Identification of Molecular Markers Differentiating *Betula papyrifera* and *B. pumila* Populations from Northern Ontario (Canada)

1Nastaran Moarefi, 2Paul Michael, 1Peter Beckett and 1,2Kabwe Nkongolo

1Department of Biology, 2Biomolecular Sciences Program, Laurentian University, Sudbury, Ontario, Canada, P3E 2C6, Canada

**Abstract:** *Betula* is a polyploid and highly polymorphic genus with several species known to hybridize very readily. This high level of introgression has resulted in hybrid populations where different species are sympatric. The main objectives of the present study were to identify putative hybrids between *B. papyrifera* and *B. pumila* in the Greater Sudbury Region (Northern Ontario) and to develop ISSR and RAPD markers that can differentiate *B. papyrifera* and *B. pumila* and their respective populations. This study revealed extensive introgression of *B. papyrifera* genes into *B. pumila* based on morphological characterization in the Greater Sudbury region (Northern Ontario) where these two species coexist. Genomic DNAs were extracted from all the collected genotypes from seven *B. papyrifera* and five *B. pumila* populations within the GSR. Additional samples from *B. pumila* from Wisconsin were also analyzed. All the DNA samples were amplified using ISSR and RAPD primers. No – species – diagnostic markers was identified because in part to a high level of polymorphic loci observed within and among populations. A close look of all the amplified products revealed a number of ISSR and RAPD diagnostic markers that differentiate *P. papyrifera* populations from different origins. Likewise, diagnostic bands distinguishing *B. pumila* from Wisconsin to the GSR population were identified.

**Keywords:** *Betula papyrifera*, *B. pumila*, ISSR, RAPD, Northern Ontario

**Introduction**

*Betula* is a polyploid and highly polymorphic genus with several species known to hybridize very readily. This high level of introgression has resulted in hybrid populations where different species are sympatric (Woodworth, 1929; Anamthawat-Jonsson and Thorsson, 2003). Hybrids between *Betula pumila* X *Betula lenta*, *Betula populifolia* X *Betula papyrifera*, *Betula uerrucosa* X *Betula papyrifera* and others have been reported (Froiland, 1952). The morphological variation in putative hybrids makes it difficult to differentiate them from their parental species on the basis of characters normally used in identification.

Bog Birch (*B. pumila*) typically inhabits wetland environments. It can also be found in low areas of sand dune habitats. This species is vulnerable to invasive species. It is typically a 1.8 to 2.5 m, densely multi-stemmed shrub but can be highly variable in leaf characters and height. It readily hybridizes with *B. alleghaniensis* and *B. papyrifera* resulting in *B. x purpusii* and *B. x sandbergii*, respectively. These hybrids are known to further cross and backcross producing potentially a myriad of intermediate characteristics. Most of these hybrids are over 3.5 m and look more tree than shrub (NOF, 2018).

White birch (*Betula papyrifera*), the most widely distributed (east to west) of all North American birches is primarily native to the cold climates of Canada and Alaska, with its range dipping down into a few of the northern U.S. states and further south in the mountains (to Colorado in the Rockies and to North Carolina in the Appalachians) (Uchytil, 1991). It is a small to medium-sized tree, often with many stems, up to 30 m tall. It grows on a variety of soils and is abundant on rolling upland terrain and floodplain sites, but it also grows on open slopes, avalanche tracks, swamp margins and in bogs. It doesn't grow well in shade and consequently it
often occurs in younger forests following a disturbance (Uchytíl, 1991; Theriault et al., 2013).

This open pollinated species is dominant in the Greater Sudbury Region in Northern Ontario (Canada) after land reclamation. In fact, this species represents 65% of all trees in the region (Theriault et al., 2013). It produces male and female flowers on the same tree in the form of catkins. Both male and female catkins lack petals enhancing B. papyrifera pollination with B. pumila flowers within the same stand. It has been hypothesized that the exchange of genetic information between Betula papyrifera and other species within the genus Betula might be a common phenomenon in open populations (Sofiev et al., 2006; 2013; Theriault et al., 2014). This should result in characteristics that are variable within species making the identification of pure and hybrid genotypes very challenging. Hence, development of species and population diagnostic molecular markers would be useful. It is also important to determine the genetic status of tree populations growing in specific ecological conditions for adaptations and reclamations purposes.

