

Development and characterization of Novel EST-based single-copy genic microsatellite DNA markers in white spruce and black spruce

Om Rajora (✉ Om.Rajora@unb.ca)

University of New Brunswick Fredericton

Ishminder K. Mann

Current Pacific Ltd

Short Report

Keywords: cDNA microsatellite markers, Genic microsatellites, Single-copy microsatellites, *Picea glauca* and *Picea mariana*, Microsatellite marker informativeness, Microsatellite inheritance

Posted Date: February 23rd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-223756/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Molecular Biology Reports on February 26th, 2021. See the published version at <https://doi.org/10.1007/s11033-021-06231-1>.

Abstract

Background Due mainly to large genome size and prevalence of repetitive sequences in the nuclear genome of spruce (*Picea*), it is very difficult to develop single-copy genomic microsatellite markers. We have developed and characterized 25 polymorphic, single-copy genic microsatellites from white spruce (*Picea glauca*) EST sequences and determined their informativeness in white spruce and black spruce (*Picea mariana*) and inheritance in black spruce. Methods and Results White spruce EST sequences from NCBI dbEST were searched for the presence of microsatellite repeats. Forty-seven sequences containing dinucleotide, trinucleotide, tetranucleotide and compound repeats were selected to develop primers. Twenty-five of the designed primer pairs yielded scorable amplicons, with single-locus patterns, and were characterized in 20 individuals each of white spruce and black spruce. All 25 microsatellites were polymorphic in white spruce and 24 in black spruce. The number of alleles at a locus ranged from 2 to 18, with a mean of 8.8 in white spruce, and from 1 to 17, with a mean of 7.6 in black spruce. The expected heterozygosity/polymorphic information content ranged from 0.10 to 0.92, with a mean of 0.67 in white spruce, and from 0 to 0.93, with a mean of 0.59 in black spruce. Conclusions Microsatellites with dinucleotide and compound repeats were more informative than those with trinucleotide and tetranucleotide repeats. Eighteen microsatellite markers polymorphic between the parents of a black spruce controlled cross inherited in a single-locus Mendelian fashion. The microsatellite markers developed can be applied for various genetics, genomics, breeding, and conservation studies and applications.

Introduction

Microsatellites or simple sequence repeats (SSR) are co-dominant, highly polymorphic and reproducible molecular genetic markers, which have been widely used for various biological, breeding, forensics and conservation studies and applications. Microsatellite markers are especially suitable for determining neutral population genetic processes, and for DNA fingerprinting and forensics applications. Numerous methods based on microsatellite-enriched and non-enriched genomic libraries have been commonly used to isolate microsatellite-containing genomic sequences in plants and animals [1]. Expressed sequence tag (EST) sequences in the NCBI database and transcriptome sequences provide an important resource for identifying microsatellite sequences and developing cDNA-based genic microsatellite markers in plants [e.g., 2,3], including conifer trees [4,5,6,7,8]. EST-derived microsatellite markers have several advantages, such as being genic markers, high prospects of being single-locus types, conserved across related species, and their ability to detect variation directly in the expressed genes.

White spruce (*Picea glauca* (Moench) Voss) and black spruce (*Picea mariana* (Mill.) B.S.P.) are widely distributed transcontinental tree species of the Canadian boreal forest and have high ecological and economic importance [9]. There are several genetics, genomics, breeding and reforestation programs for these species. Highly informative microsatellite markers, especially genic microsatellites, could facilitate these programs. White spruce genome is large (~20 Gbp) and has repetitive sequences throughout the genome [10]. Black spruce probably has similarly large and repetitive genome as inferred from its nuclear

DNA contents [11]. Such genome features make the development of single-locus and highly informative genomic microsatellite markers in white spruce and black spruce very difficult as evident from the availability of handful genomic microsatellite markers in these species. Nuclear genomic microsatellite markers have been developed in white spruce using microsatellite-enriched and non-enriched sequences [12,13], and in black spruce using microsatellite-enriched genomic libraries [14] and microsatellite-enriched genomic AFLP DNA fragments [15]. Among these methods, the AFLP-microsatellite method [15] was found to yield a higher number of highly informative single-copy microsatellite DNA markers. In addition, 25 cDNA-derived microsatellites have been developed from white spruce ESTs and characterized in several spruce species [5]. The total number of informative markers from the above studies is around 60. Thus, there is a great need to develop additional single-copy microsatellite markers, especially gene-based microsatellite markers, in white spruce and black spruce.

In the present study, we have developed and characterized 25 new cDNA-based genic microsatellite markers in white spruce and black spruce. We have evaluated informativeness of these markers in white spruce and black spruce and established their single-locus Mendelian inheritance in black spruce.

