



# Phylogeographic analysis of North American populations of the parasitic herbaceous plant *Monotropa hypopitys* L. reveals a complex history of range expansion from multiple late glacial refugia

Gemma E. Beatty and Jim Provan\*

School of Biological Sciences, Queen's  
University Belfast, 97 Lisburn Road, Belfast  
BT9 7BL, Northern Ireland, UK

## ABSTRACT

**Aim** We carried out a phylogeographic study across the range of the herbaceous plant species *Monotropa hypopitys* L. in North America to determine whether its current disjunct distribution is due to recolonization from separate eastern and western refugia after the Last Glacial Maximum (LGM).

**Location** North America: Pacific Northwest and north-eastern USA/south-eastern Canada.

**Methods** Palaeodistribution modelling was carried out to determine suitable climatic regions for *M. hypopitys* at the LGM. We analysed between 155 and 176 individuals from 39 locations spanning the species' entire range in North America. Sequence data were obtained for the chloroplast *rps2* gene ( $n = 168$ ) and for the nuclear ITS region ( $n = 158$ ). Individuals were also genotyped for eight microsatellite loci ( $n = 176$ ). Interpolation of diversity values was used to visualize the range-wide distribution of genetic diversity for each of the three marker classes. Minimum spanning networks were constructed showing the relationships between the *rps2* and ITS haplotypes, and the geographical distributions of these haplotypes were plotted. The numbers of genetic clusters based on the microsatellite data were estimated using Bayesian clustering approaches.

**Results** The palaeodistribution modelling indicated suitable climate envelopes for *M. hypopitys* at the LGM in both the Pacific Northwest and south-eastern USA. High levels of genetic diversity and endemic haplotypes were found in Oregon, the Alexander Archipelago, Wisconsin, and in the south-eastern part of the species' distribution range.

**Main conclusions** Our results suggest a complex recolonization history for *M. hypopitys* in North America, involving persistence in separate eastern and western refugia. A generally high degree of congruence between the different marker classes analysed indicated the presence of multiple refugia, with at least two refugia in each area. In the west, putative refugia were identified in Oregon and the Alexander Archipelago, whereas eastern refugia may have been located in the southern part of the species' current distribution, as well as in the 'Driftless Area'. These findings are in contrast to a previous study on the related species *Orthilia secunda*, which has a similar disjunct distribution to *M. hypopitys*, but which appears to have recolonized solely from western refugia.

## Keywords

Climate change, Cordilleran ice sheet, glaciations, Last Glacial Maximum, Laurentide ice sheet, mycoheterotroph, phylogeography.

\*Correspondence: Jim Provan, School of Biological Sciences, Queen's University Belfast, 97 Lisburn Road, Belfast BT9 7BL, Northern Ireland, UK.  
E-mail: J.Provan@qub.ac.uk

## INTRODUCTION

Climate change is not solely a contemporary issue: the Pleistocene epoch (c. 2.5 Ma–10 ka) was characterized by cyclical fluctuations of long, cool glacial periods interspersed with short, warm interglacials. The most recent major glacial period was at its most extreme during the Last Glacial Maximum (LGM; c. 21–18 ka), when large areas of the northern latitudes were covered by ice or permafrost. This resulted in an overall reduction in habitat for many species, which had to track suitable habitat south or face extirpation (Hewitt, 1996). At the height of the LGM, many species ranges, particularly those that currently occupy temperate habitats, were at their smallest. These species generally persisted in refugia south of the limits of the ice, but recent studies have suggested that for some species, more complex scenarios, including the presence of cryptic northern refugia, would best explain their glacial histories (Stewart & Lister, 2001; Bennett & Provan, 2008; Provan & Bennett, 2008; Stewart *et al.*, 2010). This combination of vicariance in separate refugia, followed by recolonization from some or all of these regions during the warmer Holocene (c. 10 ka–present), has been the primary driving force in shaping the present-day ranges of many species (Hewitt, 2003). Knowledge of how species have responded to previous changes in the Earth's climate is becoming increasingly important, as it may provide insights into future changes resulting from the current period of global warming (Hu *et al.*, 2009; Harrison & Sánchez Goñi, 2010).

Phylogeographic and palaeontological studies on European plants and animals have indicated that the southern peninsulas of Iberia, Italy and the Balkans served as refugial areas for the majority of temperate species during the LGM (Huntley & Birks, 1983; Bennett *et al.*, 1991; Taberlet *et al.*, 1998; Hewitt, 1999). In North America, however, the scenario was somewhat different, with multiple refugia identified both north and south of the Laurentide and Cordilleran ice sheets that dominated the northern part of the continent at the LGM (reviewed in Beatty & Provan, 2010). Palynological analyses have also indicated that the succession of vegetation in North America since the onset of the LGM was characterized by a 'quick-slow-quick-slow' pattern of change (Williams *et al.*, 2004). The initial rapid change due to reduction of available habitat as a result of the advance of the ice sheets was followed by a period of relative stability during the glacial period. A further rapid reorganization of the vegetation followed the deglaciation (c. 16–8 ka), with subsequent stability during the remainder of the Holocene until the anthropogenically induced changes of the last c. 500 years. As a result of this shifting dynamic regime, current-day species assemblages and distributions are often very different from those at various stages during the late Pleistocene (Thompson & Anderson, 2000; Williams *et al.*, 2004; Strong & Hills, 2005). Such changes, coupled with the fact that many herbaceous species leave only transient pollen in the soil, means that in many cases phylogeographic approaches can be particularly infor-

mative for reconstructing the past distributions of species (Hu *et al.*, 2009).

*Monotropa hypopitys* L. (syn. *M. hypopithys*: Monotropaceae), commonly referred to as pinesap or yellow bird's nest, is a distinctive waxy herbaceous perennial plant. The species is usually found in shaded woodlands in temperate regions of the Northern Hemisphere across Europe, Asia and North America, although throughout its range it is generally found scattered and at low abundance (Beatty & Provan, 2011). In North America, the species exhibits an east–west disjunct distribution (Fig. 1). In the east, it is found as far north as southern Quebec and New Brunswick, and the southern limit of its main distribution is in Tennessee and North Carolina, although it is very occasionally found further south as far as northern Florida, where the species is classified as endangered. Its distribution in central and western North America is much more fragmented: the species is found in sporadic pockets in the southern Great Plains, but has never been recorded in North and South Dakota. In the western part of its range, it can be found as far north as Alaska, extending through the Pacific Northwest, and can occasionally be found as far south as Arizona and New Mexico, usually above 1800 m where suitable forest occurs. The species is absent from Nevada and Utah. In some parts of its range, it occurs sympatrically with the only other member of the genus found in North America, *M. uniflora*, but there are no records of hybridization between the two species. *Monotropa hypopitys* primarily exists as an underground network of rhizomes for most of the year until the beginning of July, when aerial flowering shoots appear and remain until late September. The species is non-photosynthetic, and instead gains its nutrition epiparasitically from tree roots via a mycorrhizal association with fungi from the genus *Tricholoma* (Leake *et al.*, 2004). The main host species include beech (*Fagus* spp.), hazel (*Corylus* spp.) and pines (*Pinus* spp.).

The aim of the current study was to investigate the glacial history of *M. hypopitys* in North America, primarily to determine if its contemporary disjunct distribution is due to persistence in separate eastern and western refugia or, alternatively, due to vicariance of a once continuous range. A previous study on another member of the Monotropoideae, *Orthilia secunda*, which has a similar disjunct distribution, found that its current distribution is a result of recolonization from refugia solely in the western part of its range, followed by vicariance as the central mesic boreal forests were replaced by the grasslands of the Great Plains (Beatty & Provan, 2010). We used a combination of palaeodistribution modelling (Nogués-Bravo, 2009) and phylogeographic analysis of multiple genetic markers (sequence data from both the chloroplast *rps2* gene and the nuclear internal transcribed spacer (ITS) region, and eight nuclear microsatellites) to reconstruct the distributional history of *M. hypopitys* in North America from the LGM to the present. Multiple genetic markers were used because the limitations of using a single marker system to investigate phylogeographic history have been well documented (e.g. Bermingham & Moritz, 1998; Schaal *et al.*, 1998).



