Journal of Biogeography (J. Biogeogr.) (2013) 40, 1887–1898

ORIGINAL ARTICLE

Phylogeography and polyploid evolution of North American goldenrods (*Solidago* subsect. *Humiles*, Asteraceae)

Jess A. Peirson[1,2]*, Christopher W. Dick[1,2] and Anton A. Reznicek[2]

[1]*Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, Michigan, USA,* [2]*University of Michigan Herbarium, Ann Arbor, Michigan, USA*

*Correspondence: Jess A. Peirson, Department of Ecology and Evolutionary Biology, University of Michigan, 830 North University, Ann Arbor, MI 48109-1048, USA.
E-mail: peirsonj@umich.edu

ABSTRACT

Aim We analysed range-wide chloroplast DNA (cpDNA) variation in a clade of North American goldenrods (*Solidago* subsect. *Humiles*) to infer its biogeographical history and evolution. Our objectives were to: (1) examine the structuring of cpDNA diversity in this widespread species complex, (2) reconstruct Pleistocene refugia and post-glacial migration of the study species, and (3) test hypotheses relating to the frequency of polyploidization. We expected the glacial history of *Solidago* to differ markedly from that of temperate trees and forest understorey plants.

Location North America (Canada, continental USA, Mexico).

Methods 1466 bp of chloroplast intergenic spacer DNA (cpDNA) were sequenced from 368 individuals representing 72 populations of subsect. *Humiles*, which consists of the widespread *Solidago simplex* and four geographically restricted species. Estimates of N_{ST} and G_{ST} were compared as a test of phylogeographical structure, and spatial analysis of molecular variance (SAMOVA) was used to examine cpDNA variation. Rarified haplotype diversity and chromosome diversity (ploidy levels) were used to infer locations of glacial refugia and post-glacial expansion, and to determine origins of polyploidy, respectively.

Results A total of 46 haplotypes were recovered. While there was significant phylogeographical structure ($N_{ST} > G_{ST}$), cpDNA variation was not strongly partitioned across species boundaries, geography or ploidy levels, and six haplotypes were shared among species. The highest haplotype diversity was located in western North America, followed by the south-eastern USA and the formerly glaciated Great Lakes region.

Main conclusions *Solidago simplex* recolonized formerly glaciated eastern North America from refugia in western North America and near the perimeter of the ice margin. The south-eastern USA had only limited involvement in recolonization of these northern regions. The geographical disjunction and scattered positions of polyploids in the haplotype network provide evidence of multiple polyploid origins within *S. simplex*, and the restriction of endemic, polyploid taxa to post-glacial habitats provides evidence of Holocene polyploid speciation. The results highlight polyploidization as a source of adaptive genetic variation and speciation in novel and post-glacial habitats.

Keywords
Chloroplast DNA, goldenrods, haplotype sharing, Holocene biogeography, North America, phylogeography, polyploidy, post-glacial migration, *Solidago*.

Journal of Biogeography

INTRODUCTION

Quaternary climatic oscillations caused pronounced changes in species distributions and altered the genetic structure of populations across the Northern Hemisphere (Hewitt, 2000, 2003). The classical view of Northern Hemisphere Quaternary biogeography posits that repeated continental glaciations caused poleward shifts in species ranges and isolated populations in primarily southern refugia (Davis, 1983; Bennett *et al.*, 1991; Latham & Ricklefs, 1993; Williams *et al.*, 2000). During interglacial periods, many species rapidly expanded their ranges northwards. Repeated cycles of southward range contraction and northward expansion produced a genetic signature of 'southern richness' and 'northern purity' in many species (following Hewitt, 1996, 1999). Phylogeographical studies have demonstrated the importance of southern glacial refugia for many temperate trees and forest understorey plants (reviewed in Taberlet *et al.*, 1998; Soltis *et al.*, 2006). Recent studies of widespread and northern temperate and boreal plants, however, have revived early hypotheses (e.g. Braun, 1928; Marie-Victorin, 1938) that some species survived the Last Glacial Maximum (21–18 ka) in northern refugia nearer the ice margin (Stewart & Lister, 2001; Rowe *et al.*, 2004; McLachlan *et al.*, 2005; Shafer *et al.*, 2010). Phylogeographical data have also revealed evidence of admixture of haplotypes from multiple glacial refugia (Petit *et al.*, 2003; van Els *et al.*, 2012) and interspecific hybridization (Saeki *et al.*, 2011) at higher latitudes for some species, resulting in greater diversity in non-refugial areas. These cumulative data indicate that many species responded idiosyncratically to Quaternary climatic changes (Prentice, 1986; Huntley & Webb, 1989; Stewart *et al.*, 2010).

Solidago L. subsect. *Humiles* (Rydb.) Semple (Asteraceae) presents an interesting system in which to examine Quaternary biogeography and diversification across North America. The subsection has one widespread, morphologically and cytologically variable species, *Solidago simplex* Kunth, and four geographically restricted species: *Solidago arenicola* Keener & Kral, *Solidago kralii* Semple, *Solidago plumosa* Small and *Solidago spathulata* DC. It is endemic to North America and occupies four disjunct centres of distribution (Fig. 1). There is significant cytogeographical structuring within the subsection, and diploid and polyploid cytotypes co-occur regionally in the south-eastern United States and in formerly glaciated areas of eastern North America. This distribution of cytotypes has been proposed as evidence that polyploids evolved multiple times (Peirson *et al.*, 2012).

Investigation of widespread, generally non-forest herbs such as *Solidago* has the potential to uncover Pleistocene and Holocene biogeographical histories that may be shared by many other elements of the North American flora. Many previous phylogeographical studies in North America have emphasized major biome components such as forest trees (see summary and review in Jaramillo-Correa *et al.*, 2009) and forest understorey plants that, although ecologically important, comprise specialized components of the flora. Because goldenrods are perennial herbs with wind dispersed seeds, their life-history traits and dispersal capabilities certainly differ from those of large-seeded, long-lived forest trees. Therefore, we expect that their responses to Quaternary climatic changes probably differed as well.

We analysed range-wide chloroplast genetic diversity in *Solidago* subsect. *Humiles* in order to: (1) examine the structuring of chloroplast DNA (cpDNA) diversity with regard to species boundaries, geography and ploidy levels in this widespread North American species complex, (2) reconstruct Pleistocene refugia and post-glacial migration of the study species, and (3) test hypotheses relating to the frequency of polyploidization.

MATERIALS AND METHODS

Study system

Goldenrods in *Solidago* subsect. *Humiles* have resinous glands on the foliage and involucral bracts such that all members are sticky to the touch, a unique characteristic in the genus of *c.* 100 species (Semple & Cook, 2006). Members of *S.* subsect. *Humiles* generally occur in non-forest habitats and are not weeds.

Solidago arenicola is restricted to rocky or sandy riverbanks and floodplains in the Cumberland Plateau region of northern Alabama and Tennessee (Table 1). Of the four narrowly distributed species, it is the only polyploid. *Solidago kralii* is confined to sand hills along the Coastal Plain fall line in Georgia and South Carolina. *Solidago plumosa* is known from a single population on mafic rocks along the Yadkin River in Stanley Co., North Carolina. *Solidago spathulata* inhabits sand dunes along the Pacific Coast from central California to northern Oregon (Table 1). *Solidago simplex* is divided into diploid *S. simplex* subsp. *simplex* and polyploid *S. simplex* subsp. *randii*. Diploid subspecies *simplex* is widespread in montane and alpine habitats throughout the western cordillera from Alaska to Mexico (Fig. 1). Disjunct eastern diploids in the northern Great Lakes region and Gaspé, Quebec, have also been placed in subspecies *simplex* (Fig. 1). Polyploid subspecies *randii* is restricted to the Great Lakes region and Appalachian Mountains in eastern North America (Fig. 1). Four varieties are currently recognized in subspecies *randii*, with two of those endemic to the formerly glaciated Great Lakes region (Table 1).

