Nuclear microsatellite markers for population genetic studies in sugar maple (*Acer saccharum* Marsh.)

S. Khodwekar, M. Staton, M.V. Coggeshall, J.E. Carlson, O. Gailing

Abstract. A set of seven new nuclear microsatellite markers (nSSRs) was developed for sugar maple (*Acer saccharum* Marsh.) using paired-end Illumina sequencing. Out of 96 primers screened in a panel of six unrelated individuals, seven markers amplified polymorphic products. The utility of these markers, in addition to six already published microsatellites, for genetic variation and gene flow studies was assessed. Out of the seven newly developed markers, three amplified multiple fragments and were interpreted as dominant (absence/presence) markers, while four markers amplified a maximum of two amplification products per sample. The six published microsatellites and three of the four newly developed markers showed regular segregation in an open-pollinated single tree progeny. Observed heterozygosity (H_o) and expected heterozygosity (H_e) in 48 individuals from one population ranged from 0.436 to 0.917 and from 0.726 to 0.894, respectively. Dominant markers revealed 64 variable positions and moderate genetic variation within the population (H_e = 0.102, Shannon's I = 0.193). Paternity analyses in the program CERVUS at co-dominant markers showed effective dispersal of pollen in the sugar maple population both at 95% and 80% confidence levels. Dependent on the confidence level, the mean pollen dispersal distance within the population ranged from 33.25 m to 38.75 m and gene flow from outside the stand from 78% to 82%. The absence of fine-scale Spatial Genetic Structure (SGS) suggested effective dispersal of both seeds and pollen.
Keywords *Acer saccharum*, nuclear microsatellites, gene flow, spatial genetic structure, inheritance analysis.

Authors. Sudhir Khodwekar, Oliver Gailing (ogailing@mtu.edu) - School of Forest Resources and Environmental Science, Michigan Technological University, 1400 Townsend Drive, Houghton, MI 49931; Margaret Staton - Department of Entomology and Plant Pathology, The University of Tennessee, Knoxville, TN 37996; Mark V. Coggeshall - Department of Forestry, University of Missouri, Columbia, MO 65274; John E. Carlson - Department of Ecosystem Science and Management and Department of Plant Science, The Pennsylvania State University, University Park, PA 16802.

Introduction

Sugar maple (*Acer saccharum* Marsh.) is an important North American hardwood forest tree that is widespread and abundant in Northeastern North America (Godman et al. 1990). Despite its ecological and economic importance, genetic and genomic resources for this species are rather limited (Graignic et al. 2013). Here, we describe the development and application of Next-Generation Sequencing derived microsatellite markers for population genetic analyses in sugar maple.

The level and distribution of genetic variation is essential for the long-term survival and ability of organisms to adapt to changes in their environment (Krustovsky & Neale 2005). Recent research has predicted that climate change has the potential to affect species distributions, alter ecosystem functions and species interactions (Iverson et al. 2008, Whitham et al. 2008). Other important predicted consequences of global temperature increase are range shifts and northward migration of forest trees (Gunter et al. 2000, Graignic et al. 2014). Range edge populations are especially prone to geographical isolation and environmental changes. For example, sugar maple populations at the species' southern distribution edge are threatened by fragmentation as the result of climate change, and the preservation of genetic variation in these populations is a major conservation concern (Iverson et al. 2008, Graignic et al. 2013). Also, silvicultural treatments including the regulation of stand structure through harvesting operations will impact natural regeneration patterns and is expected to have considerable effects on subsequent genetic variation patterns by affecting the reproduction system of trees (Finkeldey & Ziehe 2004). These changes can be reflected in changes of fine scale genetic structure of the progeny generation within populations.

Genetic variation is ubiquitous in natural populations and is important for sugar maple to adapt to changing environments and cope with biotic and abiotic stresses such as increased drought, disruption of nutrient cycles, insect outbreaks and fungal diseases (Gibson & Wheelwright 1995, Whitham et al. 2008, Whitham et al. 2012). For example, sapstreak disease caused by the fungus *Ceratocystis virescens* (Davidson) Moreau leads to sugar maple dieback in Upper Michigan (Bal et al. 2013) which will likely be reflected in the level and distribution of genetic variation within species. The possibility to observe variation patterns at particular gene loci is of great significance in this context.

The development of highly polymorphic microsatellites can greatly enhance ecological genetic studies in sugar maple (see also Pandey et al. 2004, 2012), and genetic variation can be monitored in response to climate change and management. Despite its ecological and economic importance as a dominant overstory species in Northern hardwoods, there are few studies of genetic variation in sugar maple using highly variable genetic markers such as microsatellites (Graignic et al. 2013). The usefulness of microsatellites is due to the variability in the number of sequence repeats at a given gene locus within species and populations. Nuclear microsatellites are co-dominant, easily scored and abundant throughout the genome (Selkoe & Toonen 2006). Because of their high polymorphism they are suitable for investigations of the gene flow and mating system (Selkoe & Toonen 2006).

The main objective of the present study is the development and application of microsatellite markers to assess family structures and gene flow within a sugar maple natural population. These microsatellite markers can also be applied to monitor genetic variation, especially in peripheral populations, and used as predictive tools to characterize the long-term adaptability of these populations to predicted impacts of climate change.