**Table 1** Samples of *Monotropia hypopitys* analysed in the present study.

Country	State/Province	Location	Code	Lat/Long (N/W)	$N_{rps2}$	$N_{ITS}$	$N_{mic}$	Collector
Canada	British Columbia	Manning Park	BCMP	49.064/120.782	8	8	8	Ron Long
		Nootka	WTU214041	49.333/126.500	1	1	1	University of Washington Herbarium
	New Brunswick	Boucouteche	NBBO	46.467/64.717	8	8	8	Michael Crowell
	Ontario	Michipicoten River	ONMR	47.917/84.837	2	2	2	Michael Oldham
	Quebec	Mont-Tremblant	QCMT	46.196/74.623	8	8	8	Andre Lapointe
	Saskatchewan	Cypress Hills	SKCH	49.661/109.511	5	4	5	Glen Lee
USA	Alaska	Wrangell Island	UAFM8782	56.330/132.400	2	–	2	University of Alaska at Fairbanks
		Kadin Island	UAFM11086	56.530/132.450	5	5	5	University of Alaska at Fairbanks
		Baranof Island	UAFM43121	57.050/135.330	7	6	7	University of Alaska at Fairbanks
		Chicagof Island	UAFM44124	58.000/135.730	5	4	5	University of Alaska at Fairbanks
		Coast Mountains	UAFM44125	58.530/134.800	4	4	4	University of Alaska at Fairbanks
		Lynn Canal	UAFM27744	59.260/135.160	4	4	4	University of Alaska at Fairbanks
		Malaspina Coastal Plain	UAFM60111	60.190/142.660	1	–	1	University of Alaska at Fairbanks
		Softuk Lagoon	UAFM59072	60.210/144.650	2	2	2	University of Alaska at Fairbanks
		Chugach Mountains	UAFM82664	61.000/149.080	2	2	2	University of Alaska at Fairbanks
		Prince of Wales Island	ALPW	55.688/132.522	2	2	2	Brad Kriekhaus
	Hoonah		ALHO	58.111/135.444	1	1	1	Brad Kriekhaus
		Maine	Saltwater Cove	MESC	43.873/69.594	3	3	3
		Edgecomb	MEED	43.947/69.632	2	2	2	Jean Howe
	Massachusetts	Long Mountain	MALM	42.308/72.472	8	8	8	Karen Searcy
	Michigan	Pine Point	MIPP	45.400/84.860	7	7	8	Melanie Gunn
	Minnesota	Superior National Forest	MNSF	47.971/90.991	8	8	8	Lynden Gerdes
	Montana	Madison Range	MTMR	45.350/111.200	8	8	8	Peter Lesica
		Colfax County	NMCC	36.271/105.146	1	1	1	Ken Heil
	New Mexico	Cibola National Forest	NMCF	33.807/107.446	1	1	1	Ken Heil
		New York	Morgan Hill State Forest	NYMH	42.800/76.100	2	2	2
	Ohio	Strouds Run State Park	OHSR	39.348/82.034	2	2	2	Phil Cantino
		Glen Ebon Road	OHGE	39.402/82.188	4	2	8	Phil Cantino
	Oregon	Hood River	ORHR	45.414/121.790	8	7	8	Diana Jolles
	Rhode Island	Providence	RIPR	41.900/74.417	5	5	5	Keith Killingbeck
	South Carolina	Walhalla	WV116064	34.772/83.062	4	4	4	West Virginia University Herbarium
		Caesar's Head State Park	SCCH	35.117/82.617	8	7	8	Tim Spyra
		Tennessee	Tudor Hollow	WV116065	35.724/83.494	4	3	4
	Vermont	Greensboro	VTGR	44.593/72.303	8	8	8	Al Steiger
		Guernsey Brook	VTGB	44.327/72.414	5	5	8	Brett Engstrom
	Washington	Cedar Butte	WTU351539	47.434/121.742	1	1	1	University of Washington Herbarium
	West Virginia	Pleasant's Creek	WV117890	39.246/80.034	2	2	2	West Virginia University Herbarium
		Waiteville	WV077446	37.474/80.422	3	–	3	West Virginia University Herbarium
Wisconsin	Chaffee Creek	WICC	43.950/89.362	7	7	7	Neil Harriman	
Total					168	155	176	

$N_{rps2}$ , number of samples sequenced for the chloroplast *rps2* gene;  $N_{ITS}$ , number of samples sequenced for the nuclear ITS region;  $N_{mic}$ , number of samples genotyped for eight microsatellite loci.

5'-CTTAAACAGCTTAACAAAAGAGATG-3'). Polymerase chain reaction (PCR) was carried out on a MWG (Ebersberg, Germany) Primus thermal cycler using the following parameters: initial denaturation at 94 °C for 3 min followed by 35 cycles of denaturation at 94 °C for 1 min, annealing at 58 °C for 1 min, extension at 72 °C for 1 min and a final extension at 72 °C for 5 min. PCR was carried out in a total volume of

20 µL, containing 200 ng genomic DNA, 20 pmol of each primer, 1 × PCR reaction buffer, 200 µM each dNTP, 2.5 mM MgCl<sub>2</sub> and 0.5 U GoTaq Flexi DNA polymerase (Promega, Madison, WI, USA). 5 µL PCR product was resolved on 1.5% agarose gels and visualized by ethidium bromide staining, and the remaining 15 µL was sequenced in both directions using the BigDye sequencing kit v. 3.1 (Applied Biosystems, Foster

City, CA, USA) and run on an AB 3730XL DNA analyser (Life Technologies, Carlsbad, CA, USA).

### Nuclear ITS sequencing

In total, 155 samples were sequenced for a section of the nuclear ITS region. Primers were designed from GenBank sequence accession number AF384126 (Bidartondo & Bruns, 2001): MH-ITS-F 5'-GGTTGGCCTACCCTTTATTTT-3' and MH-ITS-R 5'-GAAGTAATCCAATCATAACTGACA-3'. Herbarium samples were amplified in two overlapping sections using MH-ITS-F and an internal reverse primer (MH-ITS-R(1), 5'-CGGGATTCTGCAATTCACAC-3'), and MH-ITS-R and an internal forward primer (MH-ITS-F(1), 5'-GTGTG-AATTGCAGAAATCCCG-3'). PCR and sequencing were carried out as described above, except that an annealing temperature of 55 °C was used.

### Microsatellite genotyping

In total, 176 individuals were genotyped for five *M. hypopitys* microsatellite loci previously described in Klooster *et al.* (2009) – Mono02, Mono15, Mono20, Mono21 and Mono22. Three additional loci developed for this study using the ISSR-cloning technique outlined in Provan & Wilson (2007) were also used (Table 2). Forward primers were modified by the addition of a 19 bp M13 tail (5'-CACGACGTTGTAAAACGAC-3') and reverse primers were modified by the addition of a 7 bp tail (5'-GTGTCTT-3'). PCR was carried out in a total volume of 10 µL containing 100 ng genomic DNA, 10 pmol of dye-labelled M13 primer (6-FAM or HEX), 1 pmol of tailed forward primer, 10 pmol reverse primer, 1 × PCR reaction buffer, 200 µM each dNTP, 2.5 mM MgCl<sub>2</sub> and 0.25 U GoTaq Flexi DNA polymerase (Promega). PCR was carried out on a MWG Primus thermal cycler using the conditions described in Klooster *et al.* (2009) and genotyping was carried out on an AB3730XL capillary genotyping system (Applied Biosystems). Allele sizes were scored in GENEMAPPER v. 4.1 (Applied Biosystems) using ROX-500 size standards and were checked by comparison with previously sized control samples.

### Data analysis

DNA sequences for the chloroplast *rps2* and nuclear ITS regions were aligned in BioEDIT v. 7.0.9.0 (Hall, 1999). Length variation at two mononucleotide repeat regions was removed

from the *rps2* alignment, as well from one repeat in the ITS alignment, as the bidirectional mutation model operating at such regions can give rise to homoplasy (Provan *et al.*, 2001). These alignments were used to construct statistical parsimony networks using the TCS software package v. 1.2.1 (Clement *et al.*, 2000). Reticulations in the networks were broken following the rules described in Pfenninger & Posada (2002).

Tests for linkage disequilibrium between pairs of microsatellite loci in each population were carried out in the program FSTAT v. 2.9.3.2 (Goudet, 2001). Levels of genetic diversity were calculated for populations with a sample size of  $n \geq 4$ . Gene diversity ( $H$ ) based on chloroplast *rps2* and nuclear ITS haplotype frequencies and expected heterozygosity ( $H_E$ ) based on nuclear microsatellite allele frequencies were calculated using the ARLEQUIN software package v. 3.01 (Excoffier *et al.*, 2005). To visualize range-wide gradients in genetic diversity for both chloroplast and nuclear markers, data interpolation was carried out using kriging in ArcGIS v. 3.2 (ESRI, Redlands, CA, USA) as described in Keller *et al.* (2010).

The software packages BAPS v. 3.2; Corander *et al.*, 2003) and STRUCTURE v. 2.2 (Pritchard *et al.*, 2000) were used to identify clusters of genetically similar populations based on the nuclear microsatellite data. Both programs use a Bayesian approach to cluster populations, but differ slightly in their approach to estimating the number of clusters ( $K$ ): BAPS uses a greedy stochastic optimization algorithm to determine  $K$ , whilst STRUCTURE derives posterior probability distributions from separate Markov chain Monte Carlo (MCMC) chains with predefined values for  $K$ . BAPS generally tends to recover more clusters than STRUCTURE, and it has been suggested that both approaches should be used, particularly where levels of differentiation between populations may be low (Latch *et al.*, 2006). For the BAPS analysis, ten replicates were run for all possible values of the maximum number of clusters ( $K$ ) up to  $K = 39$ , the number of populations, with a burn-in period of 10,000 iterations followed by 50,000 iterations. A neighbour-joining tree was constructed based on the estimated pairwise Kullback–Leibler divergences between the clusters using the NEIGHBOR program in the PHYLIP software package v. 3.57c (Felsenstein, 1995) to depict genetic relationships between clusters. For the STRUCTURE analysis, five independent runs were carried out for all values of  $K$ , the number of clusters, between 2 and 20. The program was run each time using 50,000 burn-in iterations followed by 500,000 iterations, and the most likely value of  $K$  was determined using the  $\Delta K$  statistic of Evanno *et al.* (2005).

