Journal of Biogeography (J. Biogeogr.) (2015) **42**, 339–350

ORIGINAL ARTICLE

A similar phylogeographical structure among sympatric North American birches (*Betula*) is better explained by introgression than by shared biogeographical history

Ashley M. Thomson[1,2]*, Christopher W. Dick[2] and Selvadurai Dayanandan[1]

[1]*Québec Centre for Biodiversity Science, Montréal and Department of Biology and Centre for Structural and Functional Genomics, Concordia University, Montreal, Quebec H4B 1R6, Canada,* [2]*Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109–1048, USA*

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

ABSTRACT

Aim A comparative analysis of the chloroplast (cp) DNA structure of eastern North American birches (*Betula*) was undertaken to infer the impacts of Quaternary climate change on the phylogeographical structure of these species.

Location Eastern North America.

Methods Genetic variation in chloroplast microsatellites (cpSSRs) and the *psb*A–*trn*H intergenic spacer in *Betula papyrifera*, *Betula alleghaniensis* and *Betula lenta* was analysed in samples from 65, 80 and 12 populations, respectively. Co-occurring *Betula uber*, *Betula populifolia* and *Betula cordifolia* were also sampled to examine haplotype relationships and account for potential introgression. Haplotype networks, Bayesian analysis and comparisons of R_{ST} and G_{ST} values were used to evaluate the phylogeographical structure. Genetic diversity within and among species was compared using rarefaction analysis.

Results The two most widespread species, *B. papyrifera* and *B. alleghaniensis*, showed high levels of haplotype diversity, while the Appalachian endemic *B. lenta* possessed a single haplotype. Bayesian analysis revealed three main phylogeographical groups for *B. papyrifera* and four groups for *B. alleghaniensis*, and these two species showed extensive regional haplotype sharing and a high introgression ratio.

Main conclusions We postulate that at least three separate refugia contributed to the recolonization of *B. papyrifera* and *B. alleghaniensis* within eastern North America, while *B. lenta* appears to have recolonized from a single refugium. A high haplotype diversity of *B. papyrifera* and *B. alleghaniensis* in the Great Lakes region may reflect biogeographical contact between eastern and western lineages, with the potential influence of periglacial refugia. Similar phylogeographical patterns in the distantly related *B. papyrifera* and *B. alleghaniensis* represent a geographical turnover of the same locally shared haplotypes, pointing to introgression rather than shared biogeographical history as the mechanism. Although similar phylogeographical patterns are often interpreted as evidence for common biogeographical histories, our study demonstrates that such patterns can also arise through introgression.

Keywords

***Betula*, birches, comparative phylogeography, chloroplast DNA, eastern North America, haplotype-sharing, hybridization, introgression, Quaternary climate change, refugia.**

INTRODUCTION

The genetic structure and diversity of temperate plant species are strongly influenced by vegetation changes associated with Quaternary glacial–interglacial cycles (Hewitt, 2000). The prevalent view is that unfavourable climatic conditions and habitat loss as a result of advancing ice sheets, particularly during the Last Glacial Maximum (LGM), truncated geographical ranges of species in the Northern Hemisphere and forced many temperate tree species to retreat to southern refugial locations (Bennett *et al.*, 1991; Taberlet *et al.*, 1998; Stewart *et al.*, 2010). As a result of long-term genetic isolation, different refugial populations are expected to have diverged from one another (Petit *et al.*, 2003a) and display a northwards decline in genetic diversity associated with founder events during recolonization (Hewitt, 2000). However, recent studies have provided evidence that this classical viewpoint may be overly simplistic, with species demonstrating complex and often individualistic geographical responses to past climatic fluctuations (Stewart *et al.*, 2010). For example, a number of European angiosperms possess a higher diversity at intermediate latitudes than at southern latitudes, presumably as a result of the mixing of lineages in refugial contact zones (see Petit *et al.*, 2003a, for a review). Moreover, increasing numbers of studies are reporting high genetic diversity or endemic haplotypes at relatively high northern latitudes, suggesting that numerous temperate plant species may have persisted within periglacial or cryptic northern refugia (Willis *et al.*, 2000; Stewart & Lister, 2001; Parducci *et al.*, 2012).

Comparative phylogeographical studies have been widely used to examine regional biogeography and infer the geographical distribution of evolutionary lineages, by comparing patterns of genetic variation among co-distributed taxa. While some such studies reveal concordant biogeographical patterns among closely related tree species (e.g. Premoli *et al.*, 2012), others reveal idiosyncratic patterns attributable to differing LGM distribution patterns or post-glacial recolonization dynamics (e.g. Heuertz *et al.*, 2006). A study of European birches, for example, revealed a greater number of demes within *Betula pubescens* compared with *Betula pendula,* which was attributed to the survival and earlier recolonization of *B. pubescens* at higher latitudes during the last glacial period (Maliouchenko *et al.*, 2007). Furthermore, comparative phylogeographical studies of closely related species often reveal introgression in the form of localized haplotype sharing, as has been documented in European *Quercus* (Petit *et al.*, 1993; Belahbib *et al.*, 2001). Such introgression can potentially alter species haplotype distributions through the introduction of novel haplotype variation, and should be accounted for when drawing inferences about species' biogeographical histories; for example, the finding of a distinct phylogeographical group for *Acer rubrum* in the lower Mississippi River Valley is best explained by introgression of haplotypes from *Acer saccharinum* rather than by historical biogeographical contingencies (Saeki *et al.*, 2011). To date, most comparative phylogeographical studies of North American trees have focused on widespread boreal conifers or trees of unglaciated regions (e.g. Soltis *et al.*, 2006; Jaramillo-Correa *et al.*, 2009) and there is a need for further comparative studies to understand better the glacial history of the cold-temperate angiosperms of eastern North America (but see Saeki *et al.*, 2011).

North American birches (*Betula*) are ideal for studies of comparative post-glacial biogeography. *Betula* comprises approximately 17 species in eastern North America, which share traits related to reproduction and dispersal but differ in their climatic tolerances and habitat preferences (Furlow, 1990). This study examined the phylogeography and glacial history of three *Betula* species of eastern North America: *Betula papyrifera* Marshall (paper birch), *Betula alleghaniensis* Britton (yellow birch) and *Betula lenta* L. (sweet birch). In addition, we sampled the less common species *Betula populifolia* Marshall, *Betula cordifolia* Regel and *Betula uber* (Ashe) Fernald to obtain a complete community-level sampling of *Betula* and account for potential introgression.

Betula papyrifera is the most widely distributed and cold-tolerant of the North American birches, with a distribution ranging from the boreal forest region of the USA and Canada to the northern limit of tree growth (Safford *et al.*, 1990). In comparison, *B. alleghaniensis* has a more southern distribution, occupying temperate forests from south-eastern Canada to New England and the southern Appalachians in the USA (Erdmann, 1990). *Betula lenta* is endemic to the Appalachian region of the eastern USA; it occurs from southern Maine to northern Alabama and Georgia (Lamson, 1990). *Betula populifolia* is confined to eastern Canada and the north-eastern USA, ranging from New Brunswick and Nova Scotia to Delaware, northern Pennsylvania and eastern Ontario (Furlow, 1997). *Betula cordifolia* occurs at elevations of 800–2000 m in the Canadian Maritimes and New England to Pennsylvania, North Carolina, Quebec, central Ontario, Minnesota and Michigan. *Betula uber* is represented by a single known population in south-west Virginia. A comparative summary of selected ecological and morphological characteristics of the six eastern North American *Betula* species examined here is presented in Appendix S1. The study species fall into the subgenera *Betulenta*, which contains species with generally dark, non-exfoliating bark (including *B. alleghaniensis*, *B. lenta* and *B. uber*), and *Betula*, which contains the white-barked birches (including *B. papyrifera*, *B. populifolia* and *B. cordifolia*) (DeJong, 1993). This morphological classification is supported by molecular studies that separate the white- and dark-barked birches into different clades (Jarvinen *et al.*, 2004; Li *et al.*, 2005; Schenk *et al.*, 2008).