Materials And Methods

Expressed Sequenced Tags (ESTs) with microsatellites, primers and their sequences

White spruce EST sequences were downloaded from the NCBI dbEST. The sequences were searched for the presence of microsatellite repeat motifs using Repeatmasker (<http://www.repeatmasker.org>). Attempts were made to select sequences whose gene function was annotated. Forty-seven sequences containing dinucleotide, trinucleotide, tetranucleotide and complex repeats were selected (Table 1; Table S1) for primer design. The number of sequences containing dinucleotide, trinucleotide, tetranucleotide, and complex repeats was 22, 14, 3, and 8, respectively (Table 1; Table S1). The TA repeat was the most abundant among the dinucleotide microsatellites. One sequence (RPGSE45) contained microsatellite repeats at two different places and two primer pairs were designed from this EST sequence. The primers for 48 microsatellite sequences were designed using Primer 3 program [16]. However, only one of the two primer pairs developed from the RPGSE45 sequence was used. The information on the source of the EST sequences, GenBank accessions, primer sequences and gene annotations are provided in Table 1 and Table S1. For each microsatellite marker, the M13 universal primer sequence labeled with either 700 or 800 IRDye™ (Li-Cor, Lincoln, Nebraska, USA), was added to the forward primer for detection of amplicons on Li-Cor 4200 automated sequencer (Li-Cor, Lincoln, Nebraska, USA).

PCR amplification, primer optimization and microsatellites visualization

PCR amplification, microsatellite primer testing, and primer optimization were carried out using four black spruce individuals, including two parents of a genetic mapping population. DNA was extracted from these samples using a Qiagen DNeasy Plant mini kit (Qiagen Canada, Mississauga, Ontario, Canada) and then prepared for downstream analysis.

Polymerase chain reactions (PCRs) were performed in a 10 µl mixture containing 10-20 ng of purified spruce DNA, 2 µg BSA, 1x *Taq* buffer, 0.2 mM of each dNTP, 0.025U of *Taq* DNA polymerase (MBI Fermentas, Waltham, Massachusetts, USA), 0.1 µM of reverse primer, 0.1 nM of M13 forward primer labeled with either 700 or 800 IRDye™ (Li-Cor, Lincoln, Nebraska, USA), and 1.5 mM MgCl₂. All PCRs were carried out using a PTC-200, Peltier Thermal Cycler (MJ Research, Waltham, Massachusetts, USA). Each PCR cycle had the following conditions: initial denaturation at 94°C for 5 min followed by 39 cycles each of 94°C for 30 sec, varying annealing temperatures ranging from 50°C to 65°C (determined using a 12-column gradient) for 30 sec, 72°C for 1 min, and a final extension at 72°C for 6 min.

Amplification products were denatured and electrophoresed through a 6.5% polyacrylamide gel run on a Li-Cor sequencing system IR 4200 (Li-Cor, Lincoln, NE, USA) with IRDye 700 and 800 labelled size standards (IR 50-350bp). Each gel was run for 2.5 h with the run parameters of 1400V, 40mA, 40W, 50°C plate temperature, and 16-pixel depths for collection of TIFF image files. The genotypes of individuals were scored manually, and the size of alleles was determined by using SAGA Generation 2 software (Li-Cor, Lincoln, Nebraska, USA). Alleles at a microsatellite locus were identified and named by their molecular size.

Informativeness, characterization and inheritance of microsatellite loci

We evaluated the informativeness of 25 microsatellite loci which could be amplified and resolved by genotyping 20 individuals each of white spruce and black spruce by calculating the number of alleles and their frequency, observed heterozygosity (H_O), and expected heterozygosity (H_E) using the PopGene

software [17]. We also calculated polymorphism information content (PIC) for each locus as $1 - \sum_{i=1}^n P_i^2$

where P_i is the frequency of the i^{th} allele in 20 individuals of a species examined. Thus, PIC and H_E calculations and estimates are the same. Thirteen white spruce samples originated from seven populations located at five sites in Saskatchewan: Scales Lake, Timber Cove, Snowfield Road, Prairie River, and Tee Pee Creek [details are provided in 18,19]. Seven white spruce samples originated from seven sub-populations (compartments) of the EMEND (Ecosystem Management Emulating Natural Disturbance) project located in northern Alberta approximately 90 km northwest of Peace River [20]. The black spruce samples originated from two populations at Pine Falls, four populations at Bissett, one population at The Pas and four population at Snow Lake sites in Manitoba, Canada [described in 21]. Inheritance of the polymorphic microsatellite markers was examined in a controlled cross from a three-generation outbred pedigree of black spruce [described in 22], consisting of the parents and their 90 progenies. χ^2 goodness-of-fit tests were performed to test the inheritance and segregation patterns of the microsatellite DNA markers.

Results

ESTs, microsatellite repeat motifs and microsatellite marker resolution

Twenty-two of the 47 selected EST sequences contained dinucleotide repeats, and 14 trinucleotide, three tetranucleotide and eight compound repeats. The most common repeat motif was TA, which accounted for 20 of the 22 dinucleotide-repeat and three of the eight compound-repeat microsatellites (Table 1; Table S1). From the 47 primer pairs, 30 amplified DNA and 25 yielded clear single-locus microsatellite variant patterns in black spruce and white spruce. These included nine dinucleotide, eight trinucleotide, three tetranucleotide and five compound-repeat microsatellites (Table 1). Seventeen primer pairs did not amplify any DNA amplicon and five produced complex patterns (Table S1). The source ESTs of 19 microsatellites were annotated to specific genes, whereas the source ESTs of three microsatellites were annotated to hypothetical proteins and that of three were not functionally annotated (Table 1). The optimized annealing temperature for different primer pairs varied from 50 °C to 63.9 °C (Table 1).

