CrossMark

ORIGINAL ARTICLE

Despite introgressive hybridization, North American birches (*Betula* spp.) maintain strong differentiation at nuclear microsatellite loci

Ashley M. Thomson[1,2,3] · Christopher W. Dick[3] · Ana L. Pascoini[4] ·
Selvadurai Dayanandan[1,2]

Abstract Extensive chloroplast introgression has been documented in polyploid *Betula* species of eastern North America. However, the extent to which the nuclear genomes of these species are differentiated is unknown. Therefore, we evaluated genetic differentiation among largely sympatric *Betula papyrifera*, *B. alleghaniensis*, and *B. lenta* using nuclear microsatellite markers. Principal components analysis (PCA) and analysis of molecular variation (AMOVA) were used to evaluate genetic differentiation. Bayesian model-based clustering was used to identify putatively admixed individuals. Despite a high incidence of allele sharing, all of the species were significantly differentiated even within zones of sympatry. A number of individuals were identified as possibly admixed between *B. papyrifera* and B. *alleghaniensis* and between *B. alleghaniensis* and *B. lenta*. Admixture estimates between *B. alleghaniensis* and *B. papyrifera* increased significantly moving northward into the sympatric zone, suggesting the occurrence of hybridization in previously glaciated habitats. In contrast, admixture proportions of *B. papyrifera* and *B. alleghaniensis* did not show a significant geographic trend, which points to recent ancestry as the likely cause of allele sharing between these two species. We suggest that allele sharing of *B. papyrifera* and *B. alleghaniensis* results from a combination of ongoing gene flow and historic introgression via pollen swamping during northward colonization into postglacial environments.

Keywords Allele sharing · *Betula* · Birches · Eastern North America · Hybridization · Introgression · Nuclear microsatellites

Communicated by P. Ingvarsson

This article is part of the Topical Collection on *Hybridization*

✉   Ashley M. Thomson
    amthoms@umich.edu

1   Québec Centre for Biodiversity Science, 1205 Dr. Penfield Avenue, Montréal, Canada

2   Department of Biology and Centre for Structural and Functional Genomics, Concordia University, 7141 Sherbrooke St W, Montreal, Quebec H4B 1R6, Canada

3   Present address: Department of Ecology and Evolutionary Biology, University of Michigan, 830 North University, Ann Arbor, MI 48109-1048, USA

4   New York Genome Center, 101 Avenue of the Americas, New York, NY 10013, USA

Introduction

The study of hybridization and introgression has long-fascinated evolutionary biologists because of its influence on adaptation and speciation (Stebbins 1959; Rieseberg 1995; Grant and Grant 1996). Natural hybridization and introgression are common in plants, with recent estimates indicating that hybridization occurs in approximately 10 % of plant species (Whitney et al. 2010). The evolutionary consequences of hybridization can vary widely, depending on the strength of the reproductive isolating mechanisms. When hybrid fertility is reduced, backcrosses may rarely occur and introgression among the parental taxa may be limited (e.g., Hansson et al. 2012). In other cases, hybrids may rapidly become reproductively isolated from the parental taxa through karyotypic and ecological divergence or polyploidization, thus limiting gene flow among parental species (Rieseberg and Willis 2007). Alternatively, hybridization can result in the production of

Springer

fertile or semi-fertile hybrids, which are not genetically isolated from the parents. Backcrossing can then lead to introgression and creation of novel genotypes (Oberprieler et al. 2010). If reproductive barriers among parental species are weak, substantial introgression may occur, leading to hybrid zones in which intermediate genotypes predominate (Jiggins and Mallet 2000), as observed in the *Eucalyptus* species complex in Australia (Jones et al. 2013). Alternatively, environmental selection (Dodd and Afzal-Rafii 2004) and preferential within-species mating (Lepais and Gerber 2011) may maintain morphological and molecular differentiation among parental species despite recurrent hybrid formation.

Chloroplast capture appears to occur within many plant groups (Rieseberg and Soltis 1991; Rieseberg 1995), and in temperate zones, it is often interpreted as evidence of historic hybridization associated with Pleistocene climatic fluctuations; for example, in *Acer* (Saeki et al. 2011), *Fraxinus* (Heuertz et al. 2006), *Quercus* (Dumolin-Lapègue et al. 1997), *Nothofagus* (Acosta and Premoli 2010), and *Silene* (Prentice et al. 2008). However, for many species, introgression at chloroplast DNA (cpDNA) markers is not associated with high-level introgression at nuclear loci; for example, geographical structuring of chloroplast haplotypes suggestive of introgression has been found in European *Alnus* (King and Ferris 2000) and eastern North American *Quercus* (Whittemore and Schaal 1991) though nuclear genes did not appear to be exchanged freely. There currently exist a number of hypotheses as to why organellar DNA might more readily introgress than nuclear genes (Rieseberg et al. 1995; Petit et al. 2003). However, these are mostly similar in that they relate to the influence of inter-specific pollen competition on relative rates of introgression. Rieseberg et al. (1995) suggest that when conspecific pollen is scarce, *Helianthus* species are most likely to be fertilized by heterospecific pollen. Repeated backcrossing of female offspring with the more frequent male parent may lead to recovery of individuals possessing the cytoplasmic genome of the original maternal species, but for which the majority of nuclear alleles originate from the paternal species. Petit et al. (2003) invoked a similar mechanism to explain chloroplast capture in European oaks, suggesting that during post-glacial recolonization, species may colonize a site occupied by a congener through pollen flow followed by backcrossing to the previously more-restricted species. Such post-glacial pollen swamping has been suggested to be an important mechanism of invasion of deglaciated habitats and has been used to explain observed patterns of nuclear and cpDNA discordance in a number of temperate species (Petit et al. 2003; Acosta and Premoli 2010).

The genus *Betula* provides a novel model for studies of introgression among temperate forest trees, as it is characterized by a high frequency of natural hybridization and associated polyploidy (DeJong 1993). *Betula* is composed of approximately 35 species, which are predominantly long-lived perennial trees and shrubs of mostly cold temperate and circumboreal distribution (Järvinen et al. 2004; Ashburner and Mcallister 2013). Studies of European *Betula* species have revealed widespread cpDNA haplotype sharing interpreted as evidence of cpDNA introgression (Palme et al. 2004; Maliouchenko et al. 2007; Thórsson et al. 2010), and introgression has also been confirmed by cytological and molecular studies (Thórsson et al. 2001; Anamthawat-Jónsson and Thórsson 2003; Wang et al. 2014). In eastern North America, *Betula* is composed of approximately 17 species which exhibit a wide range of morphological variation (Furlow 1990, 1997). Natural hybridization is thought to occur among numerous eastern North American *Betula* species based on reports of morphological intermediates (Sharik and Barnes 1971; Clausen 1973, 1977; Barnes et al. 1974; Furlow 1990) and a recent study which found geographically structured chloroplast haplotype sharing in *B. papyrifera* and *B. alleghaniensis* (Thomson et al. 2015). However, thus far, no studies have examined the genetic differentiation and incidence of introgression among eastern North American *Betula* species using nuclear DNA markers, possibly in part due to the difficulty in analyzing polyploid genotypes.

