# BioOne COMPLETE

# A Molecular Identification Protocol for Roots of Boreal Forest Tree Species

Authors: Randall, Morgan J., Karst, Justine, Pec, Gregory J., Davis, Corey S., Hall, Jocelyn C., et al.

# A MOLECULAR IDENTIFICATION PROTOCOL FOR ROOTS OF BOREAL FOREST TREE SPECIES[1]

MORGAN J. RANDALL[2,3,5], JUSTINE KARST[3,4], GREGORY J. PEC[3], COREY S. DAVIS[3], JOCELYN C. HALL[3], AND JAMES F. CAHILL JR.[3]

[2]Department of Integrative Biology, University of Guelph, Guelph, Ontario, N1G 2W1, Canada; [3]Department of Biological Sciences, University of Alberta, Edmonton, Alberta, T6G 2E9, Canada; and [4]Department of Renewable Resources, University of Alberta, Edmonton, Alberta, T6G 2H1, Canada

- *Premise of the study:* Roots play a key role in many ecological processes, yet our ability to identify species from bulk root samples is limited. Molecular tools may be used to identify species from root samples, but they have not yet been developed for most systems. Here we present a PCR-based method previously used to identify roots of grassland species, modified for use in boreal forests.
- *Methods:* We used repeatable interspecific size differences in fluorescent amplified fragment length polymorphisms of three noncoding chloroplast DNA regions to identify seven woody species common to boreal forests in Alberta, Canada.
- *Results: Abies balsamea, Alnus crispa, Betula papyrifera, Pinus contorta,* and *Populus tremuloides* were identifiable to species, while *Picea glauca* and *Picea mariana* were identifiable to genus. In mixtures of known composition of foliar DNA, species were identified with 98% accuracy using one region. Mixed root samples of unknown composition were identified with 100% accuracy; four species were identified using one region, while three species were identified using two regions.
- *Discussion:* This methodology is accurate, efficient, and inexpensive, and thus a valuable approach for ecological studies of roots. Furthermore, this method has now been validated for both grassland and boreal forest systems, and thus may also have applications in any plant community.

  **Key words:** boreal forest; FAFLP; molecular identification; roots; *trnL* intron; *trnT-trnL* intergenic spacer.

Knowing the species identity of plant roots is important for several central questions in ecology including linking patterns of above- to belowground richness (Frank et al., 2010; Kesanakurti et al., 2011; Hiiesalu et al., 2012; Pärtel et al., 2012), determining drivers of belowground community composition and diversity (Frank et al., 2010), and understanding plant-microbe interactions (Saari et al., 2005; Beiler et al., 2010; Toju et al., 2013). Ecological studies of belowground processes typically generate numerous soil samples that contain mixtures of tangled roots from different species. Roots, however, are notoriously difficult to identify with traditional methods as used for flowers and plant shoots. Consequently, a variety of approaches have been developed to allow species identification of root samples (Cahill and McNickle, 2011; Rewald et al., 2012). In this study, we used a PCR-based method to identify woody species from bulk root samples. Other researchers have used near-infrared spectroscopy and plant wax markers to estimate the proportion of root biomass per species in known mixtures

(Roumet et al., 2006; Lei and Bauhus, 2010). Another approach is the use of stable carbon isotopes to discriminate between $C_3$ and $C_4$ plants (Polley et al., 1992; Gealy et al., 2013). Generally these approaches have been successful at identifying roots in mixtures, but molecular techniques (e.g., restriction fragment length polymorphism [RFLP], fluorescent amplified fragment length polymorphism [FAFLP], next-generation sequencing) hold the most promise because unlike morphological, anatomical, and chemical characteristics, DNA-based molecular markers do not vary depending on biotic and abiotic conditions (Rewald et al., 2012).

The majority of DNA-based species identification techniques were developed for grassland species (Moore and Field, 2005; Frank et al., 2010; Cahill and McNickle, 2011; Haling et al., 2011; Kesanakurti et al., 2011; Taggart et al., 2011; Hiiesalu et al., 2012; Pärtel et al., 2012; Wallinger et al., 2012) with limited coverage of boreal forest species, especially trees, with the exception of two studies that examined *Pinus* and *Picea* species (Govindaraju et al., 1992; Jaramillo-Correa et al., 2003). This lack of information is problematic as roots in forests make up a considerable portion of terrestrial productivity (i.e., 2–5 kg·m$^{-2}$; Jackson et al., 1996), drive community composition by competing for soil resources (Casper and Jackson, 1997; Cahill and McNickle, 2011), and contribute to forest succession from soil bud banks (Bobowski et al., 1999). Tree roots also influence the hydrology (Nepstad et al., 1994) and nutrient cycling (Nepstad et al., 1994; Ruess et al., 2003) of forest soils through the turnover and activity of roots themselves, and also by interacting with ectomycorrhizal fungi (Simard et al., 1997; Plamboeck et al.,

2007). Our understanding of how woody roots are involved in these ecological processes is currently limited by the ability to identify tree species in the field (Beiler et al., 2010).

The ideal technique for identifying species from root samples would be time and cost effective, accurate, and could be applied to samples containing mixed species (i.e., multiple species in a single sample). Initial studies identifying woody plant roots using molecular methods were limited to analyzing small sections of roots to ensure high-quality data (Jackson et al., 1999; Linder et al., 2000). Since then, RFLP analysis has been used to identify woody species, but requires time-consuming downstream reactions using restriction enzymes (Bobowski et al., 1999; Brunner et al., 2001; Jaramillo-Correa et al., 2003; Yanai et al., 2008). Using species-specific primers has proven to be an accurate method for identifying species (McNickle et al., 2008; Mommer et al., 2008); however, designing primers can be difficult, time consuming, and is dependent on the quality of sequence data available in reference libraries (Mommer et al., 2008, 2011). More recently, high-throughput sequencing techniques (e.g., 454 pyrosequencing, metabarcoding) have been developed, but these technologies remain somewhat expensive and are often limited to shorter DNA fragments (i.e., <500 bp; Valentini et al., 2009; Burgess et al., 2011; Yoccoz et al., 2012). Similar to RFLP analysis, identification using FAFLPs relies on differences in amplified fragment sizes and circumvents sequencing, but has the added benefit of avoiding downstream reactions with restriction enzymes (Ridgway et al., 2003; Rewald et al., 2012). Using FAFLP analysis of the *trnL* intron and the *trnT-trnL* and *trnL-trnF* intergenic spacers, Taggart et al. (2011) successfully identified 80% of 95 species in a grassland community.