Microsatellite marker informativeness in white spruce and black spruce

In white spruce individuals, all 25 microsatellite loci were polymorphic; however, there was significant variation among the loci for polymorphism and informativeness (Table 2). The number of alleles ranged from two to 18, with a mean of 8.8, observed heterozygosity from 0 to 0.95 with a mean of 0.49, and PIC/expected heterozygosity ranged from 0.10 to 0.92, with a mean of 0.67. The most informative loci were *RPGSE48* and *RPGSE43* (the loci with a simple and compound TA repeats), and the least informative locus was *RPGSE17* with a trinucleotide repeat (Table 2). Twenty-four of the 25 microsatellite loci were polymorphic in black spruce individuals. The trinucleotide repeat locus *RPGSE08* was monomorphic. There was significant variation among the loci for polymorphism and informativeness in black spruce (Table 2). The number of alleles ranged from one to 17, with a mean of 7.6, observed heterozygosity ranged from 0 to 0.90, with a mean of 0.43, and PIC/ H_E ranged from 0.10 to 0.93, with a mean of 0.59 (Table 2). The most informative locus was *RPGSE34*, a locus with a simple TA repeat, and the least informative locus besides the monomorphic locus *RPGSE08* was *RPGSE17* with a trinucleotide repeat (Table 2).

In both white spruce and black spruce, significant differences ($P < 0.05$) in the number of alleles and H_E /PIC were observed among microsatellite loci classes with different repeat types (di, tri, tetra and compound). Microsatellite loci with dinucleotide and compound repeats had significantly higher informativeness for all three parameters (A , H_0 and H_E) than the microsatellite loci with trinucleotide and tetranucleotide repeats (Table 2). The trinucleotide microsatellite loci were slightly but not significantly more informative than the tetranucleotide microsatellite loci for one or two of the three parameters (Table 2). Each white spruce or black spruce individuals could be uniquely fingerprinted by their two-locus microsatellite genotypes. For the combined sample of white spruce and black spruce individuals, although the informativeness for microsatellite loci was higher than for individual species samples, the informativeness patterns were similar to those observed from individual species samples (Table 2).

Inheritance of microsatellite markers in black spruce

Eighteen microsatellite loci were polymorphic between the parents of the black spruce controlled cross (Table 3). At each of the 18 loci, the progenies of the cross segregated into two to four microsatellite

genotypic classes as expected for single-locus Mendelian inheritance patterns (Table 3; Figure S1 for *RPGSE35*).

Discussion

We have developed highly informative single-copy cDNA-based genic microsatellite DNA markers in white spruce and black spruce. From the screening of partial white spruce EST database, it was apparent that the dinucleotide TA repeat motif is most common in the white spruce EST sequences set that we searched. This result is consistent with the similar observation in loblolly pine (*Pinus taeda*) ESTs [6]. However, AT repeat motif was found to be most abundant in the Genome BC Forestry spruce EST database [5] and AG repeat motif in Norway spruce cDNA clones [4].

We establish that 24 of the 25 genic microsatellite markers we characterized in this study are informative in white spruce and black spruce. Of these, 11 (*RPGSE24*, *RPGSE24*, *RPGSE29*, *RPGSE34*, *RPGSE35*, *RPGSE40*, *RPGSE41*, *RPGSE243*, *RPGSE45*, *RPGSE46*, and *RPGSE48*) were highly informative loci in white spruce, each with over 10 alleles and H_E /PIC close to 0.90. In black spruce, 9 loci (*RPGSE02*, *RPGSE22*, *RPGSE25*, *RPGSE29*, *RPGSE34*, *RPGSE35*, *RPGSE43*, *RPGSE46*, and *RPGSE48*) with 10 or more alleles at a locus were highly informative.

Our results reveal that microsatellite loci with dinucleotide perfect and dinucleotide compound repeats are more informative than microsatellite loci with trinucleotide or tetranucleotide repeats. These results are consistent with the general view and actual observations [e.g., 23] that microsatellites with trinucleotide and tetranucleotide repeats have lower polymorphism than dinucleotide repeats. In contrast, previously, we found trinucleotide repeat microsatellite loci to be highly polymorphic in *Populus tremuloides* [24,25]. Also, there is a general view and experimental evidence that compound-repeat microsatellite loci display lower variation than simple-repeat microsatellite loci [e.g., 24,25]. However, our study demonstrates that microsatellite loci with compound repeats (*RPGSE02*, *RPGSE24*, *RPGSE43*, *RPGSE45*, and *RPGSE46*) are highly polymorphic in both white spruce and black spruce (Table 2).

Gene sequences potentially have lower mutation rates, slow evolution and selection constraints, thus higher conservation among individuals within a species and across the related species. Due to these features, EST-derived microsatellite markers are expected to show lower polymorphism than genomic microsatellites. This was somewhat true for the EST-based markers developed in this study with regard to the number of alleles, which were lower than that observed from genomic microsatellites in white spruce [12,13; Table S2] and black spruce [15; Table S2]. However, allelic diversity in our genic microsatellites was much higher than that reported for genomic microsatellites by Dobrezeniecka et al. [14] for black spruce (Table S2). Informativeness of the white spruce EST-based genic markers in our study is similar to that previously reported for white spruce EST-based markers in these two species [5; Table S2]. The average number of alleles per locus and observed and expected heterozygosity for eight trinucleotide-repeat microsatellite markers in our study is higher than that reported for EST-derived trinucleotide microsatellite markers in Norway spruce (*Picea abies*) [7; Table S2].