**Table 2** *Monotropa hypopitys* microsatellite primers developed for this study.

Locus	Repeat	Primers	Size range
MHNSR108	(GA) <sub>8</sub>	ACATTTGGGAAAATGGGAGA TTCAATGGCACGTCTTACACA	122–214 bp
MHNSR119	Complex (GA)	GGAAGTTTCTCCATCCAGGTT AGCAATCAAACACAGGACCA	142–180 bp
MHNSR135	(AG) <sub>8</sub>	CGGTTTCAGGAAACAAAACC TTGTCCGGGAATTCTCTCTC	132–150 bp

In an attempt to clarify patterns of possible gene flow between western and eastern populations, an analysis under an isolation-with-migration scenario was carried out using the IMA software package (Hey & Nielsen, 2004). The two New Mexican populations, which were represented by a single individual each, were omitted from this analysis due to their central location. The analysis was run based on the chloroplast *rps2* data, the nuclear ITS data, and six of the eight microsatellites. Loci Mono22 and MHNSR119 were removed from this analysis because they are compound repeats of different length motifs, and thus the relative contribution of each motif to the stepwise mutation process was unknown. Ten MCMC searches were run for  $10^7$  iterations, with the first  $10^6$  discarded as burn-in, and trees were sampled every 100 steps. The HKY model of sequence evolution was used for the chloroplast *rps2* and nuclear ITS, and the single stepwise mutation (SSM) model was used for the microsatellite loci.

## RESULTS

### Palaeodistribution modelling of *M. hypopitys* at the LGM

The ENM based on current climate data was a largely accurate description of the present-day distribution of *M. hypopitys* in North America (Fig. 2a). Modelling based on reconstructed climate at the LGM identified suitable regions for the persistence of *M. hypopitys* in the Pacific Northwest and in south-eastern USA (Figs 2b & 2c). Both the CCSM and MIROC models indicated suitable habitat along the west coast from the Alexander Archipelago through coastal British Columbia to Washington and Oregon, an area largely similar to the westernmost part of the current-day distribution but excluding the area around the northern Rocky Mountains. A more extensive putative refugial region in the south-east encompassed eastern Texas, through the Gulf Coast to eastern North and South Carolina, and was largely situated to the south of the present-day range, with the exception of an area of overlap in North and South Carolina. The MIROC LGM model also indicated additional suitable refugial areas in the south-west of the continent stretching from northern Mexico, through Texas and New Mexico, to eastern California and western Nevada. Based on the MAXENT clamping function, no evidence was found under either model for areas of non-analogous climate (i.e. combinations of environmental variables that are not found at present) at the LGM, which could possibly have biased the modelled locations of putative refugia.

### Chloroplast *rps2* sequencing

Removal of mononucleotide repeat length variation from the *rps2* alignment of the 168 sequences resulted in an overall alignment length of 320 bp and 42 distinct haplotypes (GenBank accession numbers JF440990–JF441031). Twenty-two individuals possessed unique haplotypes (small black circles; Fig. 3). The two main groups of haplotypes (blue/

yellow/green versus red/pink) corresponded largely to the western and eastern areas of the species' distribution, although the most common blue haplotype was also found in a few populations in the eastern part of the range (Fig. 3).

Gene diversity ( $H$ ) values calculated for populations with a sample size of  $n \geq 4$  ranged from 0.000 (many populations) to 1.00 [SCWH (South Carolina) and OHGR (Ohio); see Appendix S1 in the Supporting Information]. Interpolation of gene diversity values across the range of *M. hypopitys* identified the areas with the highest diversity as Oregon and coastal British Columbia in the west, and around the southern part of the species' range in eastern North America (Fig. 4a). Population-specific ('private') haplotypes are listed in Appendix S2.

### Nuclear ITS sequencing

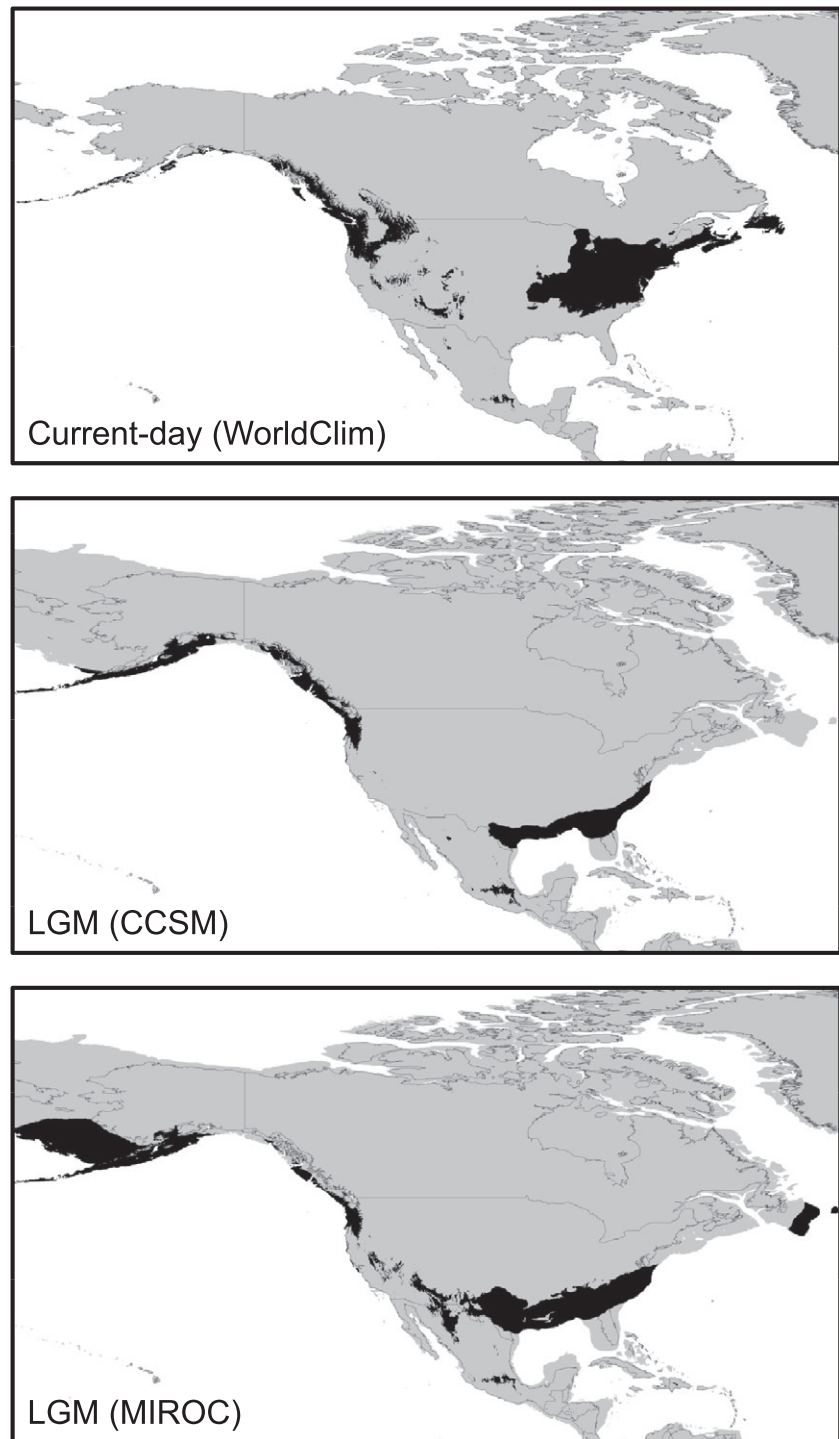
After removal of length polymorphism at a mononucleotide repeat region, the nuclear ITS alignment of 155 sequences resulted in an overall alignment length of 507 bp and 22 distinct haplotypes (GenBank accession numbers JF441032–JF441053). Twelve individuals possessed unique haplotypes (small black circles; Fig. 5). The geographical distribution of haplotypes was very similar to that of the chloroplast *rps2* haplotypes, with one group of haplotypes found predominantly in the west (blue and yellow) and another group of haplotypes (red and pink) found exclusively in the east (Fig. 5). Again, some of the northern populations in the eastern part of the range contained the common blue haplotype.

Gene diversity ( $H$ ) values calculated for populations with a sample size of  $n \geq 4$  ranged from 0.000 (many populations) to 0.810 [WICC (Wisconsin); Appendix S1]. Interpolation of gene diversity values across the range of *M. hypopitys* identified the areas with the highest diversity as Oregon in the west, and around the south-western and north-western parts of the species' range in eastern North America (Fig. 4b). Private haplotypes are listed in Appendix S2.

### Nuclear microsatellite genotyping

No significant linkage disequilibrium was detected between pairs of loci after sequential Bonferroni correction. Between 7 and 16 alleles were detected at the eight microsatellite loci studied (mean = 11.25) and levels of expected heterozygosity ( $H_E$ ) calculated for populations with a sample size of  $n \geq 4$  ranged from 0.156 [SKCH (Saskatchewan)] to 0.658 [ORHR (Oregon); Appendix S1]. Interpolation of expected heterozygosity values across the range of *M. hypopitys* identified the Pacific Northwest and the south-eastern part of the species' range as the main centres of genetic diversity, although the largely homogeneous distribution of microsatellite genetic diversity means that this is not as pronounced as in the equivalent interpolations for the chloroplast *rps2* and nuclear ITS data (Fig. 4c).