Building on the work of Taggart et al. (2011), we developed methods for FAFLP analysis of seven common woody species in the boreal forest of Alberta: *Abies balsamea* (L.) Mill., *Alnus crispa* (Aiton) Pursh, *Betula papyrifera* Marshall, *Picea glauca* (Moench) Voss, *Picea mariana* (Mill.) Britton, Sterns & Poggenb., *Pinus contorta* Douglas ex Loudon, and *Populus tremuloides* Michx. (Natural Regions Committee, 2006). We chose to focus on ectomycorrhizal host species because studies documenting ectomycorrhizal host specificity require methods to identify root tips of trees (Kennedy et al., 2003; Tedersoo et al., 2008). Yet, many researchers still rely on tracing roots to hosts or harvesting whole seedlings for identification (Sakakibara et al., 2002; Ishida et al., 2007; Bent et al., 2011). Here we outline a reliable approach to identify roots of these species by comparing FAFLP size differences of three noncoding chloroplast DNA (cpDNA) regions: the *trnT-trnL* intergenic spacer, the *trnL* intron, and the *trnL-trnF* intergenic spacer. First, we generated a reference FAFLP size key for each of the species studied, then tested the accuracy of this technique at identifying single species and in mixtures of two, four, and six species using DNA isolated from leaf tissue. Next, we used the method to distinguish among species within mixed root soil samples collected from the field. Finally, we compare the efficiency and cost of this technique to other current molecular methods.

## METHODS

***Site characteristics***—Foliar and root samples were collected from 11 sites located near Grande Prairie, Alberta, Canada (Appendix 1), within the Lower Foothills subregion of the Foothills Natural Region (Natural Regions Committee, 2006). This ecoregion is characterized by pure and mixed stands of lodgepole pine (*Pinus contorta*), as well as mixed stands of aspen (*Populus tremuloides*) and white spruce (*Picea glauca*). White birch (*Betula papyrifera*), balsam fir

(*Abies balsamea*), and black spruce (*Picea mariana*) are also common (Natural Regions Committee, 2006). The focal species in this study were chosen because of their dominance in the plant community, their ubiquitous distribution across western North America, and their ability to form ectomycorrhizas.

### Generation of FAFLP size key using foliar DNA

***Foliar sample collection and DNA isolation from leaf tissue***—Leaf tissue from 43 individuals consisting of seven common woody species (*Abies balsamea*, *Alnus crispa*, *Betula papyrifera*, *Picea glauca*, *Picea mariana*, *Pinus contorta*, and *Populus tremuloides*) was sampled in June 2011. Leaf samples were stored with Drierite (W. A. Hammond Drierite Co. Ltd., Xenia, Ohio, USA) at room temperature. Specimen vouchers are deposited at the University of Alberta Vascular Plant Herbarium (ALTA; Appendix 2). Total genomic DNA was extracted using a QIAGEN DNeasy Plant Mini Kit (QIAGEN, Mississauga, Ontario, Canada) following manufacturer's instructions and subsequently stored at −20°C.

***PCR amplification of foliar DNA***—Three regions were examined in total using DNA isolated from foliar samples. Three cpDNA noncoding regions—the *trnT-trnL* intergenic spacer, the *trnL* intron, and the *trnL-trnF* intergenic spacer—were amplified with universal primers: (i) the *trnT-trnL* intergenic spacer with primers A (5′-CATTACAAATGCGATGCTCT-3′) and B (5′-TCTACCGATTTCGCCATATC-3′), (ii) the *trnL* intron with primers C (5′-CGAAATCGGTAGACGCTACG-3′) and D (5′-GGGGATAGAGGGACTTGAAC-3′), and (iii) the *trnL-trnF* intergenic spacer with primers E (5′-GGTTCAAGTCCCTCTATCCC) and F (5′-ATTTGAACTGGTGACACGAG; Taberlet et al., 1991). One primer each per pair was labeled with a fluorescent dye (FAM_primer A: Integrated DNA Technologies, Coralville, Iowa, USA; VIC_primer C: Applied Biosystems, Life Technologies, Carlsbad, California, USA; and NED_primer E: Applied Biosystems). PCR reactions for the *trnT-trnL* intergenic spacer and the *trnL* intron totaled 25 μL: 5.0 μL 5× e2TAK buffer (Mg$^{2+}$ plus; TaKaRa Bio, Otsu, Shiga, Japan), 2.0 μL dNTP Mixture (TaKaRa Bio), 0.15 μM of each the forward and reverse primer, 1–39 ng genomic DNA, and 0.25 units of e2TAK *Taq* polymerase (TaKaRa Bio). PCR amplification for the *trnL-trnF* intergenic spacer was performed in 25-μL reactions using 0.1 μM of each forward and reverse primer, 1–39 ng of genomic DNA, 11 μL of nuclease-free water (Life Technologies), and 12.5 μL of EconoTaq PLUS 2× Master Mix (Lucigen Corporation, Middleton, Wisconsin, USA). PCR reactions were performed using an Eppendorf Mastercycler Pro gradient thermal cycler (Model 6321; Eppendorf Canada, Mississauga, Ontario, Canada). Each region had unique thermal cycler conditions (Table 1). Products were verified on a 1.5% agarose gel and visualized using ethidium bromide.

***Generation of FAFLP size key***—A reference size key of FAFLP amplicon sizes was generated by analyzing single species with three to eight replicates per species using foliar DNA (Table 2; Fig. 1). PCR products were diluted 200× by combining 396 μL of distilled H$_2$O and 2 μL of PCR product from each of the regions examined. From this dilution, 2 μL were added to 8 μL of Hi-Di Formamide and 0.3 μL of GeneScan 1200 LIZ Size Standard (Applied Biosystems). The final mixture was centrifuged for 30 s at 10,000 rpm, then denatured at 95°C for 2 min and coldsnapped to maintain single-stranded fluorescently labeled DNA. Sizes of pooled PCR amplicons were first resolved using capillary electrophoresis (ABI 3730 DNA analyzer, Applied Biosystems) and then sized with GeneMapper 4.0 software (Applied Biosystems) with GeneScan 1200 LIZ Size Standard (Applied Biosystems). Fragment sizes read by the capillary sequencer were rounded to the nearest base pair (Fig. 1).