Taxon and population sampling

A total of 368 accessions of *Solidago* subsect. *Humiles* from 72 populations (including 289 individuals from 57 populations of *S. simplex*) were collected from throughout North America (locality, sampling and voucher details are presented in Appendix S1 in Supporting Information). Sampling included all species within *S.* subsect. *Humiles* and all

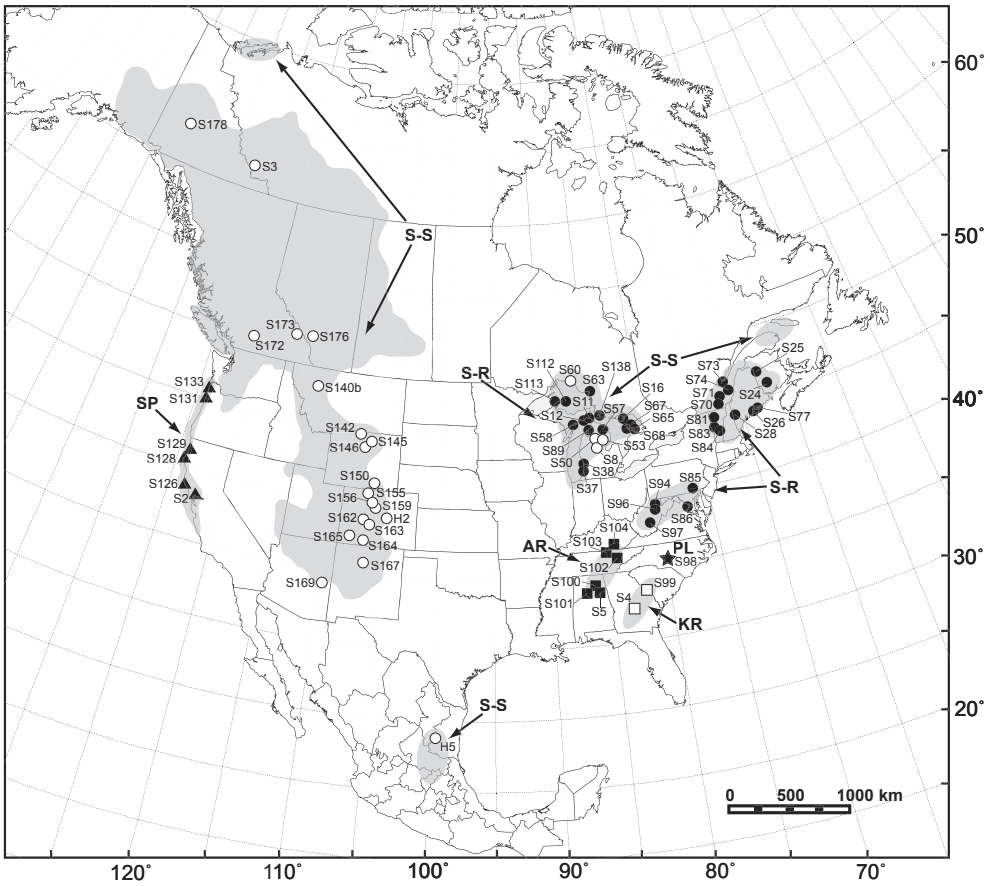

Figure 1 Distribution of *Solidago* subsect. *Humiles* in North America. Generalized taxon distributions are indicated by grey shading except for *S. plumosa*, which is indicated by a single star: *S. arenicola* (AR), *S. kralii* (KR), *S. plumosa* (PL), *S. simplex* subsp. *randii* (S-R), *S. simplex* subsp. *simplex* (S-S), *S. spathulata* (SP). Sampling localities for this study are as follows: *S. arenicola* (closed squares), *S. kralii* (open squares), *S. plumosa* (star), *S. simplex* subsp. *randii* (closed circles), *S. simplex* subsp. *simplex* (open circles), and *S. spathulata* (triangles). The Great Lakes region is the only area where taxon distributions overlap.

Table 1 Summary of natural history information for *Solidago* subsect. *Humiles*. Ploidy data are summarized from Peirson *et al.* (2012). Geographical distributions are as follows: western North America (WNA), Great Lakes Region (GR), north-eastern North America (NE), and south-eastern United States and Southern Appalachians (SEUS). *Solidago simplex* is the only species in the subsection with recognized infraspecific taxa.

Taxon	Ploidy	Geography	Habitat
S. arenicola	4*x*	SEUS	Gravel/sand riverbanks of the Cumberland Plateau
S. kralii	2*x*	SEUS	Pinelands and xeric sand hills along the SEUS Coastal Plain fall line
S. plumosa	2*x*	SEUS	Rock riverbanks along the Yadkin River
S. simplex	2*x*, 4*x*, 6*x*	WNA, GL, NE	–
subsp. *simplex*	2*x*	WNA, GL, NE	–
var. *chlorolepis*	2*x*	NE	Rock barrens and uplands of the Gaspé Peninsula
var. *nana*	2*x*	WNA	Alpine habitats in the Cascade Mountains
var. *simplex*	2*x*	WNA, GL	Rock/sand/alpine habitats in western N. America and disjunct to the Great Lakes Region
subsp. *randii*	4*x*, 6*x*	GL, NE, SEUS	–
var. *gillmanii*	4*x*	GL	Coastal sand dunes of Lake Huron and Lake Michigan
var. *ontarioensis*	4*x*	GL	Rock lakeshores of Lake Huron, Lake Michigan and Lake Superior
var. *monticola*	4*x*	NE	Rock uplands in New England and Quebec
var. *racemosa*	4*x*, 6*x*	NE, SEUS	Rock riverbanks throughout the Appalachian Mountains
S. spathulata	2*x*	WNA	Coastal sand dunes along the Pacific Ocean

infraspecific taxa within *S. simplex*, except the narrowly endemic *S. simplex* subsp. *simplex* vars *nana* and *chlorolepis*. Leaf samples from individual clones spaced at least 2 m apart were preserved in silica gel. Population vouchers were deposited at the University of Michigan Herbarium (MICH) or the University of Waterloo Herbarium (WAT). In addition

to field-sampled populations, leaf samples from two herbarium specimens of *S. simplex* var. *simplex* (samples H2 and H5 in Appendix S1) were included. Sample H5 represents the only accession from the south-western extreme of the distribution in Mexico.

DNA extraction, PCR amplification and sequencing

Total genomic DNA was isolated with the DNeasy Plant Mini Kit (Qiagen Corp., Valencia, CA, USA) following the manufacturer's protocols, except that herbarium samples were incubated overnight with β-mercaptoethanol. Several genomic samples required additional purification using glass milk technology (GeneClean, MP Biomedicals, Solon, OH, USA).

Our sequencing strategy for *Solidago* subsect. *Humiles* was twofold. First, single individuals from all 72 populations were sequenced to capture preliminary estimates of range-wide genetic diversity rather than to examine intrapopulation variation intensively (following Petit *et al.*, 2005). Second, an additional eight individuals per population from 37 of those populations that represented the taxonomic and geographical breadth of the subsection were sequenced to assess intrapopulation genetic diversity.