In the last few decades, Inter-Simple Sequence Repeat (ISSR) and Random Amplified Polymorphic DNA (RAPD) markers and other DNA technologies have been employed to analyze genetic structure of populations of several forest trees and to delineate species. DNA markers have successfully distinguished plant species and crop varieties (Mei et al., 2005; Nkongolo et al., 2005) and Mehes-Smith et al. (2007) used ISSR and RAPD marker systems to distinguish conifer species such as P. glauca and P. engelmannii, P. mariana and P. rubens, Pinus strobus and Pinus monticola from one another. But these markers have never been developed in hardwood species such as Betula.

The main objectives of this project were to identify putative hybrids between B. papyrifera and B. pumila in the GSR and to develop ISSR and RAPD markers that can differentiate B. papyrifera and B. pumila and their respective populations.

Materials and Methods

Sampling

Fresh B. papyrifera, B. pumila and putative hybrid leaf samples were collected from the targeted areas within the Greater Sudbury Region (GSR) based on leaf morphology. Five sites were selected for B. papyrifera (Capreol, St. Charles, Onaping Falls, Airport and Azilda) and six for P. pumila and potential putative hybrids samples collection (four populations from Boreal College site, one from Lasalle site and one from Lasalle extension site) (Fig. 1). Ten trees representing each population were selected for this investigation. Leaf samples were wrapped in aluminum foil, immersed in liquid nitrogen and stored at -20°C until DNA extraction.

Additional samples for B. pumila collected from Wisconsin (USA) were provided as seeds (Lot number 1819892) by the Sheffield Seed Company, New York, USA. These seeds were grown in Petawawa boxes as described in Nkongolo et al. (2005) and the DNA was extracted from three weeks old seedlings.

Molecular Analysis

Total DNA was extracted from fresh frozen leaf material using the CTAB extraction protocol as described by Theriault et al. (2013) and Kalubi et al. (2015). This is a modified Doyle and Doyle (1987) procedure that included the addition of 1% Polyvinyl Pyrrolidone (PVP) and 0.2% beta mercaptanol to the cetyl trimethylammonium bromide (CTAB) buffer solution, two additional chloroform spins prior to the isopropanol spin and no addition of RNase. After extraction, DNA was stored in a freezer at -20°C.

A total of 34 ISSR and 18 RAPD primers synthesized by Invitrogen were chosen for DNA amplification. PCR analysis was carried out following the procedure described by Vaillancourt et al. (2008) and Theriault et al. (2013). Each PCR reaction was performed using a total of 25 µl volumes which contained in a 25 µL total volume containing a master mix of 11.4 µL distilled water, 2.5 µL MgSO₄, 2.1 µL 10x buffer 0.5 µL of dNTPs (equal parts dTTP, dATP, dCTP, dGTP), 0.5 µL of ISSR primer, a Taq mix of 3.475 µL distilled water, 0.4 µL 10 x buffer and 0.125 µL Taq polymerase (Applied Biosystems) and 4 µL standardized DNA. For each primer, a negative control reaction was included where ddH₂O was added instead of DNA. All samples were covered with one drop of mineral oil to prevent evaporation and amplified with the Eppendorf Mastercycler gradient. The program was set to a hot start of 5 min at 95°C followed by 2 min at 85°C during which the Taq mix was added, then 42 cycles of 1.5 min at 95°C, 2 min at 55°C and one minute of 72°C. A final extension of 7 min at 72°C after which samples were removed from the thermocycler and placed in the -20°C freezer until further analysis. All PCR products were loaded into 2% agarose gel in 0.5X Tris-Borate-EDTA (TBE) buffer. After the DNA samples were amplified, they were separated for analysis on a 2% agarose gel in 0.5 x TBE with ethidium bromide and run at 3.14V/cm.