The BAPS analysis identified  $K = 14$  as the most likely number of genetic clusters. Genetically similar clusters tended



**Figure 2** Ecological niche modelling of suitable climate envelopes for *Monotropa hypopitys* based on current climate (top) and reconstructed climate at the Last Glacial Maximum based on the Community Climate System Model (CCSM) (middle) and Model for Interdisciplinary Research on Climate (MIROC) (bottom) models. The cumulative presence threshold based on the sensitivity-specificity sum maximization approach was 31 in both cases.

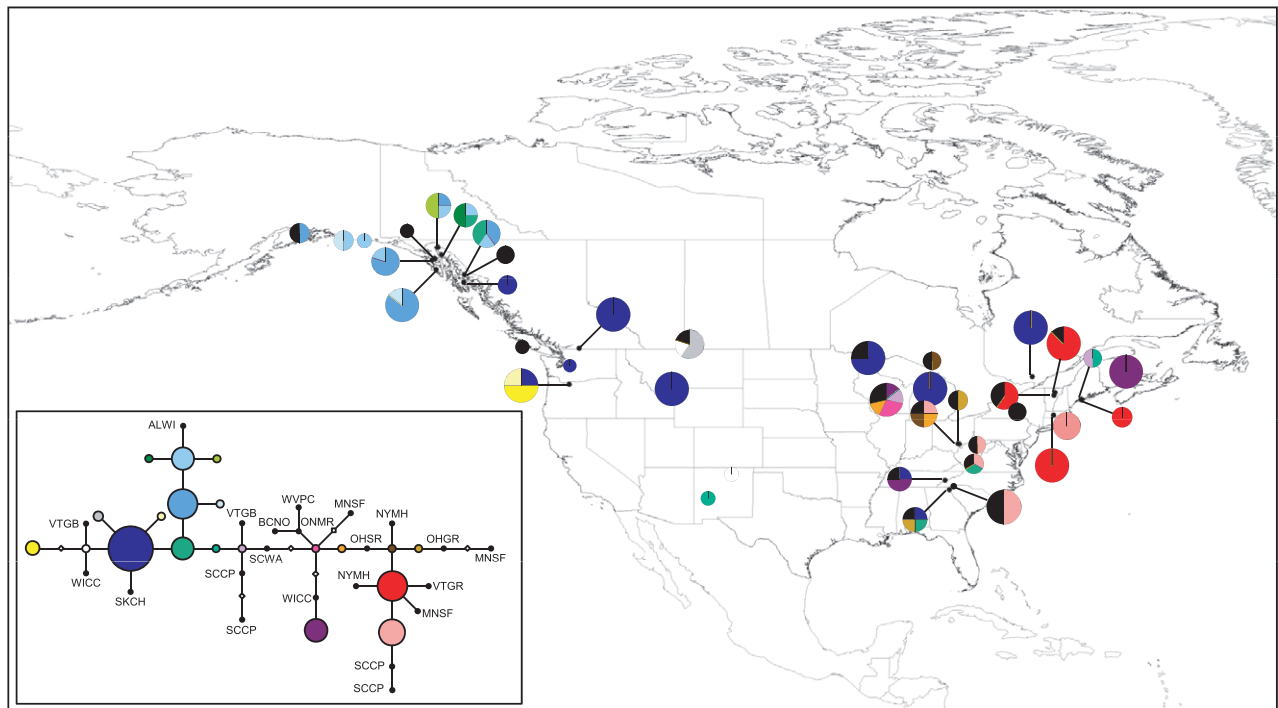
to be co-distributed geographically (Fig. 6). The STRUCTURE analysis identified  $K = 3$  as the most likely number of genetic clusters. These were also geographically localized, with individuals in the west predominantly assigned to the black cluster, and individuals in the east being predominantly assigned to one of two clusters, which exhibited a general north (white) and south (grey) distribution (Fig. 7).

The IMA analysis based on the combined chloroplast *rps2*, nuclear ITS and microsatellite data gave no significant

indication of predominantly unidirectional gene flow from west to east (0.652; 95% confidence intervals 0.135–1.275) or vice versa (1.075; 95% confidence intervals 0.345–2.135).

## DISCUSSION

The findings of the current study suggest a complex recolonization history for *Monotropa hypopitys* in North America, involving persistence in separate eastern and western refugia,



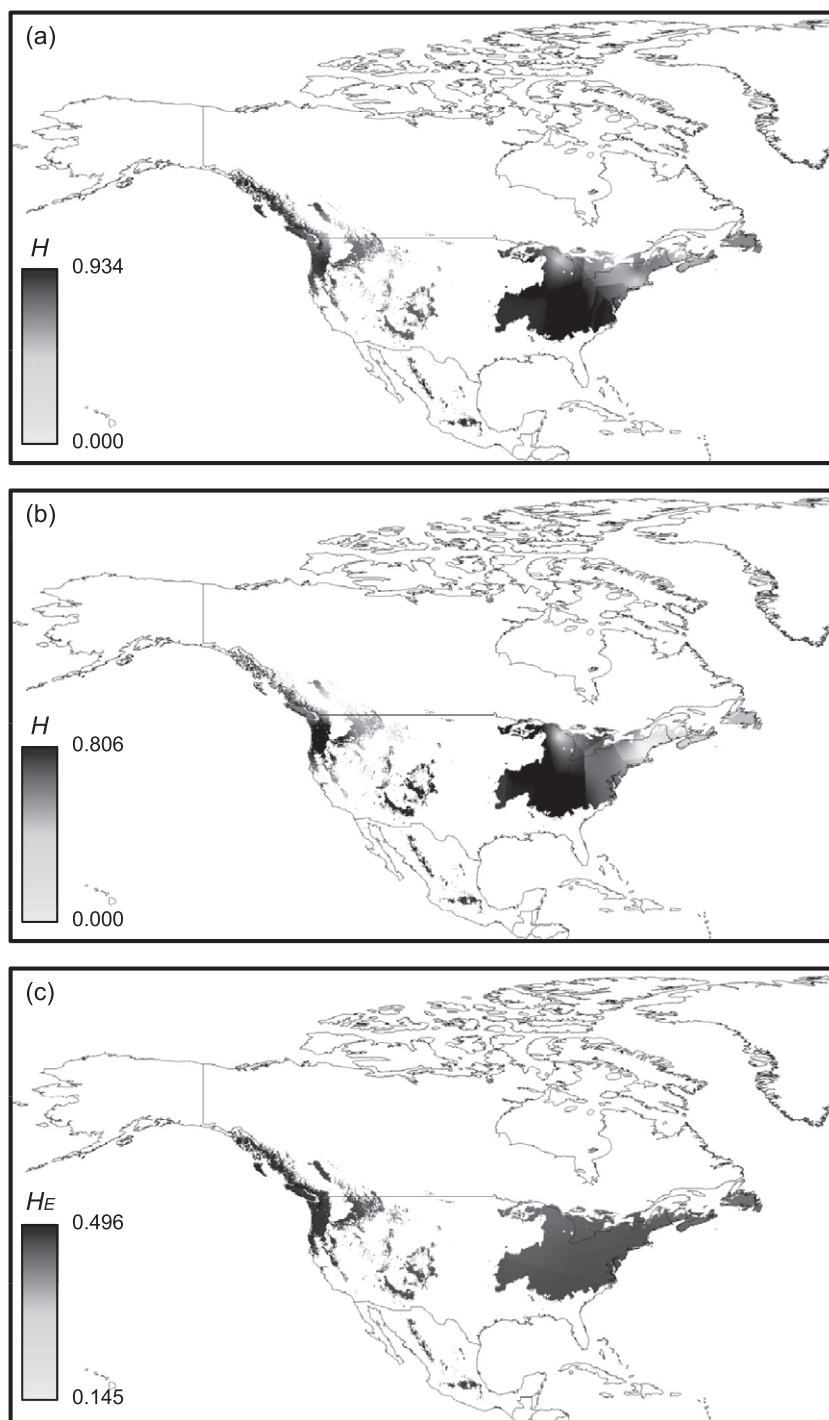
**Figure 3** Distribution of chloroplast *rps2* haplotypes of *Monotropa hypopitys*. Pie chart sizes are approximately proportional to sample size, with the smallest circles representing  $n = 1$  and the largest representing  $n = 8$ . The inset shows the phylogenetic relationships between the 42 haplotypes. Open diamonds represent missing haplotypes and small black circles represent unique haplotypes, i.e. those found in a single individual. The population of origin of each unique haplotype is indicated.

and possible long-distance dispersal and subsequent vicariance. A generally high degree of congruence between the different marker classes analysed indicated the presence of multiple refugia, with at least two western and two eastern refugia, based on levels of diversity and the occurrence of endemic haplotypes/genotypes (those found only in a single area or group of geographically proximal populations; Provan & Bennett, 2008). Previous studies on North American plant species with transcontinental distributions have usually uncovered putative refugia on both sides of the continent (Jaramillo-Correa *et al.*, 2004; Alsos *et al.*, 2005; Godbout *et al.*, 2005; Keller *et al.*, 2010; de Lafontaine *et al.*, 2010). To date, however, we are only aware of a single study on a plant that exhibits an east–west disjunct distribution, namely *Orthilia secunda* (Beatty & Provan, 2010), which, contrary to the findings of the present study, suggested recolonization from solely western refugia.