***FAFLP analysis of single and mixed species samples***—When analyzing mixed samples of DNA using PCR-based methods, false negatives (failing to detect a species known to be present) can occur both preamplification and postamplification. Prior research has indicated that false negatives can occur preamplification due to difficulties in DNA extraction, species-specific differences in amplification, and competition among primers (Kesanakurti et al., 2011; Mommer et al., 2011; Taggart et al., 2011). Taggart et al. (2011) investigated whether preamplification competition among primers affects species detection by combining template DNA prior to PCR amplification to determine the maximum number of species detectable in mixtures, and thus we did not pursue this in the current study. False negatives may also occur postamplification as fluorescently labeled amplicons are injected into capillaries. For negatively charged DNA, shorter fragments have greater mobility into the capillaries during injection and therefore may be more detectable than longer fragments. We tested for

TABLE 1. Thermocycler conditions for amplification of the *trnT-trnL* intergenic spacer, the *trnL* intron, and the *trnL-trnF* intergenic spacer (Taggart et al., 2011).

| Thermocycler conditions | *trnT-trnL* intergenic spacer | *trnL* intron | *trnL-trnF* intergenic spacer |
|---|---|---|---|
| Initial denaturation | 94°C 5 min | 94°C 5 min | 94°C 5 min |
| Cycles of denaturation, annealing, extension | 2 cycles of 94°C 45 s, 56°C 60 s, 72°C 80 s | 2 cycles of 94°C 60 s, 60°C 60 s, 72°C 80 s | 2 cycles of 94°C 60 s, 60°C 60 s, 72°C 80 s |
| | 33 cycles of 94°C 45 s, 61.5–0.3°C/cycle 60 s, 72°C 80 s | 33 cycles of 9°C 60 s, 59.6–0.4°C/cycle 60 s, 72°C 80 s | 33 cycles of 94°C 60 s, 63–0.4°C/cycle 60 s, 72°C 80 s |
| Final extension | 72°C 30 min | 72°C 30 min | 72°C 30 min |

*Note*: min = minute; s = second.

competition among different sizes of fragments by combining amplicons from different species (and thus different sizes) postamplification, using DNA isolated from leaf tissue. In the case of competition, longer fragments may not be detectable in mixtures. Samples containing one ($n = 40$), two ($n = 15$), four ($n = 15$), and six ($n = 1$) species were analyzed to determine how accurately we could detect species in mixtures. We could only include a maximum of six species in combination because the congeneric *Picea* species are indistinguishable using the two amplifiable regions in this study (see below).

In preparation for capillary electrophoresis, PCR products of one individual per species were serially diluted 400× without first quantifying. We use this approach because DNA abundance is expected to vary among species in environmental samples (Haling et al., 2011; Mommer et al., 2011). Thus, these reactions vary in DNA concentrations across samples and are consistent with the approach described above to generate the reference FAFLP size key. From the diluted products, 2 µL were added to 8 µL of Hi-Di Formamide and 0.3 µL of GeneScan 1200 LIZ Size Standard (Applied Biosystems). The final mixture was then prepared for capillary electrophoresis as previously described. Presence or absence of a species in a mixture was determined based on a set of liberal criteria as described by Taggart et al. (2011), where the presence of a peak for at least one region signifies the presence of a species in a mixture.

### FAFLP analysis of mixed root samples

*Soil sample collection and DNA isolation from root tissue*—Thirty-two soil cores (5 cm diameter, 20 cm deep) were collected in June 2012 from a subset of sites (see Appendix 1) used for leaf tissue collection. Foliar and root samples were not taken from the same individuals. The identity of the nearest tree species within 1 m of a soil core was taken to aid in root identification. Soil samples were transported on ice, then frozen at −20°C until processed. Soil samples were thawed, and fine roots were washed and sieved. Subsamples of 250 mg of fine root tissue were placed in a prechilled freeze-dryer (VirTis Freezermobile FM25XL; SP Scientific, Warminster, Pennsylvania, USA) at −45°C, lyophilized for 24 h, and twice ground to a fine powder using a mixer mill (Retsch Type MM 301; Retsch GmbH, Haan, Germany) for 1 min at 25.0 Hz. DNA extraction from roots can be sensitive to secondary metabolites found in lignified root material (Linder et al., 2000). For this reason, genomic DNA was isolated from ground fine root tissue using a hexadecyltrimethylammonium bromide (CTAB) protocol according to Roe et al. (2010) with one modification: pellets were resuspended in 50 µL of nuclease-free water (Life Technologies) with gentle agitation.

*PCR amplification and FAFLP analysis of DNA isolation of root tissue*— Despite troubleshooting efforts, amplification of the *trnT-trnL* intergenic spacer was unreliable for mixed root samples and was discarded from further analysis.

TABLE 2. Species included in this study and number of individuals sampled for foliar tissue per species.

| Family | Species | N |
|---|---|---|
| Betulaceae | *Alnus crispa* (Aiton) Pursh | 6 |
| Betulaceae | *Betula papyrifera* Marshall | 3 |
| Pinaceae | *Abies balsamea* (L.) Mill. | 6 |
| Pinaceae | *Picea glauca* (Moench) Voss | 8 |
| Pinaceae | *Picea mariana* (Mill.) Britton, Sterns & Poggenb. | 8 |
| Pinaceae | *Pinus contorta* Loudon | 7 |
| Salicaceae | *Populus tremuloides* Michx. | 5 |

*Note*: N = number of individuals.

The remaining two cpDNA noncoding regions, (i) the *trnL* intron and (ii) the *trnL-trnF* intergenic spacer, were examined using DNA isolated from mixed root samples. PCR reactions of root samples were similar to those of foliar samples (see Table 1). PCR products were verified on a 1.5% agarose gel using SYBR Safe DNA gel stain (Life Technologies). The products were prepared for capillary electrophoresis as previously described for foliar DNA, with the following modifications: the final mixture was centrifuged for 15 s at 1000 rpm (rather than 10,000 rpm), denatured at 95°C for 2 min, and then coldsnapped.

Several criteria following the work of Taggart et al. (2011) were used to indicate species presence in a mixed root sample. Two identification assumptions were used: (1) liberal, in which a species is identified as present in a sample when a known fragment size or peak (Fig. 1B, C) is found for one of the two chloroplast (cpDNA) noncoding regions, and (2) conservative, in which a species is identified as present in a sample only if known peaks are present for both chloroplast (cpDNA) noncoding regions (Fig. 1B, C). We further used a constrained approach on both the liberal and conservative assumption, where the species pool was limited to only those tree species found within a 1-m diameter of a soil core.