Within eastern North America, three of the most common and wide-ranging *Betula* species include paper birch (*B. papyrifera* Marshall), yellow birch (*B. alleghaniensis* Britton), and sweet birch (*B. lenta* L.). *B. papyrifera* is an early successional, shade-intolerant species that occurs across a broad range of site and soil types throughout its transcontinental boreal distribution (Hutnik and Cunningham 1961). *B. alleghaniensis* is most often found near streams and riverbanks within mesic and rich-mesic climax forest stands of southeastern Canada and the northeastern USA (Gilbert 1960). *Betula lenta* L. is commonly found on cool, rich-mesic slopes of climax and sub-climax forests throughout the Appalachian region of the eastern USA (Leak 1958). *B. papyrifera* and *B. alleghaniensis* are sympatric throughout a large area of southeastern Canada and the northeastern USA, and it appears that flowering times of some individuals of *B. alleghaniensis* overlap sufficiently with pollen shed of *B. papyrifera* that hybrid individuals might be formed (Clausen 1973; Barnes et al. 1974). The cross-compatibility of *B. papyrifera* and *B. alleghaniensis* has been shown by experimental crosses (Clausen 1966), and putative natural hybrids have been identified from numerous localities throughout the sympatric region based on morphology (Clausen 1973, 1977; Barnes et al. 1974). It has been suggested that hybrids might occur somewhat frequently on disturbed sites where the species overlap (Barnes et al. 1974; Clausen 1977). Moreover, *B. papyrifera* and *B. alleghaniensis* share haplotypes locally within their sympatric distribution, suggesting that these species have introgressed at some point in the past (Thomson et al. 2015). Natural hybridization between *B. lenta* and *B. alleghaniensis* might occur since these species overlap in range and flowering time and have been successfully

experimentally crossed (Clausen 1966; Sharik and Barnes 1971). The objectives of this study were to examine the genetic relationships among eastern North American *Betula* species to (i) clarify the extent to which species are genetically distinct at nuclear markers and (ii) test for evidence of hybridization and introgression at nuclear microsatellite loci.

Methods

Population sampling

Our sampling focused on sites in eastern North America from the Great Lakes States eastward, where numerous birch species may occur in the same population or within close geographic proximity. Thus, we did not consider geographic isolation of a given species as a requirement for sampling, and many of our study sites included two or more species. Mature leaf samples were collected from natural populations of *B. alleghaniensis*, *B. papyrifera*, and *B. lenta* at 20 sites in North America (Fig. 1). At each site, we aimed to sample at least 30 individuals per species, but sample sizes of 20–30 individuals per species were accepted on sites with low representation of a given species. Conspecific individuals were sampled at distances greater than 200 m from one another, while no minimum distance was used in sampling of individuals of different species that occurred on the same site. Leaf material was quick dried in silica gel to preserve DNA quality. In total, samples from 483 individuals of *B. alleghaniensis* at 16 sites, 256 individuals of *B. papyrifera* at 9 sites, and 300 individuals of *B. lenta* at 12 sites were collected (Appendix S1). *B. alleghaniensis* and *B. papyrifera* co-occurred at eight sites, *B. alleghaniensis* and *B. lenta* co-occurred at nine sites, *B. papyrifera* and *B. lenta* co-occurred at two sites, and all three species occurred together at two sites.

Laboratory procedures

DNA was extracted from 20 mg of dried leaf tissue following the modified CTAB protocol of Zeng et al. (2002). Individuals were analyzed at six polymorphic microsatellite markers developed in previous studies (Wu et al. 2002; Kulju et al. 2004; Truong et al. 2005; Tsuda et al. 2008). Primer information and thermocycling procedures are presented in Appendix S2. Amplified fragments were electrophoretically separated on the Licor 4200 DNA analyzer along with the manufacturer's 50–350 bp size standard (LiCOR Biosciences, Lincoln, NE). Fragment lengths were scored using SAGA GT 2.1 software (Licor Biosciences) and verified manually.

Data analysis

The underlying copy number of alleles at each locus could not be determined for the polyploid species *B. alleghaniensis* and *B. papyrifera*. Thus, we considered only the presence or absence of each allele size variant at each locus, and it was assumed that the underlying genotype was ambiguous with respect to allele numbers (i.e., dosage). Since *B. papyrifera* is known to have varying chromosome numbers (Grant and Thompson 1975), the ploidy level of each individual was estimated based on the maximum number of allele size variants at a locus using the POLYSAT package for R software (Clark and Jasieniuk 2011). The ploidy levels of *B. alleghaniensis* and *B. lenta* were fixed as six and two times, respectively.

The multilocus data were recorded as a binary matrix of presence/absence of each allele for each individual. Genetic diversity statistics, including the total number of alleles for each species, and the number of unique alleles per species were calculated using GenAlEx 6.41 (Peakall and Smouse 2006). Allele frequencies per locus and species were estimated in GenAlEx based on the presence/absence of unique microsatellite allele size variants. The correlation of allele frequencies among species was examined through the Pearson product-moment correlation calculated in the R statistical environment (R Development Core Team 2012). The R package adegenet (Jombart 2008) was used to calculate the percentage of shared alleles among each species.

To visualize the genetic relationships among individuals of each species, principal components analysis (PCA) was performed on the matrix of pairwise Euclidean distances among individuals calculated in GenAlEx. The Euclidean distance measure has the advantage of not making assumptions about the distribution of allele size variants nor the population genetics of the organism examined (Kloda et al. 2008). Moreover, the Euclidean distance measure does not consider the shared absence of an allele as a common characteristic. We also calculated a second genetic distance measure suitable for polyploid microsatellite data with unknown allele dosage (Bruvo et al. 2004). The Bruvo genetic distance measure uses the stepwise mutation model to determine genetic distances among individuals based on polyploid microsatellite data and is suitable for use on populations of mixed ploidy. Bruvo distance matrix calculations were performed using polysat. The PCA for both Bruvo and Euclidean distance measures was performed in R. Genetic differentiation among species was examined through analysis of molecular variance (AMOVA) (Excoffier et al. 1992), considering all species simultaneously, as well as different pairs of species groupings. AMOVA was conducted on the matrix of binary data using GenAlEx software with 9999 permutations of the data to calculate significance values. Genetic differentiation among population pairs of each species was estimated by the measure
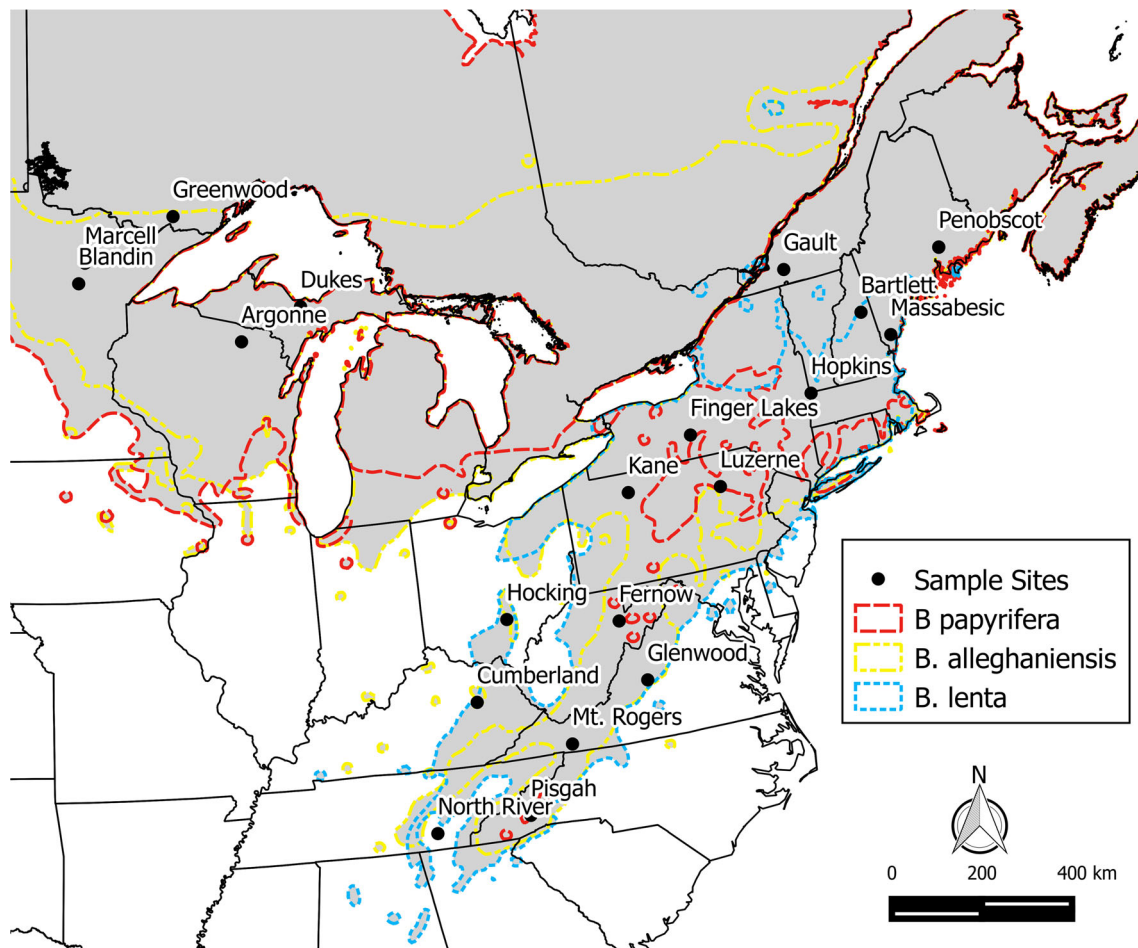
Fig. 1 Map of sampling locations showing the natural distribution and regions of sympatry of *B. papyrifera*, *B. alleghaniensis*, and *B. lenta* in eastern North America