Five µL of 1x loading buffer were added to the PCR products and 10 µL of this solution were loaded into the wells of the gel. The gel was run as described above, documented with the Bio-Rad ChemiDoc XRS system and analyzed with Image Lab Software. Only the ISSR and the RAPD primers which gave consistent profiles across the populations and also those that appeared to have diagnostic markers were selected for further analysis. The presence and absence of bands were scored as 1 or 0, respectively. Faint bands were not recorded for analysis.
Fig. 1: Locations of the soil sampling areas from the Greater Sudbury Region for Betula papyrifera, the sampling sites include Site 1: Capreol (reference); Site 2: St. Charles (reference); Site 3: Onaping Falls (reference); Site 4: Airport; Site 5: Azilda. Locations for Betula pumila and putative hybrids include site A (Sudbury Lasalle extension), site B (College Boreal) and site C (Lasalle cemetery area).

**Results and Discussion**

**Ecological Analysis**

Pure *B. papyrifera* stands were easy to identify within the GSR based on morphological characteristics. However, *P. pumila* stands showed variations in genotypes varying from pure *B. pumila* shrubs to putative *B. papyrifera x P. pumila* hybrids. We found five such putative stands within the vicinity of pure *B. papyrifera* populations at four sites (three at “College Boreal” and one at Lasalle site). This suggests that the pollen from *B. papyrifera* did fertilize readily with *P. pumila* female flowers. In all the cases, the putative hybrids showed intermediate characteristics for heights but the leaves size and shapes varied ranging from intermediate between *B. papyrifera* and *B. pumila* to similar to *B. papyrifera* (Fig. 2 and 3). *B. pumila* populations from wetland sites (One at College Boreal and one at Lasalle extension) that were isolated exhibited characteristics of true *B. pumila* species.

Cytological characteristics can be used as a tool for species identification in some cases. Chromosome number in *Betula* genus varies considerably within the species. The somatic chromosome diploid number for *B. papyrifera* is 70 or 84, rarely 56. For *B. pumila*, somatic chromosome number is usually 2n = 56 (Les, 2017). Hence, putative hybrid between *B. pumila* and *B. papyrifera* can be assessed based on chromosomes number and morphological features. In other species, pentaploids (2n = 70) hybrids known as *B. x purpusii* occur between *B. pumila* and *B. alleghaniensis* where the two species are sympatric. Likewise, *B. x sandbergii* are hybrids between *B. pumila* and *B. papyrifera* and can be found where these species coexist. Both hybrids are considered obligate aquatics.
Fig. 2: Betula samples collected from the Greater Sudbury Region; Betula pumila on the left (a), putative hybrid in the middle (b) and Betula papyrifera on the right (c)

Fig. 3: Field view of Betula pumila on the left (a), putative hybrid in the middle (b) and Betula papyrifera on the right (c)

It is known that hybridization in the genus Betula is a common phenomenon (Reznicek et al., 2011). For example, B. papyrifera hybridizes naturally with almost every other native species in the genus (Barnes et al., 1974; Clausen, 1962; Little, 1979; Viereck et al., 1983; Thorsson et al., 2001). Hybrids between B. papyrifera and shrub or small tree species include Yukon birch (B. x eastwoodiae Sarg. or B. x commixta Sarg.) with resin birch (B. glandulosa); horne birch (B. x hornei Butler or B. x beeniana A. Nels.) with dwarf arctic birch (B. nana); Sandberg birch (B. x sandbergii Britton or B. x uliginosa Dugle) with bog birch (B. pumila var. glandulifera); and Andrews birch (B. x andrewsii A. Nels. or B. x piperi Britton or B. x utahensis Britton) with water birch (B. occidentalis). The variety cordifolia is thought to be a hybrid of paper (B. papyrifera) and yellow birch (B. alleghaniensis). Blue birch (B. x caerulea or x caerulea-grandis) is a hybrid between grey birch and var. cordifolia (Brittain and Grant, 1967; Grant and Thompson, 1975). B. X caerulea is derived from B. papyrifera x B. populifolia. The hybrids between B. papyrifera and sweet (B. lenta) and river (B. nigra) birch have not been named.

Some studies have reported the usefulness of pollen morphology in birch taxonomy. Differences in pollen mean size, relative size of vestibulum and minor structural characteristics of pores have been key features of birch species identification (Blackmore et al., 2003; Karlsdóttir et al., 2008). Hence, pollen characteristics can be used to assess the introgression levels between Betula species. But, Ives (1977) who attempted to differentiate Betula nana, Betula glandulosa and Betula papyrifera based on pollen morphology was not able to separate them one from another because their pollen size characteristics form a morphological continuum. Since it is still not easy to differentiate Betula populations within
the same species using morphological, cytological and pollen characteristics, other methods have to be considered. Our focus therefore was to explore the use of molecular tools to differentiate *B. papyrifera*, *B. pumila* and their putative hybrids from the GSR.