### Western refugia

The highest levels of diversity along the western range of *M. hypopitys* for all three marker classes (chloroplast *rps2*, nuclear ITS and nuclear microsatellites) were centred around the Oregon population (ORHR), which also displayed a high percentage of private *rps2* and ITS haplotypes. Palynological analyses have indicated that pine forests, one of the main host genera for *M. hypopitys* (Williams *et al.*, 2004), persisted in the

region at the LGM. Taken alone, the ITS data would appear to suggest the classic scenario of ‘southern richness versus northern purity’ (Hewitt, 2000), with Oregon being the southern refugium which acted as the source from which the northern populations along coastal British Columbia and the Alexander Archipelago recolonized. Inspection of the distribution of chloroplast *rps2* haplotypes and the BAPS analysis of the nuclear microsatellite data, however, indicate that many of the populations in the Alexander Archipelago are genetically distinct from the Oregon population, suggesting the likely presence of a second, northerly refugium in the Pacific Northwest. The *rps2* sequence data revealed the occurrence of several related endemic haplotypes in populations exhibiting particularly high diversity in the Alexander Archipelago area, whilst the BAPS analysis indicated the presence of a second genetic cluster (depicted in light blue; Fig. 6) largely confined to the same area. The coastline of the Pacific Northwest is thought to have remained ice-free during the LGM (Warner *et al.*, 1982; Pielou, 1991) and biogeographic, fossil and phylogeographic evidence have suggested that the area served as a refugium for several plant and animal species (reviewed in Soltis *et al.*, 1997; Brunfield *et al.*, 2001). The Alexander Archipelago in particular has been highlighted previously as a potential refugial area in palaeontological and phylogeographic studies (Heaton *et al.*, 1996; Cook *et al.*, 2001). The higher resolution afforded by the chloroplast *rps2* gene, which is suspected to be under relaxed functional constraint in

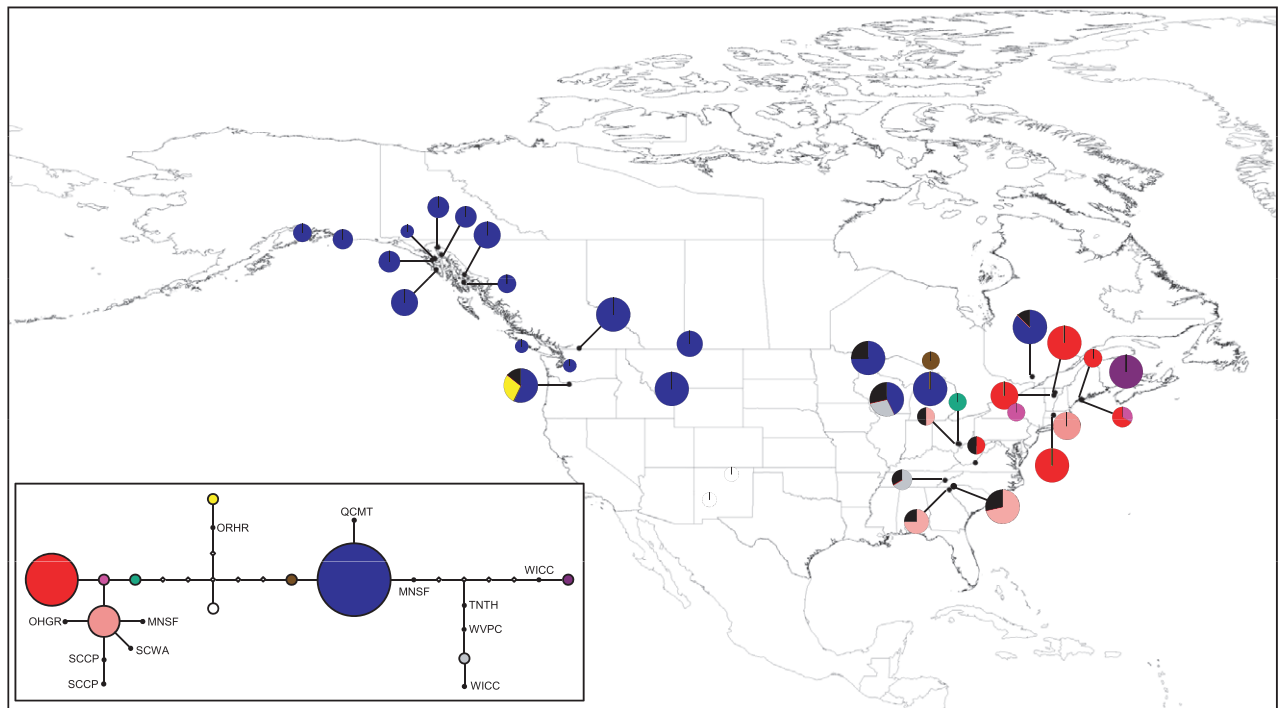


**Figure 4** Geographic gradients of genetic diversity in *Monotropa hypopitys* based on (a) chloroplast haplotype gene diversity ( $H$ ), (b) nuclear ITS gene diversity, and (c) nuclear microsatellite expected heterozygosity ( $H_E$ ). Areas of highest diversity are depicted by the darkest shading.

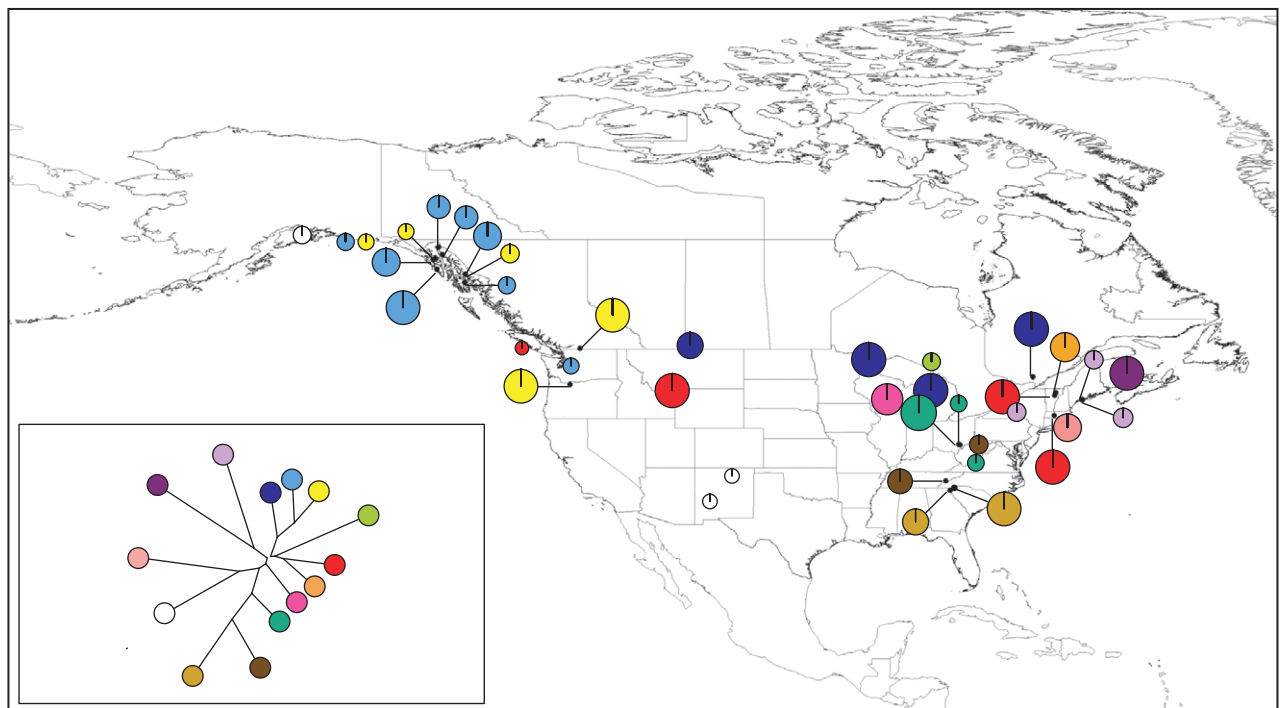
*M. hypopitys* due to loss of photosynthesis, and the BAPS analysis of the nuclear microsatellite data clarifies the likely existence of this northern refugium, as well as highlighting the pitfalls of phylogeographic inference based on a single marker. A scenario of 'leading edge' recolonization of coastal British Columbia and the Alexander Archipelago from a southern refugium, as highlighted in the region previously by Cwynar & MacDonald (1987) and Soltis *et al.* (1997), would lead to changes in allele and haplotype frequencies from south to

north. The observed endemic chloroplast *rps2* haplotypes found in the northern populations but not elsewhere, however, are characteristic of glacial refugia (Provan & Bennett, 2008). Finally, the BAPS analysis also suggested a western source for the northern Rockies population in British Columbia (BCMP), rather than recolonization from the east.