Φ_{PT}, an F_{ST} analog calculated as part of the AMOVA procedure in GenAlEx.

As another method of examining genetic structuring among the species, we used Bayesian model-based clustering analysis implemented in STRUCTURE (Pritchard et al. 2000) to examine the partitioning of individuals into species groups. STRUCTURE accommodates genetic ambiguity due to partial heterozygosity at codominant loci in polyploids through use of an algorithm which generates full genotypes for each individual based on their partial genotypes (Falush et al. 2007). We used the admixture model, with a burn-in of 100, 000 iterations and 100,000 iterations of each MCMC chain for $K = 1$–10, each with five independent runs using the correlated allele frequencies option. We also used STRUCTURE to investigate the possible occurrence of hybridization and introgression among population pairs of different species ($K = 2$), using both the admixture and no-admixture models. We evaluated the best-fit model for the data by calculating the Bayes factor (Raftery 1995) for the estimated likelihood of the admixture model to that of the no-admixture model for each species pair. The Bayes factor is a model selection criterion

calculated from Bayesian posterior probabilities that may be used to distinguish among alternative hypotheses. The results of the admixture analysis were used to identify possible natural hybrids as those having admixture proportions of <0.8 within their respective species cluster, while any individuals with admixture proportions of >0.8 were considered to be relatively genetically "pure" representatives of that species. As an additional means of investigating potential hybridization among species, we used tests of isolation by distance (IBD) to investigate the genetic similarity between population pairs of different species at different spatial scales. If introgression was occurring, it would be expected that the genetic similarity between two hybridizing species at the same locality should be greater than the genetic similarity between populations of the same two species measured at different localities (Muir and Schlötterer 2005). To test for the presence of IBD, geographic distances among population pairs of different species were compared with pairwise genetic distances (Φ_{PT} values) between them. We examined the occurrence of geographic trends in admixture by regressing individual admixture proportions obtained from STRUCTURE analysis

against latitude under the assumption that admixture proportions should show significant trends moving northward or southward into and out of sympatric regions if introgression was occurring among species. All regressions were performed in the R statistical environment.

Results

Molecular variation and genetic differentiation among species

The six microsatellite loci yielded a total of 81 alleles, with the number of alleles ranging from 10 to 16 per locus. Combinations of allele size variants across loci revealed unique microsatellite banding patterns for all sampled individuals ($N=1039$). *B. alleghaniensis* had the greatest number of alleles (77), followed by *B. papyrifera* (59) and *B. lenta* (22). In total, 18 alleles were unique to *B. alleghaniensis*, three were unique to *B. papyrifera*, and a single allele was unique to *B. lenta*. None of the observed alleles were completely species-specific. However, when comparing *B. papyrifera* and *B. alleghaniensis*, two alleles were nearly species-specific, having frequency differences of greater than 0.8 between the two species. One allele was nearly species-specific between *B. alleghaniensis* and *B. lenta*, and five alleles were nearly species-specific between *B. papyrifera* and *B. lenta*. The highest level of allele sharing (69.1 %) was found between *B. papyrifera* and *B. alleghaniensis*, while levels of allele sharing were lower for *B. alleghaniensis* and *B. lenta* (25.9 %) and *B. papyrifera* and *B. lenta* (22.2 %). Allele frequencies were significantly positively correlated between *B. alleghaniensis* and *B. papyrifera* ($r=0.3652$, $P<0.001$) and *B. alleghaniensis* and *B. lenta* ($r=0.4485$, $P<0.01$). In contrast, allele frequencies of *B. papyrifera* and *B. lenta* were not significantly correlated ($r=0.1266$, $P=0.2599$).

Principal components analyses based on Euclidean and Bruvo distance measures revealed a clear separation of *B. papyrifera*, *B. alleghaniensis*, and *B. lenta* species clusters, although a number of intermediates occurred between *B. alleghaniensis* and *B. lenta* and *B. alleghaniensis* and *B. papyrifera* (Fig. 2). The occurrence of three genetic clusters was also supported by the STRUCTURE analysis, for which $ln\ Pr(X|K)$ values increased substantially from $K=1$ to $K=3$ and then leveled off. For $K=3$, 85.3 % of *B. alleghaniensis* individuals, 98.0 % of *B. papyrifera* individuals, and 99.7 % of *B. lenta* individuals were assigned to their respective species cluster with admixture proportions ≥0.8 (Fig. 3a). Analysis of molecular variance indicated substantial variation (34.1 %) among the three species (Table 1). In contrast, variation among populations within species was low (1.4 %), and the majority of variation was harbored within populations (64.6 %). When species pairs were considered, the greatest
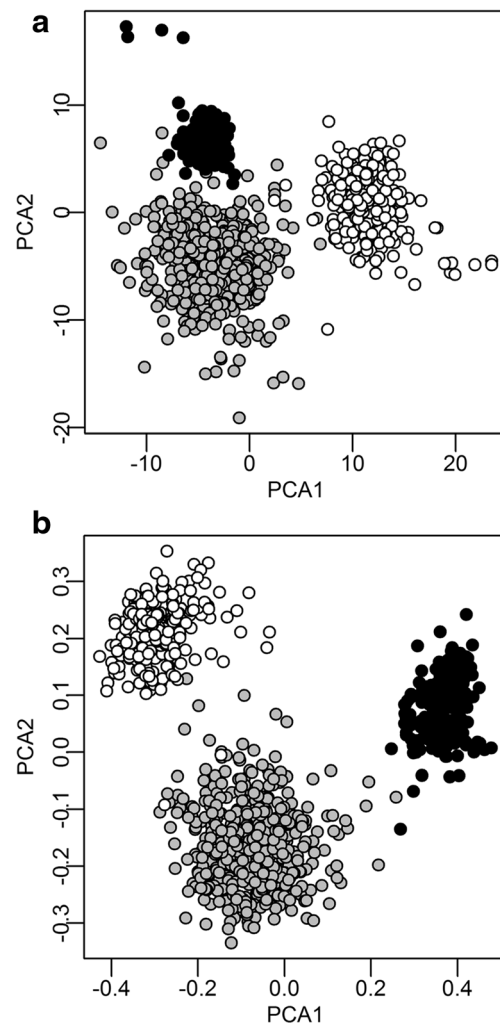
Fig. 2 Principal components analysis (*PCA*) of genetic distances between individuals of *B. alleghaniensis* (*dark gray*), *B. papyrifera* (*light gray*), and *B. lenta* (*black*) calculated using Euclidean distance (**a**) and Bruvo distance (**b**)