Two chloroplast intergenic spacers (*trn*H–*psb*A and *trn*K–*rps*16) were used for this study because of their polymorphism and consistent polymerase chain reaction (PCR) amplification [previous screening of the nuclear ribosomal internal transcribed spacer (ITS) region revealed almost no sequence variation across *Solidago*; J.A.P., unpubl. data]. The *trn*H–*psb*A spacer was amplified and sequenced using the trnH-f and psbA-3f primers of Kress *et al.* (2005), and the *trn*K–*rps*16 spacer was amplified and sequenced with primers trnK-F1 (5′-GCCGCACTTAAAAGCCGAG-3′) and rps16-R1 (5′-CCCAATGAGCCGTCTATCG-3′) designed specifically for this study. The PCR amplifications were carried out with Takara Ex Taq™ polymerase (Takara Bio Inc., Otsu, Shiga, Japan). The thermal cycle for *trn*H–*psb*A was 94 °C for 3 min, followed by 40 cycles of 94 °C for 45 s, 55 °C for 45 s and 72 °C for 3 min, with a final extension at 72 °C for 10 min. The thermal cycle for *trn*K–*rps*16 was 94 °C for 3 min, followed by 35 cycles of 94 °C for 45 s, 55 °C for 1 min and 72 °C for 3 min, with a final extension at 72 °C for 10 min. Negative controls using water instead of genomic DNA were run with each PCR mix to check for contamination. PCR products were purified using ExoSAP-IT (Affymetrix, Santa Clara, CA, USA) and sequenced using BigDye chemistry (Applied Biosystems Inc., Foster City, CA, USA) on an ABI 3730xl capillary sequencer.

DNA sequence alignment, indel scoring and concatenation

Sequence files were edited in Sequencher 4.6 (Genecodes Corp., Ann Arbor, MI, USA) and aligned automatically in Clustal X 2.0 (Larkin *et al.*, 2007) using default parameters. Final alignments were performed in Se-Al 2.0a11 (available for download at http://tree.bio.ed.ac.uk/software/seal/) following recommendations for cpDNA spacer regions (Kelchner, 2000; Borsch *et al.*, 2003). In the *trn*K–*rps*16 spacer, length variation in a poly-A region (from 9 to 14 bp) was removed because of the potential for non-homologous repeat patterns. In the *trn*H–*psb*A spacer, a variable region of mostly TA repeats (hotspot region *sensu* Borsch *et al.*, 2003; Schlaepfer *et al.*, 2008) was removed because of potential homoplasy. Also in the *trn*H–*psb*A spacer, a 23 bp inversion/indel complex in three samples from population S60 was reverse-complemented and scored as a single mutational event.

The congruence of the *trn*K–*rps*16 and *trn*H–*psb*A spacers was tested with a partition homogeneity test (the incongruence length difference test of Farris *et al.*, 1995) implemented in paup* 4.0b10 (Sinauer Associates, Sunderland, MA, USA). Because the test was not significant (100 replicates; $P = 0.45$), the regions were concatenated into a single alignment file for all subsequent analyses. Based on recommendations in Simmons *et al.* (2007), indels in the concatenated data set were coded as single mutational events using a variation of modified complex indel coding (MCIC *sensu* Müller, 2006).

Network construction and analyses

Sequence data from all 368 individuals of *Solidago* subsect. *Humiles* were analysed in a statistical parsimony framework using the program tcs 1.21 (Clement *et al.*, 2000). Haplotype identification and network reconstruction were performed using the MCIC coded data set with gaps treated as a fifth character and the default connections between haplotypes set at the 95% limit.

Genetic diversity parameters, including haplotype diversity (h), indel diversity (π_I) and nucleotide diversity (π_S), were calculated for *Solidago* subsect. *Humiles* as a whole, for each species, and for infraspecific taxa within *S. simplex* using DnaSP 4.0 (Rozas *et al.*, 2003). To compare haplotype diversity between geographical regions, the number of haplotypes was standardized by rarefying to 40 individuals for each geographical region using the program contrib 1.02 (Petit *et al.*, 1998). Four broad geographical regions were recognized (see Table 1 and Appendix S1): western North America (WNA), south-eastern Unites States and Southern Appalachians (SEUS), Great Lakes region (GL) and north-eastern North America (NE). To test for the presence of phylogeographical structure (*sensu* El Mousadik & Petit, 1996; Pons & Petit, 1996), N_{ST} (a measure of genetic differentiation that incorporates phylogenetic distance) was compared to permuted values of G_{ST} (an unordered measure of genetic differentiation that does not incorporate phylogenetic distance) with the program SPAGeDi 1.3 (Hardy & Vekemans, 2002).

To test for a signature of population expansion across *Solidago* subsect. *Humiles*, a mismatch analysis (*sensu* Harpending *et al.*, 1998) of the entire data set was conducted in

Arlequin 3.5, with 1000 replicates of Arlequin's parametric bootstrap method (Schneider & Excoffier, 1999; Excoffier & Lischer, 2010). As an additional test of population expansion, Tajima's D (Tajima, 1989) was calculated in Arlequin 3.5. The value of tau (τ) from the mismatch analysis was used to estimate the age of population expansion with the equation $t = \tau/2\mu$, where t is the number of generations and μ is the substitution rate per generation for the combined spacer data (per-site rate × length of sequence). Two per-site mutation rates were used to give a range for the timing of expansion: a slower rate of 1.0×10^{-9} substitutions/silent site/year (following Dick *et al.*, 2007) and an order of magnitude faster rate of 1.0×10^{-8} substitutions/silent site/year. A generation time of 2 years was used.

To assess the partitioning of genetic variation, an analysis of molecular variance (AMOVA; Excoffier *et al.*, 1992) was performed using Arlequin 3.5, grouping populations by species, by species but with *Solidago simplex* divided into two subspecies, by geography without regard to taxonomy (see above), by ploidy without regard to taxonomy (diploid/polyploid), and by groupings from spatial analysis of molecular variance (SAMOVA; Dupanloup *et al.*, 2002) for $K = 5$ (see below). SAMOVA was used to analyse the geographical structuring of genetic variation in *S.* subsect. *Humiles* without assignments to a priori groups. For the SAMOVA, 37 populations of *S.* subsect. *Humiles* with intrapopulation sampling ($n = 333$ individuals) were analysed using $K = 2$ to 15 simulated groups. Single accessions cannot be incorporated in SAMOVA and thus were not analysed.

RESULTS

Chloroplast DNA diversity and population expansion

The concatenated *trn*K–*rps*16 and *trn*H–*psb*A data set (including indels) had an aligned length of 1482 bp. The MCIC concatenated data set used in the statistical parsimony analysis had an aligned length of 1295 bp. Forty-six chloroplast haplotypes (GenBank accession numbers KC309690–KC310057 for *trn*H–*psb*A and KC310058–KC310425 for *trn*K–*rps*16) were recovered from 368 individuals of *Solidago* subsect. *Humiles* (Table 2, Fig. 2, Appendix S2). Fifteen individuals possessed unique haplotypes (haplotypes 1, 10, 12, 22, 25, 26, 32, 34, 39, 40, 42, 43, 44, 45 and 46). The overall topology of the statistical parsimony network and recovered groupings of haplotypes were consistent with separate maximum likelihood and Bayesian inference analyses of an expanded cpDNA haplotype data set with broader species-level sampling (Peirson, 2010).

Among the four geographical regions delimited in this study, the greatest number of haplotypes was recovered from western North America ($n = 26$). A total of 14 haplotypes were recovered from each the Great Lakes region and the south-eastern United States. The haplotype complement in north-eastern North America was considerably lower ($n = 8$). When rarefied to 40 individuals, regional haplotype diversity was as follows: western North America, 13.9; south-eastern United States, 11.5; Great Lakes region, 10.2; and north-eastern North America, 6.5.

Table 2 Chloroplast genetic diversity within *Solidago* subsect. *Humiles*. Number of sampled populations (N_P), number of sampled individuals (N_I), number of haplotypes (H_P), haplotype diversity (h), number of segregating indels (I), indel diversity (π_I), number of segregating sites (S) and nucleotide diversity (π_S) are indicated for the subsection as a whole, for each species, for infraspecific taxa within *S. simplex*, and for populations of *S. simplex* subsp. *simplex* from western (W) and eastern (E) North America.