Palaeobotanical reconstructions indicate the occurrence of two refugial *Betula* populations in eastern North America at the time of the LGM (*c.* 26.5 ka–19 ka), occupying regions of the mid-Atlantic coastal plain and north-central Louisiana (Delcourt & Delcourt, 1987; Jackson *et al.*, 1997). A third population appeared later, between 15 and 14 ka, in the American Midwest region. Starting around 18 ka, the

Atlantic coastal population began to expand westwards and southwards until it coalesced with the Midwest and Louisiana populations. From there, *Betula* populations expanded northwards following the retreat of the ice sheet. More detailed information regarding species-specific distributions has been obtained from macrofossils documenting the occurrence of *B. papyrifera* in northern Georgia at the time of the LGM. In contrast, *B. alleghaniensis* macrofossils are absent from all sites until its first appearance at 9 ka in New Hampshire (Jackson *et al.*, 2000)

We sampled chloroplast DNA (cpDNA) markers from our study taxa from across eastern North America. Our specific objectives were: (1) to examine patterns of cpDNA variation to test for individualistic or shared patterns of phylogeographical structure in *Betula*; (2) to combine genetic and fossil evidence to infer the LGM distribution and post-glacial recolonization routes for *Betula*; and (3) to explore the potential effects of historical species interactions, such as hybridization and introgression (*sensu* Saeki *et al.*, 2011), on the phylogeographical patterns.

MATERIALS AND METHODS

Population sampling

Leaf samples for DNA analyses were obtained from 1–10 individuals per species where present in each of 36 natural populations located throughout eastern Canada and the USA. Individuals separated by at least 50 m were sampled. Leaf tissue was dried in silica gel in the field. Additional samples were obtained from the seed collections of the Canadian National Tree Seed Centre (CTNSC) (Fredericton, Canada) and the British Columbia Tree Seed Centre (BCTSC) (Surrey, Canada). The seed collections spanned 75 localities, with *B. papyrifera* from 42 sites, *B. alleghaniensis* from 17 sites, *B. populifolia* from 10 sites and *B. cordifolia* from 10 sites. Because the seed collections consisted of a mix of single-tree and bulked seed, only a single seed per seed lot was used, in order to avoid genotyping multiple individuals from the same maternal tree. We also collected vegetative buds from a single individual of each of 36 provenances of *B. alleghaniensis* from a provenance trial located at the Kellogg Experimental Forest of Michigan State University (near Kellogg, MI, USA). In summary, the samples for this study represented 754 individuals from 123 populations of *Betula* from eastern North America and an additional 21 single-tree samples from the western distribution of *B. papyrifera* in British Columbia, Canada. Additional locality, sampling and voucher details are presented in Appendix S2.

Laboratory procedures

DNA was extracted from mature leaf tissues using the modified cetyltrimethylammonium bromide (CTAB) protocol of Zeng *et al.* (2002). Qiagen DNeasy extraction kits (Qiagen Inc., Valencia, CA) were used to extract DNA from vegetative buds and germinated seeds. The polymerase chain reaction (PCR) conditions and cpDNA primers used in this study are provided in Appendix S3. Haplotype variation was assessed using six chloroplast microsatellite (cpSSR) markers and the *psb*A–*trn*H intergenic spacer region. Three cpSSRs were developed for *Betula* in this study, while the remainder were obtained from the literature (Weising & Gardner, 1999). Amplification of the *psb*A–*trn*H spacer region was conducted using the primers of Sang *et al.* (1997), and PCR products were sequenced using an Applied Biosystems 3730xl DNA Analyzer (Life Technologies, Grand Island, NY, USA) Sequence chromatograms were edited and aligned using Geneious 4.8 software (Biomatters Ltd, Auckland, New Zealand). Indels were coded as single characters. We used a reduced within-population sampling intensity for sequencing compared with cpSSR genotyping because initial range-wide screening revealed that haplotypes were generally fixed across broad spatial scales. Individuals from all singleton populations were sequenced, while for multi-individual populations an average of three individuals per population was sequenced.

Data analyses

Haplotypes were defined as unique combinations of allele variants of the cpSSR and *psb*A–*trn*H sequence data. Mutational relationships among the haplotypes were inferred with a statistical parsimony network using the program tcs 1.21 (Clement *et al.*, 2000). Gaps were treated as the fifth state and the connection limit was set at eight steps. Haplotype frequencies were calculated using GenAlEx 6.41 (Peakall & Smouse, 2006) and haplotype distribution maps were constructed using ArcGIS 10.0 (ESRI, Redlands, CA, USA).

Estimates of standardized cpSSR haplotype richness were obtained for each species after rarefaction to the minimum sample size of $n = 16$ for *B. cordifolia* using contrib 1.02 (Petit *et al.*, 1998). *Betula uber* was excluded from the analysis because of its small sample size ($n = 10$) and because it possessed only one haplotype (see Results). A second calculation was made including only *B. papyrifera* and *B. alleghaniensis* with rarefaction to the minimum sample size of $n = 240$ in *B. papyrifera*. For these analyses, far western populations of *B. papyrifera* were excluded so that haplotype richness estimates included only eastern North American populations of each species. We also estimated standardized cpSSR haplotype richness for individual populations of *B. papyrifera* and *B. alleghaniensis*, considering only those populations with at least three sampled individuals. Haplotype richness values for each population were then mapped and interpolated using ArcGIS to visualize geographical patterns of cpSSR haplotype diversity within species.

Analysis of molecular variance (AMOVA) was used for assessing the partitioning cpSSR haplotype variation as implemented in Arlequin 3.1 (Excoffier *et al.*, 2005), treating each species individually as well as in species groupings. Only those populations with three or more sampled

individuals were considered (*B. papyrifera*, $n = 26$ populations; *B. alleghaniensis*, $n = 29$ populations; *B. lenta*, $n = 12$ populations) and the significance of the variance partitioning was calculated based on 10,000 random permutations of the data. We also tested for the presence of a phylogeographical structure of the cpSSR dataset by comparing permutated values of R_{ST} (synonymous with N_{ST} for microsatellite data) and G_{ST} using the program PermutCpSSR 2.0 (Pons & Petit, 1996), again considering only those populations with a minimum of three samples. To identify geographical patterns of genetic discontinuities, the population genetic structure for each species was examined using the Bayesian clustering approach implemented in baps software (Corander *et al.*, 2003, 2004). We conducted two separate sets of analyses in baps, first using the 'spatial clustering of groups' module, which considers the spatial location of populations in the determination of group structure, and second using the 'clustering of groups of individuals' module, which does not consider the geographical locations of populations. The model was run using the 'fixed K clustering' option, with individual values of K ranging from 1 to 10 and 20 replicates for each value of K. The solution that produced the highest log likelihood was selected as optimal. Corresponding among-group variation components (F_{CT} values) were calculated for the optimal baps groupings using AMOVA analysis in Arlequin software.