Materials and methods

Species description

Sugar maple is a dominant overstory species in northern hardwoods and abundant in the Northeastern part of North America (Godman et al. 1990). It is a major constituent of several forest cover types and is classified as shade tolerant hardwood tree species and considered a ''climax species'' throughout its extensive native range. It is an important commercial source of income due to its syrup and timber quality. Sugar maple is characterized by hard and strong wood, and was used as sweetener by Native Americans and early European settlers (Nesom 2006). The flowering period of sugar maple extends from late March to mid of May depending upon the geographical location and seasonal variation. Sugar maple flowers can be either perfect or staminate (pistil abortive) and are generally wind pollinated, but also pollination by bees was reported (Gabriel & Garrett 1984). The fruit is a double samara and is dispersed by wind with the potential of long distance dispersal (Guries & Nordheim 1984, Godman et al. 1990). It ripens within about 16 weeks; the large wings typically range from 20 to 27 mm in length and 7 to 11 mm in width. These samara wings allow fruits to be carried away by wind at least 100 m (Godman et al. 1990).

Marker development

Sugar maple leaf samples were collected from a location near Houghton, MI, USA (47°07'07.46''N, 88°35'16.41''W) in a suburban setting (Figure 1 - Supporting Information) and dried in silica gel in August 2012. A total of 48 adult trees were sampled. Additionally, seeds were collected from one adult tree. GPS locations for all 48 trees were recorded. Total genomic DNA was extracted from dried leaf tissues of adult trees and from 96 embryos of seed from one seed parent with the DNeasy

Plant Mini Kit following the standard protocol (Qiagen, Hilden, Germany). Approximately 1 cm² dried leaf tissue was used for DNA extraction. For the inheritance analyses, embryos were carefully separated from the seed coat prior to DNA extraction. The fruits (samaras) contained only a single seed. The embryos ranged from 5 to 8 mm in size.

We applied low coverage whole genome paired-end Illumina sequencing in ten hardwood tree species (Owusu et al. 2013, Staton et al. submitted) including sugar maple. A total of 10,152 di-, tri and tetranucleotide microsatellite motifs were identified using an SSR finder script (Staton et al. submitted). After removal of redundant sequences (identity $\geq$ 95%) using the program cap3 (Huang 1999), primer3 (Rozen & Skaletsky 2000) was used to identify primers with default settings. Primer pairs were designed for 8,012 di-, 1,665 tri- and 475 tetranucleotide read motifs. The genomic SSR data are deposited at NCBI (http://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA200681, accession PRJNA200681). In total, we selected 96 primer pairs for dinucleotide repeat microsatellites for PCR amplification.

PCR amplification

For marker development, amplification and polymorphism was assessed for 96 dinucleotide repeat SSRs in a panel of six unrelated individuals on the QIAxcel Fast Analysis System with the DNA high resolution kit (Qiagen). PCR was performed in a 15 µl reaction mix and contained 3 µl of 5× HOT FIREPol® Blend Master Mix Ready to Load (contains 10 mM MgCl₂, 0.6 units of HOT FIREPol® Taq polymerase, and 2 mM dNTPs) (Solis Bio-Dyne, Estonia), 2 µl each of 5 µM forward and reverse primers (Sigma Aldrich, St. Louis, MO and Applied Biosystems, Foster City, CA), 6 µl double deionized water (DNase and RNase free) and 2 µl DNA (~1.8 ng/µl). Amplification was done in a Peltier Thermal Cycler (GeneAmp® PCR system 2720, Applied Biosys-

tems). The PCR profile consisted of 15 minutes denaturation at 95°C, 35 cycles denaturation at 94°C for 45s, annealing at T_a for 45s (Table 1), 45s elongation at 72°C, and an extension step at 72°C for 20 minutes. The amplified PCR products were separated on agarose gels and on the QIAxcel Fast Analysis System. Polymorphic loci from this screening panel and six additional SSR loci (Graignic et al. 2013) were amplified in 48 samples from a single population and in 96 single tree progenies from one seed parent using fluorescent labeled forward primers (6-FAM, PET, NED and VIC). PCR products were separated on an ABI Prism® Genetic Analyzer 3730 with Gene-ScanTM LIZ-500 as internal size standard (Applied Biosystems) and scored with GeneMarker®

V2.6.3 (SoftGenetics). All markers were tested for marker transferability in four other maple species (*A. ginnala, A. platanoides, A. rubrum, A. saccharinum*). Three samples were collected from each of these four species, and all markers were tested for amplification and polymorphism.

Data analysis

Genetic variation assessment was conducted for individual microsatellite markers for the 48 samples using GenAlEx 6.501 (Peakall & Smouse 2006, Peakall & Smouse 2012). Observed and expected heterozygosities (H_o, H_e, respectively) (Nei 1973) and number of alleles (N_a) were calculated. HWE exact tests and pair-

Table 1 Primer sequences and characteristics of thirteen genomic microsatellite markers for *Acer saccharum*