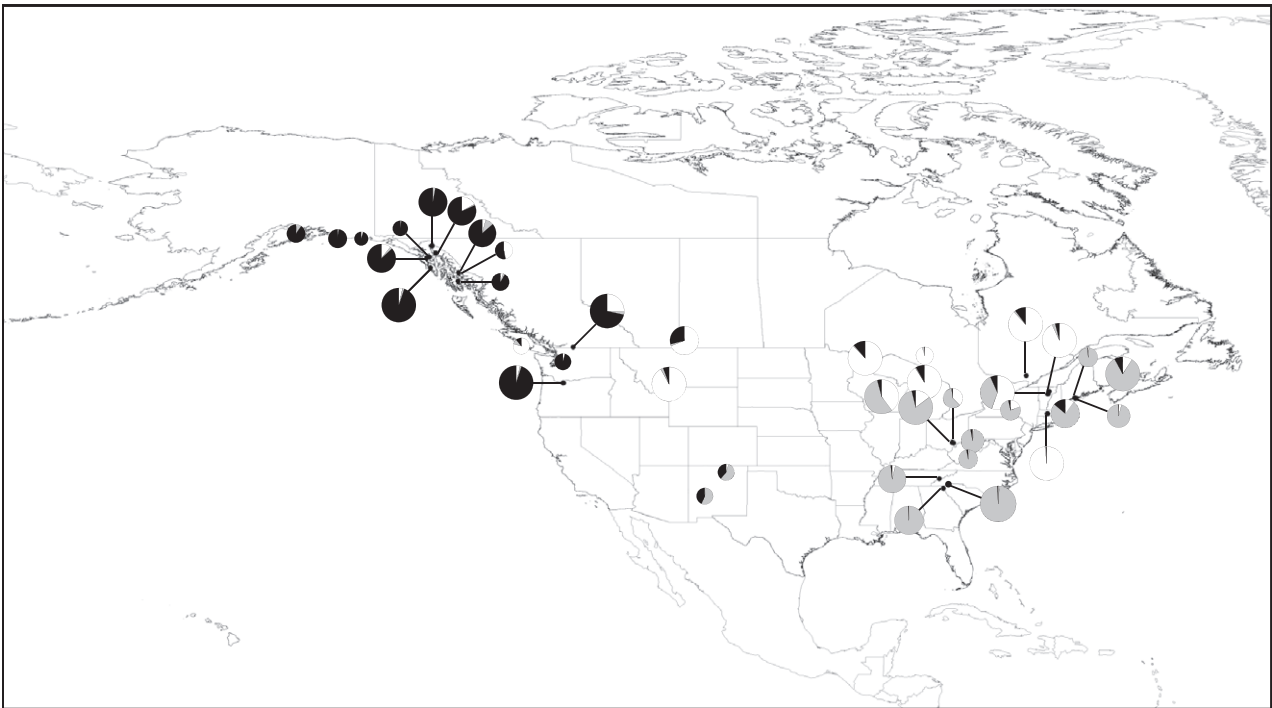
The ENMs based on both reconstructed climate data sets also indicated the presence of suitable habitat for *M. hypopitys* in Oregon and surrounding areas, as well as in the Alexander



**Figure 5** Distribution of nuclear ITS haplotypes of *Monotropa hypopitys*. Pie chart sizes are approximately proportional to sample size, with the smallest circles representing  $n = 1$  and the largest representing  $n = 8$ . The inset shows the phylogenetic relationships between the 22 haplotypes. Open diamonds represent missing haplotypes and small black circles represent unique haplotypes, i.e. those found in a single individual. The population of origin of each unique haplotype is indicated.



**Figure 6** Assignment of populations of *Monotropa hypopitys* to  $K = 14$  genetic clusters based on BAPS analysis of the nuclear microsatellite data. Each cluster is represented by a different colour. The inset shows the phylogenetic relationships between the clusters. Pie chart sizes are approximately proportional to sample size, with the smallest circles representing  $n = 1$  and the largest representing  $n = 8$ .



**Figure 7** Assignment of populations of *Monotropa hypopitys* to  $K = 3$  genetic clusters based on the STRUCTURE analysis of the nuclear microsatellite data. Each cluster is represented by a different shade (black, grey or white), and the percentage of each shade in a pie chart represents the probability of the population belonging to each of the three clusters. Pie chart sizes are approximately proportional to sample size, with the smallest circles representing  $n = 1$  and the largest representing  $n = 8$ .

Archipelago/Haida Gwaii (The Queen Charlotte Islands, British Columbia) during the LGM. Although such modelling approaches are now being used increasingly with phylogeographic analyses to determine the potential past locations of refugial areas, it should be borne in mind for the current study that a further key biotic factor, namely the co-occurrence of suitable host species for a parasite such as *M. hypopitys*, would also determine its distribution at the LGM. The fact that *M. hypopitys* has a range of tree hosts, including *Pinus*, *Abies*, *Tsuga*, *Quercus*, *Acer* and *Betula*, may make this less of a factor than in a species with narrow host specificity but, nevertheless, this factor has to be considered when interpreting the ENM results of this study, and others on parasitic species.

### Eastern refugia

In the eastern part of the North American distribution of *M. hypopitys*, all three marker classes exhibited the highest levels of diversity in the south, with a gradual decrease towards the north-east. Nevertheless, identification of the locations, and indeed number, of eastern refugia are not straightforward. The ENMs indicated suitable habitat for the species at the LGM in an area overlapping the current-day distribution in North and South Carolina, and palynological analyses indicate the presence of beech and particularly pine pollen (Williams *et al.*, 2004). The area has been previously highlighted as a refugium for many species (reviewed in Soltis *et al.*, 2006), and the

palaeodistribution modelling and phylogeographic data suggest that this is also the case for *M. hypopitys*, as populations from the surrounding region exhibited high levels of diversity and a high number of endemic haplotypes for both the *rps2* and the ITS. The absence of the common red haplotype in both the *rps2* and the ITS in these southern populations might indicate the presence of a second, more northerly refugium because this haplotype is common in New England. The BAPS analysis of the nuclear microsatellite data also indicates several endemic clusters in the New England region, but the area lacks the high levels of genetic diversity characteristic of a refugium, and the area was not identified as such by the ENM. Other phylogeographic studies, however, have suggested that such a refugium may have existed south of the limits of the Laurentide ice sheet (Jaramillo-Correa *et al.*, 2004; Godbout *et al.*, 2005; Lee-Yaw *et al.*, 2008; de Lafontaine *et al.*, 2010) and possibly on the Atlantic Coastal Plain (Wall *et al.*, 2010).

The Wisconsin population (WICC) exhibited particularly high genetic diversity, with several private *rps2* and ITS haplotypes. This could indicate a potential further refugium for *M. hypopitys* in the so-called 'Driftless Area', located in south-western Wisconsin, southern Minnesota and north-eastern Iowa. The region lacks glacial drift characteristic of retreating ice, hence its name, and there has been extensive debate as to whether it remained unglaciated at the LGM (Jacobs *et al.*, 1997; Hobbs, 1999). Phylogeographic (Jaramillo-Correa *et al.*, 2004; Rowe *et al.*, 2004; Godbout *et al.*, 2005;

Lee-Yaw *et al.*, 2008) and palaeontological (Jackson *et al.*, 2000) evidence suggests that the area acted as a refugium for several species. It is, however, also possible that this region might have been recolonized, at least in part, from a further southern refugium identified by one of the ENMs in the Texas/Mexico/New Mexico region. As *M. hypopitys* is scarce in this area, only a very limited number of samples from New Mexico were analysed in the present study, but the region has been highlighted previously as a refugium for plant species (e.g. Sosa *et al.*, 2009; Rebernik *et al.*, 2010), although these tend to be species that currently occupy the same areas, rather than those which now exhibit a more northerly distribution.

### A possible role for long-distance dispersal and vicariance?

All marker classes showed a major east–west split in haplotype/genotype distribution across the *M. hypopitys* populations studied. In general, chloroplast *rps2* and nuclear ITS haplotypes, and to a lesser extent BAPS and STRUCTURE clusters based on the nuclear microsatellite data, were only found on one side of the continent. A notable exception to this is the common haplotype designated in dark blue for both the *rps2* and ITS. As well as being found in western populations, these haplotypes were found at high frequency in populations from the northern part of the species' eastern distribution, namely those from Minnesota, Michigan and Quebec. The transcontinental distribution of these haplotypes might indicate long-distance dispersal, either eastwards from western refugia or vice versa. Such dispersal could conceivably have taken place before the start of the Holocene, as macrofossil and palynological studies have indicated a large belt of forest expanding northwards as the ice retreated (Delcourt & Delcourt, 1987; Strong & Hills, 2005; Yansa, 2006). The observed patterns of genetic variation that could differentiate between these particular scenarios are not congruent across the marker systems analysed, but several lines of evidence indicate that recolonization from western refugia may be the most likely scenario. Firstly, the *rps2* dark blue haplotype is phylogenetically closest to a group of haplotypes found solely in the west with the exception of a few singletons. Furthermore, two of these 'western' singletons are found in populations from the northern part of the eastern distribution [WICC (Wisconsin) and VTGB (Vermont)], possibly as a result of eastward migration. Secondly, the predominance of the blue ITS haplotype in the Pacific Northwest and its absence from putative eastern refugial populations (with the possible exception of the 'Driftless Area' – see above) would further indicate an eastward spread. Finally, although the MNSF (Minnesota), MIPP (Michigan) and QCMT (Quebec) populations grouped in the white cluster in the STRUCTURE analysis, the BAPS analysis grouped them in a unique cluster that was genetically very similar to the light blue and yellow 'western' clusters. The isolation-with-migration (IMA) analysis could not confirm predominantly unidirectional gene flow, either from west to east or vice versa. On the whole, however, the

balance of evidence across multiple marker types would appear to indicate some long-distance dispersal, followed by vicariance as the mesic forests that represent the preferred habitat of *M. hypopitys* were succeeded in central North America by grassland, a similar scenario to that found in the related *Orthilia secunda* (Beatty & Provan, 2010). An alternative scenario could involve the presence of the blue haplotype/BAPS cluster in both western and eastern refugia during the LGM, but the phylogenetic position of both the *rps2* haplotype and the BAPS cluster suggest a western origin. Nevertheless, the incongruence between the different marker classes and the stochastic nature of the coalescent further highlights the need to use multiple markers for accurate phylogeographic inference.

