoch (approximately 2.6–0.01 Ma) was a time of great climate change that consisted of long glacial periods separated by shorter warm interglacial periods. Climate conditions during this time were greatly variable, with more than 20 glacial cycles recorded, resulting in major alterations to the landscape (Pielou 1991; Williams et al. 1998; Hewitt 2000). During the Last Glacial Maximum (LGM; ~20–18 kya) at the end of the Wisconsin glaciation, the Laurentide ice sheet covered most of North America (Canada and the northern United States). In eastern North America, the ice margin extended south to an area comprising what is today the states of New York, Pennsylvania, and Ohio and covering all of the Great Lakes. Permafrost and tundra continued even farther south, beyond the leading edge of the ice sheet. Due to the extreme climate during this time, the distribution of species would have been greatly different from what we see today. As a result of the great expanse of the Laurentide ice sheet at the LGM, most species experienced a reduction or fragmentation in their habitat and population size and would have been confined to southern ice-free refugia in order to escape the harsh environment that made much of North America uninhabitable (Hewitt 1996).

For plants, the fossil record and pollen data suggest that as the ice sheet advanced in North America, it progressively eliminated habitable regions to the north. As a result, the deciduous forest in this region retreated south of 33°N, spanning across the region that today comprises the states of Florida and Georgia as well as the Lower Mississippi Valley (Davis 1981;

[1] Author for correspondence; e-mail: anuar.rodrigues@utoronto.ca.

Manuscript received May 2015; revised manuscript received July 2015; electronically published December 18, 2015.

Delcourt and Delcourt 1993; Jackson and Overpeck 2000; Jackson et al. 2000; Soltis et al. 2006). As temperatures rose and the glaciers melted at the end of the LGM, populations were able to expand their geographic distributions northward and recolonize new areas that became suitable (Pielou 1991). With the migration of populations northward from southern refugia, genetic patterns become evident, following a "southern richness to northern purity" scenario (Hewitt 2000). This hypothesis posits that higher genetic diversity should be found in populations that now occupy southern, previously nonglaciated regions and predicts the loss/reduction in diversity by those populations moving northward along the axes of recolonization. In North America, this is supported by phylogeographic studies that have revealed a lower genetic diversity in northern populations compared to those from the south (Hewitt 1996, 2004; McLachlan et al. 2005).

In contrast to the traditional view described above, recent phylogeographic studies have also shown that more northern, smaller, cryptic refugia may have existed (Stewart and Lister 2001; Jaramillo-Correa et al. 2004; Godbout et al. 2005; McLachlan et al. 2005). These northern regions would have been located in ice-free areas that persisted near the ice margin as well as on the peaks of mountains protruding through the ice sheet (nunataks). At these locations, isolated populations may have persisted at low population numbers that would not allow detection in the pollen record (Jackson et al. 2000). Given the increasing genetic evidence suggestive of a wider-ranging assemblage of temperate and boreal taxa at midlatitudes (Rowe et al. 2004; Soltis et al. 2006; Gonzalez et al. 2008), the boundaries that define breaks in vegetation assemblages now seem to be less well defined. These findings add to the complex history of postglacial colonization following the LGM. They affect our interpretation of how plants respond to changes in climate and whether distributional ranges can be better explained by range expansions from more northern-located refugia or by dispersal events from the traditionally hypothesized southern refugia. To discern among these hypotheses and study patterns and the tempo of postglacial history in eastern North America, species with broad distributions are needed, spanning both previously glaciated and unglaciated regions.

In a recent molecular phylogenetic study of the North American holoparasitic genus *Conopholis* (Orobanchaceae; Rodrigues et al. 2011), *Conopholis americana* (L.) Warll. was identified as one of its three major lineages and was subsequently confirmed as a distinct species by comprehensive morphometric analyses (Rodrigues et al. 2013). The distribution of *Conopholis* in eastern North America spans across (1) several known barriers to plant movement on the continent and (2) locations of glacial refugia (both known traditional locations and cryptic refugia; Soltis et al. 2006). *Conopholis americana* is primarily distributed throughout the eastern United States and adjacent Canada, from Nova Scotia to Wisconsin in the north and from Florida to Alabama in the south, with some of its populations found in southern Mexico as disjunct members of this species (Rodrigues et al. 2011, 2013). In eastern North America, these plants are found in moist, deciduous, or mixed forests attached to the roots of red oaks (*Quercus* section *Lobatae*) via haustoria.

Populations of *Conopholis* can be described as rare and isolated, at times being separated by kilometers but usually locally abundant where present. These plants do not possess floral nectaries, and they are not known to attract insect pollinators by producing a fragrance. Studies of flowers postanthesis have found the anthers of *C. americana* to be in physical contact with the stigma (Baird and Riopel 1986). This, combined with bagging experiments aimed at exploring the role of wind and insects in pollination (Baird and Riopel 1986), suggests selfing as a mode of pollination for this species. Plants such as *C. americana* that are self-fertilizing (Baird and Riopel 1986) and have relatively limited dispersal ability, especially given the reliance on their host for survival, represent great model systems that can be used to identify the location of northern refugia. Namely, such species likely cannot rely primarily on means of long-distance dispersal to explain the expanse of their geographic range seen today but instead may have existed in small populations closer to the ice margin, from which they could expand their range following the retreat of the glaciers.

The overarching goal of this study was to investigate the glacial history of *C. americana* in eastern North America. Our specific objectives were to employ phylogeographic analyses using plastid and nuclear markers along with species distributional modeling to (1) determine the genetic variation across the range of *C. americana* in eastern North America, (2) quantify the phylogeographic structure, (3) identify refugia locations and recolonization history, (4) attempt to shed further light on their breeding system, and (5) use the results as a proxy for the host range expansion (red oaks).

Material and Methods

Taxon Sampling and DNA Extraction

A total of 281 individuals from 75 populations were used in this study, covering essentially the entire range of *Conopholis americana* in eastern North America (e.g., fig. 1*b*). A complete list of collecting locations and sample sizes is provided in table 1. Total genomic DNA was extracted from fresh or silica-dried material and purified as described by Rodrigues et al. (2011).

Plastid clpP *Sequencing*

Because the plastid genome is nonrecombinant and usually only maternally inherited (Reboud and Zeyl 1994), it can be used to identify the genetic signature of maternal lineages. Also, owing to a relatively low mutation rate observed in these genomes (Wolfe et al. 1987), the majority of alleles (haplotypes) recovered are the genealogical derivatives of distinct lineages that predate postglacial colonization (McLachlan 2005). Therefore, the modern geographic distribution of plastome haplotypes is expected to largely correspond to the migration routes of expanding populations from glacial refugia, and a direct mutational relationship among the haplotypes can be detected and traced. To assess haplotype diversity of *C. americana*, we targeted the plastid *clpP* gene and its introns. Polymerase chain reactions, amplicon purification, and sequencing for all 281 sampled individuals were carried out as described in our phylogenetic study (Rodrigues et al. 2011). Newly generated sequences were deposited in GenBank under accession numbers KU000617–KU000897.