Taxon	N_P	N_I	H_P	h	I	π_I	S	π_S
Solidago subsect. *Humiles*	72	368	46	0.93	42	0.0023	28	0.0015
Solidago arenicola	6	30	7	0.81	7	0.0019	4	0.0009
Solidago kralii	2	10	3	0.51	2	0.0007	1	0.0002
Solidago plumosa	1	9	2	0.22	3	0.0005	2	0.0003
Solidago simplex	57	289	35	0.90	35	0.0021	23	0.0014
subsp. *simplex*	26	138	22	0.89	28	0.0024	14	0.0015
var. *simplex* (W)	21	109	19	0.86	28	0.0026	12	0.0015
var. *simplex* (E)	5	29	6	0.74	8	0.0012	7	0.0012
subsp. *randii*	31	151	20	0.85	20	0.0013	15	0.0014
var. *gillmanii*	6	38	5	0.70	7	0.0015	4	0.0011
var. *ontarioensis*	8	48	9	0.84	10	0.0016	8	0.0014
var. *monticola*	6	22	6	0.74	5	0.0011	6	0.0007
var. *racemosa*	11	43	12	0.79	10	0.0007	8	0.0013
Solidago spathulata	6	30	7	0.76	6	0.0015	3	0.0004

Mismatch analysis and the test of selective neutrality of the complete data set supported the hypothesis that *Solidago* subsect. *Humiles* underwent an historical population expansion (sum of squared deviation = 0.01, $P = 0.24$; Harpending's raggedness index = 0.02, $P = 0.17$; Tajima's $D = -1.4$, $P = 0.05$). The population expansion model produced an estimate of tau (τ) = 5.6 and thus an estimated age of population expansion for *S.* subsect. *Humiles* of 0.48–4.4 Ma.

Phylogeographical structure (N_{ST} versus G_{ST}) and AMOVA

N_{ST} of *Solidago* subsect. *Humiles* (0.66) was significantly higher than G_{ST} (0.62) at a global level ($P = 0.02$), and for all grouping schemes tested (Table 3), indicating significant phylogeographical structure in the data set regardless of the grouping scheme used.

AMOVA indicated generally weak partitioning of haplotype diversity with regard to taxonomic boundaries, geography and ploidy level, with the sum of intragroup and intrapopulation diversity explaining between 78.5% and 91.8% of the variation (Table 3). These results were consistent with visual inspection of the haplotype distribution data that showed haplotype sharing among species, geographical regions and ploidy levels (Fig. 3). The population structure identified by the SAMOVA ($K = 5$, see below) produced substantially more between group genetic differentiation than any of the a priori groupings (Table 3), further highlighting

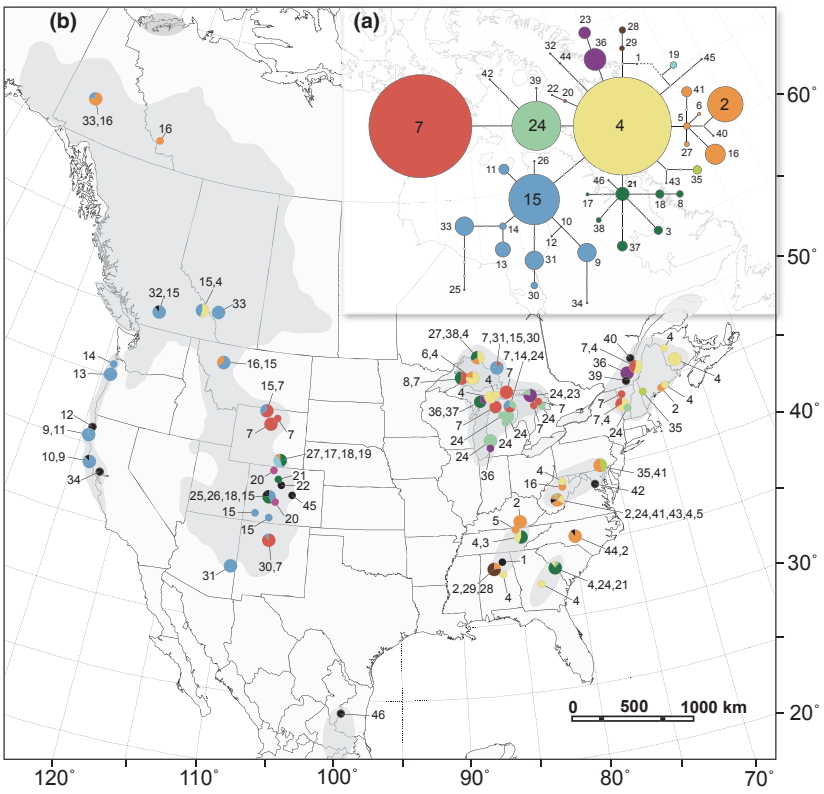

Figure 2 (a) Statistical parsimony network of 46 cpDNA haplotypes recovered from 368 accessions of *Solidago* subsect. *Humiles*. Haplotype circle area represents relative frequency. Dashed lines indicate alternative connections; hypothesized haplotypes are indicated by white squares. Haplotypes were arbitrarily assigned numbers 1–46 and colour-coded to indicate major groups. Haplotypes unique to a single individual are coloured black. (b) Geographical distribution of haplotypes in North America. The haplotype complement at each sampling locality is indicated; haplotypes are ordered from least to most common. Generalized taxon distributions are indicated by grey shading as in Fig. 1.

Table 3 Results of the analyses of molecular variance (AMOVA) and the test of phylogeographical structure ($N_{ST} > G_{ST}$) of cpDNA sequence data from *Solidago* subsect. *Humiles* populations. Groupings were as follows: (a) by species; (b) by species but with *S. simplex* divided into two subspecies; (c) by geography with four regions of North America recognized (see text and Table 1 and Appendix S1 for regions); (d) by ploidy (diploid versus polyploid); and (e) by groupings from SAMOVA (for $K = 5$).

Grouping	d.f.	Fixation indices	% of variation	*P*	N_{ST}	G_{ST}	*P*
(a)	4	$F_{CT} = 0.22$	21.5	0.003	0.50	0.33	< 0.001
	32	$F_{SC} = 0.63$	49.2	< 0.001	–	–	–
	296	$F_{ST} = 0.71$	29.3	< 0.001	–	–	–
(b)	5	$F_{CT} = 0.16$	16.2	0.001	0.44	0.30	0.002
	31	$F_{SC} = 0.62$	51.7	< 0.001	–	–	–
	296	$F_{ST} = 0.68$	32.1	< 0.001	–	–	–
(c)	3	$F_{CT} = 0.14$	13.6	< 0.001	0.18	0.12	0.012
	33	$F_{SC} = 0.62$	53.9	< 0.001	–	–	–
	296	$F_{ST} = 0.67$	32.5	< 0.001	–	–	–
(d)	1	$F_{CT} = 0.08$	8.2	< 0.001	0.11	0.04	0.002
	35	$F_{SC} = 0.65$	59.4	< 0.001	–	–	–
	296	$F_{ST} = 0.68$	32.4	0.005	–	–	–
(e)	4	$F_{CT} = 0.48$	48.1	< 0.001	0.73	0.53	< 0.001
	32	$F_{SC} = 0.51$	26.3	< 0.001	–	–	–
	296	$F_{ST} = 0.74$	25.6	< 0.001	–	–	–

F_{CT} = differentiation among groups of populations, F_{SC} = differentiation among populations within groups, F_{ST} = differentiation among individuals within populations.