variation was found between *B. papyrifera* and *B. lenta* (51.4 %). Differentiation between *B. alleghaniensis* and *B. lenta* was somewhat lower (31.2 %), and the lowest variation was found between *B. alleghaniensis* and *B. papyrifera* (27.3 %).

Genetic admixture and inferred hybridization

The approximate Bayes factor favoring the admixture to no-admixture model approached infinity for the STRUCTURE analysis of *B. papyrifera* and *B. alleghaniensis*, and the analysis of *B. alleghaniensis* and *B. lenta* (ratio of probability for admixture to no-admixture model=1:0). As Bayes factors greater than 100:1 are considered definitive (Raftery 1995), the analyses more strongly support the occurrence of admixture than genetic isolation between *B. papyrifera* and *B. alleghaniensis* and between *B. alleghaniensis* and *B. lenta*.
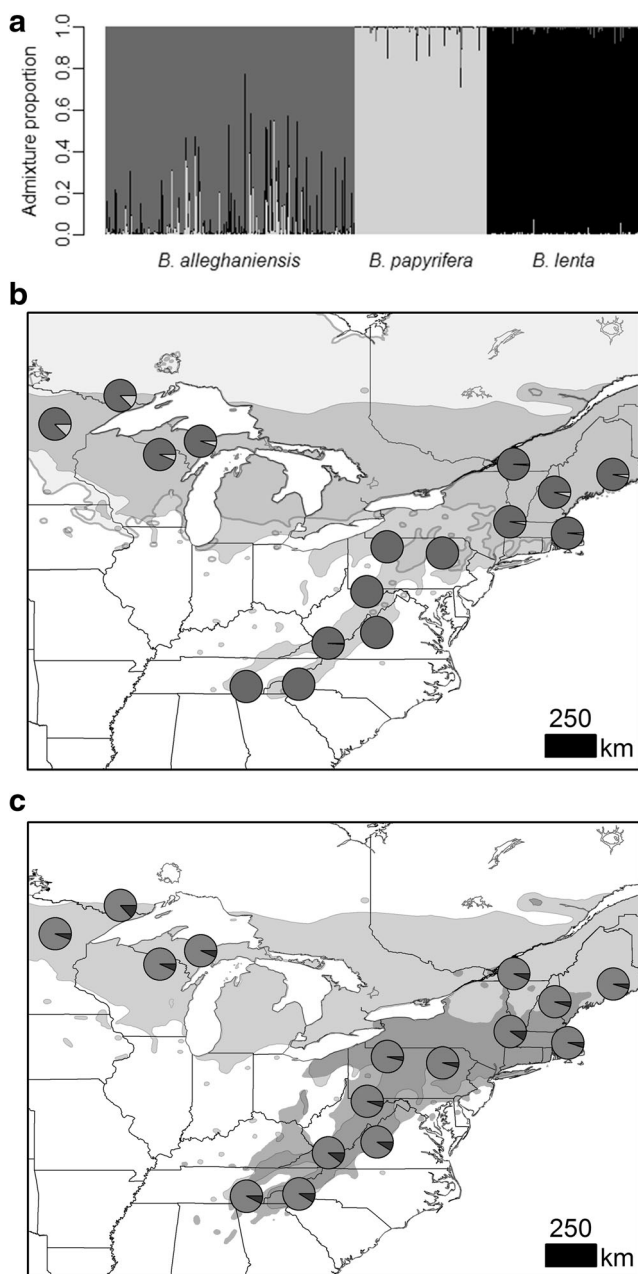
Fig. 3 **a** Assignment of individual genomes to species clusters identified by STRUCTURE analysis for $K=3$ using the admixture model. **b** Admixture of *B. papyrifera* alleles into populations of *B. alleghaniensis*. **c** Admixture of *B. lenta* alleles into populations of *B. alleghaniensis*. Pie charts represent the proportion of alleles at each sampling location originating from each species, with *B. papyrifera* represented as light grey, *B. alleghaniensis* as medium grey, and *B. lenta* as dark grey.

The admixture model for $K=2$ inferred a total of 32 individuals as having mixed ancestry between *B. papyrifera* and *B. alleghaniensis* (admixture proportion <0.8 within their respective species cluster). Of these, three individuals had values close to 0.5 and were classified as possible F$_1$ hybrids. Inferred hybrids were found in nine of eleven sympatric populations (Appendix S3). In contrast, no inferred hybrids were found in populations sampled outside of the sympatric distribution of

B. papyrifera and *B. alleghaniensis*. A total of 60 individuals were inferred as having mixed ancestry between *B. alleghaniensis* and *B. lenta* based on admixture proportions obtained from STRUCTURE analysis for $K=2$. Of these, three were classified as possible F$_1$ hybrids while the remaining were classified as possible backcrosses with admixture proportions ranging from 0.227 to 0.782 (average=0.662). Admixed individuals were found at each of the eleven sympatric sampling locations and at six of the eight allopatric locations (Appendix S4). The frequency distribution of admixture proportions for *B. papyrifera* and *B. alleghaniensis* at eight shared sites indicates a bimodal distribution with most individuals being classified as either pure *B. papyrifera* or *B. alleghaniensis* (Appendix S5a). In total, 18.75 % of the inferred hybrids were genetically more similar to *B. papyrifera* (admixture proportion of <0.5) while 81.25 % were genetically more similar to *B. alleghaniensis* (admixture proportion of >0.5). Interestingly, two of the hybrids which were genetically more similar to *B. papyrifera* were classified as *B. alleghaniensis* on the basis of morphology, while one individual that was genetically more similar to *B. alleghaniensis* was classified as *B. papyrifera* based on its morphology. The frequency distribution of admixture proportions for *B. alleghaniensis* and *B. lenta* at nine shared sites indicates a bimodal distribution with a relatively small proportion of admixed individuals. Of the individuals inferred to be admixed between *B. alleghaniensis* and *B. lenta*, 85.7 % were genetically most similar to *B. alleghaniensis*, while 14.3 % were genetically most similar to *B. lenta* (Appendix S5b). Five of the admixed individuals which were genetically most similar to *B. lenta* were identified as *B. alleghaniensis* based on morphology.