To examine the incidence of haplotype sharing between *B. papyrifera* and *B. alleghaniensis*, we calculated the introgression ratio (IG) (Belahbib *et al.*, 2001) and modified introgression ratio (IG*) (Thórsson *et al.*, 2010), which quantify the similarity of haplotype compositions among species at sites where they co-occur. We also compared introgression ratios calculated for pairs of populations of the same species at different sites to introgression ratios calculated between two species at shared sites (IGR) (Palme *et al.*, 2004). While IG and IG* do not distinguish whether haplotype sharing is the result of shared ancestry or introgression, significant negative correlations between IGR and geographical distance indicate the presence of isolation-by-distance resulting from local admixture. The significance of the correlations between IGR and geographical distance was tested using Mantel tests with 1000 permutations of the data. All statistical tests were conducted using R software (R Core Team, 2012). Haplotype sharing was not examined among other species because analyses showed that either they did not share haplotypes or because sample sizes were insufficient.

RESULTS

Haplotype distributions

The combined cpSSR and *psb*A–*trn*H sequence polymorphisms revealed 30 cpDNA haplotypes for the six North American *Betula* species (Fig. 1a). Haplotype definitions are presented in Appendix S3. In *B. papyrifera*, most of the eastern range, from central Ontario to New England and the Maritimes, was dominated by H8, H10 and related haplotypes (Fig. 1b). A second group composed of H17 and H19 was found along the Atlantic coast of Canada. *Betula papyrifera* populations in western Ontario and the Great Lakes states contained a mixture of endemic H22, H29 and related haplotypes, as well as a low frequency of eastern haplotypes (H8 and related haplotypes). Haplotypes H20, H24 and H25 were common in southern British Columbia, while H27 was common in northern British Columbia. The remaining haplotypes (H1, H3, H7, H11, H26, H28 and H30) were unique to single individuals.

Of 14 haplotypes recovered for *B. alleghaniensis*, eight were shared with *B. papyrifera* and had similar geographical distributions (Fig. 1c). Haplotypes H8, H10, H14 and related formed a large eastern group. H17 and H19 occurred along the Atlantic coast. Populations west of the Great Lakes were composed predominantly of H22 and H29, and H8 at a lower frequency. In, *B. alleghaniensis*, the cpDNA haplotype distribution revealed a separate southern Appalachian haplotype group (H9 and H15) that was not evident in *B. papyrifera*. Haplotypes H4, H5, H16, H18 and H19 were unique to single individuals. The haplotype distribution for *B. lenta* differed from that of *B. papyrifera* and *B. alleghaniensis*, as a single haplotype (H23) was found throughout its geographical range (Fig. 1d). CpSSR haplotype distributions for *B. populifolia* and *B. cordifolia* (Fig. 1e,f) were similar to those of sympatric *B. papyrifera* and *B. alleghaniensis,* with populations composed mostly of H8, H10 and closely related haplotypes. *Betula uber* (Fig. 1g) was fixed for the single haplotype (H23) that was found in *B. lenta*.

Eight of the 20 haplotypes found in *B. papyrifera* were shared with *B. alleghaniensis,* while four of the six haplotypes found in *B. populifolia* were shared with *B. papyrifera* and *B. alleghaniensis,* and two of the three haplotypes found in *B. cordifolia* were shared with *B. papyrifera* and *B. alleghaniensis*. Generally, the shared haplotypes had similar geographical distributions within the species in which they occurred; H8 and H10 were shared within the eastern study area, while H22 and H29 were shared within the western Great Lakes region, H6 and H14 within the central-eastern Great Lakes region, and H17 and H19 within the Atlantic coastal region.

Chloroplast DNA diversity

Standardized estimates after rarefaction ($n = 16$) revealed that cpSSR haplotype richness was similar for *B. papyrifera* (3.37) and *B. alleghaniensis* (3.33), and lower for *B. populifolia* (2.61) and *B. cordifolia* (1.00). Allelic richness for *B. lenta* was 0.00 because it possessed only one haplotype. When only eastern North American populations of *B. papyrifera* and *B. alleghaniensis* were considered, rarefaction analysis ($n = 240$) revealed a slightly higher haplotype richness for *B. papyrifera* (9.00) than *B. alleghaniensis* (7.46). Estimates of haplotype diversity were similar for *B. papyrifera* and *B. alleghaniensis,* with values of 0.61 and 0.60, respectively. Interpolated maps of cpSSR haplotype richness for *B. papyrifera* and *B.*