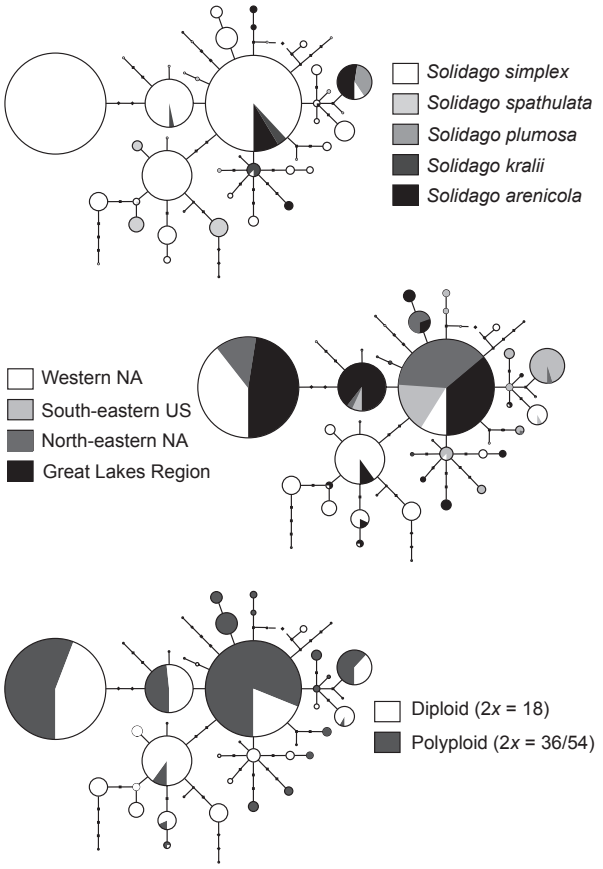

Figure 3 Statistical parsimony cpDNA haplotype networks of *Solidago* subsect. *Humiles*, showing distribution of haplotypes among species (top), geographical regions of North America (NA) (middle), and ploidy levels (bottom). See text for description of geographical regions. Ploidy data are from Peirson *et al.* (2012).

the relatively shallow levels of taxonomic, geographical and ploidy-related structuring of genetic diversity (Fig. S1 in Appendix S3; see population assignments to SAMOVA groups 1–5 in Appendix S1).

Spatial analysis of molecular variance

The optimal number of groups indicated by SAMOVA was $K = 5$ ($F_{CT} = 0.48$; 0 singleton populations). Higher levels of K produced modestly higher fixation indices but also included single-population groups and were thus suboptimal (Table S1 in Appendix S3). The genetic structure (Fig. S1 in Appendix S3) identified by the SAMOVA ($K = 5$) included one large grouping of 26 populations that comprised multiple species and ploidy levels and populations in all four geographical regions (Group 1). Three of the smaller groupings ranged in size from 2–3 populations per group and were composed of regionally proximal sets of populations. Groups 2 and 4 included *Solidago simplex* var. *simplex* and *S. spathulata* from western North America, while Group 3 included *S. plumosa* and one population of *S. arenicola* in the south-eastern United States. Group 5 comprised two populations of diploid *S. simplex* var. *simplex* from the northern Rocky Mountains and New Mexico and two populations of tetraploid *S. simplex* var. *gillmanii* from the Great Lakes region.

DISCUSSION

Structuring of genetic diversity

Tests of phylogeographical structure indicated that there was significant structuring of cpDNA diversity within *Solidago* subsect. *Humiles*, but AMOVA revealed that groupings of populations by taxonomy, geography and ploidy explained relatively little of the partitioning of genetic diversity within the subsection. Of the grouping schemes tested with AMOVA, the partitioning based on ploidy explained the smallest amount of variation ($F_{CT} = 8.2\%$). SAMOVA (at $K = 5$) produced substantially higher between-group measures of differentiation than any of the a priori groupings but revealed shallow geographical structuring of cpDNA variation. Twenty-six of the 37 populations in the SAMOVA were incorporated into a widely distributed group that included populations of multiple species and ploidy levels from across North America (Fig. S1).

Consistent with the weak partitioning of cpDNA diversity, our range-wide cpDNA data revealed that six haplotypes were shared among species in *Solidago* subsect. *Humiles* (Fig. 3). Broader species-level sampling in *Solidago* has also uncovered haplotype sharing among more distantly related species from different subsections (Schlaepfer *et al.*, 2008; Peirson, 2010; Laureto & Barkman, 2011). The second most common haplotype in our study, haplotype 4, was present in *S. simplex* (both 2*x* and 4*x* individuals), *S. arenicola* and *S. kralii*; and it was widely distributed throughout eastern North America and into south-western Canada. This haplotype has also been found in *S. hispida* and *S. rigida*, species in different subsections of the genus (Peirson, 2010). Haplotype 4 occupies a central position in the network (Fig. 2), suggesting that it is likely to be an ancestral haplotype.

The patterns of cpDNA diversity emerging from *Solidago*, including the sharing of likely ancestral haplotypes across species, are similar to patterns found in recently diversified groups like North American rock cresses (*Boechera*, Brassicaceae; Kiefer *et al.*, 2009; Kiefer & Koch, 2012) and New World barleys (*Hordeum*, Poaceae; Jakob & Blattner, 2006). Past hybridization and chloroplast capture have also been important factors in the structuring of cpDNA in some species, such as red and silver maple (*Acer rubrum* and *A. saccharinum*), which initially diverged in the Pliocene but share cpDNA haplotypes in areas of syntopy (Saeki *et al.*, 2011). It is likely that both the retention of ancestral polymorphism and gene flow have been factors in haplotype sharing among species in *S.* subsect. *Humiles*.

Weak partitioning of cpDNA diversity, sharing of haplotypes between species, and a potentially recent population

expansion (0.48–4.4 Ma) all support treating *Solidago* subsect. *Humiles* as a species complex and including these closely related taxa in a single phylogeographical analysis (similar to the approach used by Scotti-Saintagne *et al.*, 2013).

Holocene biogeography

European and North American phylogeographical studies have highlighted the importance of southern refugia for many temperate trees and forest understorey plants (reviewed in Taberlet *et al.*, 1998; Soltis *et al.*, 2006), but patterns emerging from phylogeographical studies of widespread and northern-temperate and boreal plants suggest that the late Pleistocene and Holocene histories of some species included survival in northern/periglacial refugia (e.g. McLachlan *et al.*, 2005; Ronikier *et al.*, 2008; Beatty & Provan, 2010, 2011; Michl *et al.*, 2010; Allen *et al.*, 2012). Our results indicate that *Solidago simplex* survived glaciation in multiple refugia in different regions of the continent and that formerly glaciated areas of eastern North America are actually a mixing ground for migrants from western North America, northern refugia around the perimeter of the ice sheets, and possibly from the south-eastern United States.

The genetic structure identified by the SAMOVA grouped all but four populations of *Solidago* subsect. *Humiles* in eastern North America into the widespread group of 26 populations. Yet despite the grouping from SAMOVA and the fact that glaciated eastern North America and the south-eastern United States harboured similar numbers of haplotypes, the haplotype complements between the regions were quite different. Of the 14 haplotypes found in the Great Lakes region, only haplotypes 4 and 24 were present in the south-eastern United States. Haplotype 4 is an ancestral haplotype that occupies the centre of the network and is common and widely distributed taxonomically and geographically. Thus the distribution of this haplotype provides little explanatory power, but it is possible that individuals with this haplotype recolonized glaciated areas from the south. Haplotype 24 is common in Great Lakes region *S. simplex* populations but is rare in the south-eastern United States, recovered from single individuals of *S. kralii* (South Carolina) and *S. simplex* var. *racemosa* (West Virginia). Of the eight haplotypes found in north-eastern North America, four were shared with the south-eastern United States. Of these, only haplotype 4 was common in north-eastern North America.