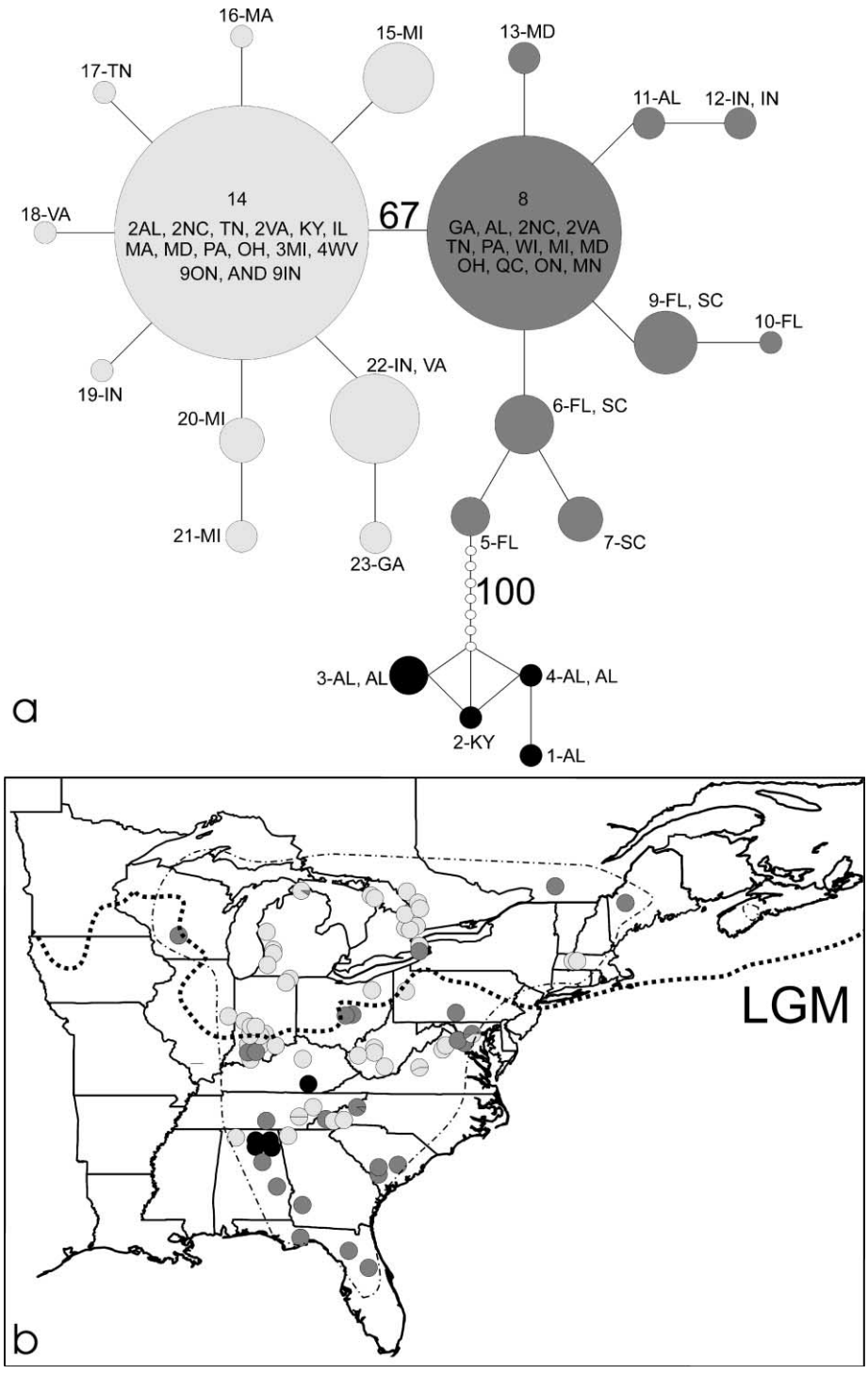

Fig. 1 Plastid haplotype reconstruction and its distribution for *Conopholis americana* in eastern North America. *a*, Statistical parsimony network shows the relationships between the 23 recovered haplotypes; the size of each circle is proportional to the number of individuals found to have that particular haplotype. Haplotypes are numbered, and their provenance (state or province) is indicated. See table 1 for further details on haplotype identity and major groups to which they belong. Lines connecting haplotypes represent a single mutational step, while small open circles represent unsampled/extinct haplotypes. Numbers adjacent to lines represent measures of bootstrap support. Shades of haplotypes correspond to the three major groups recovered; bootstrap values supporting their separation are shown (≥60%). *b*, Map showing the distribution of the plastid haplotypes. Dashed line spanning east to west shows the approximate location of the ice margin at the Last Glacial Maximum (LGM). Area inside dash/dot pattern shows the current distribution of *C. americana* in eastern North America.

Microsatellite Genotyping

It has become increasingly clear that conclusions drawn from results based solely on a single nonrecombining region/gene of the plastid genome can potentially be misleading (Schaal et al. 1998; Brito and Edwards 2009). To make more reliable inferences of both past population history and present population structure, phylogeographic studies are moving toward using markers from the organellar genomes along with multiple unlinked nuclear markers in combination with species-distribution modeling. The same 281 individuals (table 1) were genotyped for 10 unlinked microsatellite loci developed and characterized for *C. americana* (SSR6, SSR9, SSR22, SSR27, SSR33, SSR42, SSR43, SSR49, SSR51, SSR56) following the methods for amplification and scoring detailed by Rodrigues et al. (2012).

Analyses of Population Diversity and Structure

Plastid sequences were aligned manually in Se-Al, version 2.0a11 (Rambaut 2002). Gaps in the alignment were treated as missing data. However, indels were coded, and binary codes were appended to the nucleotide sequences. A statistical parsimony haplotype network was constructed using TCS, version 1.21 (Clement et al. 2000). Support for relationships among major lineages was inferred from nonparametric bootstrapping (Felsenstein 1985) implemented in PAUP*, version 4.0b10 (Swofford 2002), using 500 pseudoreplicates (each with 20 random sequence addition cycles), tree bisection and reconnection branch swapping, and MULTREES option turned off (DeBry and Olmstead 2000). Levels of within-population plastid diversity were calculated for populations using the Arlequin software package, version 3.5.1.2 (Excoffier and Lischer 2010): the number of polymorphic sites (S), the number of haplotypes (h), and the nucleotide diversity (π; specifically, for populations with a sample size of $N \geq 4$).

The structure of the nuclear microsatellite data was explored using the Bayesian approach implemented in Bayesian Analysis of Population Structure (BAPS), version 5.3 (Corander et al. 2003). Using BAPS, groups of genetically similar populations are identified that have restricted gene flow between them. In preliminary analyses, 10 replicates were run for all possible numbers of clusters (K) up to a maximum of 75, the number of populations sampled in our study. We found that 23–35 clusters were supported by the data, with 35 being the optimal number of clusters. Therefore, the final runs were performed with the maximum possible number of clusters set to 35. To infer the relationships between the optimal number of clusters recovered, the distances between clusters obtained in BAPS were used to produce a neighbor-joining (NJ) tree in PAUP* (Swofford 2002).

To test the significance of the groups recovered using both TCS for the plastid-sequenced data and BAPS for the microsatellite genotyped data, we performed an analysis of molecular variance (AMOVA) implemented in Arlequin (Excoffier and Lischer 2010). Each data set was partitioned into the groups recovered: (1) plastid-sequenced data partitioned by TCS major and (2) microsatellite genotyped data grouped according to major BAPS lineages. In addition to partitioning the data in this manner, we also divided both data sets according to geographic barriers: (1) populations currently found north versus south of where the ice margin was located at the LGM and (2) populations found north, on, and south of the Appalachian mountain range.

As a further test of phylogeographic structure throughout the entire range of *C. americana* in eastern North America, we compared measures of genetic differentiation among populations using the program SPADS (Dellicour and Mardulyn 2014). The G_{ST} parameter is an unordered measure of genetic variation that depends on the frequencies of haplotypes and does not incorporate their relationships. The N_{ST} parameter, on the other hand, takes into account both the haplotype frequencies and the genetic distance between haplotypes. Parameters that account for the relative degree of similarity between haplotypes (N_{ST}) make better use of the data inherent in haplotype data compared to measures that account only for haplotype frequencies (G_{ST}; Petit et al. 2005). As a result, a significantly higher N_{ST} value compared to G_{ST} indicates that haplotypes that are more closely related tend to be found in a given geographical area (Pons and Petit 1996).