Up to 18 of the 20 (90%) black spruce and 17 of the 20 (85%) white spruce individuals could be uniquely fingerprinted from their single-locus microsatellite genotypes. All individuals of white spruce or black spruce could be uniquely fingerprinted by their genotypes at two microsatellite loci. These results clearly show that the SSR markers reported in this manuscript could provide an excellent molecular tool for fingerprinting and identification of white spruce and black spruce individuals and clones.

Our study demonstrates that the microsatellite variants at each of the 18 loci, polymorphic between the parents of the controlled cross, are controlled by a single nuclear locus because they conformed to single-locus inheritance patterns (Table 3). Significant deviation from the expected 1:1 segregation ratio for *RPGSE05* could be by chance or some biological mechanisms. Segregation distortion has been commonly observed in backcrosses and F_2 progeny of plants [26]. The black spruce cross that we used was from a three-generation outbred pedigree, equivalent to a F_2 cross.

Several features of the novel genic microsatellite markers that we have developed, such as single-copy, high informativeness and single-locus co-dominant inheritance patterns, makes them highly suitable for various applications in genetics, genomics, DNA fingerprinting and tree forensic, breeding, and genetic resource conservation programs in white spruce and black spruce. Five of these markers have already been successfully used in white spruce population genetics studies [19,20]. The microsatellite markers reported here could potentially be used in other spruce species.

Conclusions

We have developed 25 new reliable and highly informative single-copy genic microsatellites in white spruce and black spruce from white spruce EST sequences. Microsatellites with dinucleotide and compound repeats were more informative than microsatellites with trinucleotide and tetranucleotide repeats. The markers inherited co-dominantly in a single-locus Mendelian pattern. The new set of genic microsatellite markers provides an excellent genetic/genomic resource that could be used for diverse biological, genetics, DNA fingerprinting, genomics, tree breeding, and genetic conservation studies and applications in white spruce, black spruce and possibly other spruce species.

Declarations

Acknowledgements

We thank Nicholas Buckley and Shayne Sangster for laboratory assistance. The research was funded by a Natural Sciences and Engineering Research Council of Canada (NSERC) Strategic Project grant (STPGP 234783 – 00) to O.P. Rajora. I.K. Mann was financially supported by NSERC Strategic Grants funds (STPGP 234783 – 00) to O.P. Rajora and Dalhousie University Graduate Scholarship. O.P. Rajora held the Stora Enso Senior Chair in Forest Genetics and Biotechnology at Dalhousie University, which was supported by Stora Enso Port Hawkesbury Ltd., and the Senior (Tier 1) Canada Research Chair in Forest and Conservation Genomics and Biotechnology at UNB, which was supported by the Canada Research Chair Program (CRC950-201869).

Funding

The research was funded by a Natural Sciences and Engineering Research Council of Canada (NSERC) Strategic Project grant (STPGP 234783 – 00) to O.P. Rajora. I.K. Mann was financially supported by NSERC Strategic Grants funds (STPGP 234783 – 00) to O.P. Rajora and Dalhousie University Graduate Scholarship.

Authors Contributions

OPR conceived the idea, provided overall study directions, procured funding and prepared, finalized and revised the manuscript. IKM identified the EST sequences with microsatellites, designed the primers, conducted the lab work and analyzed the data under the direction of OPR. Both authors read and approved the final manuscript.

Compliance with Ethical Standards

Conflict of interest or competing interest

The authors declare that they have no conflict of interest or competing interests.

Ethics Approval

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

Consent to Participate

Not applicable

Consent to Publish

The authors consent to publish this manuscript.

References

1. Zane L, Bargelloni L, Patarnello T (2002) Strategies for microsatellite isolation: a review. *Mol Ecol* 11:1-16.
2. Scott KD, Egger P, Seaton G, Rossetto M, Ablett EM, et al. (2000) Analysis of SSRs derived from grape ESTs. *Theor Appl Genet* 100:723–726.
3. Thiel T, Michalek W, Varshney RK, Graner A (2003) Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theor Appl Genet* 106: 411-422.
4. Scotti I, Magni F, Fink R, Powell W, Binelli G, Hedley PE (2000) Microsatellite repeats are not randomly distributed within Norway spruce (*Picea abies* K.) expressed sequences. *Genome* 43: 41-46.