## RESULTS

***Generation of FAFLP size key using foliar DNA***—We were able to amplify DNA from 40 of the 43 individuals sampled. From single species samples, PCR amplicon sizes were obtained for all seven species for the *trnT-trnL* intergenic spacer, the *trnL* intron, and the *trnL-trnF* intergenic spacer using foliar DNA (Fig. 1). Intraspecific variation in FAFLP size was observed in one or more regions in all seven species (Fig. 1). Intraspecific variation in fragment length was observed in *Betula papyrifera*, *Picea mariana*, *Pinus contorta*, and *Populus tremuloides* for the *trnT-trnL* intergenic spacer (Fig. 1A); *Picea glauca*, *P. mariana*, and *P. tremuloides* for the *trnL* intron (Fig. 1B); and in *Abies balsamea*, *Alnus crispa*, *P. contorta*, *P. glauca*, and *P. mariana* for the *trnL-trnF* intergenic spacer (Fig. 1C). Interspecific variation was observed between *P. glauca* and *P. mariana* in both the *trnT-trnL* intergenic spacer and the *trnL* intron (Fig. 1A, B), and between *A. crispa*, *B. papyrifera*, *P. glauca*, and *P. mariana* in the *trnL-trnF* intergenic spacer (Fig. 1C). Additionally, interspecific variation was observed between *A. crispa* and *P. contorta* in the *trnL-trnF* intergenic spacer (Fig. 1C). Unique fragment sizes were found for five of the seven species studied for the *trnT-trnL* intergenic spacer (Fig. 1A; *A. crispa*, *A. balsamea*, *B. papyrifera*, *P. contorta*, and *P. tremuloides*), four species for the *trnL* intron (Fig. 1B; *A. crispa*, *B. papyrifera*, *P. contorta*, and *P. tremuloides*), and three species for the *trnL-trnF* intergenic spacer (*A. balsamea*, *P. contorta*, and *P. tremuloides*). Five species had unique fragment lengths using two of the three regions (*trnT-trnL* intergenic spacer and *trnL-trnF* intergenic spacer); the remaining *Picea* species had identical sizes for all three regions studied (Fig. 1). In summary, we were able to identify all six genera studied from two regions; however, we were unable to discern between

**A**   *trnT-trnL* intergenic spacer



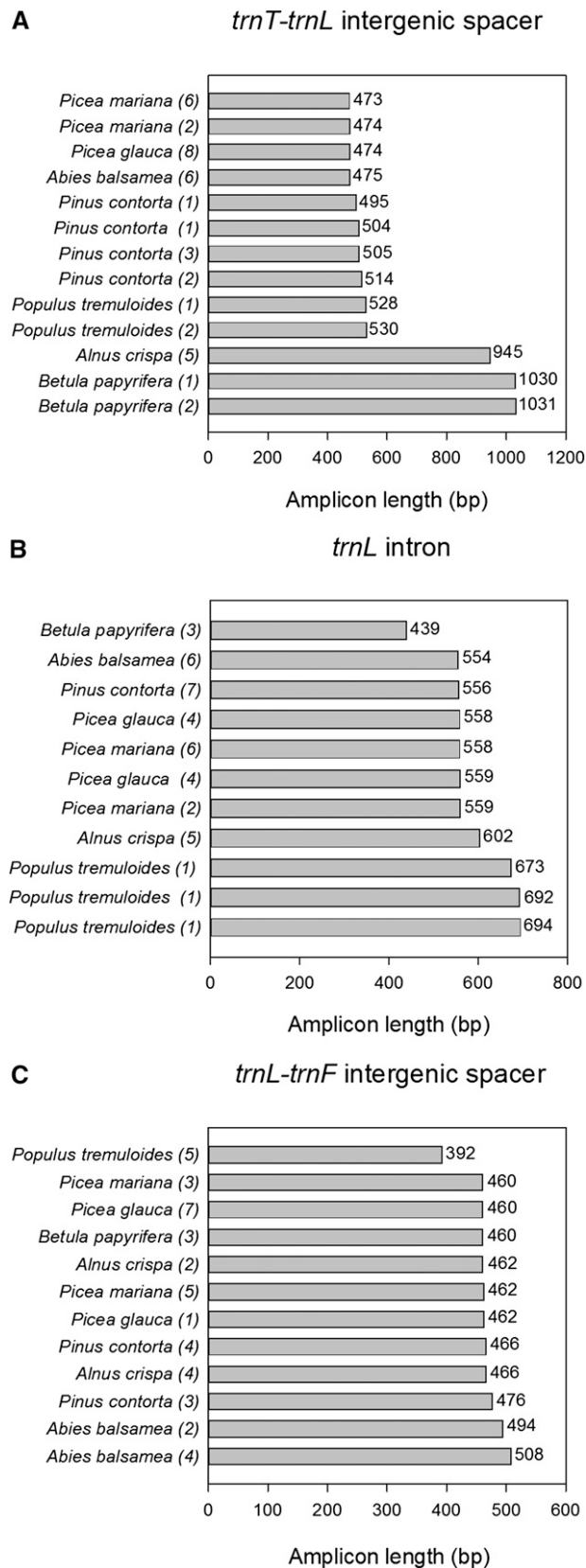**B**   *trnL* intron



**C**   *trnL-trnF* intergenic spacer



Fig. 1.   Fluorescent amplicon lengths for (A) the *trnT-trnL* intergenic spacer, (B) *trnL* intron, and (C) the *trnL-trnF* intergenic spacer, rounded to the nearest base pair for each species based on foliar samples.

congeneric *Picea* species. Using a single region (the *trnL* intron), we successfully identified species in mixtures containing two, four, and six species with a 0% false-positive rate and a 2% false-negative rate.

***FAFLP analysis of mixed root samples***—Across all 32 mixed root samples, we were able to detect the presence of a tree genus within a 1-m diameter in 100% of the soil cores using only one of the two amplifiable regions (*trnL* intron, 0% false-negative rate; Fig. 2). This number decreased to a rate of 88% when trying to discern both *Picea* species. When using the presence of known peaks in both the *trnL* intron and *trnL-trnF* intergenic spacer regions (i.e., conservative criteria), our detection rate for a known species within a 1-m diameter of a soil core decreased to 25% (a 75% false-negative rate; Fig. 2). However, using the liberal criteria, we were able to identify known tree species with a 100% success rate when looking only at known peaks of the *trnL-trnF* intergenic spacer region (*P. tremuloides* = 392; *P. contorta* = 476; *A. balsamea* = 494, 508) and integrating them with the peaks of the *trnL* intron region.