Locus name	Forward primer sequence (5'-3') Reverse primer sequence (5'-3')	Repeat motif	5' dye	Amplicon size (bp)	Size range (bp)	T_a (⁰C)	Accession number
AS05	GGAGGTGATGAGAGAATCTGG GAATGTGTGGCTGTGATTGG	$(AT)_9$	VIC	124	100-147	58	PRJNA200681
AS06*	GTCCACTGAAAGGAGATCGG ACCCAGAATTCCAGAAACCC	$(GA)_{11}$	6-FAM	114	98-156	58	PRJNA200681
AS09*	CAACGTCTCATGCAAATATCG TGTACCACAACAATGCAGGG	$(CA)_{14}$	6-FAM	149	128-155	58	PRJNA200681
AS11	AACATGTGCGGTCTTCTCC AGAAGGGTCATCGACGTGC	$(CT)_9$	VIC	154	89-152	58	PRJNA200681
AS47	CCATGGCCATCTATCTCTCC TGAATCTCCATTGCTGATCG	$(CT)_{12}$	6-FAM	128	106-139	56	PRJNA200681
AS48	CCAAATGACACCAAGATCACC GTTGAATCCCTGTCCCTTCC	$(GA)_{19}$	NED	115	90-113	56	PRJNA200681
AS127* [b]	TCACCGGAAAGGACAAGTGC CCAATCGAACCTAAACGGCG	$(TC)_{17}$	6-FAM	107	100-120	56	PRJNA200681
Am116 [a]	AACGCTACCGACTTCGCCAACT TGGAGGTCAAGTGCTGGAAACAA	$(AG)_6$	PET		230-273	56	AB303350
SM14 [a]	TTTTATGTAGAGCAACTCAACCCA TATCTGCTGCTTGACATGAACTT	$(GA)_{15}$	NED		72-112	56	KC751436
SM21A [a]	TAGTTGTGCACCAACCATGC TCCATCAAAACGCTGCTATG	$(GAT)_{14}$	VIC		179-243	56	KC751437
SM22 [a]	CCAGAGCTTGAAGAAAATGTACG GGTAAGGGGTTGTTTATGCAAT	$(GA)_{13}$	PET		295-323	56	KC751438
SM27 [a,]	TTTCTACTTTAGAGATGGAAACGG CCCTAAATCCCAAATCAGTGAA	$(TC)_{10}$	PET		242-260	56	KC751440
SM37 [a]	TGGGAGAAAATTAACATAGGATTT AGACTTGACTCTCCTAATCTTGGTG	$(AG)_{14}$	NED		174-198	56	KC751444

Note. T_a – annealing temperature, * amplifies multiple loci (multi banding pattern), [a] SSRs published in Graignic et al. (2013), [b] Reverse primer sequence (5'-3') labeled fluorescently multi-banding.

wise linkage disequilibrium (LD) were calculated with GENEPOP 4.2 (Raymond & Rousset 1995, Rousset 2008). For the three markers that revealed a multi-banding pattern (Table 1) Shannon's information index I (Brown & Weir 1983) and Nei's gene diversity H_e (Nei 1973) were calculated in GenAlEx. The frequency of the null allele at each locus was estimated by the square root method (Nei 1987).

Paternity exclusion probabilities for candidate pollen parents were calculated at co-dominant markers using the computer program CERVUS 3.0.7 (Marshall et al. 1998). Paternity analysis was carried out in CERVUS using log-likelihood ratios (LOD score) for the paternity assignment on the basis of observed multilocus genotypes at a strict (95%) and relaxed (80%) confidence level. Simulations of paternity analysis were completed to determine confidence levels of LOD scores (Pakkad et al. 2008, Pluess et al. 2009). The parameters for the simulation analyses were: number of cycles: 100,000; number of candidate parents: 200; proportion of candidate fathers sampled: 25%; proportion of loci typed: 99%; loci mistyped: 1%; error rate in likelihood calculations: 0.010 and the minimum number of typed loci: 4.

Pairwise distances between trees (m) were calculated in GenAlEx based on GPS coordinates.

Spatial genetic structure (SGS) was characterized in the population by spatial autocorrelation analysis using SPAGeDi 1.4c (Hardy & Vekemens 2002). Ten distance classes were selected ensuring an equal number of pairwise comparisons in each distance class (113 comparisons) at an approximate increment of 15 m per distance class. To assess the association between genetic similarity and geographic distance, regression analyses of pairwise kinship coefficients (F_{ij}) (Loiselle et al. 1995) on geographical distances were performed and a jackknife method was used to estimate standard errors. Lower and upper 95% confidence intervals were established by 10,000 permutations of individuals among distance classes. Correlograms of pairwise kinship coefficients (F_{ij}) were plotted against distance and 95% confidence intervals are shown (Figure 2).

The Sp statistic was derived from the kinship coefficient F_{ij} between individuals of the first distance class and its regression slope b_f as $Sp = -b_f/(1- F_{ij})$.

To calculate the census density (CD) as individuals per hectare according to Stefenon et al. (2008), a polygon was drawn to represent the population area using Google Earth Pro (Source: ''Houghton'', MI, USA, 47°07'07.46''N, 88°35'16.41''W. Google Earth Pro. May 9, 2014. November 26, 2014). The total area was 4393.81 m² (0.493 hectare) and the census density was 109.25 individuals per hectare.

An inheritance analysis was performed for all markers using single tree progeny following the approach by Gillet and Hattemer (1989). The following assumptions were tested. For a heterozygous seed parent A_iA_j each offspring has either allele A_i or A_j. Additionally, under regular segregation and absence of selection and other evolutionary factors, the following absolute genotype frequencies are expected in the progeny: $N(A_iA_j) = N(A_iA_i) + N(A_jA_j)$, and $N(A_iA_k) = N(A_jA_k)$, with $k \neq i, j$.