Our data indicate that the south-eastern United States was not the primary source of post-glacial migrants for formerly glaciated areas of eastern North America. The Great Lakes region had a higher genetic similarity to western North America ($n = 7$ shared haplotypes) and north-eastern North America ($n = 4$ shared haplotypes) than to the south-eastern United States. The presence of four 'western' haplotypes (14, 15, 30, 31 in Figs 2–3) in the northern Great Lakes region supports biogeographical hypotheses of a migration corridor connecting the western Cordillera and the Great Lakes region (Fernald, 1925; Marie-Victorin, 1938; Marquis & Voss, 1981), and the SAMOVA confirmed a linkage between the glaciated Great Lakes region and the Rocky Mountain region, south of the Cordilleran ice sheet. Our data do not support Ringius & Semple's (1987) hypothesis that diploid *S. simplex* recolonized glaciated eastern North America from a refugium in Beringia (consistent with patterns found in some other widespread boreal species, e.g. Breen *et al.*, 2012; Wróblewska, 2012). None of the haplotypes recovered from north-western Canada were present in the Great Lakes region. Our sampling in western Canada, however, was not extensive enough to fully explain patterns of Holocene recolonization there.

Our analyses identified haplotypes that were restricted to formerly glaciated regions of eastern North America (haplotypes 6, 8, 23, 37, 38 in the Great Lakes region; 39, 40 in north-eastern North America; and 36 shared by both regions). These data from *Solidago simplex* support the possibility that glaciated eastern North America was recolonized in part by migrants that survived glaciation in northern refugia closer to the ice margin, where the species no longer occurs. This pattern in *S. simplex* is congruent with findings from other widespread and northern-temperate and boreal plant and animal taxa in North America (e.g. Dobes *et al.*, 2004; McLachlan *et al.*, 2005; Beatty & Provan, 2010, 2011; van Els *et al.*, 2012). Dobes *et al.* (2004) hypothesized a cryptic refugium near the Great Lakes region for a widespread montane and boreal rockcress complex (*Arabis drummondii–A. holboellii* complex, Brassicaceae), and Beatty & Provan (2011) suggested that saprotrophic pinesap (*Monotropa hypopithys*, Monotropaceae) survived glaciation in multiple refugia, possibly including the unglaciated 'Driftless Area' of the upper Midwestern United States. Together these studies suggest that northern refugia were important contributors to post-glacial colonization in North America, especially for temperate and boreal herbs.

Evolution of polyploidy

Our phylogeographical data rejected a single origin of polyploidy in *Solidago simplex* and indicated more generally that polyploids in *S.* subsect. *Humiles* probably formed numerous times. Polyploids in *Solidago* subsect. *Humiles* harboured 24 chloroplast haplotypes that do not form a single clade; 8 of those haplotypes are shared with diploids while 16 are present in polyploids only (Fig. 3). From a theoretical standpoint, a single origin of polyploidy would result in a single ancestral chloroplast haplotype, and descendants of the original polyploid should form a single lineage. Therefore, the lack of monophyly of polyploids in our study strongly suggests that they evolved multiple times. While it is also possible that gene flow and chloroplast capture from other tetraploid *Solidago* species created the patterns uncovered in *S.* subsect. *Humiles*, our data do not strongly support that scenario. Of the 16 polyploid-only haplotypes recovered in our study, only two have been found in goldenrod species

outside of *S.* subsect. *Humiles* (Peirson, 2010). Peirson *et al.* (2012) also argued that the recurrent origin of polyploids in *S. simplex* was likely, in part because of the rare occurrence of cryptic tetraploids in otherwise diploid populations. They suggested that while those data did not indicate high levels of recurrent polyploidy, they did show that polyploid formation in *S. simplex* could occur on contemporary ecological timescales.

The likely recurrent evolution of polyploids in *Solidago* subsect. *Humiles* is consistent with results from studies of other goldenrod species. Schlaepfer *et al.* (2008) concluded that multiple polyploidization events were significantly more likely than a single origin of tetraploids in *S. gigantea*, and they inferred as many as seven independent origins in eastern North America. Similarly, Halverson *et al.* (2008) rejected the hypothesis that polyploid cytotypes in *S. altissima* had a single, ancient origin. While genus-wide phylogenetic data are largely lacking for *Solidago* (see discussion in Beck *et al.*, 2004), recurrent formation of polyploids seems probable in a number of species (Peirson *et al.*, 2012). These results from *Solidago* are consistent with numerous molecular phylogenetic studies that have demonstrated that the recurrent formation of polyploid lineages is the norm in many plants (e.g. Rebernig *et al.*, 2010; Symonds *et al.*, 2010; Artyukova *et al.*, 2011).

While our demographic analyses placed the main population expansion in *S.* subsect. *Humiles* between the late Pleistocene and early Pliocene (well before the Wisconsinan glacial period hypothesis of Ringius & Semple, 1987), our phylogeographical data also support the evolution of polyploid taxa within a Holocene timeframe. Phylogeographical results revealed that diploid *S. simplex* in the Great Lakes region shared haplotypes with the two tetraploid varieties that are endemic to post-glacial habitats there. In addition, the mixed-ploidy populations found by Peirson *et al.* (2012) were all in the formerly glaciated Great Lakes region. The presence of newly formed, cryptic tetraploids and the sharing of haplotypes between diploids and endemic polyploids suggest that this formerly glaciated region is an active zone of recurrent polyploid formation. It also suggests that the two edaphic endemic polyploids in the Great Lakes region evolved *in situ* during the Holocene. These data for *Solidago* support the hypothesis of polyploidization as a source of adaptive genetic variation and speciation in novel and post-glacial habitats (Stebbins, 1974; Ramsey, 2011; Parisod, 2012).

CONCLUSIONS

Our results indicate that members of *Solidago* subsect. *Humiles* survived glaciation in multiple refugia. Formerly glaciated areas of eastern North America, like the Great Lakes region, are a mixing ground for *Solidago* and were most likely recolonized by migrants from western North America and from refugia near the perimeter of the ice margin. Refugia in the south-eastern United States had only limited involvement in recolonization of these more northerly regions. The glacial biogeography of *S. simplex* is congruent with patterns emerging from other North American temperate and boreal herbs and does not adhere to palaeoecological models of glacial survival in isolated southern refugia or to phylogeographical patterns found in many temperate trees and forest understorey species. These results highlight the dynamic and individualistic nature of post-glacial recolonization, and they also emphasize the importance of taxon selection when attempting to elucidate the broader picture of Holocene biogeography in North America.

The results of this study of *S.* subsect. *Humiles* also raise the question of whether Quaternary climatic cycles and associated vegetation dynamics promoted or hindered differentiation and speciation in this complex of goldenrods. The likely recurrent formation of polyploid lineages supports the hypothesis of adaptive, polyploid diversification in the group, and the probable Holocene evolution of edaphic endemic polyploids in post-glacial habitats in the Great Lakes region suggests rapid differentiation since the Last Glacial Maximum. To address this question more thoroughly, however, will require a well-resolved *Solidago* phylogeny and additional genetic data in the *S. simplex* species complex. Recent advances in next-generation sequencing hold significant promise for recently evolved and cytologically complex groups such as *Solidago* (Hudson, 2008; Emerson *et al.*, 2010). An approach similar to the one used by Griffin *et al.* (2011) to examine the evolution of polyploid Australian alpine grasses will be essential to untangling the complicated evolution of *Solidago simplex* and other species complexes in *Solidago*.

ACKNOWLEDGEMENTS

We thank B. Bennett and J. Semple for providing leaf material, A. Larson for assistance in the field, and the National Park Service and numerous state agencies for granting permission to collect specimens on public lands. We acknowledge the curators and staff of the following herbaria: ALA, ARIZ, ASU, BRIT, BRY, COLO, GH, MAINE, MO, MONTU, MSC, MT, NMC, NYS, QFA, RM, TEX, UBC, UMBS, UTC, VT, WIS and WVA. Financial assistance was provided to J. Peirson by The University of Michigan, The Michigan Botanical Foundation, The Hanes Fund for Botanical Research, The American Society of Plant Taxonomists, The American Philosophical Society and The National Science Foundation (Doctoral Dissertation Improvement Grant DEB-0808242).