Although we targeted tree species, several mixed root soil samples contained "extra" species (*n* = 5) characterized by different FAFLP sizes than those included in the size key based on foliar samples. These may have been roots of shrubs and herbaceous species, which a size profile had not been created for, as the goals of this project were not dependent upon the use of these additional species. Alternatively, the "extra" species detected could indicate additional intraspecific variation of the species included that was not captured in this study, possibly due to the small number of foliar samples.

## DISCUSSION

We were successful in developing an accurate method to detect the presence of boreal tree species from mixed DNA samples. This general protocol has now been used in both grassland and forest systems, facilitated by universal plant primers for PCR (Taberlet et al., 1991). This method has promising applications in the field and to additional plant communities beyond the boreal forest and grassland regions.

In mixed DNA samples isolated from leaf tissue, we were able to identify five out of seven species in mixtures 98% of the time (2% false-negative rate) using the *trnT-trnL* intergenic spacer. In mixed DNA samples isolated from root tissue, we were able to identify four out of seven species using the *trnL* intron and three out of seven species using the *trnL-trnF* intergenic spacer and both the *trnL* intron and the *trnL-trnF* intergenic spacer regions combined. The general increase in the false-negative rate when including the *trnL-trnF* region in analyzing mixed roots appears to be because of: (1) increased interspecific variation within the *trnL-trnF* intergenic spacer, and (2) the intraspecific variation found between both *Picea* species as well as *B. papyrifera*, *P. contorta*, and *A. crispa*. However, we were able to distinguish among genera with 100% success and encountered no false positives using only the *trnL* intron region, which contained little intraspecific variation.

We were unable to discern among *Picea* species, although the genus *Picea* was differentiated from the other five genera. Many existing molecular techniques have issues with discriminating among closely related species (Bobowski et al., 1999; Linder et al., 2000; Brunner et al., 2001; Frank et al., 2010; Burgess et al., 2011; Kesanakurti et al., 2011; Mommer et al.,
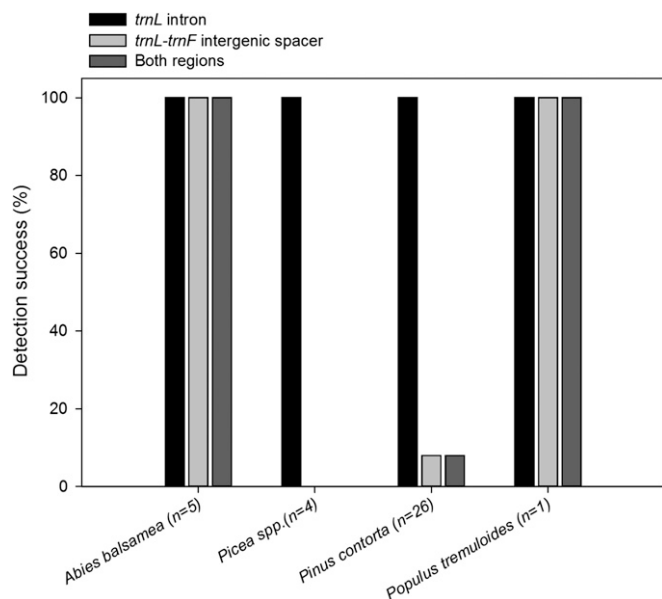
Fig. 2. Detection success of boreal tree species (*Abies balsamea*, *Picea* spp., *Pinus contorta*, and *Populus tremuloides*) using the *trnL* intron, the *trnL-trnF* intergenic spacer, or both regions based on mixed root samples collected in June 2012 from sites near Grande Prairie, Alberta, Canada.

2011; Taggart et al., 2011). Including another region will increase the likelihood of discerning among congeneric species and many universal plant primers have been described to facilitate this (e.g., Shaw et al., 2005, 2007).

Our method has several important advantages over other molecular techniques. Compared to RFLP, FAFLP analysis saves time and labor by avoiding restriction digests; without accounting for time spent on DNA extraction or subsequent data analysis, each mixed sample took approximately three to five minutes of handling time. Run times for sequencing range between 0.5 h and 14 d (Glenn, 2011; Shokralla et al., 2012), compared to capillary electrophoresis that takes between 1–2 h per run. Our method is inexpensive (approximately US$0.50 per sample) compared to sequencing that costs between US$0.10 per megabase (Mb) and US$10/Mb, although next-generation sequencing (NGS) platforms are quickly becoming more affordable (Glenn, 2011). Other than a slight increase in cost for additional reagents needed for analyzing mixed samples, fragment sizing does not incur a large increase in cost when more regions are added because additional amplicons can be labeled with different fluorescent dyes. For a comprehensive comparison of cost and performance among different NGS platforms, see Glenn (2011) and Shokralla et al. (2012). Using a capillary sequencer, up to four regions can be analyzed simultaneously. Moreover, the use of universal primers requires no prior knowledge of sequences to design species-specific primers. Sequence information is superfluous when the objective is to determine species identity, in which case PCR-based approaches save time, labor, and resources.

A limitation of PCR-based methods is that they often require successful amplification of two or more regions for each species. This is more relevant to studies where a greater number of species are present (e.g., Taggart et al., 2011; De Barba et al., 2014). Another requirement for PCR-based methods is a complete reference database. Similar to how sequencing studies

require reference databases such as BOLD or GenBank, FAFLP analysis requires a database of FAFLP sizes for all species in a community. This is important because unknown species may have identical amplicon lengths when using FAFLP analysis to identify mixed root samples from the field. Consequently, PCR-based techniques used on mixed samples often have problems with numerous false positives (detecting a species known to be absent), especially in species-rich plant communities (Taggart et al., 2011; Rewald et al., 2012). Creating a reference database by surveying plant species aboveground greatly reduces the occurrence of false positives (Taggart et al., 2011). Regardless, there is smaller risk of encountering false positives when identifying mixed samples from boreal forests because they are relatively species-poor compared to other plant communities. Anatomical and morphological characteristics can be used to filter out roots of shrubs and herbaceous species from samples, so a reference database of those species may not be necessary in studies focusing on tree roots.