Distribution Modeling

Ecological niche modeling provides us with the tools necessary to determine the location(s) of suitable habitats for species and define potential distributional ranges at the LGM. In combination with traditional molecular phylogeographic data, it can be used to offer an independent, more objective, and more spatially defined hypothesis for the geographic distributions and patterns of species in the past (Waltari et al. 2007). To determine regions with relatively suitable climate niches for lineages/populations of *C. americana* in eastern North America at present-day conditions and at the LGM (ca. 21 kyBP), we took advantage of WorldClim climate data (Hijmans et al. 2005) available for 19 bioclimatic factors. These environmental conditions summarize aspects of climate that may be particularly relevant for determining species distribution and their limits. Employing the maximum entropy approach implemented in Maxent, version 3.3.3 (Phillips et al. 2006; Phillips and Dudik 2008), we used the data to predict where individuals of this species are most likely to occur. Maxent generates ecological niche models utilizing presence-only species records and contrasts them with pseudoabsence data sampled from the remainder of the study area. Layers were trimmed to the area surrounding North America and projected across the same dimensions after modeling. Present-day species occurrence data for *C. americana* in eastern North America (640 entries) were downloaded from the Global Biodiversity Information Facility data portal (http://www.gbif.org). Prior to analyses, the data were mapped, and any points that were not in the geographic range of *C. americana* specific to eastern North America were removed from the data set. In addition, duplicate sample records were removed to avoid the effects of spatial autocorrelation. For the present niche model predictions, we used the 19 bioclimatic variables from the WorldClim data set with a 2.5-min spatial resolution (Hijmans et al. 2005). For the LGM climate, data layers representative of that time were derived from the community climate system model (Collins et al. 2004) at the same resolution (2.5 min). Climate variables were evaluated with ENMTools (Warren et al. 2010) to minimize model overfitting and remove correlated variables. The following six variables were retained for analyses: annual mean temperature, minimum

Table 1

Collection and Label Information for *Conopholis americana* Populations Used in This Study

State/province and county	Accession label	Sample size	Latitude (°)	Longitude (°)	Plastid haplotype	Microsatellite genotype	Plastid statistics		
							S	*h*	π
Alabama:									
Lee	SS.05.01	5	32.5228	85.4969	P2; 8	M1; 21	0	1	0
Madison	SS.05.79	3	34.7293	86.5505	P1; 1, 4	M2 & M3; 5, 11	2	2	…
Madison	SS.05.80	1	34.7209	86.5320	P1; 4	M3; 11	0	1	…
Marshal	SS.06.98	2	34.5577	86.2089	P2; 11	M1; 15	0	1	…
Lauderdale	SS.06.103	1	34.8113	87.3574	P3; 14	M2; 25	0	1	…
Jackson	SS.06.161	1	34.7090	86.0074	P1; 3	M2; 25	0	1	…
Jackson	SS.06.162	2	34.6971	86.0247	P1; 3	M1 & M2; 15, 6	0	1	…
Florida:									
Marion	AC.FL.JS	2	29.1778	81.7164	P2; 6	M1; 12	0	1	…
Alachua	AC.FL.SF	5	29.7299	82.4348	P2; 9, 10	M1; 3, 12	7	2	.001846
Wakulla	SS.06.48	3	30.1325	84.3570	P2; 5	M1; 20	0	1	…
Georgia:									
Early	SS.06.24	3	31.4639	84.9229	P2; 8	M1; 13	0	1	…
Dade	SS.06.80	2	34.8407	85.4829	P3; 23	M2; 25	0	1	…
Illinois:									
Vermillion	AC.IL.KSP	9	40.1246	87.7352	P3; 14	M2; 24	0	1	0
Indiana:									
Monroe	SS.03.29	1	39.0202	86.3713	P3; 14	M3; 2	0	1	…
Monroe	SS.03.30	1	39.0246	86.3607	P3; 14	M3; 7	0	1	…
Monroe	SS.03.31	1	39.0348	86.3213	P3; 19	M2; 25	0	1	…
Lawrence	SS.04.80	1	38.7319	86.4175	P3; 14	M2; 25	0	1	…
Perry	SS.04.83	1	37.9925	86.5938	P3; 22	M2; 25	0	1	…
Martin	SS.04.89	1	38.6676	86.7159	P3; 14	M3; 19	0	1	…
Crawford	SS.04.93	1	38.3689	86.6414	P2; 12	M2; 25	0	1	…
Crawford	SS.04.94	1	38.3706	86.6469	P2; 12	M3; 19	0	1	…
Parke	SS.04.96	1	39.8916	87.2032	P3; 14	M3; 19	0	1	…
Monroe	SS.04.102	1	39.2039	86.5298	P3; 14	M3; 7	0	1	…
Monroe	SS.04.109	1	39.1955	86.5201	P3; 14	M3; 7	0	1	…
Steuben	SS.04.170	1	41.6850	85.0074	P3; 14	M3; 19	0	1	…
Steuben	SS.09.28	9	41.7123	85.0258	P3; 14	M3; 16, 19	0	1	0
Clarke	AC.IN.CCF	10	38.4856	85.8326	P3; 22	M2; 17	0	1	0
Kentucky:									
McCreary	SS.03.11	1	36.6912	84.4697	P1; 2	M2; 25	0	1	…
Montgomery	AC.KY.MC	2	38.0716	83.9347	P3; 14	M2; 25	0	1	…
Maine:									
Franklin	SS.09.38	6	44.7554	70.0740	P2; 8	M3; 8	0	1	0
Maryland:									
Montgomery	AC.MD.MT	20	39.0006	77.2101	P2 & P3; 8, 13, 14	M2 & M3; 9, 11	10	3	.003125
Massachusetts:									
Hampshire	AC.MA.MN	14	42.3051	72.5127	P3; 14	M2; 23	0	1	0
Hampshire	AC.MA.RK	1	42.3061	72.4965	P3; 16	M3; 8	0	1	…
Michigan:									
Ottawa	SS.05.82	1	42.7958	86.1008	P3; 14	M3; 19	0	1	…
Van Buren	SS.09.30	8	42.3331	86.2992	P3; 20, 21, 22	M3; 19	4	3	.001325
Allegan	SS.09.31	6	42.6989	86.1954	P3; 14	M3; 19	0	1	0
Muskegon	SS.09.32	8	43.4105	86.3287	P3; 14	M3; 19	0	1	0
Cheboygan	SS.09.37	11	45.5505	84.6674	P2 & P3; 8, 15	M2; 22	11	2	.001318
North Carolina:									
Macon	SS.06.64	3	35.2064	83.4206	P3; 14	M3; 2	0	1	…
Swain	SS.06.146	2	35.5053	83.6761	P2; 8	M1; 14	0	1	…
Jackson	SS.06.160	2	35.4239	83.0848	P3; 14	M2; 25	0	1	…
Madison	AC.NC.L	10	35.7333	82.8697	P2; 8	M2; 25	0	1	0
Ohio:									
Licking	SS.06.173	1	40.0736	82.5193	P2; 8	M3; 8	0	1	…
Granville	SS.06.174	1	40.0705	82.5330	P2; 8	M3; 8	0	1	…
Summit	SS.09.25	11	41.2609	81.5693	P3; 14	M3; 1, 8	0	1	0
Ontario:									
Simcoe	SS.05.02	1	44.3991	79.8561	P3; 14	M3; 8	0	1	…
Halton	SS.05.94	2	43.4257	79.8817	P3; 14	M3; 8	0	1	…

Table 1 (*Continued*)