Results

Initially 96 primer pairs were tested on a set of DNA samples from six individuals in the source population. Out of these 96 primer pairs, 72 primer pairs successfully amplified products in the expected size range, but only seven produced a clear discernible banding pattern on agarose gels and on the QIAxcel Fast Analysis System using the high resolution kit for microsatellite analysis. These markers were selected as potential polymorphic microsatellites for further testing on an ABI Prism® Genetic Analyzer 3730. Three markers amplified more than one gene locus and were inter-

preted as dominant (absence/presence) markers, while four markers amplified a maximum of two products per individual and were interpreted as co-dominant markers (see Figure 2 - Supporting Information). Also, all six published markers (Graignic et al. 2013) showed a pattern consistent with the amplification of a single gene locus. Using the diversity panel of 48 individuals from one population, all ten co-dominant microsatellite markers showed relatively high levels of polymorphism with number of alleles per locus ranging between 9 and 16. Observed heterozygosity (H_o) ranged from 0.436 to 0.917 and expected heterozygosity (H_e) ranged from 0.726 to 0.894. The inbreeding coefficient, F_{is}, was estimated for each locus in the population using GENEPOP, and the F_{is} values ranged from -0.0581 to 0.4992 with three loci showing values significantly different from zero (α=0.05) after Bonferroni corrections (Table 2). Out of forty five marker pairs, five marker pairs showed significant linkage disequilibrium at the 5% level and no marker pair was significant at the 1% level.

Shannon's information index I and gene diversity for dominant markers were 0.193 and 0.102, respectively, with 64 variable positions in 48 adults.

Results of the inheritance analysis for the ten putative single-locus markers using single tree progenies of one heterozygous seed parent are shown in Tables 3 and 4. Table 3 includes the comparisons for absolute genotype frequencies of the heterozygous genotype $N(A_iA_j)$ and of both homozygous genotypes $N(A_iA_i)$ and $N(A_jA_j)$. Chi-square tests of goodness-of-fit for observed and expected frequencies revealed significant deviations between observations and expectations for AS05 (p < 0.0001), SM21A (p = 0.012) and AS47 (p = 0.029). Table 4 includes absolute frequencies of heterozygous genotypes $N(A_iA_k)$ and $N(A_jA_k)$ ($k \neq i, j$) and results of the Chi-square test of goodness-of-fit for observed and expected frequencies. Only AS05 showed significant deviation from the expected 1:1 ratio in this test. The observed

frequencies indicated regular segregation patterns for the other nine markers. Consequently, AS05 was excluded from the gene flow and SGS analyses.

Gene flow analyses using the computer program CERVUS showed effective dispersal of pollen in the sugar maple population both at the 95% and 80% confidence levels (Figure 1 A, B). The nine highly polymorphic single-locus microsatellite markers showed high exclusion probabilities for paternity analysis (Table 2); the total paternity exclusion probability for all loci was 0.999. At the 95% strict threshold 22% and at the 80% relaxed threshold 18% of the seeds were assigned to their putative fathers within the stand. Thus, the minimum estimates for gene flow from outside the plot are 78% and 82% under the strict and relaxed thresholds, respectively. Some seeds also resulted from the self-pollination (2%). Figure 1 shows the average gene flow distance in the population stand ranging from 0 m (self-pollination) to 81 m, with a mean pollen dispersal distance ranging from 33.25 m to 38.75 m at 95% and 80% confidence intervals, respectively.

Spatial genetic structure analysis using SPAGeDi detected non-significant SGS, corroborating effective dispersal by both seeds and pollen (Figure 2). Consequently the estimate for the Sp statistic was very low (0.00576). Gene dispersal estimates did not reach convergence.

Marker transferability tests showed that all thirteen markers were transferable to four species of *Acer* (*A. rubra, A. saccharinum, A. platanoides, A. ginnala*). Most of the markers showed a multi-banding pattern. However, AS47 and SM37 amplified a single polymorphic locus in *A. ginnala*, and markers Am116 and SM14 showed a single-locus pattern in *A. platanoides*. Finally, SM21A was heterozygous for the same alleles in *A. saccharinum*.

Table 2 Genetic variation parameters and exclusion probabilities at codominant markers in *Acer saccharum*

Locus name	N	N_a	H_o	H_e	F_{is}	Exclusion probability
[a]AS05	38	13	0.436	0.854	0.4992*	0.7251
AS11	48	15	0.521	0.726	0.2922*	0.5562
AS47	48	14	0.771	0.864	0.1179	0.7309
AS48	48	15	0.792	0.852	0.0815*	0.7321
Am116	47	14	0.729	0.781	0.0774	0.5908
SM14	48	16	0.750	0.880	0.1576	0.7738
SM21A	48	11	0.917	0.858	-0.0581	0.7173
SM22	48	13	0.813	0.894	0.1019	0.7940
SM27	48	9	0.708	0.757	0.0750	0.5615
SM37	48	9	0.646	0.750	0.1487	0.5553

Note. N - number of samples, N_a - number of alleles, H_o - observed heterozygosity, H_e - expected heterozygosity, *: significantly different from Hardy-Weinberg proportions ($\alpha = 0.05$) after Bonferroni corrections. [a]: this marker showed irregular segregation and likely amplified more than one gene locus (see Table 3).

Table 3 Segregation within the progenies of the putatively heterozygous seed tree

Marker	Maternal genotype (bp)	Frequency of offspring genotypes				
	A_i/A_j	A_iA_i	A_jA_j	A_iA_j	χ^2	p
AS05	124/138	0	5	42	32.44*	<0.0001
AS11	120/140	16	12	35	0.77	0.38
AS47	124/126	1	3	13	4.76*	0.03
AS48	93/111	6	1	2	2.78	0.10
Am116	231/241	13	19	23	1.47	0.22
SM14	72/86	5	1	2	2.00	0.16
SM21A	201/207	2	2	15	6.37*	0.01
SM22	311/315	4	5	8	0.06	0.77
SM27	248/256	6	11	24	1.20	0.27
SM37	178/180	21	11	31	0.02	0.89

Note. A_iA_j: genotype of the seed tree. A_iA_i and A_jA_j: frequency of putatively homozygous progenies. A_iA_j: frequency of putatively heterozygous progenies. Chi square test to analyze statistical significance between expected and observed segregation frequencies ($\chi2$-test; $df = 1$) and exact probabilities p.