Admixture between *B. papyrifera* and *B. alleghaniensis* was significantly correlated to latitude ($P<0.001$), with introgression of *B. papyrifera* alleles into *B. alleghaniensis* increasing from south to north (Figs. 3b and 4a). However, no significant trend of admixture values with respect to latitude was observed between *B. alleghaniensis* and *B. lenta* ($P=0.388$) (Figs. 3c and 4b). Linear regression of Φ_{PT} values between population pairs of *B. papyrifera* and *B. alleghaniensis* revealed a weak but significant trend of IBD ($P=0.027$) (Appendix S6a). Regressions of Φ_{PT} values against distance between populations of *B. alleghaniensis* and *B. lenta* revealed a contrasting pattern, with a significant trend of increasing genetic similarity with increasing distance between populations ($P=0.001$) (Appendix S6b).

Discussion

Genetic differentiation among species

The molecular data presented herein suggest that *B. alleghaniensis*, *B. papyrifera*, and *B. lenta* are genomically differentiated, despite moderate to high incidence of microsatellite allele sharing. Although no species-specific alleles were

Table 1　Analysis of molecular variance among species

Species grouping	Source of variation	df	Variation (%)	Fixation indices	P value
B. papyrifera/B. alleghaniensis/ B. lenta	Among species	2	34.1	Φ_{RT}=0.3408	<0.001
	Among populations	34	1.4	Φ_{PR}=0.0205	<0.001
	Within populations	1002	64.6	Φ_{PT}=0.3544	<0.001
B. papyrifera/B. alleghaniensis	Among species	1	27.3	Φ_{RT}=0.2725	<0.001
	Among populations	23	1.5	Φ_{PR}=0.0212	<0.001
	Within populations	714	71.2	Φ_{PT}=0.2880	<0.001
B. alleghaniensis/B. lenta	Among species	1	31.2	Φ_{RT}=0.3121	<0.001
	Among populations	26	1.3	Φ_{PR}=0.0187	<0.001
	Within populations	755	67.5	Φ_{PT}=0.3249	<0.001
B. papyrifera/B. lenta	Among species	1	51.4	Φ_{RT}=0.5142	<0.001
	Among populations	19	1.2	Φ_{PR}=0.0241	<0.001
	Within populations	535	47.4	Φ_{PT}=0.5259	<0.001

Φ_{RT} differentiation among groups of populations, Φ_{PR} differentiation among populations within groups, Φ_{PT} differentiation among populations

observed, we found a number of alleles with large frequency differences among species resulting in strong genetic clustering in the PCA and significant among-species variation detected by AMOVA. xMoreover, the STRUCTURE analysis was able to partition the majority of the individuals into their respective species clusters with a high level of confidence and without inclusion of prior population information. These results suggest that the markers used herein can be used to distinguish among the three species, even in absence of information regarding morphology. Of the different species pairs, B. papyrifera and B. lenta were the most highly differentiated based on microsatellite allele frequencies (Φ_{PT}=0.5259), while B. papyrifera and B. alleghaniensis were the least differentiated (Φ_{PT}=0.2880). The high incidence of allele sharing is similar to that of a study of single nucleotide polymorphism (SNP) variation in European birches, which found that the majority of alleles were shared among B. nana, B. pendula, and B. pubescens (Wang et al. 2013). In that study, only 4 of 719 loci (0.6 %) harbored alleles unique to B. nana. Similar to our study, nuclear microsatellite analysis of European birches also revealed clear differentiation among the species despite the occurrence of allele sharing (Wang et al. 2014).

Evidence for hybridization and introgression

Our analysis revealed a number of individuals whose genotypes were of mixed ancestry between B. papyrifera and B. alleghaniensis and between B. alleghaniensis and B. lenta. Although allele sharing may point to the possible occurrence of hybridization and introgression, it is important to distinguish between allele sharing resulting from shared ancestry versus introgression. The shared ancestry hypothesis assumes a recent divergence between species and large effective population sizes, allowing for alleles to be shared among

species in the absence of gene flow (Muir and Schlötterer 2005). Incomplete lineage sorting (ILS) is thought to be common among tree species due to their large effective population sizes and long generation times (Bouillé and Bousquet 2005; Chen et al. 2010). We suspect that shared ancestry is responsible for some of the allele sharing of B. alleghaniensis and B. lenta, since these species belong to the same subgenus (DeJong 1993) and are classified as sister species based on molecular phylogenies (Järvinen et al. 2004; Schenk et al. 2008). It is uncertain the extent to which shared ancestry might contribute to the allele sharing of B. papyrifera and B. alleghaniensis, since the phylogenetic relationship between these two species is currently unclear. Most studies suggest that B. papyrifera and B. alleghaniensis belong to two separate clades on the basis of morphology and molecular markers (DeJong 1993; Järvinen et al. 2004; Li et al. 2005; Schenk et al. 2008). However, both B. papyrifera and B. alleghaniensis are suspected allopolyploids, and it has been suggested that both species may have at least one parental species belonging to the white-barked birch clade (DeJong 1993; Järvinen et al. 2004). In the case of a close evolutionary relationship between B. papyrifera and B. alleghaniensis, we would expect some role of ILS.

While introgression and ILS can both result in allele sharing between species, the geographic distribution of shared alleles is expected to differ under the scenarios (Muir and Schlötterer 2005). Specifically, introgression may be associated with a detectable pattern of IBD between heterospecific populations, whereas such a pattern should not arise from ILS. Also, generally higher rates of introgression might be expected in zones of sympatry compared with zones of allopatry, as has been observed for Pinus (Ye et al. 2002), Alnus (Bousquet et al. 1990), Coffea (Mahe et al. 2007), and Silene (Minder and Widmer 2008). In this study, we found evidence for a significant cross-specific pattern of IBD among population pairs of
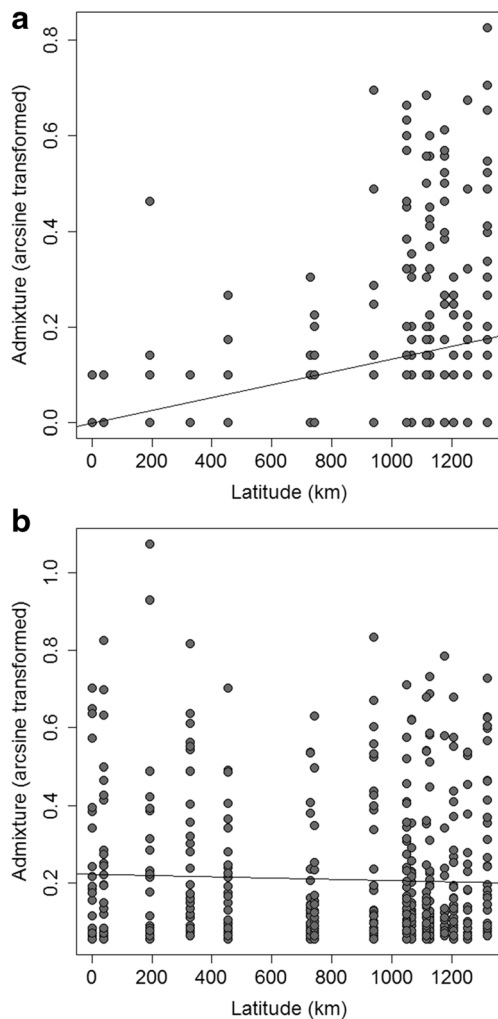
Fig. 4 Regression of admixture proportions against latitude for populations of *B. alleghaniensis*. **a** Introgression of *B. papyrifera* alleles into *B. alleghaniensis*. **b** Introgression of *B. lenta* alleles into *B. alleghaniensis*. Arcsine-transformed admixture proportions obtained for each individual from STRUCTURE are represented by circles. Higher admixture values represent individuals with greater proportions of alleles derived from another species

B. papyrifera and *B. alleghaniensis*, suggesting that these species may exchange alleles through local gene flow. Moreover, *B. alleghaniensis* individuals were found to be significantly more admixed moving from the zone of allopatry in the south towards the zone of sympatry in the north. In contrast, no significant pattern of IBD could be detected between *B. alleghaniensis* and *B. lenta,* and admixture proportions varied little between the zones of allopatry and sympatry. Several *B. alleghaniensis* individuals identified as putatively admixed with *B. lenta* originated from far outside the zone of sympatry, which might suggest that these individuals represent "false positives" identified by the STRUCTURE analysis. Based on the absence of a strong geographic pattern of admixture between *B. alleghaniensis* and *B. lenta*, we conclude that allele

sharing between these species is better explained by ILS than by introgression.