**Molecular Analysis**

All the genomic DNA samples were tested to assess their degradation level. They were run in a 1% agarose gel with 0.5x TBE (Tris-Borate-EDTA) buffer. All DNA samples showed large molecular weight bands at the top of the gel. This indicated that they were not degraded and were suitable for PCR amplifications.

The ISSR and RAPD primers used are described in Table 1. Out of the 52 primers screened, 22 ISSR and 8 RAPD primers generated amplification products. The level of polymorphic loci between species was around 98%. This level of intra-population polymorphism was > 75% for the selected primers. Figure 4 and 5 depict an ISSR and a RAPD profile showing this high level of polymorphism. Theriault *et al.* (2013) reported a level of polymorphism ranging from 30% to 79% among *B. papyrifera* populations from the GSR. However, the primers and some of the populations used were not the same as in the present study. Hao *et al.* (2015) revealed that polymorphic frequency of the alleles ranged from 17% to 100% with a mean of 55.85% in *B. papyrifera* populations they analyzed using SSR (microsatellite) primers. Tran *et al.* (2014) reported a level of polymorphic loci ranging from 44% to 65% with ISSR primers and 61% to 72% with RAPD primers for *Quercus rubra* from the GSR. Kalubi *et al.* (2015) reported a polymorphism from 51% to 67% in *A. rubrum* from the GSR based on ISSR analysis. No species – diagnostic ISSR or RAPD markers could be identified because in part of the high level of genetic variation observed in the present study. This lack of differentiation between the two species can also be attributed to the high level of gene integration between them.

A close look of all the amplified products revealed population – diagnostic markers that differentiate the *P. papyrifera* populations from St Charles from other populations. Likewise, diagnostic markers distinguishing *B. pumila* from Wisconsin and Sudbury were identified.

RAPD primer (RAPD UBC 402) revealed a population diagnostic marker for *Betula pumila* whereas four ISSR primers (SC ISSR 10, ISSR UBC 844, ISSR UBC 827 and ISSR HB 12) generated population diagnostic markers for both *B. pumila* and *B. papyrifera*. Some populations – diagnostic markers are illustrated in Fig. 6 to 8. For *Betula papyrifera*, primer SC ISSR 10 generated diagnostic markers of ~550bp and 475bp differentiating the St. Charles population from other samples from the GSR (Fig. 6). It also generated a diagnostic marker (~750 bp) distinguishing *B. pumila* from Sudbury from Wisconsin’s. The ISSR UBC 827 primer amplification profile showed three markers of ~550 bp, ~600bp and ~750bp that distinguish St Charles population from other samples. It generated a marker (~700bp) that differentiated *B. Pumila* from Wisconsin and Sudbury. Other markers depicting differences between St Charles population with other populations include ISSR UBC 844 (~525 bp, ~550 bp and ~400 bp bands) and ISSR HB12 (~550 bp, ~650 bp and ~750 bp). These two primers (ISSR UBC 844 and ISSR HB12) amplified also PCR products of ~575 bp and ~1200 bp that distinguish *B. pumila* population from Wisconsin from the Sudbury’s (Fig. 7).