Table 4 Segregation within the progenies of the putatively heterozygous seed tree

Marker	Maternal genotype (bp)	Frequency of offspring genotypes			
	A_i/A_j	A_iA_k	A_jA_k	χ^2	p
AS05	124/138	1	37	34.11*	0.0001
AS11	120/140	15	11	0.61	0.43
AS47	124/126	37	42	0.32	0.57
AS48	93/111	47	39	0.74	0.39
Am116	231/241	22	19	0.22	0.64
SM14	72/860	37	50	1.94	0.16
SM21A	201/207	40	37	0.12	0.68
SM22	311/315	34	44	1.28	0.26
SM27	248/256	24	31	0.89	0.35
SM37	178/180	19	11	2.13	0.14

Note. A_iA_j: genotype of the seed tree. A_iA_k and A_jA_k: frequency of putatively heterozygous progenies. Chi square test to analyze statistical significance between expected and observed segregation frequencies ($\chi2$-test; $df = 1$) and exact probabilities p.

Discussion

Low coverage whole genome sequencing proved to be useful to generate microsatellite resources for ten North American hardwood species including sugar maple (Owusu et al. 2013, Staton et al. submitted). To date, locus-specific markers have been developed for two of the species, sugar maple (present study) and honeylocust (*Gleditsia triacanthos* L.) (Owusu et al. 2013). While most primers produced fragments in the expected size range, the percentage of polymorphic and single-locus markers was low as compared to microsatellites developed from microsatellite-enriched genomic libraries for related species using Sanger sequencing (Pandey et al. 2004). Thus, out of 96 markers only seven were polymorphic and three of them showed regular segregation in the inheritance analysis (present study). In honeylocust, 14 locus-specific markers were developed after testing 144 primer pairs using the same procedure as in sugar maple (Owusu et al. 2013). Additional
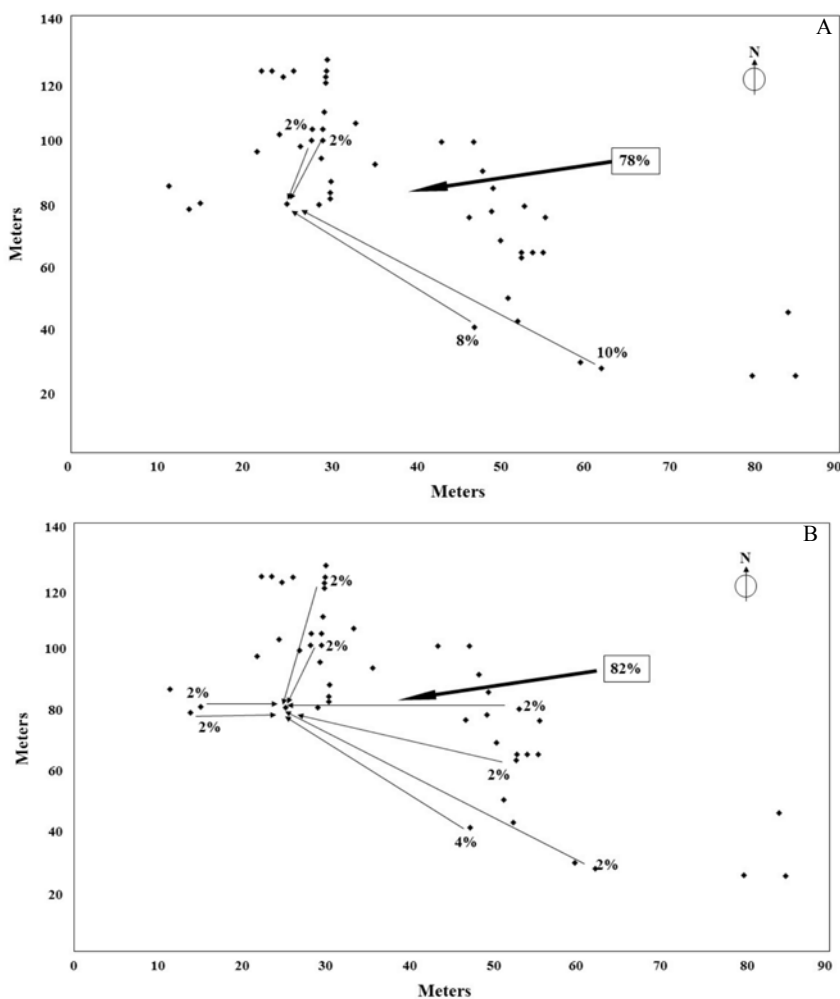