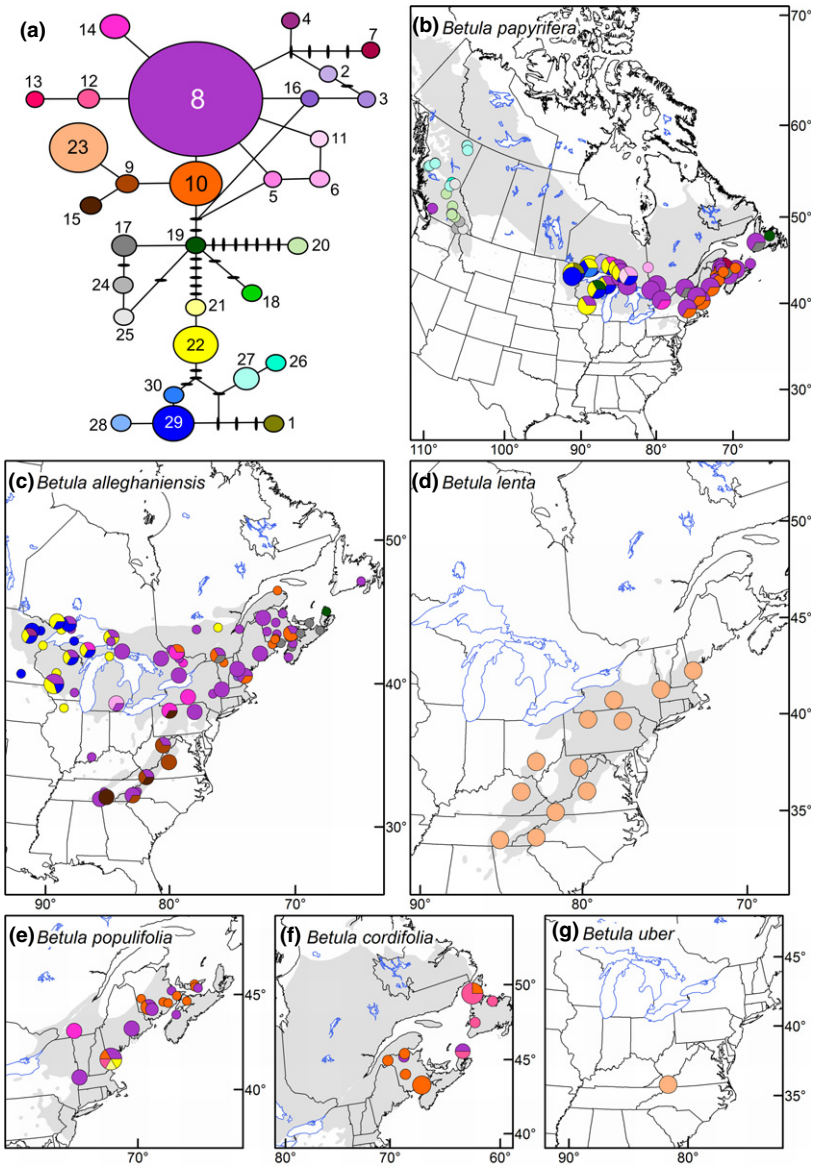

Figure 1 (a) Statistical parsimony network of 30 chloroplast (cp) DNA haplotypes from North American populations of *Betula* species. Coloured circles represent the observed haplotypes and the size of the circles is proportional to the haplotype frequency. Black ellipses on lines connecting haplotypes represent inferred haplotypes. The geographical distribution of cpDNA haplotypes was recovered for (b) *B. papyrifera*, (c) *B. alleghaniensis*, (d) *B. lenta*, (e) *B. populifolia*, (f) *B. cordifolia* and (g) *B. uber*.

alleghaniensis revealed that haplotype diversity was highest in the central-western Great Lakes region for both species (Fig. 2). There were also areas of moderate diversity located in the Canadian Maritimes range of *B. papyrifera* and in the Ontario/Quebec range of *B. alleghaniensis*.

Partitioning of molecular variation and phylogeographical structure

Analyses of molecular variance indicated a significant differentiation in cpSSR haplotype frequencies among populations of *B. papyrifera* (56.2%) and *B. alleghaniensis* (61.3%) (Table 1). Population genetic subdivision (G_{ST}) was strong in both species, but was somewhat higher for *B. alleghaniensis* (0.601) than *B. papyrifera* (0.526). In contrast, *B. lenta* did not demonstrate variation within or among populations. R_{ST} was significantly higher than G_{ST} in *B. papyrifera* (G_{ST} = 0.526, R_{ST} = 0.691, $P < 0.05$) and *B. alleghaniensis* (G_{ST} = 0.601, R_{ST} = 0.725, $P < 0.05$) indicating the presence of a phylogeographical structure in both species.

Marginal log-likelihood values and fixation indices obtained from the spatial and non-spatial BAPS analysis are available from the authors upon request. Spatial Bayesian clustering analysis of the combined cpSSR and *psb*A–*trn*H dataset revealed the optimal group structure for *B. papyrifera* when populations were grouped into five clusters (corresponding F_{CT} = 0.686). The largest group (group 1) was composed of populations occupying the eastern range of *B. papyrifera*, from New England and the Maritimes to the central Great Lakes region, and included a single disjunct population from southern British Columbia (Fig. 3a). Groups 2 and 3 were smaller and included populations from the central Great Lakes region and western Great Lakes region, respectively, while groups 4 and 5 included southern and

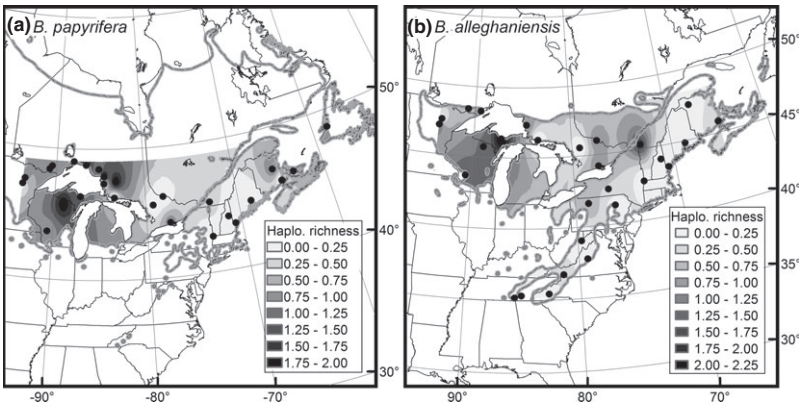

Figure 2 Mapped chloroplast microsatellite (cpSSR) haplotype richness for eastern North American populations of (a) *Betula papyrifera* and (b) *B. alleghaniensis*. Values were determined by interpolation of haplotype richness values for individual sampled populations using ArcGIS. Only those populations with a minimum of three sampled individuals were included in the analysis. The interpolated grid is clipped to the extent of the sampled populations to avoid extrapolation outside the study area. Other species were not analysed because they did not demonstrate cpSSR variation or because sample sizes were not sufficient.

Table 1 Analysis of molecular variance (AMOVA) and tests of phylogeographical structure (R_{ST} versus G_{ST}) for eastern North American *Betula papyrifera* and *B. alleghaniensis* based on chloroplast microsatellite (cpSSR) markers. Only populations with three or more samples were included in the analysis.

Species	Source of variation	d.f.	Variation (%)	*P*-value	G_{ST}	R_{ST}	$R_{ST} > G_{ST}$ *P*-value
B. papyrifera	Among populations	25	56.2	< 0.001	0.526	0.691	0.012
	Within populations	196	43.8		–	–	–
B. alleghaniensis	Among populations	28	61.3	< 0.001	0.601	0.725	<0.001
	Within populations	232	38.7		–	–	–

G_{ST}, measure of differentiation that makes use only of allelic frequency differences among populations.
R_{ST}, measure of differentiation for which genetic distances as a result of differing microsatellite repeat numbers are taken into account.
$R_{ST} > G_{ST}$, *P*-value for test of phylogeographic structure.

northern British Columbia populations, respectively. For *B. alleghaniensis* the optimal spatial group structure occurred for $K = 4$ (corresponding $F_{CT} = 0.764$). The largest group (group 1) spanned most of the eastern distribution of *B. alleghaniensis* from the central Great Lakes region to the western Maritimes, New England, and south to the Appalachians (Fig. 3b). Group 2 was smaller and more regionally restricted, consisting of several populations from the eastern Maritimes, while group 3 included the westernmost *B. alleghaniensis* populations from the central and western Great Lakes region. In comparison, population clusters identified by BAPS without consideration of spatial reference information resulted in a greater number of population groups for both *B. papyrifera* ($K = 7$; Fig. 3c) and *B. alleghaniensis* ($K = 5$; Fig. 3d) and somewhat differing geographical patterns. Specifically, the non-spatial clustering for *B. papyrifera* included an additional population group in the north-central Great Lakes region and an additional group in southern British Columbia. For *B. alleghaniensis*, the non-spatial BAPS analysis recognized Appalachian populations containing endemic haplotypes (H9 and H15) as belonging to a distinct cluster, whereas these populations were not recognized as distinct based on the spatial BAPS analysis.