**Table 1:** Nucleotide sequence of ISSR and RAPD primers used on *B. papyrifera* and *B. pumila* populations

<table>
<thead>
<tr>
<th>Primers</th>
<th>Primer sequence (5’-3’)</th>
<th>G + C content (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISSR primers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ISSR HB 12</td>
<td>CAC CACCCAC GC</td>
<td>72.73</td>
</tr>
<tr>
<td>ISSR HB 13</td>
<td>GAG GAGGAG GC</td>
<td>72.73</td>
</tr>
<tr>
<td>ISSR UBC 827</td>
<td>ACA CAC ACA CAC ACA CG</td>
<td>52.94</td>
</tr>
<tr>
<td>ISSR UBC 834</td>
<td>AGA GAG AGA GAG AGA GYT</td>
<td>44.44</td>
</tr>
<tr>
<td>ISSR UBC 835</td>
<td>AGA GAG AGA GAG AGA GYC</td>
<td>50.00</td>
</tr>
<tr>
<td>ISSR UBC 844</td>
<td>CTC TCT TCT TCT TCT TCT</td>
<td>50.00</td>
</tr>
<tr>
<td>ISSR UBC 873</td>
<td>GAC AGA CAG ACA GAC A</td>
<td>50.00</td>
</tr>
<tr>
<td>ISSR 17898A</td>
<td>CAC ACA CAC ACA AG</td>
<td>50.00</td>
</tr>
<tr>
<td>SC ISSR 3</td>
<td>GAC GAGGACGAC G</td>
<td>69.23</td>
</tr>
<tr>
<td>SC ISSR 10</td>
<td>CTT CTCTCTCTCTCTT CCTCCT CCTCCT CCTCTCTCTCT CT</td>
<td>51.42</td>
</tr>
<tr>
<td>RAPD primers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAPD Grasse 2</td>
<td>GTG GTC CGC A</td>
<td>70.00</td>
</tr>
<tr>
<td>RAPD Grasse 3</td>
<td>GTG GCC GCC G</td>
<td>90.00</td>
</tr>
<tr>
<td>RAPD Grasse 4</td>
<td>GAG GCG CTG C</td>
<td>80.00</td>
</tr>
<tr>
<td>RAPD Grasse 5</td>
<td>CGC CCC CAG T</td>
<td>80.00</td>
</tr>
<tr>
<td>RAPD UBC 270</td>
<td>TGC GCG CGG G</td>
<td>90.00</td>
</tr>
<tr>
<td>RAPD UBC 337</td>
<td>TCC CGA ACC G</td>
<td>70.00</td>
</tr>
<tr>
<td>RAPD UBC 402</td>
<td>CCC GCC GTT G</td>
<td>80.00</td>
</tr>
</tbody>
</table>
Fig. 4: Primer SC ISSR 10 amplified profiles of genomic DNA from *Betula papyrifera* and *Betula pumila* in 2% agarose gel. Where lane 1 is the 1Kb+ DNA ladder, lane 2 is the blank, lane 3 is *B. pumila* from Wisconsin, lane 4 is *B. pumila* from Wisconsin. Lanes 5-14 represent *B. papyrifera* (where lanes 5-6 are samples from Capreol, lanes 7-8 from St. Charles, lanes 9-10 from Onaping Falls, lanes 11-12 from Airport and lanes 13-14 from Azilda). Note the distinct profiles of St. Charles samples (lanes 7 and 8).

Fig. 5: Primer RAPD UBC 270 amplified profiles of genomic DNA from *Betula papyrifera* and *Betula pumila* in 2% agarose gel; Where lane 1 is the blank, lane 2 is the 1 kb+ DNA ladder, Lane 3 and 4 are *B. pumila* from Wisconsin; Lanes 5-9 represent *B. papyrifera* samples (where lane 5 are samples from Capreol, lane 6 from St. Charles, lane 7 from Onaping falls, lane 8 from Airport, lane 9 from Azilda). Note the distinct profile of St. Charles sample (lanes 6).
Fig. 6: Primer SC ISSR 10 amplified profiles of genomic DNA from *Betula papyrifera* and *Betula pumila* in 2% agarose gel; (a) Where lane 1 is the 1Kb+ DNA ladder and lanes 2-6 represent amplified pooled DNA from *B. papyrifera* (where lane 2 are samples from Capreol, lane 3 from St. Charles, lane 4 from Onaping falls, lane 5 from Airport, lane 6 from Azilda). Lane 7 is the blank. The arrows show population - diagnostic bands for St. Charles (lane 3) in *B. papyrifera*. b) Lane 1 is the 1kb+ DNA ladder, lane 2 represents amplified pooled *B. pumila* DNA from Sudbury, lane 3 is the amplified pooled *B. pumila* DNA from Wisconsin USA (lot number: 1819892) and lane 4 is the blank. The arrow reveals a population-diagnostic band for *B. pumila* from Wisconsin (lane 3).