Figure 1 Gene flow at the 95% (A) and 80% confidence level (B)

studies are necessary to evaluate whether the relatively low number of polymorphic and locus-specific markers are specific for the species studied or are a peculiarity of the sequencing technique. Thus, AT-rich, relatively short motifs were preferentially recovered from Illumina sequencing for all ten species (Owusu et al. 2013, Staton et al. submitted). The relatively low number of polymorphic loci that were discovered in sugar maple may be partly related to the bias toward short and AT-rich motifs (Staton et al. submitted). In the present study, primers were developed based on low coverage genome sequencing (without enrichment step) for dinucleotide repeat motifs with repeat numbers ranging from 8 to 25 (mean of 16.8). As a comparison, pyrosequencing of an enriched microsatellite library in sugar maple resulted in the development of a moderate number of single-locus polymorphic markers (18 new next-generation sequencing derived markers out of 90 markers tested) when motifs with the highest number of repeats were targeted (10 - 25 repeats) (Graignic et al. 2013). The relatively low read coverage for reconstructed fragments with SSR motifs in the present study (~ 1X coverage, Staton et al. submitted) may

have resulted in wrong base calls in the primer binding sites and in unspecific PCR amplification. Also, amplification and sequencing errors might have resulted in an overestimation of repeat numbers for microsatellite regions. Thus, a 20 fold sequencing coverage was suggested to compensate for sequencing errors (Dohm et al. 2008).

In summary, low coverage genome sequencing was very cost-effective with regard to the generation of sequenced reads with microsatellite motifs (Staton et al. submitted), but a comparatively large number of markers needed to be tested for the development of locus-specific polymorphic markers for population genetic analyses.

In the present study we assessed the suitability of six published (Graignic et al. 2013) and of seven newly developed nuclear microsatellites for population genetic analyses. Specifically, we tested for Hardy-Weinberg proportions and regular segregation using open pollinated single tree progeny (Gillet & Hattemer 1989). Based on these results we selected nine polymorphic and easy-to interpret markers to analyze fine scale genetic structure and gene flow in one sugar maple population. In addition to
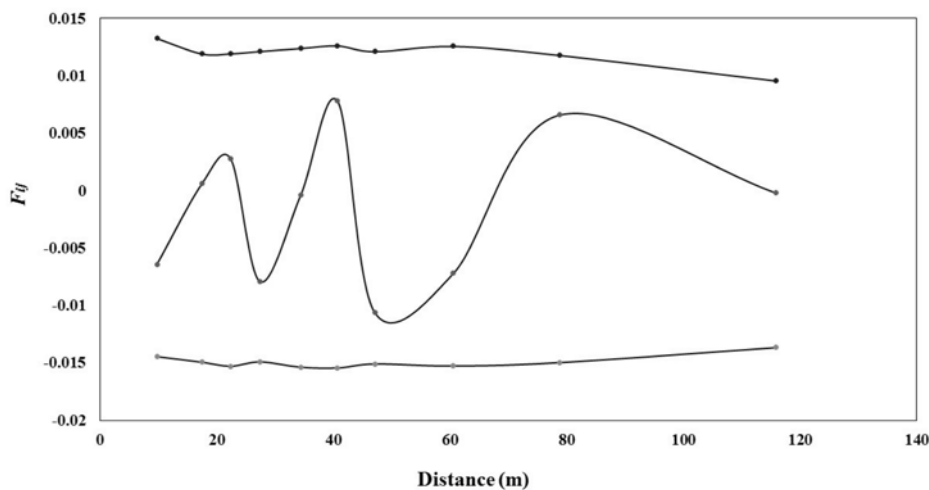

Figure 2 Correlograms of kinship coefficient measures (F_{ij}) plotted against distance

other published microsatellites for sugar maple (Graignic et al. 2013), these markers can serve as tools for future population genetic analyses. Also, all markers are transferable to four other species of *Acer* and some of them amplify single polymorphic loci opening up opportunities for population genetic studies across different maple sections.

Gene flow analyses based on paternity assignment in CERVUS showed very low gene flow between trees within the sampled stand and high gene flow from outside the stand (78%, 82%) both at the strict and relaxed confidence levels, suggesting a prevalence of long-distance pollen dispersal. Within the sampled stand, contributions of individual pollen parents to the progeny were ≤ 10% at the strict confidence level (95%) and ≤ 4% at the relaxed level, and most neighboring trees did not contribute to the progeny. The sexual system (e.g. dioecy, monoecy) and differences in flowering time are important determinants for the patterns of gene flow within a population and may result in uneven contributions of pollen donors to the progeny and in the development of spatial genetic structure (SGS) at different spatial scales (Young & Merriam 1994).

Since sugar maple trees can be monoecious (male and female flowers are produced on the same tree) or functionally dioecious (separate male and female trees) (Stephens 1973, Geburek 1993, Nesom 2006), self-pollination is likely limited. Even in monoecious trees selfing rate is expected to be comparatively low since most female and male flowers come into bloom at different times (Gabriel 1968). Accordingly, only 2% of the seeds were found to be originating from self-pollination in the present study.

In summary, gene flow patterns as inferred by paternity analysis suggested effective means of pollen dispersal in sugar maple and no preferential mating between neighboring trees as would have been expected for solely insect-pollinated species. These results sug-