### ACKNOWLEDGEMENTS

We are extremely grateful to everybody who provided samples for this project (listed in Table 1). Jordan Metzgar (University of Alaska Fairbanks Herbarium), Donna Ford-Werntz (West Virginia University Herbarium) and David Giblin (University of Washington Herbarium) provided valuable herbarium specimens. We also thank Matt Lundy and Neil Reid, who provided valuable advice on GIS and modelling, and two anonymous referees and the editor, Brett Riddle, for helpful comments and suggestions on an earlier version of the manuscript. Gemma Beatty's PhD research is funded by the Department of Agriculture and Rural Development, Northern Ireland.

### REFERENCES

- Alsos, I.G., Engelskjøn, T., Gjelty, L., Taberlet, P. & Brochmann, C. (2005) Impact of ice ages on circumpolar molecular diversity: insights from an ecological key species. *Molecular Ecology*, **14**, 2739–2753.
- Beatty, G.E. & Provan, J. (2010) Refugial persistence and postglacial recolonization of North America by the cold-tolerant herbaceous plant *Orthilia secunda*. *Molecular Ecology*, **19**, 5009–5021.
- Beatty, G.E. & Provan, J. (2011) High clonal diversity in threatened peripheral populations of the yellow bird's nest (*Hypopitys monotropa*; syn. *Monotropa hypopitys*). *Annals of Botany*, DOI: 10.1093/aob/mcr003.
- Bennett, K.D. & Provan, J. (2008) What do we mean by 'refugia'? *Quaternary Science Reviews*, **27**, 2449–2455.
- Bennett, K.D., Tzedakis, P.C. & Willis, K.J. (1991) Quaternary refugia of north European trees. *Journal of Biogeography*, **18**, 103–115.
- Bermingham, E. & Moritz, C. (1998) Comparative phylogeography: concepts and applications. *Molecular Ecology*, **7**, 367–369.
- Bidartondo, M.I. & Bruns, T.D. (2001) Extreme specificity in epiparasitic Monotropeae (Ericaceae): widespread phylogenetic and geographic structure. *Molecular Ecology*, **10**, 2285–2295.

- Booth, G.D., Niccolucci, M.J. & Schuster, E.G. (1994) *Identifying proxy sets in multiple linear regression: an aid to better coefficient interpretation*. Research paper INT-470. United States Department of Agriculture Forest Service, Ogden, UT.
- Brunsfeld, S.J., Sullivan, J., Soltis, D.E. & Soltis, P.S. (2001) Comparative phylogeography of northwestern North America: a synthesis. *Integrating ecology and evolution in a spatial context* (ed. by J. Silvertown and J. Antonovics), pp. 319–339. Blackwell Science, Oxford.
- Cantor, S.B., Sun, C.C., Tortolero-Luna, G., Richards-Kortum, R. & Follen, M. (1999) A comparison of C/B ratios from studies using receiver operating characteristic curve analysis. *Journal of Clinical Epidemiology*, **52**, 885–892.
- Clement, M., Posada, D. & Crandall, K.A. (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1659.
- Cook, J.A., Bidlack, A.L., Conroy, C.J., Demboski, J.R., Fleming, M.A., Runck, A.M., Stone, K.D. & MacDonald, S.O. (2001) A phylogeographic perspective on endemism in the Alexander Archipelago of southeast Alaska. *Biological Conservation*, **97**, 215–227.
- Corander, J., Waldmann, P. & Sillanpää, M.J. (2003) Bayesian analysis of genetic differentiation between populations. *Genetics*, **163**, 367–374.
- Cwynar, L.C. & MacDonald, G.M. (1987) Geographical variation of lodgepole pine in relation to population history. *The American Naturalist*, **126**, 463–469.
- Delcourt, P.A. & Delcourt, H.R. (1987) Late-Quaternary dynamics of temperate forests: applications of paleoecology to issues of global environmental change. *Quaternary Science Reviews*, **6**, 129–146.
- Evanno, G., Regnaud, S. & Goudet, J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.
- Excoffier, L., Laval, L.G. & Schneider, S. (2005) Arlequin, version 3.0: an integrated software package for population genetic data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.
- Felsenstein, J. (1995) *PHYLIP (Phylogeny Inference Package), version 3.57c*. Department of Genetics, University of Washington, Seattle.
- Godbout, J., Jaramillo-Correa, J.P., Beaulieu, J. & Bosquet, J. (2005) A mitochondrial DNA minisatellite reveals the post-glacial history of jack pine (*Pinus banksiana*), a broad-range North American conifer. *Molecular Ecology*, **14**, 3497–3512.
- Goudet, J. (2001) *FSTAT, version 2.9.3. A program to estimate and test gene diversities and fixation indices*. Available at: <http://www2.unil.ch/popgen/softwares/fstat.htm>.
- Hall, T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, **41**, 95–98.
- Harrison, S.P. & Sánchez Goñi, M.F. (2010) Global patterns of vegetation response to millennial scale variability and rapid climate change during the last glacial period. *Quaternary Science Reviews*, **29**, 2957–2980.
- Heaton, T.H., Talbot, S.L. & Shields, G.F. (1996) An ice age refugium for large mammals in the Alexander Archipelago, southeastern Alaska. *Quaternary Research*, **46**, 86–192.
- Hewitt, G.M. (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, **58**, 247–276.
- Hewitt, G.M. (1999) Post-glacial recolonisation of European biota. *Biological Journal of the Linnean Society*, **68**, 87–112.
- Hewitt, G.M. (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.
- Hewitt, G.M. (2003) Ice ages: their impact on species distributions and evolution. *Evolution on planet Earth* (ed. by L.J. Rothschild and A.M. Lister), pp. 339–361. Academic Press, London.
- Hey, J. & Nielsen, R. (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics*, **167**, 747–760.
- Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *Journal of Climatology*, **25**, 1965–1978.
- Hobbs, H. (1999) Origin of the Driftless Area by subglacial drainage – a new hypothesis. *Geological Society of America Special Papers*, **337**, 93–102.
- Hu, F.S., Hampe, A. & Petit, R.J. (2009) Palaeoecology meets genetics: deciphering past vegetational dynamics. *Frontiers in Ecology and the Environment*, **7**, 371–379.
- Hultén, E. & Fries, M. (1986) *Atlas of North European vascular plants north of the Tropic of Cancer*. Koeltz Scientific Books, Koenigstein.
- Huntley, B. & Birks, H.J.B. (1983) *An atlas of past and present pollen maps for Europe 0–13,000 years ago*. Cambridge University Press, Cambridge.
- Jackson, S.T., Webb, R.S., Anderson, K.H., Overpeck, J.T., Webb, T., III, Williams, J.W. & Hansen, B.C.S. (2000) Vegetation and environment in Eastern North America during the Last Glacial Maximum. *Quaternary Science Reviews*, **19**, 489–508.
- Jacobs, P.M., Knox, J.C. & Mason, J.A. (1997) Preservation and recognition of middle and early Pleistocene loess in the Driftless Area, Wisconsin. *Quaternary Research*, **47**, 147–154.
- Jaramillo-Correa, J.P., Beaulieu, J. & Bosquet, J. (2004) Variation in mitochondrial DNA reveals multiple distant glacial refugia in black spruce (*Picea mariana*), a transcontinental North American conifer. *Molecular Ecology*, **13**, 2735–2747.
- Keller, S.R., Olson, M.S., Silim, S., Schroeder, W. & Tiffin, P. (2010) Genomic diversity, population structure, and migration following rapid range expansion in the Balsam Poplar, *Populus balsamifera*. *Molecular Ecology*, **19**, 1212–1226.
- Klooster, M.R., Hoenle, A.W. & Culley, T.M. (2009) Characterization of microsatellite loci in the myco-heterotrophic plant *Monotropa hypopitys* (Ericaceae) and amplification in related taxa. *Molecular Ecology Resources*, **9**, 219–221.
- de Lafontaine, G., Turgeon, J. & Payette, S. (2010) Phylogeography of white spruce (*Picea glauca*) in eastern North