5. Rungis D, Berube Y, Zhang J, Ralph S, Ritland CE, Ellis BE, Douglas C, Bohlman J, Ritland K (2004) Robust simple sequence repeat markers for spruce (*Picea* spp.) from expressed sequence tags. *Theor Appl Genet* 109: 1283-1294.
6. Liewlaksaneeyanawin C, Ritland CE, El-Kassaby YA, Ritland K (2004) Single-copy, species-transferable microsatellite markers developed from loblolly pine ESTs. *Theor Appl Genet* 109: 361-369.
7. Fluch S, Burg A, Kopecky D, Homolka A, Spiess N, Vendramin GG (2011) Characterization of variable EST SSR markers for Norway spruce (*Picea abies* L.). *BMC Research Notes* 4:401.
8. Duan D, Jia Y, Yang J, Li Z-H (2017) Comparative transcriptome analysis of male and female conelets and development of microsatellite markers in *Pinus bungeana*, an endemic conifer in China. *Genes* 8, 393; doi:10.3390/genes8120393.
9. Hosie RC (1979) Native trees of Canada. Fitzhenry and Whiteside Ltd., Don Mills, Ontario.
10. Birol I, Raymond A, Jackman SD, Pleasance S, Coope R, et al (2013) Assembling the 20 Gb white spruce (*Picea glauca*) genome from whole-genome shotgun sequencing data. *Bioinformatics* 29:1492-1497.
11. Murray, BG (1998) Nuclear DNA amounts in gymnosperms. *Ann Bot* 82: 3–15.
12. Rajora OP, Rahman MH, Dayanandan S, Mosseler A (2001) Isolation, characterization, inheritance and linkage of microsatellite DNA markers in white spruce (*Picea glauca*) and their usefulness in other spruce species. *Mol Gen Genet* 264: 871-882.
13. Hodgetts RB, Aleksiuik MA, Brown A, Clarke C, Macdonald E, Nadeem s, Khasa D (2001) Development of microsatellite markers for white spruce (*Picea glauca*) and related species. *Theor Appl Genet* 102: 1252-1258.
14. Dobrzezeniecka S, Nkongolo KK, Michael P, Wyss S, Mehes M (2009) Identification and characterization of microsatellite markers for genetic analysis of black spruce (*Picea mariana* (Mill.) populations. *Silvae Genetica* 58:168-172.
15. Shi Y-Z, Forneris N, Rajora OP (2014) Highly informative single-copy nuclear microsatellite DNA markers developed using an AFLP-SSR approach in black spruce (*Picea mariana*) and red spruce (*P. rubens*). *PLoS One* 9(8): e103789. DOI: 10.1371/journal.pone.0103789.
16. Rozen S, Skaletsky H (2000) Primer 3 on the WWW for general users and for biologist programmers. In: Clifton NJ (ed). *Methods in Molecular Biology*. Humana Press, Totowa, New Jersey, pp 365–386.
17. Yeh FC, Boyle TJB (1997) Population genetic analysis of codominant and dominant markers and quantitative traits. *Belg J Bot* 129: 157.
18. Rajora OP (1999) Genetic biodiversity impacts of silvicultural practices and phenotypic selection in white spruce. *Theor Appl Genet* 99: 954-961.
19. Fageria MS, Rajora OP (2014) Effects of silvicultural practices on genetic diversity and population structure of white spruce in Saskatchewan. *Tree Genetics and Genomes* 10:287-296.
20. Fageria MS, Rajora OP (2013) Effects of harvesting of increasing intensities on genetic diversity and population structure of white spruce. *Evolutionary Applications* 6:778-794; doi: 10.1111/eva.12064.

21. Rajora OP, Pluhar SA (2003) Genetic diversity impacts of forest fires, forest harvesting, and alternative reforestation practices in black spruce (*Picea mariana*). *Theor Appl Genet* 106:1203–1212.
22. Kang B-Y, Mann IK, Major JE, Rajora OP (2010) Near-saturated and complete genetic linkage map of black spruce (*Picea mariana*). *BMC Genomics* 11: 515.
23. Gadaleta A, Mangini G, Mule G, Blanco A (2007) Characterization of dinucleotide and trinucleotide EST-derived microsatellites in the wheat genome. *Euphytica* 153: 73-85.
24. Dayanandan S, Rajora OP, Bawa KS (1998) Isolation and characterization of microsatellites in trembling aspen (*Populus tremuloides*). *Theor Appl Genet* 96: 950-956.
25. Rahman MH, Dayanandan S, Rajora OP (2000) Microsatellite DNA markers in *Populus tremuloides*. *Genome* 43: 293-297.
26. Zamir D, Tadmor Y (1986) Unequal segregation of nuclear genes in plants. *Bot Gaz* 147: 355-358.

Tables

Table 1. Microsatellite loci, source sequences and their GenBank accession numbers, gene annotations, primers, repeat motifs, and optimized annealing temperatures

Locus	Source	GenBank Accession	Forward (F) and reverse (R) primers	Repeat motif	Annealing temperature	Gene targeted
<i>RPGSE02</i>	<i>Picea glauca</i> cDNA clone WS0079_H08 3', mRNA sequence	Ace: CN480904 gi: 46562408	F:M13*- GAGAATGCAGGAGGCAAGAC R:CTTACTTACCGTGCCAATGC	(CTG) ₃ G(CTG) ₃ G(CTG) ₃	61.7	Homeotic protein BEL1
<i>RPGSE03</i>	<i>Picea glauca</i> cDNA clone WS0032_M17 3', mRNA sequence	Ace: CN480906 gi: 46562410	F:M13*- AGCTAACTGGACTGGGACCTT R:CCGCACATGATATCCACAAG	(TTA) ₆	50	Putative Beta glycosidase
<i>RPGSE04</i>	Male strobili developmental sequence <i>Picea glauca</i> cDNA clone GQ0012a_D05 5', mRNA sequence	Ace: CK442213 gi: 40777567	F:M13*- CTTGATTTTTGGCGATCGTT R:GAACCGGAGGAGATGGACTA	(CCG) ₆	58.9	Gibberellin 12-oxidase
<i>RPGSE05</i>	Male strobili developmental sequence <i>Picea glauca</i> cDNA clone GQ0013d_C02 5', mRNA sequence	Ace: CK442392 gi: 40777946	F:M13*- CCGATTCAGGCAAGAGAATC R:TCACTGGCCACAGTTTATCG	(GAA) ₆	52.5	-
<i>RPGSE08</i>	Male strobili developmental sequence <i>Picea glauca</i> cDNA clone GQ0014b_C06 5', mRNA sequence	Ace: CK443173 gi: 40779582	F:M13*- TCTCAAGAGAGGACGGAGGA R:GCATTCTGAGAGCCTTGCTT	(GAA) ₆	52.5	Hypothetical protein (<i>A. thaliana</i>)
<i>RPGSE10</i>	Male strobili developmental sequence <i>Picea glauca</i> cDNA clone GQ0011d_C01 3', mRNA sequence	Ace: CK441912 gi: 40776941	F:M13*- ATACGTTGGCGTTTCCGTCT R:TGAGGGCTTATGGACTACGC	(GGA) ₆	63.9	Cytochrome P450
<i>RPGSE11</i>	Male strobili developmental sequence <i>Picea glauca</i> cDNA clone GQ0011a_E05 5', mRNA sequence	Ace: CK442036 gi: 40777203	F:M13*- AGGAGCCTTCGAGTGAAGTG R:GCTCGTTAGCTCAACCCAAA	(TCC) ₇	52.5	Early light inducible protein