Partitioning of cpDNA variation among species

When *B. papyrifera*, *B. alleghaniensis* and *B. lenta* were considered jointly, AMOVA revealed that the majority of cpSSR haplotype variation was partitioned among species (40.7%), while 35.2% of the variation occurred among populations within species and 24.1% of the variation occurred within populations (Table 2). However, variation among the polyploid species *B. papyrifera* and *B. alleghaniensis* was low (0.3%) and non-significant. When *B. papyrifera* and *B. alleghaniensis* were analysed at the 17 sites at which they co-occurred, a total of eight cpSSR haplotypes was found and five of those were shared between them. The IG ratio calculated for *B. papyrifera* and *B. alleghaniensis* across the 17 pooled sites was 0.98, while IG ratios for individual sites ranged from 0.36 to 1.12, with an across-site average of 0.89, indicating that the two species generally had similar haplotype frequencies on sites where they co-occurred (data are available from the authors upon request). Correlations of IGR with geographical distance were significant and negative, indicating the occurrence of isolation-by-distance between populations of *B. papyrifera* and *B. alleghaniensis* with *B. papyrifera* as the focal species ($r = –0.458$,

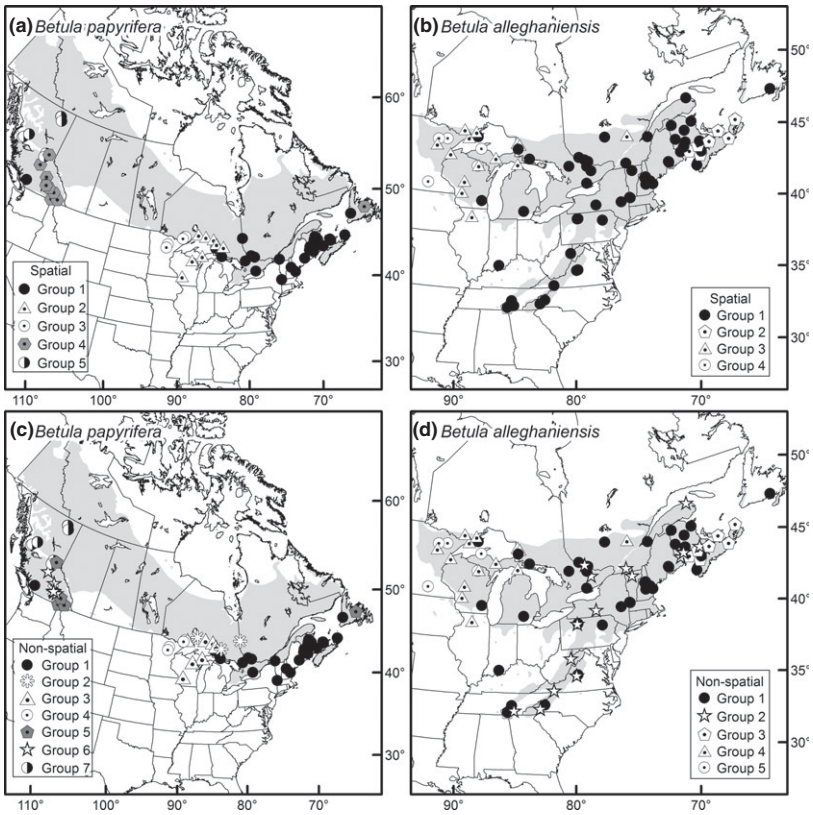

Figure 3 Geographical groupings from BAPS spatial clustering analysis of eastern North American *Betula* populations: (a) *B. papyrifera* (K = 5, F_{CT} = 0.686) and (b) *B. alleghaniensis* (K = 4, F_{CT} = 0.764). Geographical groupings from BAPS non-spatial clustering analysis of North American *Betula* populations: (c) *B. papyrifera* (K = 7, F_{CT} = 0.751) and (d) *B. alleghaniensis* (K = 5, F_{CT} = 0.727). The geographical range of each species is indicated by light grey shading. Other species were not analysed because they did not demonstrate cpSSR variation or because sample sizes were not sufficient.

$P < 0.001$) and with *B. alleghaniensis* as the focal species (r = –0.558, $P < 0.001$).

DISCUSSION

Phylogeographical patterns and Pleistocene biogeography

This study has revealed generally similar spatial patterns of haplotype variation among species despite some differences in haplotype compositions. Haplotype distributions and BAPS analyses indicated at least three separate glacial lineages for *Betula* in eastern North America: (1) a large eastern group occurring from the southern Appalachians to New England and the Maritimes, Quebec and eastern Ontario; (2) a western group located in the Great Lakes region; and (3) a geographically restricted eastern group overlapping with the large eastern group along the Atlantic coast of Canada and the northern USA. These groups mostly correspond well with the LGM fossil pollen and macrofossil reconstructions for *Betula* in eastern North America that indicate the presence of low-density *Betula* populations in the mid-Atlantic coastal

Table 2 Analysis of molecular variance among eastern North American *Betula* species based on chloroplast microsatellite (cpSSR) haplotype frequencies. Only populations with three or more samples were included in the analysis.

Species grouping	Source of variation	d.f.	Variation (%)	Fixation indices	*P*-value
B. papyrifera/ *B. alleghaniensis*/ *B. lenta*	Among species	2	40.7	F_{CT} = 0.40710	< 0.001
	Among populations	64	35.2	F_{SC} = 0.59289	< 0.001
	Within populations	536	24.1	F_{ST} = 0.75862	< 0.001
B. papyrifera/ *B. alleghaniensis*	Among species	1	0.3	F_{CT} = 0.00324	0.3287
	Among populations	53	58.7	F_{SC} = 0.58919	< 0.001
	Within populations	428	41.0	F_{ST} = 0.59052	< 0.001
B. papyrifera/ *B. lenta*	Among species	1	61.8	F_{CT} = 0.61821	< 0.001
	Among populations	36	21.7	F_{SC} = 0.56943	< 0.001
	Within populations	304	16.4	F_{ST} = 0.83561	< 0.001
B. alleghaniensis/ *B. lenta*	Among species	1	62.3	F_{CT} = 0.62250	< 0.001
	Among populations	39	23.4	F_{SC} = 0.61842	< 0.001
	Within populations	340	14.4	F_{ST} = 0.85595	< 0.001

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

plain, northern Georgia, the Midwest and north-central Louisiana between 18 and 15 ka (Delcourt & Delcourt, 1987; Jackson *et al.*, 1997, 2000).

One large haplotype group, consisting of H8 and related haplotypes, predominated across much of the eastern range of *B. papyrifera* and *B. alleghaniensis*, and haplotypes of this group were also frequent within sympatric populations of *B. populifolia* and *B. cordifolia*. This haplotype group corresponds geographically to group 1 identified by BAPS spatial analysis for *B. alleghaniensis* and *B. papyrifera,* containing populations from the southern Appalachians to New England, the Maritimes and into eastern Ontario. We suggest that this distinct eastern lineage originated from an Appalachian or mid-Atlantic coastal refugium, as fossil pollen and macrofossil reconstructions indicate the presence of *Betula* populations in the mid-Atlantic coastal plain and southern Appalachia at the time of the LGM (Delcourt & Delcourt, 1987; Jackson *et al.*, 1997, 2000). In *B. alleghaniensis*, a group of endemic haplotypes (H9 and H15) was found in the southern Appalachians. These haplotypes have a close phylogenetic relationship with the sympatric and widespread haplotype H8. Populations containing this southern Appalachian haplotype group were not recognized as a distinct cluster based on the spatial BAPS analysis, but were grouped with other populations of mixed haplotype composition in the non-spatial BAPS analysis. Although this pattern could arise from genetic drift, i.e. allele-surfing during post-glacial recolonization (Excoffier *et al.*, 2009), it could indicate that the Appalachian region was colonized from two or more eastern refugial populations, as has been suggested for several co-distributed species, including hemlock (*Tsuga canadensis*) (Lemieux *et al.*, 2011), *Narceus* millipedes (Walker *et al.*, 2009) and spotted salamanders (*Ambystoma maculatum*) (Zamudio & Savage, 2003). Our analyses identified a distinct western lineage containing haplotypes mostly endemic to the Great Lakes region (H22 and H29). This lineage corresponds geographically with spatial BAPS groups 2 and 3 for *B. papyrifera* and groups 3 and 4 for *B. alleghaniensis*. The presence of endemic haplotypes within this region suggests that the western lineage recolonized from a distinct refugium located west of the Appalachians. Currently, the refugial origins of this lineage are unclear, as potential western refugial locations fall outside the current distribution of *B. papyrifera* and *B. alleghaniensis* and representative samples could not be obtained. We postulate that the western lineage may have originated from a Mississippian refugium, or from periglacial refugia in mid-western USA, which have been identified as sources of recolonization of Great Lakes populations of other eastern North American plant species (McLachlan *et al.*, 2005; de Lafontaine *et al.*, 2010; Peirson *et al.*, 2013). Notably, fossil pollen data for *Betula* suggests the existence of low-density populations in north-central Louisiana at 18,000 yr BP (Delcourt & Delcourt, 1987) and in the Midwest between 15,000 and 14,000 yr BP (Jackson *et al.*, 1997).