Possible isolating mechanisms

Our study did not find strong evidence for introgression between *B. papyrifera* and *B. lenta*, suggesting that the two species are reproductively isolated. This could be due to pre-zygotic isolation, since *B. papyrifera* and *B. lenta* overlap only within a small area of New England and the northern Appalachians, which should limit their opportunity for gene flow. Ploidy differences between diploid *B. lenta* and polyploid *B. papyrifera* might contribute to post-zygotic barriers, since large differences in chromosome number are often associated with reproductive isolation (Hersch-Green 2012). The apparent lack of hybridization between *B. lenta* and *B. alleghaniensis* is somewhat less expected, since these species are closely related and sympatric throughout a large region of the northeastern USA and flowering times overlap sufficiently that hybridization might be expected to occur (Clausen 1973). While there are few reports of natural hybrids between *B. alleghaniensis* and *B. lenta* (Harger et al. 1917), controlled crosses between them have been made successfully (Sharik and Barnes 1971). Hybrids resulting from controlled crosses were partially fertile, and pollen grain abortion of the hybrids did not differ markedly from the parents. However, germination success was significantly reduced in the hybrid, and hybrid seedlings demonstrated significantly reduced height and diameter in comparison with the parents. Thus, the reproductive isolation of *B. alleghaniensis* and *B. lenta* may be due largely due to genomic incompatibilities resulting in the reduced viability of hybrid offspring.

Despite finding evidence for admixture among *B. papyrifera* and *B. alleghaniensis*, most individuals of *B. alleghaniensis* (85.3 %) and *B. papyrifera* (98.0 %) possessed pure or nearly genetically pure genotypes suggesting that species integrity is largely maintained by pre-zygotic and/ or post-zygotic barriers. For example, pollen competition can act as a partial reproductive barrier in *Helianthus* and *Quercus*, limiting the frequency of hybrid formation (Rieseberg et al. 1995, 1998; Lepais et al. 2009; Lepais and Gerber 2011). We expect that pollen competition may play a role in the reproductive isolation of *B. papyrifera* and *B. alleghaniensis*, since paternity analysis of other North American *Betula* species has revealed that most offspring (98 %) result from conspecific matings (Williams et al. 1999). Clausen (1973) reported that 32 of 55 *B. alleghaniensis* seed sources planted at a provenance trial in Wisconsin contained putative hybrids of *B. alleghaniensis* and *B. papyrifera*. The provenance of these hybrids ranged from Georgia to Quebec and Newfoundland, but hybrids were most frequent in three sources originating northwestern Minnesota, northern Illinois, and central Iowa

which occur at the extreme western limit of *B. alleghaniensis*. The high proportion of hybrid seedlings in those provenances was attributed to the greater availability of *B. papyrifera* pollen due to the fact that the stands of origin contained a greater number of individuals of *B. papyrifera* than *B. alleghaniensis*. In our study, we found that the greatest proportion of hybrid individuals originated from the extreme western distribution of *B. alleghaniensis*, supporting the idea that increased frequency of hybridization in that area may be associated with the relatively greater availability of *B. papyrifera* pollen. While *B. alleghaniensis* generally flowers later than *B. papyrifera*, flowering times can overlap such that some flowers of *B. alleghaniensis* are receptive during pollen shed of *B. papyrifera* (Barnes et al. 1974). Clausen (1973) postulated that reports of natural hybrids may be infrequent due to the fact that hybrid trees have not been recognized as such or that the hybrid seedlings may have difficulty in becoming established and competing with the parental species. In fact, all putative natural hybrids reported by Barnes et al. (1974) originated from disturbed habitats characterized by mineral soil, abundant light, and little competition from the parental species or other vegetation.

Historic and contemporary introgression

We propose that the admixture of *B. papyrifera* and *B. alleghaniensis* observed in this study is likely to result from both historic and ongoing introgression. The process of post-glacial pollen swamping, whereby the nuclear genome of an expanding population of one species becomes replaced over time through gene swamping via pollen from a more abundant species, has been invoked to explain patterns of introgression in *Quercus* (Petit et al. 2003) and *Nothofagus* (Acosta and Premoli 2010). In Europe, introgression of *B. nana* alleles into *B. pubescens* has been attributed to hybridization during Holocene range shifts (Wang et al. 2014). In *B. papyrifera* and *B. alleghaniensis*, local haplotype sharing has been suggested to result from introgression during post-glacial recolonization (Thomson et al. 2015). Under this scenario, we propose that introgression between *B. papyrifera* and *B. alleghaniensis* may have occurred via pollen swamping of *B. alleghaniensis* onto *B. papyrifera*. The macrofossil record for *Betula* in eastern North America indicates that *B. papyrifera* began to recolonize deglaciated habitats much earlier than *B. alleghaniensis,* which apparently did not begin to expand until relatively late in the Holocene (Jackson et al. 1997). Since hybridization is most likely to take place when the availability of conspecific pollen is low (Rieseberg et al. 1995), we might expect that female flowers produced by newly dispersed, low density *B. papyrifera* populations might have been somewhat frequently pollinated by *B. alleghaniensis*. Under this scenario, repeated backcrossing of hybrids to *B. alleghaniensis* could produce individuals with a predominantly *B. alleghaniensis* nuclear

genome and *B. papyrifera* plastids. This scenario would account for the relatively low level of nuclear DNA introgression in comparison with the high levels of cpDNA introgression previously observed (Thomson et al. 2015).

While historic introgression might play a significant role in influencing patterns of genetic similarity between *B. papyrifera* and *B. alleghaniensis*, the presence of putative early-generation hybrids observed in populations surveyed in this study suggests that hybrid formation may be ongoing. Our study revealed a small number of individuals with admixture proportions near 0.5, suggesting that these individuals represent F_1 or early-generation hybrids, which we would not expect to find if genetic similarity of *B. papyrifera* and *B. alleghaniensis* were only the result of historic gene flow. Moreover, we found a significant pattern of IBD among sympatric populations, which would not be expected in the absence of ongoing gene flow. In comparison with a study of European birches that found up to 10 % of Icelandic populations were hybrids (Anamthawat-Jónsson and Thórsson 2003), we found up to 13 % of individuals surveyed in one sympatric western population were probable hybrids. Although we might expect the majority of hybrids might be fully or partially sterile as a result of ploidy differences between the parental species, a small percentage of them might produce viable seed (Barnes et al. 1974), and thus might act as a conduit for gene ongoing gene flow between *B. papyrifera* and *B. alleghaniensis*.