<i>RPGSE13</i>	Female <i>cones developmental sequence</i> <i>Picea glauca</i> cDNA clone GQ0025c_G06 5', mRNA sequence	Ace: CK440825	F:M13*- CTGTCAGTGACCCTGCAAGA	(TCC) ₇	52.5	Auxin-induced protein 1
		gi: 40774649	R:TGCAGCACTCGTTCTCACTT			
<i>RPGSE17</i>	Non-lignified secondary xylem from mature trees <i>Picea glauca</i> cDNA clone GQ0043_O08 3', mRNA sequence	Ace: CK436417	F:M13*- TCACTGTGGGCAACTTACCA	(TCG) ₆	63.9	Expressed protein
		gi: 40765246	R:TCCCTCAAAAACACGGAGAC			
<i>RPGSE22</i>	Secondary xylem of girdled saplings <i>Picea glauca</i> cDNA clone GQ0072_P09 3', mRNA sequence	Ace: CK444999	F:M13*- CCAGCTCAGGTGAATCCTTC	(TA) ₉₊₃	50	-
		gi: 40781419	R:TGGAGATGGTCCGACTTAGG			
<i>RPGSE24</i>	<i>Picea glauca</i> cDNA clone WS0016_O09 3', mRNA sequence	Ace: CN480894	F:M13*- TGCTCTCTTATTCGGGCTTC	(TA) ₃ TC(TA) ₄ GA(TA) ₈	50.4	Remorin family protein
		gi: 46562398	R:AAGAACAAGGCTTCCCAATG			
Locus	Source	GenBank Accession	Forward (F) and reverse (R) primers	Repeat motif	Annealing temperature	Gene targeted
<i>RPGSE25</i>	<i>Picea glauca</i> cDNA clone WS0019_M09 3', mRNA sequence	Ace: CN480895	F:M13*- CATGGCTGCAGCTGTAGTGT	(TA) ₁₉	58.9	-
		gi: 46562399	R:TTGGAGTGCTCCCTTGAAGT			
<i>RPGSE29</i>	<i>Picea glauca</i> cDNA clone WS0022_B15 3', mRNA sequence	Ace: CN480899	F:M13*- TGGCTTTTATTCCAGCAAG	(GA) ₁₁	54.2	Unknown protein (<i>A. thaliana</i>)
		gi: 46562403	R:GCCAGATTTTGCAAAGTGGA			
<i>RPGSE34</i>	<i>Picea glauca</i> cDNA clone WS00716_F13 3', mRNA sequence	Ace: CN480905	F:M13*- CCAATTTGGTCCAATCTAGCA	(GA) ₁₀	50	Chloroplast nucleod DNA binding protein
		gi: 46562409	R:GGATGTGTTTTGGAGGGTTG			

<i>RPGSE35</i>	<i>Picea glauca</i> cDNA clone WS0033_A18 3', mRNA sequence	Ace: CN480907 gi: 46562411	F:M13*- TGGCTCTCATCCAGAAAAGAA R:GGCTGCTCTCTTATCCGTTTT	(TA) ₂₆	54.2	Hypothetical protein
<i>RPGSE40</i>	Male strobili developmental sequence <i>Picea glauca</i> cDNA clone GQ0011c_B11 5', mRNA sequence	Ace: CK442712 gi: 40778611	F:M13*- TTTCCAAGTGCCTACCACCT R:TTCATTTTGAGCTGCATGGT	(TA) ₂₁	50	Cytochrome B561
<i>RPGSE41</i>	Male strobili developmental sequence <i>Picea glauca</i> cDNA clone GQ0011c_B11 3', mRNA sequence	Ace: CK442714 gi: 40778615	F:M13*- TGAAGGAACACCGTAGCGTA R:GCGGGTGAGAGGAATATCAT	(TA) ₂₁	63.9	Cytochrome B561
<i>RPGSE43</i>	Male strobili developmental sequence <i>Picea glauca</i> cDNA clone GQ0013b_D09 3', mRNA sequence	Ace: CK443040 gi: 40779303	F:M13*- TGCATTACAAGCACCTCAGA R:TGCTGTGCCTGCCTATACAT	(TA) ₉ TC(TG) ₅	58.9	LEA protein
<i>RPGSE44</i>	Female cones developmental sequence <i>Picea glauca</i> cDNA clone GQ0024d_H04 3', mRNA sequence	Ace: CK440928 gi: 40774857	F:M13*- GAATGAAGGCAGGGAAACAG R:CCTGGGATTTCAAGTTGAGG	(TA) ₉	61.7	NDR1/ HIN1 like protein
<i>RPGSE45</i>	Vegetative buds developmental sequence <i>Picea glauca</i> cDNA clone GQ0032a_C01 3', mRNA sequence	Ace: CK439419 gi: 40771748	F:M13*- GCAGCACTGGCAACATTCTA R:ACACGGAGACAAATCGCTTG	(TA) ₈ T(TA) ₅	51.3	Unknown protein (<i>O. sativa</i>)
<i>RPGSE46</i>	Vegetative buds developmental sequence <i>Picea glauca</i> cDNA clone GQ0033c_B03	Ace: CK438739 gi:	F:M13*- GGAATAAAATGGCAGGTGGA R:GCCTGCAGTAGTTGGCAGA	(GA) ₉ AA(GA) ₇	50	Putative UNC-50