State/province and county	Accession label	Sample size	Latitude (°)	Longitude (°)	Plastid haplotype	Microsatellite genotype	Plastid statistics		
							S	*h*	π
Township of Archipelago	SS.05.194	2	45.3401	80.0457	P3; 14	M3; 8	0	1	…
Halton	SS.06.170	2	43.5066	79.9594	P3; 14	M3; 8	0	1	…
Peel	SS.08.03	6	43.5524	79.6636	P3; 14	M3; 10	0	1	0
Bruce	SS.08.04	6	45.2311	81.5983	P3; 14	M3; 19	0	1	0
Bruce	SS.08.05	6	45.2005	81.5331	P3; 14	M3; 19	0	1	0
Lincoln	SS.09.05	2	43.1348	79.1575	P3; 14	M3; 8	0	1	…
Lincoln	SS.09.08	6	42.9095	79.2748	P2; 8	M3; 8	0	1	0
Simcoe	SS.09.39	4	44.8497	79.9978	P3; 14	M3; 8	0	1	0
Simcoe	SS.09.40	3	44.7583	79.8281	P3; 14	M3; 8	0	1	…
Pennsylvania:									
Butler	SS.07.42	1	40.9434	80.0875	P3; 14	M3; 8	0	1	…
Franklin	AC.PA.MSF	9	39.8246	77.5160	P2; 8	M3; 11	0	1	0
Quebec:									
Vallée-du-Richelieu	SS.07.80	1	45.5491	73.3569	P2; 8	M3; 8	0	1	…
South Carolina:									
Hampton	SS.06.53	4	32.8321	81.1755	P2; 7	M1; 12	0	1	0
Banberg	SS.06.54	5	33.0480	81.0970	P2; 9	M2; 25	0	1	0
Dorchester	SS.06.63	4	33.0636	80.6167	P2; 9	M1; 4	0	1	0
Tennessee:									
Franklin	SS.05.81	1	35.0533	86.2732	P2; 8	M1; 15	0	1	…
Blount	SS.06.127	2	35.7256	83.5083	P3; 14	M2; 25	0	1	…
Blount	SS.06.133	2	35.6317	83.9435	P3; 14, 17	M1 & M2; 14, 25	12	2	…
Virginia:									
Rappahannock	SS.07.57	1	38.8194	78.1801	P3; 18	M2; 25	0	1	…
Shenandoah	SS.07.58	1	38.6904	78.3323	P3; 14	M2; 25	0	1	…
Rockbridge	AC.VA.NB	7	37.6776	79.5073	P3; 14, 22	M3; 11, 18, 19	2	2	.000757
Fairfax	AC.VA.TRP	5	38.9579	77.1627	P2; 8	M3; 11	0	1	0
Fairfax	AC.VA.TRP.2	9	38.9474	77.2692	P2; 8	M3; 11	0	1	0
West Virginia:									
Kanawha	SS.04.71	1	38.3172	81.6680	P3; 14	M2; 25	0	1	…
Kanawha	SS.04.72	1	38.2529	81.6568	P3; 14	M2; 25	0	1	…
Summers	SS.04.75	1	37.5421	80.9599	P3; 14	M2; 25	0	1	…
Cabell	AC.WV.BFL	2	38.3230	82.3785	P3; 14	M2; 25	0	1	…
Wisconsin:									
Sauk	AC.WI.DL	6	43.4283	89.7274	P2; 8	M3; 19	0	1	0

Note. For each population, plastid haplotypes are labeled (1–23), and the major group to which they belong is indicated (P1–P3). Likewise, microsatellite genotypes are labeled (1–25), and the major genotype cluster in which they were found is indicated (M1–M3). Accompanying plastid nucleotide diversity statistics are provided. h = number of haplotypes; S = number of polymorphic sites; π = nucleotide diversity (an ellipsis indicates that this value was not calculated due to small population size).

temperature of coldest month, mean temperature of wettest quarter, mean temperature of driest quarter, mean temperature of coldest quarter, and annual precipitation. In Maxent, the models were run using the default convergence setting (10^{-5}) with 1000 iterations, using 25% of the localities for model training. Maxent outputs a continuous surface value ranging from 0 to 1, indicating regions of potentially suitable climate niches where individuals/populations of the species could be found. When projected onto the reconstructed LGM data, it can be used to identify potential refugial locations.

Results

Plastid Sequencing and Analysis

We amplified and sequenced the plastid *clpP* gene and its introns for all 281 individuals of *Conopholis* from 75 populations. Sequences were readily alignable and resulted in an overall alignment length of 1553 bp. Scoring indels resulted in an additional 13 characters that were appended to the nucleotide matrix. Most of the populations were fixed for a single haplotype, while nucleotide diversity (π) for populations with a sample size of 4 or greater ranged from 0.000 (many populations) to 0.003125 (accession AC.MD.MT from Montgomery County, Maryland; table 1). The network constructed from the combined data (nucleotides plus indels) revealed a total of 23 distinct haplotypes. Based on a combination of bootstrap support, presence of unambiguous characters, and substantial branch length subtending them, we divided these haplotypes into three major groups (plastid groups P1, P2, and P3; table 1; fig. 1) and shaded them black, dark gray, and light gray, respectively, in figure 1. There were clear geographical differences in haplotype frequencies of the three major groups. Plastid group P3 is centered to the north, while group P2 is centered to the south. These

two haplotype groups each have a broad distribution, with a large area of overlap in their range (fig. 1*b*). The third group, P1, comprises four very distinct and unique haplotypes found in the central overlapping region, specifically limited in distribution to the southern tip of the Appalachian Plateau (northeastern Alabama and southeastern Kentucky). Populations sampled from across the Appalachians are the most diverse, with representatives of all three haplotype groups present. The states of Florida and South Carolina along with the southern ranges of Alabama and Georgia (all of which are found south of the Appalachians) are found to have populations of *Conopholis americana* belonging only to haplotype group P2. With respect to the LGM boundary, both haplotype groups P3 and P2 are found north of the ice margin, but those belonging to P3 dominate the northern range. None of the haplotypes belonging to group P1 are found north of the LGM line.

A high proportion of variation resulted from differences between populations ($G_{ST} = 0.8775$ and $N_{ST} = 0.9410$). The N_{ST} parameter was calculated to determine whether closely related haplotypes would cluster according to their geographical location. Significant phylogeographical structure was detected ($N_{ST} - G_{ST} = 0.0635$; $P = 0.0010$).

Microsatellite Genotyping and Analysis

Of the 10 loci targeted in this study, there was only one locus with >5% missing data (SSR42). The number of alleles per locus ranged from 2 to 11, with a mean value of 6.5. No significant linkage disequilibrium was detected between any of 45 pairs of microsatellite loci tested, in accordance with our previous results (Rodrigues et al. 2012) based on a smaller sampling. Levels of observed and expected heterozygosity ranged from 0 to 0.375 and from 0 to 0.4, respectively. Bayesian analysis of population structure resulted in 25 genetically distinct clusters. The unrooted NJ dendrogram (fig. 2*a*) produced from the distance matrix obtained in BAPS revealed the relationship between these genotype groups. Based on a combination of branch lengths and geographical distribution, we divided these haplotypes into three major groups (microsatellite groups M1, M2, and M3) and shaded them black, dark gray, and light gray, respectively, in figure 2. Unlike the plastid case, microsatellite data show only a narrow zone of overlap between these three groups, and each group has a relatively broad distribution. Namely, genotype group M1 dominates the very southern range of *C. americana* in eastern North America, group M2 has a distribution that is more concentrated in the central region, and populations from group M3 are the predominant genotype in the north. None of these three groups is unique and distinct to a very specific and localized region. On the other hand, similar to what has been seen with the plastid data, populations from the Blue Ridge Mountains and the Appalachians in general are the most diverse, with all three groups represented in that region. Of the 28 populations north of the LGM line, 24 belong to group M3 (light gray) and only four belong to the genotypes of group M2 (dark gray). None of the populations belonging to group M1 (black) are found north of the ice margin; this group does not extend beyond southern Tennessee and southwestern North Carolina. When the NJ dendrogram is rooted by the midpoint method, the distribution of these three groups largely follows a latitudinal subdivision, where southern genotypes from group M1 are sister to central and northern genotype groups M2 and M3, respectively.

The AMOVA confirmed that the best regional differentiation is based on plastid haplotype groupings. Comparison of clusters recovered using both TCS and BAPS revealed that a significant proportion of the genetic variance was explained by differences among groups when the plastid data was partitioned according to the three major plastid groups recovered following TCS analyses ($F_{CT} = 0.89$). The microsatellite data sorted by BAPS clusters recovered resulted in $F_{CT} = 0.19$. When geographic boundaries were compared, we observed high and significant structure in the plastid data (LGM: $F_{CT} = 0.34$; Appalachians: $F_{CT} = 0.38$) compared to the microsatellite data (LGM: $F_{CT} = 0.07$; Appalachians: $F_{CT} = 0.15$). In the cases when considering the boundaries, most of the variation was partitioned among populations within regions as opposed to between regions.