A pair of closely related and regionally restricted cpDNA haplotypes (H17 and H19) was recovered along the Atlantic coastal region in both *B. papyrifera* and *B. alleghaniensis*. Populations from that region formed a distinct genetic cluster corresponding to BAPS spatial group 2 for *B. alleghaniensis*. The presence of these endemic haplotypes suggests that this area may have been populated from a refugium located in or near this area. Although there is little fossil data to support the existence of an Atlantic refugium, endemic haplotypes of jack pine (*Pinus banksiana*) and black spruce (*Picea mariana*) have been found within the same area (Jaramillo-Correa *et al.*, 2004, 2009; Godbout *et al.*, 2005; Gérardi *et al.*, 2010). Godbout *et al.* (2005) suggest that Atlantic populations may have recolonized from a cryptic refugium located on an exposed Atlantic shelf. Together, these studies suggest that a cryptic Atlantic refugium might have harboured low-density populations of *Betula* and other boreal–temperate plant species at the time of the LGM.

Although it is sometimes postulated that recently colonized areas should harbour reduced levels of genetic diversity as a result of founder effects during recolonization (e.g. Hewitt, 2000), in the present study the areas of highest genetic diversity for *Betula* were observed in previously glaciated areas. Populations of the central-western Great Lakes region harboured the highest levels of genetic diversity for both *B. papyrifera* and *B. alleghaniensis.* These populations contained haplotypes characteristic of both the western (H22 and H29) and eastern (H8) lineages, and thus the Great Lakes region appears to represent a zone of biogeographical contact between eastern and western glacial lineages. It is also possible that the diversity of the central-western Great Lakes region may be partially attributable to introgression, considering that this region represents a zone of natural hybridization between *B. papyrifera* and *B. alleghaniensis* (Barnes *et al.*, 1974).

Comparison of phylogeographical patterns among species

Betula papyrifera and *B. alleghaniensis* both possessed relatively high levels of haplotype diversity and displayed similar phylogeographical patterns within eastern North America. In comparison, *B. lenta* was the outlier in that it was fixed for a single haplotype and lacked phylogeographical structure. The finding of just one haplotype in *B. lenta* indicates that the current distribution may derive from a single glacial lineage, which presumably would have originated from the Appalachian region. Thus the low cpDNA diversity of *B. lenta* might be attributed to a restricted LGM range and associated low population size leading to haplotype fixation through genetic drift. Indeed, low levels of cpDNA diversity found in a number of boreal–temperate trees are attributed to small LGM population sizes (Demesure *et al.*, 1996; Echt *et al.*, 1998; Vendramin *et al.*, 2008). In contrast, genetic structure analyses for *B. papyrifera* and *B. alleghaniensis* suggest that these species may derive from at least three separate eastern North American lineages. Thus we suggest that the relatively high haplotype diversity of *B. papyrifera* and *B. alleghaniensis*

may be attributed to relatively broad LGM distributions translating into larger population sizes and greater population connectivity in the glacial and post-glacial landscapes.

This study uncovered regionally structured haplotype sharing between *B. papyrifera* and *B. alleghaniensis*, which also shared haplotypes with *B. populifolia* and *B. cordifolia* throughout their sympatric distribution. Similar patterns of local haplotype sharing have been documented in numerous plant groups and are often attributed to chloroplast capture as a result of historic hybridization and introgression within shared glacial refugia or during post-glacial expansion (e.g. Prentice *et al.*, 2008). Under this scenario, only a few natural hybrids need to become established to have a large impact on species allele frequencies, if hybridization is followed by repeated backcrossing (Rieseberg *et al.*, 1995; Petit *et al.*, 2003b). In the present study, the finding of regionally shared haplotypes among *B. papyrifera* and *B. alleghaniensis* might also indicate chloroplast capture as a result of past hybridization and introgression. Other potential sources of haplotype sharing, including shared ancestry and homoplasy, must also be considered, but are unlikely to explain the observed regional haplotype sharing because they are expected to result in a random distribution of shared haplotypes across a geographical range (Palme *et al.*, 2004). In contrast, haplotype sharing as a result of introgression should result in similar haplotype composition among species within geographical regions, as was found for *B. papyrifera* and *B. alleghaniensis* in this study (IG ratio 0.89). While haplotypes H8 and H10 are probably the oldest haplotypes and might therefore be shared through ancestral polymorphism, haplotypes H6, H14 and H19 are less frequent and peripheral, suggesting that they are unlikely to be shared through common ancestry and thus might be shared as a result of introgression. Also, while shared ancestry should result in a widespread distribution of shared haplotypes or shared haplotypes that occur in different geographical regions for different species (Palme *et al.*, 2004), all of the haplotypes shared among *Betula* species in this study had restricted geographical distributions that were concordant among species (with the exception of H8, which was also found in a single western North American population of *B. papyrifera*).

Additional evidence suggesting that haplotype sharing might be the result of introgression is the lack of significant differentiation between haplotype frequencies of *B. papyrifera* and *B. alleghaniensis*, which could suggest a lack of reproductive isolation between the two species. The regression of IGR values against geographical distance revealed significant isolation-by-distance between population pairs of *B. papyrifera* and *B. alleghaniensis,* which suggests the occurrence of local gene flow. Moreover, IGR values were on average less than 1, indicating that haplotype compositions were more strongly influenced by geography than by species identity. Together, these results suggest an important role of introgression in influencing the phylogeographical structure and haplotype sharing of *B. papyrifera* and *B. alleghaniensis*, similar to that reported for European birches (Palme *et al.*, 2004; Thórsson *et al.*, 2010).

In the Appalachian region, *B. lenta* and *B. uber* shared H23, for which they were both fixed, while *B. lenta* did not share haplotypes with the remaining *Betula* species. We suggest that shared ancestry is potentially responsible for the haplotype sharing of *B. lenta* and *B. uber*, as morphological and molecular studies indicate a close phylogenetic relationship between the two species (Sharik & Ford, 1984; Li *et al.*, 2005) and it has also been suggested that *B. uber* represents a variant of *B. lenta* (Sharik, 1990).