- America reveals contrasting ecological trajectories. *Journal of Biogeography*, **37**, 741–751.
- Latch, E.K., Dharmarajan, G., Glaubitz, J.C. & Rhodes, O.E., Jr (2006) Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. *Conservation Genetics*, **7**, 295–302.
- Leake, J.R., McKendrick, S.L., Bidartondo, M. & Read, D.L. (2004) Symbiotic germination and development of the myco-heterotroph *Monotropa hypopitys* in nature and its requirement for locally distributed *Tricholoma* spp. *New Phytologist*, **163**, 405–423.
- Lee-Yaw, J.A., Irwin, J.T. & Green, D.M. (2008) Postglacial range expansion from northern refugia by the wood frog, *Rana sylvatica*. *Molecular Ecology*, **17**, 867–884.
- Nogués-Bravo, D. (2009) Predicting the past distribution of species climatic niches. *Global Ecology and Biogeography*, **18**, 521–531.
- Pfenninger, M. & Posada, D. (2002) Phylogeographic history of the land snail *Candidula unifasciata* (Helicellinae, Stylommatophora): fragmentation, corridor migration and secondary contact. *Evolution*, **56**, 1776–1788.
- Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.
- Pielou, E.C. (1991) *After the Ice Age: the return of life to glaciated North America*. University of Chicago Press, Chicago.
- Pritchard, J.K., Stephens, M. & Donnelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.
- Provan, J. & Bennett, K.D. (2008) Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology and Evolution*, **23**, 564–571.
- Provan, J. & Wilson, P.J. (2007) Development of microsatellites for the peat moss *Sphagnum capillifolium* using ISSR cloning. *Molecular Ecology Notes*, **7**, 254–256.
- Provan, J., Powell, W. & Hollingsworth, P.M. (2001) Chloroplast microsatellites: new tools for studies in plant ecology and systematics. *Trends in Ecology and Evolution*, **16**, 142–147.
- Rebernick, C.A., Weiss-Schneeweiss, H., Schneeweiss, G.M., Schonswetter, P., Obermayer, R., Villasenor, J.L. & Stuessy, T. (2010) Quaternary range dynamics and polyploidy evolution in an arid brushland plant species (*Melampodium cinerium*, Asteraceae). *Molecular Phylogenetics and Evolution*, **54**, 594–606.
- Rowe, K.C., Heske, E.J., Brown, P.W. & Paige, K.N. (2004) Surviving the ice: northern refugia and postglacial recolonization. *Proceedings of the National Academy of Sciences USA*, **101**, 10355–10359.
- Schaal, B.A., Hayworth, D.A., Olsen, K.M., Rauscher, J.T. & Smith, W.A. (1998) Phylogeographic studies in plants: problems and prospects. *Molecular Ecology*, **7**, 465–474.
- Soltis, D.E., Gitzendanner, M.A., Strenge, D.D. & Soltis, P.S. (1997) Chloroplast DNA intraspecific phylogeography of plants from the Pacific Northwest of North America. *Plant Systematics and Evolution*, **206**, 353–373.
- Soltis, D.E., Morris, A.B., McClachlan, J.S., Manos, P.S. & Soltis, P.S. (2006) Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology*, **15**, 4261–4293.
- Sosa, V., Ruiz-Sanchez, E. & Rodriguez-Gomez, F.C. (2009) Hidden phylogeographic complexity in the Sierra Madre Oriental: the case of the Mexican tulip poppy *Hunnemannia fumarifolia* (Papaveraceae). *Journal of Biogeography*, **36**, 18–27.
- Stewart, J.R. & Lister, A.M. (2001) Cryptic northern refugia and the origins of the modern biota. *Trends in Ecology and Evolution*, **16**, 608–613.
- Stewart, J.R., Lister, A.M., Barnes, I. & Dalén, L. (2010) Refugia revisited: individualistic responses of species in space and time. *Proceedings of the Royal Society B: Biological Sciences*, **277**, 661–671.
- Strong, W.L. & Hills, L.V. (2005) Late-glacial and Holocene palaeovegetation zonal reconstruction for central and north-central North America. *Journal of Biogeography*, **32**, 1043–1062.
- Taberlet, P., Fumigalli, L., Wust-Saucy, A.G. & Cosson, J.F. (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**, 453–464.
- Thompson, R.S. & Anderson, K.H. (2000) Biomes of western North America at 18,000, 6,000 and 0 <sup>14</sup>C yr BP reconstructed from pollen and packrat midden data. *Journal of Biogeography*, **27**, 555–584.
- Wall, W.A., Douglas, N.A., Xiang, Q.Y., Hoffmann, W.A., Wentworth, T.R. & Hohmann, M.G. (2010) Evidence for range stasis during the latter Pleistocene for the Atlantic Coastal Plain endemic genus *Pyxidanthera* Michaux. *Molecular Ecology*, **19**, 4302–4314.
- Warner, B., Mathewes, R. & Clague, J. (1982) Ice-free conditions on the Queen Charlotte Islands, British Columbia, at the height of the late Wisconsin glaciation. *Science*, **218**, 675–677.
- Williams, J.W., Shuman, B.N., Webb, T., Bartlein, P.J. & Leduc, P.L. (2004) Late-Quaternary vegetation dynamics in North America: scaling from taxa to biomes. *Ecological Monographs*, **74**, 309–334.
- Yansa, C.H. (2006) The timing and nature of Late Quaternary vegetation changes in the northern Great Plains, USA and Canada: a re-assessment of the spruce phase. *Quaternary Science Reviews*, **25**, 263–281.

## SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

**Appendix S1** Diversity values for chloroplast *rps2*, nuclear ITS and microsatellites of *Monotropa hypopitys* by population.

**Appendix S2** Occurrence of private haplotypes for chloroplast *rps2* and nuclear ITS of *Monotropa hypopitys* by population.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

## BIOSKETCHES

**Gemma Beatty** is studying for a PhD under the supervision of Jim Provan, comparing how post-glacial recolonization and range-edge effects have shaped the genetic diversity of several Northern Hemisphere Monotropoideae species. This PhD represents a continuation of her Master's research into range-edge effects in *Orthilia secunda*.

**Jim Provan** is a Senior Lecturer in Evolutionary Genetics at Queen's University Belfast. His research interests focus on examining the effects of threats to biodiversity such as climate change and habitat loss on populations at the genetic level.

---

Editor: Brett Riddle

## SUPPORTING INFORMATION

### Phylogeographic analysis of North American populations of the parasitic herbaceous plant *Monotropa hypopitys* L. reveals a complex history of range expansion from multiple late glacial refugia

Gemma E. Beatty and Jim Provan

*Journal of Biogeography*

**Appendix S1** Diversity values for chloroplast *rps2*, nuclear ITS and microsatellites of *Monotropa hypopitys* by population. A dash indicates zero. (NC – not calculated due to small sample size  $n < 4$ ). Population codes refer to those in Table 1 in the main paper.

Population	Diversity		
	<i>rps2</i> ( $H$ )	ITS ( $H$ )	Microsatellites ( $H_E$ )
BCMP	-	-	0.460
NBBO	-	-	0.207
QCMT	-	0.250	0.260
SKCH	0.700	-	0.156
UAFM11086	0.800	-	0.450
UAFM43121	0.286	-	0.497
UAFM44124	0.400	-	0.384
UAFM44125	0.833	-	0.415
UAFM27744	0.833	-	0.384
MALM	-	-	0.449
MIPP	-	-	0.195
MNSF	0.250	0.464	0.279
MTMR	-	-	0.360
OHGE	1.000	NC	0.553
ORHR	0.714	0.667	0.658
RIPR	-	-	0.511
WV116064	1.000	0.500	0.504
SCCH	0.786	0.524	0.577
WV116065	0.833	NC	0.513
VTGR	0.250	-	0.497
VTGB	0.600	-	0.570
WICC	0.952	0.810	0.473

**Appendix S2** Occurrence of private haplotypes for chloroplast *rps2* and nuclear ITS of*Monotropa hypopitys* by population. Numbers in parentheses represent percentage of private

haplotypes in each population. Population codes refer to those in Table 1 in the main paper.

Population	Private haplotypes	
	<i>rps2</i>	ITS
BCMP	-	-
WTU214041	1 (100%)	-
NBBO	-	1 (100%)
ONMR	1 (50%)	1 (100%)
QCMT	-	1 (13%)
SKCH	1 (20%)	-
UAFM8782	2 (100%)	-
UAFM11086	-	-
UAFM43121	-	-
UAFM44124	-	-
UAFM44125	-	-
UAFM27744	-	-
UAFM60111	-	-
UAFM59072	-	-
UAFM82664	1 (50%)	-
ALPW	-	-
ALHO	1 (100%)	-
MESC	-	-
MEED	-	-
MALM	-	-
MIPP	-	-
MNSF	1 (25%)	2 (25%)
MTMR	-	-
NMCC	-	-
NMCF	-	-
NYMH	2 (100%)	-
OHSR	1 (50%)	1 (100%)
OHGE	1 (25%)	1 (50%)
ORHR	2 (75%)	2 (43%)
RIPR	-	-
WV116064	1 (25%)	1 (25%)
SCCH	4 (50%)	2 (29%)
WV116065	1 (25%)	1 (33%)
VTGR	1 (13%)	-
VTGB	1 (40%)	-
WTU351539	-	-
WV117890	1 (50%)	1 (50%)
WV077446	1 (33%)	-
WICC	3 (57%)	2 (29%)