Distribution Modeling

The geographic distribution of *C. americana* in eastern North America based on current climate data was well modeled by the ecological niche models (fig. 3*a*), as evidenced by the high area under the curve score of 0.982 and the good match between the predicted and the observed current distribution. The modeled distribution accurately shows highly probable areas of habitat extending from as far south as central Florida, north to Nova Scotia, west to Wisconsin, and south to Alabama. When the models were projected onto past reconstructed climate layers at the LGM, suitable regions for the persistence of populations were identified as highly probable in various areas of the south, north-central Florida, coastal Louisiana, the southeastern border of Texas, and the Mexican state of Tamaulipas. In addition, a separate and more northern location with high suitability scores was the tri-state area in the southern reach of the Blue Ridge Mountains where Georgia, South Carolina, and North Carolina border. Relatively habitable locations (with suitability scores between 0.25 and 0.50) extend farther north to straddle the LGM line in southern Indiana and Ohio as well as northern Virginia, West Virginia, and Maryland.

Discussion

The results of our study shed light on the phylogeographic structure of *Conopholis americana* in eastern North America and are suggestive of this plant's breeding system. Most of the populations are fixed for a particular plastid haplotype or microsatellite genotype (table 1). Of the 75 populations surveyed, only seven populations have individuals belonging to more than one haplotype. According to the microsatellite data, the maximum value for observed heterozygosity was 0.375, with an average value of 0.042 across all populations. Of the 75 populations sampled, 67 are fixed for a particular genotype cluster. In addition, the inbreeding coefficient (F_{IS}) was 0.88. This is a measure of the extent of genetic inbreeding within subpopulations, and such a high value is in agreement with the life history and previous bagging experiments (Baird

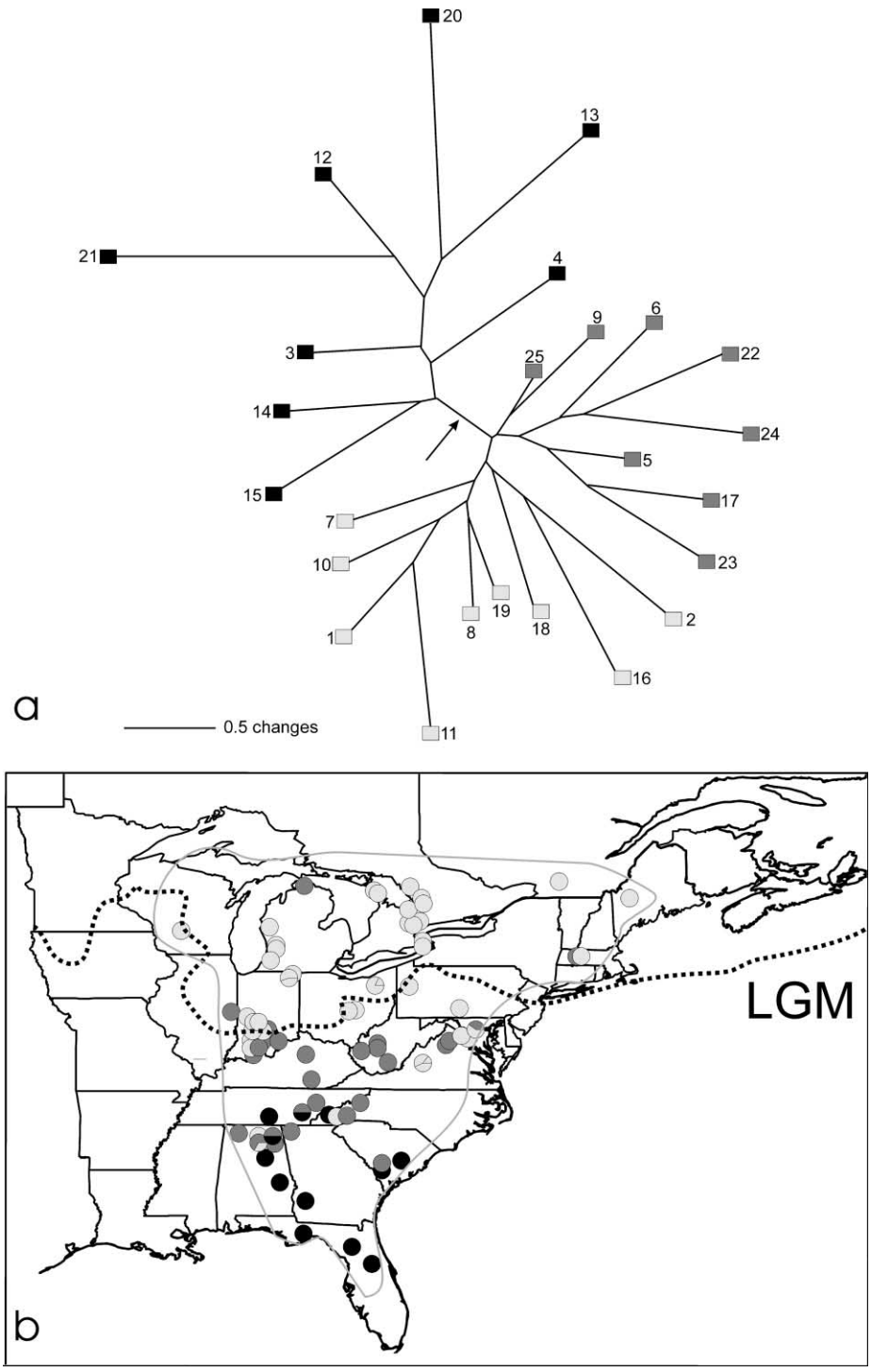

Fig. 2 Microsatellite dendrogram and its distribution for *Conopholis americana* in eastern North America. *a*, Unrooted neighbor-joining dendrogram produced from the distance matrix obtained in BAPS showing the relationship between the genotypes. Shading corresponds to the three major cluster groups recovered based on a combination of branch lengths and geographic distribution. The arrow represents the placement of the root according to the midpoint rooting. *b*, Range map showing the distribution of the microsatellite clusters. Dashed line spanning east to west shows the approximate location of the ice margin at the Last Glacial Maximum (LGM). Area inside the solid light gray outline shows the current distribution of *C. americana* in eastern North America.