Advances in sequencing technology have made NGS faster and more affordable. As a result, NGS is beginning to have applications across a multitude of different disciplines, such as studying plant belowground richness (Hiiesalu et al., 2012), diet analysis (De Barba et al., 2014), and root-fungal associations (Toju et al., 2013). For example, Toju et al. (2013) used 454 pyrosequencing to identify fungal species and their associated plant hosts from root samples, but this approach required two PCRs per root sample, thus adding to overall costs. One recent study used a combination of metabarcoding and NGS to identify plant and animal DNA to investigate diet components of brown bears (De Barba et al., 2014). By amplifying multiple DNA markers (the *trnL* intron, internal transcribed spacer region 1 [ITS1], and internal transcribed spacer region 2 [ITS2]) simultaneously, they streamlined the PCR process and increased species resolution for Poaceae, Asteraceae, Cyperaceae, and Rosaceae. The ITS1 and ITS2 regions were included because species resolution using the *trnL* intron region was limited for these families. Ultimately, knowledge of local plant distributions aboveground was still used to identify 28 out of 60 plant species (De Barba et al., 2014).

While NGS methods will likely become more readily used and accessible, it may not be the most appropriate choice for all studies. A major challenge to NGS is data processing because these approaches generate millions of reads per run, which requires substantial computing demands for storage and analysis. Moreover, the researcher is required to learn a range of data pipelines that are often unique to different sequencing platforms (Harrison and Kidner, 2011; Egan et al., 2012). These challenges aside, these studies highlight the future potential of NGS technologies as a method of identifying roots, yet future work is needed to further simplify the process of data filtering and analysis. We suggest that while sequencing is critical and appropriate for some investigations, the FAFLP method remains a useful tool for ecological studies generating hundreds of samples, where specific sequence information is not required.

The FAFLP method described here is straightforward, quick, and requires minimal resources. Furthermore, the use of universal plant primers permits the application of this method to plant communities worldwide (Taberlet et al., 1991). We acknowledge that the verification procedure used does not include two important aspects for application to the field: variable abundances of species in mixtures and species-specific biases in DNA extraction. Instead, our goal was to show proof of concept

of this method. Nevertheless, this protocol was effective at identifying woody boreal forest species from roots and can easily be applied to field-based studies.

## LITERATURE CITED

BEILER, K. J., D. M. DURALL, S. W. SIMARD, S. A. MAXWELL, AND A. M. KRETZER. 2010. Architecture of the wood-wide web: *Rhizopogon* spp. genets link multiple Douglas-fir cohorts. *New Phytologist* 185: 543–553.

BENT, E., P. KIEKEL, R. BRENTON, AND D. L. TAYLOR. 2011. Root-associated ectomycorrhizal fungi shared by various boreal forest seedlings naturally regenerating after a fire in interior Alaska and correlation of different fungi with host growth responses. *Applied and Environmental Microbiology* 77: 3351–3359.

BOBOWSKI, B. R., D. HOLE, P. G. WOLF, AND L. BRYANT. 1999. Identification of roots of woody species using polymerase chain reaction (PCR) and restriction fragment length polymorphism (RFLP) analysis. *Molecular Ecology* 8: 485–491.

BRUNNER, I., S. BRODBECK, U. BÜCHLER, AND C. SPERISEN. 2001. Molecular identification of fine roots of trees from the Alps: Reliable and fast DNA extraction and PCR-RFLP analyses of plastid DNA. *Molecular Ecology* 10: 2079–2087.

BURGESS, K. S., A. J. FAZEKAS, P. R. KESANAKURTI, S. W. GRAHAM, B. C. HUSBAND, S. G. NEWMASTER, D. M. PERCY, ET AL. 2011. Discriminating plant species in a local temperate flora using the *rbcL+matK* DNA barcode. *Methods in Ecology and Evolution* 2: 333–340.

CAHILL JR., J. F., AND G. G. McNICKLE. 2011. The behavioral ecology of nutrient foraging by plants. *Annual Review of Ecology Evolution and Systematics* 42: 289–311.

CASPER, B. B., AND R. B. JACKSON. 1997. Plant competition underground. *Annual Review of Ecology Evolution and Systematics* 28: 545–570.

DE BARBA, M., C. MIQUEL, F. BOYER, C. MERCIER, D. RIOUX, E. COISSAC, AND P. TABERLET. 2014. DNA metabarcoding multiplexing and validation of data accuracy for diet assessment: Application to omnivorous diet. *Molecular Ecology Resources* 14: 306–323.

EGAN, A. N., J. SCHLUETER, AND D. M. SPOONER. 2012. Applications of next-generation sequencing in plant biology. *American Journal of Botany* 99: 175–185.

FRANK, D. A., A. W. PONTES, E. M. MAINE, J. CARUANA, R. RAINA, S. RAINA, AND J. D. FRIDLEY. 2010. Grassland root communities: Species distributions and how they are linked to aboveground abundance. *Ecology* 91: 3201–3209.

GEALY, D., K. MOLDENHAUER, AND S. DUKE. 2013. Root distribution and potential interactions between allelopathic rice, sprangletop (*Leptochloa* spp.), and barnyardgrass (*Echinochloa crus-galli*) based on $^{13}C$ isotope discrimination analysis. *Journal of Chemical Ecology* 39: 186–203.

GLENN, T. C. 2011. Field guide to next-generation DNA sequencers. *Molecular Ecology Resources* 11: 759–769.

GOVINDARAJU, D., P. LEWIS, AND C. CULLIS. 1992. Phylogenetic analysis of pines using ribosomal DNA restriction fragment length polymorphisms. *Plant Systematics and Evolution* 179: 141–153.

HALING, R. E., R. J. SIMPSON, A. C. McKAY, D. HARTLEY, H. LAMBERS, K. OPHEL-KELLER, S. WIEBKIN, ET AL. 2011. Direct measurement of roots in soil for single and mixed species using a quantitative DNA-based method. *Plant and Soil* 348: 123–137.

HARRISON, N., AND C. A. KIDNER. 2011. Next-generation sequencing and systematics: What can a billion base pairs of DNA sequence data do for you? *Taxon* 60: 1552–1566.

HIIESALU, I., M. ÖPIK, M. METSIS, L. LILJE, J. DAVISON, M. VASAR, M. MOORA, ET AL. 2012. Plant species richness belowground: Higher richness and new patterns revealed by next-generation sequencing. *Molecular Ecology* 21: 2004–2016.

ISHIDA, T. A., K. NARA, AND T. HOGETSU. 2007. Host effects on ectomycorrhizal fungal communities: Insight from eight host species in mixed conifer–broadleaf forests. *New Phytologist* 174: 430–440.