	3', mRNA sequence	40770308				
<i>RPGSE48</i>	Cambium and phloem from mature trees <i>Picea glauca</i> cDNA clone GQ0064_B13 3', mRNA sequence	Ace: CK435272	F:M13*- TTTCTCTGGGGATGATCCAG	(TA) ₁₄	50.4	Hypothetical protein
		gi: 40762826	R:GCACGAGAGGGCATCTAACT			
<i>RPGSE49</i>	<i>Picea glauca</i> cDNA clone WS0011_P12 3', mRNA sequence	Ace: CN480892	F:M13*- CCTGATTCTGTCTTCCTTACCG	(GGAA) ₅	63.9	EF-hand Ca ²⁺ binding protein
		gi: 46562396	R:AAAGGAAGGGCTTAGCTTGC			
<i>RPGSE50</i>	Secondary xylem of girdled saplings <i>Picea glauca</i> cDNA clone GQ0071_A03 3', mRNA sequence	Ace: CK444187	F:M13*- GAAAAGCGGTCCTGGAAAGT	(GCTG) ₅	52.5	Probable aquaporin
		gi: 40780607	R:CGGCGTATCACCAATACGTT			
<i>RPGSE52</i>	Secondary xylem of girdled saplings <i>Picea glauca</i> cDNA clone GQ0072_G15 3', mRNA sequence	Ace: CK444729	F:M13*- TGGCACGATAGTGAAGACCT	(GCTG) ₅	52.5	Probable aquaporin
		gi: 40781149	R:AGCCCACCAGTTTGGCAATA			

Table 2. Informativeness of microsatellite loci in white spruce, black spruce and the total white spruce and black spruce sample. A, number of alleles; H_O, observed heterozygosity; H_E, expected heterozygosity; PIC, polymorphic information content.

Locus	White Spruce				Black Spruce				Total Sample			
	A	Allele Size Range	H _O	H _E /PIC	A	Allele Size Range	H _O	H _E /PIC	A	Allele Size Range	H _O	H _E /PIC
PGSE02	7	185-211	0.70	0.59	11	185-209	0.70	0.85	12	185-211	0.70	0.82
PGSE03	8	231-267	0.56	0.73	5	228-246	0.78	0.70	9	228-267	0.67	0.75
PGSE04	9	220-256	0.60	0.83	7	217-247	0.55	0.52	11	217-256	0.58	0.74
PGSE05	2	259-262	0.30	0.32	2	259-262	0.30	0.38	2	259-262	0.30	0.35
PGSE08	2	220-223	0.00	0.10	1	223	0.00	0	2	220-223	0.00	0.05
PGSE10	5	187-208	0.30	0.45	5	192-208	0.45	0.54	6	187-208	0.38	0.51
PGSE11	6	460-484	0.60	0.74	5	466-484	0.37	0.51	8	460-484	0.49	0.72
PGSE13	3	226-232	0.60	0.57	4	220-232	0.45	0.69	4	220-232	0.53	0.65
PGSE17	2	173-176	0.20	0.18	2	176-182	0.05	0.05	3	173-182	0.13	0.12
PGSE22	7	208-230	0.55	0.55	12	208-234	0.60	0.85	12	208-234	0.58	0.74
PGSE24	13	222-226	0.74	0.89	7	222-260	0.58	0.70	14	222-260	0.66	0.87
PGSE25	15	204-262	0.53	0.90	13	206-270	0.15	0.89	21	204-270	0.33	0.93
PGSE29	13	242-288	0.75	0.90	10	242-288	0.60	0.81	17	242-288	0.68	0.91
PGSE34	12	258-284	0.95	0.85	17	256-292	0.90	0.93	18	256-292	0.92	0.91
PGSE35	12	151-211	0.42	0.83	10	165-189	0.63	0.80	16	151-211	0.53	0.87
PGSE40	14	223-257	0.21	0.90	9	223-251	0.35	0.75	16	223-257	0.28	0.87
PGSE41	15	228-266	0.21	0.91	9	232-258	0.55	0.75	18	228-266	0.38	0.88
PGSE43	18	170-230	0.55	0.92	17	170-232	0.50	0.90	24	170-232	0.53	0.93
PGSE44	3	220-224	0.30	0.27	2	222-226	0.05	0.05	4	220-226	0.18	0.16
PGSE45	11	246-302	0.74	0.87	7	268-310	0.47	0.65	14	246-310	0.61	0.84
PGSE46	13	193-221	0.85	0.90	13	193-221	0.80	0.88	14	193-221	0.83	0.90
PGSE48	18	236-260	0.58	0.91	11	226-260	0.37	0.87	23	226-260	0.47	0.93
PGSE49	3	255-263	0.15	0.58	3	259-289	0.05	0.14	4	255-289	0.10	0.47
PGSE50	5	175-211	0.30	0.51	5	175-199	0.20	0.38	6	175-211	0.25	0.65
PGSE52	4	169-182	0.40	0.55	3	162-177	0.20	0.26	5	162-182	0.30	0.57
Means												
Overall	8.8		0.48	0.67	7.6		0.43	0.59	11.3		0.46	0.69
Dinucleotide	12.1		0.50	0.78	10.3		0.47	0.74	16.1		0.48	0.80
Trinucleotide	4.6		0.40	0.49	3.9		0.39	0.42	5.6		0.39	0.49
Tetranucleotide	4.0		0.28	0.55	3.7		0.15	0.20	5.0		0.22	0.56
Compound	12.4		0.72	0.76	11.0		0.61	0.80	15.6		0.67	0.87