Conclusion

Despite moderate to high levels of allele sharing among species, this study demonstrates that *B. alleghaniensis*, *B.papyrifera*, and *B. lenta* are genetically differentiated at nuclear microsatellite markers. Allele sharing between *B. alleghaniensis* and *B. lenta* appears to be more consistent with the hypothesis of shared ancestry than introgression since admixture proportions did not change between the zone of sympatry and the zone of allopatry. The strong genetic differentiation of *B. papyrifera* and *B. lenta* can likely be explained in terms of pre-zygotic barriers to interbreeding as well potential genomic incompatibilities due to differing ploidy levels. Although allele sharing as a result of common ancestry cannot be entirely ruled out, the significant geographic pattern of allele sharing between *B. papyrifera* and *B. alleghaniensis* points to local gene flow through introgression as the most likely cause. We propose that this pattern of introgression is likely the result of both historic and ongoing asymmetric gene flow between the species, resulting in introgression of *B. papyrifera* alleles into the nuclear genome of *B. alleghaniensis*.

References

Acosta CM, Premoli AC (2010) Evidence of chloroplast capture in South American *Nothofagus* (subgenus *Nothofagus*, Nothofagaceae). Mol Phylogenet Evol 54:235–242. doi:10.1016/j.ympev.2009.08.008

Anamthawat-Jónsson K, Thórsson AT (2003) Natural hybridisation in birch: triploid hybrids between *Betula nana* and *B. pubescens*. Plant Cell Tissue Organ Cult 75:99–107

Ashburner K, Mcallister HA (2013) The genus *Betula*: a taxonomic revision of birches. Royal Botanic Gardens, Kew, UK

Barnes BV, Dancik BP, Sharik TL (1974) Natural hybridization of yellow birch and paper birch. For Sci 20:215–221

Bouillé M, Bousquet J (2005) Trans-species shared polymorphisms at orthologous nuclear gene loci among distant species in the conifer *Picea* (Pinaceae): implications for the long-term maintenance of genetic diversity in trees. Am J Bot 92:63–73. doi:10.3732/ajb.92. 1.63

Bousquet J, Cheliak WM, Wang J, Lalonde M (1990) Genetic divergence and introgressive hybridization between *Alnus sinuata* and *A. crispa* (Betulaceae). Plant Syst Evol 170:107–124. doi:10.1007/BF00937853

Bruvo R, Michiels NK, D'Souza TG, Schulenburg H (2004) A simple method for the calculation of microsatellite genotype distances irrespective of ploidy level. Mol Ecol 13:2101–2106. doi:10.1111/j. 1365-294X.2004.02209.x

Chen J, Källman T, Gyllenstrand N, Lascoux M (2010) New insights on the speciation history and nucleotide diversity of three boreal spruce species and a Tertiary relict. Heredity (Edinb) 104:3–14. doi:10. 1038/hdy.2009.88

Clark LV, Jasieniuk M (2011) POLYSAT: an R package for polyploid microsatellite analysis. Mol Ecol Resour 11:562–566. doi:10.1111/j. 1755-0998.2011.02985.x

Clausen KE (1966) Studies of compatibility in *Betula*. Joint Proceedings of the Second Genetics Workshop of the Society of American Foresters and the Seventh Lake States Forest Tree Improvement Conference. U.S. Department of Agriculture Forest Service, North Central Forest Experiment Station, St. Paul, MN, pp 48–52

Clausen KE (1973) Genetics of yellow birch. USDA Forest Service Research Paper WO-18. U.S. Department of Agriculture, Washington, DC

Clausen KE (1977) The yellow×paper birch hybrid—a potential substitute for yellow birch on problem sites. Proceedings of the Thirteenth Lake States Forest Tree Improvement Conference; Gen. Tech. Rep.

NC-50. U.S. Department of Agriculture, Forest Service, North Central Forest Experiment Station., St. Paul, MN, pp 166–171

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

Dodd RS, Afzal-Rafii Z (2004) Selection and dispersal in a multispecies oak hybrid zone. Evolution 58:261–269. doi:10.1111/j.0014-3820. 2004.tb01643.x

Dumolin-Lapègue S, Demesure B, Fineschi S, Le Corre V, Petit RJ (1997) Phylogeographic structure of white oaks throughout the European continent. Genetics 146(4):1475–1487

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131: 479–491

Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. Mol Ecol Notes 7:574–578. doi:10.1111/j.1471-8286.2007. 01758.x

Furlow JJ (1990) The genera of Betulaceae in the southeastern United States. J Arnold Arbor 71:1–67

Furlow JJ (1997) Betulaceae. In: Flora of North America Committee (ed) Flora of North America North of Mexico, vol 3. Oxford University Press, New York, pp 507–538

Gilbert AM (1960) Silvical characteristics of yellow birch (*Betula alleghaniensis*). Station paper no. 134. U.S. Department of Agriculture Forest Service, North Central Forest Experiment Station, Upper Darby, PA

Grant PR, Grant BR (1996) Speciation and hybridization in island birds. Philos Trans R Soc B Biol Sci 351:765–772

Grant WF, Thompson BK (1975) Observations on Canadian birches, *Betula cordifolia*, *B. neoalaskana*, *B. populifolia*, *B. papyrifera*, and *B.×caerulea*. Can J Bot 53:1478–1490. doi:10.1139/b75-181

Hansson B, Tarka M, Dawson DA, Horsburgh GJ (2012) Hybridization but no evidence for backcrossing and introgression in a sympatric population of great reed warblers and clamorous reed warblers. PLoS One 7:e31667. doi:10.1371/journal.pone.0031667

Harger EB, Graves CB, Eames EH et al (1917) Additions to the flora of Connecticut. Rhodora 19:119–130

Hersch-Green EI (2012) Polyploidy in Indian paintbrush (*Castilleja*; Orobanchaceae) species shapes but does not prevent gene flow across species boundaries. Am J Bot 99:1680–1690. doi:10.3732/ ajb.1200253

Heuertz M, Carnevale S, Fineschi S, Sebastiani F, Hausman JF, Paule L, Vendramin GG (2006) Chloroplast DNA phylogeography of European ashes, *Fraxinus* sp. (Oleaceae): roles of hybridization and life history traits. Mol Ecol 15(8):2131–2140

Hutnik RJ, Cunningham FE (1961) Silvical characteristics of paper birch (*Betula papyrifera*). Station paper no. 141. U.S. Department of Agriculture Forest Service, North Central Forest Experiment Station, Upper Darby, PA

Jackson ST, Overpeck JT, Webb T et al (1997) Mapped plant-macrofossil and pollen records of late quaternary vegetation change in eastern North America. Quat Sci Rev 16:1–70. doi:10.1016/S0277-3791(96)00047-9

Järvinen P, Palmé A, Morales LO et al (2004) Phylogenetic relationships of *Betula species* (Betulaceae) based on nuclear ADH and chloroplast matK sequences. Am J Bot 91:1834–1845. doi:10.3732/ajb.91. 11.1834

Jiggins CD, Mallet J (2000) Bimodal hybrid zones and speciation. Trends Ecol Evol 15:250–255

Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403–1405

Jones RC, Steane DA, Lavery M et al (2013) Multiple evolutionary processes drive the patterns of genetic differentiation in a forest tree species complex. Ecol Evol 3:1–17. doi:10.1002/ece3.421

King RA, Ferris C (2000) Chloroplast DNA and nuclear DNA variation in the sympatric alder species, *Alnus cordata* (Lois.) Duby and *A. glutinosa* (L.) Gaertn. Biol J Linn Soc 70:147–160. doi:10.1111/j.1095-8312.2000.tb00205.x