Fig. 7: Primer ISSR HB12 amplified profiles of genomic DNA from *Betula papyrifera* and *Betula pumila* in 2% agarose gel; (a) Where lane 1 is the 1kb+ DNA ladder. Lanes 2-6 represent the amplified pooled DNA from *B. papyrifera* (where lane 2 are samples from Capreol, lane 3 from St. Charles, lane 4 from Onaping falls, lane 5 from Airport, lane 6 from Azilda. Lane 7 is a blank. The arrows reveal population diagnostic bands for St. Charles (lane 3) in *B. papyrifera*. b) Where lane 1 is the 1kb+ DNA ladder. Lane 2 represents amplified pooled *B. pumila* DNA from Sudbury and lane 3 is amplified pooled *B. pumila* DNA from Wisconsin USA (lot number: 1819892). The arrow shows a population-diagnostic band for *B. pumila* Sudbury (lane 2).
RAPD UBC 402 primer generated two markers of ~675bp and ~350 bp that differentiated *Betula pumila* populations from Wisconsin from Sudbury’s (Fig. 8).

All the observed *Betula papyrifera* diagnostic markers were observed in the St. Charles population. However, the diagnostic marker for *Betula pumila* was observed either in the populations from Wisconsin or the GSR.

In previous studies, Kalubi et al. (2015) identified an ISSR marker that differentiated red maple (*Acer rubrum*) populations from metal contaminated sites and uncontaminated areas in Northern Ontario. Likewise, Theriault et al. (2014) characterized two population–diagnostic ISSR markers in *B. papyrifera* from Northern Ontario that showed difference between St. Charles population and other populations from the Greater Sudbury areas. This result along with the observations made in the present study suggested that *B. papyrifera* population from St. Charles represent a distinct population from other GSR populations which might be all from a common source. Since no-species diagnostic markers were detected in the present study, the validation of putative *P. pumila* x *B. papyrifera* hybrids using molecular was not possible.

Thomson et al. (2015) reported that despite a high incidence of allele sharing among *B. papyrifera*, *B. alleghaniensis* and *B. lenta*, all of the species were significantly differentiated even within zones of sympatry using nuclear microsatellite markers. They identified putatively admixed individuals using Bayesian model-based clustering in their study. This theoretical classification of hybridity is usually not informative. Genomic *in situ* hybridization would be the most accurate approach to characterize putative hybrids and to determine the level of introgression between two species (Nkongolo et al., 2009; Silva and Souza, 2013).

For the present study, since *B. pumila* and *B. papyrifera* coexist in close proximity in GSR, the
hybridization might result in the development of fertile or semi-fertile hybrids. Backcrossing of these progenies with the parents can then lead to introgression and creation of novel genotypes (Oberprieler et al., 2010). The other scenario described by Rieseberg and Willis (2007) is that hybrids may through karyotypic and ecological divergence or polyploidization become reproducitively isolated from the parents resulting in limited gene flow among parental species. In fact, in some species, pollen competition can act as a partial reproductive barrier limiting the frequency of hybrid formation (Rieseberg et al., 1995; 1998; Lepais et al., 2009; Lepais and Gerber, 2011). This last alternative is unlikely in the GSR because of a large pollen production and the lack of pollination barriers between the two species and their hybrids in sampled areas. This is documented by large frequencies of hybrid genotypes.

Conclusion

In conclusion, this study revealed extensive introgression of B. papyrifera genes into B. pumila based on morphological characterization in the Greater Sudbury areas where these two species are present. ISSR and RAPD analyses confirmed a high level of polymorphic loci in populations from the two species making the development of species-specific molecular markers challenging. Population-diagnostic markers were identified between populations from different origins.

Acknowledgement

This study was supported by the Natural Sciences and Engineering Council of Canada (NSERC). Thanks to Dr. Gabriel Theriault for assistance with B. papyrifera sampling and processing.

Author’s Contributions

Nastaran Moarefi: Conduct all the molecular work and data analysis.

Paul Michael: Monitor molecular experiences and data analysis

Peter Beckett: Coordinare field sampling and hybrid identification

Kabwe Nkongolo: Coordinate the project and wrote the manuscript.

Conflict of Interest

Authors declare that they have no conflict of interest.

Compliance with Ethical Standards

This article does not contain any studies with human subjects or animals performed by any of the authors.

References