Differing ploidy numbers among the *Betula* species in this study may play a role in influencing patterns of introgression, as differing ploidy levels are expected to act as a major barrier to gene flow among species (Hersch-Green, 2012). In contrast, gene flow is more likely to occur when species have the same number of chromosomes. Thus we expect that the apparent cpDNA introgression observed between *B. papyrifera* and *B. alleghaniensis* is facilitated by similar chromosome numbers (*B. papyrifera* = 5×, 6×; *B. alleghaniensis* = 6×), whereas cpDNA introgression between diploid *B. lenta* and polyploids *B. alleghaniensis* and *B. papyrifera* could be impeded by different chromosome numbers and associated post-zygotic incompatibilities.

CONCLUSIONS

The high level of chloroplast haplotype diversity observed in *B. papyrifera* and *B. alleghaniensis* in this study could be attributed to their relatively broad LGM distributions translating into relatively large population sizes in the glacial and post-glacial landscape. In contrast, *B. lenta* was fixed for one haplotype, suggesting that a restricted LGM distribution contributed to population bottlenecks. The high haplotype diversity of *B. papyrifera* and *B. alleghaniensis* of the central-western Great Lakes region may be the result of biogeographical contact between eastern and western glacial lineages but introgression could also contribute to the haplotype diversity of this region. *Betula papyrifera* and *B. alleghaniensis* shared local haplotypes, suggesting past introgression within glacial refugia or during post-glacial migration, whereas *B. lenta* did not share haplotypes with the polyploid birches. The inclusion of *B. populifolia* and *B. cordifolia* revealed that these species also share localized haplotypes with *B. papyrifera* and *B. alleghaniensis* across their sympatric distribution, while *B. uber* shared its single haplotype with *B. lenta*. Such haplotype sharing across multiple co-distributed species suggests that hybridization and introgression within LGM habitats and during post-glacial expansion has probably played an important role in shaping the patterns of phylogeographical structure of eastern North American birches.

ACKNOWLEDGEMENTS

We thank three anonymous referees for their insightful comments and suggestions for improving the manuscript. We thank the many personnel associated with the state parks, national forests and experimental forests of the eastern United

States who provided assistance in locating sites, obtaining sample collections and providing directions and lodging for the various field sites visited as part of this study. We thank Fred Huber of the George-Jefferson National Forest for assisting with *B. uber* sampling. We are grateful to Dale Simpson and Bernard Daigle of the CNTSC, as well as Dave Kolotelo of the BCTSC for providing seed collections. We also thank Greg Kowawleski and Paul Bloese of Michigan State University for assistance in obtaining yellow birch provenance samples and Burton Barnes of the University of Michigan, whose expertise regarding eastern North American birches was instrumental in this study. We thank the various organizations that provided funding, including the National Research Council of Canada, who provided support in the form of a Canada Graduate Scholarship and Michael Smith Foreign Study supplement to A.M.T., a Discovery Grant to S.D., and the US National Science Foundation grant (DEB 0640379) to C.W.D. Additional support was provided by Concordia University, the Quebec Centre for Biodiversity Research and the Quebec Fonds de Recherche Nature et Technologies.

REFERENCES

Barnes, B.V., Dancik, B.P. & Sharik, T.L. (1974) Natural hybridization of yellow birch and paper birch. *Forest Science*, **20**, 215–221.

Belahbib, N., Pemonge, M.H., Ouassou, A., Sbay, H., Kremer, A. & Petit, R.J. (2001) Frequent cytoplasmic exchanges between oak species that are not closely related: *Quercus suber* and *Q. ilex* in Morocco. *Molecular Ecology*, **10**, 2003–2012.

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

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

Corander, J., Waldmann, P. & Sillanpää, M.J. (2003) Bayesian analysis of genetic differentiation between populations. *Genetics*, **163**, 367–374.

Corander, J., Waldmann, P., Marttinen, P. & Sillanpää, M.J. (2004) BAPS 2: enhanced possibilities for the analysis of genetic population structure. *Bioinformatics*, **20**, 2363–2369.

DeJong, P.C. (1993) An introduction to *Betula*: its morphology, evolution, classification and distribution, with a survey of recent work. *Betula: Proceedings of the IDS Betula Symposium* (ed. by D. Hunt), pp. 7–18. Richmond, UK.

Delcourt, P.A. & Delcourt, H.R. (1987) *Long-term forest dynamics of the temperate zone: a case study of late-Quaternary forests in eastern North America*. Springer-Verlag, New York.

Demesure, B., Comps, B. & Petit, R.J. (1996) Chloroplast DNA phylogeography of the common beech (*Fagus sylvatica* L.) in Europe. *Evolution*, **50**, 2515–2520.

Echt, C.S., DeVerno, L.L., Anzidei, M. & Vendramin, G.G. (1998) Chloroplast microsatellites reveal population genetic diversity in red pine, *Pinus resinosa* Ait. *Molecular Ecology*, **7**, 307–316.

Erdmann, G.G. (1990) Yellow birch (*Betula alleghaniensis* Britton). *Silvics of North America,* Vol. 2, *Hardwoods* (ed. by R.M. Burns and B.H. Honkola), pp. 133–147. Agriculture Handbook No. 654, United States Department of Agriculture, Washington, DC.

Excoffier, L., Laval, G. & Schneider, S. (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.

Excoffier, L., Foll, M. & Petit, R.J. (2009) Genetic consequences of range expansions. *Annual Review of Ecology and Systematics*, **40**, 481–501.

Furlow, J.J. (1990) The genera of Betulaceae in the southeastern United States. *Journal of the Arnold Arboretum*, **71**, 1–67.

Furlow, J.J. (1997) Betulaceae. *Flora of North America north of Mexico,* Vol. 3 (ed. by Flora of North America Editorial Committee), pp. 507–538. Oxford University Press, New York.

Gérardi, S., Jaramillo-Correa, J.P., Beaulieu, J. & Bousquet, J. (2010) From glacial refugia to modern populations: new assemblages of organelle genomes generated by differential cytoplasmic gene flow in transcontinental black spruce. *Molecular Ecology*, **19**, 5265–5280.

Godbout, J., Jaramillo-Correa, J.P., Beaulieu, J. & Bousquet, J. (2005) A mitochondrial DNA minisatellite reveals the postglacial history of jack pine (*Pinus banksiana*), a broad-range North American conifer. *Molecular Ecology*, **14**, 3497–3512.

Hersch-Green, E.I. (2012) Polyploidy in Indian paintbrush (*Castilleja*; Orobanchaceae) species shapes but does not prevent gene flow across species boundaries. *American Journal of Botany*, **99**, 1680–1690.

Heuertz, M., Carnevale, S., Fineschi, S., Sebastiani, F., Hausman, J.F., Paule, L. & Vendramin, G.G. (2006) Chloroplast DNA phylogeography of European ashes, *Fraxinus* sp. (Oleaceae): roles of hybridization and life history traits. *Molecular Ecology*, **15**, 2131–2140.

Hewitt, G. (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.

Jackson, S.T., Overpeck, J.T., Webb, T. III, Keattch, S.E. & Anderson, K.H. (1997) Mapped plant-macrofossil and pollen records of late Quaternary vegetation change in eastern North America. *Quaternary Science Reviews*, **16**, 1–70.