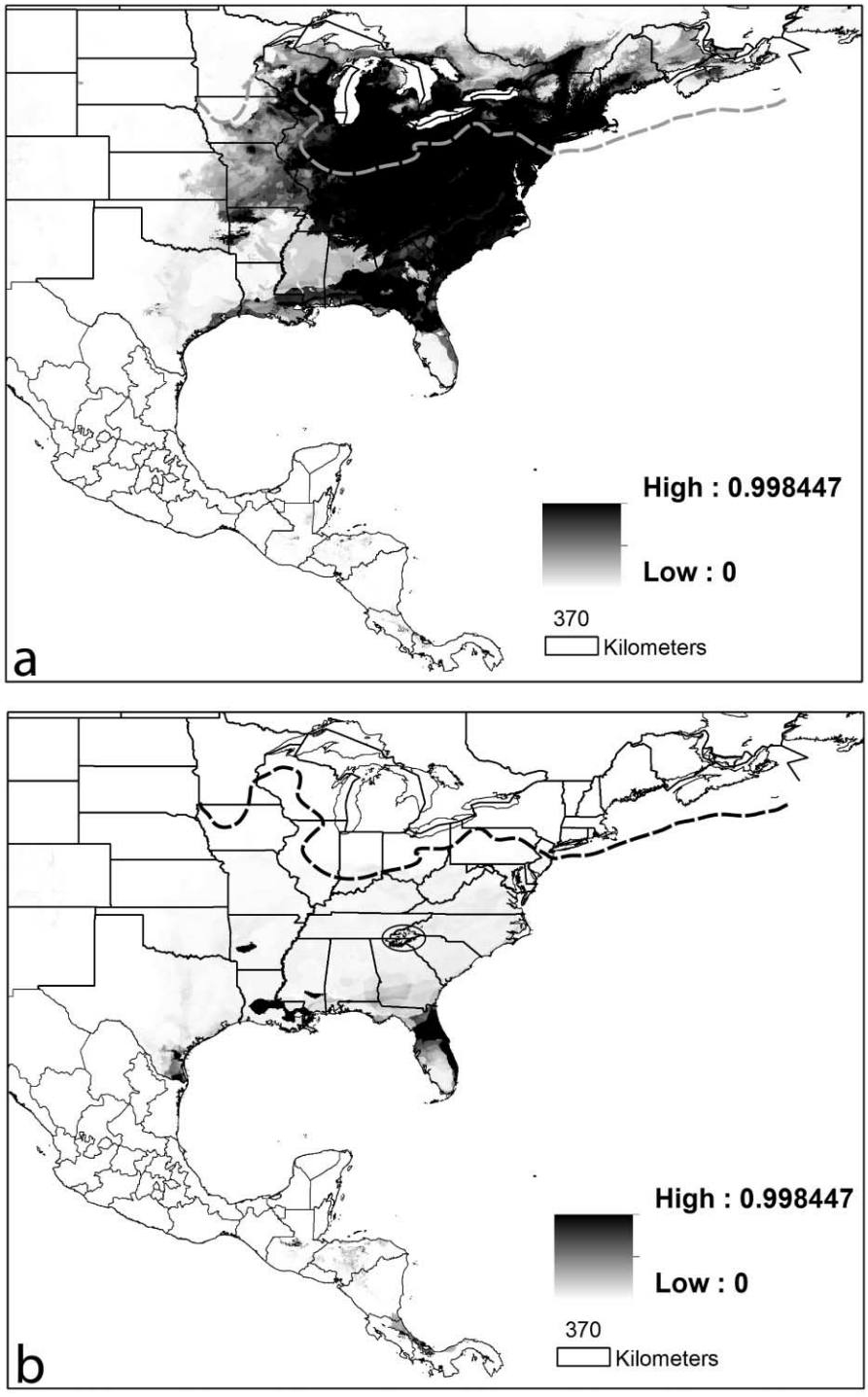

Fig. 3 Ecological niche models showing regions with suitable climate envelopes for *Conopholis americana* in eastern North America based on scenarios for current (*a*) and past (*b*) climate. Past climate (ca. 21 kya) is reconstructed based on the community climate system model. Regions of suitability range from 0 (white) to 1 (black). Dashed line shows the approximate location of the ice margin at the Last Glacial Maximum. Circled region shows the northern location with high suitability scores; see the main text for further discussion.

and Riopel 1986), showing that members of this species are highly self-fertilizing.

If populations of *C. americana* existed in separate refugia during the LGM, each harboring separate haplotypes/genotypes, we would expect to recover genetic differentiation between regions when populations are clustered according to the major haplotype/genotype lineages. Our analyses show that populations of *C. americana* are indeed geographically struc-

tured in eastern North America. Phylogeographic structure is detected when the distribution of phylogenetically related haplotypes contributes to the overall geographic structure of a species (El Mousadik and Petit 1996). Based on the combination of the number of plastid haplotype and microsatellite genotype groups and their locations in addition to the phylogeographical structure that was detected by means of haplotype-identity permutations (N_{ST} [0.9410] > G_{ST} [0.8775]; $P < 0.01$), we infer the persistence of a minimum of two glacial refugia at the LGM from which the populations we see today have likely originated. Such a high degree of structure between populations is mostly likely a result of the high frequency of plastid haplotypes P8 and P14 and their derivatives in the southern and northern parts of the species' range, respectively. We identified a southern refugium located in north-central Florida and southern Alabama. The multilocus microsatellite pattern observed is concordant with the plastid model. Plastid haplotypes within group P2 (haplotypes 5–13) are primarily found distributed to the south, though a few are found in more northern populations located in a previously glaciated region. Microsatellite genotypes belonging to group M1 also dominate the southern landscape (genotypes shaded black in fig. 2) in the same locations as haplotypes from group P2 (genotypes shaded dark gray in fig. 1). In addition, these genotypes harbor high genetic diversity, as evident by the long branches within microsatellite group M1 (fig. 2*a*). Given that populations from Florida, South Carolina, and the southern ranges of Alabama and Georgia (fig. 1) cluster together in both data sets, a migration route from the deep south is suggested, possibly out of Florida and the Lower Mississippi Valley. The location of such a southern refugium is posited for several other species (*Acer rubrum* [McLachlan et al. 2005], *Liriodendron tulipifera* [Sewell et al. 1996], *Sagittaria latifolia* [Mylecraine et al. 2004], *Trillium cuneatum* [Gonzales et al. 2008]) and is in agreement with where the hosts of *C. americana*, the red oaks, are presumed to have survived during the LGM along with other temperate hardwood taxa (Delcourt and Delcourt 1993; Jackson and Overpeck 2000).

In addition, we identified a region in the southern Appalachian Mountains, near the southern tip of the Blue Ridge Mountains, as the location of a second, more northern refugium for *C. americana* during the LGM. At that time, the southern Appalachians were believed to have been dominated by boreal forest. Today, populations belonging to neither plastid haplotype group P3 nor the microsatellite groups M2 and M3 are found south of the southern Appalachians but instead are the most common haplotypes/genotypes found in the central and northern ranges of this species (figs. 1, 2). Such a pattern in geographic distribution for these haplotypes and genotypes suggests their persistence in a more northern refugium during the LGM. Their occurrence in the Appalachians today is not reflective of the species' most recent northward migration from the southern Florida refugium following the retreat of the ice margin. Instead, a few relictual populations could have survived in the southern Appalachians during the LGM, at which point they would become geographically and genetically isolated from the southern refugium, resulting in their unique genetic signal that is found across the northern spread of the species today. In the plastid haplotype network (fig. 1*a*), the character that differentiates group P3 from P2 (light gray and dark gray, respectively) is an eight-nucleotide deletion. These eight nucleotides are present in the haplotypes of groups P2 and P1 (black in fig. 1) in addition to the other two species of *Conopholis* and the sister genus *Epifagus* (data not shown). The deletion of this nonrepetitive sequence of nucleotides in these haplotypes represents a strong and unique character that is unlikely to be homoplastic (Kelchner and Wendel 1996; Graham et al. 2000). It is the defining character that supports the separation of populations/individuals belonging to haplotype group P3 (light gray in fig. 1) from those originating from the southern refugium. This finding of a more northern refugium located in the southern Appalachians is consistent with that suggested for other plant and animal taxa (e.g., Church et al. 2003; McLachlan et al. 2005; Walker et al. 2009; Jackson and Austin 2010).

Further evidence for the existence of two separate refugia is the starburst pattern observed from the plastid haplotype network. Such a pattern is an expected signature of a species that has recently expanded from a single geographic source (Avise 2009). In this case, we observe two starburst patterns, one for plastid group 2, shaded dark gray, and another for plastid group 3, shaded light gray (fig. 1*a*). This indicates a recent expansion from two separate geographic sources, where the common and widespread haplotypes 8 (dark gray) and 14 (light gray) are the ancestral conditions from which the other haplotypes were more recently derived and are still rare. The relatively high diversity found across the Appalachians likely represents a secondary contact zone between populations following recolonization between the two previously separated refugia. As populations from the southern refugium migrated northward at the end of the LGM, they would have encroached on the geographic range harboring populations from the northern refugium. This central region of eastern North America along the Appalachians is where we see a mixture of the different haplotype/genotype groups (figs. 1*b*, 2*b*). However, with the retreat of the glaciers, range expansions and recolonization northward would primarily involve populations at the leading edge in this region. Populations from the established northern refugium (light gray groups) are likely to block the range expansion of the related southern refugium populations (leading edge hypothesis; Hewitt 1996; Swenson and Howard 2005). As a result, we expect a poleward decrease in genetic diversity within and among populations (Hewitt 2000; Hampe and Petit 2005). Our study supports such a hypothesis, whereby populations derived from the southern and northern refugia are both found along the Appalachian Mountains, while in the north, populations are primarily derived from the northern refugium (i.e., the leading edge).