REFERENCES

Allen, G.A., Marr, K.L., McCormick, L.J. & Hedba, R.J. (2012) The impact of Pleistocene climate change on an ancient arctic-alpine plant: multiple lineages of disparate history in *Oxyria digyna*. *Ecology and Evolution*, **2**, 649–665.

Artyukova, E.V., Kozyrenko, M.M., Kholina, A.B. & Zhuravlev, Y.N. (2011) High chloroplast haplotype diversity in the endemic legume *Oxytropis chankaensis* may result from independent polyploidization events. *Genetica*, **139**, 221–232.

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) 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. *Journal of Biogeography*, **38**, 1585–1599.

Beck, J.B., Nesom, G.L., Calie, P.J., Baird, G.I., Small, R.L. & Schilling, E.E. (2004) Is subtribe Solidagininae (Asteraceae) monophyletic? *Taxon*, **53**, 691–698.

Bennett, K.D., Tzedakis, P.C. & Willis, K.J. (1991) Quaternary refugia of north European trees. *Journal of Biogeography*, **18**, 103–115.

Borsch, T., Hilu, K.W., Quandt, D., Wilde, V., Neinhuis, C. & Barthlott, W. (2003) Noncoding plastid *trnT-trnF* sequences reveal a well resolved phylogeny of basal angiosperms. *Journal of Evolutionary Biology*, **16**, 558–576.

Braun, E.L. (1928) Glacial and post-glacial plant migrations indicated by relic colonies of southern Ohio. *Ecology*, **9**, 284–302.

Breen, A.L., Murray, D.F. & Olson, M.S. (2012) Genetic consequences of glacial survival: the late Quaternary history of balsam poplar (*Populus balsamifera* L.) in North America. *Journal of Biogeography*, **39**, 918–928.

Clement, M., Posada, D. & Crandall, K.A. (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1659.

Davis, M.B. (1983) Holocene vegetational history of the eastern United States. *Late-Quaternary environments of the United States* (ed. by H.E.J. Wright), pp. 166–181. University of Minnesota Press, Minneapolis, MN.

Dick, C.W., Bermingham, E., Lemes, M.R. & Gribel, R. (2007) Extreme long-distance dispersal of the lowland tropical rainforest tree *Ceiba pentandra* L. (Malvaceae) in Africa and the Neotropics. *Molecular Ecology*, **16**, 3039–3049.

Dobes, C.H., Mitchell-Olds, T. & Koch, M.A. (2004) Extensive chloroplast haplotype variation indicates Pleistocene hybridization and radiation of North American *Arabis drummondii*, *A.* × *divaricarpa*, and *A. holboellii* (Brassicaceae). *Molecular Ecology*, **13**, 349–370.

Dupanloup, I., Schneider, S. & Excoffier, L. (2002) A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology*, **11**, 2571–2581.

El Mousadik, A. & Petit, R.J. (1996) Chloroplast DNA phylogeography of the argan tree of Morocco. *Molecular Ecology*, **5**, 547–555.

van Els, P., Cicero, C. & Klicka, J. (2012) High latitudes and high genetic diversity: phylogeography of a widespread boreal bird, the gray jay (*Perisoreus canadensis*). *Molecular Phylogenetics and Evolution*, **63**, 456–465.

Emerson, K.J., Merz, C.R., Catchen, J.M., Hohenlohe, P.A., Cresko, W.A., Bradshaw, W.E. & Holzapfel, C.M. (2010) Resolving postglacial phylogeography using high-throughput sequencing. *Proceedings of the National Academy of Sciences USA*, **107**, 16196–16200.

Excoffier, L. & Lischer, H.E.L. (2010) Arlequin suite ver. 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.

Excoffier, L., Smouse, P.E. & Quattro, J.M. (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial-DNA restriction data. *Genetics*, **131**, 479–491.

Farris, J.S., Källersjö, M., Kluge, A.G. & Bult, C. (1995) Constructing a significance test for incongruence. *Systematic Biology*, **44**, 570–572.

Fernald, M.L. (1925) Persistence of plants in unglaciated areas of boreal America. *Memoirs of the American Academy of Arts and Sciences*, **15**, 241–342.

Griffin, P.C., Robin, C. & Hoffmann, A.A. (2011) A next-generation sequencing method for overcoming the multiple gene copy problem in polyploid phylogenetics, applied to *Poa* grasses. *BMC Biology*, **9**, 19.

Halverson, K., Heard, S.B., Nason, J.D. & Stireman, J.O. (2008) Origins, distribution, and local co-occurrence of polyploid cytotypes in *Solidago altissima* (Asteraceae). *American Journal of Botany*, **95**, 50–58.

Hardy, O.J. & Vekemans, X. (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, **2**, 618–620.

Harpending, H.C., Batzer, M.A., Gurven, M., Jorde, L.B., Rogers, A.R. & Sherry, S.T. (1998) Genetic traces of ancient demography. *Proceedings of the National Academy of Sciences USA*, **95**, 1961–1967.

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 re-colonization 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 of planet Earth* (ed. by L.J. Rothschild and A.M. Lister), pp. 339–361. Academic Press, London.

Hudson, M.E. (2008) Sequencing breakthroughs for genomic ecology and evolutionary biology. *Molecular Ecology Resources*, **8**, 3–17.

Huntley, B. & Webb, T., III (1989) Migration: species' response to climate variations caused by changes in the earth's orbit. *Journal of Biogeography*, **16**, 5–19.

Jakob, S.S. & Blattner, F.R. (2006) A chloroplast genealogy of *Hordeum* (Poaceae): long-term persisting haplotypes,

incomplete lineage sorting, regional extinction, and the consequences for phylogenetic inference. *Molecular Biology and Evolution*, **23**, 1602–1612.

Jaramillo-Correa, J.P., Beaulieu, J., Khasa, D.P. & Bousquet, J. (2009) Inferring the past from the present phylogeographic structure of North American forest trees: seeing the forest for the genes. *Canadian Journal of Forest Research*, **39**, 286–307.

Kelchner, S.A. (2000) The evolution of non-coding chloroplast DNA and its application in plant systematics. *Annals of the Missouri Botanical Garden*, **87**, 482–498.

Kiefer, C. & Koch, M.A. (2012) A continental-wide perspective: the genepool of nuclear encoded ribosomal DNA and single-copy gene sequences in North American *Boechera* (Brassicaceae). *PLoS ONE*, **7**, e36491.

Kiefer, C., Dobes, C., Sharbel, T.F. & Koch, M.A. (2009) Phylogeographic structure of the chloroplast DNA gene pool in North American *Boechera* – a genus and continental-wide perspective. *Molecular Phylogenetics and Evolution*, **52**, 303–311.

Kress, W.J., Wurdack, K.J., Zimmer, E.A., Weigt, L.A. & Janzen, D.H. (2005) Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences USA*, **102**, 8369–8374.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J. & Higgins, D.G. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**, 2947–2948.

Latham, R.E. & Ricklefs, R.E. (1993) Continental comparisons of temperate-zone tree species diversity. *Species diversity in ecological communities* (ed. by R.E. Ricklefs and D. Schluter), pp. 294–314. University of Chicago Press, Chicago.

Laureto, P.J. & Barkman, T.J. (2011) Nuclear and chloroplast DNA suggest a complex single origin for the threatened allopolyploid *Solidago houghtonii* (Asteraceae) involving reticulate evolution and introgression. *Systematic Botany*, **36**, 209–226.

Marie-Victorin, F. (1938) Phytogeographical problems of eastern Canada. *The American Midland Naturalist*, **19**, 489–558.

Marquis, R.J. & Voss, E.G. (1981) Distributions of some western North American plants disjunct in the Great Lakes region. *Michigan Botanist*, **20**, 53–82.