Kloda JM, Dean PDG, Maddren C et al (2008) Using principle component analysis to compare genetic diversity across polyploidy levels within plant complexes: an example from British Restharrows (*Ononis spinosa* and *Ononis repens*). Heredity (Edinb) 100:253–260. doi:10.1038/sj.hdy.6801044

Kulju KKM, Pekkinen M, Varvio S (2004) Twenty-three microsatellite primer pairs for *Betula pendula* (Betulaceae). Mol Ecol Notes 4:471–473. doi:10.1111/j.1471-8286.2004.00704.x

Leak WB (1958) Silvical characteristics of sweet birch (*Betula lenta*). Station paper no. 113. U.S. Department of Agriculture Forest Service, North Central Forest Experiment Station, Upper Darby, PA

Lepais O, Gerber S (2011) Reproductive patterns shape introgression dynamics and species succession within the European white oak species complex. Evolution 65:156–70. doi:10.1111/j.1558-5646.2010.01101.x

Lepais O, Petit RJ, Guichoux E et al (2009) Species relative abundance and direction of introgression in oaks. Mol Ecol 18:2228–2242. doi:10.1111/j.1365-294X.2009.04137.x

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

Mahe L, LePierres D, Combes MC et al (2007) Introgressive hybridization between the allotetraploid *Coffea arabica* and one of its diploid ancestors, *Coffea canephora*, in an exceptional sympatric zone in New Caledonia. Genome 50:316–324. doi:10.1139/G07-011

Maliouchenko O, Palmé AE, Buonamici A et al (2007) Comparative phylogeography and population structure of European *Betula* species, with particular focus on *B. pendula* and *B. pubescens*. J Biogeogr 34:1601–1610. doi:10.1111/j.1365-2699.2007.01729.x

Minder AM, Widmer A (2008) A population genomic analysis of species boundaries: neutral processes, adaptive divergence and introgression between two hybridizing plant species. Mol Ecol 17:1552–1563. doi:10.1111/j.1365-294X.2008.03709.x

Muir G, Schlötterer C (2005) Evidence for shared ancestral polymorphism rather than recurrent gene flow at microsatellite loci differentiating two hybridizing oaks (*Quercus* spp.). Mol Ecol 14:549–561. doi:10.1111/j.1365-294X.2004.02418.x

Oberprieler C, Barth A, Schwarz S, Heilmann J (2010) Morphological and phytochemical variation, genetic structure and phenology in an introgressive hybrid swarm of *Senecio hercynicus* and *S. ovatus* (Compositae, Senecioneae). Plant Syst Evol 286:153–166. doi:10.1007/s00606-010-0295-0

Palme AE, 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*. Mol Ecol 13:167–178. doi:10.1046/j.1365-294X.2003.02034.x

Peakall R, Smouse PE (2006) GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295. doi:10.1111/j.1471-8286.2005.01155.x

Petit RJ, Bodenes C, Ducousso A et al (2003) Hybridization as a mechanism of invasion in oaks. New Phytol 161:151–164. doi:10.1046/j.1469-8137.2003.00944.x

Prentice HC, Malm JU, Hathaway L (2008) Chloroplast DNA variation in the European herb *Silene dioica* (red campion): postglacial migration and interspecific introgression. Plant Syst Evol 272:23–37. doi:10.1007/s00606-007-0629-8

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

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Raftery AE (1995) Bayesian model selection in social research. Sociol Methodol 25:111–163

Rieseberg LH (1995) The role of hybridization in evolution: old wine in new skins. Am J Bot 82:944–953

Rieseberg LH, Soltis DE (1991) Phylogenetic consequences of cytoplasmic gene flow in plants. Evol Trends Plants 5:65–84

Rieseberg LH, Willis JH (2007) Plant speciation. Science 317:910–4. doi:10.1126/science.1137729

Rieseberg LH, Desrochers AM, Youn SJ (1995) Interspecific pollen competition as a reproductive barrier between sympatric species of *Helianthus* (Asteraceae). Am J Bot 82:515–519

Rieseberg LH, Baird SJE, Desrochers AM (1998) Patterns of mating in wild sunflower hybrid zones. Evolution (N Y) 52:713–726. doi:10.2307/2411266

Saeki I, Dick CW, Barnes BV, 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. J Biogeogr 38(5):992–1005

Schenk MF, Thienpont C-L, Koopman WJM et al (2008) Phylogenetic relationships in *Betula* (Betulaceae) based on AFLP markers. Tree Genet Genomes 4:911–924. doi:10.1007/s11295-008-0162-0

Sharik TL, Barnes BV (1971) Hybridization in *Betula alleghaniensis* Britt. and *B. lenta* L.: a comparative analysis of controlled crosses. For Sci 17:415–424

Stebbins GL (1959) The role of hybridization in evolution. Proc Am Philos Soc 103:231–251

Thomson AM, Dick CW, Dayanandan S (2015) A similar phylogeographical structure among sympatric North American birches (*Betula*) is better explained by introgression than by shared biogeographical history. J Biogeogr 42:339–350. doi:10.1111/jbi.12394

Thórsson AT, Salmela E, Anamthawat-Jónsson K (2001) Morphological, cytogenetic, and molecular evidence for introgressive hybridization in birch. J Hered 92:404–408. doi:10.1093/jhered/92.5.404

Thórsson AT, 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. J Biogeogr 37:2098–2110. doi:10.1111/j.1365-2699.2010.02353.x

Truong C, Palme AE, Felber F, Naciri-Graven Y (2005) Isolation and characterization of microsatellite markers in the tetraploid birch, *Betula pubescens* ssp. *tortuosa*. Mol Ecol Notes 5:96–98. doi:10.1111/j.1471-8286.2004.00848.x

Tsuda Y, Ueno S, Ide Y, Tsumura Y (2008) Development of 14 EST-SSRs for *Betula maximowicziana* and their applicability to related species. Conserv Genet 10:661–664. doi:10.1007/s10592-008-9608-z

Wang N, Thomson M, Bodles WJA et al (2013) Genome sequence of dwarf birch (*Betula nana*) and cross-species RAD markers. Mol Ecol 22:3098–3111. doi:10.1111/mec.12131

Wang N, Borrell JS, Bodles WJA, Kuttapitiya A (2014) Molecular footprints of the Holocene retreat of dwarf birch in Britain. 2771–2782. doi: 10.1111/mec.12768

Whitney KD, Ahern JR, Campbell LG et al (2010) Patterns of hybridization in plants. Perspect Plant Ecol Evol Syst 12:175–182. doi:10.1016/j.ppees.2010.02.002

Whittemore AT, Schaal BA (1991) Interspecific gene flow in sympatric oaks. Proc Natl Acad Sci 88:2540–2544. doi:10.1073/pnas.88.6.2540

Williams JH, Friedman WE, Arnold ML (1999) Developmental selection within the angiosperm style: using gamete DNA to visualize interspecific pollen competition. Proc Natl Acad Sci 96:9201–9206. doi:10.1073/pnas.96.16.9201

Wu B, Lian C, Hogetsu T (2002) Development of microsatellite markers in white birch (*Betula platyphylla* var. *japonica*). Mol Ecol Notes 2: 413–415. doi:10.1046/j.1471-8278

Ye TZ, Yang R, Yeh FC (2002) Population structure of a lodgepole pine (*Pinus contorta*) and jack pine (*P. banksiana*) complex as revealed by random amplified polymorphic DNA. Genome 540:530–540. doi:10.1139/G02-016

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