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. *Science*, **19**, 489–508.

Jaramillo-Correa, J.P., Beaulieu, J. & Bousquet, 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.

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.

Jarvinen, P., Palme, A., Morales, L.O., Lannenpaa, M., Keinanen, M., Sopanen, T. & Lascoux, M. (2004) Phylogenetic relationships of *Betula* species (Betulaceae) based on nuclear *ADH* and chloroplast *matK* sequences. *American Journal of Botany*, **91**, 1834–1845.

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.

Lamson, N.I. (1990) Sweet birch (*Betula lenta* L.). *Silvics of North America,* Vol. 2, *Hardwoods* (ed. by R.M. Burns and B.H. Honkola), pp. 148–152. Agriculture Handbook No. 654, United States Department of Agriculture, Washington, DC.

Lemieux, M.J., Beaulieu, J. & Bousquet, J. (2011) Chloroplast DNA polymorphisms in eastern hemlock: range-wide genogeographic analyses and implications for gene conservation. *Canadian Journal of Forest Research*, **41**, 1047–1059.

Li, J., Shoup, S. & Chen, Z. (2005) Phylogenetics of *Betula* (Betulaceae) inferred from sequences of nuclear ribosomal DNA. *Rhodora*, **107**, 69–86.

Maliouchenko, O., Palmé, A.E., Buonamici, A., Vendramin, G.G. & Lascoux, M. (2007) Comparative phylogeography and population structure of European *Betula* species, with particular focus on *B. pendula* and *B. pubescens*. *Journal of Biogeography*, **34**, 1601–1610.

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

Palme, A.E., Su, Q., Palsson, S. & Lascoux, M. (2004) Extensive sharing of chloroplast haplotypes among European birches indicates hybridization among *Betula pendula*, *B. pubescens* and *B. nana*. *Molecular Ecology*, **13**, 167–178.

Parducci, L., Jørgensen, T., Tollefsrud, M.M. *et al.* (2012) Glacial survival of boreal trees in northern Scandinavia. *Science*, **335**, 1083–1086.

Peakall, R. & Smouse, P.E. (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288–295.

Peirson, J.A., Dick, C.W. & Reznicek, A.A. (2013) Phylogeography and polyploid evolution of North American goldenrods (*Solidago* subsect. *Humiles*, Asteraceae). *Journal of Biogeography*, **40**, 1887–1898.

Petit, R.J., Kremer, A. & Wagner, D.B. (1993) Geographic structure of chloroplast DNA polymorphisms in European oaks. *Theoretical and Applied Genetics*, **87**, 122–128.

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

Petit, R.J., Bodenes, C., Ducousso, A., Roussel, G. & Kremer, A. (2003b) Hybridization as a mechanism of invasion in oaks. *New Phytologist*, **161**, 151–164.

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

Premoli, A.C., Mathiasen, P., Acosta, M.C. & Ramos, V.A. (2012) Phylogeographically concordant chloroplast DNA divergence in sympatric *Nothofagus* s.s. How deep can it be? *New Phytologist*, **193**, 261–275.

Prentice, H.C., Malm, J.U. & Hathaway, L. (2008) Chloroplast DNA variation in the European herb *Silene dioica* (red campion): postglacial migration and interspecific introgression. *Plant Systematics and Evolution*, **272**, 23–37.

R Core Team (2012) *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna.

Rieseberg, L.H., Desrochers, A.M. & Youn, S.J. (1995) Interspecific pollen competition as a reproductive barrier between sympatric species of *Helianthus* (Asteraceae). *American Journal of Botany*, **82**, 515–519.

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**, 992–1005.

Safford, L.O., Bjorkbom, J.C. & Zasada, J.C. (1990) Paper birch (*Betula papyrifera* Marsh.). *Silvics of North America,* Vol. 2, *Hardwoods* (ed. by R.M. Burns and B.H. Honkola), pp. 158–171. Agriculture Handbook No. 654, United States Department of Agriculture, Washington, DC.

Sang, T., Crawford, D. & Stuessy, T.F. (1997) Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). *American Journal of Botany*, **84**, 1120–1136.

Schenk, M.F., Thienpont, C.-L., Koopman, W.J.M., Gilissen, L.J.W.J. & Smulders, M.J.M. (2008) Phylogenetic relationships in *Betula* (Betulaceae) based on AFLP markers. *Tree Genetics & Genomes*, **4**, 911–924.

Sharik, T.L. (1990) *Virginia round-leaf birch recovery plan*. United States Fish and Wildlife Service, Newton Corner, MA.

Sharik, T.L. & Ford, R.H. (1984) Variation and taxonomy of *Betula uber*, *B. lenta*, and *B. alleghaniensis*. *Brittonia*, **36**, 307–316.

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.

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.

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

Thórsson, A.T., Pálsson, S., Lascoux, M. & Anamthawat-Jónsson, K. (2010) Introgression and phylogeography of *Betula nana* (diploid), *B. pubescens* (tetraploid) and their triploid hybrids in Iceland inferred from cpDNA haplotype variation. *Journal of Biogeography*, **37**, 2098–2110.

Vendramin, G.G., Fady, B., González-Martínez, S.C., Hu, F.S., Scotti, I., Sebastiani, F., Soto, A. & Petit, R.J. (2008) Genetically depauperate but widespread: the case of an emblematic Mediterranean pine. *Evolution*, **62**, 680–688.

Walker, M.J., Stockman, A.K., Marek, P.E. & Bond, J.E. (2009) Pleistocene glacial refugia across the Appalachian Mountains and coastal plain in the millipede genus *Narceus*: evidence from population genetic, phylogeographic, and paleoclimatic data. *BMC Evolutionary Biology*, **9**, 25.

Weising, K. & Gardner, R.C. (1999) A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. *Genome*, **42**, 9–19.

Willis, K.J., Rudner, E. & Sümegi, P. (2000) The full-glacial forests of central and southeastern Europe. *Quaternary Research*, **53**, 203–213.

Zamudio, K.R. & Savage, W.K. (2003) Historical isolation, range expansion, and secondary contact of two highly divergent mitochondrial lineages in spotted salamanders (*Ambystoma maculatum*). *Evolution*, **57**, 1631–1652.

Zeng, J., Zou, Y.-P., Bai, J.-Y. & Zheng, H.-S. (2002) Preparation of total DNA from "recalcitrant plant taxa". *Acta Botanica Sinica*, **44**, 694–697.

SUPPORTING INFORMATION

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

Appendix S1 A summary of the life-history traits of six eastern North American *Betula* species.

Appendix S2 Location information and GenBank accession numbers for the eastern North American *Betula* population samples.

Appendix S3 Primer information, thermocycling procedures and haplotype definitions.

BIOSKETCH

Ashley M. Thomson is a post-doctoral research fellow in the Department of Ecology and Evolutionary Biology at the University of Michigan. She is interested in the biogeography, population genetics and patterns of adaptive variation of boreal–temperate trees and the systematics and evolution of the eastern North American birches (*Betula*, Betulaceae).

Author contributions: A.M.T., C.W.D. and S.D. designed the project; A.M.T. collected and analysed the data; A.M.T. led the writing; and C.W.D. and S.D. contributed to the writing.

Editor: Mark Carine