JACKSON, R. B., J. CANADELL, J. R. EHLERINGER, H. A. MOONEY, O. E. SALA, AND E. D. SCHULZE. 1996. A global analysis of root distributions for terrestrial biomes. *Oecologia* 108: 389–411.

JACKSON, R. B., L. A. MOORE, W. A. HOFFMANN, W. T. POCKMAN, AND C. R. LINDER. 1999. Ecosystem rooting depth determined with caves and DNA. *Proceedings of the National Academy of Sciences, USA* 96: 11387–11392.

JARAMILLO-CORREA, J. P., J. BOUSQUET, J. BEAULIEU, N. ISABEL, M. PERRON, AND M. BOUILLÉ. 2003. Cross-species amplification of mitochondrial DNA sequence-tagged-site markers in conifers: The nature of polymorphism and variation within and among species in *Picea*. *Theoretical and Applied Genetics* 106: 1353–1367.

KENNEDY, P. G., A. D. IZZO, AND T. D. BRUNS. 2003. There is high potential for the formation of common mycorrhizal networks between understorey and canopy trees in a mixed evergreen forest. *Journal of Ecology* 91: 1071–1080.

KESANAKURTI, P. R., A. J. FAZEKAS, K. S. BURGESS, D. M. PERCY, S. G. NEWMASTER, S. W. GRAHAM, S. C. H. BARRETT, ET AL. 2011. Spatial patterns of plant diversity below-ground as revealed by DNA barcoding. *Molecular Ecology* 20: 1289–1302.

LEI, P., AND J. BAUHUS. 2010. Use of near-infrared reflectance spectroscopy to predict species composition in tree fine-root mixtures. *Plant and Soil* 333: 93–103.

LINDER, C. R., L. A. MOORE, AND R. B. JACKSON. 2000. A universal molecular method for identifying underground plant parts to species. *Molecular Ecology* 9: 1549–1559.

McNICKLE, G. G., J. F. CAHILL JR., AND M. K. DEYHOLOS. 2008. A PCR-based method for the identification of the roots of 10 co-occurring grassland species in mesocosm experiments. *Botany* 86: 485–490.

MOMMER, L., C. A. M. WAGEMAKER, H. DE KROON, AND N. J. OUBORG. 2008. Unravelling below-ground plant distributions: A real-time polymerase chain reaction method for quantifying species proportions in mixed root samples. *Molecular Ecology Resources* 8: 947–953.

MOMMER, L., A. J. DUMBRELL, C. A. M. WAGEMAKER, AND N. J. OUBORG. 2011. Belowground DNA-based techniques: Untangling the network of plant root interactions. *Plant and Soil* 348: 115–121.

MOORE, L. A., AND C. B. FIELD. 2005. A technique for identifying the roots of different species in mixed samples using nuclear ribosomal DNA. *Journal of Vegetation Science* 16: 131–134.

NATURAL REGIONS COMMITTEE. 2006. Natural regions and subregions of Alberta. Compiled by D. J. Downing and W. W. Pettapiece. Pub. No. T/852. Government of Alberta, Edmonton, Alberta, Canada.

NEPSTAD, D. C., C. R. DE CARVALHO, E. A. DAVIDSON, P. H. JIPP, P. A. LEFEBVRE, G. H. NEGREIROS, E. D. DA SILVA, ET AL. 1994. The role of deep roots in the hydrological and carbon cycles of Amazonian forests and pastures. *Nature* 372: 666–669.

PÄRTEL, M., I. HIIESALU, M. ÖPIK, AND S. D. WILSON. 2012. Below-ground plant species richness: New insights from DNA-based methods. *Functional Ecology* 26: 775–782.

PLAMBOECK, A. H., T. E. DAWSON, L. M. EGERTON-WARBURTON, M. NORTH, T. D. BRUNS, AND J. I. QUEREJETA. 2007. Water transfer via ectomycorrhizal fungal hyphae to conifer seedlings. *Mycorrhiza* 17: 439–447.

POLLEY, H. W., H. B. JOHNSON, AND H. S. MAYEUX. 1992. Determination of root biomasses of three species grown in a mixture using stable isotopes of carbon and nitrogen. *Plant and Soil* 142: 97–106.

REWALD, B., C. MEINEN, M. TROCKENBRODT, J. E. EPHRATH, AND S. RACHMILEVITCH. 2012. Root taxa identification in plant mixtures: Current techniques and future challenges. *Plant and Soil* 359: 165–182.

RIDGWAY, K. P., J. M. DUCK, AND J. P. W. YOUNG. 2003. Identification of roots from grass swards using PCR-RFLP and FFLP of the plastid *trn*L (UAA) intron. *BioMed Central Ecology* 3: 8.

ROE, A. D., A. V. RICE, S. E. BROMILOW, J. E. K. COOKE, AND F. A. H. SPERLING. 2010. Multilocus species identification and fungal DNA barcoding: Insights from blue stain fungal symbionts of the mountain pine beetle. *Molecular Ecology Resources* 10: 946–959.

ROUMET, C., C. PICON-COCHARD, L. A. DAWSON, R. JOFFRE, R. MAYES, A. BLANCHARD, AND M. J. BREWER. 2006. Quantifying species composition in roots mixtures using two methods: Near-infrared reflectance spectroscopy and plant wax markers. *New Phytologist* 170: 631–638.

RUESS, R. W., R. L. HENDRICK, A. J. BURTON, K. S. PREGITZER, B. SVEINBJORNSSÖN, M. F. ALLEN, AND G. E. MAURER. 2003. Coupling fine root dynamics with ecosystem carbon cycling in black spruce forests of interior Alaska. *Ecological Monographs* 73: 643–662.

SAARI, S. K., C. D. CAMPBELL, J. RUSSELL, I. J. ALEXANDER, AND I. C. ANDERSON. 2005. Pine microsatellite markers allow roots and ectomycorrhizas to be linked to individual trees. *New Phytologist* 165: 295–304.

SAKAKIBARA, S. M., M. D. JONES, M. GILLESPIE, S. M. HAGERMAN, M. E. FORREST, S. W. SIMARD, AND D. M. DURALL. 2002. A comparison of ectomycorrhiza identification based on morphotyping and PCR-RFLP analysis. *Mycological Research* 106: 868–878.

SHAW, J., E. B. LICKEY, J. T. BECK, S. B. FARMER, W. LIU, J. MILLER, K. C. SIRIPUN, ET AL. 2005. The tortoise and the hare II: Relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany* 92: 142–166.