Comparing the past and present ecological niche distribution models for *C. americana* (fig. 3), we notice a substantial reduction in availability of suitable habitat for population at the LGM. At the LGM, highly suitable habitats for *C. americana* east of the Mississippi River were located in only two regions (fig. 3*b*). The first is the north-central region of Florida and the southern portion of Mississippi and Louisiana along the Gulf Coast, while a second location with high probability scores is in the Blue Ridge Mountains of the southern Appalachians (circled region in fig. 3*b*). This finding is consistent with the presumed location of where the hosts of *Conopholis* and other temperate deciduous hardwood species (*Populus*, *Quercus*, *Alnus*, *Betula*) are believed to have survived during the most recent glacial cycle (Delcourt and Delcourt 1993; Jackson et al. 2000; Soltis et al. 2006). The molecular data also support this (as dis-

cussed above), resulting in an overall agreement between where the genetic diversity is observed, the genetic signatures of glacial refugia, and where the most likely suitable habitats for populations are located following our LGM distribution modeling. The convergence of ecological niche models at the LGM and molecular evidence from plastid and nuclear sources provides strong support for the existence of midlatitudinal LGM refugia in eastern North America.

The detection and location of a unique and genetically distinct plastid haplotype group (P1, black haplotypes; table 1; fig. 1) provide some clues to the relationships between populations of *C. americana* in eastern North America and the disjunct members in southern Mexico. This infrequent haplotype group (present in only five of 75 populations and eight of 281 individuals) likely represents relictual retention of the ancestral haplotype. A molecular phylogenetic study of *Conopholis* (Rodrigues et al. 2011) that used two of these five sampled populations (SS.05.79 and SS.03.11 only) found that these particular populations are more closely related to disjunct members in southern Mexico than to populations found in eastern North America. Such an east-west split between the eastern United States and eastern Mexico, observed in more than 50 species of plants, is presumed to have occurred during the late Miocene to mid-Pleistocene (Wood 1972; Graham 1999). These particular five populations today still retain the distinctive and ancestral genetic signature of south Mexican populations (see number of steps separating haplotypes shaded black from those shaded dark gray in fig. 1*a*).

Finally, it should be pointed out that a previous study using the holoparasite *Epifagus virginiana*, the monotypic sister genus to *Conopholis* that also exhibits similar intraspecific east-west disjunction in North America, found that the southern and midwestern regions contained higher allelic richness compared to the north and that population differentiation was greatest in the south (Tsai and Manos 2010). The results of our study are similar to that recovered in *Epifagus*. However, unlike their case, where the definition of regions was driven in part by the knowledge of the single host and its location as the ice margin retreated following the LGM, the particular species of red oaks that are the hosts to *C. americana* in eastern North America are unknown. Even though the distribution of oaks during the LGM at approximately 20 kya is well established based on isopollen maps (Jackson et al. 2000), it is difficult to distinguish between the different species of oak based on pollen grains (Bennett 1983). As a result, the records of fossil oak pollen deposits provide an indirect proxy only for the presence of oak communities and not for the particular species of oaks present at any given time. The results of our study on the locations of glacial refugia and the genetic diversity of *C. americana* can be used as a proxy for the location of glacial refugia and the range expansion of those species of red oaks that are the hosts for *Conopholis* in this region. A study focusing on the chloroplast DNA variation in one species of red oak (*Quercus rubra*) in North America found weak phylogeographic structure and no spatial structure of genetic diversity (Magni et al. 2005). One haplotype was present in 75% of the sampled trees and was the most dominant haplotype north of where the ice margin was located at the LGM. The phylogenetic relationship between haplotypes also exhibited a starburst-like pattern with populations in the southern Appalachians, showing more diversity and harboring some rare haplotypes, suggestive of a glacial refugium being located in that region, much the same as for *C. americana*. The range of *C. americana* in eastern North America, however, extends farther south, beyond the distribution of this particular species of red oak (cf. fig. 1 in Magni et al. 2005; Rodrigues et al. 2011). This further supports the notion that *C. americana* parasitizes more than one species of red oaks. Other related *Quercus* species of red oaks whose ranges overlap with that of *C. americana* and go beyond the distribution of *Q. rubra* are *Q. coccineae*, *Q. falcata*, *Q. ilicifolia*, *Q. imbricaria*, *Q. marilandica*, *Q. pagoda*, *Q. palustris*, *Q. phellos*, and *Q. velutina* (Aldrich et al. 2003; Samuelson and Hogan 2003).

Conclusions

In summary, this study utilizes both plastid and nuclear data in addition to paleodistribution modeling and identifies two geographic regions where populations of *Conopholis americana* in eastern North America persisted through the LGM. It provides support for a notion that populations have existed in separate and isolated refugia from which they expanded their range following the retreat of the ice. The recovery of a distinct southern lineage is in agreement with the location of a previously proposed southern glacial refugium spanning across north-central Florida, southern Georgia and Alabama, and the Lower Mississippi Valley. The second lineage is dominant across the present northern range and is hypothesized to have been located in the southern extent of the Blue Ridge Mountains of the southern Appalachians at the LGM. Following the retreat of the glaciers, populations from the more northern refugium were the primary players at the leading edge of the northward migration. As a result, their haplotypes/genotypes are the most prevalent in the north, especially in the previously glaciated regions. In addition, the diversity seen across the southern Appalachian Mountains is congruent with the hypothesis that this is the area where populations derived from the southern and northern refugia come together. Future work in this group should focus on identifying the particular species of red oaks that are the hosts for *C. americana*. If a specific species (or limited set of species) can be ascertained, a similar study can be conducted to determine whether this host(s) also exhibits a comparable (1) LGM history and (2) present-day geographic structure.

Acknowledgments

We thank Robert Brown and Tanya Kenesky (University of Toronto Mississauga Library) for their help using ArcMap to create, view, edit, and analyze geospatial data as well as Alison Colwell (Yosemite National Park) for plant collections and helpful discussions. We also warmly thank Christopher Campbell, Leslie Goertzen, Steven Hill, Peter Kuhlman, Maria Kuzmina, and Jeffrey Morawetz for supplying plant material. Financial support from a Natural Sciences and Engineering Research Council of Canada Discovery grant (326439) and a University of Toronto Connaught New Staff Matching grant to S. Stefanović are gratefully acknowledged.

Literature Cited

Aldrich PR, GR Parker, CH Michler, J Romeo-Severson 2003 Whole-tree silvic identifications and microsatellite genetic structure of a red oak species complex in an Indiana old-growth forest. Can J For Res 33:2228–2237.

Avise JC 2009 Phylogeography: retrospect and prospect. J Biogeogr 36:3–15.

Baird VW, JL Riopel 1986 Life history studies of *Conopholis americana* (Orobanchaceae). Am Midl Nat 116:140–151.

Bennett KD 1983 Postglacial population expansion of forest trees in Norfolk, UK. Nature 303:164–167.

Brito PH, SV Edwards 2009 Multilocus phylogeography and phylogenetics using sequence-based markers. Genetica 135:439–455.

Church SA, AM Kraus, JC Mitchell, DR Church, DR Taylor 2003 Evidence of multiple Pleistocene refugia in the post glacial expansion of the eastern tiger salamander, *Ambystoma tigrinum tigrinum*. Evolution 57:372–383.

Clement M, D Posada, KA Crandall 2000 TCS: a computer program to estimate genetic genealogies. Mol Ecol 9:1657–1659.

Collins WD, CM Bitz, ML Blackmon, GB Bonan, CS Bretherton, JA Carton, P Chang, et al 2004 The community climate system model: CCSM3. J Clim 19:2122–2143.

Corander J, P Waldmann, MJ Sillanpää 2003 Bayesian analysis of genetic differentiation between populations. Genetics 163:367–374.

Davis MB 1981 Quaternary history and the stability of forest communities. Pages 132–153 *in* DC West, HH Shugart, DB Botkin, eds. Forest succession: concepts and application. Springer, New York.

DeBry RW, RG Olmstead 2000 A simulation study of reduced tree-search effort in bootstrap resampling analysis. Syst Biol 49:171–179.