Table 3. Chi square goodness-of-fit tests for segregation of microsatellite loci in the progeny obtained from three-generation outbred pedigree in Black spruce.

Locus	Parental genotypes	No. of Progeny	Expected Ratio	Observed distribution of progeny genotypes	The Chi square value
<i>RPGSE02</i>	204/193 (AB) X 193 /193 (AA)	89	1:1	AA(45):AB(44)	0.01 (n.s.)
<i>RPGSE03</i>	240/237 (AB) X 237/231 (BC)	90	1:1:1:1	AB(22):AC(14):BB(30):BC(24)	5.82 (n.s.)
<i>RPGSE04</i>	231/223 (BC) X 232/220 (AD)	90	1:1:1:1	AB(23):BD(14):AC(31):CD(22)	6.44 (n.s.)
<i>RPGSE05</i>	262/262 (AA) X 262/259 (AB)	90	1:1	AA (34): AB (56)	5.38 (s.*)
<i>RPGSE08</i>	Monomorphic				
<i>RPGSE10</i>	208/0 (A0) X 0/0 (0/0)	90	1:1	A0(48) : 00 (42)	0.40 (n.s.)
<i>RPGSE11</i>	478/478 (AA) X 478/475 (AB)	90	1:1	AA (52): AB (38)	2.18 (n.s.)
<i>RPGSE13</i>	231/231 (BB) X 226/231 (AB)	90	1:1	BB (57): AB(43)	0.18 (n.s.)
<i>RPGSE17</i>	Monomorphic				
<i>RPGSE22</i>	210/208 (AB) X 210/0 (A0)	90	1:2:1	AB (24): AA+A0 (39): B0 (27)	1.80 (n.s.)
<i>RPGSE24</i>	228/224 (AB) X 224/224 (BB)	90	1:1	AB (45): BB(45)	0.90 (n.s.)
<i>RPGSE25</i>	210/0 (A0) X 206/0 (B0)	90	1:1:1:1	A0(27): B0(24): AB(15): 00(24)	3.60 (n.s.)
<i>RPGSE29</i>	278/246 (AC) X 266/0 (B0)	90	1:1:1:1	AB(18): A0(22): BC(21): C0(29)	2.89 (n.s.)
<i>RPGSE34</i>	276/270 (AB) X 268/266 (CD)	90	1:1:1:1	AC(22): AD(27):BC(20): BD(21)	1.29 (n.s.)
<i>RPGSE35</i>	177/171 (BC) X 185/171 (AC)	90	1:1:1:1	AB(20): AC(26): BC(19): CC(25)	1.64 (n.s.)
<i>RPGSE40</i>	229/0 (A0) X 223/0 (B0)	90	1:1:1:1	AB(26): A0(23): B0(20): 00(21)	0.93 (n.s.)
<i>RPGSE41</i>	236/0 (A0) X 0/0 (00)	90	1:1	A0 (48): 00 (42)	0.40 (n.s.)
<i>RPGSE43</i>	200/192 (AB) X 190/0 (C0)	90	1:1:1:1	A0(25):B0(17):AC(28):BC(20)	3.24 (n.s.)
<i>RPGSE44</i>	Monomorphic				
<i>RPGSE45</i>	302/300 (AB) X 302/302 (AA)	90	1:1	AB(38): AA(52)	2.18 (n.s.)
<i>RPGSE46</i>	213/195 (AD) X 209/197 (BC)	90	1:1:1:1	AB(20): AC(18): CD(26): BD(26)	2.27 (n.s.)
<i>RPGSE48</i>	Monomorphic				
<i>RPGSE49</i>	Monomorphic				
<i>RPGSE50</i>	Monomorphic				
<i>RPGSE52</i>	Monomorphic				

n.s. - non significant

s.* - significant at 5% level of significance

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [RajoraMannSpruceESTSSRSupplementaryMaterial.docx](#)