SHAW, J., E. B. LICKEY, E. E. SCHILLING, AND R. L. SMALL. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: The tortoise and the hare III. *American Journal of Botany* 94: 275–288.

SHOKRALLA, S., J. L. SPALL, J. F. GIBSON, AND M. HAJIBABAEI. 2012. Next-generation sequencing technologies for environmental DNA research. *Molecular Ecology* 21: 1794–1805.

SIMARD, S. W., D. A. PERRY, M. D. JONES, D. D. MYROLD, D. M. DURALL, AND R. MOLINA. 1997. Net transfer of carbon between ectomycorrhizal tree species in the field. *Nature* 388: 579–582.

TABERLET, P., L. GIELLY, G. PAUTOU, AND J. BOUVET. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* 17: 1105–1109.

TAGGART, J. M., J. F. CAHILL JR., G. G. MCNICKLE, AND J. C. HALL. 2011. Molecular identification of roots from a grassland community using size differences in fluorescently labelled PCR amplicons of three cp-DNA regions. *Molecular Ecology Resources* 11: 185–195.

TEDERSOO, L., T. JAIRUS, B. M. HORTON, K. ABARENKOV, T. SUVI, I. SAAR, AND U. KÕLJALG. 2008. Strong host preference of ectomycorrhizal fungi in a Tasmanian wet sclerophyll forest as revealed by DNA barcoding and taxon-specific primers. *New Phytologist* 180: 479–490.

TOJU, H., H. SATO, S. YAMAMOTO, K. KADOWAKI, A. S. TANABE, S. YAZAWA, O. NISHIMURA, AND K. AGATA. 2013. How are plant and fungal communities linked to each other in belowground ecosystems? A massively parallel pyrosequencing analysis of the association specificity of root-associated fungi and their hosts. *Ecology and Evolution* 3: 3112–3124.

VALENTINI, A., F. POMPANON, AND P. TABERLET. 2009. DNA barcoding for ecologists. *Trends in Ecology & Evolution* 24: 110–117.

WALLINGER, C., A. JUEN, K. STAUDACHER, N. SCHALLHART, E. MITTERRUTZNER, E. STEINER, B. THALINGER, AND M. TRAUGOTT. 2012. Rapid plant identification using species- and group-specific primers targeting chloroplast DNA. *PLoS ONE* 7: e29473.

YANAI, R. D., M. C. FISK, T. J. FAHEY, N. L. CLEAVITT, AND B. B. PARK. 2008. Identifying roots of northern hardwood species: Patterns with diameter and depth. *Canadian Journal of Forest Research* 38: 2862–2869.

YOCCOZ, N. G., K. A. BRÅTHEN, L. GIELLY, J. HAILE, M. E. EDWARDS, T. GOSLAR, H. VON STEDINGK, ET AL. 2012. DNA from soil mirrors plant taxonomic and growth form diversity. *Molecular Ecology* 21: 3647–3655.

APPENDIX 1. GPS coordinates of 11 collection sites near Grande Prairie, Alberta, Canada.

| Site name | GPS coordinates |
| --- | --- |
| 070A[a] | 54°44′24.24″N, 118°58′10.08″W |
| 070 | 54°44′24.06″N, 118°58′11.10″W |
| 069[a] | 54°44′12.96″N, 118°57′15.42″W |
| 061[a] | 54°39′27.36″N, 118°59′50.16″W |
| 062[a] | 54°40′7.14″N, 118°59′12.66″W |
| KB1[a] | 54°33′34.53″N, 118°39′48.50″W |
| 057 | 54°38′15.18″N, 118°59′28.32″W |
| 056[a] | 54°38′43.74″N, 118°58′35.46″W |
| 059 | 54°38′44.94″N, 118°59′7.26″W |
| 062A | 54°40′10.68″N, 118°59′17.04″W |
| 063[a] | 54°40′12.36″N, 118°59′29.94″W |

[a] Indicates sites used for collection of mixed root soil samples.

APPENDIX 2. Voucher and location information for species used in this study. All voucher specimens were collected near Grande Prairie, Alberta, Canada, on 25 June 2011.

| Species | Voucher specimen accession no.[a] | Collection ID | Geographic coordinates |
|---|---|---|---|
| *Abies balsamea* | 133046 | JK-1 | 54°44′24.24″N, 118°58′10.08″W |
| *Picea glauca* | 133047 | JK-2 | 54°44′24.24″N, 118°58′10.08″W |
| *Populus tremuloides* | 133048 | JK-3 | 54°44′24.24″N, 118°58′10.08″W |
| *Betula papyrifera* | 133049 | JK-4 | 54°44′24.24″N, 118°58′10.08″W |
| *Betula papyrifera* | 133050 | JK-5 | 54°44′12.96″N, 118°57′15.42″W |
| *Picea mariana* | 133051 | JK-6 | 54°44′12.96″N, 118°57′15.42″W |
| *Picea mariana* | 133052 | JK-7 | 54°44′12.96″N, 118°57′15.42″W |
| *Alnus crispa* | 133053 | JK-8 | 54°39′27.36″N, 118°59′50.16″W |
| *Picea glauca* | 133054 | JK-9 | 54°39′27.36″N, 118°59′50.16″W |
| *Picea mariana* | 133055 | JK-10 | 54°39′27.36″N, 118°59′50.16″W |
| *Picea glauca* | 133056 | JK-11 | 54°44′24.06″N, 118°58′11.10″W |
| *Abies balsamea* | 133057 | JK-12 | 54°44′24.06″N, 118°58′11.10″W |
| *Pinus contorta* | 133058 | JK-13 | 54°44′24.06″N, 118°58′11.10″W |
| *Alnus crispa* | 133059 | JK-14 | 54°40′7.14″N, 118°59′12.66″W |
| *Pinus contorta* | 133060 | JK-15 | 54°40′7.14″N, 118°59′12.66″W |
| *Abies balsamea* | 133061 | JK-16 | 54°40′7.14″N, 118°59′12.66″W |
| *Picea mariana* | 133062 | JK-17 | 54°40′7.14″N, 118°59′12.66″W |
| *Betula papyrifera* | 133063 | JK-18 | 54°33′34.53″N, 118°39′48.50″W |

*Note*: JK = Justine Karst, collector.

[a] Voucher specimens were deposited at the University of Alberta Herbarium (ALTA), Edmonton, Alberta, Canada.