Delcourt HR, PA Delcourt 1993 Paleoclimates, paleovegetation, and paleofloras during the late Quaternary. Pages 71–94 *in* Flora of North America. Oxford University Press, New York.

Dellicour S, P Mardulyn 2014 spads 1.0: a toolbox to perform spatial analyses on DNA sequence data sets. Mol Ecol Resour 14 (3):647–51.

El Mousadik A, RJ Petit 1996 Chloroplast DNA phylogeography of the argan tree of Morocco. Mol Ecol 5:547–555.

Excoffier L, HEL Lischer 2010 Arlequin suite version 3.5: a new series of programs to perform population genetic analyses under Linux and Windows. Mol Ecol Resour 10:564–567.

Felsenstein J 1985 Confidence limits on phylogenies—an approach using bootstrap. Evolution 39:783–791.

Godbout J, JP Jaramillo-Correa, J Beaulieu, J Bosquet 2005 A mitochondrial DNA minisatellite reveals the postglacial history of jack pine (*Pinus banksiana*), a broad-range North American conifer. Mol Ecol 14:3497–3512.

Gonzales E, JL Hamrick, S-M Chang 2008 Identification of glacial refugia in south-eastern North America by phylogeographic analyses of a forest understorey plant, *Trillium cuneatum*. J Biogeogr 35:844–852.

Graham A 1999 The Tertiary history of the northern temperate element in the northern Latin America biota. Am J Bot 86:32–38.

Graham SW, PA Reeves, ACE Burns, RG Olmstead 2000 Microstructural changes in noncoding chloroplast DNA: interpretation, evolution, and utility of indels and inversions in basal angiosperm phylogenetic inference. Int J Plant Sci 161(suppl.):S83–S96.

Hampe A, RJ Petit 2005 Conserving biodiversity under climate change: the rear edge matters. Ecol Lett 8:461–467.

Hewitt GM 1996 Some genetic consequences of ice ages, and their role in divergence and speciation. Biol J Linn Soc 58:247–276.

——— 2000 The genetic legacy of the Quaternary ice ages. Nature 405:907–913.

——— 2004 Genetic consequences of climatic oscillations in the Quaternary. Philos Trans R Soc B 359:183–195.

Hijmans RJ, SE Cameron, JL Parra, PG Jones, A Jarvis 2005 Very high resolution interpolated climate surfaces for global land areas. Int J Climatol 25:1965–1978.

Jackson ND, CC Austin 2010 The combined effects of rivers and refugia generate extreme cryptic fragmentation within the common ground skink (*Scincella lateralis*). Evolution 64:409–428.

Jackson ST, JT Overpeck 2000 Responses of plant populations and communities to environmental changes of the late Quaternary. Paleobiology 26:194–220.

Jackson ST, RR Webb, KH Anderson, JT Overpeck, T Webb III, JW Williams, BCS Hansen 2000 Vegetation and environment in eastern North America during the Last Glacial Maximum. Quat Sci Rev 19:489–508.

Jaramillo-Correa JP, J Beaulieu, J Bosquet 2004 Variation in mitochondrial DNA reveals multiple distant glacial refugia in black spruce (*Piceae mariana*), a transcontinental North American conifer. Mol Ecol 13:2735–2747.

Kelchner SA, JF Wendel 1996 Hairpins create minute inversions in non-coding regions of chloroplast DNA. Curr Genet 30:259–262.

Magni CR, A Ducousso, H Caron, RJ Petit, A Kremer 2005 Chloroplast DNA variation of *Quercus rubra* L. in North America and comparison with other Fagaceae. Mol Ecol 14:513–524.

McLachlan JS, JS Clark, PS Manos 2005 Molecular indicators of tree migration capacity under rapid climate change. Ecology 86: 2088–2098.

Mylecraine KA, JE Kuser, PE Smouse, GL Zimmermann 2004 Geographic allozymes variation in Atlantic white-cedar *Chamaecyparis thyoides* (Cupressaceae). Can J For Res 34:2443–2454.

Petit RJ, J Duminil, S Fineschi, A Hampe, D Salvini, DD Vendramin 2005 Comparative organization of chloroplast, mitochondrial, and nuclear diversity in plant populations. Mol Ecol 14:689–701.

Phillips SJ, RP Anderson, RE Schapire 2006 Maximum entropy of spatial geographic distributions. Ecol Model 190:231–259.

Phillips SJ, M Dudik 2008 Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography 31:161–175.

Pielou EC 1991 After the ice age: the return of life to glaciated North America. University of Chicago Press, Chicago.

Pons O, RJ Petit 1996 Measuring and testing genetic differentiation with ordered versus unordered alleles. Genetics 144:1237–1245.

Rambaut A 2002 Se-Al sequence alignment editor. Version 2.0a11. University of Oxford.

Reboud X, C Zeyl 1994 Organelle inheritance in plants. Heredity 72:132–140.

Rodrigues A, S Shaya, TA Dickinson, S Stefanović 2013 Morphometric analyses and taxonomic revision of the North American holoparasitic genus *Conopholis* (Orobanchaceae). Syst Bot 38:795–804.

Rodrigues AG, AEL Colwell, S Stefanović 2011 Molecular systematics of the parasitic genus *Conopholis* (Orobanchaceae) inferred from plastid and nuclear sequences. Am J Bot 98:896–908.

——— 2012 Development and characterization of polymorphic microsatellite markers for *Conopholis americana* (Orobanchaceae). Am J Bot 99:e4–e6.

Rowe KC, EJ Heske, PW Brown, KN Paige 2004 Surviving the ice: northern refugia and postglacial colonization. Proc Natl Acad Sci USA 101:10355–10359.

Samuelson LJ, ME Hogan 2003 Forest trees: a guide to the southeastern and mid-Atlantic regions of the United States. Von Hoffmann, Upper Saddle River, NJ.

Schaal BA, DA Hayworth, KM Olsen, JT Rauscher, WA Smith 1998 Phylogeographic studies in plants: problems and prospects. Mol Ecol 7:465–474.

Sewell MM, RP Clifford, MW Chase 1996 Intraspecific chloroplast DNA variation and biogeography of North American *Liriodendron* L. (Magnoliaceae). Evolution 50:1147–1154.

Soltis DE, AB Morris, JS McLachlan, PS Manos, PS Soltis 2006 Comparative phylogeography of unglaciated eastern North America. Mol Ecol 15:4261–4293.

Stewart JR, AM Lister 2001 Cryptic northern refugia and the origins of northern biota. Trends Ecol Evol 16:608–613.

Swenson NG, DJ Howard 2005 Clustering of contact zones, hybrid zones, and phylogeographic breaks in North America. Am Nat 166:581–591.

Swofford DL 2002 PAUP*: phylogenetic analysis using parsimony (*and other methods). Version 4.0b10. Sinauer, Sunderland, MA.

Tsai Y-HE, PS Manos 2010 Host density drives the postglacial migration of the tree parasite *Epifagus virginiana*. Proc Natl Acad Sci USA 107:17035–17040.

Walker MJ, AK Stockman, PL Marek, JE Bond 2009 Pleistocene glacial refugia across the Appalachian Mountains and coastal plain in the millipede genus *Narceus*: evidence from population genetics, phylogeographic, and paleoclimatic data. BMC Evol Biol 9:25.

Waltari E, RJ Hijmans, AT Peterson, AS Nyari, SL Perkins, RP Guralnick 2007 Locating Pleistocene refugia: comparing phylogeographic and ecological niche model predictions. PLoS ONE 7:e563–e573.

Warren DL, RE Glor, M Turelli 2010 ENMTools: a toolbox for comparative studies of environmental niche models. Ecography 33:607–611.

Williams M, D Dunkerley, P De Decker, P Kershaw, J Chappel 1998 Quaternary environments. Oxford University Press, New York.

Wolfe KH, WH Li, PM Sharp 1987 Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc Natl Acad Sci USA 84:9054–9058.

Wood CEJ 1972 Morphology and phytogeography: the classical approach to the study of disjunctions. Ann Mo Bot Gard 59:107–124.