McLachlan, J.S., Clark, J.S. & Manos, P.S. (2005) Molecular indicators of tree migration capacity under rapid climate change. *Ecology*, **86**, 2088–2098.

Michl, T., Huck, S., Schmitt, T., Liebrich, A., Haase, P. & Büdel, B. (2010) The molecular population structure of the tall forb *Cicerbita alpina* (Asteraceae) supports the idea of cryptic glacial refugia in central Europe. *Botanical Journal of the Linnean Society*, **164**, 142–154.

Müller, K. (2006) Incorporating information from length-mutational events into phylogenetic analysis. *Molecular Phylogenetics and Evolution*, **38**, 667–676.

Parisod, C. (2012) Polyploids integrate genomic changes and ecological shifts. *New Phytologist*, **193**, 297–300.

Peirson, J.A. (2010) *Biogeography, ecology, and evolution of the endemic vascular flora of the glaciated Great Lakes region: a case study of the* Solidago simplex *species complex*. PhD Thesis, The University of Michigan, Ann Arbor, MI.

Peirson, J.A., Reznicek, A.A. & Semple, J.C. (2012) Polyploidy, infraspecific cytotype variation, and speciation in goldenrods: the cytogeography of *Solidago* subsect. *Humiles* (Asteraceae) in North America. *Taxon*, **61**, 197–210.

Petit, R.J., El Mousadik, A. & Pons, O. (1998) Identifying populations for conservation on the basis of genetic markers. *Conservation Biology*, **12**, 844–855.

Petit, R.J., Aguinagalde, I., Beaulieu, J.L., Bittkau, C., Brewer, S., Cheddadi, R., Ennos, R., Fineschi, S., Grivet, D., Lascoux, M., Mohanty, A., Muller-Starck, G., Demesure, B., Palme, A., Martin, J.P., Rendell, S. & Vendramin, G.G. (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. *Science*, **300**, 1563–1565.

Petit, R.J., Duminil, J., Fineschi, S., Hampe, A., Salvini, D. & Vendramin, G.G. (2005) Comparative organization of chloroplast, mitochondrial, and nuclear diversity in plant populations. *Molecular Ecology*, **14**, 689–701.

Pons, O. & Petit, R.J. (1996) Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics*, **144**, 1237–1245.

Prentice, I.C. (1986) Vegetation responses to past climate variation. *Vegetatio*, **67**, 131–141.

Ramsey, J. (2011) Polyploidy and ecological adaptation in wild yarrow. *Proceedings of the National Academy of Sciences USA*, **108**, 7096–7101.

Rebernig, C.A., Weiss-Schneeweiss, H., Schneeweiss, G.M., Schönswetter, P., Obermayer, R., Villaseñor, J.L. & Stuessy, T.F. (2010) Quaternary range dynamics and polyploid evolution in an arid brushland plant species (*Melampodium cinereum*, Asteraceae). *Molecular Phylogenetics and Evolution*, **54**, 594–606.

Ringius, G.S. & Semple, J.C. (1987) Cytogeography of the *Solidago spathulata–glutinosa* complex (Compositae, Astereae). *Canadian Journal of Botany*, **65**, 2458–2462.

Ronikier, M., Costa, A., Aguilar, J.F., Feliner, G.N., Küpfer, P. & Mirek, Z. (2008) Phylogeography of *Pulsatilla vernalis* (L.) Mill. (Ranunculaceae): chloroplast DNA reveals two evolutionary lineages across central Europe and Scandinavia. *Journal of Biogeography*, **35**, 1650–1664.

Rowe, K.C., Heske, E.J., Brown, P.W. & Paige, K.N. (2004) Surviving the ice: northern refugia and postglacial colonization. *Proceedings of the National Academy of Sciences USA*, **101**, 10355–10359.

Rozas, J., Sánchez-DelBarrio, J.C., Messequer, Z. & Rozas, R. (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, **19**, 2496–2497.

Saeki, I., Dick, C.W., Barnes, B.V. & Murakami, N. (2011) Comparative phylogeography of red maple (*Acer rubrum* L.) and silver maple (*Acer saccharinum* L.): impacts of

habitat specialization, hybridization, and glacial history. *Journal of Biogeography*, **38**, 1–14.

Schlaepfer, D.R., Edwards, P.J., Widmer, A. & Billeter, R. (2008) Phylogeography of native ploidy levels and invasive tetraploids of *Solidago gigantea*. *Molecular Ecology*, **17**, 5245–5256.

Schneider, S. & Excoffier, L. (1999) Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics*, **152**, 1079–1089.

Scotti-Saintagne, C., Dick, C.W., Caron, H., Guichoux, E., Vendramin, G.G., Buonanici, A., Duret, A., Sire, P., Valencia, R., Lemes, M.R., Gribel, R. & Scotti, I. (2013) Phylogeography of a species complex of lowland Neotropical rain forest trees (*Carapa*, Meliaceae). *Journal of Biogeography*, **40**, 676–692.

Semple, J.C. & Cook, R.E. (2006) Solidago. *Flora of North America: Asteraceae* (ed. by F.O.N.A.E. Committee), pp. 107–166. Oxford University Press, New York.

Shafer, A.B., Cullingham, C.I., Côté, S.D. & Coltman, D.W. (2010) Of glaciers and refugia: a decade of study sheds new light on the phylogeography of northwestern North America. *Molecular Ecology*, **19**, 4589–4621.

Simmons, M.P., Müller, K. & Norton, A.P. (2007) The relative performance of indel-coding methods in simulations. *Molecular Phylogenetics and Evolution*, **44**, 724–740.

Soltis, D.E., Morris, A.B., McLachlan, J.S., Manos, P.S. & Soltis, P.S. (2006) Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology*, **15**, 4261–4293.

Stebbins, G.L. (1974) *Flowering plants: evolution above the species level*. Belknap Press, Cambridge, MA.

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.

Symonds, V.V., Soltis, P.S. & Soltis, D.E. (2010) Dynamics of polyploid formation in *Tragopogon* (Asteraceae): recurrent formation, gene flow, and population structure. *Evolution*, **64**, 1984–2003.

Taberlet, P., Fumagalli, L., Wust-Saucy, A.G. & Cosson, J.F. (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**, 453–464.

Tajima, F. (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.

Williams, J.W., Webb, T., III, Richard, P.H. & Newby, P. (2000) Late Quaternary biomes of Canada and the eastern United States. *Journal of Biogeography*, **27**, 585–607.

Wróblewska, A. (2012) The role of disjunction and postglacial population expansion on phylogeographical history and genetic diversity of the circumboreal plant *Chamaedaphne calyculata*. *Biological Journal of the Linnean Society*, **105**, 761–775.

SUPPORTING INFORMATION

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

Appendix S1 Voucher and sampling information for *Solidago* subsect. *Humiles* populations used in this study.

Appendix S2 Species- and population-level haplotype distributions within *Solidago* subsect. *Humiles*.

Appendix S3 Fixation indices and number of singleton populations (Table S1) and geographical groupings from SAMOVA (Fig. S1).

BIOSKETCHES

Jess A. Peirson is a post-doctoral research fellow in the Department of Ecology and Evolutionary Biology at The University of Michigan. He is interested in the Holocene biogeography and evolution of the North American flora and the evolution and systematics of goldenrods (*Solidago*, Asteraceae) and spurges (*Euphorbia*, Euphorbiaceae).

Christopher W. Dick is an associate professor in the Department of Ecology and Evolutionary Biology at The University of Michigan. He is broadly interested in plant ecology and evolution, and has a longstanding interest in the biogeographical history of Neotropical forests.

Anton A. Reznicek is Curator of Vascular Plants at the University of Michigan Herbarium. His primary interests are in the systematics and evolution of sedges (especially *Carex*, Cyperaceae) and the biogeography, floristics, systematics and conservation of the Great Lakes Region flora.

Editor: Miles Silman