gest that wind pollination is predominant and sufficient for pollination (Gabriel & Garrett 1984). However, a spatial clustering of related individuals can also result from spatially restricted seed dispersal. In the present study, spatial autocorrelation analyses showed no significant SGS in any distance class, suggesting random and long-distance dispersal of both pollen and seeds. Since the stand is affected by sugar maple dieback, mortality could also have resulted in a secondary disintegration of family structures. In accordance with our results, most allozyme loci showed non-significant SGS at small distance classes. Thus, spatial autocorrelation analyses in six natural sugar maple populations showed significant autocorrelation coefficients (Morans *I*) in three out of nine allozymes in only two populations (Geburek 1993). Significant spatial aggregation of alleles was found for seven alleles for a total of 30 tests in three natural populations (Perry & Knowels 1991). In seedling cohorts of fragmented and continuous sugar maple forests, no association between genetic and geographic distance and no SGS at smaller distance classes were found at ten allozyme loci (Young & Merrium 1994). However, spatial aggregation of genotypes at different larger distance classes suggested that SGS might be influenced by the spatial patterns of flowering times within a population (Gabriel 1968, Young & Merrium 1994). In the present study, we found significant SGS between 60 m and 140 m for only a single marker (SM14). Since the ability to detect SGS is positively correlated with the number of polymorphic markers analyzed (Cavers et al. 2005), inclusion of a much larger number of SSR loci might detect weak SGS at some distance classes. However, significant SGS at a similar or smaller number of polymorphic SSRs was found in other outcrossing and wind-pollinated species (Stefenon et al. 2008) including *Acer pseudoplatanus* (Pandey et al. 2012).

Conclusions

In addition to other recently developed microsatellite makers, the newly developed and tested markers proved to be useful for the assessment of genetic variation and structure, and for gene flow analyses. They will be an important tool to monitor the level of genetic variation especially at the southern range edge of the species for which a decrease in abundance is predicted as the result of warming climates (Iverson et al. 2008, Graignic et al. 2013).

Acknowledgements

References

Bal T.L., Richter D.L., Storer A.J., Jurgensen M.F., 2013. The relationship of the sapstreak fungus, *Ceratocystis virescens*, to sugar maple dieback and decay in Northern Michigan. American Journal of Plant Sciences 4: 436-443. DOI: 10.4236/ajps.2013.42A056.

Brown A.H.D., Weir B.S., 1983. Measuring genetic variability in plant populations. In: Tanksley S.D., Orton T.J. (eds). Isozymes in Plant Genetics and Breeding, Part A. Elsevier, Amsterdam, pp 219-239.

Cavers S., Degen B., Caron H., Lemes M.R., Margis R., Salgueiro F., Lowe A.J., 2005. Optimal sampling strategy for estimation of spatial genetic structure in tree populations. Heredity 95: 281-289. DOI: 10.1038/sj.hdy.6800709.

Dohm J.C., Lottaz C., Borodina T., Himmelbauer H., 2008. Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. Nucleic Acids Research 36: E1005. DOI: 10.1093/nar/gkn425.

Finkeldey R., Ziehe M., 2004. Genetic implications of silvicultural regimes. Forest Ecology and Management 197: 231-244. DOI: 10.1016/j.foreco.2004.05.036.

Gabriel W.J., 1968. Dichogamy in *Acer saccharum*. Botanical Gazzette 129: 334-338. DOI: 10.1086/336453.

Gabriel W.J., Garrett P.W., 1984. Pollen vectors in sugar maple (*Acer saccharum*). Canadian Journal of Botany 62: 2889-2890. DOI: 10.1139/b84-385.

Geburek T., 1993. Are genes randomly distributed over space in mature populations of sugar maple (*Acer saccharum* Marsh.)? Annals of Botany 71: 217-222. DOI: 10.1006/anbo. 1993.1027.

Gibson P.J., Wheelwright N.T., 1995. Genetic structure in a population of a tropical tree *Ocotea tenera* (Lauraceae): influence of avian seed dispersal. Oecologia 103: 49-54. DOI: 10.1007/BF00328424.

Gillet E., Hattemer H.H., 1989. Genetic analysis of isozyme phenotypes using single tree progenies. Heredity 63: 135-143. DOI: 10.1038/hdy.1989.84.

Godman R.M., Yawney H.W., Tubbs C.H., 1990. *Acer saccharum* Marsh. Sugar maple. In: Burns R.M., Honkala B.H. (eds.) Silvics of North America, Agricultural handbook 654 volume 2. Hardwoods- USDA Forest Service 2: 78-91.

Graignic N., Tremblay F., Bergeron Y., 2013. Development of polymorphic nuclear microsatellite markers in sugar maple (*Acer saccharum* Marsh.) using cross-species transfer and SSR-enriched shotgun pyrosequencing. Conservation Genetic Resources 5: 845-848. DOI: 10.1007/s12686-013-9923-7.

Graignic N., Tremblay F., Bergeron Y., Williams J., 2014. Geographical variation in reproductive capacity of sugar maple (*Acer saccharum* Marshall) northern peripheral populations. Journal of Biogeography 41: 145-157. DOI: 10.1111/jbi.12187.

Gunter L.E.G., Tuskan A.C., Gunderson A., Norby R.J., 2000. Genetic variation and spatial structure in sugar maple (*Acer saccharum* Marsh.) and implications for predicted global-scale environmental change. Global Change Biology 6: 335-344. DOI: 10.1046/j.1365-2486.2000.00313.x.

Guries R.P., Nordheim E.V., 1984. Flight characteristics and dispersal potential of maple samaras. Forest Science 30: 434-440.

Hardy O.J., Vekemens X., 2002. SPAGeDi: A versatile computer program to analyse spatial genetic structure at individual or population levels. Molecular Ecology Notes 2: 618-620. DOI: 10.1046/j.1471-8286.2002.00305.x.

Huang X., 1999. CAP3: A DNA sequence assembly program. Genome Research 9: 868-877. DOI: 10.1101/gr.9.9.868.

Iverson L.R., Prasad A.M., Matthews S.N., Peters M., 2008. Estimating potential habitat for 134 eastern US tree species under six climate scenarios. Forest Ecology and Management 254: 390-406. DOI: 10.1016/j.foreco.2007.07.023.

Krustovsky K.V., Neale D.B., 2005. Forest genomics and new molecular approaches to measuring and conserving adaptive genetic diversity in forest trees, In: T. Geburek and J. Turok (eds.), Conservation and Management of Forest Genetic Resources in Europe, Arbora Publishers, Zvolen, pp 369-390.

Loiselle B.A., Sork V.L., Nason G., Graham C., 1995.

Spatial genetic structure of a tropical understory shrub. American Journal of Botany 82: 1420-1425. DOI: 10.2307/ 2445869.

Marshall T.C., Slate J., Kruuk L.E.B., Pemberton J.M., 1998. Statistical confidence for likelihood-based paternity inference in natural populations. Molecular Ecology 7: 639-655. DOI: 10.1046/j.1365-294x.1998.00374. x.

Nei M., 1973. Analysis of gene diversity in subdivided populations. Proceedings of the National Academy of Sciences USA 70: 3321-3323. DOI: 10.1073/ pnas.70.12.3321.

Nei M., 1987. Molecular Evolutionary Genetics. Columbia University Press, New York.

Nesom G., 2006. Plant Guide, sugar maple. United States Department Department of Agriculture Natural Resources Conservation Service (http://plants.usda.gov/ plantguide/pdf/pg_acsa3.pdf).

Owusu S.A., Staton M., Jennings T.N., Schlarbaum S., Coggeshall M.V., Romero-Severson J., Carlson J.E., Gailing O., 2013. Development of genomic microsatellites in *Gleditsia triacanthos* (Fabaceae) using Illumina sequencing. Applications in Plant Sciences 1 (12): 1300050. DOI: 10.3732/apps.1300050.

Pakkad G., Ueno S., Yoshimaru H., 2008. Gene flow pattern and mating system in a small population of *Quercus semiserrata* Roxb. (Fagaceae). Forest Ecology and Management 255: 3819-3826. DOI: 10.1016/ j.foreco.2008.03.017.

Pandey M., Gailing O., Fischer D., Hattemer H.H., Finkeldey R., 2004. Characterization of microsatellite markers in sycamore (*Acer pseudoplatanus* L.). Molecular Ecology Notes 4: 253-255. DOI: 10.1111/j.1471-8286.2004.00633.x.

Pandey M., Gailing O., Hattemer H.H., Finkeldey R., 2012. Fine-scale spatial genetic structure of sycamore maple (*Acer pseudoplatanus* L.). European Journal of Forest Research 131: 739-746. DOI: 10.1007/s10342-011-0546-9.

Peakall R., Smouse P. E., 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28: 2537-2539. DOI: 10.1093/bioinformatics/bts460.

Peakall R.O.D., Smouse P.E., 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6: 288-295. DOI: 10. 1111/j.1471-8286.2005.01155.x.

Perry D.J., Knowels P., 1991. Spatial genetic structure within three sugar maple (*Acer saccharum* Marsh) stands. Heredity 66: 137-142. DOI: 10.1038/hdy.1991.17.

Pluess A.R., Sork V.L., Dolan B., Davis F.W., Grivet D., Merg K., Papp J., Smouse P.E., 2009. Short distance pollen movement in a wind-pollinated tree, *Quercus lobata* (Fagaceae). Forest Ecology and Management 258: 735-744. DOI: 10.1016/j.foreco.2009. 05.014.

Raymond M., Rousset F., 1995. GENEPOP (Version 1.2): Population Genetics Software for Exact Tests and Eucemenicism. Journal of Heredity 86: 248 - 249.

Rousset F., 2008. genepop'007: a complete re-implementation of the genepop software for Windows and Linux. Molecular Ecology Resources 8: 103-106. DOI: 10.1111/j.1471-8286. 2007.01931.x.

Rozen S., Skaletsky H. , 2000. Primer3 on the WWW for general users and biologist programmers. Methods in Molecular Biology 132: 365-386.

Selkoe K.A., Toonen R.J., 2006. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. Ecology Letters 9:615-629. DOI: 10.1111/j.1461-0248.2006.00889.x.

Stefenon V.M., Gailing O., Finkeldey R., 2008. The role of gene flow in shaping genetic structures of the subtropical conifer species *Araucaria angustifolia*. Plant Biology 10:356-364. DOI: 10.1111/j.1438-8677.2008.00048. x.

Stephens H.A., 1973. Woody plants of the North Central plains. Lawrence, KS: The University Press of Kansas. 530 p.

Whitham T.G., Difazio S.P., Schweitzer J.A., Shuster S.M., Allan G.J., Bailey J.K., Woolbright S.A., 2008. Extending genomics to natural communities and ecosystems. Science 320: 492-495. DOI: 10.1126/science.1153918

Whitham T.G., Gehring C.A., Lamit L.J., Wojtowicz T., Evans L.M., Keith A.R., Smith D.S., 2012. Community specificity: life and afterlife effects of genes. Trends in Plant Science 17: 271-281. DOI: 10.1016/ j.tplants.2012.01.005.

Young A.G., Merrium H.G., 1994. Effects of forest fragmentation on the spatial genetic structure in *Acer saccharum* Marsh. (sugar maple) populations. Heredity 72: 201-208. DOI: 10.1038/hdy.1994.27.

Supporting Information

The online version of the article includes Supporting Information.

Figure 1. Sample location (total area = 4393.81 m^2)

Figure 2. Electropherograms showing marker amplification patterns